



# MEGI

---

**Mestrado em Estatística e Gestão de Informação**  
Master Program in Statistics and Information Management

## **The market system demand curve - a fundamental distortion**

Ricardo Costa Mendes

Dissertation proposal as a partial requisite to obtain the  
Master degree in Statistics and Information Management

NOVA Information Management School  
Instituto Superior de Estatística e Gestão de Informação  
Universidade Nova de Lisboa

**NOVA Information Management School**  
**Instituto Superior de Estatística e Gestão de Informação**  
Universidade Nova de Lisboa

# **THE MARKET SYSTEM DEMAND CURVE - A FUNDAMENTAL DISTORTION**

By

Ricardo Costa Mendes

Dissertation as a partial requisite to obtain the Master degree in Statistics and Information Management, with a specialization in Information Analysis and Management

**Advisors:** Prof. Ana Cristina Costa and Prof. Joshua Farley

November 2017

## ABSTRACT

Since the mid-1970s, the global economy has been dominated by the spread of capitalist market economies, growing inequality and increasing ecological degradation. The latter may be the most serious of these trends. Human economic activities have reached a level that is bound to instigate irreversible change to the global environment, creating conditions likely to be less conducive to human development. The market system demand curve is distorted if inequality is sufficiently great and the purchasing power has a greater impact on allocation than preferences. If we attempt to internalize the ecological costs of essential resources into their market prices, driving up price, the best case scenario is that the poor reduce consumption by more than the rich, even though the rich have been the primary drivers of ecological degradation. The thesis focuses on the food allocation and demand curve distortion. Mainstream economists argue that it is impossible to objectively compare marginal utility across individuals, and the best we can do is equate willingness to pay with utility. However, food consumption is a physiological need, and it is quite possible to objectively compare the marginal utility it provides to different individuals. Certainly, a malnourished person gains more from additional food than an overfed one. A comprehensive econometric modeling of an aggregated and two-staged food demand systems is carried out for one hundred-seventy-seven countries. The data is retrieved from the 2011 round of the World Bank International Comparison Program. In the first stage, the Florida preference independence model is applied to the main broad groups of goods and services. In the second stage, a conditional demand system for food subcategories is estimated using the Florida Slutsky model. An Inaccuracy measure and the Storable decomposition are used in the outliers detection procedure. The system equations are corrected for both groupwise and scale heteroskedasticity. The iterated nonlinear seemingly unrelated regression (ITNLSUR) is applied in the final estimation, while the iterated nonlinear least squares (ITNLLS) produces the initial values. The Gauss-Newton method is used to approximate the maximums of the objective functions. Expenditures elasticities, Slutsky own price elasticities, Frisch own price elasticities and Cournot own price elasticities constitute the estimate of the elasticities structure. Expenditure and marginal expenditure shares are the most valued direct outcome of the models. In a perfect allocation system food marginal shares would be equal for every country. The discrepancies shown are an indicator of the market distortion. A redistribution towards poorer countries would increase total utility. Even if a Pareto optimum is in place in every economy (normally it is not), the solution captured by the model seems to be far from a global optimum. It is of utmost importance to know what are the implications on the real income of the poor if ecological thresholds are put in place through a market based mechanism. The Cournot elasticity estimates make evident that the poorer countries have more elastic demand curves, resulting primarily from the impact of increasing prices on real income, since there are no substitutes for food. This means that in an unequal market economy, if market based instruments are used to reduce the ecological degradation caused by food production, the poor will reduce consumption by a much greater percentage than the rich in response to price increases. Since the rich are responsible for far more ecological degradation than the poor, this outcome is highly perverse. This distortion is associated with the food shares and the marginal food shares that are higher in less affluent countries.

## **KEYWORDS**

Market demand distortion; Ecological economics; Market based mechanism; Florida preference independence model; Florida Slutsky model; NLLS; NLSUR; Food international consumption patterns;

# INDEX

1. Introduction .....	7
1.1. Background and problem identification .....	8
1.2. Study objectives .....	10
2. Literature review .....	11
2.1. Two-stage budgeting.....	11
2.2. Demand models .....	11
2.2.1. Working’s model .....	12
2.2.2. Florida Slutsky model .....	12
2.2.3. Florida Preference Independence model.....	14
2.2.4. First stage expenditure and price elasticities .....	16
2.2.5. Second stage expenditure and price elasticities .....	16
2.3. Econometrics .....	17
2.3.1. Estimation methods .....	17
2.3.2. Minimization methods .....	22
2.3.3. Outliers detection in cross-country demand analysis.....	24
2.3.4. Heteroskedasticity modeling .....	25
3. International comparison program data .....	28
4. Methodologic framework .....	30
5. Results and discussion .....	31
5.1. First stage and second stage aggregation .....	31
5.2. Outliers .....	32
5.3. First stage results .....	34
5.4. Second stage results.....	37
5.5. First stage discussion.....	38
5.6. Second stage discussion .....	42
6. Conclusions.....	45
7. Limitations and recommendations for future works .....	47
8. Bibliography.....	48
9. Annexes .....	52
9.1. Annex I – first stage dataset .....	52
9.2. Annex II – second stage dataset .....	55
9.3. Annex III – SAS code for final first stage ITSUR estimation.....	58



## 1. INTRODUCTION

Since the mid-1970s, the global economy has been dominated by three trends: the spread of capitalist market economies (Frieden, 2006); growing economic inequality (Piketty, 2014); and increasing ecological degradation driven by economic activity (Rockström et al., 2009; W. Steffen, Grinevald, Crutzen, & McNeill, 2011).

The spread of capitalism has been widely praised and strenuously promoted by national governments, trade organizations, international organizations (in particular, the IMF and World Bank), free trade agreements, academia, and others. Academics have provided strong theoretical justifications for markets. The dominant theories claim that in a free market economy, the decentralized, utility-maximizing decisions of producers and consumers drive the economy towards an economic equilibrium in which it would be impossible to make anyone better off without harming others, a condition known as Pareto efficiency or Pareto optimality. This outcome is based on the realistic assumption that individuals experience diminishing marginal utility, and the more questionable assumption that they maximize their own subjective utility within a budget constraint by allocating their income so that the last dollar spent on any good or service provides the same marginal utility (Akerlof & Shiller, 2015; Keen, 2011; Thaler, 2015). Based on decentralized information and free choice, the market supposedly allocates factors of production to those products with the greatest value, then apportions those products to those who value them most, as measured by willingness to pay (Hayek, 1945). If we accept willingness to pay as an objective measure of utility received, and if all consumers in a competitive market economy pay the same price for a given commodity, then markets will equalize marginal utility per dollar spent across all consumers, thus maximizing total utility as well. Unfortunately, market optimality, in theory and practice, is seriously undermined by both inequality and environmental degradation.

Inequality was long ignored by neoclassical economists. Many early economists recognized that if individuals experience diminishing marginal utility, then rich people presumably experience lower marginal utility from consumption than the poor, suggesting that redistributing income from the rich to the poor would increase total utility (Marshall, 1890; Pigou, 1932). However, at the turn of the 20th century, Vilfredo Pareto, for whom the market outcome of Pareto efficiency is named, argued that it is not possible to meaningfully quantify utility, or compare the utility experienced by different individuals, who experience pleasure and pain in different ways. One cannot therefore claim that redistribution would improve utility, and economists should not be concerned with inequality (Pareto, 1971). During the 1960s, when Nobel laureate Joseph Stiglitz entered the field, he found that “most [economists] were unconcerned about inequality; the dominant school worshiped at the feet of (a misunderstood) Adam Smith, at the miracle of the efficiency of the market economy.” At the time inequality in the US was at its nadir. Even after soaring inequality over the subsequent four decades, another Nobel laureate, Robert Lucas, could acknowledge the US current “staggering and unprecedented income inequality”, but also argue that we should pay it no mind, since free markets’ “limitless potential of increasing production” leads to “sustained, exponential growth in living standards”; in fact, he writes, “of the tendencies that are harmful to sound economics, the most seductive, and in my opinion the most poisonous, is to focus on questions of distribution.” (Lucas, 2004). Economic growth would solve the problem of poverty.

Steadily increasing inequality however has forced economists to pay attention. The richest eight individuals in the world now control the same wealth as the poorest 3.75 billion (OXFAM, 2017). In the United States, the top 1% has captured over half the increase in GDP since the great recession (Saez, 2016). Picketty's (2014) bestselling *Capital in the 21st Century* argued not only that inequality was reaching levels not seen in centuries, but also that capitalism was the cause. In a capitalist economy, returns on capital generally exceed the growth rate of the economy as a whole, inevitably concentrating wealth in the hands of the few in the absence of government intervention. It is no longer plausible to argue that we cannot be certain that an additional \$1,000 of income would improve utility more for a destitute person than a rich one. The only remaining justification for extreme inequality is that it increases economic growth, and that a larger pie will benefit everyone. However, there is also abundant evidence that relative wealth matters more than absolute wealth, and increasing income for the few makes reduces subjective utility for the many (Frank, 2005, 2007; Lane, 2001). Furthermore, compelling evidence shows that inequality contributes to a host of social and individual ills, ranging from homicide rates to obesity, and even the rich citizens of unequal countries may be worse off than lower income individuals in more equal ones (Wilkinson & Pickett, 2009).

Ecological degradation may be the most serious of these trends. Human economic activities have reached a level that is bound to instigate irreversible change to the global environment, creating conditions likely to be less conducive to human development. The thresholds for climate change, biodiversity and the nitrogen cycle have already been exceeded and the thresholds for freshwater, land use change, ocean acidification and global phosphorous cycle are being approached (Rockström et al., 2009; Will Steffen et al., 2015). This presents serious challenges for market economies for several reasons. First, economic growth is the main driver of ecological degradation. Not only can we no longer rely on continued growth to end poverty, we will likely need to reduce the physical size of the economy in order to return to the safe operating space for the global economy (Kallis, Kerschner, & Martinez-Alier, 2012; Martinez-Alier, 2016; Wackernagel et al., 2002). Given finite resources created by nature, we must pay attention to their equitable distribution. Second, ecological degradation is destroying ecosystem services, most of which are non-excludable and/or non-rival. Markets are not viable for non-excludable resources, and are inefficient for non-rival ones (Farley & Costanza, 2010). Third, all economic production requires raw materials from nature, most of which alternatively serve as the structural building blocks of ecosystems, and energy, primarily fossil fuels. Removing ecosystem structure and emitting waste, particularly from fossil fuels, both degrade global ecosystem services. Since all economic production has negative impacts on others, all economic activities improve the welfare of some and leave others worse off, and Pareto efficiency is a meaningless criterion.

### **1.1. BACKGROUND AND PROBLEM IDENTIFICATION**

This thesis will focus on an issue that illustrates the importance of these three trends: the allocation of food in a market economy. Food production poses one of the most serious threats to global ecosystems (Foley et al., 2011; Rockström et al., 2009; Tilman, Balzer, Hill, & Befort, 2011) but is essential for human wellbeing. Essential resources with limited possibilities for substitution exhibit highly price inelastic demand, meaning that demand is insensitive to changes in price, and a small decrease in supply will therefore lead to a large increase in price. The lower the share of the budget for which a resource accounts, the more inelastic the demand, which means that the rich are particularly insensitive to price increases. Nevertheless, market-based instruments are being



increasingly used as a conceptualized framework to tackle the latent incompatibility between economic activity and ecological thresholds (McCauley, 2006). Because markets weight preferences by purchasing power, market-based “solutions” are particularly problematic in a highly unequal world. If inequality is sufficiently great, purchasing power will have a greater impact on allocation than preferences. If we attempt to internalize the ecological costs of essential resources into their market prices, driving up price, the best case scenario is that the poor reduce consumption by more than the rich, even though the rich have been the primary drivers of ecological degradation. The worst case scenario is that the poor cannot purchase enough essential resources to survive, while the rich fail to even notice rising prices. This is precisely what happened when food prices soared in 2007-2008 (Farley, Schmitt Filho, Burke, & Farr, 2015).

While mainstream economists have historically argued that it is impossible to objectively compare marginal utility across individuals, this is not true for food consumption, which is a physiological need with scientifically measurable marginal benefits. Certainly, a malnourished person gains more utility from additional food than an overfed one. This thesis defines an economic demand curve as the marginal value of an essential commodity for any quantity consumed. The marginal value of essential resources becomes immeasurably large for an individual as the quantity consumed approaches the minimum required for survival. For food, this might be about 1200-1500 calories per day for an average female and male, respectively. Marginal value approaches zero as all physiological and psychological needs are satisfied, and arguably goes below zero as additional consumption makes people less healthy, above perhaps 3000-3500 calories per day for an active female and male, respectively. Since physiological needs for food are roughly similar for all groups of individuals, the aggregate economic demand curve is just the individual one scaled up.

In contrast, the individual’s market demand curve tells how much a consumer will buy of any given commodity at any given price, and is therefore a measure of consumer preferences weighted by purchasing power. The aggregate market demand curve is the sum of the individual curves. In theory, markets allocate resources to their highest value uses, and as prices increase, markets will ensure that lowest value uses are eliminated first. For food, energy, and other essential resources, we can objectively state that the highest value use is satisfying basic needs, and the lowest value use is food waste. However, with an unequal distribution of wealth and income, markets will prioritize the preferences of the rich over those of the poor. With sufficiently unequal income, markets will allocate food to the rich who throw 30-40% in the garbage (Gunders, 2012; Gustavsson, J. Cederberg, Sonesson, Otterdijk, & Meybeck, 2011) rather than to the poor who need it to stave off malnutrition (FAO, IFAD, & WFP, 2015). For example, when food prices skyrocketed in the 2007-2008 and 2011-2012 food crises, consumption in the wealthiest countries was essentially unchanged, while it declined in the poorest countries (Farley, Schmitt Filho, Burke, & Farr, 2015). This empirical evidence reveals a fundamental and endogenous bias in the market resource allocation process.

The commonly accepted aim of economics is the efficient allocation of resources, where ‘efficient’ implies utility maximizing. It is widely acknowledged that market prices fail to reflect many ecological costs and benefits, leading to excessive ecological degradation. Among economists, the widely accepted solution is to ‘internalize’ ecological costs into market prices, which would dramatically increase the price of food. Before pursuing such an option, it is important to empirically examine how markets allocate food and other essential resources. Estimating the price and income elasticity of

demand for food in different countries will allow us to determine by how much different populations will reduce food consumption in response to price increases, and hence evaluate the impact on food security. Furthermore, in spite of ecological limits to growth, endless economic growth remains a near universal goal among both politicians and economists, and is included in the UN's Sustainable Development Goals. As people rise into the global middle class, they demand more grain fed animal protein. This not only threatens to worsen the ecological impacts of agriculture, it also increases the demand for staple grains, and hence their price. Unequal growth may make it more difficult for the basic needs of the poor to compete against the luxury preferences of the rich. Estimating the price and income elasticity of demand for different foods will allow us determine the impacts of unequal economic growth on food security.

## **1.2. STUDY OBJECTIVES**

Is market allocation addressing the claims of wealthier consumers with an inherent minor marginal utility before poorer consumers with an inherent major marginal utility? If so, what are the fundamental factors behind that distortion?

The thesis objective is to develop the market demand distortion concept and contribute to the ongoing debate over the inadequacies of market-based instruments on tackling the ecological thresholds issue.

A comprehensive econometric modeling of a two staged aggregated and food demand systems will be carried out (Barnett & Serletis, 2008; Seale, Regmi, & Bernstein, 2003). The parametric demand modelling literature will be reviewed. The data will be retrieved from the 2011 rounds of the World Bank International Comparison Program.

The market demand distortion is more obvious when the utility function is directly linked to a metabolic process strongly uncorrelated to cultural, anthropological, psychological or conceptual aspects. In this case the similarity of the individual utility functions is such that the optimum solution should not be far from elasticities parity. This fact was the main reason food was chosen.

## **2. LITERATURE REVIEW**

### **2.1. TWO-STAGE BUDGETING**

Two-stage budgeting occurs when the consumer allocates the expenditure in two stages. At the first stage, the total expenditure is allocated to broad groups of goods and services. Here, each group expenditure depends on both the total expenditure and on some appropriate group price index. At the second stage, first stage group expenditures are further allocated to either individual or subgroups of goods and services. Second stage expenditures depend on both the group expenditure and on the prices within. These two allocations have to be made in such a way that guarantees a final result identical to an one step allocation with full information (Deaton & Muellbauer, 1980).

The second stage is intrinsically linked to the assumption of weak separability of preferences. (Henri Theil, 1980) named them blockwise dependent preferences. Under weak separability, preferences within a group are independent of the quantities that are consumed elsewhere in other groups and therefore the substitutes and complements should be kept together. As a result, each group has a sub-utility function dependent on its own expenditure and own individual prices or subgroups price indexes, specific group demand systems can be conceptualized and separately estimated with potential huge gains in terms of degrees of freedom and, at last, the global utility can be interpreted as a combination of the subsequent sub-utilities functions. Although there isn't any restriction on price substitution within groups, the price substitution between groups is equal to all elements of the same group and so a group level Slutsky matrix can be calculated (Deaton & Muellbauer, 1980).

In the first stage, it is common to postulate strong or additive separability of preferences. (Henri Theil, 1980) named it block independence. Under strong separability the utility function is a mere addition of the subsequent sub-utilities functions and the Slutsky matrix includes neither complements nor inferior goods and services. Despite its large restrictiveness, there are substantial gains in terms of degrees of freedom. Empirical evidence strongly disapproves its usage in low aggregated groups. Even in highly aggregated groups substantial care must be taken. In particular, close related goods should join the same group (Deaton & Muellbauer, 1980; Seale & Regmi, 2006).

### **2.2. DEMAND MODELS**

In the first stage, the Florida Preference Independence model is used to estimate an aggregate demand system over nine categories of goods and services: food, beverages and tobacco; clothing and footwear; education; gross rent and fuel; house furnishings and operations; medical care; transport and communications; recreation; other expenditures. In the second stage, the Florida Slutsky model is used instead to estimate a food demand system over eight food sub-categories: bread and cereals; meat; fish; dairy products; oils and fats; fruits and vegetables; beverages and tobacco; other food products.

The Florida Slutsky model is an extension of Working's model and the Florida Preference Independence model is a particular case of the Florida Slutsky model (Muhammad, Meade, Regmi, & Seale, 2011; Seale & Regmi, 2006; Henri Theil, 1980; Henry Theil, Chung, & L. Seale, 1989; Henry Theil, Suhm, & Meisner, 1981; Working, 1943).

Both demand models assume constant tastes across countries. The impact of this restriction should be minimized by broad groups of goods and services that are more bound to encompass significant

international tastes fluctuations (Reimer & Hertel, 2004).

### 2.2.1. Working's model

The Working's model was first applied to the analysis of U.S. household expenditure data in the 1930's. The model main assumption is that each household faces the same price vector so that the quantity demanded only depends on expenditure (Working, 1943). The Working's model can be expressed as follows,

$$w_i = \alpha_i + \beta_i \cdot \log E + \varepsilon_i \quad i = 1, \dots, n \quad (1)$$

$W_i = \frac{E_i}{E}$  is the budget share of good  $i$ ,  $E_i = p_i \cdot q_i$  is the expenditure on good  $i$  and  $E = \sum_1^n p_i \cdot q_i$  is the total consumption expenditure,  $p_i$  is the price of good  $i$ ,  $q_i$  is the quantity of good  $i$ , and  $\varepsilon_i$  is a residual term.

The following additivity constraints are met,

$$\sum_1^n \alpha_i = 1 \text{ and } \sum_1^n \beta_i = 0 \quad (2)$$

Let us multiply both sides of equation (1) by total expenditure,

$$E_i = \alpha_i \cdot E + \beta_i \cdot \log E \cdot E \quad (3)$$

Then differentiate equation (3) with respect to  $E$ ,

$$\vartheta_i = \frac{dE_i}{dE} = \alpha_i + \beta_i \cdot \log E + \beta_i \quad (4)$$

$$\vartheta_i = \frac{dE_i}{dE} = w_i + \beta_i \quad (5)$$

$\vartheta_i$  is the marginal share of good  $i$ . It measures the increase in expenditure on good  $i$  induced by a unitary increase in total expenditure, ceteris paribus. Income elasticities are simply the product of  $\frac{dE_i}{dE}$  and  $\frac{E}{E_i}$  or the ratio between the marginal share and the budget share of good  $i$ , that is,

$$\frac{\vartheta_i}{w_i} = \frac{dE_i}{dE} \cdot \frac{E}{E_i} = 1 + \frac{\beta_i}{w_i} \quad (6)$$

Through equation (6) good  $i$  is a luxury good if  $\beta_i > 0$  as its income elasticity is greater than unity. On the other hand, good  $i$  is a necessity if  $\beta_i < 0$  as its income elasticity is less than unity. If  $\beta_i = 0$ , good  $i$  has a unitary elasticity.

### 2.2.2. Florida Slutsky model

In international cross countries demand analysis, the price vector varies freely. As Working's model presupposes the same price vector, (Henry Theil et al., 1989) augmented it by assuming prices as explanatory variables.

Following (Henry Theil et al., 1989), let  $\tilde{W}_i^c$  be the budget share of good  $i$  at the geometric mean price  $\bar{p}_i$ , for  $i = 1, \dots, n$ , and at the observed real expenditure per capita of country  $c$  ( $Q_c$ ), where  $c$  represents the countries ( $c = 1, \dots, N$ ). Accordingly, the equation (1) is modified as follows,

$$\tilde{w}_i^c = \alpha_i + \beta_i \cdot \log Q_c + \varepsilon_i^c \quad (7)$$

Where  $\varepsilon_i^c$  is a disturbance. Therefore, the observed budget share of good  $i$  in country  $c$  can be written

as,

$$wic = ai + \beta i . \log Qc + (wic - \widetilde{wic}) + \varepsilon ic \quad (8)$$

Where  $(wic - \widetilde{wic})$  represents the variation in the budget share of good  $i$  as the price vector changes from  $\bar{p}i$  to  $p ic$ . The Florida Slutsky model introduces it through total differentiation.

Hence, differentiate  $wi = \frac{p i . qi}{E}$  as  $wi = e^{(\log pi + \log qi - \log E)}$ ,

$$dwi = wi . d(\log pi) + wi . d(\log qi) - wi . d(\log E) \quad (9)$$

A change in the budget share can be subdivided in a price, a quantity and an income component. Add and subtract  $wi . d(\log P)$  to the right-hand side of equation (9) and obtain:

$$dwi = wi . (d(\log pi) - d(\log P)) + wi d(\log qi) - wi . (d(\log E) - d(\log P)) \quad (10)$$

Where  $d(\log P)$  is the Divisia price index,

$$d(\log P) = \sum_1^n wi . d(\log pi) \quad (11)$$

Noting that  $d(\log E) - d(\log P) = d(\log Q)$  replace  $d(\log E) - d(\log P)$  in equation (10) by  $d(\log Q)$  which is the Divisia volume index,

$$d(\log Q) = \sum_1^n wi . d(\log qi) \quad (12)$$

The term  $wi . d(\log qi)$  is the dependent variable of the general differential demand equation that defines the Rotterdam demand model (Henry Theil et al., 1989),

$$wi . d(\log qi) = \theta i . d(\log Q) + \sum_{j=1}^n \pi ij d(\log pj) \quad (13)$$

Where the  $\pi ij$  is the Slutsky price coefficient and  $\theta i$  is the marginal share,

$$\pi ij = \frac{\partial qi(p, y)}{\partial pj} + \frac{\partial qi(p, y)}{\partial y} . qj(p, y) \quad (14)$$

Where  $qi(p, y)$  and  $qj(p, y)$  are the Walrasian demand functions of good  $i$  and  $j$  respectively and  $y$  is wealth.

The Slutsky matrix  $[\pi ij]$  corresponds to the Hessian of the expenditure function  $e(p, u)$  relative to prices or alternatively to the Jacobian matrix of the Hicksian demand function vector. Its elements are compensated price effects (Mas-Colell, Whinston, & Green, 2006).

If the neoclassical rational preferences axioms hold, the Walrasian demand functions would be continuous, differentiable, homogeneous of degree zero, observing the Walras law, and having a symmetric and negative semidefinite Slutsky matrix  $[\pi ij]$  of rank  $n-1$ . The negative semidefiniteness implies that the elements of the diagonal, the compensated own price effects, are nonpositive. The homogeneity of degree zero implies a Slutsky matrix that is singular with rank  $n-1$  (Mas-Colell et al.,

2006).

In the Florida Slutsky Model both the Slutsky symmetry property,  $\pi_{ij} = \pi_{ji}$  and the demand homogeneity property,  $\sum_{j=1}^n \pi_{ij} = 0$  for  $i = 1, \dots, n$  are imposed. Meanwhile, the additivity constraints also imply that  $\sum_{i=1}^n \pi_{ij} = 0$  for  $j = 1, \dots, n$ .

We can express equation (13) in terms of equation (8),

$$dwi = wi.[d(\log pi) - d(\log P)] + \vartheta i.d(\log Q) + \sum_{j=1}^n \pi_{ij}.d(\log p_j) - wi.d(\log Q) \quad (15)$$

As real income is fixed at  $Q_c$ ,  $d(\log Q)$  is null,

$$dwi = wi.[d(\log pi) - d(\log P)] + \sum_{j=1}^n \pi_{ij}.d(\log p_j) \quad (16)$$

Replace  $w_i$  with the initial  $\alpha_i + \beta_i.\log Q_c$ , the budget share of good  $i$  at the geometric mean price  $\bar{p}_i$  and interpret  $d(\log pi)$  as the difference between  $\log p_{ic}$  and  $\log \bar{p}_i = \frac{1}{N} \sum_{c=1}^N \log p_{ic}$ , the log of the geometric mean price of the good  $i$ . By applying the mean value theorem of calculus, we can now substitute equation (16) into equation (8) and obtain the Florida-Slutsky model,

$$wic = \alpha_i + \beta_i.qc + (\alpha_i + \beta_i.qc). \left\{ \log \frac{p_{ic}}{\bar{p}_i} - \sum_{j=1}^n (\alpha_j + \beta_j.qc). \log \frac{p_{jc}}{\bar{p}_j} \right\} + \sum_{j=1}^n \pi_{ij}. \log \frac{p_{jc}}{\bar{p}_j} + \varepsilon_{ic} \quad (17)$$

Where  $qc = \log Q_c$ .

The Florida Slutsky model has a linear real income term, a quadratic pure price term, and a linear substitution term. The parameters are  $\alpha_i$ ,  $\beta_i$  and  $\pi_{ij}$ .

It is applied to estimate the food and nonalcoholic beverages conditional demand system in the second stage under assumption of weak separability of preferences.

### 2.2.3. Florida Preference Independence model

The Florida Preference Independence (PI) Model assumes preferences independence. Under strong separability or additivity of preferences, the utility function corresponds to the addition of the  $n$  goods sub-utility functions. The marginal utility of good  $i$  is independent from the quantity of good  $j$  and so the Hessian of the utility function and its inverse are both diagonal matrices.

The marginal budget shares are given by

$$\theta_{ij} = \frac{(\mu_{.pi} \cdot \mu_{ij} \cdot p_j)}{\phi \cdot E} \quad (18)$$

Where  $\mu_{ij}$  is the term  $(i,j)$  of the Hessian of the utility function,  $E$  is the total expenditure and  $\phi$  is the income flexibility, that is the reciprocal of the income elasticity of the marginal utility of income:

$$\phi^{-1} = \frac{\frac{\partial^2 \mu}{\partial y^2}}{\frac{\frac{\partial \mu}{\partial y}}{y}} \quad (19)$$

As the Walras law is supposed to hold, income and expenditure are used here interchangeably. The estimation procedure uses also actual expenditures.

In the Florida PI Model, the elements of the Slutsky matrix can be expressed as follows:

$$\pi_{ij} = \phi \cdot (\theta_{ij} - \theta_i \cdot \theta_j) \quad (20).$$

Since the Hessian of the utility function  $[\mu_{ij}]$  is diagonal, the matrix  $[\theta_{ij}]$  of the marginal budget shares is also diagonal. We can rewrite (20) as follows:

$$\pi_{ij} = \phi \cdot \theta_i \cdot (1 - \theta_i); \quad i = j$$

$$\pi_{ij} = -\phi \cdot \theta_i \cdot \theta_j; \quad i \neq j \quad (21).$$

From (21) we conclude that complements are not possible and that for the negative semidefiniteness of the Slutsky matrix to hold  $\phi$  must be negative.

Thus, the Florida Slutsky model can be simplified to

$$wic = \alpha_i + \beta_i \cdot qc + (\alpha_i + \beta_i \cdot qc) \cdot \left\{ \log \frac{p_{ic}}{\bar{p}_i} - \sum_{j=1}^n (\alpha_j + \beta_j \cdot qc) \cdot \log \frac{p_{jc}}{\bar{p}_j} \right\} + \phi \cdot \theta_i \cdot \log \frac{p_{ic}}{\bar{p}_i} - \phi \cdot \theta_i \cdot \sum_{j=1}^n \theta_j \cdot \log \frac{p_{jc}}{\bar{p}_j} + \epsilon_{ic} \quad (22).$$

(Henry Theil et al., 1989) demonstrate that

$$\theta_{ic} = \alpha_i + \beta_i \cdot \check{q}c \quad (23)$$

Where  $\check{q}c = (1 + qc)$ .

The simplification of the Florida Slutsky model leads us to the Florida PI model

$$wic = \alpha_i + \beta_i \cdot qc + (\alpha_i + \beta_i \cdot qc) \cdot \left\{ \log \frac{p_{ic}}{\bar{p}_i} - \sum_{j=1}^n (\alpha_j + \beta_j \cdot qc) \cdot \log \frac{p_{jc}}{\bar{p}_j} \right\} + \phi \cdot ((\alpha_i + \beta_i \cdot \check{q}c) \cdot \left\{ \log \frac{p_{ic}}{\bar{p}_i} - \sum_{j=1}^n (\alpha_j + \beta_j \cdot \check{q}c) \cdot \log \frac{p_{jc}}{\bar{p}_j} \right\} + \epsilon_{ic} \quad (24).$$

The Florida PI model has a linear real income term, a quadratic pure price term, and a cubic substitution term. The parameters are each  $\alpha_i$  and  $\beta_i$  plus  $\phi$ .

The Florida PI model is applied to estimate the broad categories demand system of the first stage. Despite the assumption of strong separability being very restrictive, it decreases substantially the number of parameters delivering a solution with more degrees of freedom. Nevertheless, note that in a clear partition between essential and non-essential resources the strong separability is totally inappropriate because the greatest role non-essentials can have is to be complements of essentials and under strong separability complementarity is not possible. While the marginal utilities of essentials have an existence on their own the marginal utilities of non-essentials are inexistent if the adequate physiological threshold of the former is not met.

### 2.2.4. First stage expenditure and price elasticities

The first stage (aggregate) expenditure and price elasticities are calculated through the Florida PI model.

The expenditure elasticities are calculated as follows,

$$\eta_{gc} = \frac{\theta_{gc}}{\tilde{w}_{gc}} = 1 + \frac{\beta g}{\tilde{w}_{gc}} \quad (25)$$

Where g refers to the first stage group.

Three own price elasticities are calculated:

The Frisch own price elasticity presupposes that changes in prices are compensated to ensure a constant marginal utility of income,

$$F_{gc} = \phi \cdot \frac{\theta_{gc}}{\tilde{w}_{gc}} = \phi \cdot \left(1 + \frac{\beta g}{\tilde{w}_{gc}}\right) = \phi \cdot \eta_{gc} \quad (26)$$

The Slutsky (compensated) own price elasticity presupposes that changes in prices are compensated to ensure a constant real income,

$$S_{gc} = \frac{\phi \cdot (\tilde{w}_{gc} + \beta g) \cdot (1 - \tilde{w}_{gc} - \beta g)}{\tilde{w}_{gc}} = F_{gc} \cdot (1 - \tilde{w}_{gc} - \beta g) \quad (27)$$

The Cournot (uncompensated) own price elasticity comes as follows,

$$C_{gc} = \frac{\phi \cdot (\tilde{w}_{gc} + \beta g) \cdot (1 - \tilde{w}_{gc} - \beta g)}{\tilde{w}_{gc}} - (\tilde{w}_{gc} + \beta g) = S_{gc} - (\tilde{w}_{gc} + \beta g) \quad (28)$$

The Cournot elasticity is the most relevant to real-world consumers, who are rarely guaranteed a constant real income or constant marginal utility of income in response to price changes.

### 2.2.5. Second stage expenditure and price elasticities

The second stage disaggregated conditional expenditure and price elasticities are calculated through the Florida Slutsky model. Conditional expenditures assume that the expenditure allocated to a block (aggregated group) will be reallocated among commodities within the block in response to a price change within those commodities.

The unconditional expenditure and price elasticities are computed combining first and second stage expenditure and price elasticities. Unconditional expenditures assume that consumers will reallocate expenditures across blocks.

The conditional expenditure elasticities are calculated as follows,

$$\check{\eta}_{ic} = \frac{\check{\theta}_{ic}}{\check{w}_{ic}} = 1 + \frac{\check{\beta}_i}{\check{w}_{ic}} \quad (29)$$

The unconditional expenditure elasticities are computed as the product of (25) and (29),

$$\eta^{U}_{ic} = \eta_{gc} \cdot \check{\eta}_{ic} \quad i \in g \quad (30)$$

The conditional Frisch own price elasticity comes as follows,



$$\ddot{F}ic = \mathcal{F}gc \cdot \frac{\ddot{\theta}ic}{\ddot{w}ic} \quad i \in g \quad (31)$$

The unconditional Frisch own price elasticity is,

$$\mathcal{F}^U ic = \phi \cdot \eta^U ic \quad (32)$$

The conditional Slutsky own price elasticity is calculated as,

$$\ddot{S}ic = \frac{\ddot{\pi}iu}{\ddot{w}ic} \quad (33)$$

The unconditional Slutsky own price elasticity is,

$$S^U ic = \phi \cdot \ddot{\eta}ic \cdot \ddot{\eta}ic \frac{\ddot{w}ic}{\ddot{w}gc} \cdot (1 - \eta gc \cdot \tilde{w}gc) + \frac{\ddot{\pi}iu}{\ddot{w}ic} \quad i \in g \quad (34)$$

The conditional Cournot own price elasticity is,

$$\ddot{C}ic = \frac{\ddot{\pi}iu}{\ddot{w}ic} - \ddot{w}ic = \ddot{S}ic - \ddot{w}ic \quad (35)$$

In turn, the unconditional Cournot own price elasticity comes as follows,

$$C^U ic = S^U ic - \ddot{\eta}ic \cdot \ddot{w}ic \cdot \eta gc \cdot \tilde{w}gc \quad i \in g \quad (36)$$

## 2.3. ECONOMETRICS

### 2.3.1. Estimation methods

#### 2.3.1.1. General Nonlinear Simultaneous Equations model

(SAS, 2014) notation is nearly followed throughout the econometric literature review.

The nonlinear model specification can be written as

$$\tilde{\varepsilon}_t = \tilde{q}(\tilde{y}_t, \tilde{x}_t, \tilde{\theta})$$

$$\tilde{z}_t = Z(\tilde{x}_t)$$

For obs.  $t = 1, \dots, n$  (37).

The disturbance vector  $\tilde{\varepsilon}_t \in \mathbb{R}^g$ , where  $g$  is the system equations number, has the following general properties:

1.  $E(\tilde{\varepsilon}_t) = \tilde{0}$  (38)
2.  $E(\tilde{\varepsilon}_t \cdot \tilde{\varepsilon}_t') = \tilde{\Sigma}$  (39).

The  $\tilde{q} \in \mathbb{R}^g$  is a real function vector of  $\tilde{y}_t \in \mathbb{R}^g$ ,  $\tilde{x}_t \in \mathbb{R}^l$  and  $\tilde{\theta} \in \mathbb{R}^p$ , where  $l$  is the number of exogenous variables and  $p$  is the number of parameters.

The  $\tilde{z}_t \in \mathbb{R}^k$  is a vector of instruments.

The  $(gxg)$  matrix  $\tilde{\Sigma}$  is the disturbances variance-covariance matrix across equations.

In the general simultaneous equations models, there is reverse causation and the disturbance vector  $\tilde{\varepsilon}_t \in \mathbb{R}^g$  is correlated with the right-hand side endogenous variables  $\tilde{y}_t \in \mathbb{R}^g$ . The violation of the respective Gauss-Markov assumption<sup>1</sup> produces biased Ordinary Least Squares (OLS), Feasible Generalized Least Squares (FGLS) and Seemingly Unrelated Regression (SUR) estimators (Kennedy, 2008).

### 2.3.1.1. Nonlinear Linear Least Squares (NLLS)

The NLLS can be applied if the structural equations obey the Gauss-Markov theorem conditions, the model function is smooth and the objective function Hessian is positive definite somewhere (local minimum) in the unknown parameters space. In case the disturbances density is gaussian, the NLLS coincides with the maximum likelihood estimator. In other cases, the NLLS is consistent and converges asymptotically to a gaussian distribution (Amemiya, 1985).

The iterated version of the NLLS (ITNLLS) iterates the estimation of  $\tilde{\Sigma}$  ( $\tilde{S}$ ). The applied minimization method algorithm stops estimation when both parameters and  $\tilde{S}$  convergence criteria are simultaneously met (SAS, 2014).

The NLLS estimates  $\hat{\theta}$  are achieved by minimizing the following objective function:

$$\hat{S}_n = \tilde{r}' \cdot \tilde{r} / n \quad (43)$$

Where  $\tilde{r} \in \mathbb{R}^{n \cdot g}$  is a vector of residuals for the  $g$  equations stacked together.

The ITNLLS objective function is:

$$\hat{S}_n = \tilde{r}' \cdot (\text{diag}(\tilde{S})^{-1} \otimes \tilde{I}) \cdot \tilde{r} / n \quad (44)$$

Where  $\tilde{I} \in \mathbb{R}^n$  is an identity matrix,  $\tilde{S}$  is a  $(g \times g)$  matrix that estimates  $\tilde{\Sigma}$ ,  $\otimes$  is the Kronecker product and  $\tilde{I}$  is an  $(n \times n)$  identity matrix.

The NLLS and ITNLLS  $(p \times p)$  variance-covariance matrix of  $\hat{\theta}$  comes as follows:

$$\hat{\delta}_{\hat{\theta}} = (\tilde{X}' \cdot (\text{diag}(\tilde{S})^{-1} \otimes \tilde{I}) \cdot \tilde{X})^{-1} \quad (45)$$

Where  $\tilde{X}$  is an  $(n \times p)$  matrix of residuals partial derivatives with respect to the parameters.

### 2.3.1.2. Nonlinear Seemingly Unrelated Regression (NLSUR)

If each structural equation obeys the Gauss-Markov theorem conditions and equation disturbances are not independent from each other, the Seemingly Unrelated Regression (SUR) can be applied instead of OLS.

The SUR method consists in estimating by OLS the  $\tilde{\Sigma}$  matrix and then apply to the full system the Joint Feasible Generalized Least Squares (JFGLS). Bear in mind that the efficiency gain relative to OLS is only

---

<sup>1</sup> In the Classical Regression Model (CRM) the Gauss-Markov theorem assumptions hold and the OLS is therefore the best linear unbiased estimator. In the simultaneous equations model at least one endogenous variable is an independent variable that cannot be considered fixed in repeated samples, violating assumption four of the CRM and biasing the OLS estimator. In the nonlinear case, the simultaneous bias is apparent whenever the Jacobian of the disturbances relative to the parameters are not independent from at least one endogenous variable (Amemiya, 1985; Dhrymes, 1994).

materialized if the sample is large (Pindyck & Rubinfeld, 1998).

The NLSUR and SUR are alike. The NLSUR uses the Nonlinear Least Squares (NLLS) and the Nonlinear Joint Feasible Generalized Least Squares (NLJFGLS) to estimate the  $\tilde{\Sigma}$  matrix and then the system global equation which is formed by stacking together all equations.

The Iterated Nonlinear Least Squares (ITNLLS) iterates the estimation of  $\tilde{\Sigma}$  ( $\tilde{S}$ ), applying the NLJFGLS until it converges.

The NLSUR estimates  $\hat{\theta}$  are achieved by minimizing the following objective function:

$$\hat{S}_n = \tilde{r}' \cdot (\tilde{S}_{NLLS}^{-1} \otimes \tilde{I}) \cdot \tilde{r} / n \quad (46)$$

Where *NLLS* refers to the Nonlinear Least Squares estimator.

The ITNLSUR objective function is:

$$\hat{S}_n = \tilde{r}' \cdot (\tilde{S}^{-1} \otimes \tilde{I}) \cdot \tilde{r} / n \quad (47)$$

The NLSUR and ITNLSUR ( $p \times p$ ) variance-covariance matrix of  $\hat{\theta}$  comes as follows:

$$\hat{\delta}_{\hat{\theta}} = (\tilde{X}' \cdot (\tilde{S}^{-1} \otimes \tilde{I}) \cdot \tilde{X})^{-1} \quad (48)$$

### 2.3.1.3. Nonlinear Two Stage Least Squares (NL2SLS)

The Two Stage Least Squares (2SLS) estimator is a single equation estimation method.

The Theil's interpretation of the 2SLS estimator split the procedure as follows (Dhrymes, 1994):

1. Undertake an OLS regression to each right-hand side endogenous variable in the equation on the available instrumental variables. To the instrumental set of variables should belong all exogenous variables in the simultaneous equation model and other hypothetical variables which are uncorrelated with the disturbance vector  $\tilde{\varepsilon}_t \in \mathbb{R}^d$  and highly correlated with each endogenous variable. Compute the estimated values for each of these endogenous variables.
2. In the relevant equation, replace each endogenous right-hand side variable by its first stage estimated values and proceed to the second stage OLS estimation of the entire equation.

In the NL2SLS, it is the first derivatives of the model with respect to the parameters that are replaced with estimated values. (Amemiya, 1985) called it the Best Nonlinear Two Stage Least Squares (BNL2S). In this matter, it can be said that the 2SLS is no more than a special case of the general approach embedded in the NL2SLS.

The iterated version of the NL2SLS (ITNL2SLS) encompasses the iterated estimation of  $\tilde{\Sigma}$  ( $\tilde{S}$ ). Here the applied minimization method algorithm stops estimation when both parameters and  $\tilde{S}$  convergence criteria are simultaneously met (SAS, 2014).

The NL2SLS is biased in finite samples, but consistent in large samples. Its asymptotical efficiency is dependent on the spherical quality of the disturbances  $\varepsilon_{gt}$  (homoskedasticity and inexistence of autocorrelation). The NL2SLS gaussian asymptotical distribution is not coincidental with the maximum likelihood estimator distribution (Amemiya, 1985; Dhrymes, 1994).

The NL2SLS estimates  $\hat{\theta}$  are computed by minimizing the following objective function:

$$\hat{S}_n = \tilde{r}' \cdot (\tilde{I} \otimes \tilde{W}), \frac{\tilde{r}}{n} \quad (49)$$

Where:

- $\tilde{W} = \tilde{Z} \cdot (\tilde{Z}' \cdot \tilde{Z})^{-1} \cdot \tilde{Z}'$  is the  $(n \times n)$  projection matrix;
- $\tilde{Z}$  is an  $(n \times k)$  matrix of instrumental variables;

The ITNL2SLS objective function to be minimized is:

$$\hat{S}_n = \tilde{r}' \cdot (\text{diag}(\tilde{S})^{-1} \otimes \tilde{W}), \frac{\tilde{r}}{n} \quad (50)$$

The NL2SLS and the ITNL2SLS  $(p \times p)$  variance-covariance matrix of  $\hat{\theta}$  comes as follows:

$$\hat{\delta}_{\theta} = (\tilde{X}' \cdot (\text{diag}(\tilde{S})^{-1} \otimes \tilde{W}) \cdot \tilde{X})^{-1} \quad (51)$$

### 2.3.1.4. Nonlinear Three Stage Least Squares (NL3SLS)

The Three Stage Least Squares (3SLS) is a system estimation method as two or more equations are jointly estimated. It uses the actual correlation among the disturbances across equations to produce more precise parameter estimates. It has the inconveniences of requiring more data and being more sensitive to model specification errors.

The 3SLS estimator procedure can be split as follows (Kennedy, 2008):

1. Apply the 2SLS to estimate the equations;
2. Use the 2SLS structural equations residuals to estimate  $\tilde{\Sigma}$ , the variance-covariance matrix across equations;
3. Apply the SUR estimator to the transformed system, that is, the JFGLS to the transformed global equation using the 2SLS estimate for  $\tilde{\Sigma}$ . In the transformed system each right-hand endogenous variable is replaced by its 2SLS first stage estimated values.

In the NL3SLS, NL2SLS and NLSUR are used instead of 2SLS and SUR (Amemiya, 1985). As with the NL2SLS, the NL3SLS is a general approach encompassing the 3SLS.

In the NL3SLS, it is the first derivatives of the model with respect to the parameters that are replaced with estimated values. (Amemiya, 1985) called it the Best Nonlinear Three Stage Least Squares (BNL3S).

The iterated version of the NL3SLS (ITNL3SLS) encompasses iterated estimation of  $\tilde{\Sigma}$  ( $\tilde{S}$ ). In the third stage, the Iterated Nonlinear Seemingly Unrelated Regression (ITNLSUR) is applied. As for the ITNL2SLS, the applied minimization method algorithm stops estimation when both parameters and  $\tilde{S}$  convergence criteria are simultaneously met (SAS, 2014).

Although the NL3SLS is biased in finite samples, it is both consistent and asymptotically more efficient than NL2SLS. Its asymptotical efficiency is dependent on the spherical quality of the disturbances  $\varepsilon_{gt}$  (homoskedasticity and inexistence of autocorrelation). As for the NL2SLS, the NL3SLS gaussian asymptotical distribution does not depend on the disturbances density. This robustness is very attractive (Amemiya, 1985; Dhrymes, 1994).

The NL3SLS estimates  $\hat{\theta}$  are computed by minimizing the following objective function:

$$\hat{S}_n = \tilde{r}' \cdot (\hat{S}_{NL2SLS}^{-1} \otimes \tilde{W}) \cdot \frac{\tilde{r}}{n} \quad (52)$$

The ITNL3SLS objective function to be minimized is:

$$\hat{S}_n = \tilde{r}' \cdot (\hat{S}^{-1} \otimes \tilde{W}) \cdot \frac{\tilde{r}}{n} \quad (53)$$

The NL3SLS and the ITNL3SLS ( $p \times p$ ) variance-covariance matrix of  $\hat{\theta}$  comes as follows:

$$\hat{\delta}_{\hat{\theta}} = (\tilde{X}' \cdot (\hat{S}^{-1} \otimes \tilde{W}) \cdot \tilde{X})^{-1} \quad (54)$$

### 2.3.1.5. Nonlinear Full Information Maximum Likelihood (NLFIML)

The NLFIML is considered a system estimation method because all equations are jointly estimated. The NLFIML assumes that the disturbances vector  $\tilde{\varepsilon}_t \in \mathbb{R}^g$  follows a multivariate normal density with  $E(\tilde{\varepsilon}_t) = \tilde{0}$  and  $E(\tilde{\varepsilon}_t \cdot \tilde{\varepsilon}_t') = \tilde{\Sigma}$  that is full characterized by the  $\tilde{\theta} \in \mathbb{R}^p$  and  $\tilde{\sigma}^2 \in \mathbb{R}^g$  parameters vectors. As the  $\tilde{\sigma}^2 \in \mathbb{R}^g$  can be expressed as a function of  $\tilde{\theta} \in \mathbb{R}^p$ , a concentrated multivariate normal density is used. The NLFIML consists in estimating the  $\tilde{\theta} \in \mathbb{R}^p$  that maximizes the concentrated likelihood function over the parameter space  $\Theta \subset \mathbb{R}^p$ . The concentrated likelihood function is the joint normal density of  $\tilde{\varepsilon}_t \in \mathbb{R}^g$  expressed as a function of  $\tilde{\theta} \in \mathbb{R}^p$  and  $\tilde{\Sigma}(\tilde{\theta})$ . Due to computational convenience, the log likelihood function is normally used in the estimation procedure (Amemiya, 1985; SAS, 2014).

The NLFIML just guarantees that the estimates upon convergence correspond to a likelihood function local maximum. It ensures the global maximum only if the likelihood function is strictly concave (Amemiya, 1985; Dhrymes, 1994).

The NLFIML estimates of  $\hat{\theta}$  and  $\hat{\sigma}^2$  are computed by maximizing the following likelihood function:

$$I_n(\tilde{\theta}, \tilde{\sigma}^2) = -\frac{n \cdot g}{2} \cdot \text{Ln}(2\pi) + \sum_{t=1}^n \text{Ln} \left( \left| \frac{\partial \tilde{q}(\tilde{y}_t, \tilde{x}_t, \tilde{\theta})}{\partial \tilde{y}_t'} \right| \right) - \frac{n}{2} \cdot \text{Ln}(|\tilde{\Sigma}(\tilde{\sigma}^2)|) \\ - \frac{1}{2} \cdot \text{Tr} \left( \tilde{\Sigma}(\tilde{\sigma}^2)^{-1} \cdot \sum_{t=1}^n \tilde{q}(\tilde{y}_t, \tilde{x}_t, \tilde{\theta}) \cdot \tilde{q}'(\tilde{y}_t, \tilde{x}_t, \tilde{\theta}) \right) \quad (55)$$

Where the straight vertical parenthesis and the  $Tr$  function represent the determinant and the trace of the argument matrix respectively.

The  $\tilde{\Sigma}(\tilde{\theta})$  variance-covariance matrix across equations can be expressed as follows:

$$\tilde{\Sigma}(\tilde{\theta}) = \frac{1}{n} \cdot \sum_{t=1}^n \tilde{q}(\tilde{y}_t, \tilde{x}_t, \tilde{\theta}) \cdot \tilde{q}'(\tilde{y}_t, \tilde{x}_t, \tilde{\theta}) \quad (56)$$

The concentrated log likelihood function is reached using  $\tilde{\Sigma}(\tilde{\theta})$  instead of  $\tilde{\Sigma}(\tilde{\sigma}^2)$  in equation (48):

$$I_n(\tilde{\theta}) = -\frac{n \cdot g}{2} \cdot (1 + \text{Ln}(2\pi)) + \sum_{t=1}^n \text{Ln} \left( \left| \frac{\partial \tilde{q}(\tilde{y}_t, \tilde{x}_t, \tilde{\theta})}{\partial \tilde{y}_t'} \right| \right) - \frac{n}{2} \cdot \text{Ln}(|\tilde{\Sigma}(\tilde{\theta})|) \quad (57)$$

Take note that the developed SAS estimation code minimizes the symmetric of the concentrated log likelihood function.

The FIML ( $p \times p$ )  $\hat{\theta}$  variance-covariance matrix estimator is (Wooldridge, 2002):

$$\hat{\delta}_{\hat{\theta}} = (-\tilde{H}_{\hat{\theta}})^{-1} \quad (58)$$

Where  $\tilde{H}_{\hat{\theta}}$  is the Hessian of the concentrated log likelihood function.

$$\tilde{H}_{\hat{\theta}} = \frac{\partial^2 I_n(\hat{\theta})}{\partial \hat{\theta} \cdot \partial \hat{\theta}'} \quad (59)$$

As the expected Hessian when evaluated at the true parameters values is equal to the symmetric of the expected gradient vector outer product, another estimator can be used:

$$\hat{\delta}_{\hat{\theta}} = (\tilde{s}_{\hat{\theta}} \cdot \tilde{s}_{\hat{\theta}}')^{-1} \quad (60)$$

Where  $\tilde{s}_{\hat{\theta}}$  is the gradient vector of the concentrated log likelihood function.

$$\tilde{s}_{\hat{\theta}} = \frac{\partial I_n(\hat{\theta})}{\partial \hat{\theta}} \quad (61)$$

If the  $\tilde{\varepsilon}_t$  have a jointly normal density, the NLFIML is consistent, asymptotically normal and has a smaller variance-covariance matrix than NL3SLS. However, if not the case, the NLFIML is generally not consistent (Amemiya, 1985).

To test for the normality of the disturbances, a Shapiro-Wilk test (Shapiro & Wilk, 1965) is carried out for each equation:

$$H_0: \varepsilon_t \sim N(0, \sigma^2) \quad (62)$$

$$H_1: \text{Not } H_0 \quad (63)$$

After sorting the equation residuals  $r_i$  in ascending order, the following test statistic is computed:

$$W = \frac{b^2}{\sum_{i=1}^n (r_i - \bar{r})^2} \quad (64)$$

$$b = \sum_{i=1}^{\frac{n}{2}} a_i \cdot (r_{n-i+1} - r_i) \text{ if } n \text{ is even} \quad (65)$$

$$b = \sum_{i=1}^{\frac{n}{2}} a_i \cdot (r_{n-i+1} - r_i) \text{ if } n \text{ is odd} \quad (66)$$

$$\text{Reject } H_0 \text{ if } W_{obs} < W_{critical} \quad (67)$$

Where  $a_i$  and  $W_{critical}$  are tabulated and if  $n$  is odd the median observation is dropped.

### 2.3.2. Minimization methods

The nonlinear optimization methods can be subdivided in local and global methods. Local methods just guarantee local optimization. Although global methods such as simulating annealing or genetic algorithms can reach the global optimum, their huge computation effort reinforces the local methods attractiveness (Nelles, 2001).

The class of gradient based algorithms is a class of local optimization methods. It can be synthesized by the following expression:

$$\tilde{\theta}_k = \tilde{\theta}_{k-1} - \eta_{k-1} \cdot \tilde{p}_{k-1} \quad \text{with} \quad \tilde{p}_{k-1} = \tilde{R}_{k-1} \cdot \tilde{s}_{k-1} \quad (68)$$

Where  $k$  represents actual iteration,  $\eta$  the step size,  $\tilde{p}$  the  $(p \times 1)$  direction vector,  $\tilde{R}$  the  $(p \times p)$  scaling and rotation matrix and  $\tilde{s}$  the  $(p \times 1)$  objective function gradient vector.

As the objective functions,  $\hat{S}_n$  in least squares and  $-I_n(\tilde{\theta})$  in maximum likelihood, are meant to decrease in each iteration, the scaling and rotation matrix  $\tilde{R}$  ought to be positive definite.

The gradient based algorithms differ from one another in the specific choices that are made on the step size  $\eta$  and the scaling and rotation matrix  $\tilde{R}$ .

### 2.3.2.1. Newton method

In the Newton method,  $\eta_{k-1}$  and  $\tilde{R}_{k-1}$  are defined as follows (Amemiya, 1985; Nelles, 2001; Wooldridge, 2002):

$$\tilde{R}_{k-1} = \tilde{H}_{k-1}^{-1} \quad (69)$$

$$\eta_{k-1} = 1 \quad (70)$$

Where  $\tilde{H}_{k-1}$  is the objective function Hessian relative to vector parameter  $\tilde{\theta}_{k-1}$ .

Setting the step size to one is the optimal choice for a linear optimization problem that just requires one iteration to reach the solution. It is a direct consequence of the second order Taylor expansion of the objective function. However, in a nonlinear problem it is not generally possible to reach the solution in a single iteration. As a fixed step size here can be too small or too large, it is therefore common to use a line search in each iteration. Its aim is to find the optimal step size  $\eta_{k-1}$  that minimizes the objective function in the direction of  $\tilde{p}_{k-1}$ . (Nelles, 2001).

The Newton method requires at each iteration the computation of the objective function second derivatives relative to the parameters and the subsequent inversion of the Hessian, restricting its application to simple problems with low computational effort. Another non-neglectable drawback is the fact that in each iteration the objective function decreases only if the Hessian  $\tilde{H}_{k-1}$  is positive definite. In the optimum neighborhood that is always true. However, in the first iterations it can easily fail (Nelles, 2001; Wooldridge, 2002).

### 2.3.2.2. Gauss-Newton method

The Gauss-Newton method, an adaptation of the general Newton method, was especially designed to calculate the nonlinear least squares estimator (Amemiya, 1985; Nelles, 2001).

In the nonlinear least squares minimization problem the objective function gradient and Hessian relative to the parameters can be expressed as follows (Nelles, 2001):

$$\tilde{s}_{k-1} = 2 \cdot \tilde{J}'_{k-1} \cdot \tilde{Q}r_{k-1}(\tilde{y}_t, \tilde{x}_t, \tilde{\theta}) \quad (71)$$

Where  $\tilde{J}_{k-1}$  is the  $(ng \times p)$  Jacobian Matrix of the model residuals relative to the parameters and  $\tilde{Q}r_{k-1}$  is the  $(ng \times 1)$  packed residuals vector both evaluated at iteration  $k - 1$ .

$$\tilde{H}_{k-1} = 2 \cdot \tilde{J}'_{k-1} \cdot \tilde{J}_{k-1} + 2 \cdot \tilde{S}^q \quad (72)$$

$$\tilde{S}^q = \sum_{i=1}^{n.g} qr_{k-1,i}(\tilde{y}_t, \tilde{x}_t, \tilde{\theta}) \cdot \tilde{T}_{i,k-1} \quad (73)$$

Where  $qr_{k-1,i}$  is the  $i^{\text{th}}$  element of the  $\tilde{Q}r_{k-1}$  vector and  $\tilde{T}_{i,k-1}$  is the  $(pxp)$   $i^{\text{th}}$  residual Hessian relative to the parameters both evaluated at iteration  $k - 1$ .

In the Gauss-Newton method the residuals  $qr_{k-1,i}(\tilde{y}_t, \tilde{x}_t, \tilde{\theta})$  are assumed to be small and therefore it neglects the  $\tilde{S}^q$  matrix when setting the scaling and rotation matrix  $\tilde{R}_{k-1}$ :

$$\tilde{s}_{k-1} = \tilde{J}'_{k-1} \cdot \tilde{Q}r_{k-1}(\tilde{y}_t, \tilde{x}_t, \tilde{\theta}) \quad (74)$$

$$\tilde{R}_{k-1} = (\tilde{J}'_{k-1} \cdot \tilde{J}_{k-1})^{-1} \quad (75)$$

$$\eta_{k-1} = 1 \quad (76)$$

$$\tilde{\theta}_k = \tilde{\theta}_{k-1} - \eta_{k-1} \cdot \tilde{p}_{k-1} \quad \text{with} \quad \tilde{p}_{k-1} = (\tilde{J}'_{k-1} \cdot \tilde{J}_{k-1})^{-1} \cdot \tilde{J}'_{k-1} \cdot \tilde{Q}r_{k-1}(\tilde{y}_t, \tilde{x}_t, \tilde{\theta}) \quad (77)$$

As the scale and direction remain equivalent, the factor 2 is dropped in (74) and (75).

As long as  $\tilde{S}^q \rightarrow 0$ , the Gauss-Newton method has the same properties of the general Newton method without demanding the computation of second derivatives. The classical version sets the step size  $\eta_{k-1}$  to 1. However, it is common to perform a line search at each iteration to optimize the step size  $\eta_{k-1}$  (Amemiya, 1985; Nelles, 2001).

### 2.3.2.3. Levenberg-Marquardt method

The Gauss-Newton method cannot be applied if the matrix  $\tilde{J}'_{k-1} \cdot \tilde{J}_{k-1}$  in (77) is singular or ill conditioned. The smaller the least eigen value of  $\tilde{J}'_{k-1} \cdot \tilde{J}_{k-1}$  is, the worse is the convergence rate (Nelles, 2001).

The Levenberg-Marquardt method overcomes the problems associated with an ill conditioned  $\tilde{J}'_{k-1} \cdot \tilde{J}_{k-1}$  matrix, modifying the equation (77) as follows:

$$\tilde{\theta}_k = \tilde{\theta}_{k-1} - \eta_{k-1} \cdot \tilde{p}_{k-1} \quad \text{with} \quad \tilde{p}_{k-1} = (\tilde{J}'_{k-1} \cdot \tilde{J}_{k-1} \cdot \alpha_{k-1} \cdot \tilde{I}_p)^{-1} \cdot \tilde{J}'_{k-1} \cdot \tilde{Q}r_{k-1}(\tilde{y}_t, \tilde{x}_t, \tilde{\theta}) \quad (78)$$

Where  $\tilde{I}_p$  is a  $(pxp)$  identity matrix.

The Gauss-Newton method second order approximation is powerful when close to the optimum and a small  $\alpha_{k-1}$  should be adequate. However, when far from the optimum the convergence of the Gauss-Newton method may not occur and a larger  $\alpha_{k-1}$  should be chosen to guarantee a positive definite scale and rotation matrix  $\tilde{R}_{k-1}$  and a descent direction for the objective function.

Initially a positive value for  $\alpha_{k-1}$  should be chosen. At each further iteration,  $\alpha_{k-1}$  normally decreases as the optimum is bound to be approached. However, even in further iterations and whenever a downhill direction is not achievable  $\alpha_{k-1}$  will always increase. (Nelles, 2001).

### 2.3.3. Outliers detection in cross-country demand analysis

The information inaccuracy measure and the Strobe decomposition can be used to detect outliers in a cross-country demand analysis (Henri Theil, 1996). The formulation is as follows:

$$I_c = \sum_{i=1}^n I_{i,c} \quad i \text{ is a consumption group and } c \text{ is a country} \quad (79)$$



$$I_{i,c} = \widehat{W}_{i,c} - W_{i,c} + W_{i,c} \cdot \text{Log} \left( \frac{W_{i,c}}{\widehat{W}_{i,c}} \right) \quad (80)$$

$I_c$  is the information inaccuracy measure,  $I_{i,c}$  is the  $(i, c)$  Stoble component,  $W_{i,c}$  and  $\widehat{W}_{i,c}$  are respectively the actual and estimated expenditure share.

If the fit is perfect and  $W_{i,c}$  is equal to  $\widehat{W}_{i,c}$  for every  $(i, c)$ , the information inaccuracy measure  $I_c$  is zero. As the fit precision decreases, the  $I_c$  increases. Since the estimated  $\widehat{W}_{i,c}$  can be negative in the Florida Models, sometimes the  $I_c$  cannot be calculated. In these cases, the Stoble component is replaced with a value that is indifferent in terms of outlier classification.

### 2.3.4. Heteroskedasticity modeling

The presence of heteroskedasticity has a negative and non-negligible impact on the asymptotic properties of the least squares estimators. It is therefore useful to perform statistical tests for heteroskedasticity and if it is the case to model and modify the estimation procedure accordingly (Greene, 2012).

#### 2.3.4.1. Statistical tests for heteroskedasticity

In the Simultaneous Equations Model, a specific statistical test is undertaken for each equation.

##### 2.3.4.1.1. White's general test

$$H_0: \sigma_i^2 = \sigma^2 \quad \forall i \quad (81)$$

$$H_1: \text{Not } H_0 \quad (82)$$

The test is undertaken for each equation as follows (Greene, 2012; SAS, 2014; White, 1980):

- A. Regress the square of equation residuals  $r_i^2$  on a constant and on every unique variable (column) contained in the  $(n^2 \times p^2)$   $\tilde{J} \otimes \tilde{J}$  matrix where the  $(n \times p)$   $\tilde{J}$  is the Jacobian of the residuals  $r_i$  relative to the estimated parameters. In the linear simultaneous equations model the columns of  $\tilde{J}$  are constituted by all unique variables, by their squares and cross products.
- B. Compute the test statistic  $n \cdot R^2$  in the regression of  $r_i^2$ . Use the density under null, a Chi-squared with  $m - 1$  degrees of freedom where  $m$  is the number of parameters, and perform the test.

The White's test is general and very sensitive to model specification errors. In case the null is rejected, it does not deliver any clue about the actual form of the heteroskedasticity (Greene, 2012; White, 1980).

##### 2.3.4.1.2. The Breusch-Pagan Lagrange Multiplier test

$$\sigma_i^2 = \sigma^2 \cdot (\alpha_0 + \tilde{\alpha}' \cdot \tilde{Z}_i) \quad \forall i \quad (83)$$

$$H_0: \tilde{\alpha}' = \tilde{0} \quad (84)$$

$$H_1: \text{Not } H_0 \quad (85)$$

Where  $\tilde{\alpha}'$  and  $\tilde{Z}_i$  are  $(p \times 1)$  parameters and independent variables vectors respectively.

The test statistic is (Greene, 2012; SAS, 2014):

$$LM = \frac{1}{v} \cdot (\tilde{r}_g - \bar{r}_g \cdot \tilde{\mathbf{1}})' \cdot \tilde{\mathbf{Z}} \cdot (\tilde{\mathbf{Z}}' \cdot \tilde{\mathbf{Z}})^{-1} \cdot \tilde{\mathbf{Z}}' (\tilde{r}_g - \bar{r}_g \cdot \tilde{\mathbf{1}}) \quad (86)$$

Where  $\tilde{r}_g$  is the  $(nx1)$  vector of square residuals,  $\bar{r}_g$  is the mean square residual,  $\tilde{\mathbf{1}}$  is a  $(nx1)$  vector of ones,  $\tilde{\mathbf{Z}}$  is a  $(n \times p)$  matrix of independent variables observations and

$$v = \frac{1}{n} \cdot \sum_{i=1}^n \left( r_{g,i}^2 - \frac{\tilde{r}_g' \cdot \tilde{r}_g}{n} \right)^2 \quad (87)$$

As the Breusch-Pagan test is sensitive to the assumption of normality of the original disturbance term, a more robust estimator  $v$  of the  $e_{g,i}^2$  is used to compute the Lagrange multiplier statistic. Under the null, the test statistic has a Chi-squared asymptotic density with  $p$  degrees of freedom. The Breusch-Pagan test can be applied to various variance models (Greene, 2012).

### 2.3.4.2. Heteroskedasticity model

(Henry Theil et al., 1989) found group heteroskedasticity in their ICP cross-country demand analysis. The countries were divided in two groups. The first group included the countries that participated in phase II and phase III of the ICP and the second group the countries that didn't participate in neither. Fitting the Florida PI model to each group separately, the estimated covariance matrix of the second group was almost two-fold larger than the estimated covariance matrix of the first group.

Following (Seale & Regmi, 2006), three groups of heteroskedasticity were formed. The group I includes the countries that joined the ICP in the first three phases. The group III includes the countries that were added in 2011. The group II includes the remaining countries.

The final step is to evaluate the T-tests for the statistical significance of the variance parameterization.

As in cross section econometric analysis the existence of scale heteroskedasticity is very often, the country per capita expenditure natural logarithm  $q_t$  was introduced in the modeling.

The heteroskedasticity model is:

$$\tilde{\varepsilon}_t = \tilde{q}(\tilde{y}_t, \tilde{x}_t, \tilde{\theta}) \quad (88)$$

$$\tilde{\varepsilon}_t = \tilde{H}_t \cdot \tilde{\varepsilon}_t \quad (89)$$

$$\tilde{H}_t = \begin{bmatrix} \sqrt{h_{t,1}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{h_{t,g}} \end{bmatrix} \quad (90)$$

$$\sigma_{t,i}^2 = \sigma_i^2 \cdot (1 + \phi_1 \cdot D1 + \phi_2 \cdot D2 + \phi_3 \cdot q_t) + v_{t,i}$$

$$h_{t,i} = (\sigma_{t,i}^2)^{-1} \quad (91)$$

$$\tilde{\varepsilon}_t \sim N(\tilde{\mathbf{0}}, \tilde{\Sigma}) \quad (92)$$

$D1$  and  $D2$  are dummy variables that address the group heteroskedasticity.

It is assumed a normal density to the homoskedastic disturbance term  $\tilde{\varepsilon}_t$ . In the NLSUR, NL2SLS and

NL3SLS the Nonlinear Feasible Generalized Least Squares (NLFGLS) is used instead of NLLS whenever applicable and a pseudolikelihood procedure is undertaken in the heteroskedasticity model estimation (SAS, 2014).

In the NLFIML, the heteroskedasticity modeling changes the concentrated log likelihood function to:

$$I_n(\tilde{\theta}, \tilde{\varphi}) = -\frac{n \cdot g}{2} \cdot (1 + \text{Ln}(2\pi)) + \sum_{t=1}^n \text{Ln} \left( \left| \frac{\partial \tilde{q}(\tilde{y}_t, \tilde{x}_t, \tilde{\theta})}{\partial \tilde{y}_t'} \right| \right) - \frac{1}{2} \cdot \sum_{t=1}^n \sum_{i=1}^g \left( \text{Ln}(h_{t,i}) + \frac{q_i(\tilde{y}_t, \tilde{x}_t, \tilde{\theta})^2}{h_{t,i}} \right) \quad (93)$$

### 3. INTERNATIONAL COMPARISON PROGRAM DATA

The expenditure and price data are retrieved from 2011 round of the World Bank International Comparison Program (ICP). The ICP 2011 gathered 199 countries and its ultimate objective is to allow international comparisons for the different GDP components. As the national GDPs are computed in local currency prices, a conversion to a common unit must be made before a direct comparison can be undertaken. The use of exchange rates is not satisfactory. Two reasons assist here. Firstly, the non-traded goods and services are not taken-into-account. Most construction, government and market services for instance do not act on the international trade stage. As non-traded goods and services prices are normally higher in high income economies, this distortion is bound to lead to the overestimation of the gap between low-income and high-income countries. Secondly, exchange rates are susceptible of being influenced by erratic short-term capital flows such as speculation, interest rates and monetary policy, decoupling its pricing from medium term international trade economic fundamentals. To tackle these issues, the ICP delivers Purchasing Power Parities (PPPs). A PPP is equal to the local currency units needed to buy a specific basket of goods and services that can be exactly bought in the base country with just one unit of its currency. As the baskets used comprise non-traded goods and services, PPPs are supposed to be sounder currency converters than exchange rates. PPPs are first calculated for individual goods and services, and then for the various levels of aggregation up to GDP (International Comparison Program, 2011; World Bank, n.d.).

PPPs are not comparable because they are denominated in national currencies. Comparing the price levels of two countries requires the conversion of PPPs to the same currency unit. Here, the ICP standardizes the PPPs into Price Level Indexes (PLIs) by dividing them by the U.S. dollar exchange rate. As we use them as research inputs, it is advisable to retain that PLIs changes are normally instigated by exchange rates fluctuations. Finally, the real expenditures per capita research inputs come from using PPPs to convert nominal into real expenditures (International Comparison Program, 2011).

The ICP split GDP into 155 basic headings. Within each basic heading, individual product PPPs are calculated for each pair of countries. The ICP uses multilateral PPPs ensuring that direct and indirect computation via a base country produces the same result. Furthermore, a change in the group of countries included will induce a change in each pair PPP (International Comparison Program, 2011).

There are several multilateral PPPs computational methods. The choice criteria focus on two main characteristics. Firstly, PPPs are transitive whenever the PPP between two countries does not change with the type of computation, either direct or indirect through a third country. Secondly, PPPs are base country invariant if the PPP between two countries is independent from the base country choice (International Comparison Program, 2011).

From basic headings aggregation level on, the PPPs are weighted by expenditures. In each step two PPP are computed, first using the weights of the base country (Laspeyres index) and then using the weights of the other (Paasche index). The ICP takes the geometric mean of the two aggregated PPPs (Fisher index) to undertake the comparisons, allowing countries to be treated symmetrically. More precisely, the multilateral PPP is the geometric mean of the direct and indirect Fisher indexes. However, symmetry of treatment does not go with additivity. Indeed, the real expenditure computed from PPPs does not equal to the simple sum of each real expenditure belonging to the very same

aggregate (International Comparison Program, 2011).

Ensuring the PPPs and real expenditures per capita are compatible throughout the first and the second stage demand systems estimation requires using an additive method. Thus, the Paasche PPPs are computed for the first stage.

The ICP analytical categories that are aggregated in the first stage demand system are Food and nonalcoholic beverages; Alcoholic beverages, tobacco and narcotics; Clothing and footwear; Housing, water, electricity, gas and other fuels; Furnishings, household equipment and maintenance; Health; Transport; Communication; Recreation and culture; Education; Restaurants and hotels; Miscellaneous goods and services.

In the second stage, the aggregation involves the twenty-nine ICP Food and Nonalcoholic Beverages basic headings.

First and second stage datasets are shown in 9.1 Annex I and 9.2 Annex II respectively.

## 4. METHODOLOGIC FRAMEWORK

The estimation general procedure encompasses an estimation method, the algorithm initial values and the minimization method used to approximate the minimum of the objective function. The Levenberg-Marquardt method is used in case the Gauss Newton method fails to converge.

The additivity constraint  $\sum_1^n Wi = 1$  implies that the cross equations covariance matrix is singular. To overcome this issue, one equation is dropped from estimation and their estimates are computed from the output given by the others. (A. P. Barten, 1969) proved that the estimates are invariant to the chosen equation.

In the right side of the structural model equations there is not any dependent variable. Therefore, the NL2SLS and NL3SLS are not appropriate.

When not explicitly mentioned, the level of significance of the statistical tests is 5%.

The code was developed in SAS 9.4.

The estimation of the two-stage demand system for food and non-alcoholic beverages from the ICP 2011 data has the following methodologic steps:

- A. Choose the first stage broad categories of goods and services by aggregating the ICP analytical categories. Then choose the second stage subcategories of ICP Food and Non- alcoholic beverages analytical category by aggregating Food and Non- alcoholic beverages basic headings;
- B. Fit the Florida-PI model to the first stage aggregated data using the ITNLSUR. Take the ITNLLS to estimate the initial parameters values. Drop from the dataset those countries with an information inaccuracy measure greater than 0.1 (Seale & Regmi, 2006);
- C. Repeat the estimation procedure of the latter step without outliers in the dataset. Compute the mean of square residuals for each ICP joining date groups and across equations. Compare the results and decide about the final heteroskedastic group aggregation;
- D. Repeat the estimation procedure again including now the heteroskedasticity correction described in section 2.3.4.2 Heteroskedasticity model. Analyze the T-test of statistical significance of the estimated parameters and evaluate the goodness of the group and the scale heteroskedasticity model (SAS code in Annex III);
- E. Observe the empirical distributions of the residuals produced in the latter step and perform a Shapiro-Wilk test for the normality of the equations disturbances. In case the null is not rejected, fit the Florida-PI model to the first stage aggregated data without outliers using NLFIML. Take as initial values the estimates of the latter step. Include the heteroskedasticity model;
- F. Assume the same outliers and heteroskedastic model of first stage and fit the Florida Slutsky model to the second stage subcategories. Use the ITNLSUR, taking ITNLLS to estimate the initial parameters values.
- G. Observe the empirical distributions of the residuals produced in the latter step and perform a Shapiro-Wilk test for the normality of the equations disturbances. In case the null is not rejected, fit the Florida-Slutsky model to the second stage aggregated data without outliers using NLFIML. Take as initial values the estimates of the latter step and include the heteroskedasticity model;
- H. With first stage and second stage estimates compute the countries expenditure and price elasticities of the food subcategories.

## 5. RESULTS AND DISCUSSION

### 5.1. FIRST STAGE AND SECOND STAGE AGGREGATION

Table I shows the first stage aggregation:

First stage group	ICP analytical categories	
1	1	Food and nonalcoholic beverages
2	3	Clothing and footwear
3	4	Housing, water, electricity, gas and other fuels
4	5	Furnishings, household equipment and maintenance
5	6	Health
6	7	Transport
	8	Communication
7	10	Education
8	2	Alcoholic beverages, tobacco, and narcotics
	9	Recreation and culture
	11	Restaurants and hotels
	12	Miscellaneous goods and services

The joining of Transport and Communication into the 6<sup>th</sup> group is natural. The group 8<sup>th</sup> is made of luxury categories. The assumption of strong separability of preferences does not seem to be contradicted by any specific feature of the groups structure.

As the second stage aggregation is undertaken under weak separability of preferences, the restrictions on grouping are minor. Table II shows the second stage groups:

Table II		
Second stage group	ICP food and nonalcoholic beverages basic headings	
1	1101111	Rice
	1101112	Other cereals, flour and other products
	1101181	Sugar
2	1101113	Bread
	1101114	Other bakery products
	1101115	Pasta products
3	1101121	Beef and veal
	1101122	Pork
	1101123	Lamb, mutton and goat
	1101124	Poultry
	1101125	Other meats and meat preparations
	1101131	Fresh, chilled or frozen fish and seafood
	1101132	Preserved or processed fish and seafood
4	1101141	Fresh milk
	1101142	Preserved milk and other milk products
	1101143	Cheese
	1101144	Eggs and egg-based products
	1101151	Butter and margarine
	1101153	Other edible oils and fats
5	1101161	Fresh or chilled fruit
	1101171	Fresh or chilled vegetables other than potatoes
	1101172	Fresh or chilled potatoes
6	1101162	Frozen, preserved or processed fruit and fruit-based products
	1101173	Frozen, preserved or processed vegetables and vegetable-based products
7	1101182	Jams, marmalades and honey
	1101183	Confectionery, chocolate and ice cream
	1101191	Food products nec
	1101211	Coffee, tea and cocoa
	1101221	Mineral waters, soft drinks, fruit and vegetable juices

Group I is the most important: staple grains, plus sugar. While grains are primary source of nutrition for most people, sweetness is among the most universally preferred flavors (Reed & McDaniel, 2006). The ongoing debate over whether or not sugar is addictive (Avena, Rada, & Hoebel, 2008). is at the very least strong evidence that it exhibits highly price inelastic demand, and should therefore be grouped with staple grains. Group IV groups together dairy products and eggs, as well as oil, which was done because preliminary examination of the data suggested similar consumption patterns.

## 5.2. OUTLIERS

Table III shows the twenty-seven outliers that were detected (see section 2.3.3).



C	Countries	First stage groups							Inaccuracy
		I	II	III	IV	V	VI	VII	
7	Cameroon	0.0025	0.0039	0.0797	0.0071	0.0143	0.0010	0.0143	0.1228
11	Comoros	0.0005	0.0748	0.0431	0.0046	0.0143	0.0143	0.0007	0.1522
15	Djibouti	0.0244	0.0209	0.0421	0.0005	0.0143	0.0135	0.0004	0.1159
16	Egypt, Arab Rep.	0.0211	0.0010	0.0376	0.0005	0.0070	0.0803	0.0002	0.1478
18	Ethiopia	0.0482	0.0001	0.0062	0.0089	0.0098	0.0143	0.0128	0.1003
20	Gambia, The	0.0015	0.0032	0.0891	0.0600	0.0294	0.0768	0.0057	0.2657
24	Kenya	0.0133	0.0374	0.0563	0.0000	0.0035	0.0043	0.0015	0.1164
25	Lesotho	0.0405	0.0253	0.0006	0.0121	0.0143	0.0059	0.0062	0.1049
26	Liberia	0.5615	0.0149	0.0208	0.0001	0.0143	0.0080	0.0233	0.6430
28	Malawi	0.0017	0.0514	0.0010	0.0112	0.0011	0.0011	0.0448	0.1124
40	Seychelles	0.0468	0.0000	0.0002	0.0005	0.0143	0.0472	0.0014	0.1105
41	Sierra Leone	0.0068	0.0046	0.0930	0.0227	0.0322	0.0210	0.0067	0.1870
56	Fiji	0.0000	0.0570	0.0136	0.0063	0.0032	0.0264	0.0019	0.1083
57	Hong Kong SAR, China	0.0034	0.0001	0.0053	0.0000	0.0002	0.2765	0.0002	0.2856
63	Maldives	0.0464	0.0143	0.0699	0.0009	0.0000	0.2023	0.0143	0.3481
72	Thailand	0.0011	0.0023	0.1790	0.0011	0.0001	0.0030	0.0036	0.1903
110	Malta	0.0007	0.0013	0.1114	0.0009	0.0001	0.0001	0.0001	0.1146
112	Montenegro	0.0113	0.0222	0.0476	0.0059	0.0143	0.0001	0.0547	0.1561
119	Russian Federation	0.0053	0.0069	0.2930	0.0002	0.0004	0.0012	0.0143	0.3213
147	Aruba	0.0437	0.0006	0.0351	0.0002	0.0143	0.0019	0.0143	0.1101
148	Bahamas, The	0.0787	0.0037	0.0175	0.0012	0.0005	0.0052	0.0019	0.1087
149	Barbados	0.0249	0.0143	0.0462	0.0143	0.0143	0.0451	0.0143	0.1733
151	Bermuda	0.0008	0.0541	0.0099	0.0000	0.0109	0.0888	0.0005	0.1651
157	Montserrat	0.0160	0.0143	0.0001	0.0008	0.0645	0.0471	0.0143	0.1571
161	Sint Maarten	0.0678	0.0004	0.0383	0.0011	0.0038	0.0001	0.0143	0.1258
164	Turks and Caicos Islands	0.0987	0.0018	0.0584	0.0022	0.0080	0.0239	0.0067	0.1998
176	Georgia	0.0023	0.0843	0.0512	0.0033	0.0109	0.0004	0.0046	0.1568

### 5.3. FIRST STAGE RESULTS

Table IV				
Nonlinear ITSUR Parameter Estimates				
Parameter	Estimate	Standard deviation	T value	Approx
				Pr >  t
$\alpha_1$	0.0821	0.0062	13.16	<.0001
$\alpha_2$	0.0447	0.0021	21.59	<.0001
$\alpha_3$	0.2347	0.0073	32.21	<.0001
$\alpha_4$	0.0534	0.0020	26.92	<.0001
$\alpha_5$	0.0775	0.0074	10.52	<.0001
$\alpha_6$	0.1659	0.0048	34.28	<.0001
$\alpha_7$	0.0186	0.0031	6.01	<.0001
$\alpha_8$	0.3231	0.0078	41.71	<.0001
$\beta_1$	-0.1332	0.0046	-28.98	<.0001
$\beta_2$	-0.0031	0.0015	-2.06	0.0415
$\beta_3$	0.0337	0.0051	6.66	<.0001
$\beta_4$	0.0007	0.0015	0.51	0.6096
$\beta_5$	0.0205	0.0055	3.74	0.0003
$\beta_6$	0.0176	0.0036	4.9	<.0001
$\beta_7$	-0.0016	0.0021	-0.78	0.4363
$\beta_8$	0.0655	0.0046	14.31	<.0001
$\phi$	-0.9153	0.0279	-32.85	<.0001
$h_1$	14.2997	0.3961	36.1	<.0001
$h_2$	25.5463	1.1043	23.13	<.0001
$h_3$	-6.7129	0.1555	-43.18	<.0001
$\delta_1$	0.0037	0.0000	105.84	<.0001
$\delta_2$	0.0034	0.0001	36.02	<.0001
$\delta_3$	0.0037	0.0000	90.56	<.0001
$\delta_4$	0.0036	0.0001	31.77	<.0001
$\delta_5$	0.0039	0.0000	96.73	<.0001
$\delta_6$	0.0037	0.0001	73.5	<.0001
$\delta_7$	0.0039	0.0002	24.65	<.0001

The estimated  $\alpha$ s, which are the United States budget shares at the countries price indexes geometric mean, compare quite well with the actual US budget shares.

Needs are characterized by negative  $\beta$ s, and wants by positive ones. As expected, the first (food and nonalcoholic beverages) and the second (clothing and footwear) groups, which are formed mainly by needs, have negative estimated  $\beta$ s. The fourth (furnishings, household equipment and maintenance) and the seventh group (education) have estimated  $\beta$ s that are not statistically significant, indicating expenditure elasticities around unity. This is quite acceptable because both needs and wants can be found in the groups and they are less essential than the first and second groups. The third (housing,

water, electricity, gas and other fuels), the fifth (health), the sixth (transport and communication) and the eighth (a set of luxury items) groups have positive estimated  $\beta$ s. The initial units of housing, water and energy are obviously needs, not wants. However, once basic needs are satisfied, all of these do become luxury goods. Excessively large homes, the energy required to power them, and drinking water used for swimming pools, lawns and flush toilets are clearly not physiological needs. Empirical evidence suggests that house size is often a positional good, used to convey status (Wei, Zhang, & Liu, 2017). Similar results were reached by (Muhammad et al., 2011).

The estimate of  $\phi$  (income flexibility) is negative as expected and guarantees the negative semidefiniteness of the Slutsky matrix. The value is comparable to that found by (Muhammad et al., 2011; Seale & Regmi, 2006).

Table V below shows the results of the heteroskedasticity tests.

Table V					
Heteroskedasticity Test					
Equation	Test	Statistic	DF	Pr > Chi Sq	Variables
I	White's Test	140.10	58	<.0001	Cross of all vars
	Breusch-Pagan	1.84	1	0.18	1, q
II	White's Test	143.90	59	<.0001	Cross of all vars
	Breusch-Pagan	0.00	1	0.95	1, q
III	White's Test	144.50	61	<.0001	Cross of all vars
	Breusch-Pagan	10.86	1	0.00	1, q
IV	White's Test	136.60	57	<.0001	Cross of all vars
	Breusch-Pagan	1.01	1	0.31	1, q
V	White's Test	150.00	58	<.0001	Cross of all vars
	Breusch-Pagan	4.20	1	0.04	1, q
VI	White's Test	142.60	59	<.0001	Cross of all vars
	Breusch-Pagan	4.57	1	0.03	1, q
VII	White's Test	145.30	59	<.0001	Cross of all vars
	Breusch-Pagan	3.38	1	0.07	1, q

The estimates for the heteroskedasticity model parameters are all statistical significant. The parameters  $h_1$  and  $h_2$  refer to the heteroskedasticity group II and group III respectively and the parameter  $h_3$  is associated with the scale heteroskedasticity (see section 2.3.4). The estimates of  $\delta_i$  refer to homoskedastic standard deviation. The  $h_1$  and  $h_2$  estimates are positive and statistical significant as expected (Seale & Regmi, 2006). The  $h_3$  estimate is negative highlighting the fact that more affluent countries tend to have low variances.

The Maximum likelihood procedure was not undertaken because the Shapiro Wilk tests (Table VI below) to the final ITSUR residuals had rejected the null for all system equations.

Table VI		
Shapiro Wilk W Test		
Equation	Value	Prob
one	0.96	0.0002
two	0.94	<.0001
three	0.97	0.0026
four	0.95	<.0001
five	0.48	<.0001
six	0.93	<.0001
seven	0.93	<.0001

## 5.4. SECOND STAGE RESULTS

Parameter	Estimate	Standard deviation	T value	Approx	Parameter	Estimate	Standard deviation	T value	Approx
				Pr >  t					Pr >  t
$\alpha_1$	0.0817	0.0058	14.01	<.0001	$\delta_6$	0.0076	0.0003	29.94	<.0001
$\alpha_2$	0.1110	0.0044	25.38	<.0001	$\pi_{11}$	-0.1375	0.0130	-10.57	<.0001
$\alpha_3$	0.2675	0.0088	30.39	<.0001	$\pi_{12}$	-0.0147	0.0094	-1.56	0.1207
$\alpha_4$	0.1576	0.0057	27.69	<.0001	$\pi_{13}$	-0.0268	0.0106	-2.53	0.0124
$\alpha_5$	0.1410	0.0053	26.43	<.0001	$\pi_{14}$	0.0431	0.0112	3.85	0.0002
$\alpha_6$	0.0487	0.0038	12.71	<.0001	$\pi_{15}$	0.0356	0.0105	3.38	0.0009
$\alpha_7$	0.1926	0.0092	20.87	<.0001	$\pi_{16}$	-0.0167	0.0080	-2.11	0.0368
$\beta_1$	-0.0898	0.0102	-8.83	<.0001	$\pi_{22}$	-0.0724	0.0152	-4.75	<.0001
$\beta_2$	0.0230	0.0075	3.06	0.0027	$\pi_{23}$	0.0917	0.0090	10.24	<.0001
$\beta_3$	0.0355	0.0168	2.12	0.0361	$\pi_{24}$	0.0134	0.0120	1.12	0.2663
$\beta_4$	0.0060	0.0086	0.70	0.4874	$\pi_{25}$	0.0021	0.0114	0.19	0.8528
$\beta_5$	-0.0164	0.0100	-1.64	0.1036	$\pi_{26}$	0.0013	0.0101	0.13	0.9005
$\beta_6$	0.0172	0.0057	3.00	0.0032	$\pi_{33}$	-0.2013	0.0213	-9.46	<.0001
$\beta_7$	0.0245	0.0159	1.54	0.1264	$\pi_{34}$	0.0653	0.0105	6.22	<.0001
$h_1$	36.2486	1.1167	32.46	<.0001	$\pi_{35}$	0.0257	0.0107	2.40	0.0176
$h_2$	33.6687	2.4539	13.72	<.0001	$\pi_{36}$	-0.0044	0.0072	-0.61	0.5414
$h_3$	-3.7705	0.0416	-90.69	<.0001	$\pi_{44}$	-0.1834	0.0192	-9.56	<.0001
$\delta_1$	0.0097	0.0002	43.11	<.0001	$\pi_{45}$	-0.0273	0.0134	-2.04	0.0431
$\delta_2$	0.0063	0.0001	50.95	<.0001	$\pi_{46}$	0.0286	0.0113	2.53	0.0126
$\delta_3$	0.0072	0.0001	100.07	<.0001	$\pi_{55}$	-0.1384	0.0168	-8.25	<.0001
$\delta_4$	0.0073	0.0001	51.77	<.0001	$\pi_{56}$	0.0238	0.0092	2.58	0.0108
$\delta_5$	0.0053	0.0001	77.05	<.0001	$\pi_{66}$	-0.0652	0.0130	-5.00	<.0001

The estimate of  $\alpha_7$  is 0.1926 and contrasts with the actual US budget share for the seventh second stage group of luxury food items of 0.348. To a significant extent, the other shares are reasonable comparable. The higher level of heterogeneity found in the second stage dataset can explain this kind of divergence.

As expected, the first (unprocessed grains and sugar) group has a negative  $\beta$  estimate. The fourth (dairy, eggs and fats), the fifth (fresh vegetables and fruits) and the seventh (luxury items) group have  $\beta$ s estimates that are not statistical significant, indicating expenditure elasticities around unity. The seventh group  $\beta$  is positive as expected but should also be statistically significant. This contradiction is perhaps associated with some developed world addiction to high calories foods that induces a buying behavior closer to that related to needs instead of wants. The second (processed grains), the third (meat and fish) and the sixth (processed vegetables and fruits) have positive  $\beta$ s estimates. Similar results were reached by (Muhammad et al., 2011).

The symmetry of the Slutsky matrix is imposed. The estimates of its diagonal elements  $\pi_{ii}$  are negative and the matrix is negative semidefinite as expected. Therefore, the compensated law of demand is obeyed (Mas-Colell et al., 2006).

As for first stage, the heteroskedasticity model parameters are all statistical significant. In the second stage the Group II variance is higher than Group III.

The Maximum likelihood procedure was not undertaken because the Shapiro Wilk tests (Table VIII below) to the final ITSUR residuals had rejected the null for all system equations.

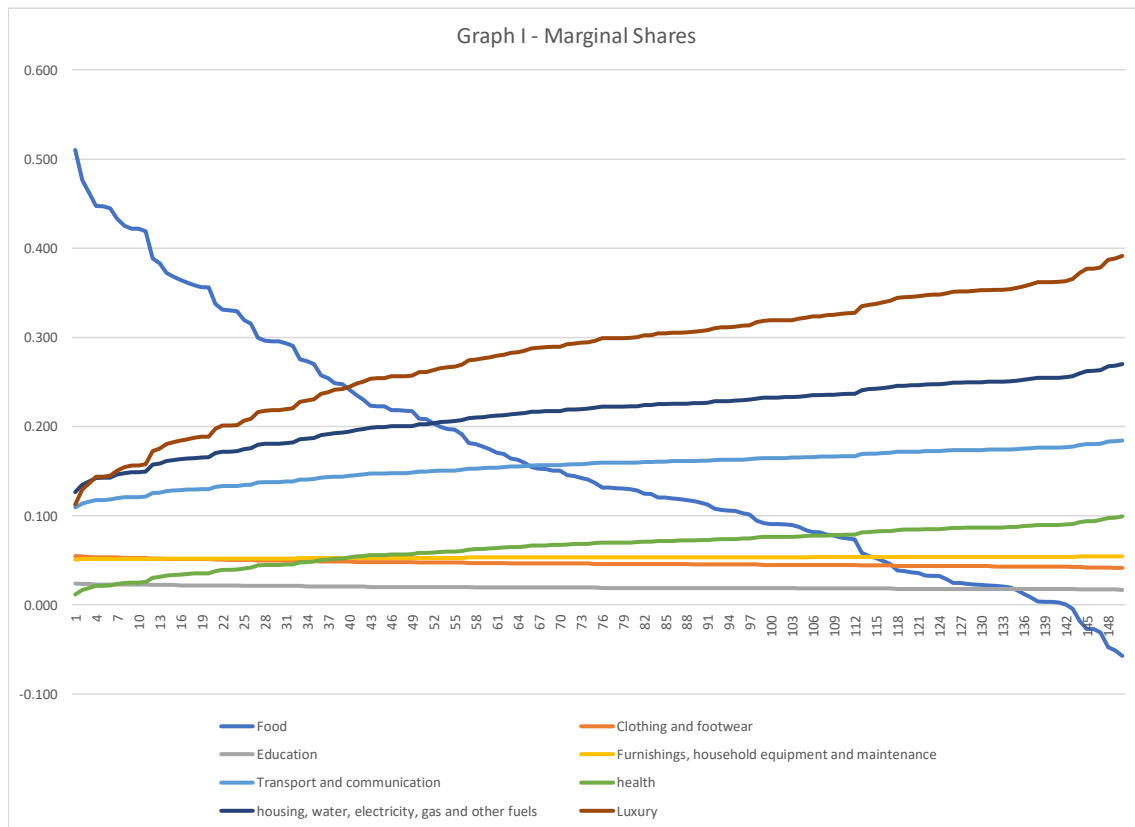
Shapiro Wilk W Test		
Equation	Value	Prob
one	0.91	<.0001
two	0.84	<.0001
three	0.81	<.0001
four	0.88	<.0001
five	0.91	<.0001
six	0.75	<.0001

### 5.5. FIRST STAGE DISCUSSION

Income category	#	Food and nonalcoholic beverages	Clothing and footwear	Housing, water, electricity, gas and other fuels	Furnishings, household equipment and maintenance	Health	T & C	Education	Luxuries
Low	53	0.466	0.053	0.127	0.050	0.035	0.107	0.026	0.135
Middle	49	0.260	0.051	0.183	0.055	0.042	0.168	0.024	0.217
High	48	0.144	0.051	0.229	0.055	0.042	0.159	0.022	0.297

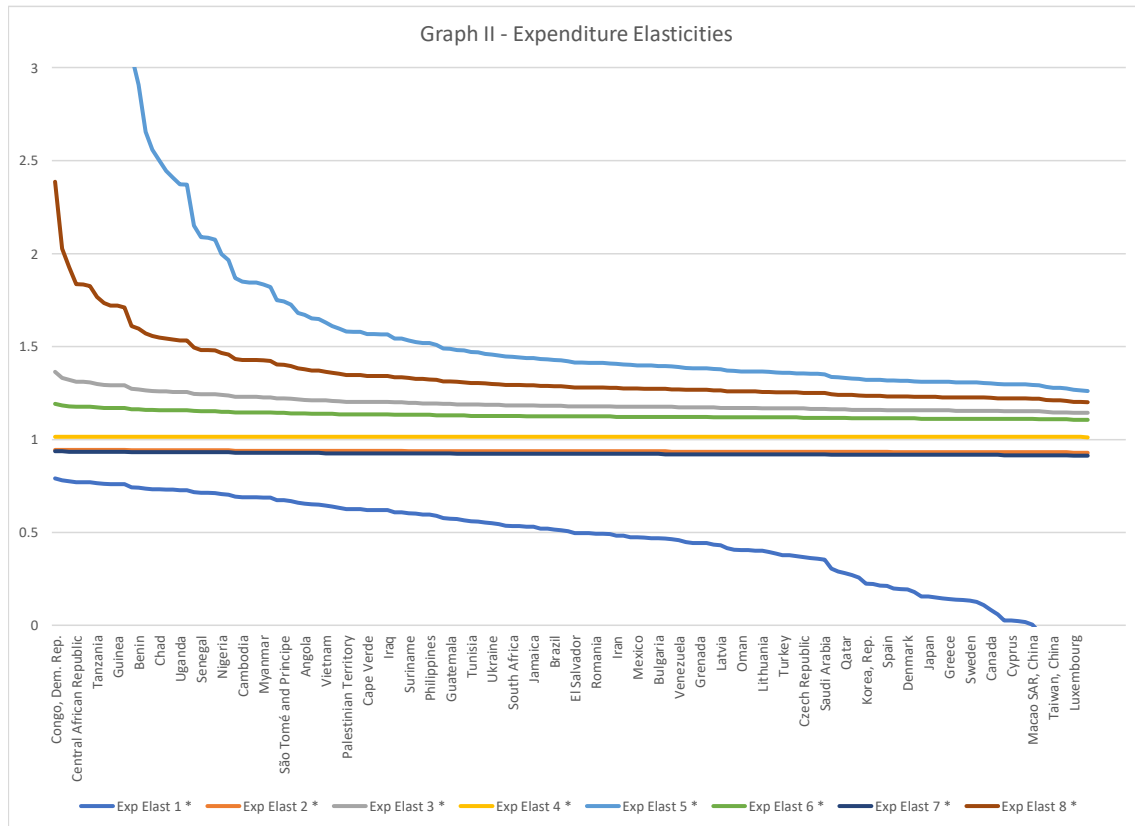
The geometric mean prices budget shares estimates are all plausible. The group III (housing, water, electricity, gas and other fuels) appears as a luxury category. Clearly, up to a minimum required level, shelter, water and energy are essential and non-substitutable. However, once basic needs are satisfied, it is reasonable to view additional consumption as a luxury, as discussed in section 5.3. In fact, (Steinberger, 2016) shows that the high correlation between lower energy and lower Human Development Index (HDI) abates as energy use increases and at higher energy use, there is evidence of saturation and no statistically significant dependency is found. After achieving energy use levels of about 1/4 -1/3 of those in the US, there seems to be negligible real benefits.

Perhaps the most important implication of these results concerns efforts to achieve ecological sustainability. Agriculture, which accounts for the vast majority of food production, is one of the greatest threats to global ecosystems (Foley et al., 2011; Godfray et al., 2010). Achieving ecological sustainability will require a reduction in these negative impacts. In an unequal capitalist economy, Pigouvian taxes that internalize the ecological costs of food production will directly increase food prices, while regulations that reduce the use of agrochemicals or agricultural lands will reduce food production, in response to which prices will rise. In either case, the real need of imposing limits to a growing economy and its implications on increasing food prices will have greater impact on the poor than the rich because the food shares of the former are higher, and rising food prices will severely reduce their real income.



Graph I plots the geometric mean price indexes marginal shares estimates against real expenditures per capita. The graphic shows that the proportion of each additional dollar of income spent on wants (luxury goods and services) will increase with affluence, while the share spent on needs (essential goods and services) will decrease. The graphic shows that luxuries shares increase with affluence and essential goods and services decrease. The only anomaly is the richest countries negative marginal shares for food. The Florida demand models were developed following the Chicago School assumption that tastes are constant across countries. The models use only one Slutsky matrix for every country across a seventy-fold range of real expenditures. Perhaps, in future improvements to the model, the dataset should be split into affluent and less affluent countries to acknowledge tastes differences across countries. As affluence increases the food buying decision becomes less important and the consumer can be unwilling to make the effort that is required to achieve a rational decision. (Keen, 2011) notes that there are over 1030 possible options for filling our shopping carts in a typical supermarket. If we could compare the utility of a trillion carts per second, it would take a billion years to choose the optimal one. Perfect optimization is obviously impossible, and people simply strive to 'satisfice', or find an acceptable rather than optimal market basket (Simon, 1956). As income increases, the opportunity costs of comparing options also rises, and it becomes rational to spend less time comparing marginal utilities. Behavior will increasingly deviate from conventional assumptions. In addition, richer societies are facing overfeeding issues and appear to be in a process of adjustment to a healthier nutrition. Finally, since the geometric mean food price index is lower than the rich countries' actual food price indexes, it is reasonable to expect that actual marginal shares are much higher.

In a perfect allocation system food marginal shares would be equal for every country. The discrepancies are an indicator of the market distortion. A redistribution towards poorer countries would increase total utility. Even if a pareto optimum is in place in every economy (normally it is not), the solution captured by the model seems to be far from a global optimum.



Graph II plots expenditure elasticities against real expenditures per capita. Expenditure elasticities for wants should be greater than one, and for needs, less than one, as seen in this figure. The expenditure elasticities are higher in poorer countries. This reflects diminishing marginal utility of expenditure and the non-homothetic preferences, both embedded in the linear-logarithm specification of the Working's model. Expenditure elasticities are particularly high for health and luxuries in the poorest countries, precisely because actual expenditures are so low. Even a tiny nominal increase in expenditures is a large percentage increase.

When the preferences are homothetic the  $\beta_i$  is zero, the shares are equal for every country and the expenditure elasticities are unitary and independent from affluence. That's the reason behind the flat curves that describes the second (processed grains), the fourth (furnishings, household equipment and maintenance) and the seventh group (education).

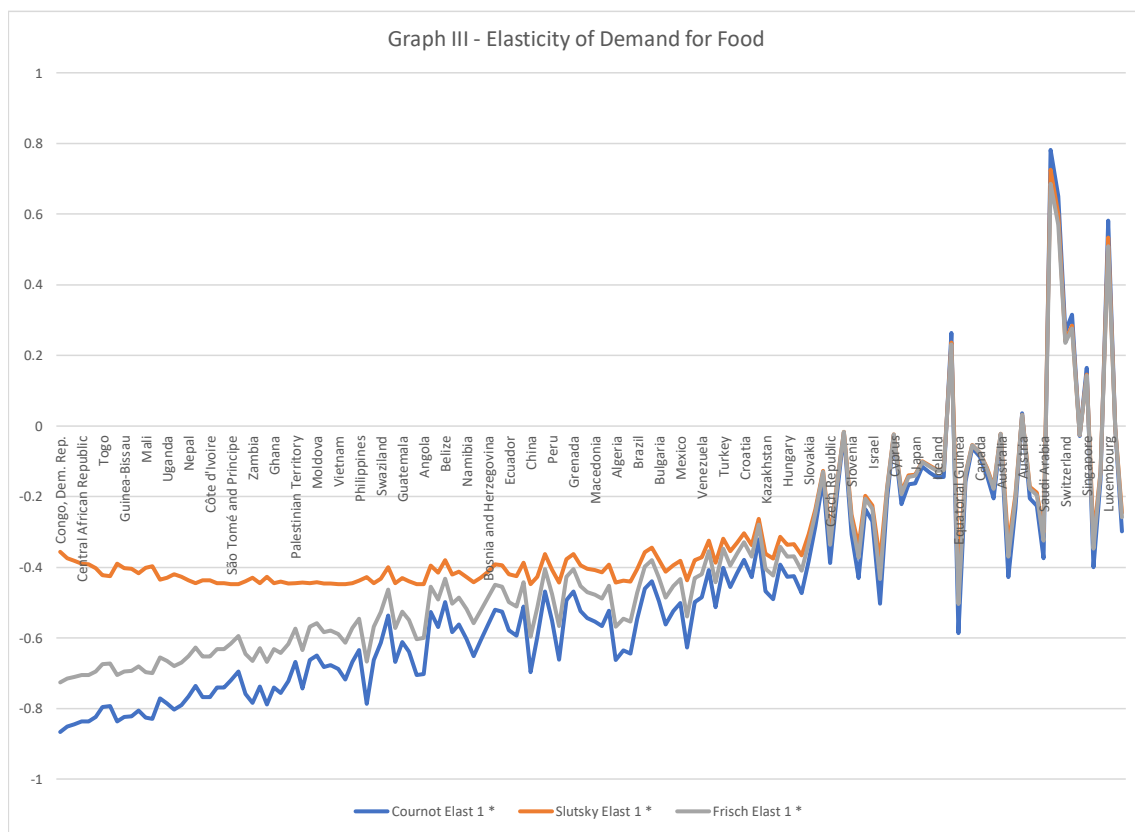
Estimated geometric mean price index budget shares for group V (health) were negative in some of the poorest countries (DR Congo, Niger and Burundi). One possible reason is the actual health care price index in those countries being lower than the geometric mean of reference. Nevertheless, there are significant discrepancies in health care systems across countries, which range from public sector provision, through public sector provision supplemented by private provision of luxury benefits, to



mainly private provision. Furthermore, the health care sector is strongly characterized by natural monopolies, legal barriers to entry, and government enforced monopolies on intellectual property (Farley & Farley, 2013). With such public intervention and market concentration, price indexes are far from perfect for ascertaining real costs of health systems across countries. This heterogeneity and the imposing restriction on tastes can explain the negative shares that were found.

Some richer countries have negative expenditure elasticities for the group I (food). This issue is due to negative marginal shares that were discussed above.

In a utility maximizing allocation system, food expenditure elasticities would be equal for every country. A redistribution towards poorer countries would increase total utility. Discrepancies in food expenditure elasticities is clear evidence of a non-utility maximizing allocation, defined here as a market distortion.



Graph III plots the price elasticities of demand for food against real expenditure per capita. The Slutsky elasticities are compensated elasticities derived from the Hicksian demand function where real income and total utility remains constant. The Frisch elasticities are compensated elasticities where marginal utility of income remains constant. The Cournot elasticities are uncompensated elasticities derived from the Walrasian demand function where nominal income is constant and real income and total utility vary freely. Following the descending order of income restrictions, the Slutsky elasticity is the highest, the Frisch elasticity comes next and the Cournot elasticity last.

Positive demand elasticities for food for some rich countries are in line with the negative expenditure elasticities found and discussed above.

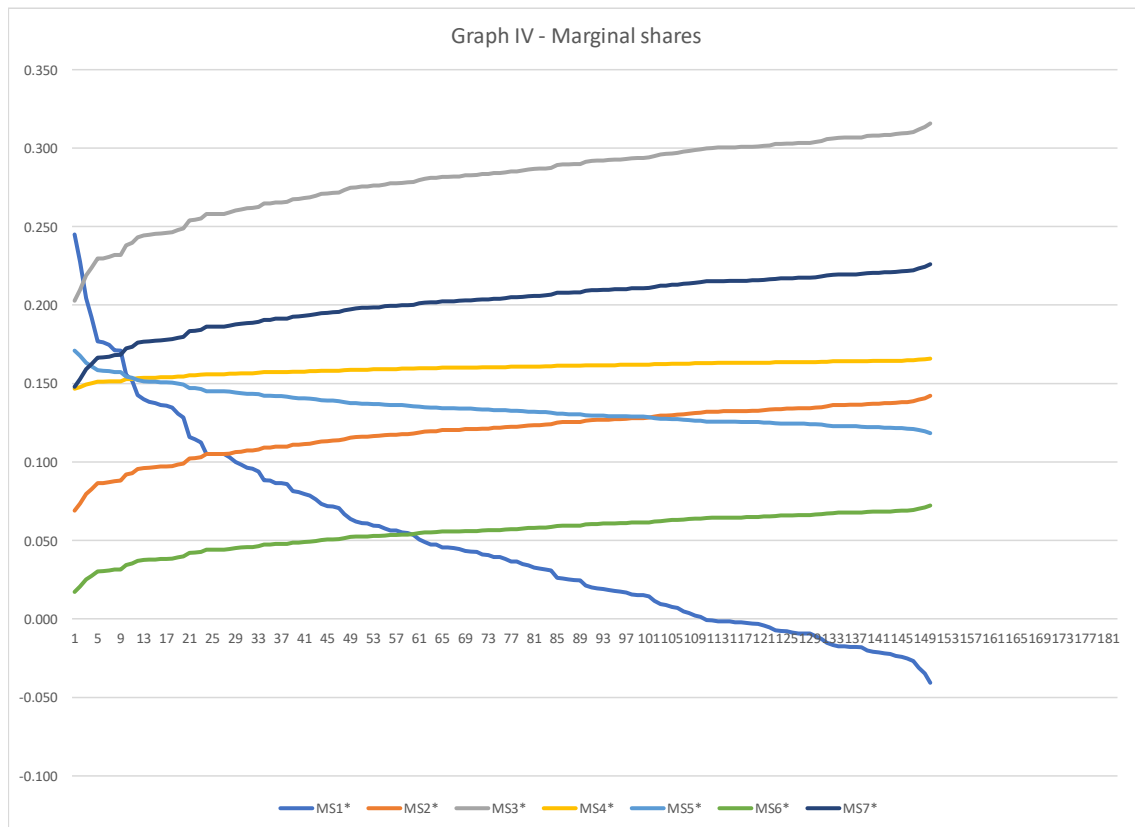
In real life, there are few mechanisms available to ensure that real income, total utility or marginal utility remain constant as the prices of food or other commodities increase. To ascertain the consequences of imposing ecological thresholds on the economic system, whose main impact would be increasing food prices, Cournot elasticities are the most relevant.

It is of utmost importance to know what are the implications on the real income of the poor if ecological thresholds are put in place through a market based mechanism. The Cournot elasticity estimates make evident that the poorer countries have more elastic demand curves, resulting primarily from the impact of increasing prices on real income, since there are no substitutes for food. This means that in an unequal market economy, if market based instruments are used to reduce the ecological degradation caused by food production, the poor will reduce consumption by a much greater percentage than the rich in response to price increases. Since the rich are responsible for far more ecological degradation than the poor, this outcome is highly perverse. This distortion is associated with the food shares and the marginal food shares that are higher in less affluent countries.

## 5.6. SECOND STAGE DISCUSSION

Income category	#	rice, cereals, sugar	bread, baked, pasta	meats and seafood	eggs, milk, butter and fats	fresh fruits and veggies	froz fruits veg	luxury
Low	53	0.244	0.062	0.228	0.136	0.171	0.030	0.129
Middle	49	0.102	0.107	0.258	0.165	0.159	0.033	0.176
High	48	0.057	0.117	0.272	0.166	0.153	0.039	0.196

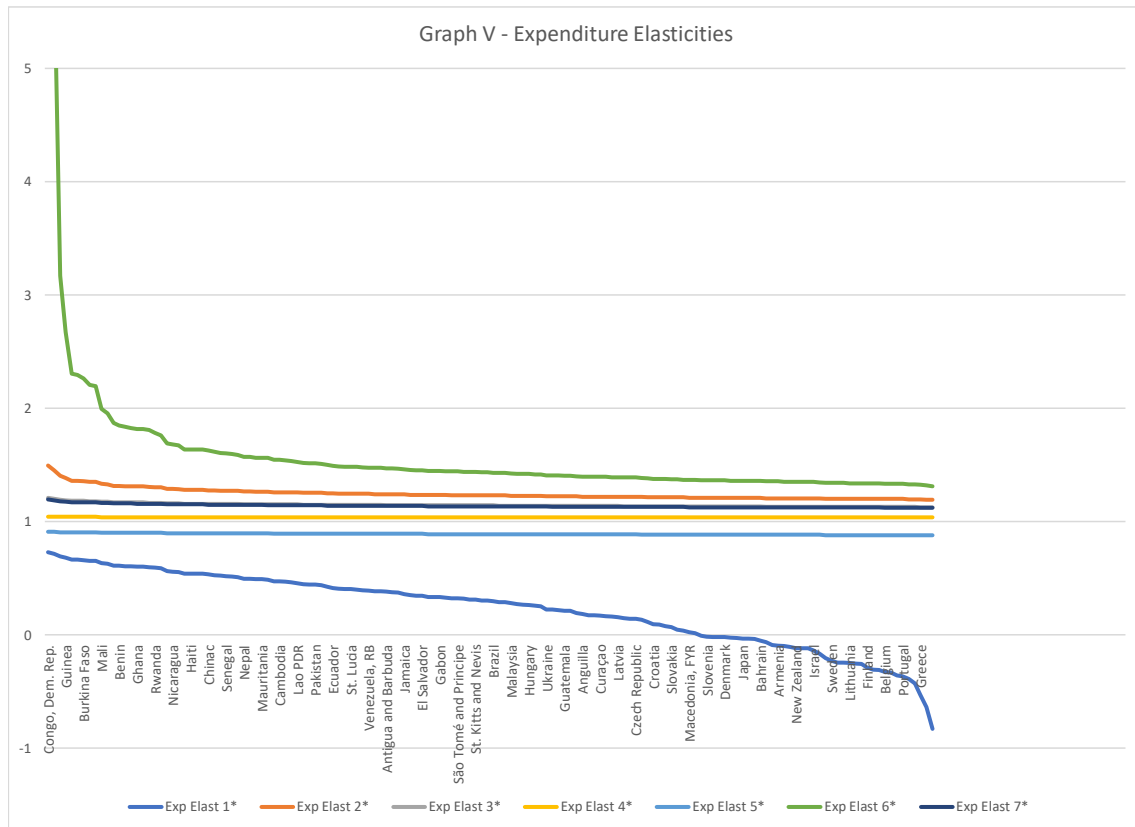
The geometric mean prices budget shares estimates are all plausible. Within the first stage food category, the second stage subcategory most directed implicated in the process of imposing ecological limits to the economic system is rice, cereals and sugar where the poorer have a higher budget share. Again, the impact on the real income of increasing grain prices would be stronger in poorer countries and perhaps even more intense than can be perceived by first stage conclusions.



Graph IV plots the geometric mean price indexes marginal shares estimates against real expenditures per capita. The graphic shows that luxuries shares increase with affluence and essential goods and services decrease. The only anomaly is the richest countries' negative marginal shares for second stage group I (unprocessed grains and sugar). The models use only one Slutsky matrix for every country in a seventy-fold range of real expenditures. The geometric mean price index is lower than the rich countries actual food price indexes, it is therefore reasonable to expect that actual marginal shares are higher. Another alternative is that the negative marginal shares relate to the obesity epidemic and the shift away from starches and sugars that is just getting underway. For example, a recent study found significant declines in the consumption of sweetened beverages in the US in recent years (Bleich, Vercaemmen, Koma, & Li, 2017).

In a perfect allocation system all second stage groups would have equal marginal shares for every country. For more essential groups, such as group I (unprocessed grains and sugar) and group V (fresh vegetables and fruits) that despite no statistically significant has a negative estimated  $\beta$ , the marginal shares decreases with affluence. Here, a redistribution towards poorer countries would increase total utility. For less essential groups, the marginal shares increase with affluence. Yet a redistribution toward the richer will not increase total utility. In fact, needs take precedence to wants. Needs marginal utility equilibrium across countries and consumers should be achieved first and only then some room can be given to wants production and consumption. Theoretically, it can be postulated through the two following assumptions. First, the marginal utility of needs is highest for low levels of income and is ever decreasing. Secondly, the marginal utility of wants is zero whenever the marginal utility of needs is positive, reaches a maximum when marginal utility of needs is zero and then becomes

ever decreasing. Thus, in every circumstance supplying wants when needs have positive marginal utilities is an overwhelming indication of market distortion.



Graph V plots conditional expenditure elasticities against real expenditures per capita. They are all plausible. The conditional elasticities show that group VI (processed vegetables and fruits) and group II (processed grains) have the highest expenditure elasticity for the poorest nations and are luxury goods. Demand for processed fruits and vegetables, processed grains and group VII (miscellaneous luxury foods) are all elastic with respect to income. This is especially true for lower income countries, where current expenditures are low, and even a small nominal increase is a large percentage increase. Demand for staple grains (plus sugar) and fresh fruits and vegetables are inelastic with respect to income, showing their role as necessities. Some richer countries have negative expenditure elasticities for the group I (unprocessed grains and sugar). This issue is due to negative marginal shares that were discussed above.

## 6. CONCLUSIONS

Perhaps the most important implication of these results concerns efforts to achieve ecological sustainability. Agriculture, which accounts for the vast majority of food production, is one of the greatest threats to global ecosystems (Foley et al., 2011; Godfray et al., 2010). Achieving ecological sustainability will require a reduction in these negative impacts. In an unequal capitalist economy, Pigouvian taxes that internalize the ecological costs of food production will directly increase food prices, while regulations that reduce the use of agrochemicals or agricultural lands will reduce food production, in response to which prices will rise. In either case, the real need of imposing limits to a growing economy and its implications on increasing food prices will have greater impact on the poor than the rich because the food shares of the former are higher, and rising food prices will severely reduce their real income.

In a perfect allocation system food marginal shares would be equal for every country. The discrepancies shown are an indicator of the market distortion. A redistribution towards poorer countries would increase total utility. Even if a pareto optimum is in place in every economy (normally it is not), the solution captured by the model seems to be far from a global optimum.

In a utility maximizing allocation system, food expenditure elasticities would be equal for every country. A redistribution towards poorer countries would increase total utility. Discrepancies in food expenditure elasticities is clear evidence of a non-utility maximizing allocation, defined here as a market distortion.

It is of utmost importance to know what are the implications on the real income of the poor if ecological thresholds are put in place through a market based mechanism. The Cournot elasticity estimates make evident that the poorer countries have more elastic demand curves, resulting primarily from the impact of increasing prices on real income, since there are no substitutes for food. This means that in an unequal market economy, if market based instruments are used to reduce the ecological degradation caused by food production, the poor will reduce consumption by a much greater percentage than the rich in response to price increases. Since the rich are responsible for far more ecological degradation than the poor, this outcome is highly perverse. This distortion is associated with the food shares and the marginal food shares that are higher in less affluent countries.

In a perfect allocation system all second stage groups would have equal marginal shares for every country. For more essential groups, such as group I (unprocessed grains and sugar) and group V (fresh vegetables and fruits), that despite no statistically significant has a negative estimated  $\beta$ , the marginal shares decreases with affluence. Here, a redistribution towards poorer countries would increase total utility. For less essential groups, the marginal shares increase with affluence. Yet a redistribution toward the richer will not increase total utility. In fact, needs take precedence to wants. Needs marginal utility equilibrium across countries and consumers should be achieved first and only then some room can be given to wants production and consumption. Theoretically, it can be postulated through the two following assumptions. First, the marginal utility of needs is highest for low levels of income and is ever decreasing. Secondly, the marginal utility of wants is zero whenever the marginal utility of needs is positive, reaches a maximum when marginal utility of needs is zero and then becomes

ever decreasing. Thus, in every circumstance supplying wants when needs have positive marginal utilities is an overwhelming indication of market distortion.

## **7. LIMITATIONS AND RECOMMENDATIONS FOR FUTURE WORKS**

The expenditure and consumption data retrieved from the World Bank International Comparison Program (ICP) does not come without problems. Poor data can be found in some African countries. As consumption of home-produced food in some low-income countries is normally understated, significant inaccuracies in the food expenditure estimates can prevail. The health care expenditures suffer from differences in the provision systems. More public state intervention usually leads to underestimated national expenditures.

The ICP methodology does not preview the fundamental partition between needs and wants. The ICP broad categories are a mix of needs and wants. Even the food category is not free from wants as it is the case for more affluent countries. The ecological economics framework does not fit perfectly with World Bank methodological decisions.

The Florida models assumption of constant tastes across countries is too restrictive. In further works a split between affluent and nonaffluent countries should be carried out. Nevertheless, the lack of degrees of freedom especially in the second stage estimation can then be an issue.

Based on the ICP basic headings expenditures, a more adequate aggregation can be done that tries to separate better needs from wants.

## 8. BIBLIOGRAPHY

- A. P. Barten. (1969). Maximum likelihood estimation of a complete system of demand equations. *European Economic Review*, 1(1), 7–73. Retrieved from [https://econpapers.repec.org/article/eeeeecrev/v\\_3a1\\_3ay\\_3a1969\\_3ai\\_3a1\\_3ap\\_3a7-73.htm](https://econpapers.repec.org/article/eeeeecrev/v_3a1_3ay_3a1969_3ai_3a1_3ap_3a7-73.htm)
- Akerlof, G. A., & Shiller, R. J. (2015). *Phishing for phools : the economics of manipulation and deception*. Princeton University Press. Retrieved from <https://press.princeton.edu/titles/10534.html>
- Amemiya, T. (1985). *Advanced Econometrics*. Cambridge, Massachusetts: Harvard University Press.
- Avena, N. M., Rada, P., & Hoebel, B. G. (2008). Evidence for sugar addiction: Behavioral and neurochemical effects of intermittent, excessive sugar intake. *Neuroscience & Biobehavioral Reviews*, 32(1), 20–39. <http://doi.org/10.1016/j.neubiorev.2007.04.019>
- Barnett, W. A., & Serletis, A. (2008). Consumer preferences and demand systems. *Journal of Econometrics*, 147(2), 210–224. <http://doi.org/10.1016/j.jeconom.2008.09.009>
- Bleich, S. N., Vercammen, K. A., Koma, J. W., & Li, Z. (2017). Trends in Beverage Consumption Among Children and Adults, 2003-2014. *Obesity*. <http://doi.org/10.1002/oby.22056>
- Deaton, A., & Muellbauer, J. (1980). *Economics and consumer behavior*. Cambridge University Press.
- Dhrymes, P. (1994). *Topics in Advanced Econometrics - Linear and Nonlinear Simultaneous Equations* (Vol. II). Springer Verlag.
- FAO, IFAD, & WFP. (2015). *The State of Food Insecurity in the World: Meeting the 2015 international hunger targets: taking stock of uneven progress*. FAO, IFAD and WFP. Rome. <http://doi.org/14646E/1/05.15>
- Farley, J., & Costanza, R. (2010). Payments for ecosystem services: From local to global. *Ecological Economics*, 69(11), 2060–2068. <http://doi.org/10.1016/j.ecolecon.2010.06.010>
- Farley, J., & Farley, T. (2013). Box Insert: The Economics of Health Care. *Solutions Journal*, 4(48).
- Farley, J., Schmitt Filho, A., Burke, M., & Farr, M. (2015). Extending market allocation to ecosystem services: Moral and practical implications on a full and unequal planet. *Ecological Economics*. <http://doi.org/10.1016/j.ecolecon.2014.06.021>
- Foley, J. a., Ramankutty, N., Brauman, K. a., Cassidy, E. S., Gerber, J. S., Johnston, M., ... Zaks, D. P. M. (2011). Solutions for a cultivated planet. *Nature*, 478(7369), 337–342. <http://doi.org/10.1038/nature10452>
- Frank, R. H. (2005). Positional Externalities Cause Large and Preventable Welfare Losses. *American Economic Review*, 95(2), 137–141. <http://doi.org/10.1257/000282805774670392>
- Frank, R. H. (2007). *Falling behind : how rising inequality harms the middle class*. University of California Press. Retrieved from <https://www.ucpress.edu/book.php?isbn=9780520280526>
- Frieden, J. A. (2006). *Global capitalism : its fall and rise in the twentieth century*. W.W. Norton.
- Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., ... Toulmin, C. (2010). Food Security: The Challenge of Feeding 9 Billion People. *Science*, 327(5967), 812–818.



<http://doi.org/10.1126/science.1185383>

- Greene, W. W. H. . (2012). *Econometric analysis*. Prentice Hall (Vol. 97). New York: Pearson.  
<http://doi.org/10.1198/jasa.2002.s458>
- Gunders, D. (2012). Wasted: How America is losing up to 40 percent of its food from farm to fork to landfill. *NRDC Issue Paper*, (August), 1–26. <http://doi.org/12-06-B>
- Gustavsson, J. Cederberg, C., Sonesson, U., Otterdijk, R. v., & Meybeck, A. (2011). *Global food losses and food waste - Extent, causes and prevention*. SAVE FOOD: An initiative on Food Loss and Waste Reduction. Rome: FAO. <http://doi.org/10.1098/rstb.2010.0126>
- Hayek, F. (1945). The Price System as a Mechanism for Using Knowledge. *American Economic Review*, (35), 519–530.
- International Comparison Program. (2011). Purchasing Power Parities and Real Expenditures of World Economies Summary of Results and Findings of the 2011 International Comparison Program.
- Kallis, G., Kerschner, C., & Martinez-Alier, J. (2012). The economics of degrowth. *Ecological Economics*, 84, 172–180. <http://doi.org/10.1016/j.ecolecon.2012.08.017>
- Keen, S. (2011). *Debunking economics : the naked emperor dethroned?* Zed Book.
- Kennedy, P. (2008). *A guide to econometrics*. Blackwell Pub.
- Lane, R. E. (2001). *The Loss of Happiness in Market Democracies*. New Haven: Yale University Press. Retrieved from <https://isps.yale.edu/research/publications/isps01-004>
- Lucas, R. E. (2004). *The Industrial Revolution: Past and Future*. 2003 Annual Report Essay. Retrieved from <https://www.minneapolisfed.org/publications/the-region/the-industrial-revolution-past-and-future>
- Marshall, A. (1890). *Principles of Economics*. New York: The Macmillan company.
- Martinez-Alier, J. (2016). Socially sustainable economic degrowth. In J. Farley & D. Malghan (Eds.), *Beyond uneconomic growth* (pp. 280–301). Edward Elgar Publishing. Retrieved from [https://econpapers.repec.org/bookchap/elgeechap/15668\\_5f13.htm](https://econpapers.repec.org/bookchap/elgeechap/15668_5f13.htm)
- Mas-Colell, A., Whinston, M. D., & Green, J. R. (2006). *Microeconomic theory*. Oxford University Press.
- McCauley, D. J. (2006). Selling out on nature. *Nature*, 443(7107), 27–28.  
<http://doi.org/10.1038/443027a>
- Muhammad, A., Meade, B., Regmi, A., & Seale, J. (2011). International Evidence on Food Consumption Patterns An Update Using 2005 International Comparison Program Data. *TB-1929*. U.S. Dept. of Agriculture, Econ. Res. Serv. March 2011. Revised Februari 2013.
- Nelles, O. (2001). *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*. Springer Verlag.
- OXFAM. (2017). An Economy for the 1%. *OXFAM Briefing Papers*, 210(January 2016), 1–44.  
<http://doi.org/ISBN 978-1-78077-993-5>

- Pareto, V. (1971). *Manual of political economy*. A.M. Kelley.
- Pigou, A. C. (1932). *The Economics of Welfare* (4th ed.). London: Macmillan.
- Piketty, T. (2014). *Capital in the twenty-first century*.
- Pindyck, R. S., & Rubinfeld, D. L. (1998). *Econometric models and economic forecasts*. Irwin/McGraw-Hill.
- Reed, D. R., & McDaniel, A. H. (2006). The Human Sweet Tooth. *BMC Oral Health*, 6(Suppl 1), S17. <http://doi.org/10.1186/1472-6831-6-S1-S17>
- Reimer, J., & Hertel, T. (2004). Estimation of International Demand Behaviour for Use with Input-Output Based Data. *Economic Systems Research*, 16(4), 347–366. <http://doi.org/10.1080/0953531042000304245>
- Rockström, J., Steffen, W., Noone, K., Persson, Ä., Chapin, F. S., Lambin, E., ... Foley, J. (2009). Planetary boundaries: Exploring the safe operating space for humanity. *Ecology and Society*, 14(2). <http://doi.org/10.1038/461472a>
- Saez, E. (2016). Striking it Richer: The Evolution of Top Incomes in the United States (Updated with 2015 preliminary estimates). <http://doi.org/10.1177/2043610616665033>
- SAS. (2014). *SAS/ETS 13.2 User's Guide*. Cary, NC, USA: SAS Institute Inc.
- Seale, J., & Regmi, A. (2006). MODELING INTERNATIONAL CONSUMPTION PATTERNS by James L. Seale, Jr. \*, (4).
- Seale, J., Regmi, A., & Bernstein, J. (2003). Evidence on Food Consumption Patterns. *Report from the Economic Research Service of the United States Department of Agriculture. Technical Bulletin Number 1904.*, (1904).
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3–4), 591–611. <http://doi.org/10.1093/biomet/52.3-4.591>
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*. US: American Psychological Association. <http://doi.org/10.1037/h0042769>
- Steffen, W., Grinevald, J., Crutzen, P., & McNeill, J. (2011). The Anthropocene: conceptual and historical perspectives. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1938), 842–867. <http://doi.org/10.1098/rsta.2010.0327>
- Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., Bennett, E. M., ... Sörlin, S. (2015). Planetary boundaries: guiding human development on a changing planet. *Science (New York, N.Y.)*, 347(6223). <http://doi.org/10.1126/science.1259855>
- Steinberger, J. K. (2016). *UNDP Human Development Reports - Energising Human Development. Human Developments Reports*. Retrieved from <http://hdr.undp.org/en/content/energising-human-development>
- Thaler, R. H. (2015). *Misbehaving : the making of behavioral economics*.
- Theil, H. (1980). *The system-wide approach to microeconomics*. Univ. of Chicago Press [u.a.].
- Theil, H. (1996). *Advanced Studies in Theoretical an Applied Econometrics: Studies in Global*

*Econometrics*. Kluwer Academic Publishers.

- Theil, H., Chung, C.-F., & L. Seale, J. (1989). *Advances in econometrics: international evidence on consumption patterns*. Greenwich, CT and London: JAI Press.
- Theil, H., Suhm, F. E., & Meisner, J. F. (1981). *International consumption comparisons : a system-wide approach*. North Holland Pub. Co.
- Tilman, D., Balzer, C., Hill, J., & Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Pnas*, *108*(50), 20260–4. <http://doi.org/10.1073/pnas.1116437108>
- Wackernagel, M., Schulz, N. B., Deumling, D., Linares, A. C., Jenkins, M., Kapos, V., ... Randers, J. (2002). Tracking the ecological overshoot of the human economy. *Proceedings of the National Academy of Sciences*, *99*(14), 9266–9271. <http://doi.org/10.1073/pnas.142033699>
- Wei, S.-J., Zhang, X., & Liu, Y. (2017). Home ownership as status competition: Some theory and evidence. *Journal of Development Economics*, *127*, 169–186. <http://doi.org/10.1016/j.jdeveco.2016.12.001>
- White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, *48*, 817–838.
- Wilkinson, R. G., & Pickett, K. (2009). *The spirit level : why greater equality makes societies stronger*. New York: Bloomsbury Press.
- Wooldridge, J. M. (2002). Econometric Analysis of Cross Section and Panel Data, *58*(2), 752. <http://doi.org/10.1515/humr.2003.021>
- Working, H. (1943). Statistical Laws of Family Expenditure. *Journal of the American Statistical Association*, (38), 43–56.
- World Bank. (n.d.). World Bank International Comparison Project. Retrieved from [http://siteresources.worldbank.org/ICPEXT/Resources/ICP\\_2011.html](http://siteresources.worldbank.org/ICPEXT/Resources/ICP_2011.html)















### 9.3. ANNEX III – SAS CODE FOR FINAL FIRST STAGE ITSUR ESTIMATION

```

Proc model data=thesis.FirstStageData2 outparms = thesis.parmsFirstFinal;
  endogenous W1 W2 W3 W4 W5 W6 W7;
  exogenous PLI1 PLI2 PLI3 PLI4 PLI5 PLI6 PLI7 q;
  parameters alfa1 alfa2 alfa3 alfa4 alfa5 alfa6 alfa7
             beta1 beta2 beta3 beta4 beta5 beta6 beta7 fi h1 h2 h3;
  ID C Countries Group;
  label W1='Food Nonalc. beverages Share' W2='Clothing Footwear Share'
        W3='Housing Water Electr. Gas Oth. fuels Share' W4='Furnishings Household equip.
Maintenance Share'
        W5='Health Share' W6='Transport Communication Share'
        W7='Education Share'
        PLI1='Food Nonalc. beverages PLI' PLI2='Clothing Footwear PLI'
        PLI3='Housing Water Electr. Gas Oth. fuels PLI' PLI4='Furnishings Household equip.
Maintenance PLI'
        PLI5='Health Share PLI' PLI6='Transport Communication PLI'
        PLI7='Education Share'
        q='LN Real Expenditure pc' fi='Income flexibility';
  estimate 'alfa8' 1-(alfa1+alfa2+alfa3+alfa4+alfa5+alfa6+alfa7),
           'beta8'
(beta1+beta2+beta3+beta4+beta5+beta6+beta7),/OUTEST=thesis.estimatesFirstFinal;

  eq.one=W1-(alfa1+beta1*q+(alfa1+beta1*q)*(gmPLI1-
((alfa1+beta1*q)*gmPLI1+(alfa2+beta2*q)*gmPLI2
+(alfa3+beta3*q)*gmPLI3+(alfa4+beta4*q)*gmPLI4+(alfa5+beta5*q)*gmPLI5
+(alfa6+beta6*q)*gmPLI6+(alfa7+beta7*q)*gmPLI7))
+fi*(alfa1+beta1*(1+q))*(gmPLI1-
((alfa1+beta1*(1+q))*gmPLI1+(alfa2+beta2*(1+q))*gmPLI2
+(alfa3+beta3*(1+q))*gmPLI3+(alfa4+beta4*(1+q))*gmPLI4+(alfa5+beta5*(1+q))*gmPLI5
+(alfa6+beta6*(1+q))*gmPLI6+(alfa7+beta7*(1+q))*gmPLI7));

  H.one=(sigma1**2)+(sigma1**2)*h1*D1+(sigma1**2)*h2*D2+(sigma1**2)*h3*q;

  eq.two=W2-(alfa2+beta2*q+(alfa2+beta2*q)*(gmPLI2-
((alfa1+beta1*q)*gmPLI1+(alfa2+beta2*q)*gmPLI2
+(alfa3+beta3*q)*gmPLI3+(alfa4+beta4*q)*gmPLI4+(alfa5+beta5*q)*gmPLI5
+(alfa6+beta6*q)*gmPLI6+(alfa7+beta7*q)*gmPLI7))
+fi*(alfa2+beta2*(1+q))*(gmPLI2-
((alfa1+beta1*(1+q))*gmPLI1+(alfa2+beta2*(1+q))*gmPLI2
+(alfa3+beta3*(1+q))*gmPLI3+(alfa4+beta4*(1+q))*gmPLI4+(alfa5+beta5*(1+q))*gmPLI5
+(alfa6+beta6*(1+q))*gmPLI6+(alfa7+beta7*(1+q))*gmPLI7));

  H.two=(sigma2**2)+(sigma2**2)*h1*D1+(sigma2**2)*h2*D2+(sigma2**2)*h3*q;

  eq.three=W3-(alfa3+beta3*q+(alfa3+beta3*q)*(gmPLI3-
((alfa1+beta1*q)*gmPLI1+(alfa2+beta2*q)*gmPLI2
+(alfa3+beta3*q)*gmPLI3+(alfa4+beta4*q)*gmPLI4+(alfa5+beta5*q)*gmPLI5

```

$$\begin{aligned}
& +(\alpha_6+\beta_6q)*gmPLI6+(\alpha_7+\beta_7q)*gmPLI7)) \\
& +fi*(\alpha_3+\beta_3(1+q))*(gmPLI3- \\
((\alpha_1+\beta_1(1+q))*gmPLI1+(\alpha_2+\beta_2(1+q))*gmPLI2 \\
& +(\alpha_3+\beta_3(1+q))*gmPLI3+(\alpha_4+\beta_4(1+q))*gmPLI4+(\alpha_5+\beta_5(1+q))*gmPLI5 \\
& +(\alpha_6+\beta_6(1+q))*gmPLI6+(\alpha_7+\beta_7(1+q))*gmPLI7)); \\
H.three=(\sigma_3^{**2})+(\sigma_3^{**2})*h1*D1+(\sigma_3^{**2})*h2*D2+(\sigma_3^{**2})*h3*q; \\
eq.four=W4-(\alpha_4+\beta_4q+(\alpha_4+\beta_4q)*(gmPLI4- \\
((\alpha_1+\beta_1q)*gmPLI1+(\alpha_2+\beta_2q)*gmPLI2 \\
& +(\alpha_3+\beta_3q)*gmPLI3+(\alpha_4+\beta_4q)*gmPLI4+(\alpha_5+\beta_5q)*gmPLI5 \\
& +(\alpha_6+\beta_6q)*gmPLI6+(\alpha_7+\beta_7q)*gmPLI7)) \\
& +fi*(\alpha_4+\beta_4(1+q))*(gmPLI4- \\
((\alpha_1+\beta_1(1+q))*gmPLI1+(\alpha_2+\beta_2(1+q))*gmPLI2 \\
& +(\alpha_3+\beta_3(1+q))*gmPLI3+(\alpha_4+\beta_4(1+q))*gmPLI4+(\alpha_5+\beta_5(1+q))*gmPLI5 \\
& +(\alpha_6+\beta_6(1+q))*gmPLI6+(\alpha_7+\beta_7(1+q))*gmPLI7)); \\
H.four=(\sigma_4^{**2})+(\sigma_4^{**2})*h1*D1+(\sigma_4^{**2})*h2*D2+(\sigma_4^{**2})*h3*q; \\
eq.five=W5-(\alpha_5+\beta_5q+(\alpha_5+\beta_5q)*(gmPLI5- \\
((\alpha_1+\beta_1q)*gmPLI1+(\alpha_2+\beta_2q)*gmPLI2 \\
& +(\alpha_3+\beta_3q)*gmPLI3+(\alpha_4+\beta_4q)*gmPLI4+(\alpha_5+\beta_5q)*gmPLI5 \\
& +(\alpha_6+\beta_6q)*gmPLI6+(\alpha_7+\beta_7q)*gmPLI7)) \\
& +fi*(\alpha_5+\beta_5(1+q))*(gmPLI5- \\
((\alpha_1+\beta_1(1+q))*gmPLI1+(\alpha_2+\beta_2(1+q))*gmPLI2 \\
& +(\alpha_3+\beta_3(1+q))*gmPLI3+(\alpha_4+\beta_4(1+q))*gmPLI4+(\alpha_5+\beta_5(1+q))*gmPLI5 \\
& +(\alpha_6+\beta_6(1+q))*gmPLI6+(\alpha_7+\beta_7(1+q))*gmPLI7)); \\
H.five=(\sigma_5^{**2})+(\sigma_5^{**2})*h1*D1+(\sigma_5^{**2})*h2*D2+(\sigma_5^{**2})*h3*q; \\
eq.six=W6-(\alpha_6+\beta_6q+(\alpha_6+\beta_6q)*(gmPLI6- \\
((\alpha_1+\beta_1q)*gmPLI1+(\alpha_2+\beta_2q)*gmPLI2 \\
& +(\alpha_3+\beta_3q)*gmPLI3+(\alpha_4+\beta_4q)*gmPLI4+(\alpha_5+\beta_5q)*gmPLI5 \\
& +(\alpha_6+\beta_6q)*gmPLI6+(\alpha_7+\beta_7q)*gmPLI7)) \\
& +fi*(\alpha_6+\beta_6(1+q))*(gmPLI6- \\
((\alpha_1+\beta_1(1+q))*gmPLI1+(\alpha_2+\beta_2(1+q))*gmPLI2 \\
& +(\alpha_3+\beta_3(1+q))*gmPLI3+(\alpha_4+\beta_4(1+q))*gmPLI4+(\alpha_5+\beta_5(1+q))*gmPLI5 \\
& +(\alpha_6+\beta_6(1+q))*gmPLI6+(\alpha_7+\beta_7(1+q))*gmPLI7)); \\
H.six=(\sigma_6^{**2})+(\sigma_6^{**2})*h1*D1+(\sigma_6^{**2})*h2*D2+(\sigma_6^{**2})*h3*q; \\
eq.seven=W7-(\alpha_7+\beta_7q+(\alpha_7+\beta_7q)*(gmPLI7- \\
((\alpha_1+\beta_1q)*gmPLI1+(\alpha_2+\beta_2q)*gmPLI2 \\
& +(\alpha_3+\beta_3q)*gmPLI3+(\alpha_4+\beta_4q)*gmPLI4+(\alpha_5+\beta_5q)*gmPLI5 \\
& +(\alpha_6+\beta_6q)*gmPLI6+(\alpha_7+\beta_7q)*gmPLI7))
\end{aligned}$$

```

+fi*(alfa7+beta7*(1+q))*(gmPLI7-
((alfa1+beta1*(1+q))*gmPLI1+(alfa2+beta2*(1+q))*gmPLI2
+(alfa3+beta3*(1+q))*gmPLI3+(alfa4+beta4*(1+q))*gmPLI4+(alfa5+beta5*(1+q))*gmPLI5
+(alfa6+beta6*(1+q))*gmPLI6+(alfa7+beta7*(1+q))*gmPLI7));

```

```

H.seven=(sigma7**2)+(sigma7**2)*h1*D1+(sigma7**2)*h2*D2+(sigma7**2)*h3*q;

```

```

fit one two three four five six seven

```

```

ESTDATA=thesis.parmsFitFirstFinalITOLS
ITSUR PRL=LR
METHOD= GAUSS
VARDEF=DF
MISSING=DELETE
OUTRESID
OUT=thesis.ResidualsFirstFinal
OUTEST=thesis.parmsFitFirstFinalITSUR OUTCOV
OUTS=thesis.EstimatedEqErrorsFirstFinal
BREUSCH=(1 q)
NORMAL WHITE
PRINTALL
ITPRINT
CONVERGE=(0.001,0.001)
NESTIT;

```

```

quit;

```