ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF ARTS AND SOCIAL SCIENCES

ESTIMATION OF FOOD DEMAND IN TURKEY USING QUADRATIC ALMOST IDEAL DEMAND SYSTEM

M.A. THESIS

Ali Furkan KALAY

Department of Economics
Economics M.A. Program

DECEMBER 2018

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF ARTS AND SOCIAL SCIENCES

ESTIMATION OF FOOD DEMAND IN TURKEY USING QUADRATIC ALMOST IDEAL DEMAND SYSTEM

M.A. THESIS

Ali Furkan KALAY
(412161005)

Department of Economics
Economics M.A. Program

Thesis Advisor: Prof. Dr. Sencer ECER

DECEMBER 2018

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ SOSYAL BİLİMLER ENSTİTÜSÜ

KARASEL İDEALE YAKLAŞIK TALEP SİSTEMİ İLE TÜRKİYE’DE GIDA TALEBİ TAHMİNİ

YÜKSEK LİSANS TEZİ

Ali Furkan KALAY
(412161005)

İktisat Anabilim Dalı
İktisat Yüksek Lisans Programı

Tez Danışmanı: Prof. Dr. Sencer ECER

ARALIK 2018

Ali Furkan KALAY, a M.A. student of ITU Graduate School of Arts and Social Sciences student ID 412161005, successfully defended the thesis/dissertation entitled "ESTIMATION OF FOOD DEMAND IN TURKEY USING QUADRATIC ALMOST IDEAL DEMAND SYSTEM" which he prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor:	**Prof. Dr. Sencer ECER** Istanbul Technical University	....
Jury Members:	**Prof. Dr. Bülent GÜLOĞLU** Istanbul Technical University	....
	Asst. Prof. Gökhan ÖVENÇ Istanbul University	....

Date of Submission: **16 November 2018**
Date of Defense: **11 December 2018**

To my wife and family,

FOREWORD

This thesis is written as completion of Master of Economics at Istanbul Technical University. It examines the Food Demand in Turkey by utilizing the Quadratic Almost Ideal Demand System in the context of Microeconomic Theory.

I would like to thank my thesis advisor, Prof. Dr. Sencer Ecer, for his contributions and wisdom. He has motivated and guided me all along with and before my Master education.

I received help from many people while writing my thesis. Prof. Dr. Bülent Güloğlu advised about my econometric estimation. My colleague, A. G. Buğra Akyüz, helped me to overcome some mathematical issues. I acquired many ideas from my other colleagues and faculty. I am grateful to all for their sincerity and help.

I, of course, owe a debt of gratitude to my parents and wife. They have encouraged me to develop myself and to chase my goals.

November 2018

Ali Furkan Kalay
Research Assistant

TABLE OF CONTENTS

Page

FOREWORD .... **ix**
TABLE OF CONTENTS .... **xi**
ABBREVIATIONS .... **xiii**
SYMBOLS .... **xv**
LIST OF TABLES .... **xvii**
LIST OF FIGURES .... **xix**
SUMMARY .... **xxi**
ÖZET .... **xxiii**
1. INTRODUCTION **1**
2. DEMAND THEORY **5**
2.1 Neoclassical Demand Theory 5
2.2 Demand Systems 11
2.2.1 Almost ideal demand estimation 11
2.2.2 Quadratic almost ideal demand estimation 16
2.2.3 Demographic variables 18
3. DATA **21**
3.1 Categorization 21
3.2 Data Manipulation 22
3.3 Prices 23
3.4 Descriptive Statistics 23
4. ECONOMETRIC ESTIMATION **29**
4.1 Estimation of the Model 29
4.2 Tests 31
4.3 Elasticities 32
4.3.1 Income elasticity 32
4.3.2 Price elasticity 32
5. CONCLUSION **35**
5.1 Results 35
5.2 Future Work 35
REFERENCES **37**
APPENDICES **41**
APPENDIX A 42
CURRICULUM VITAE **43**

ABBREVIATIONS

AI	**:**	Almost Ideal
AID	**:**	Almost Ideal Demand
AIDS	**:**	Almost Ideal Demand System
CG	**:**	Commodity Group
Coef.	**:**	Coefficient
EASI	**:**	Exact Affine Stone Index
IFGLS	**:**	Iterated Feasible Generalized Least Squares
LA/AIDS	**:**	Linear Approximated Almost Ideal Demand System
ML	**:**	Machine Learning
n.e.c.	**:**	Not Elsewhere Classified
QUAI	**:**	Quadratic Almost Ideal
QUAID	**:**	Quadratic Almost Ideal Demand
QUAIDS	**:**	Quadratic Almost Ideal Demand System
TURKSTAT	**:**	Turkish Statistical Institute

SYMBOLS

p	:	Price vector of commodities
m	:	Income level
v	:	Indirect utility level
W	:	Walrasian
H	:	Hicksian
q	:	Consumption quantity
ϵ	:	Price elasticity
μ	:	Income elasticity
w	:	Expenditure share
u	:	Utility level
α	:	Alpha coefficient of AIDS
β	:	Linear income coefficient
λ	:	Quadratic income coefficient
z	:	Vector of demographic variables
ρ	:	Scalar parameters of demographic variables
η	:	Demographic variables' effect vector

LIST OF TABLES

		Page
Table 1.1 :	Comparison of two elasticity estimations in Turkey for some products.	3
Table 3.1 :	Descriptive Statistics of Food Groups	24
Table 4.1 :	R^2 of Commodity Groups predictions	29
Table 4.2 :	QUAIDS estimation coefficients and their standard errors	30
Table 4.3 :	Scalar Coefficients	30
Table 4.4 :	Likelihood Ratio Test Results	31
Table 4.5 :	Income Elasticity of Commodity Groups	32
Table 4.6 :	Income Elasticity Simulation Results with Specification Trials	32
Table 4.7 :	Uncompensated Price Elasticities	33
Table 4.8 :	Compensated Price Elasticities	33

LIST OF FIGURES

Page

Figure 3.1 : Weighted Price Levels for food groups (monthly) 24
Figure 3.2 : Changes in same descriptive parameters between 2008 and 2016.... 26
Figure 3.3 : Monthly Expenditure Shares of Commodity Groups between 2008 and 2016 27

ESTIMATION OF FOOD DEMAND IN TURKEY USING QUADRATIC ALMOST IDEAL DEMAND SYSTEM

SUMMARY

Demand estimation is crucial for people, companies, and governments. It is important not only for predicting the future quantities, but is also important to understand the consumer behaviour. Because, understanding the consumer behaviour helps to decision-making process.

This thesis estimates the food demand in Turkey for the years between 2008 and 2016. Because, drastic increases in inflation rates in recent years is mostly attributed to the food prices. Besides, demand for food is relatively more inelastic compared to the other commodity groups. So, food demand in Turkey is an important subject for the welfare of the Turkish people.

One can make inferences on the consumers' well-being by observing the revealed preferences which is a subject of the Microeconomic Theory. By understanding the Neoclassical Demand Theory, one can better understand the consumers' behaviours (and preferences). Therefore, I examine the Neoclassical Demand Theory before estimating the food demand in Turkey.

I use Almost Ideal Demand System to estimate food demand, because it satisfies many desirable assumptions of the Neoclassical Demand Theory. However, I make estimations using Quadratic AIDS instead of AIDS. Because QUAIDS assumes that consumers have quadratic Engel Curve which is more appropriate for the demand systems. On the other hand, AIDS and QUAIDS are nonlinear models; but in order to make estimations faster, it is converted into a linear model called Linear Approximated AIDS. However, LA/AIDS has some econometric problems like endogeneity; therefore I estimate QUAIDS nonlinearly.

One of the distinctive feature of this study is that I estimate food demand after aggregation of household surveys. I convert surveys into time series data. By doing so, I notice unusual behaviour of the consumers during the sacrifice feasts; I estimate the model faster; and I do not need to apply censored demand techniques to reduce heterogeneity. However, the number of observations is reduced to 108 from (almost) 100.000 by aggregation. Thus, I make robustness checks at some points.

The results carry conviction and give important intuitions about the current debates like "bread pricing" and "meat imports". "Meat and Fish" and "Sugar, jam, honey, chocolate, and confectionary" commodity groups have high income elasticity. Thus, welfare effects can differ for the low and high income people. The difference between compensated and uncompensated price elasticity for "Bread and Cereals", "Fruit and Vegetables", and "Meat and Fish" commodity groups are very high. This means, consumers cannot substitute these commodity groups. Thus, price increases for these commodity groups yield high welfare loss.

In order to make better welfare analysis, one should estimate the food supply and supply elasticities of the commodity groups. And further methodological researches can be conducted on the machine learning algorithms. ML algorithms offer better performance of prediction accuracy and precision. However, there is a huge literature gap on interpretability of the ML algorithms. ML can be used for demand systems as well.

KARASEL İDEALE YAKLAŞIK TALEP SİSTEMİ İLE TÜRKİYE'DE GIDA TALEBİ TAHMİNİ

ÖZET

Talep tahmini insanlar, şirketler ve hükumetler için kritik bir konudur. Talep tahmini sadece gelecekteki talep miktarlarını öngörmek için değil, aynı zamanda tüketici davranışlarını anlamak için de önemlidir. Çünkü tüketici davranışlarını anlamak, karar verme süreçlerinde yardımcı olmaktadır.

Bu tez, 2008 ve 2016 yılları arasındaki gıda talebini tahmin etmektedir. Çünkü, son yıllarda enflasyon oranında yaşanan artış, büyük ölçüde gıda fiyatları ile ilişkilendirilmektedir. Diğer yandan, gıda talebi diğer emtialara göre, nispeten, daha az fiyat esnekliği göstermektedir. Bu yüzden, Türkiye'deki gıda talebi Türk halkının refah seviyesi için önemli bir unsurdur.

Bir kişi, tüketicilerin refahı hakkında tüketici tercihlerine bakarak çıkarımlar yapabilir. Bu yaklaşım mikro ekonomik teorinin konularından birisidir. Neoklasik talep teorisini anlamak, tüketici davranışlarının (ve tercihlerinin) daha iyi anlaşılmasını sağlayacaktır. Bu yüzden tezde, Türkiye'de gıda talebini tahmin etmeden önce, bu bağlamda neoklasik talep teorisinin incelemekteyim.

Gıda talebini tahmin etmek için İdeale Yaklaşık Talep Sistemini (AIDS) kullanmaktayım, çünkü bu model neoklasik talep teorisinde yer alan birçok önemli varsayımı sağlamaktadır. Ama, gıda talebinin ekonometrik tahminini AIDS yerine, Karesel AIDS (QUAIDS) kullanarak yapmaktayım. Çünkü QUAIDS tüketicilerin karesel Engel eğrisine sahip olduğu varsayımı kullanmaktadır, ki bu talep sistemleri için daha uygun bir yöntemdir. Diğer yandan, hem AIDS hem de QUAIDS doğrusal olmayan modellerdir. Literatürde genellikle sonuca daha hızlı yakınsamak için, LA/AIDS olarak adlandırılan doğrusal bir modele dönüştürülerek kullanılmaktadır. Ama, LA/AIDS modelinin içsellik (endogeneity) gibi ekonometrik sorunları vardır. Bu yüzden, ben bu tezde QUAIDS modelini doğrusallaştırma olmadan tahmin etmekteyim.

Bu çalışmanın özgün yönlerinden birisi de gıda talebini tüketici anketlerini birleştirme (aggregation) ile zaman serisine dönüştürerek tahmin etmemdir. Bu şekilde yaparak, özellikle kurban bayramındaki tüketicilerin farklı davranışlarını fark ediyorum; modeli çok daha hızlı tahmin ediyorum; ve heterojenliği azaltmak için sansürlü talep teknikleri uygulamama gerek kalmıyor. Ama gözlem sayısın birleştirme ile (neredeyse) 100.000'den 108'e düşmektedir. Bu yüzden bazı noktalarda, modele sağlamlık (robustness) testleri yapmaktayım.

Çıkan sonuçlar "ekmek fiyatları" ve "et ithalatı" gibi güncel tartışmalar hakkında önemli görüler vermektedir. "Et ve Balık" ve "Şeker, pekmez, bal, çikolata, ve şekerleme" ürün gruplarında yüksek gelir esnekliği görülmektedir. Bu yüzden, bu ürün gruplarında yüksek ve düşük gelirli insanlar için farklı refah etkileri gözlemlenebilir. Telafi edilmemiş ve telafi edilen fiyat esnekliği arasındaki fark ise "Ekmek ve Tahıl", "Meyve ve Se-

bzeler” ve “Et ve Balık” ürün gruplarında çok yüksek gözlenmektedir. Bu durum, tüketicilerin fiyat artışını diğer ürün grupları ile ikame edemediklerini manasına gelmektedir. Dolayısıyla, bu ürün gruplarındaki fiyat artışı yüksek refah kaybına neden olmaktadır.

Daha etkin bir refah analizi yapmak için, arz ve arz esnekliğinin de tahmin edilmesi gerekmektedir. Ayrıca bu konuda, makine öğrenmesi üzerine daha fazla metodolojik araştırılma yapılabilir. Makine öğrenmesi algoritmaları tahmin doğruluğu ve kesinliği hususunda daha iyi performans göstermektedir. Ama, makine öğrenmesi algoritmalarının yorumlanabilmesi konusunda ciddi bir literatür eksiği vardır. Talep tahmini sistemlerinde makine öğrenmesi yöntemleri kullanabilir.

1. INTRODUCTION

Demand models are useful for estimating price and income elasticity, making inferences about substitute and complementary goods, measuring welfare loss, among other things. So one should note that in this research, the phrase of "Demand model" implies "understanding the consumer behaviour" rather than "predicting the future demand quantities". This is also why this research investigates the neoclassical demand theory.

This research estimates the food demand in Turkey. Because, in recent years, there is a drastic increase in inflation rate in Turkey; and food prices are blamed to be main reason. Food demand is relatively more inelastic compared to other commodity groups. Therefore high inflation rates might lead to high welfare loss, especially for poor people. On the other hand, there are many studies estimating food demand in Turkey. By this means, we can make robustness check of this research's estimation results.

There are many demand systems like Rotterdam model (Theil, 1965), the translog model (Christensen, Jorgenson, & Lau, 1975), BLP (Berry, 1994), EASI (Lewbel & Pendakur, 2009), and so on. I use Deaton & Muellbauer's (1980) Almost Ideal Demand Estimation System to estimate the food demand in Turkey. It, by construction, satisfies many desirable assumptions of consumer behaviour. Besides, AIDS is one of the most popular demand systems in the literature. It has several extensions like Inverse AIDS (Eales & Unnevehr, 1994), Generalized AIDS (Alston, Chalfant, & Piggott, 2001), and Quadratic AIDS (Banks, Blundell, & Lewbel, 1997). Herein I conduct econometric estimation with the QUAI demand model by controlling for demographic parameters.

AIDS' functional form is nonlinear, and this makes its estimation relatively difficult. Thus, Deaton & Muellbauer (1980) replace the nonlinear component (price index) with a linear function that approximates it. The approximated model is referred as Linear Approximated AIDS and different approximation techniques for LA/AIDS are proposed later (Moschini, 1995). Most early studies used the LA/AIDS instead of (the original) AIDS.

There are different methodologies using AIDS and its extensions in the literature for food demand estimation. Jabarin (2005) estimate meat demand using LA/AIDS. Ulubasoglu, Mallick, Wadud, Hone, & Haszler (2016) estimate using the household expenditure survey in Australia by adding regional parameters to the model. Verbeke & Ward (2001) investigate the impact of negative TV press and advertising on food demand using AIDS. One distinctive feature of their study is that they incorporated non-linear AIDS by reconstructing the translog price index. Colchero, Salgado, Unar-Munguia, Hernandez-Avila, & Rivera-Dommarco (2015) estimate the price elasticity for beverages and soft drinks in Mexico by use of LA/AIDS with Laspeyres price index [1].

Food demand estimations in Turkey have been studied with similar methodologies with minor differences. Sengul & Tuncer (2005) estimate the food demand by separating the poverty levels using LA/AIDS with Stone's Price Index. They use TURKSTAT's household consumption expenditure survey of 1994. Separating poverty levels allow to reduce the heterogeneity and to estimate the demands for different income-levels. But low-income families have "zero" consumption for some foods, so they use Shonkwiler & Yen's (1999) two step estimation method to estimate the censored demand. They also include the socioeconomic parameters in the econometric model like age, settlement sizes, education and dummy variables for months. Akbay, Boz, & Chern (2007) use the same econometric model to estimate food demand but they increase the number of food groups, and they do not consider poverty levels. Bilgic & Yen (2013) use more homogeneous food groups to estimate food demand and they used Shonkwiler & Yen's (1999) two step estimation.

Demand estimations yield versatile results for income and price elasticities. For example, Armagan & Akbay (2008) use a similar methodology to estimate the demand for animal products using a household expenditure survey in Aydin, Turkey. Interestingly, their study finds more inelastic demand. This result might be the consequence of biased data, because people in Aydin wealthier than the average of Turkey. However, other studies show diversified results as well. On the other hand, LA/AIDS estimations are likely to be not robust because of two main reasons: (1) AIDS assumes a linear Engel curve [2] and linear approximation methods might lead to price endogeneity problem. In order to better understand the consumer behaviour, more robust estimations must be done. Table 1.1 compares the elasticity estimations of Armagan & Akbay (2008) and

[1] Laspeyres price index considered to be better approximates the translog price values than Stone's price index which is used by Deaton & Muellbauer (1980)

[2] By construction, AIDS is a rank-two demand system. So, it assumes a linear Engel Curve.

Bilgic & Yen (2013) for some products:

Table 1.1: Comparison of two elasticity estimations in Turkey for some products.

	Bilgic & Yen (2013)	Armagan & Akbay (2008)
Milk	-1.229	-0.311
Fish	-0.932	-0.432
Yogurt	-1.337	-0.300
Cheese	-0.712	-0.526

İpek & Akyazı (2015) criticize that demand estimations in Turkey done without controling the demand ranks as Lewbel (1991) proposed. İpek & Akyazı suggest to do ranking test before choosing a demand system or to use a non-parametric demand system like Exact Affine Stone Index [3] (Lewbel & Pendakur, 2009). However, Banks et al. (1997) show that Engel curves unlikely to be linear. So using rank-two demand systems (AIDS is a rank-two demand system) gives meaningful results only in rare cases and quadratic rank-three demand systems [4] should be enough for the most.

Theoretically, one of the very problematic issues is using price indices to linearize the econometric model. For instance, Stone's price index is criticized for allowing price endogeneity in the econometric model. The alternative price indexes are not perfect though they are slightly better (Henningsen, 2017).

There is a considerable literature for demand estimation for food in Turkey. Most of the studies were conducted with TURKSTAT's household survey. However there is no study using the recent household surveys, between 2012 and 2016. This research have three contributions to the literature: (1) Investigating the consistency of demand estimations with Neoclassical Microeconomic Theory; and (2) estimating (nonlinear) Quadratic AIDS instead of LA/AIDS.

Chapter 2 investigates the Neoclassical Demand Theory, and shows the derivations of AIDS and QUAIDS in the same context. Chapter 3 explains the data and methodology used for food demand estimation. Chapter 4 evaluates the econometric estimation results. Chapter 5 remarks the opinions for this and the future researches.

[3] EASI

[4] Assumes a quadratic Engel curve

2. DEMAND THEORY

This section summarizes some parts (that relate to AIDS) of the neoclassical demand theory from Mas-Colell, Whinston, & Green (1995) and Jehle & Reny (2011); and shows the derivation of AIDS and QUAIDS in the same context.

2.1 Neoclassical Demand Theory

The preferences of individuals are represented via utility functions which allow using calculus methods. A preference relation, $\succsim$, is represented with a utility function, $u(.)$, when:

$$u(x^0) \geq u(x^1) \iff x^0 \succsim x^1 \qquad \forall x^0, x^1 \in X = \mathbb{R}^L_+ \tag{2.1}$$

The existence of the utility function for a preference relation depends on properties of $\succsim$. If a rational preference is continuous on its consumption bundle, X, then there exists a continuous utility function, $u(x)$, representing the rational preference (Mas-Colell et al., 1995). Note that, existence of the utility function depends on the *rationality* condition.

However, existence of utility function that represents a preference relation does not imply $u(.)$ is unique. Say, there exists a relation such that $v = u + 3$. If $u(.)$ is a utility function representing the preference relation, $\succsim$, then $v(.)$ is a utility function representing the relation $\succsim$; because the condition 2.1 is satisfied for $v(.)$ as well. Thus, the utility functions are invariant to the strictly increasing transformations. The non-uniqueness property of utility functions allow using different functional forms representing the same preferences (Jehle & Reny, 2011).

It is possible that a utility function satisfying all required properties is continuous but undifferentiable such as the Leontief utility function. However, demand systems require differentiable utility functions to be able to make mathematical conversions (Mas-Colell et al., 1995). In short, *differentiability* is not a required but a desired property of utility functions.

It is generally difficult to derive the *utility function* of an individual but only revealed preferences can be observed. A rational consumer is expected to *maximize* his or her utility with a limited budget (m) and existing prices (p). So *consumer's problem* can be reduced to an optimization problem with a budget constraint. The Utility Maximization Problem (UMP):

$$\max_{x \in \mathbb{R}^n_+} u(x) \qquad s.t. \qquad p.x \leq m. \tag{2.2}$$

Solution of the problem gives consumer's preferences who is a *utility maximizer*. The solution is called *Walrasian Demand* or *Marshallian Demand*: $x^*(p, m)$. Proposition given by Mas-Colell et al. (1995, Proposition 3.D.2.) as follows:

Proposition: A continuous utility function $u(.)$ representing a locally non-satiated preference relation $\succsim$ is defined on a consumption set $X \in \mathbb{R}^L_+$. Then Walrasian demand corresponding the utility function $u(.)$ holds the following properties:

1. *Homogeneity degree of zero*: $x(ap, am) = x(p, m)$.
2. *Walras' Law* or *Budget Exhaustion Condition*: $p.x = m, \quad \forall x \in x(p, m)$
3. *Uniqueness*: The preference relation is assumed to be *convex* and therefore $u(.)$ is quasiconcave. This implies the $x(p, m)$ is convex and it has a unique solution (x^*).

The utility function defined on consumption bundle is called the *direct utility function*, because utility is the result of consumption bundle itself. The parameters rendering the consumption bundle (x^*) are prices (p) and income (m). However, the same problem problem can be defined in a different way:

$$v(p, m) = \max u(x) \qquad s.t. \quad p.x \leq m \tag{2.3}$$

An *indirect utility function* can be defined using the solution of the utility maximization problem (equation 2.2) x^* :

$$v(p, m) = u(x^*(p, m))$$

So parameters defining utility are the prices and the income, indirectly. The properties of indirect utility function defined on continuous utility function, $u(x)$, are as follows: (Mas-Colell et al., 1995):

1. *Homogeneity degree of zero*: $v(ap, am) = v(p, m)$
2. Strictly increasing in m and nonincreasing in p_l for any l
3. *Quasiconvex:* for any $\bar{v}$ the set $\{(p, w) : v(p, w) \leq \bar{v}\}$ is convex. Say there exists two extreme budget sets for p and p'. This property says that consumers prefer two extreme budget sets for the prices p and p' to the average the two, $\frac{p+p'}{2}$ (Jehle & Reny, 2011).
4. Continuous

Additionally, the indirect utility function has a property called "Roy's Identity". It implies that Walrasian Demand for a commodity is the negative ratio of partial derivatives of indirect utility respect to price of the commodity and the income level. So, if $v(p, m)$ is differentiable at (p, m) and $\partial v(p, m)/\partial m \neq 0$ then there is the relationship:

$$x_i(p, m) = -\frac{\partial v(p, m)/\partial p_i}{\partial v(p, m)/\partial m}, \qquad i = 1, ..., n. \tag{2.4}$$

Roy's identity constitutes an important step for derivation of AIDS.

The consumers' problem can be approached from another perspective. Remember the definition of indirect utility (equation 2.3). An *expenditure function* is defined that minimizes the expenditures while maintaining a minimum level of utility:

$$e(p, u) = \min_x p.x \qquad s.t. \qquad u(x) \geq u \tag{2.5}$$

If $u(.)$ is a strictly increasing and continuous utility function then properties of $e(p, u)$ are as follow(Mas-Colell et al., 1995):

1. Homogeneous of degree 1 in p
2. Strictly increasing in u and increasing in p_l for any commodity, l
3. Concave in p: $e(p', u) \geq \lambda e(p, u) + (1 - \lambda)e(p'', u)$ where $p'' > p' > p$ and $0 < \lambda < 1$.

4. Continuous

Expenditure function has another property called the *Shephard's lemma*: $e(p,u)$ is differentiable in p at (p^0, u^0) with $p^0 \gg 0$, and:

$$\frac{\partial e(p^0, u^0)}{\partial p_i} = h_i(p^0, u^0), \qquad i = 1, ..., n. \tag{2.6}$$

Where h_i is the Hicksian Demand.

The consumer's problem is reduced to a maximization problem in equation 2.2. However, by definition in equation 2.5, the consumer's problem changes and becomes an expenditure minimization problem which have a utility constraint. So the expenditure is minimized by maintaining a minimum utility level. The solution of the expenditure minimization problem can be used to derive the Compensated Demand or Hicksian Demand. The expenditure function have two parameters: u and p. Let's say $v(p,m) = u$ and the price levels had changed: $v(p', y) = u'$. Hicksian demand measures the compensated demand by answering the question: What would be the consumption bundle if we keep utility level u constant while prices are p'?

Properties of Hicksian demand function $h(p,u)$ (Mas-Colell et al., 1995, Proposition 3.E.3) are as follows:

1. *Homogeneity of degree zero in* p: So $h(ap,u) = h(p,u)$. This implies price levels does not change the compensated demand level, but only the relative prices are important.

2. *No excess utility*: $x \in h(p,u)$ implies $u(x) = u$.

3. *Convexity (Uniqueness)*: If preference relation is convex then $h(p,u)$ is convex; and if preference relation is strictly convex then quasiconcavity of $u(.)$ implies $h(p,u)$ has a unique solution.

So far, we approached the consumers' problem from two perspectives: *Utility maximization* and *cost minimization*. The solutions are the Walrasian Demand and the Hicksian Demand, respectively. So if we set the constraint of minimization problem in 2.5 as the solution of the maximization problem then it would give the Walrasian demand:

$$h(p, v(p,m)) = x(p,m) \tag{2.7}$$

Similarly we can derive the Hicksian Demand from Walras' Demand.

$$x(p, e(p, u)) = h(p, u) \tag{2.8}$$

The Slutsky Equation: Assume $u(.)$ is a continuous and strictly increasing utility function representing the preference relation, $\succsim$, and $u = v(p, m)$ for all p and m. Recall the equation 2.8 and take the partial derivative of $h_l(p, u^*) = x_l(p, e(p, u^*))$ respect to p_k:

$$\frac{\partial h_l(p, u^*)}{\partial p_k} = \frac{\partial x_l(p, e(p, u^*))}{\partial p_k} + \frac{\partial x_l(p, e(p, u^*))}{\partial m}\frac{\partial e(p, u^*)}{\partial p_k} \tag{2.9}$$

Replace $\partial e(p, u^*)/\partial p_k$ with $h_k(p, u^*) = h_k(p, v(p, m)) = x_k(p, m)$ using equations 2.6 and 2.7, respectively. After some rearrangements we have the Slutsky Equation:

$$\frac{\partial h_l(p, u)}{\partial p_k} = \frac{\partial x_l(p, m)}{\partial p_k} + x_k(p, m)\frac{\partial x_l(p, m)}{\partial m} \tag{2.10}$$

$$\underbrace{\frac{\partial x_l(p, m)}{\partial p_k}}_{\text{Total Effect}} = \underbrace{\frac{\partial h_l(p, u)}{\partial p_k}}_{\text{Substitution Effect}} \underbrace{- x_k(p, m)\frac{\partial x_l(p, m)}{\partial m}}_{\text{Income Effect}} \tag{2.11}$$

Check out for the explanations in the equation 2.11. Remember the logic of Hicksian demand that holding utility level constant, how price changes affect the demand. So $\partial h_l(p, u)/\partial p_k$ measures the substitution effect (change in demand of commodity l) when the price of commodity k is changed. On the other hand, $\partial x_l(p, w)/\partial p_k$ measures the total effect. The remainder is the measure of income effect.

Again remember the Shephard's lemma 2.6 and the properties of expenditure function that it is concave. By Shephard's Lemma Hicksian Demand is defined as $h_i(p, u) = \partial e(p, u)/\partial p_i$ and therefore $\partial h_i(p, u)/\partial p_i$ can be written as $\partial^2 e(p, u)/\partial p_i^2$. Therefore concavity of the expenditure function implies negativeness of own-substitution:

$$\frac{\partial h_i(p, u)}{\partial p_i} = \frac{\partial^2 e(p, u)}{\partial p_i^2} \leq 0 \quad \forall i = 1, 2, 3..., n \tag{2.12}$$

This property called the *Negative Own-Substitution Terms* (Jehle & Reny, 2011). Let's define the substitution matrix:

$$\sigma(p,u) = \begin{bmatrix} \frac{\partial h(p,u)}{\partial p_1} & \cdots & \frac{\partial h(p,u)}{\partial p_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial h(p,u)}{\partial p_1} & \cdots & \frac{\partial h(p,u)}{\partial p_n} \end{bmatrix} \tag{2.13}$$

All terms in the diagonal matrix are negative as explained in the equation 2.12. Therefore, substitution matrix is negative semi-definite. Also substitution terms are symmetric such that (Jehle & Reny, 2011, Theorem 1.15):

$$\partial h_i(p,u)/\partial p_j = \partial h_j(p,u)/\partial p_i \tag{2.14}$$

Let's rewrite the Negative Own-Substitute Terms matrix of Walrasian demand which is called the Slutsky matrix:

$$s(p,u) = \begin{bmatrix} \frac{\partial x_1(p,m)}{\partial p_1} + \frac{\partial x_1(p,m))}{\partial m} x_1(p,m) & \cdots & \frac{\partial x_1(p,m)}{\partial p_n} + \frac{\partial x_1(p,m))}{\partial m} x_n(p,m) \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n(p,m)}{\partial p_1} + \frac{\partial x_n(p,m))}{\partial m} x_1(p,m) & \cdots & \frac{\partial x_n(p,m)}{\partial p_n} + \frac{\partial x_n(p,m))}{\partial m} x_n(p,m) \end{bmatrix} \tag{2.15}$$

By definition of Slutsky equation Slutsky matrix have the same properties with substitution matrix. Therefore Slutsky matrix is negative semidefinite and symmetric as substitution matrix. (Jehle & Reny, 2011, Theorem 1.16)

Recall the Slutsky equation 2.11 and follow the steps:

$$\begin{aligned} \frac{\partial x_l(p,m)}{\partial p_k} &= \frac{\partial h_l(p,u)}{\partial p_k} - x_k(p,m)\frac{\partial x_l(p,m)}{\partial m} \\ \frac{\partial x_l(p,m)}{\partial p_k} p_k &= \frac{\partial h_l(p,u)}{\partial p_k} p_k - \frac{p_k x_k(p,m)}{m}\frac{\partial x_l(p,m)}{\partial m} m \end{aligned} \tag{2.16}$$

The expenditure share of commodity k is defined as follows:

$$\frac{p_k q_k(p,m)}{m} = w_k \tag{2.17}$$

Substituting 2.17 in equation 2.16 and dividing all terms with x_l will result:

$$\frac{\partial x_l(p,m)}{\partial p_k}\frac{p_k}{x_l} = \frac{\partial h_l(p,u)}{\partial p_k}\frac{p_k}{x_l} - w_k \frac{\partial x_l(p,m)}{\partial m}\frac{m}{x_l}$$

The result is, respectively, the relation between uncompensated price elasticity (ϵ_{lk}^{W}), compensated price elasticity (ϵ_{lk}^{H}), expenditure share (w_k) and income elasticity (μ_l):

$$\epsilon_{lk}^{W} = \epsilon_{lk}^{H} - w_k \mu_l \tag{2.18}$$

2.2 Demand Systems

2.2.1 Almost ideal demand estimation

Deaton & Muellbauer (1980) had proposed a demand system that is compatible with the neoclassical demand theory called the Almost Ideal Demand System. AIDS and its extensions are one of the most popular demand systems used in the literature. So far, basic concepts of the neoclassical demand theory are explained. This section elaborates on the AID model from given perspective of Neoclassical Demand Theory.

The derivation of the AIDS is based on the PIGLOG preferences defined by (Muellbauer, 1975). Of course, the PIGLOG preferences treat market demand as if they are the outcome of rational consumers' decisions. PIGLOG is represented by an expenditure function in form of:

$$\ln e(p, u) = (1 - u) \ln a(p) + u \ln b(p) \tag{2.19}$$

The utility level u lies between 0 and 1, representing the "subsistence" and "bliss" levels respectively. Therefore $a(p)$ is referred as cost of subsistence, and $b(p)$ is referred as cost of bliss. The terms, $a(p)$ and $b(p)$ are positive linearly homogeneous functions and their functional form must allow expenditure function to be flexible. It it must hold the properties given and must be differentiable. Deaton & Muellbauer (1980) take the functional forms for $a(p)$ and $b(p)$ to be:

$$\ln a(p_t) = a_0 + \sum_i a_i \ln p_{it} + \frac{1}{2} \sum_i \sum_j \gamma_{ij}^{*} \ln p_{it} \ln p_{jt} \tag{2.20}$$

$$\ln b(p_t) = \ln a(p_t) + \beta_0 \prod_k p_k^{\beta_k} \tag{2.21}$$

Where γ^{*} measures the cross-price effects.

Substituting the equations 2.20 and 2.21 inside the PIGLOG preferences in equation 2.19 gives the expenditure function:

$$\ln e(p_t, u_t) = \underbrace{a_0 + \sum_i \ln p_{it} + \frac{1}{2}\sum_i \sum_j \gamma^*_{ij} \ln p_{it} \ln p_{jt}}_{\ln a(p_t)} + u_t \underbrace{\beta_0 \prod_i p_{it}^{\beta_i}}_{\ln b(p_t) - \ln a(p_t)} \tag{2.22}$$

Applying *Shephard's Lemma* (see equation 2.6) to the expenditure function gives the Hicksian Demand, $h(p, u)$. If u is replaced with indirect utility function $v(p, m)$ then Hicksian Demand transforms into Walrasian Demand as explained in equation 2.7. The indirect utility can be derived from the equation 2.22 by leaving u alone:

$$u = v(p, m) = \frac{\ln m - \ln a(p)}{\ln b(p) - \ln a(p)} \tag{2.23}$$

Where $e(p_t, u_t)$ is treated as an exogenous variable represented with m. By applying Shephard's lemma in logarithmic functional forms will give the *expenditure shares* as shown:

$$\frac{\partial \ln e(p, u)}{\partial \ln p_i} = \frac{p_i q_i}{e(p, u)} = w_i \tag{2.24}$$

From the equation 2.19 w_i is:

$$\begin{aligned}
w_i &= \frac{\partial \ln a(p)}{\partial \ln p_i} - u\frac{\partial \ln a(p)}{\partial \ln p_i} + u\frac{\partial \ln b(p)}{\partial \ln p_i} \\
&= \frac{\partial \ln a(p)}{\partial \ln p_i} - u\frac{\partial \ln a(p)}{\partial \ln p_i} + u\frac{\partial(\ln a(p) + \beta_0 \prod_i p_{it}^{\beta_i})}{\partial \ln p_i} \\
&= \frac{\partial \ln a(p)}{\partial \ln p_i} - u\frac{\partial \ln a(p)}{\partial \ln p_i} + u\frac{\partial \ln a(p)}{\partial \ln p_i} + u\frac{\partial(\beta_0 \prod_i p_{it}^{\beta_i})}{\partial \ln p_i} \\
&= \frac{\partial \ln a(p)}{\partial \ln p_i} + u\frac{\partial(\beta_0 \prod_i p_{it}^{\beta_i})}{\partial \ln p_i} \\
&= a_i + \sum_j \gamma_{ij} \ln p_{jt} + \beta_i u \beta_0 \prod_i p_{it}^{\beta_i}
\end{aligned}$$

where $\gamma_{ij} = \frac{1}{2}(\gamma^*_{ij} + \gamma^*_{ji})$. The functional form above represents the expenditure share for Hicksian Demand. By substituting u with indirect utility given in equation 2.23 expenditure share for Walrasian Demand is derived:

$$\begin{aligned} w_i &= a_i + \sum_j \gamma_{ij} \ln p_j + \beta_i \Big(\frac{\ln m - \ln a(p)}{\ln b(p) - \ln a(p)} \beta_0 \prod_i p_{it}^{\beta_i} \Big) \\ &= a_i + \sum_j \gamma_{ij} \ln p_j + \beta_i \Big(\frac{\ln m - \ln a(p)}{\beta_0 \prod_i p_{it}^{\beta_i}} \beta_0 \prod_i p_{it}^{\beta_i} \Big) \\ &= a_i + \sum_j \gamma_{ij} \ln p_j + \beta_i \ln \Big(\frac{m_i}{a(p)} \Big) \end{aligned}$$

Given derivations result the AIDS model:

$$w_i = a_i + \sum_j \gamma_{ij} \ln p_j + \beta_i \ln \Big(\frac{m_i}{a(p)} \Big) \tag{2.25}$$

There are some constraints for consistency originating from the Neoclassical Demand Theory.

By intuition summation of all expenditure shares must be 1. Adding-up conditions ensure that $\sum_i w_i = 1$:

$$\sum_i a_i = 1\,; \quad \sum_i \beta_i = 0\,; \quad \sum_i \gamma_{ij} = 0 \quad \forall j \tag{2.26}$$

The rational consumer's decisions are boxed on the relative prices not the nominal prices which requires the imposition of given constraint (Recall the Homogeneity Degree of Zero property):

$$\sum_j \gamma_{ij} = 0 \quad \forall i \tag{2.27}$$

However applicability of the Shephard's lemma depends on the *Slutsky symmetry* condition. Therefore, cross-price effect parameters, γ, must be symmetric:

$$\gamma_{ij} = \gamma_{ji} \quad \forall\, i, j \tag{2.28}$$

With given constraints (2.26, 2.27, and 2.28) and w_i the econometric model can be estimated:

$$w_{it} = a_i + \sum_j \gamma_{ij} \ln p_{jt} + \beta_i \ln \left(\frac{m_i t}{a(p_t)} \right) + err_{it} \tag{2.29}$$

However the econometric model above is nonlinear because of translog price index ($a(p_t)$). To make linear econometric estimations Deaton & Muellbauer (1980) use a linear price parameter, Stone's Price index instead of $a(p)$:

$$\ln a(p_t^S) = \sum w_{it} \ln p_{it} \tag{2.30}$$

Deaton & Muellbauer (1980) empirically show that Stone's Price Index approximates the translog price index with given equation 2.30. By replacing the translog price with the price index, they derived LA-AIDS econometric model to make estimations easier:

$$w_{it} = a_i^S + \sum_j \gamma_{ij}^S \ln p_{jt} + \beta_i^S \ln \left(\frac{m_i t}{a(p_t^S)} \right) + err_{it}^S \tag{2.31}$$

Including the $a(p_t^S)$ to the econometric model instead of $a(p_t)$ causes endogeneity by placing w_{it} both sides of the equation. Moschini (1995) proposes alternative price indexes: Paasche, Laspeyres, and Tornqvist. It looks like Laspeyres index relatively performs better than others but it still remains as a controversial issue (Henningsen, 2017) in three aspects: (1) Units of measurement changes the results, (2) there is approximation errors, and (3) theoretical inconsistency of LA-AIDS[5].

Moving away from the debates, it is unnecessary to use LA-AIDS, because advances in software technology allow to solve nonlinear econometric models easily. One can use translog price, $a(p_t)$, non-linearly without external price indices.

One of the very important output of the demand estimations are the elasticities. It gives important clues about the consumer behaviours. Walrasian (Uncompensated) price elasticity of demand is:

[5]Theoretically, demand functions must be integrable that yield a utility function. Under certain conditions, LaFrance (2004) solved the integrability problem of LA-AIDS.

$$\epsilon_{ij}^{W} = \frac{\partial q_i}{\partial p_j}\frac{p_j}{q_i} \tag{2.32}$$

The expansion of the expenditure share (equation 2.17) contains the quantity variable, q_i, therefore AIDS model (equation 2.25) for commodity i can be written as follows:

$$q_i = \frac{m_i}{p_i}\left[a_i + \sum_j \gamma_{ij}\ln p_j + \beta_i \ln\left(\frac{m_i}{a(p)}\right)\right] \tag{2.33}$$

Putting AIDS quantity, q_i, into the the equation 2.32:

$$\begin{aligned}
\frac{\partial q_i}{\partial p_j}\frac{p_j}{q_i} &= \frac{m_i p_j}{p_i q_i}\frac{\partial\left[a_i + \sum_j \gamma_{ij}\ln p_j + \beta_i \ln\left(\frac{m_i}{a(p)}\right)\right]}{\partial p_j} \\
&= \frac{p_j}{w_i}\left[\frac{\gamma_{ij}}{p_j} - \beta_i \frac{\partial \ln a(p)}{\partial p_j}\right] \\
&= \frac{p_j}{w_i}\left[\frac{\gamma_{ij}}{p_j} - \frac{\beta_i}{p_j}\left(a_j + \sum_k \gamma_{ik}\ln p_k\right)\right] \\
&= \frac{1}{w_i}\left[\gamma_{ij} - \beta_i\left(a_j + \sum_k \gamma_{ik}\ln p_k\right)\right] \\
&= \frac{\gamma_{ij} - \beta_i\left(a_j + \sum_k \gamma_{ik}\ln p_k\right)}{w_i}
\end{aligned}$$

If $i = j$ then the own price elasticity becomes:

$$\epsilon_{ij}^{W} = \frac{\gamma_{ij} - \beta_i\left(a_j + \sum_k \gamma_{ik}\ln p_k\right)}{w_i} - 1 \tag{2.34}$$

The own-price elasticity is has -1 additionaly because of extra derivations in $\ln a(p)$. Through the Slutsky equation compensated price elasticity (see equation 2.18) can be calculated.

The income elasticity calculated in the same way:

$$\mu_i = \frac{\partial q_i}{\partial m_i}\frac{m_i}{q_i} = \frac{m_i}{q_i}\frac{\partial\left[\frac{m_i}{p_i}\left(a_i + \sum_j \gamma_{ij}\ln p_j + \beta_i \ln\left(\frac{m_i}{a(p)}\right)\right)\right]}{\partial m_i}$$
$$= \frac{m_i}{q_i}\left(\frac{w_i}{p_i} + \frac{m_i}{p_i}\frac{\beta_i}{m_i}\right)$$
$$= 1 + \frac{\beta_i}{w_i} \tag{2.35}$$

2.2.2 Quadratic almost ideal demand estimation

The functional form of AIDS assumes a linear income effect, in other words, assumes that Engel curve's shape is linear. However, increase or decrease rate of the demand for some goods may not be linear in terms of income. In fact as Banks, Blundell, & Lewbel (1997) shows Engel curves' shapes are nonlinear in most cases. AIDS becomes an appropriate demand system if and only if all commodities (or commodity groups) have linear Engel Curves. Otherwise, it would lead to un-robust estimations.

According to Banks et al. (1997), the simplest form of expenditure function consistent with the empirically observed Engel curves is as follows:

$$w_i = A_i(p) + B_i(p)\ln x + C_i(p)g(x) \tag{2.36}$$

where $g(x)$ is some continuous function. Banks et al. (1997) states that the only indirect utility function that is consistent with the expenditure function given as the equation 2.36 is:

$$\ln V(p,m) = \left[\left\{\frac{\ln m - \ln a(p)}{b(p)}\right\}^{-1} - \lambda(p)\right]^{-1} \tag{2.37}$$

The functional form of the translog price is the same with the equation 2.20. However, the price aggregator, $b(p)$, has a different functional form than AIDS. There is no difference in terms of mathematical model (Compare with the equation 2.21):

$$b(p) = \prod_{i=1}^{k} p_i^{\beta_i} \tag{2.38}$$

The function λ makes the demand system quadratic which is differentiable and homogeneous degree of zero in p:

$$\lambda(p) = \sum_{i=1}^{k} \lambda_i \ln p_i \tag{2.39}$$

In addition to the constraints of AIDS (The adding-up condition: 2.26, Homogeneity condition: 2.27, and Symmetry condition 2.28) the quadratic function have another constraint:

$$\sum_{i=1}^{k} \lambda_i = 0 \tag{2.40}$$

Instead of applying Shephard's Lemma as in the AIDS, Roy's Identity (equation 2.4) can be used to transform the indirect utility (equation 2.37) into the Walrasian demand. Applying indirect utilities in logarithmic form yields the expenditure shares as follows:

$$= -\frac{\partial \ln v(p,m)/\partial \ln p_i}{\partial \ln v(p,m)/\partial \ln m_i} = -\frac{\partial v(p,m)/\partial p_i}{\partial v(p,m)/\partial m} \frac{p_i}{m_i} = w_i \tag{2.41}$$

To solve the Roy's Identity equation 2.37 is simplified:

$$\ln v(p,m) = \frac{\ln x}{\ln x \lambda(p) + b(p)} \tag{2.42}$$

Expanding the equation 2.41 yields (see Appendix A for steps):

$$w_i = a_i + \sum_k \gamma_{ik} \ln p_k + \beta_i ln(x) + \frac{\lambda_i}{b(p)} \ln^2 x \tag{2.43}$$

Knowing that $\ln x = \ln m - \ln a(p) = \ln\left(\frac{m}{a(p)}\right)$, the QUAIDS equation is:

$$w_i = a_i + \sum_j \gamma_i \ln p_j + \beta_i \ln\left(\frac{m}{a(p)}\right) + \frac{\lambda_i}{b(p)} \left[\ln\left(\frac{m}{a(p)}\right)\right]^2 \tag{2.44}$$

Mind the difference between AIDS and QUAIDS (see equations 2.44 and 2.29). Income parameters became quadratic with new variables. This implies if $\lambda_i = 0\ \forall i$ then Engel curve is linear. That is why AIDS is a special case of QUAIDS.

Partial derivative of equation 2.44 respect to $\ln m$ and $\ln p_j$ will give price and expenditure elasticity respectively as in equation 2.35 :

$$\mu_i = \beta_i + \frac{2\lambda_i}{b(p)}\left(\ln\left[\frac{m}{a(p)}\right]\right) \tag{2.45}$$

$$\epsilon_{ij} = \gamma_{ij} - \mu_i\Big(a_j + \sum_k \gamma_{jk}\ln p_k\Big) - \frac{\beta_j\lambda_i}{b(p)}\left[\ln\left(\frac{m}{a(p)}\right)\right]^2 \tag{2.46}$$

The uncompensated demand elasticity is derived with equation 2.18 as done for AIDS.

2.2.3 Demographic variables

The demand for any good, apart from the income and prices, depends on some personal/household characteristics. Two households, even if they have the same income level, could behave in different manner depending on demographic differences. However, these household characteristics are not accounted in the AI or QUAI demand systems. Thereby, Ray (1983) transformed the expenditure function to account for the demographic variables:

$$e(p, z, u) = m_0(p, z, u) \times e^R(p, u) \tag{2.47}$$

where $e^R(p, u)$ is the reference expenditure function scaled by the demographic variables, $m_0(p, z, u)$. The scale function consists of two components:

$$m_0(p, z, u) = \bar{m}_0(z) \times \phi(p, z, u) \tag{2.48}$$

Ray (1983) defined the $\bar{m}_0(z)$ as:

$$\bar{m}_0(z) = 1 + \rho' z \tag{2.49}$$

z is the vector of demographic variables and ρ is the vector of parameters to be estimated that represents the demographic characteristics. ρ measures the linear/average increase in household expenditure and does not consider the different consumption patterns. The tricky part is the functional form of $\phi(p, z, u)$. Applying Shepard's Lemma (2.6), to the equation 2.47 yields:

$$\begin{aligned} w_{it} &= \frac{\partial \ln e(p, z_t, u_t)}{\partial \ln p_i} \\ &= \frac{\partial \bar{m}_0(z_t)}{\partial \ln p_i} + \frac{\partial \phi(p, z_t, u_t)}{\partial \ln p_i} + \frac{\partial \ln e^R(p, u_t)}{\partial \ln p_i} \\ &= w_i^R(u, p) + \frac{\partial \phi(p, z_t, u_t)}{\partial \ln p_i} \end{aligned} \tag{2.50}$$

So the second term, $\phi(.)$, expresses the expenditure differences between households via demographic characteristics. Functional form of $\phi(.)$ must be homogeneous degree of zero in prices and must have non-negative values (Ray, 1983). I use the functional form from Poi (2002):

$$\ln \phi(p, z_t, u_t) = \frac{\prod_{j=1}^{k} p_j^{\beta_j} (\prod_{j=1}^{k} p_j^{\eta_j' z} - 1)}{1/u - \sum_{j=1}^{k} \lambda_j \ln p_j} \tag{2.51}$$

After specifying the functional form of $\phi(.)$, w_{it} can be represented in the vector form, w_i, as in the equation 2.50:

$$w_i = a_i + \sum_j \gamma_{ij} \ln p_j + (\beta_i + \eta_i' z) \ln \left(\frac{m}{\bar{m}_0(z) a(p)} \right) + \frac{\lambda_i}{b(p) c(p, z)} \left[\ln \left(\frac{m}{\bar{m}_0(z) a(p)} \right) \right]^2 \tag{2.52}$$

The parameter $c(p, z)$ is:

$$\prod_{j=1}^{k} p_j^{\eta_j' z}$$

After employing the demographic characteristics we have another constraint on the top of that $\sum_{j=1}^{k} \eta_{rj} = 0$ for all demographic characteristics $r = 1, 2, ..., s$.

Elasticities are calculated in the same way with AI and QUAI demand systems. Taking the partial derivatives of the equation 2.52 with respect to $\ln m$ and $\ln p_j$, respectively, yields:

$$\mu_i = \beta_i + \eta_i' z + \frac{2\lambda_i}{b(p)c(p,z)}\left(\ln\left[\frac{m}{\bar{m}_0(z)a(p)}\right]\right) \tag{2.53}$$

$$\begin{aligned}\epsilon_{ij} = \gamma_{ij} - &\left(\beta_i + \eta_i' z + \frac{2\lambda_i}{b(p)c(p,z)}\Big(\ln\Big[\frac{m}{\bar{m}_0(z)a(p)}\Big]\Big)\right)\left(a_j + \sum_k \gamma_{jk}\ln p_k\right) \\ &- \frac{(\beta_j + \eta_j' z)\lambda_i}{c(p,z)b(p)}\Big[\ln\Big(\frac{m}{\bar{m}_0(z)a(p)}\Big)\Big]^2\end{aligned} \tag{2.54}$$

3. DATA

This thesis estimates the food demand using Turkish Statistical Institute's Household Consumption Surveys between 2008-2016 (TURKSTAT, n.d.). TURKSTAT makes household surveys plannings every year. The period for the surveys are one month. So at the beginning of the year they plan to make surveys with $12n$ households, and every month they make surveys with n families.

3.1 Categorization

TURKSTAT (2015) uses the COICOP classification system, Classification of Individual Consumption by Purpose. There are 12 categories of expenditure survey: (1) Food and non-alcoholic beverages; (2) Alcoholic beverages, tobacco and narcotics, (3) Clothing and footwear; (4) Housing, water, electricity, gas and other fuels; (5) Furnishings, household equipment and routine household maintenance; (6) Health; (7) Transport; (8) Communication; (9) Recreation and culture; (10) Education; (11) Restaurants and hotels; (12) Miscellaneous goods and services.

There are 4-levels in each categories. I will use the first category's subcategories. The first level within the Food and non-alcoholic beverages category have two subcategories: (1) Food and (2) Non-alcoholic beverages. Only the first subcategory, Food, is used for this research. Food subcategory consists of 9 food groups:

1. Bread and cereals
2. Meat
3. Fish and seafood
4. Milk, cheese, and eggs
5. Oils and fats
6. Fruit
7. Vegetables
8. Sugar, jam, honey, chocolate, and confectionery

9. Food products n.e.c.

For instance all goods numbered 111x are in the Bread and cereals category or 112x are in the Meat category.

I aggregate the expenditures in the third level. In other words, there are 9 food groups to be analyzed. Apart form the expenditure data, household and individual characteristics are recorded in separate databases. These databases contains information like household size, number of rooms, if an individual smokes within that household, income of the household and individuals, and so on.

TURKSTAT surveys specific number of households monthly where the number is predetermined for that year. By construction, data is cross-sectional but it can be transformed into time series data by aggregating the households within the same month and year. There are more than 100,000 surveys done between 2008 and 2016, however transforming them into Time Series data results in 108 observations. There are two main reason justifying this transformation: (1) Zero-consumption causes heterogeneity and it requires to use censored demand methods as Akbay et al. (2007) did; (2) Household surveys do not contain the prices of expenditures separately but only the total expenditure is available. Prices of the goods can only be obtained through TURKSTAT's price index, which is recorded monthly.

There are pros and cons of this transformation. The worst consequence is the decrease in the number of observations. Secondly, some of the demographic characteristics cannot be parameterized or causes restrictions on their use. However, one of the major problems of the demand estimation, heterogeneity, is reduced by aggregation. Another obvious advantage is that it can be estimated much faster. Because nonlinear estimations require much more time than LA/AIDS.

3.2 Data Manipulation

TURKSTAT used, specifically, the COICOP_HBS classification system between 2002 and 2014. However they made minor changes and used a different classification system named COICOP (V.2011) after 2015 (TURKSTAT, n.d.). Therefore combining the surveys with different classification system created minor inconsistencies within the groups. Particularly, ”Food products n.e.c. ” category suffers from that inconsistency; because it has relatively smaller share in expenditure budget. Also some new products were included and the definitions had been changed, therefore, in both the expenditure shares and the price levels there are sudden jumps and drops. Consequently, this category is dropped from the demand estimation.

Estimations give meaningful results with the remaining 8 categories. However, with 8 categories, QUAIDS model (without demographic variables) estimates 60 parameters. Meanwhile including n demographic variables increases the number of parameters by $n \times 8 + n$. This might cause some econometric problems for estimation like identification problem. Herewith, the number of groups are reduced to 6 by combining the categories Meat and Fish; and categories Fruit and Vegetables. Thus, I estimate food demand for six categories: (1) Bread and Cereals, (2) Meat and Fish, (3) Milk, Cheese and eggs, (4) Oils and Fats, (5) Fruits and Vegetables, and (6) Sugar, jam, honey, and others.

3.3 Prices

I received the price levels from TURKSTAT. The prices for the commodities are given with the COICOP classification system. Thus, price levels and household survey commodity groups are consistent. However, no prices were given for the food groups. Thus, I estimate the weighted prices for each food category. Figure 3.1 implies:

- Price levels had increased for some commodities faster than others like "Meat and Fish" and "Oils and fats".
- Prices increase rate of "Fruit and Vegetables" seems increased since 2012.
- Seasonality can be observed for "Meat and Fish" and "Fruit and Vegetables" categories.

3.4 Descriptive Statistics

Within the 6 food categories, "Meat and Fish" have the highest variance as seen in table 3.1. There are two major reasons: (1) Seasonality of the fish prices, and (2) Meat consumption during the Sacrifice Feasts. Second reason has significant impacts on expenditure shares, which can be observed in the average monthly expenditure per individual (see peaks in the figure 3.3). It appears that jumps of "Meat and Fish" and "Sugar, jam, honey, chocolate, and confectionary" consumption is the reason of increase in expenditures during the Sacrifice feasts (See figures 3.2 and 3.3). The consumption of other groups does not decrease nominally during the sacrifice feasts but their expenditure shares decrease. Consequently, the sacrifice-feast-effect cannot be neglected, because

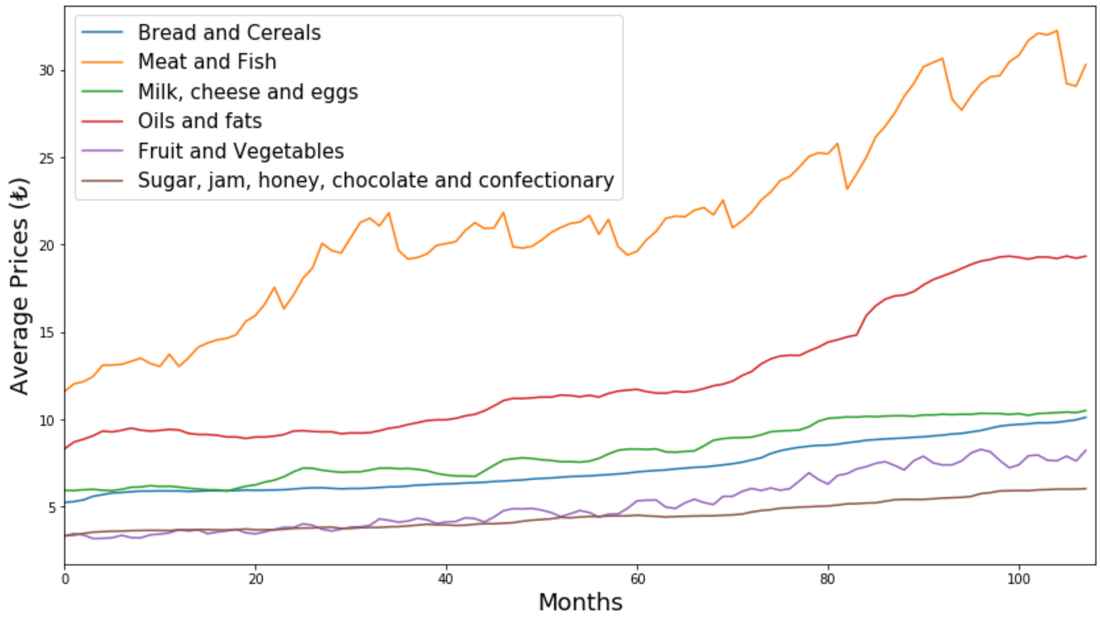

Figure 3.1: Weighted Price Levels for food groups (monthly)

it leads to too much disturbance in the data. This is why this dates will be employed as dummy variable in the demographic characteristics. The existing studies have not controlled the sacrifice feast effect. Though sacrifice feast is not coherent with the Gregorian calendar, date changes so slowly in years that its effect on consumer behaviour is pretty much compensated by other seasonality parameters like months or seasons in the existing studies. However, sacrifice feast effect and seasonality must be distinguished. Otherwise seasonality parameters become biased.

Table 3.1: Descriptive Statistics of Food Groups

Variable	Mean	Std. Dev.	Min	Max
Bread and Cereals	0.23	0.03	0.13	0.28
Meat and Fish	0.19	0.08	0.12	0.49
Milk, cheese and eggs	0.16	0.02	0.09	0.20
Oils and fats	0.06	0.01	0.04	0.08
Fruit and Vegetables	0.29	0.03	0.18	0.34
Sugar, jam, honey, ...	0.08	0.01	0.06	0.11

As stated in the many reports, Turkish population is aging (TC KALKINMA BAKANLIĞI, 2014) which can be observed from two aspects in the household surveys. Firstly, the average age is increasing [6]. And secondly, the average household size is decreasing every year (see Figure 3.2). Age is used as demographic variable in many demand

[6]The years 2008 and 2009 excluded from the graphs because age is recorded as categorical variable within those years.

system analysis (Akbay et al., 2007; Dybczak et al., 2014; Sengul & Tuncer, 2005), but I do not use it because of inconsistency between years. The explanatory effect of the age variable can be partially compensated by the household size variable. There is a significant correlation between average household size and the average age (see figure 3.2)[7].

[7] The correlation between the age and average household size is -0.7817 after 2010.

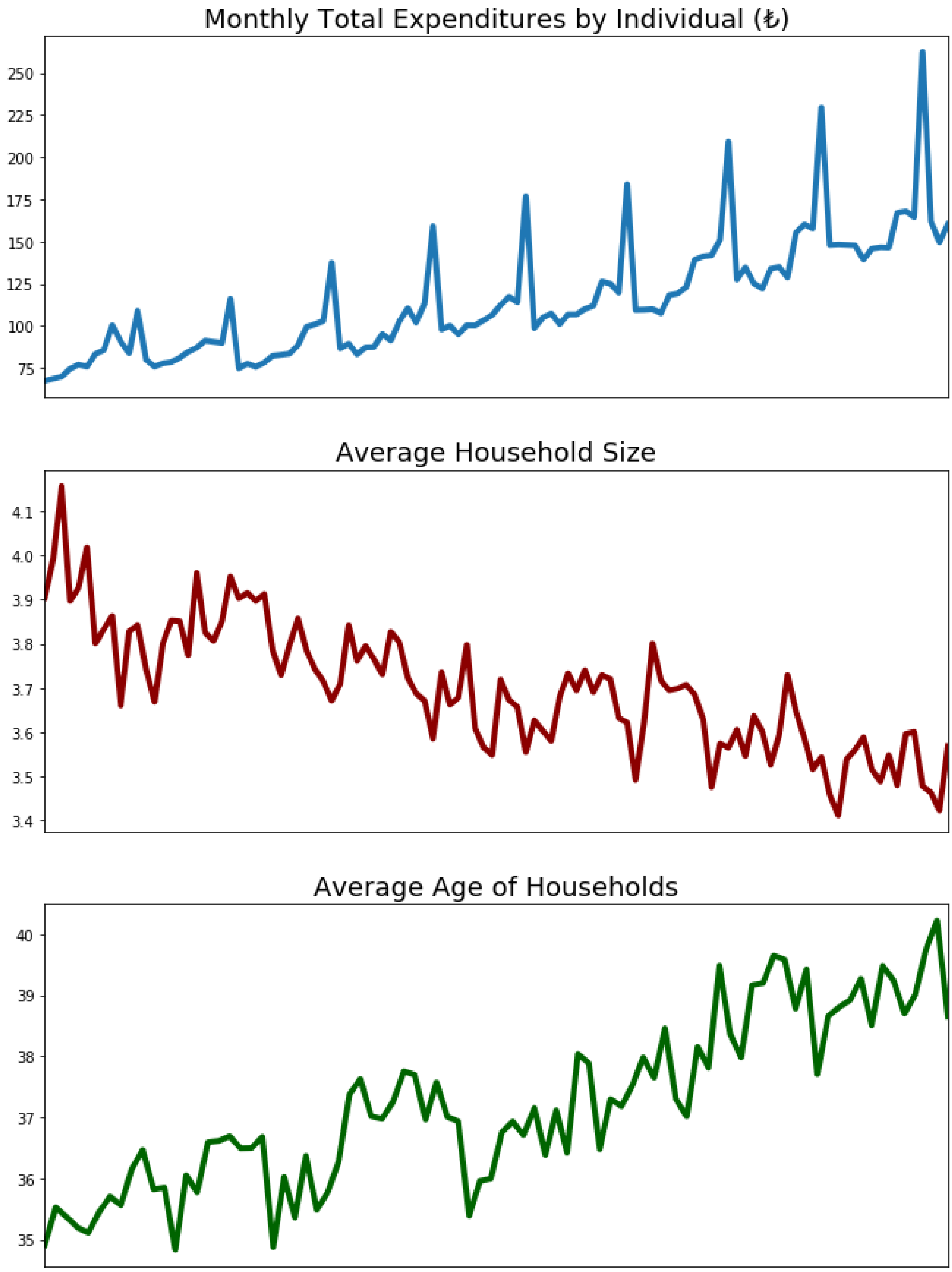

Figure 3.2: Changes in same descriptive parameters between 2008 and 2016

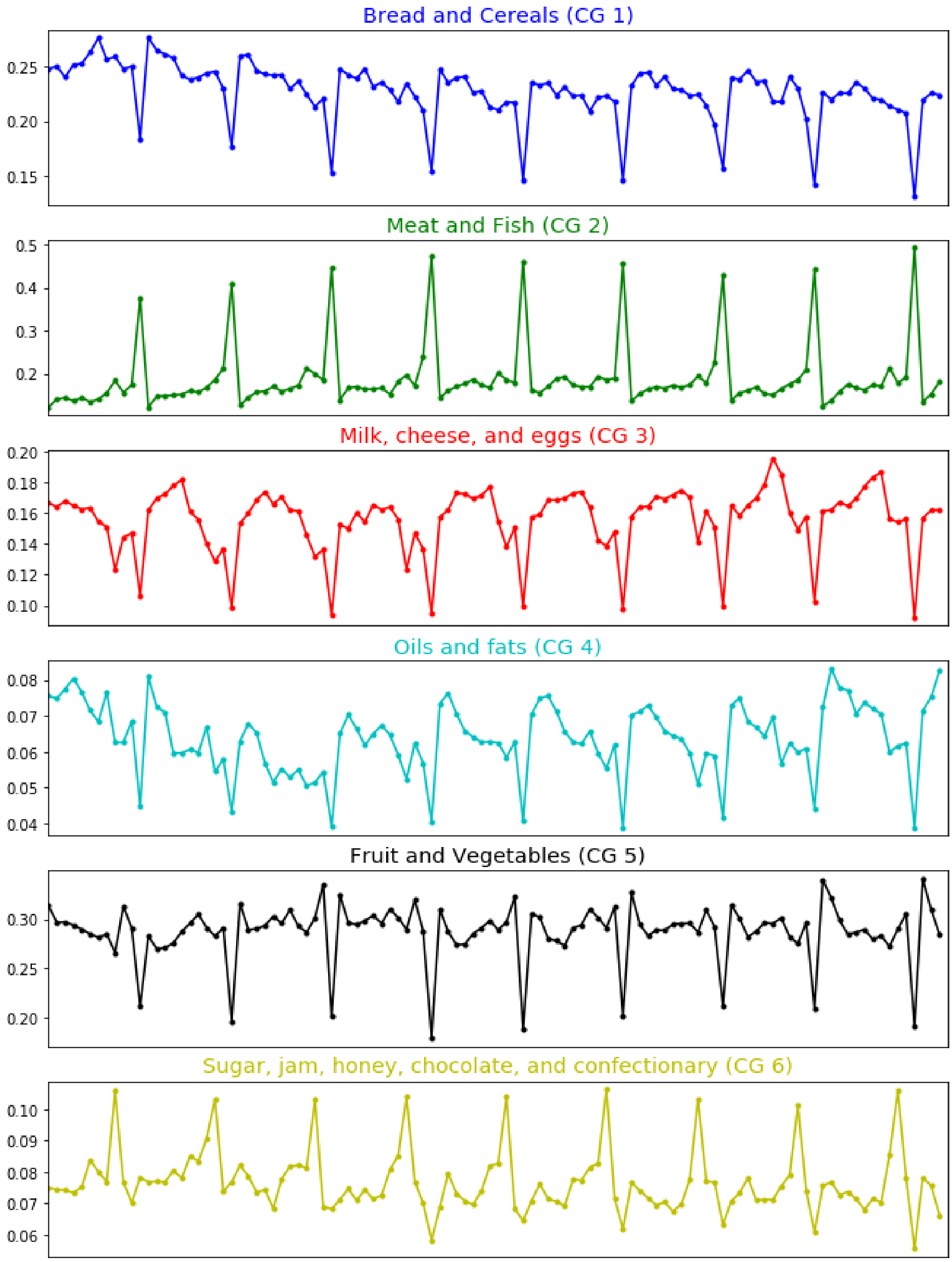

Figure 3.3: Monthly Expenditure Shares of Commodity Groups between 2008 and 2016

4. ECONOMETRIC ESTIMATION

As explicated in the previous sections I estimate the QUAIDS model by employing the Ray's 1983 expenditure function with demographic variables. I use Poi's (2012) user-written STATA package, "quaids", to estimate the equation 2.52.

Table 4.1 gives the R^2 values of predictions for each commodity groups. Apparently, the model cannot explain the variation well in the Commodity Group 6. So, further attention should be given for interpretations related with the Commodity Group 6.

Table 4.1: R^2 of Commodity Groups predictions

CG 1	CG 2	CG 3	CG 4	CG 5	CG 6
89%	97%	91%	82%	82%	51%

Three demographic variable are used in the estimation. The first demographic variable is the average household size. Household size is expected to have an effect on both the amount of expenditure and the expenditure behaviour of households. Secondly, a dummy variable for the date (months) of the sacrifice feasts is added. It is the most important demographic variable for the robustness of the demand estimation. As seen in the Figure 3.3 household behaviour changes dramatically during this months. Thirdly, the seasonality dummy variable is added to separate the seasonal expenditure behaviours. Ideally, there should be 11 seasonality variable for the months. However, it leads to $6 \times 11 + 1$ extra parameters to estimate. The given estimation cannot be performed with 108 observations. So, seasonality variable comprises the Fall and Winter months as dummy alone.

4.1 Estimation of the Model

The α coefficients are not meaningful alone which could be negative either.

β_i coefficients give intuitions about the income effect. As income increases expenditure shares of commodity groups 1, 2, and 6 decrease; commodity groups 2, 4, and 5 decrease. Interpreting the λ, and β coefficients together gives further and accurate intuition. For instance, interpretation of the commodity group 2 is reversed, because the quadratic term has a considerable and bigger value with positive sign. This means as income increases commodity group 2's expenditure share increases quadratically. However, commodity groups 1, 3, 5, and 6 have quadratic term that are statistically not significant. Therefore, their Engel curves are linear. Commodity group 4's expenditure share increases as income increases but in a diminishing rate and even decreasing after a point.

Table 4.2: QUAIDS estimation coefficients and their standard errors

Coef.	CG 1	CG 2	CG 3	CG 4	CG 5	CG 6
α_i	0.3160	0.0893***	0.2500***	0.0946***	0.2310***	0.0191
	(0.024)	*(0.022)*	*(0.026)*	*(0.015)*	*(0.024)*	*(0.030)*
β_i	-0.1394	-0.1647	0.0158	0.1102**	0.3181**	-0.1400
	(0.096)	*(0.147)*	*(0.088)*	*(0.044)*	*(0.154)*	*(0.098)*
γ_{i1}	0.1436***	-0.0688***	-0.0038	-0.0309***	-0.0145	-0.0256
	(0.030)	*(0.009)*	*(0.017)*	*(0.011)*	*(0.013)*	*(0.025)*
γ_{i2}	-0.0688***	0.0944***	-0.0310***	-0.0372***	0.0367***	0.0058
	(0.009)	*(0.014)*	*(0.007)*	*(0.004)*	*(0.012)*	*(0.007)*
γ_{i3}	-0.0038	-0.0310***	-0.0136	0.0067	0.0170	0.0247
	(0.017)	*(0.007)*	*(0.020)*	*(0.010)*	*(0.011)*	*(0.018)*
γ_{i4}	-0.0309***	-0.0372***	0.0067	0.0366***	-0.0019	0.0267**
	(0.011)	*(0.004)*	*(0.010)*	*(0.009)*	*(0.008)*	*(0.012)*
γ_{i5}	-0.0145	0.0367***	0.0170	-0.0019	-0.0301*	-0.0072
	(0.013)	*(0.012)*	*(0.011)*	*(0.008)*	*(0.017)*	*(0.014)*
γ_{i6}	-0.0256	0.0058	0.0247	0.0267**	-0.0072	-0.0245
	(0.025)	*(0.007)*	*(0.018)*	*(0.012)*	*(0.014)*	*(0.029)*
λ_i	-0.0639	0.2154***	0.0006	-0.0849***	-0.0735	0.0063
	(0.041)	*(0.062)*	*(0.047)*	*(0.032)*	*(0.055)*	*(0.037)*
$\eta_{i,sac}$	-0.0628	0.2576*	-0.0156	-0.0195	-0.1271**	-0.0326***
	(0.039)	*(0.139)*	*(0.025)*	*(0.025)*	*(0.053)*	*(0.007)*
$\eta_{i,hhs}$	0.0283	0.0510	-0.0495**	-0.0225*	-0.0734*	0.0661***
	(0.023)	*(0.040)*	*(0.023)*	*(0.012)*	*(0.039)*	*(0.025)*
$\eta_{i,season}$	0.0479***	-0.1015***	0.0637**	0.0287***	0.0245	-0.0632***
	(0.017)	*(0.034)*	*(0.029)*	*(0.008)*	*(0.017)*	*(0.020)*

Table 4.3 summarizes the scalar coefficients for the demographic variables (recall the equation 2.49).

Table 4.3: Scalar Coefficients

Scalars (ρ)	Coef.	Std. Err.	z	P>z
Sacrifice Feast months	0.0117	0.04	0.288595	0.968
Fall and Winter Seasons	0.2889	5.38	0.053692	0.00
Average Household Size	-0.4600	-3.36	0.136847	0.001

4.2 Tests

QUAIDS have only one extra parameter in addition to AIDS, which is the quadratic term, λ. Thus, there exists 6 additional terms for all the commodity groups. Two models can be compared by looking significance of these terms.

The relevance of demographic variables should be tested as well. These tests can be done either by Wald Test or by Likelihood Ratio test. The likelihood ratio test is more troublesome because it requires a separate estimation (restricted) apart from the main model (unrestricted). However it is easy to make separate estimations in our case.

The Likelihood ratio test is done by comparing the log-likelihood values of restricted and unrestricted estimations as formulated below (Wooldridge, 2013):

$$\text{Likelihood Ratio} = 2(\mathcal{L}_{ur} - \mathcal{L}_r) \tag{4.1}$$

where $\mathcal{L}_{ur}$ is the log-likelihood value of unrestricted model which is 1916.3392 for our econometric model. $\mathcal{L}_r$ is the log-likelihood value of restricted models that we are testing. The idea of the test is that excluding any parameters from the model does not increase the log-likelihood value but it may not decrease either. If there is no significant difference between $\mathcal{L}_r$ and $\mathcal{L}_{ur}$ then it means that excluded terms are not improving the model.

Log-likelihood value of the AIDS model with all demographic variables is 1902.501; QUAIDS without sacrifice feast dummy is 1872.210; QUAIDS without average household size value is 1905.083; and QUAIDS without seasonality dummy is 1863.542.[8] Table 4.4 shows that all tests are statistically significant.

Table 4.4: Likelihood Ratio Test Results

Restriction	$\mathcal{L}_{ur}$	$2(\mathcal{L}_{ur} - \mathcal{L}_r)$	*df*	p-value
Without quadratic term (AIDS)	1902.501	27.6762	6	0.0001
Without sacrifice feast dummy	1872.210	88.258	7	0.0000
Without average household size value	1905.083	22.5124	7	0.0021
Without seasonality dummy	1863.542	105.595	7	0.0000

[8] Demographic variables have the degree of freedom 7 because of the additional scalar term, ρ

4.3 Elasticities

4.3.1 Income elasticity

Table 4.5 refers the income elasticity estimations for commodity groups. All goods are "Normal Goods". "Meat and Fish" and "Sugar, jam, honey, chocolate, and confectionary" have very high income elasticity. Interestingly, "Milk, cheese and eggs" commodity group has very low income elasticity.

Table 4.5: Income Elasticity of Commodity Groups

CG 1	CG 2	CG 3	CG 4	CG 5	CG 6
0.6662	1.8986	0.2251	0.6267	0.9646	1.6453

Therefore, I make a robustness check by estimating the model for different a_0 [9] values between 0 and 15. In total, 80 estimations conducted. Table 4.6 shows that mean of the income elasticity of Commodity group 3 is approximately 0.1605. It ranges between 0.1064 and 0.2474. Also, other parameters do not have too much variance in income elasticity estimations either.

Table 4.6: Income Elasticity Simulation Results with Specification Trials

	CG 1	**CG 2**	**CG 3**	**CG 4**	**CG 5**	**CG 6**
mean	0.6856	1.8956	0.1605	0.5567	0.9963	1.8652
std	0.0178	0.0657	0.0315	0.0287	0.0269	0.0635
min	0.6236	1.5952	0.1064	0.5357	0.9448	1.6354
25%	0.6711	1.8423	0.1457	0.5423	0.9930	1.8338
50%	0.6901	1.8846	0.1483	0.5453	1.0022	1.8471
75%	0.6975	1.9605	0.1927	0.5567	1.0086	1.8758
max	0.7351	1.9655	0.2474	0.6943	1.1169	2.0356

4.3.2 Price elasticity

Table 4.7 shows the uncompensated price elasticities across the commodity groups [10]. High uncompensated price elasticities may imply that the commodity group is luxury. So the commodity groups 3, 5, and 6 can be luxury commodities. As expected "Bread and Cereals" and "Fats and Oil" commodities are less elastic. "Meat and Fish" category has relatively high own-price elasticity.

[9] a_0 is not estimated but is given as an exogenous parameter to make estimation. Its value is given between 0 and 10 by rule of thumb.

[10] Bold numbers show own-price elasticities.

Table 4.7: Uncompensated Price Elasticities

	CG 1	CG 2	CG 3	CG 4	CG 5	CG 6
CG 1	**-0.2992**	-0.2601	0.0350	-0.1169	0.0191	-0.0942
CG 2	-0.5606	**-0.6283**	-0.3082	-0.2534	-0.0506	-0.0268
CG 3	0.1892	-0.0663	**-0.9113**	0.0979	0.3599	0.2089
CG 4	-0.3941	-0.5263	0.1578	**-0.3937**	0.0892	0.4538
CG 5	-0.0448	0.1314	0.0594	-0.0037	**-1.0961**	-0.0212
CG 6	-0.5707	-0.0723	0.1274	0.2862	-0.3724	**-1.3742**

Table 4.8 show the compensated price elasticities. Remember the Hicksian demand (see equation 2.7). The compensated price elasticity holds consumer utility (or real income) constant. In other words, it is an hypothetical measure unlike the uncompensated price elasticity. Thus, one can make an important inference on demand by comparing the compensated and uncompensated demand elasticities: The bigger the difference between them, higher the welfare loss[11]. Because, this means consumers cannot "compensate" the price increases via substitution. "Bread and Cereals", "Fruit and Vegetables", and especially "Meat and Fish" commodity groups have considerable differences. So, these commodity groups does not have proper substitutes.

Table 4.8: Compensated Price Elasticities

	CG 1	CG 2	CG 3	CG 4	CG 5	CG 6
CG 1	**-0.1366**	-0.1239	0.1462	-0.0713	0.2248	-0.0392
CG 2	-0.1456	**-0.2807**	-0.0246	-0.1370	0.4743	0.1137
CG 3	0.2168	-0.0432	**-0.8924**	0.1057	0.3949	0.2183
CG 4	-0.2549	-0.4097	0.2529	**-0.3546**	0.2653	0.5009
CG 5	0.1765	0.3168	0.2106	0.0584	**-0.8161**	0.0538
CG 6	-0.1222	0.3035	0.4339	0.4121	0.1950	**-1.2223**

The tax increases eventually yield welfare losses. However there are two points that can be clarified: (1) Does tax increase yield high tax income, and (2) how much welfare loss does it cause? If tax increases do not yield high income, then it is an inefficient action. Because, the new tax policy does not provide revenue by causing welfare loss for the consumers. The second question matters, because the governments does not want to cause too much welfare loss.

"Bread and Cereals" category has low uncompensated price elasticity. Thus, increasing tax rates for "Bread and Cereals" yields high tax incomes. However, it has considerably lower compensated price elasticity. This means, tax increases cause high welfare losses

[11] However, one should not forget that we cannot estimate the welfare loss appropriately without estimating supply.

for consumers. In other words, tax increase creates high deadweight loss for “Bread and Cereals”. The difference between compensated and uncompensated price elasticity is even higher for “Meat and Fish”. The consumers cannot compensate the price increases with other commodities, thus meat imports (by decreasing the price levels) substantially increases welfares of consumers. Tax increases for “Milk, cheese, and egg” is relatively elastic. So, tax increases does not yield high income. The difference between uncompensated and compensated is not high, thus tax increases does not increase welfare losses as much as the first two commodity groups. “Oils and Fats” has low price elasticity for both uncompensated and compensated price elasticities. Tax increases yield high incomes and does not cause too much welfare loss. “Fruit and Vegetables” and “Sugar, jam, honey, chocolate, and confectionary” commodity groups have relatively high price elasticities with small differences. Thus, high tax rates yield low income but does not cause too much welfare loss. Because, they are substitutable for the consumers.

5. CONCLUSION

5.1 Results

Policies for the "Bread and Cereals" and "Meat and Fish" commodity groups are crucial for society's welfare. The government should take actions regarding the decreasing price levels. Especially, "Meat and Fish" has very high income elasticity. The regarding actions are more important for the low-income people.

The tax increases for the commodity groups 3, 5, and 6 do not yield substantial revenues. Also it might cause high welfare loss (inefficiency) for the suppliers since the consumers are elastic.

Besides, there could be other parameters for the taxation of the foods. For instance, the government may want to create incentives for health foods. Therefore, decreasing the "Meat" prices might be unwise.

5.2 Future Work

For better policy recommendations, supply of the given food commodities should be estimated. We cannot comprehensively interpret the consequences of the policy decisions. For instance, increasing tax rates for the "Milk, cheese, and eggs" might lead to bankruptcy of the small firms.

Further methodological researches should be conducted on Machine Learning algorithms. Bajari, Nekipelov, Ryan, & Yang (2015) compared the ML learning algorithms for demand estimation. However, it should be referred as "prediction" rather than "estimation". Because, explained ML algorithms are not integrated with the Microeconomics theory and they do not offer interpretable tools to make causal inferences. However, There are promising econometric studies on ML. Athey, Tibshirani, & Wager (2016) introduces Generalized Random Forest as an extension to random forests. This method is used for non-parametric quantile regression, conditional average partial effect estimation, and heterogeneous treatment effect estimation via IV. Wager & Athey (2017) propose a method to use random forest for identifying heterogeneous treatment effects like personalized medicine. Chernozhukov et al. (2016) theorized a methodology that utilizes ML algorithms in econometric context called Double Machine Learning. It allows to make inference about treatment effect in the presence of high-dimensional control variables. Moreover, Chernozhukov, Goldman, Semenova, & Taddy (2017) used Double Machine learning method to make “demand estimation” and it can estimate the demand elasticities.

REFERENCES

Akbay, C., Boz, I., & Chern, W. S. (2007). Household food consumption in turkey. *European Review of Agricultural Economics*, *34*(2), 209–231.

Alston, J. M., Chalfant, J. A., & Piggott, N. E. (2001). Incorporating demand shifters in the almost ideal demand system. *Economics Letters*, *70*(1), 73–78.

Armagan, G., & Akbay, C. (2008). An econometric analysis of urban households' animal products consumption in turkey. *Applied Economics*, *40*(15), 2029–2036.

Athey, S., Tibshirani, J., & Wager, S. (2016). Generalized random forests. *arXiv preprint arXiv:1610.01271*.

Bajari, P., Nekipelov, D., Ryan, S. P., & Yang, M. (2015). Machine learning methods for demand estimation. *American Economic Review*, *105*(5), 481–85.

Banks, J., Blundell, R., & Lewbel, A. (1997). Quadratic engel curves and consumer demand. *Review of Economics and statistics*, *79*(4), 527–539.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 242–262.

Bilgic, A., & Yen, S. T. (2013). Household food demand in turkey: A two-step demand system approach. *Food Policy*, *43*, 267–277.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., & Newey, W. K. (2016). *Double machine learning for treatment and causal parameters* (cemmap working paper, Centre for Microdata Methods and Practice No. CWP49/16). Retrieved from `http://hdl.handle.net/10419/149795`

Chernozhukov, V., Goldman, M., Semenova, V., & Taddy, M. (2017). Orthogonal machine learning for demand estimation: High dimensional causal inference in dynamic panels. *arXiv preprint arXiv:1712.09988*.

Christensen, L. R., Jorgenson, D. W., & Lau, L. J. (1975). Transcendental logarithmic utility functions. *The American Economic Review*, *65*(3), 367–383.

Colchero, M., Salgado, J., Unar-Munguia, M., Hernandez-Avila, M., & Rivera-Dommarco, J. (2015). Price elasticity of the demand for sugar sweetened beverages and soft drinks in mexico. *Economics & Human Biology*, *19*, 129–137.

Deaton, A., & Muellbauer, J. (1980). An almost ideal demand system. *The American economic review*, *70*(3), 312–326.

Dybczak, K., Tóth, P., & Vonka, D. (2014). Effects of price shocks on consumer demand: Estimating the quaids demand system on czech household budget survey data. *Finance a Uver*, *64*(6), 476.

Eales, J. S., & Unnevehr, L. J. (1994). The inverse almost ideal demand system. *European Economic Review*, *38*(1), 101–116.

Henningsen, A. (2017). *Demand analysis with the "almost ideal demand system" in r: Package miceconaids.*

İpek, E., & Akyazı, H. (2015). Talep sistemlerinin ranki: Türkiye uygulamalari için bir literatür eleştirisi. *Global Journal of Economics and Business Studies*(8), 01–09.

Jabarin, A. S. (2005). Estimation of meat demand system in jordan: an almost ideal demand system. *International Journal of consumer studies*, *29*(3), 232–238.

Jehle, G., & Reny, P. (2011). *Advanced microeconomic theory*. Financial Times/Prentice Hall. Retrieved from `https://books.google.com.tr/books?id=1c2tbwAACAAJ`

LaFrance, J. T. (2004). Integrability of the linear approximate almost ideal demand system. *Economics Letters*, *84*(3), 297–303.

Lewbel, A. (1991). The rank of demand systems: theory and nonparametric estimation. *Econometrica: Journal of the Econometric Society*, 711–730.

Lewbel, A., & Pendakur, K. (2009). Tricks with hicks: The easi demand system. *American Economic Review*, *99*(3), 827–63.

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory* (Vol. 1). Oxford university press New York.

Moschini, G. (1995). Units of measurement and the stone index in demand system estimation. *American journal of agricultural economics*, *77*(1), 63–68.

Muellbauer, J. (1975). Aggregation, income distribution and consumer demand. *The Review of Economic Studies*, *42*(4), 525–543.

Poi, B. P. (2002). *Three essays in applied econometrics* (Doctoral thesis). University of Michigan. (Chapter 2: Dairy policy and consumer welfare)

Poi, B. P. (2012). Easy demand-system estimation with quaids. *Stata Journal*, *12*(3), 433.

Ray, R. (1983). Measuring the costs of children: an alternative approach. *Journal of Public Economics*, *22*(1), 89–102.

Sengul, S., & Tuncer, İ. (2005). Poverty levels and food demand of the poor in turkey. *Agribusiness*, *21*(3), 289–311.

Shonkwiler, J. S., & Yen, S. T. (1999). Two-step estimation of a censored system of equations. *American Journal of Agricultural Economics*, *81*(4), 972–982.

TC KALKINMA BAKANLIĞI. (2014). *Onuncu kalkınma planı (2014-2018) yaşlanma özel ihtisas komisyonu raporu.* (Haz.: Komisyon)

Theil, H. (1965). The information approach to demand analysis. *Econometrica: Journal of the Econometric Society*, 67–87.

TURKSTAT. (n.d.). *Household budget survey micro data set.* `http://www.turkstat.gov.tr/UstMenu.do?metod=bilgiTalebi`. (Surveys between 2008 and 2016)

TURKSTAT. (2015). *Classifications used.* `http://www.tuik.gov.tr/MicroVeri/HBI_2015/english/meta-data/method/classifications-used/index.html`.

Ulubasoglu, M., Mallick, D., Wadud, M., Hone, P., & Haszler, H. (2016). Food demand elasticities for australia. *Australian Journal of Agricultural and Resource Economics*, *60*(2), 177–195.

Verbeke, W., & Ward, R. W. (2001). A fresh meat almost ideal demand system incorporating negative tv press and advertising impact. *Agricultural Economics*, *25*(2-3), 359–374.

Wager, S., & Athey, S. (2017). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*(just-accepted).

Wooldridge, J. M. (2013). *Introductory econometrics: A modern approach*. South-Western, Cengage Learning.

APPENDICES

APPENDIX A: Derivation of Roy's Identity for QUAIDS

APPENDIX A

Remember that $\ln x = \ln m - \ln a(p)$. Expanding the equation 2.41:

$$w_i = -\frac{\dfrac{\partial \ln x}{\partial \ln p_i}\dfrac{1}{\ln x\lambda(p)+b(p)}}{\dfrac{\partial \ln x}{\partial \ln m_i}\dfrac{1}{\ln x\lambda(p)+b(p)} - \dfrac{\ln x}{(\ln x\lambda(p)+b(p))^2}\left[\dfrac{\partial \ln x}{\partial \ln m_i}\lambda(p)\right]}$$

$$+\frac{\dfrac{\ln x}{(\ln x\lambda(p)+b(p))^2}\left[\dfrac{\partial \ln x}{\partial \ln p_i}\lambda(p)+\dfrac{\partial \lambda(p)}{\partial \ln p_i}\ln x+\dfrac{\partial b(p)}{\partial \ln p_i}\right]}{\dfrac{\partial \ln x}{\partial \ln m_i}\dfrac{1}{\ln x\lambda(p)+b(p)} - \dfrac{\ln x}{(\ln x\lambda(p)+b(p))^2}\left[\dfrac{\partial \ln x}{\partial \ln m_i}\lambda(p)\right]}$$

$$= -\frac{\dfrac{\partial \ln x}{\partial \ln p_i} - \dfrac{\ln x}{\ln x\lambda(p)+b(p)}\left[\dfrac{\partial \ln x}{\partial \ln p_i}\lambda(p)+\dfrac{\partial \lambda(p)}{\partial \ln p_i}\ln x+\dfrac{\partial b(p)}{\partial \ln p_i}\right]}{\dfrac{\partial \ln x}{\partial \ln m_i} - \dfrac{\ln x}{\ln x\lambda(p)+b(p)}\left[\dfrac{\partial \ln x}{\partial \ln m_i}\lambda(p)\right]}$$

$$= -\frac{-\dfrac{\partial \ln a}{\partial \ln p_i} - \dfrac{\ln x}{\ln x\lambda(p)+b(p)}\left[-\dfrac{\partial \ln a}{\partial \ln p_i}\lambda(p)+\dfrac{\partial \lambda(p)}{\partial \ln p_i}\ln x+\dfrac{\partial b(p)}{\partial \ln p_i}\right]}{1-\dfrac{\ln x\lambda(p)}{\ln x\lambda(p)+b(p)}}$$

$$= -\frac{-\dfrac{\partial \ln a}{\partial \ln p_i}\left[1-\dfrac{\ln x\lambda(p)}{\ln x\lambda(p)+b(p)}\right]}{\dfrac{\ln x\lambda(p)+b(p)-\ln x\lambda(p)}{\ln x\lambda(p)+b(p)}}$$

$$-\frac{-\dfrac{\partial \lambda(p)}{\partial \ln p_i}\dfrac{\ln^2 x}{(\ln x\lambda(p)+b(p))} - \dfrac{\partial b(p)}{\partial \ln p_i}\dfrac{\ln x}{(\ln x\lambda(p)+b(p))}}{\dfrac{\ln x\lambda(p)+b(p)-\ln x\lambda(p)}{\ln x\lambda(p)+b(p)}}$$

$$= -\frac{-\dfrac{\partial \ln a}{\partial \ln p_i}\left[\ln x\lambda(p)+b(p)-\ln x\lambda(p)\right] - \dfrac{\partial \lambda(p)}{\partial \ln p_i}\ln^2 x - \dfrac{\partial b(p)}{\partial \ln p_i}\ln x}{\ln x\lambda(p)+b(p)-\ln x\lambda(p)}$$

$$= -\frac{-\dfrac{\partial \ln a}{\partial \ln p_i}b(p) - \dfrac{\partial \lambda(p)}{\partial \ln p_i}\ln^2 x - \dfrac{\partial b(p)}{\partial \ln p_i}\ln x}{b(p)}$$

$$= \frac{\partial \ln a}{\partial \ln p_i} + \frac{\partial b(p)}{\partial \ln p_i}\frac{\ln x}{b(p)} + \frac{\partial \lambda(p)}{\partial \ln p_i}\frac{\ln^2 x}{b(p)}$$

$$= a_i + \sum_k \gamma_{ik}\ln p_k + \beta_i ln(x) + \frac{\lambda_i}{b(p)}\ln^2 x$$

CURRICULUM VITAE

Name Surname : Ali Furkan KALAY

Place and Date of Birth : Antalya, 10 June 1994

E-Mail : kalaya@itu.edu.tr

EDUCATION

- **B.Sc.** : 2016, Istanbul Technical University, Industrial Engineering
- **M.A.** : 2018, Istanbul Technical University, Economics

PROFESSIONAL EXPERIENCE AND REWARDS

- 2017-2018 Research Assistant at Istanbul Technical University