

Nonparametric welfare and demand analysis with unobserved individual heterogeneity

Citation for published version (APA):

Cosaert, S., & Demuyne, T. (2014). *Nonparametric welfare and demand analysis with unobserved individual heterogeneity*. Maastricht University, Graduate School of Business and Economics. GSBE Research Memoranda No. 010 <https://doi.org/10.26481/umagsb.2014010>

Document status and date:

Published: 01/01/2014

DOI:

[10.26481/umagsb.2014010](https://doi.org/10.26481/umagsb.2014010)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Sam Cosaert, Thomas Demuynck

**Nonparametric welfare and
demand analysis with
unobserved individual
heterogeneity**

RM/14/010

GSBE

Maastricht University School of Business and Economics
Graduate School of Business and Economics

P.O. Box 616
NL- 6200 MD Maastricht
The Netherlands

Nonparametric welfare and demand analysis with unobserved individual heterogeneity

Sam Cosaert^{*}

Thomas Demuyck[†]

Abstract

In this paper, we combine elementary revealed preference principles and nonparametric estimation techniques in order to obtain nonparametric bounds on the distribution of the money metric utility over a population of heterogeneous households. The main benefit of our approach is that it is independent of any functional specification on the household utility functions, which means that our results are robust against parametric specification errors. We further demonstrate that our methodology can be used to establish bounds on the distribution of the demand function for counterfactual price regimes. In order to demonstrate the relevance of our approach, we illustrate our findings using a repeated cross-sectional household consumption data set.

JEL-codes: D12, C14

Keywords: unobserved heterogeneity, stochastic revealed preference, household demand

1 Introduction

This paper presents a framework to construct nonparametric bounds on the distribution of the money metric utility function while taking into account individual unobserved heterogeneity. Our approach combines elementary revealed preference concepts (in particular the Weak Axiom of Revealed Preference) with nonparametric (kernel) estimation techniques. In this manner, our approach remains independent of any parametric specification on the underlying household utility functions or on the unobserved heterogeneity distribution. We further demonstrate how the framework can be used to establish bounds on the distribution of the demand functions in counterfactual price regimes. An illustration using the Consumer Expenditure Survey, a US cross-sectional household consumption data set demonstrates the practical usefulness of our results.

^{*}K.U.Leuven, E.Sabbelaan 53, 8500 Kortrijk, Belgium, email: sam.cosaert@kuleuven.be

[†]Corresponding author, Maastricht University, Tongersestraat 53, 6711 LM Maastricht, Netherlands, email: t.demuyck@maastrichtuniversity.nl

Motivation Demand analysis provides a powerful tool to analyse behavioural responses and welfare effects due to price and income variations. In a typical demand study, the researcher first estimates the parameters of some parametric demand system,¹ and subsequently uses these estimates to calculate the associated indirect utilities. This ‘parametric’ approach has two major shortcomings. The first shortcoming results from the fact that the outcome is sensitive to the specific functional structure chosen by the researcher. Imposing the wrong functional form can therefore severely bias the resulting analysis. The second shortcoming is that it treats the problem of individual (unobserved) heterogeneity in an very add-hoc manner. In a typical consumer data set, we observe individuals or households only once. Given this data limitation, it is often assumed that similar looking individuals have similar preferences. Many demand studies therefore model a household’s demand to equal a rational systematic component, from a common utility function across all (similar looking) households, and a household specific additive error term capturing the unobserved heterogeneity or taste variation. By controlling for various observable characteristics (like household size), it is hoped that the issue of heterogeneity across the households is adequately addressed by including such additive error term. However, this assumption disregards the finding that individuals who look very similar may actually differ dramatically in their actual choice behaviour.² Moreover, as Lewbel (2001) showed, imposing additivity of the unobserved heterogeneity is a strong assumption whose implications come very close to enforcing a representative agent assumption.³ To summarize, we see that different people (although they may look the same) have different tastes and, consequentially, behave differently. In order to take this into account, it is crucial to allow for non-additive unobserved heterogeneity.

Literature overview In order to deal with aforementioned two problems, one can distinguish between two approaches. The first approach looks at the nonparametric differential ‘smooth’ restriction that can still be established in a heterogeneous population. These usually take the form of population level generalizations of Slutsky symmetry, negativity and homogeneity. Recent examples that follow this approach are Hoderlein (2011), Blundell, Horowitz, and Pairey (2013), Hausman and Newey (2013), and Hoderlein and Vanhems (2013). A second approach, which is the one followed in this paper, is to rely on revealed preference theory. Revealed preference theory was initiated by Samuelson (1938), Houthakker (1950) and further developed in several seminal contributions by Afriat (1967), Diewert (1973) and Varian (1982). The main aim of revealed preferences theory is to establish (combinatorial) restrictions on observed demand behaviour of a certain individual or household such that it is consistent with the classical model of utility maximization subject to a budget constraint. One of its main advantages is that it imposes no functional restrictions on the underlying utility function, except for some regularity conditions like local non-satiation.

Revealed preference theory, as it was initially developed has two main problems. First, from

¹Popular parametric demand systems are the Translog (Christensen, Jorgenson, and Lau, 1975), the Almost Ideal (Deaton and Muellbauer, 1980), or the Quadratic Almost Ideal (Banks, Blundell, and Lewbel, 1997) demand system.

²Unobserved heterogeneity is often seen as an main reason why demand estimations on cross sectional data typically have low r-squared values.

³See also Brown and Walker (1989) and McElroy (1987) for a discussion of other issues when taking into account unobserved heterogeneity in demand analysis.

an empirical point of view, the method does not really seem to provide very tight bounds. The main reason for this is that relative price variations usually tend to be quite small in comparison to income variation. This implies that budget hyperplanes often do not cross. We refer to Bronars (1987) and Varian (1982) for a discussion of this problem. The second problem is that revealed preference theory is not well suited to deal with unobserved individual heterogeneity. As a result, most of its applications remain confined to a few panel consumption data sets, where the same household or individual is observed over multiple periods.

The first problem has been the subject of several recent studies that apply revealed preference theory to repeated cross sectional data by combining insights from revealed preference theory with nonparametric estimation techniques (see Blundell (2005); Blundell, Browning, and Crawford (2003, 2007, 2008) and Blundell, Browning, Cherchye, Crawford, De Rock, and Vermeulen (2012)). The main contribution from this literature is that it shows how to use nonparametric Engel curve demand estimates as an input for revealed preference analysis. If we assume that households in the same time period and location face the same relative prices, then the nonparametric Engel curves estimate the mean (or average) expansion paths for each price regime. The availability of these expansion paths greatly improves the nonparametric bounds on various welfare related concepts and on the counterfactual demand estimates that can be obtained using revealed preference techniques.

A remaining drawback of this approach is the way it deals with the issue of unobserved heterogeneity. Given that the Engel curve estimates are obtained from a mean regression, the methodology is subject to Lewbel (2001)'s critique in the sense that imposing revealed preference restrictions on the mean Engel curve estimates comes very close to imposing a representative consumer assumption. Given this, the approach does not fully address the individual heterogeneity problem. Moreover, despite the fact that the procedure has the potential to produce tight bounds on the 'representative' money metric utility and demand functions, it does not give us any information concerning the distribution of these functions across the heterogeneous population.

A useful extension of revealed preference theory that explicitly takes into account individual heterogeneity is Stochastic Revealed Preference Theory, initiated by McFadden and Richter (1971) and Falmagne (1978).⁴ We refer to McFadden (2005) for an overview of the literature. Stochastic revealed preference takes as input the entire distribution of demand behaviour over a heterogeneous population of households for a finite number of budget sets.⁵ Therefore, it is well suited to deal with the issue of unobserved heterogeneity. The literature has put forward several rationality axioms (e.g. the Axiom of Revealed Stochastic Preference and the Weak Axiom of Stochastic Revealed Preference) that provide conditions on the distributions of choices such that a population of individuals are consistent with rational choice theory, which postulates that individuals are preference maximizers. Although the literature is mainly theoretical, several recent papers have started to develop statistical tests to verify whether the stochastic revealed preference axioms are satisfied in reality. Hoderlein and Stoye (2013) derive a statistical procedure to in-

⁴See also Block and Marschak (1959), McFadden (1975), Fishburn (1978), Cohen (1980), Barberá and Pattanaik (1986), Fishburn and Falmagne (1989), Cohen and Falmagne (1990), Fishburn (1992), Bandyopadhyay, Dasgupta, and Pattanaik (1999) for other contributions.

⁵A second interpretation of stochastic revealed preference theory is that the demand behaviour is generated by a single household with a random utility function.

fer bounds on the fraction of the population that violate the Weak Axiom of Stochastic Revealed Preference. Kitamura and Stoye (2013) derive a statistical test to verify whether a population of heterogeneous households satisfies the Axiom of Stochastic Revealed Preference for a finite collection of budget sets, thereby explicitly taking into account that preferences are transitive. Finally, Kawaguchi (2012) derives several procedures to test the validity of various axioms of revealed stochastic preference. Interestingly, these studies find little evidence that the stochastic revealed preference restrictions are violated. The main difference between these papers and ours is the focus. While above mentioned contributions mainly deal with testing whether the axioms imposed by the stochastic revealed preference literature hold, we are more interested in the restrictions that these axioms impose on the resulting distribution of the money metric utility and demand functions. In the terminology of Varian (1982); while above papers deal with the *testing* of a theory, we concentrate on the *recovery* of the underlying structure of the model.

Also closely related to our paper is the recent paper of Blundell, Kristensen, and Matzkin (2011). These authors focus on the issue of unobserved heterogeneity in a two goods setting. In particular, they tackle the problem of individual unobserved heterogeneity by using nonparametric quantile demand estimates in combination with standard revealed preference tests (i.e. SARP). In this respect, Stoye and Hoderlein (2013) recently showed that in a two goods setting, imposing the usual revealed preference axioms on the quantile demands is equivalent to imposing the Axiom of Stochastic Revealed Preference on the entire data set.⁶ Unfortunately, their methodology is not easily extendible to a setting with more than two goods, whereas our analysis requires no limitation on the number of goods under consideration.

Contribution The main contribution of this paper to derive nonparametric bounds on the money metric utility functions and the demand functions without imposing any functional structure on the household utility functions and under very weak requirements on the unobserved heterogeneity structure. As such, we avoid the problem that our results might be biased because of a wrong functional specification or because the households do not satisfy the ‘representative agent’ condition. We establish our results by combining elementary stochastic revealed preference theory and nonparametric estimation techniques. Our framework not only allows us to derive bounds on the mean of the money metric utility and demand functions, but on the entire distribution of these functions over the heterogeneous population. This provides important additional information concerning the distribution of welfare and demand over the population.

In order to obtain our results, we exploit the Weak Axiom of Revealed Preferences (WARP) applied to a population of heterogeneous households. Although this axiom is weaker than the revealed preference axioms that exploit transitivity (e.g. the Strong Axiom of Revealed Preference), we nevertheless show that it is powerful enough to establish narrow bounds. We demonstrate the usefulness of our results by applying it to the Consumer Expenditure survey, a US cross sectional consumption data set.

⁶The analysis is greatly simplified by the fact that in a two goods setting, the Weak Axiom of (Stochastic) Revealed Preference coincides with the Strong Axiom of (Stochastic) Revealed Preference, i.e. imposing transitivity implies no additional testable implications, see Rose (1958).

Outline In section two, we set out our framework and we present the necessary notation, concepts and definitions for the remaining part of the paper. Section 3 establishes the theoretical results that provide the bounds on the distribution of the money metric utility function and the demand functions. Section 4 contains our empirical application. We discuss estimation, statistical inference and we present several results. Section 5 concludes and points towards future research.

2 Notation and Definitions

In this section, we set out our basic framework and we introduce the notation and definitions that are needed in order to establish the results in the following sections.

2.1 Set up

We consider an economy with a large (infinite) number of different households. Each household, h , is endowed with a utility function which we denote by $u(\mathbf{q}^h; \mathbf{a}^h, \boldsymbol{\nu}^h)$. This utility function depends on a (column) vector of consumed goods $\mathbf{q}^h \in \mathbb{R}_+^n$, where n is the number of goods, a vector of observable household specific attributes \mathbf{a}^h , e.g. household composition, and a household specific vector of unobservable attributes $\boldsymbol{\nu}^h$, capturing unobserved preference heterogeneity. In order to decide how much to consume, the household maximizes its utility function subject to a household budget constraint,

$$\mathbf{q}(\mathbf{p}, x^h; \mathbf{a}^h, \boldsymbol{\nu}^h) = \arg \max_{\mathbf{q}} u(\mathbf{q}; \mathbf{a}^h, \boldsymbol{\nu}^h) \text{ s.t. } \mathbf{p}\mathbf{q} \leq x^h.$$

Here we denote by $\mathbf{p} \in \mathbb{R}_{++}^n$ a (row) vector of strict positive prices and by $x^h \in \mathbb{R}_{++}^n$ the total household disposable income. We assume that the solution of this optimization problem gives a system of n continuous demand function $\mathbf{q}^h = \mathbf{q}(\mathbf{p}, x^h; \mathbf{a}^h, \boldsymbol{\nu}^h)$ which depends on the vector of prices, the income and the household observable and unobservable attributes. We assume that the utility function is strict quasi-concave and twice continuously differentiable such that the demand functions are well defined and continuous. For a price vector \mathbf{p}_t and an expenditure level x , we denote by (\mathbf{p}_t, x) , the budget set consisting of all bundles \mathbf{q} such that $\mathbf{p}_t\mathbf{q} \leq x$.

We treat $\boldsymbol{\nu}$ as a random vector. Using $F(\boldsymbol{\nu}|\cdot)$ to denote the conditional distribution of the unobserved preference attributes over the population of households, we impose the following assumption.

Assumption 1. *For all income levels x and prices \mathbf{p} ,*

$$F(\boldsymbol{\nu}|\mathbf{p}, x, \mathbf{a}) = F(\boldsymbol{\nu}|\mathbf{a}).$$

Assumption 1 requires that the vector of unobserved attributes is independent of prices and income, conditional on all observable attributes. This ‘independence of budget sets’ condition is frequently used in the literature.⁷ If we interpret $\boldsymbol{\nu}$ as a vector of preference parameters, Assumption 1 encompasses the idea, common in consumer demand, that preferences do not vary with

⁷See for example, Lewbel (2001, equation 4), Hausman and Newey (2013, Assumption 1) and Blundell, Kristensen, and Matzkin (2011, Assumption A.1).

prices and income. For notational convenience, we omit from now (until section 4.2) the dependence on the observable attributes \mathbf{a} , taking into account that every expression is valid conditional on some particular value of this vector.

For the remaining part of the paper, it will be more useful to work with the indirect utility function $v(\mathbf{p}, x^h; \boldsymbol{\nu}^h)$ which gives the maximal utility that household h can obtain at prices \mathbf{p} and income x^h . The indirect utility function is defined from the direct utility function by,

$$v(\mathbf{p}, x^h; \boldsymbol{\nu}^h) = u(\mathbf{q}(\mathbf{p}, x^h; \boldsymbol{\nu}^h); \boldsymbol{\nu}^h).$$

The indirect utility function is strictly increasing in the level of disposable income x^h . If we invert the indirect utility function $v(\mathbf{p}, x; \boldsymbol{\nu}^h)$, with respect to x^h , we obtain the expenditure function $e(\mathbf{p}, u^h; \boldsymbol{\nu}^h)$ which gives the minimal outlay for household h to reach utility level u^h at prices \mathbf{p} . Finally, using the expenditure function, we can define the money metric utility function,

$$\mu(\mathbf{p}_v, \mathbf{p}_t, x^h; \boldsymbol{\nu}^h) \equiv e(\mathbf{p}_v, v(\mathbf{p}_t, x^h; \boldsymbol{\nu}^h); \boldsymbol{\nu}^h).$$

The money metric utility $\mu(\mathbf{p}_v, \mathbf{p}_t, x^h; \boldsymbol{\nu}^h)$ gives the minimal amount of expenditure that household h needs to make at prices \mathbf{p}_v to be equally well off as it would have been when facing prices \mathbf{p}_t and income x^h . The money metric utility lies at the basis of many cost of living indices. In particular, given two price vectors \mathbf{p}_t and \mathbf{p}_v and some reference budget (\mathbf{p}, x) , the Konüs cost of living index, describing the price increase from \mathbf{p}_t to \mathbf{p}_v , is defined as,

$$\frac{\mu(\mathbf{p}_v, \mathbf{p}, x; \boldsymbol{\nu})}{\mu(\mathbf{p}_t, \mathbf{p}, x; \boldsymbol{\nu})}$$

There are two natural choices for \mathbf{p} , namely \mathbf{p}_t and \mathbf{p}_v . Setting \mathbf{p} equal to the initial price \mathbf{p}_t gives the Laspreyes-Konüs cost of living index,

$$\frac{\mu(\mathbf{p}_v, \mathbf{p}_t, x; \boldsymbol{\nu})}{x}$$

Setting \mathbf{p} equal to the final price vector \mathbf{p}_v gives the Paasche-Konüs cost of living index,

$$\frac{x}{\mu(\mathbf{p}_t, \mathbf{p}_v, x; \boldsymbol{\nu})}$$

Both indices can be used to describe the increase in the cost necessary to maintain the same living standard over time. Their distributions can easily be constructed provided we know the distribution of the money metric utility function. The money metric utility also provides a cardinalisation of the utility function in the sense that for any reference price vector \mathbf{p} and for any two budgets (\mathbf{p}_t, x) and (\mathbf{p}_v, y) :

$$\mu(\mathbf{p}, \mathbf{p}_t, x; \boldsymbol{\nu}^h) \geq \mu(\mathbf{p}, \mathbf{p}_v, y; \boldsymbol{\nu}^h) \iff v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu}^h).$$

As such, the difference in the money metric can be used as a measure for the welfare change between two different budgets: if (\mathbf{p}_t, x) is the old budget and (\mathbf{p}_v, y) is the new one, then the welfare change can be measured by,

$$\mu(\mathbf{p}, \mathbf{p}_v, y; \boldsymbol{\nu}^h) - \mu(\mathbf{p}, \mathbf{p}_t, x; \boldsymbol{\nu}^h)$$

Again, there are two obvious choices for the base price vector \mathbf{p} , namely \mathbf{p}_t or \mathbf{p}_v . The first leads to the concept of the equivalent variation,

$$EV = \mu(\mathbf{p}_t, \mathbf{p}_v, y; \boldsymbol{\nu}^h) - x.$$

The second gives the compensating variation,

$$CV = y - \mu(\mathbf{p}_v, \mathbf{p}_t, x; \boldsymbol{\nu}^h).$$

Revealed preferences The analysis in the following sections depends on a very simple revealed preference idea. Fix a household h and consider two distinct budgets (\mathbf{p}_t, x) and (\mathbf{p}_v, y) . If the household is utility maximizing, then the following condition must hold,

$$\text{If } x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h) \text{ then } v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) > v(\mathbf{p}_v, y; \boldsymbol{\nu}^h). \quad (1)$$

The reasoning behind the condition is simple, if $x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$, then the consumed bundle $\mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$ at the budget (\mathbf{p}_v, y) was also feasible when $\mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu}^h)$ was chosen. Given that the household is utility maximizing and that the budget sets are distinct, it follows that $u(\mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu}^h); \boldsymbol{\nu}^h) > u(\mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h); \boldsymbol{\nu}^h)$, or equivalently, $v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) > v(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$.

It is easy to see that condition (1) implies the Weak Axiom of Revealed Preference (see Samuelson (1938)), which states that for any two distinct budgets (\mathbf{p}_t, x) and (\mathbf{p}_v, y) ,

$$\text{If } x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h) \text{ then } y < \mathbf{p}_v \mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu}^h).$$

3 Nonparametric bounds

In this section we show how to use basic revealed preference restrictions, in particular condition 1, together with information on the distribution of $\mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu})$ in order to establish bounds on the distribution of the money metric utility function and the mean demand functions. As a first partial result, we demonstrate the possibility to obtain bounds on the proportion of households in the economy that prefer a certain budget over another.

Observational assumptions We depart from the observational restrictions imposed by a repeated cross sectional household consumption dataset, where different households face the same prices in each cross section. This gives us a data structure that consists of a limited set of different price regimes, and for each price regime a large number of consumption bundles which are obtained from a random sample of households in the economy. We denote by $T = \{1, \dots, |T|\}$, the set of cross sections. The price vector corresponding to cross section $t \in T$ is denoted by \mathbf{p}_t .

Given that different households face distinct expenditure levels, it is possible to observe (or estimate) the distribution of the random consumption bundles $\mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu})$ for every cross sectional price vector \mathbf{p}_t , ($t \in T$) and for any level of expenditure x . Actually, none of our results will require us to estimate the distributions of $\mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu})$ but it will be easier to conceptualize things if we assume that these distributions are known. Estimation will be discussed in section 3. We assume that $\mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu})$ has a continuous density function which is strict positive on its domain.

Before we start, let us introduce one last piece of notation. Let $A(\boldsymbol{\nu})$ represent a collection of conditions involving the random vector $\boldsymbol{\nu}$. We use the notation $\Pr(A(\boldsymbol{\nu}))$ as a shorthand for the following probability,

$$\Pr(A(\boldsymbol{\nu})) = \int \mathbb{1}[A(\boldsymbol{\nu})]dF(\boldsymbol{\nu}),$$

where $\mathbb{1}[\cdot]$ is the binary indicator function which equals one if and only if the term between brackets is true. $\Pr[A(\boldsymbol{\nu})]$ gives us the fraction of the households for which the statement $A(\boldsymbol{\nu}^h)$ is true. Equivalently, it gives us the probability that $A(\boldsymbol{\nu}^h)$ holds if we draw at random a household h from the population.

Using this notation, we further require that there is sufficient variation of preferences and demand in the population such that for any two distinct budgets (\mathbf{p}_t, x) and (\mathbf{p}_v, y) ,

$$\begin{aligned} \Pr[x = \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu})] &= 0, \text{ and} \\ \Pr[v(\mathbf{p}_t, x; \boldsymbol{\nu}) = v(\mathbf{p}_v, y; \boldsymbol{\nu})] &= 0 \end{aligned}$$

This will allow us to freely interchange strict and weak inequalities.

3.1 Bounds on population preferences

In this section, we show how to establish bounds on the proportion of populations that prefer a certain budget (\mathbf{p}_t, x) over another budget (\mathbf{p}_v, y) . Given the notation introduced above, this proportion is given by,

$$\Pr[v(\mathbf{p}_t, x; \boldsymbol{\nu}) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu})].$$

The following shows how to obtain bounds on this proportion using only information on the distribution of $\mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu})$. Consider the fraction of households for which $x \geq \mathbf{p}_t \mathbf{q}^h(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$.

$$r_{t,v}(x, y) = \Pr[x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu})]$$

We claim that this number gives a lower bound on the fraction of households that prefer the budget (\mathbf{p}_t, x) over the budget (\mathbf{p}_v, y) .

Lemma 1.

$$r_{t,v}(x, y) \leq \Pr[v(\mathbf{p}_t, x; \boldsymbol{\nu}) > v(\mathbf{p}_v, y; \boldsymbol{\nu})].$$

Proof. Given that all households are rational, we know from condition 1 that for all values $\boldsymbol{\nu}^h$ of the random vector:

$$\text{If } x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h) \text{ then, } v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) > v(\mathbf{p}_v, y; \boldsymbol{\nu}^h).$$

This means that $\mathbb{1}[x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h)] \leq \mathbb{1}[v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) > v(\mathbf{p}_v, y; \boldsymbol{\nu}^h)]$. Integrating both sides over all values of the random vector $\boldsymbol{\nu}$ and using Assumption 1, we obtain,

$$\begin{aligned} \Pr[x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu})] &\leq \Pr[v(\mathbf{p}_t, x; \boldsymbol{\nu}) > v(\mathbf{p}_v, y; \boldsymbol{\nu})], \\ \iff r_{t,v}(x, y) &\leq \Pr[v(\mathbf{p}_t, x; \boldsymbol{\nu}) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu})]. \end{aligned}$$

□

Figure 1 illustrates the reasoning behind the Lemma in the two goods case. The figure gives two budget sets (\mathbf{p}_v, y) and (\mathbf{p}_t, x) . Given budget exhaustion, consumption within each budget set is distributed over the respective budget lines. All consumption bundles on the blue segment capture the consumption bundles of the households that satisfy $x > \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$. As such, the mass of households that consume on this line segment is equal to $r_{t,v}(x, y)$. What the Lemma says is that this fraction is smaller than the proportion of households for which $v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$. If this would not be the case, then there would be an individual with $v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) < v(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$ which consumes a bundle on the blue line segment when facing the budget (\mathbf{p}_v, y) . However, this is impossible because this consumption bundle is in the interior of the budget set (\mathbf{p}_t, x) , which means, using condition (1), that $v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) > v(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$.

[Figure 1 about here.]

Given above lemma, and the fact that,

$$\Pr [v(\mathbf{p}_t, x; \boldsymbol{\nu}) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu})] + \Pr [v(\mathbf{p}_v, y; \boldsymbol{\nu}) \geq v(\mathbf{p}_t, x; \boldsymbol{\nu})] = 1,$$

We immediately obtain the upperbound,

$$\Pr [v(\mathbf{p}_t, x; \boldsymbol{\nu}) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu})] \leq 1 - r_{v,t}(y, x).$$

For both lower and upper bounds to be valid, it should be the case that for all $t, v \in T$ and all expenditure levels x, y ,

$$r_{t,v}(x, y) + r_{v,t}(y, x) \leq 1.$$

This condition is equivalent to the Weak Axiom of Stochastic Revealed Preference applied to our setting (see Bandyopadhyay, Dasgupta, and Pattanaik (1999, 2002, 2004)). Hoderlein and Stoye (2013) and Kawaguchi (2012) recently developed (among other things) a statistical test that verifies (for two given budgets (\mathbf{p}_t, x) and (\mathbf{p}_v, y)) whether this condition is satisfied. If we go back to Figure 1, the condition states that the sum of the mass of households on the blue line segment and the mass of households on the red line segment must be smaller than 1. If this would not be the case, then there would be a household which is on the blue segment when the budget is (\mathbf{p}_v, y) and on the red segment when the budget is equal to (\mathbf{p}_t, x) . However, this implies that the household violates the Weak Axiom of Revealed Preference (i.e. $x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$ and $y \geq \mathbf{p}_v \mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu}^h)$).

The condition also shows that our bounds on $\Pr [v(\mathbf{p}_t, x; \boldsymbol{\nu}) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu})]$ will be tighter, the closer $r_{t,v}(x, y) + r_{v,t}(y, x)$ is to one. In particular, if the sum equals one, then $\Pr [v(\mathbf{p}_t, x; \boldsymbol{\nu}) > v(\mathbf{p}_v, y; \boldsymbol{\nu})]$ will be exactly identified.

There are two potential issues that may arise. First of all, it may happen that $r_{t,v}(x, y) + r_{v,t}(y, x)$ is larger than one, in which case the bounds cannot be simultaneously satisfied. Alternatively, it may happen that $r_{t,v}(x, y) + r_{v,t}(y, x)$ is considerably smaller than one, in which case the range may be too large to contain much useful information. Whether one of those problems arises is obviously an empirical matter. However, it may nevertheless be useful to discuss each of the issues a bit more in detail and to present some potential solutions.

Incompatible bounds A first problem arises if,

$$r_{t,v}(x, y) + r_{v,t}(y, x) > 1,$$

for some budgets (\mathbf{p}_t, x) and (\mathbf{p}_v, y) . In such case, we know that there is at least one household in the population that violates the Weak Axiom of Revealed Preference, hence, we should reject individual rationality of the households in the population. In order to remedy the situation, we see two possible solutions. A first solution is to allow a certain fraction of the population to violate the Weak Axiom of Revealed Preference, i.e. a certain subset of the population is considered to be irrational. Applying this solution would amount to subtracting a certain percentage, that equals the fraction of irrational households, from $r_{t,v}(x, y)$ and $r_{v,t}(y, x)$, thereby widening the range of possible values for $\Pr [v(\mathbf{p}_t, x; \boldsymbol{\nu}) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu})]$.

A second solution is to relax the rationality constraints for all households simultaneously. In revealed preference theory, such relaxation is usually conceived by using a ‘goodness-of-fit’ measure. The basic idea here is that households may not ‘exactly’ pass the revealed preference restriction but are still very close to passing them. The way to proceed is to consider an extended version of the basic revealed preference conditions that focuses on nearly optimizing behaviour rather than exact optimising behaviour. See also Varian (1990) for a general discussion on the usefulness of considering such nearly optimizing behaviour in empirical revealed preference analysis. Here, we consider one way in which this can be accomplished. We consider an adaptation of an early proposal of Afriat (1973) for revealed preference tests in a non-stochastic setting to our specific setting. In particular, we capture optimization error by a so-called Afriat index $e \in [0, 1]$. For a given value of e , the new rationality criterion adjusts condition (1) in the following way,

$$\text{If } e \cdot x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h) \text{ then } v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) > v(\mathbf{p}_v, y; \boldsymbol{\nu}^h).$$

When comparing the budget (\mathbf{p}_t, x) to another budget (\mathbf{p}_v, y) the Afriat index e reduces the expenditure level x towards $e \cdot x$. In other words, we now check whether behaviour is rational if allowing the household to waste as much as $(1 - e)$ of the income x by making irrational choices. In other words, we only require the household to prefer the bundle $\mathbf{q}(\mathbf{p}_t, x, \boldsymbol{\nu})$ over the bundle $\mathbf{q}(\mathbf{p}_v, y, \boldsymbol{\nu})$ if the latter is in the budget $(\mathbf{p}_t, e \cdot x)$. As such, wasting/irrational behaviour can be also be regarded as sub-optimizing behaviour, we thus verify whether behaviour is rational if we account for an optimization error equal to $(1 - e)$. Lowering the value of e will lead to a less strict test. Using this Afriat index, we can construct the following probabilities,

$$r_t^e(x, y) = \Pr [e \cdot x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu})].$$

The number $r_t^e(x, y)$ is increasing in e and $r_{t,v}^0(x, y) = 0$. Given this, there will always be a value of $e \in [0, 1]$ such that,

$$r_{t,v}^e(x, y) + r_{v,t}^e(y, x) \leq 1.$$

The analysis could then proceed by replacing $r_t(x, y)$ by the numbers $r_t^{e^*}(x, y)$ where e^* is either fixed a priori or coincides with the largest number for which this inequality holds.

Uninformative bounds A second problem arises if $r_{t,v}(x, y) + r_{v,t}(y, x)$ is considerably below 1. In such cases, the range of values for $\Pr[v(\mathbf{p}_t, x; \boldsymbol{\nu}) > v(\mathbf{p}_v, y; \boldsymbol{\nu})]$ will be too wide and, therefore, not very informative. An approach to tighten the bounds is to impose a stronger stochastic revealed preference condition. In the construction of $r_{t,v}(x, y)$ above, we only used information concerning the two budget sets (\mathbf{p}_t, x) and (\mathbf{p}_v, y) . In some cases, however, it is possible to include information on additional budget sets to obtain tighter bounds. One such tightening relies on the fact that for any three distinct numbers a, b and c it is always the case that,

$$\Pr(a > c) \geq \Pr(a > b) + \Pr(b > c) - 1.$$

The reasoning behind the inequality is simple. The probability that c is larger than b is given by $1 - \Pr(b > c)$. As such, $\Pr(a > b > c)$ is bounded from below by $\Pr(a \geq b) - (1 - \Pr(b > c))$. The conclusion then follows from the fact that $\Pr(a > c) \geq \Pr(a > b > c)$. Rewriting above condition shows that it is equivalent to the famous triangle inequality.

$$\Pr(b > c) < \Pr(b > a) + \Pr(a > c).$$

The triangle inequality has first been noted by Guilbaud (1953) and has been popularized by Marschak (1960). The inequality is one of the key condition in the literature on binary probability systems. This literature, which is closely related to the literature on stochastic revealed preference theory, tries to characterize all collections of binary probabilities over a finite set of alternatives that are induced by probability distributions over the family of linear orders (preference relations) on this set.

If we apply above condition to our setting and use the previously established bounds, we obtain that for all $t, v, w \in T$ and all incomes x, y, z ,

$$\begin{aligned} \Pr[v(\mathbf{p}_t, x; \boldsymbol{\nu}) > v(\mathbf{p}_v, y; \boldsymbol{\nu})] &\geq \Pr[v(\mathbf{p}_t, x; \boldsymbol{\nu}) > v(\mathbf{p}_w, z; \boldsymbol{\nu})] + \Pr[v(\mathbf{p}_w, z; \boldsymbol{\nu}) > v(\mathbf{p}_v, y; \boldsymbol{\nu})] - 1, \\ &\geq r_{t,w}(x, z) + r_{w,v}(z, y) - 1. \end{aligned}$$

In cases where $r_{t,v}(x, y)$ is lower than

$$\max_{w,z} \{r_{t,w}(x, z) + r_{w,v}(z, y) - 1\},$$

this improves the lower bound on $\Pr[v(\mathbf{p}_t, x; \boldsymbol{\nu}) > v(\mathbf{p}_v, y; \boldsymbol{\nu})]$. Of course this tightening of the bounds can be iterated until no further improvements are possible. If the range is still too wide, further tightening could still be obtained by using other, though more elaborate ‘binary probability system’ conditions. See, for example, Fishburn (1992) for an overview of the various kind of conditions that could be imposed. Although the triangle inequality potentially improves the bounds, we found that for our application, it does not give any significant improvements. The main reason is probably that, in our application, the bounds are already quite narrow. Given this, we abstain from implementing it in the empirical analysis.

3.2 Bounding the money metric utility

In this section, we show how to use above results to bound the distribution of the money metric utility function $\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu})$ for some price vectors \mathbf{p}_0 and \mathbf{p}_t corresponding to the prices of

two cross sections in the data set and for a particular level of income x_0 . Let us first focus on the upper bounds.

Upper bounds Let us fix a cross sectional price vector \mathbf{p}_0 and an income level x_0 . For any number $\pi \in (0, 1)$ and any cross section $t \in T$, let $h_t(\pi)$ solve the following condition,

$$\begin{aligned}\pi &= \Pr [h_t(\pi) \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_0, x_0; \boldsymbol{\nu})], \\ &= r_{t,0}(h_t(\pi), x_0)\end{aligned}$$

The value of $h_t(r)$ corresponds to the π th quantile of the random variable $\mathbf{p}_t \mathbf{q}(\mathbf{p}_0, x_0; \boldsymbol{\nu})$. From Lemma 1, we know that π is lower than the fraction of the households that prefer the budget $(\mathbf{p}_t, h_t(\pi))$ over the budget (\mathbf{p}_0, x_0) .

$$\begin{aligned}\pi &\leq \Pr [v(\mathbf{p}_t, h_t(\pi); \boldsymbol{\nu}) \geq v(\mathbf{p}_0, x_0; \boldsymbol{\nu})], \\ &= \Pr [h_t(\pi) \geq \mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu})].\end{aligned}$$

The second line is obtained by inverting the indirect utility function $v(\mathbf{p}_t, h_t(\pi); \boldsymbol{\nu})$ with respect to its second argument. This can be done by the fact that the indirect utility function is strictly increasing in income.

Let us denote by $m_t(\pi)$ the quantile function of $\mu(\mathbf{p}_t, \mathbf{p}_0, x_0)$, i.e. for all $\pi \in (0, 1)$

$$\Pr [\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu}) \leq m_t(\pi)] = \pi.$$

Then, using previously established result, we have that,

$$\begin{aligned}\Pr [\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu}) \leq m_t(\pi)] &= \pi \leq \Pr [\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu}) \leq h_t(\pi)], \\ \iff m_t(\pi) &\leq h_t(\pi).\end{aligned}$$

The last line uses the assumption that cumulative distribution function of $\mu(\mathbf{p}_t, \mathbf{p}_0, x; \boldsymbol{\nu})$ is strictly increasing on its support. This result shows that $h_t(\pi)$ is an upper bound on the π th quantile of the distribution of the money metric utility function. Using these upper bounds on the quantiles; we can also derive an upper bound on the mean value of the money metric utility. Let M be the mean of the function $\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu})$. We have that:

$$\begin{aligned}M &= \int_0^\infty \mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu}) dF(\mu(\mathbf{p}_t, \mathbf{p}_0, x_0, \boldsymbol{\nu})), \\ &= \int_0^1 m_t(\pi) d\pi \leq \int_0^1 h_t(\pi) d\pi.\end{aligned}$$

In practice, we compute the values of $h_t(\pi)$ for a finite grid of values $\pi_0, \pi_1, \dots, \pi_n$ with $\pi_0 = 0$ and $\pi_n = 1$.⁸ This allows us to approximate this upper bound by,

$$\int_0^1 h_t(\pi) d\pi \leq \sum_{i=1}^n (\pi_n - \pi_{n-1}) h_t(\pi).$$

The finer the grid, the better the approximation.

⁸The upperbound $h_t(1)$ can be set to the minimal income such that the budget hyperplane for $(\mathbf{p}_t, h_t(1))$ lies above the hyperplane for (\mathbf{p}_0, x_0) .

Lower bound We use a similar procedure to compute lower bounds on the quantiles. Let $\ell_t(\pi)$ solve the following equality,

$$\begin{aligned} 1 - \pi &= \Pr [x_0 > \mathbf{p}_0 \mathbf{q}(\mathbf{p}_t, \ell_t(\pi); \boldsymbol{\nu})], \\ &= r_{0,t}(x_0, \ell_t(\pi)) \end{aligned}$$

Then,

$$\begin{aligned} 1 - \pi &\leq \Pr [v(\mathbf{p}_0, x_0; \boldsymbol{\nu}) > v(\mathbf{p}_t, \ell_t(\pi); \boldsymbol{\nu})], \\ &= \Pr [\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu}) > \ell_t(\pi)], \\ &= 1 - \Pr [\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu}) \leq \ell_t(\pi)] \end{aligned}$$

As before, let $m_t(\pi)$ be the π th quantile of the distribution of the money metric utility $\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu})$. We have that,

$$\begin{aligned} \Pr [\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu}) \leq m_t(\pi)] &= \pi \geq \Pr [\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu}) \leq \ell_t(\pi)], \\ \iff m_t(\pi) &\geq \ell_t(\pi) \end{aligned}$$

This shows that $\ell_t(\pi)$ is a lower bound for the quantile $m_t(\pi)$. The mean M is then bounded from below by the quantity $\int_0^1 \ell_t(\pi) d\pi$ which can be approximated by $\sum_{n=0}^{N-1} \ell_t(\pi_n)(\pi_{n+1} - \pi_n)$.⁹

3.3 Bounds on demand correspondences

In this section, we show how to adapt above framework in order to establish bounds on the quantiles of the demand functions for unobserved, counterfactual, price regimes \mathbf{p}_0 and expenditure levels x_0 , i.e. \mathbf{p}_0 does not necessarily correspond to the observed price vector of a certain cross section.

Consider a function $f : \mathbb{R}_+^n \rightarrow \mathbb{R} : \mathbf{q} \mapsto f(\mathbf{q})$. In this section, we will provide upper bounds on the quantiles of the distribution of the random variable $f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu}))$. The function $f(\mathbf{q})$ can encompass various interesting measures. For example, if we want to bounds the expenditure share on one of the goods, we can use the function $f(\mathbf{q}) = \frac{1}{x_0} p_{0,j} q_j$, where $p_{0,j}$ is the price of good j in vector \mathbf{p}_0 , q_j is the quantity of good j in vector \mathbf{q} and x_0 is the expenditure level.

The focus on upper bounds is not restrictive given that we can always use information on upperbounds to construct lowerbounds. In order to see this, let $-m(1 - \pi)$ be the $(1 - \pi)$ th quantile of the variable $-f(\mathbf{q}(\mathbf{p}_0, x_0; \boldsymbol{\nu}))$ and let $-g(1 - \pi)$ be its upper bound. We then have

⁹The lowerbound $\ell_t(0)$ can be set to the maximal income such that the hyperplane for the budget set $(\mathbf{p}_t, \ell_t(0))$ lies below the hyperplane for (\mathbf{p}_0, x_0) .

that,

$$\begin{aligned}
1 - \pi &= \int \mathbb{1}[-f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu})) \leq -m(1 - \pi)] dF(\boldsymbol{\nu}), \\
&\leq \int \mathbb{1}[-f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu})) \leq -g(1 - \pi)] dF(\boldsymbol{\nu}), \\
\iff \pi &\geq 1 - \int \mathbb{1}[-f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu})) \leq -g(1 - \pi)] dF(\boldsymbol{\nu}), \\
&= \int \mathbb{1}[-f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu})) > -g(1 - \pi)] dF(\boldsymbol{\nu}), \\
&= \int \mathbb{1}[f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu})) \leq g(1 - \pi)] dF(\boldsymbol{\nu}).
\end{aligned}$$

As such, we see that $g(1 - \pi)$ provides a lowerbound on the π th quantile of $f(\mathbf{q}(\mathbf{p}_0, x_0; \boldsymbol{\nu}))$. For example, we can establish a lowerbound on the π th quantile of $f(\mathbf{q}) = \frac{1}{x_0} p_{0,j} q_j$ by constructing an upperbound on the $(1 - \pi)$ th quantile of $-\frac{1}{x_0} p_{0,j} q_j (= \sum_{i \neq j} \frac{1}{x_0} p_{0,i} q_i - 1)$.

For every cross section t , we previously defined the value $\ell_t(1 - \pi)$ that satisfied the following condition,

$$\begin{aligned}
\pi &= \Pr[x_0 > \mathbf{p}_0 \mathbf{q}(\mathbf{p}_t, \ell_t(1 - \pi); \boldsymbol{\nu})], \\
&= r_{0,t}(x_0, \ell_t(1 - \pi)).
\end{aligned}$$

The value of $\ell_t(\pi)$ can be obtained using information on x_0 , \mathbf{p}_0 and the distribution of $\mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu})$ alone, which we assumed to be known. For the next step, we use the Weak Axiom of Stochastic Revealed Preference, which requires that,

$$\begin{aligned}
r_{t,0}(\ell_t(1 - \pi), x_0) + r_{0,t}(x_0, \ell_t(1 - \pi)) &\leq 1, \\
\iff r_{0,t}(x_0, \ell_t(1 - \pi)) &\leq 1 - r_{t,0}(\ell_t(1 - \pi), x_0).
\end{aligned}$$

Let $m(\pi)$ be the π th quantile of the distribution function of the random variable $f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu}))$. We have that,

$$\begin{aligned}
\Pr[f(\mathbf{q}(\mathbf{p}_0, x_0; \boldsymbol{\nu})) \leq m(\pi)] &= \pi = r_{0,t}(x_0, \ell_t(1 - \pi)), \\
&\leq 1 - r_{t,0}(\ell_t(1 - \pi), x_0), \\
&= \Pr[\ell_t(1 - \pi) \leq \mathbf{p}_t \mathbf{q}(\mathbf{p}_0, y_0; \boldsymbol{\nu})] \\
&\leq \Pr\left[f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu})) \leq \max_{\mathbf{q}} f(\mathbf{q}) \text{ s.t } \ell_t(1 - \pi) \leq \mathbf{p}_t \mathbf{q} \text{ and } \mathbf{p}_0 \mathbf{q} = x_0\right]
\end{aligned}$$

The last inequality follows from the fact that whenever $\ell_t(1 - \pi) \leq \mathbf{p}_t \mathbf{q}(\mathbf{p}_0, y_0; \boldsymbol{\nu})$ holds, then $f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu})) \leq \max_{\mathbf{q}} f(\mathbf{q}) \text{ s.t } \ell_t(1 - \pi) \leq \mathbf{p}_t \mathbf{q} \text{ and } \mathbf{p}_0 \mathbf{q} = x_0$ must also hold. In order to see this, assume on the contrary that $f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu}))$ is larger than $f(\mathbf{q})$ for all vectors \mathbf{q} where $\mathbf{p}_0 \mathbf{q} = x_0$ and $\ell_t(1 - \pi) \leq \mathbf{p}_t \mathbf{q}$. Then, given that $\mathbf{p}_0 \mathbf{q}(\mathbf{p}_0, x_0; \boldsymbol{\nu}) = x_0$, it must be that $\ell_t(1 - \pi) > \mathbf{p}_t \mathbf{q}(\mathbf{p}_0, x_0; \boldsymbol{\nu})$, a contradiction.

Above result shows that,

$$m(\pi) \leq \max_{\mathbf{q}} f(\mathbf{q}) \text{ s.t } \ell_t(1 - \pi) \leq \mathbf{p}_t \mathbf{q} \text{ and } \mathbf{p}_0 \mathbf{q} = x_0,$$

for all cross sections t . In practice, we compute this right hand side for every cross section t and then take the lowest value across all cross sections as the upper bound. If f is a linear function, then the right hand side is a simple linear programming problem which can be solved efficiently.

The construction of the bounds in the simple two goods setting is illustrated in Figure 2. There are three budget lines corresponding to (\mathbf{p}_0, x_0) , $(\mathbf{p}_t, \ell_t(1 - \pi))$ and $(\mathbf{p}_v, \ell_v(\pi))$. The incomes $\ell_t(1 - \pi)$ and $\ell_v(\pi)$ are chosen such that the mass of households on the green line segment (where $x_0 > \mathbf{p}_0 \mathbf{q}(\mathbf{p}_t, \ell_t(1 - \pi), \boldsymbol{\nu})$) is equal to π and the mass of households on the red line segment (where $x_0 > \mathbf{p}_0 \mathbf{q}(\mathbf{p}_v, \ell_v(\pi), \boldsymbol{\nu})$) is equal to $(1 - \pi)$.

The quantity \bar{q}_2 is the maximum value of good 2 that corresponds to a bundle on the budget (\mathbf{p}_0, x_0) (where $\mathbf{p}_0 \mathbf{q} = x_0$) and $\ell_t(\pi) \leq \mathbf{p}_t \mathbf{q}$. From the result above, we know that this value gives an upperbound on the π th quantile of the distribution of $q_2(\mathbf{p}_0, x_0; \boldsymbol{\nu})$. Given that there are only two goods, this upperbound immediately gives a lowerbound on the $(1 - \pi)$ th quantile of $q_1(\mathbf{p}_0, x_0; \boldsymbol{\nu})$, given by \underline{q}_1 . Similarly, \bar{q}_1 gives an upperbound on the $(1 - \pi)$ th quantile of $q_1(\mathbf{p}_0, x_0; \boldsymbol{\nu})$, while \underline{q}_2 gives a lowerbound on the π th quantile of $q_2(\mathbf{p}_0, x_0; \boldsymbol{\nu})$. As such, the π th quantile of $q_2(\mathbf{p}_0, x_0; \boldsymbol{\nu})$ is bounded by the quantities \bar{q}_2 and \underline{q}_2 .

[Figure 2 about here.]

Using the bounds on the quantiles of the demand functions, we can compute bounds on the mean of the demand function by using a similar procedure as for the money metric utility function.

4 Application

In this section, we discuss the empirical implementation of the theoretical bounds that were established in the previous section. We first present our estimation procedure for the measures $r_{t,v}(x, y)$, $\ell_t(\pi)$ and $h_t(\pi)$. Next, we discuss a modification of the estimator with better finite sample properties and we show how we control for observed heterogeneity and endogeneity of the total expenditures. We also briefly discuss the issue of statistical inference on bounds. Finally, we present some empirical results.

4.1 Estimation procedure

The construction of the bounds in the previous section assumed that we know the distribution of the variables $\mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu})$ for every cross sectional price \mathbf{p}_t and every income level x . Given these distributions it is easy to obtain the quantities $r_{t,v}(x, y) = \Pr[x > \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu})]$, which form the main building blocks for our bounds. In practice, however, these probabilities need to be estimated. We propose a kernel estimator. For the estimation of the numbers $\ell_t(\pi)$ and $h_t(\pi)$ which are used as bounds for the money metric utility function and the construction for the bounds on the demand functions we propose a plug-in estimator.

Consider the v th cross section, $v \in T$. Assume that this cross section contains a sample of n observed household demand bundles $\{\mathbf{Q}_{i,v}\}_{i \leq n}$ where i corresponds to a particular observation. We denote by $\{X_{i,v}\}_{i \leq n}$ the corresponding expenditure levels ($X_{i,v} = \mathbf{p}_v \mathbf{Q}_{i,v}$). We assume that the sample $\{\mathbf{Q}_{i,v}\}_{i \leq n}$ is i.i.d drawn from the random vector \mathbf{q}_v . We denote by x_v the random variable $\mathbf{p}_v \mathbf{q}_v$. We denote by $\mu(\cdot)$ the distribution function of x_v .

Consider the value $r_{t,v}(x, y) = \Pr [x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu})]$. This value corresponds to the cdf of the random variable $\mathbf{p}_t \mathbf{q}_v$, conditional on the value $x_v = y$,

$$\begin{aligned} r_{t,v}(x, y) &= \int \mathbb{1} [x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu})] dF(\boldsymbol{\nu}), \\ &= \int \mathbb{1} [x \geq \mathbf{p}_t \mathbf{q}_v] dF(\mathbf{q}_v | x_v = y), \end{aligned}$$

where $F(\mathbf{q}_v | x_v = y)$ is the conditional cdf of \mathbf{q}_v given the level of expenditure $x_v = y$. This expression is equal to the conditional mean of the indicator function $\mathbb{1} [x \geq \mathbf{p}_t \mathbf{q}_v]$,

$$\int \mathbb{1} [x \geq \mathbf{p}_t \mathbf{q}_v] dF(\mathbf{q}_v | x_v = y) = \mathbb{E} \{ \mathbb{1} [x - \mathbf{p}_t \mathbf{q}_v \leq 0] | x_v = y \}.$$

We can express this conditional mean in the following way,

$$\mathbb{1} [x - \mathbf{p}_t \mathbf{Q}_{i,v} \geq 0] = g(\ln(X_{i,v})) + \varepsilon_{i,v},$$

where $\mathbb{E}(\varepsilon_{i,v} | x_v = y) = 0$ for all values of y . The quantity of interest is given by the value of $g(\ln(y))$. A straightforward Nadaraya-Watson Kernel estimator is given by,

$$\hat{r}_{t,v}(x, y) = \frac{\frac{1}{nh} \sum_{i=1}^n \mathbb{1} [x \geq \mathbf{p}_t \mathbf{Q}_{i,v}] k \left(\frac{\ln(X_{i,v}) - \ln(y)}{h} \right)}{\frac{1}{nh} \sum_{i=1}^n k \left(\frac{\ln(X_{i,v}) - \ln(y)}{h} \right)}.$$

where h is the bandwidth and $k(\cdot)$ is a symmetric Kernel function that satisfies $\int k(v) dv = 1$ and $\int v k(v) dv = 0$.¹⁰ We take the log of expenditure as metric.

If for $n \rightarrow \infty$, $h \rightarrow 0$ and $nh \rightarrow \infty$, then the estimator $\hat{r}_{t,v}(x, y)$ consistently estimates $r_{t,v}(x, y)$. If in addition (i) $\mu(y) > 0$, (ii) $r_{t,v}(x, y) \in (0, 1)$, (iii) the functions $\mu(\cdot)$ and $r_{t,v}(\cdot, \cdot)$ are sufficiently smooth¹¹ and (iv) $nh^7 \rightarrow 0$ then the estimator $\hat{r}_{t,v}(x, y)$ has the following asymptotic distribution (see, for example Li and Racine (2007)).

$$\sqrt{nh} \frac{1}{V(x, y)^{1/2}} [\hat{r}_{t,v}(x, y) - r_{t,v}(x, y) - B(x, y)] \rightarrow N(0, 1).$$

where $B(x, y) = h^2 \kappa_2 \left[\frac{1}{2} \frac{\partial^2 r_{t,v}(x, y)}{\partial y^2} + \frac{\partial r_{t,v}(x, y)}{\partial y} \frac{\partial \mu(y)}{\partial y} \frac{1}{\mu(y)} \right]$ is the asymptotic bias and $V(x, y) = r_{t,v}(x, y) [1 - r_{t,v}(x, y)] \kappa / \mu(y)$ is the asymptotic variance. Here $\kappa_2 = \int v^2 k(v) dv$ and $\kappa =$

¹⁰In practice, we use the Gaussian kernel.

¹¹The exact condition is that $\mu(y)$ and $r_{t,v}(x, y)$ have continuous second order derivatives with respect to y .

$\int k(v)^2 dv$. As usual with nonparametric kernel estimators, the bias, $B(y, x)$, does not disappear asymptotically when using the optimal bandwidth $h = O(n^{1/5})$. One possible solution is to undersmooth.

The estimates for $h_t(\pi)$ and $\ell_t(\pi)$ are computed as the solution to the following conditions,

$$\begin{aligned}\pi &= \hat{r}_{t,0}(\hat{h}_t(\pi), x_0), \\ 1 - \pi &= \hat{r}_{0,t}(x_0, \hat{\ell}_t(\pi)).\end{aligned}$$

This is done using standard binary search algorithms. In order for this algorithm to work, we assume that $\hat{r}_{0,t}(x_0, \hat{\ell}_t(\pi))$ is decreasing in $\hat{\ell}_t(\pi)$. This assumption is (asymptotically) valid if all goods are normal (i.e. all demand functions are increasing in income).¹²

The estimator $\hat{h}_t(\pi)$ is equivalent to a conditional quantile kernel estimator. This estimator is consistent for $h_t(\pi)$ as long as for $n \rightarrow \infty$, $h \rightarrow 0$ and $nh \rightarrow \infty$. If $r_{t,v}(h_t(\pi), x_0)$ is continuously differentiable in $h_t(\pi)$, then under the same conditions as for the estimator $\hat{r}_{t,v}(x, y)$, it can be shown that its asymptotic distribution is given by (see, for example Li and Racine (2007, section 6.3)),

$$\sqrt{nh} \frac{\left| \frac{r_{t,0}(h_t(\pi), x_0)}{\partial h_t(\pi)} \right|}{V(h_t(\pi), x_0)^{1/2}} \left(\hat{h}_t(\pi) - h_t(\pi) - \frac{B(h_t(\pi), x_0)}{\left| \frac{\partial r_{t,0}(h_t(\pi), x_0)}{\partial h_t(\pi)} \right|} \right) \rightarrow^d N(0, 1).$$

Using a proof similar to the one of Cai (2002, Theorem 2), we show in the appendix that if for $n \rightarrow \infty$, $h \rightarrow 0$ and $nh \rightarrow \infty$, the estimator $\hat{\ell}_t(\pi)$ is consistent for $\ell_t(\pi)$ and that under some additional conditions its asymptotic distribution is given by,¹³

$$\sqrt{nh} \frac{\left| \frac{r_{0,t}(x_0, \ell_t(\pi))}{\partial \ell_t(\pi)} \right|}{V(x_0, \ell_t(\pi))^{1/2}} \left(\hat{\ell}_t(\pi) - \ell_t(\pi) + \frac{B(x_0, \ell_t(\pi))}{\left| \frac{\partial r_{0,t}(x_0, \ell_t(\pi))}{\partial \ell_t(\pi)} \right|} \right) \rightarrow^d N(0, 1).$$

The estimates for the bounds on the demand functions are computed by substituting the estimated values $\hat{\ell}_t(\pi)$ for the values of $\ell_t(\pi)$ in the linear programming problems. As such, the resulting estimators are determined as the minimum over a finite set of values which are themselves the solution of a linear maximization problem that contain the estimates $\hat{\ell}_t(\pi)$ as a parameter. Although this estimator is consistent, by the continuous mapping theorem, we are unaware of any results that discusses the asymptotic distribution of this (or similarly constructed) estimator.

4.2 Adjustments

We make several adjustments to the estimator $\hat{r}_{t,v}(x, y)$ presented above.

¹²See also Blundell, Browning, and Crawford (2003) for a similar assumption.

¹³In particular, we require that $r_{t,v}(x_0, \ell_t(\pi))$ has strict negative partial derivative with respect to $\ell_t(\pi)$ and that $\sqrt{nh^3} \rightarrow \infty$.

Boundary problems The estimator $\hat{r}_{t,v}(x, y)$ has the undesirable property that it may give an estimates strictly between zero and one even when the budget sets (\mathbf{p}_t, x_t) and (\mathbf{p}_v, x_v) do not intersect. In order to see this, assume that $x > \mathbf{p}_t \mathbf{q}$ for all \mathbf{q} for which $\mathbf{p}_v \mathbf{q} = y$, i.e. the budget set defined by price income (\mathbf{p}_t, x) lies strictly above the budget defined by (\mathbf{p}_v, y) . In this setting, it may still be the case that the indicator function $\mathbb{1}[x \geq \mathbf{p}_t \mathbf{Q}_{i,v}]$ is zero for some observations i . From this, it follows that the Kernel estimator will also be strictly below one although the true value of $\pi_{t,v}(x, y)$ is clearly equal to one.

In order to avoid this boundary problem, we reformulate the probability to be estimated in the following way,

$$\begin{aligned} \int \mathbb{1}[x \geq \mathbf{p}_t \mathbf{q}_v] dF(\mathbf{q}_v | x_v = y) &= \int \mathbb{1}[x \mathbf{p}_v \mathbf{s}_v \geq \mathbf{p}_t y \mathbf{s}_v] dF(\mathbf{s}_v | x_v = y), \\ &= \int \mathbb{1}[(x \mathbf{p}_v - y \mathbf{p}_t) \mathbf{s}_v \geq 0] dF(\mathbf{s}_v | x_v = y), \end{aligned}$$

Where \mathbf{s}_v is the random vector of consumption shares, $\mathbf{s}_v = \mathbf{q}_v / x_v$, and we used the identity $\mathbf{p}_v \mathbf{s}_v = 1$ and the fact that, conditional on $x_v = y$, $\mathbf{q}_v = y \mathbf{s}_v$. If we denote the realizations of \mathbf{s}_v by $\mathbf{S}_{v,i} = \mathbf{Q}_{v,i} / X_{v,i}$, we can estimate this using the Nadaraya-Watson estimator,

$$\frac{\frac{1}{nh} \sum_{i=1}^n \mathbb{1}[(x \mathbf{p}_v - y \mathbf{p}_t) \mathbf{S}_{i,v} \geq 0] k\left(\frac{\ln(X_{i,v}) - \ln(y)}{h}\right)}{\frac{1}{nh} \sum_{i=1}^n k\left(\frac{\ln(X_{i,v}) - \ln(y)}{h}\right)}.$$

Although this estimator could still over or underestimate the true proportion slightly, the bias should be considerably less. Also, the estimator has the advantage that it is either zero or one if the two budget lines do not intersect. In order to see this, assume that the two budgets do not intersect, i.e. there is no bundle \mathbf{q} such that

$$\begin{aligned} \mathbf{p}_t \mathbf{q} &= x \text{ and,} \\ \mathbf{p}_v \mathbf{q} &= y \end{aligned}$$

Now, assume that the estimator is somewhere strictly between zero and one. This means that there exist observations i and j such that:

$$\begin{aligned} (x \mathbf{p}_v - y \mathbf{p}_t) \mathbf{s}_{i,v} &< 0 \text{ and,} \\ (x \mathbf{p}_v - y \mathbf{p}_t) \mathbf{s}_{j,v} &\geq 0 \end{aligned}$$

As both left hand sides are continuous functions of the shares vector vector, we can use the intermediate value function and show the existence of a vector \mathbf{s} such that:

$$\begin{aligned} (x \mathbf{p}_v - y \mathbf{p}_t) \mathbf{s} &= 0, \\ \iff \frac{\mathbf{p}_v}{y} \mathbf{s} &= \frac{\mathbf{p}_t}{x} \mathbf{s} \end{aligned}$$

If we define $\alpha = \frac{y}{\mathbf{p}_v \mathbf{s}} = \frac{x}{\mathbf{p}_t \mathbf{s}}$, and let $\tilde{\mathbf{q}} = \alpha \mathbf{s}$, we have that $\mathbf{p}_t \tilde{\mathbf{q}} = x$ and $\mathbf{p}_v \tilde{\mathbf{q}} = y$, a contradiction.

Semiparametric adjustment We also adjust the kernel estimator $\hat{r}_{t,v}(x, y)$ by including a semi-parametric specification. We have two reasons to do this. First of all, given the data limitations, we would like to allow our estimator to depend on the vector of unobserved covariates, \mathbf{a} , without fully conditioning on each of its values. Next, we need to take into account the fact that total expenditures are probably endogenous. We follow Blundell, Browning, and Crawford (2008), and consider the following semiparametric modification,

$$\mathbb{1}[(x\mathbf{p}_t - y\mathbf{p}_v)\mathbf{S}_{i,v} \geq 0] = g(\ln(X_{i,v}) - \phi(\mathbf{A}'_{i,v}\theta)) + \mathbf{A}'_{i,v}\gamma + \varepsilon_{i,v},$$

where $\mathbf{A}_{i,v}$ be the observed household composition of observation i in cross section v . The function $\phi(\mathbf{A}'_{i,v}\theta)$ can be interpreted as the log of a general equivalence scale for household i , and $\mathbf{A}'_{i,v}\gamma$, documents the way in which observable demographic differences across households impacts the left hand side. Similar to Blundell et al. (2008) we use an estimate of the general equivalence scale $\phi(\mathbf{A}'_{i,v}\theta)$ taken from the Organisation for Economic Co-operation and Development (OECD) scales.

In order to control for the endogeneity of the expenditure level x_v , Blundell, Browning, and Crawford (2008) suggest to use a control function approach based on the two step semiparametric estimator (this is a linearised version of the procedure set out by Newey, Powell, and Vella (1999)). In a first step, we obtain the residuals from a regression of the log of total expenditure on all exogenous variables in the model and on an excluded instrument. We take the log of (equivalent) labour income as an instrument. In the second step, we conduct a semiparametric regression of $\mathbb{1}[(x\mathbf{p}_t - y\mathbf{p}_v)\mathbf{S}_{i,v} \geq 0]$ on $g(\ln(X_{i,v}) - \phi(\mathbf{A}'_{i,v}\theta))$, $\mathbf{A}'_{i,v}\gamma$ and $\hat{v}_{i,v}$, where $\hat{v}_{i,v}$ are the residuals from this first stage regression.

4.3 Inference on bounds

The methodology outlined in section 3 provides nonparametric bounds on various parameters of interest (e.g. the quantiles of the money metric utility). In the previous section, we have also shown how these bounds can be consistently estimated. However, given that the bounds are based on finite sample estimates, we are confronted with the issue of statistical inference, in particular, the construction of confidence intervals. Given that our estimates only provide bounds, this problem fits in the literature that deals with the construction of confidence intervals for partially identified estimators. We refer to the several recent papers by Imbens and Manski (2004); Chernozhukov, Hong, and Tamer (2007); Stoye (2009); Chernozhukov, Lee, and Rosen (2013) and in particular to the recent paper of Hoderlein and Stoye (2013) who consider the problem of constructing confidence intervals in a setting which is similar to ours.

Given that the main contribution of this paper is not on statistical inference, we will only briefly discuss this issue and instead refer to above mentioned papers for more details on how to construct confidence intervals in our setting. As an example of how such construction could look like, consider the case of the quantile $m_t(\pi)$ which gives the π th quantile of the money metric utility $\mu(\mathbf{p}_t, p_0, x_0; \boldsymbol{\nu})$. Using the results above, we know that

$$m_t(\pi) \in \Theta_0 = [\ell_t(\pi); h_t(\pi)]$$

In practice, however, we only have estimates $\hat{\ell}_t(\pi)$ and $\hat{h}_t(\pi)$. There are two kinds of intervals that can be constructed. The first is a confidence interval for the interval Θ_0 , i.e. a set CI_α such that,

$$\lim_{nh \rightarrow \infty} \Pr(\Theta_0 \subseteq CI_\alpha) = \alpha.$$

A second kind of interval, CI_α^m , constructs an interval for the quantile $m_t(\pi)$ itself in the sense that,

$$\lim_{nh \rightarrow \infty} \Pr(m_t(\pi) \in CI_\alpha^m) = \alpha.$$

One important results from the literature (see, for example, Imbens and Manski (2004)) is that,

$$\lim_{nh \rightarrow \infty} \Pr(m_t(\pi) \in CI_\alpha) \geq \lim_{nh \rightarrow \infty} \Pr(m_t(\pi) \in CI_\alpha^m) = \alpha$$

As such, any confidence set for the interval, Θ_0 , is also a (conservative) confidence interval for the parameter $m_t(\pi)$. An interval CI_α can be constructed in the following way, provided that the estimates of the upper and lower bounds are asymptotically normally distributed with zero asymptotic bias.¹⁴ Let $[\hat{\ell}_t(\pi) - \frac{c_\alpha \hat{\sigma}_l}{\sqrt{nh}}, \hat{\ell}_t(\pi) + \frac{c_\alpha \hat{\sigma}_l}{\sqrt{nh}}]$ be an asymptotic $\alpha\%$ confidence interval for the lowerbound $\ell_t(\pi)$ and let $[\hat{h}_t(\pi) - \frac{c_\alpha \hat{\sigma}_h}{\sqrt{nh}}, \hat{h}_t(\pi) + \frac{c_\alpha \hat{\sigma}_h}{\sqrt{nh}}]$ be an asymptotic $\alpha\%$ confidence interval for the upperbound $h_t(\pi)$, where $\hat{\sigma}_l$ and $\hat{\sigma}_h$ are consistent estimates of the standard errors of the asymptotic distribution of the lower and upper bound and where c_α is chosen such that,

$$\Phi(c_\alpha) - \Phi(-c_\alpha) = \alpha,$$

where $\Phi(\cdot)$ is the standard normal probability distribution. Then, using a simple Bonferroni argument, we know that,

$$\left[\hat{\ell}_t(\pi) - \frac{c_\alpha \hat{\sigma}_l}{\sqrt{nh}}, \hat{h}_t(\pi) + \frac{c_\alpha \hat{\sigma}_h}{\sqrt{nh}} \right],$$

is a conservative asymptotic $\alpha\%$ confidence interval for Θ_0 . The standard deviations $\hat{\sigma}_l$ and $\hat{\sigma}_h$ could be estimated by using an asymptotic consistent plug-in estimator when the asymptotic distribution is analytically available, or from a bootstrap procedure, provided its asymptotical validity.

4.4 Data Description

We illustrate our approach by using a data sample from the Consumer Expenditure Survey (CEX) a repeated cross section. We use data on consumption decisions by US households from 1994 to 2007 (14 years). It is important to note that the consumer expenditures are derived from the diary survey (and not from the interview data). The diary data seem well-suited for (static) demand analysis. First of all, given that we focus on non-durable consumption, which is customary in static

¹⁴For kernel estimators, we could get the bias to converge to zero by undersmoothing.

demand analysis, information on the purchase of big, durable items is unnecessary. Second, for non-durable commodities, the diary survey invites respondents to indicate their consumption in a two-week period. Because this period is relatively short, respondents should be able to recall their expenditures. We follow Blundell et al. (2008) by focussing our attention to three broad expenditure categories, namely, food, other non-durables and services.¹⁵ As the diary survey reports expenditures on a two-week basis, we convert these to yearly equivalents. Converting two-week expenditures to yearly data poses an important problem of seasonality. Therefore, we deseasonalise using a dummy regression approach. Specifically, the expenditures on each category (reported for two weeks) are regressed on month dummies. Residuals from this regression (which can be interpreted as the variation in expenditures which can not be explained by seasonality or by months) are added to the mean expenditures for each category in order to construct deseasonalized expenditures. Observations with negative total expenditures are dropped. As mentioned above, we also take into account that variation in expenditures can be driven by the household composition, e.g. the number of adults or the number of kids living in the family. Therefore, we deflate total expenditures as well as total income by an OECD equivalence scale.

For the empirical analysis, we restrict attention to (i) households who have completed the two-week diary, (ii) households who are not living in student housing, (iii) households who are vehicle owners (to include fuel expenses), (iv) households where both members works at least 17 hours, (v) households where both members are not self-employed, (vi) households where the age of the reference person is at least 21 and finally we restrict attention to (vii) households that consist of a husband, a wife and possibly children. As a final step we also remove some outlier observations. In particular, we removed observations for which rescaled total expenditures are not within 3 standard deviations from the mean, rescaled total expenditures are among the 5 per cent lowest or 5 per cent highest expenditures, expenditure shares on each category are not within 3 standard deviations from the mean or expenditure shares are close to 0. On average, we are left with 2163 observations per cross-section with a minimum of 1775 observations in 1994 and a maximum of 2379 observations in 2007. Figure 3 plots the evolution of the mean consumption shares of the three goods over the considered periods. Price information is obtained from the Bureau of Labor Statistics.

[Figure 3 about here.]

4.5 Empirical results

In this section, we provide the results of several exercises. Due to limited space, we need to restrict our analysis to some particular base years and some reference income levels. Additional results are available from the authors upon request.

¹⁵See Appendix B for a list of the different goods used for the construction of the aggregates.

Bounds on the mean cost of living Let us first show how our bounds perform with respect to the computation of the mean of the Laspreyes-Konüs cost of living index,

$$\int \frac{\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu})}{x_0} dF(\boldsymbol{\nu}) = \frac{1}{x_0} \int \mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu}) dF(\boldsymbol{\nu}).$$

The Laspreyes-Konüs price index measures the income that one would need, relative to the income in period 0, in order to be as equally good of as in the initial period. We take 1994 as the reference year which means that \mathbf{p}_0 corresponds to the price vector in the year 1994. We choose x_0 as the (OECD equivalence scale deflated) median expenditure level in 1994. The bounds on the cost of living that we obtain using our procedure are given in the last column of Table 1. The table also reports values for various other prices indices like the Laspreyes (L), the Paasche (P) and the Tornqvist price index (T). We also provide information on three other nonparametric bounds. The first are the Lerner bounds which are obtained from the fact that:

$$\min_j \left\{ \frac{p_{t,j}}{p_{0,j}} \right\} \leq \frac{\mu(\mathbf{p}_t, \mathbf{p}_0, x_0)}{x_0} \leq \max_j \left\{ \frac{p_{t,j}}{p_{0,j}} \right\}.$$

The bounds by (Pollak, 1971) improve upon this by replacing the upper bound by the Laspreyes price index.

$$\min_j \left\{ \frac{p_{t,j}}{p_{0,j}} \right\} \leq \frac{\mu(\mathbf{p}_t, \mathbf{p}_0, x_0)}{x_0} \leq \frac{\mathbf{p}_t \mathbf{q}_0}{x_0}.$$

The second to last column gives the bounds that are obtained by using the procedure set out by Blundell, Browning, and Crawford (2003). This method first estimates nonparametric Engel curves and subsequently uses these estimates in combination with revealed preference restrictions to establish nonparametric bounds. We would like to emphasize that there is a clear conceptual difference between the bounds of Blundell, Browning, and Crawford (2003), and ours. Their procedure provides bounds on the cost of living that corresponds to some kind of ‘representative individual’ whose demand functions equal the mean demand functions over the population. Our bounds, on the other hand, correspond to bounds of the mean cost of living over all households within the population. Although it is reassuring to see that both procedures give very similar results, this does not have to be the case in general.

[Table 1 about here.]

Distribution of the price and quantity indices Let us now have a look at the bounds on the quantiles of this cost of living index over the population. Figure 4 provides bounds on the quantiles of the Laspreyes-Konüs cost of living index, for the 10th (blue), 30th (red), 50th (black), 70th (red) and 90th (blue) percentile. Upper and lower bounds on a particular quantiles are presented by the same colour. Again the base year is 1994 and the reference income is given by the median expenditure level in this year. In general, we see that the bounds on the quantiles are quite narrow. The width of the distribution for a particular year depends to a large extent on the difference in relative slopes between the base year (\mathbf{p}_0/x_0) and the evaluation year (\mathbf{p}_t/x_t). The closer the

relative prices, the narrower the difference between the largest and smallest cost of living for the particular year. The reason is that the distribution is naturally bounded between the minimum and maximum values of y/x_0 for which the budget hyperplanes corresponding to (\mathbf{p}_t, y) and (\mathbf{p}_0, x_0) do not intersect. We see that the distribution is the narrowest in the year 2000 and the widest in the year 2002 giving differences in cost of living up to more than 5 percentage points between the 10th and 90th percentile. One noticeable feature about the figure is that there seems to be a considerable amount of heterogeneity in the population although the width of the distribution remains more or less constant for the latter 5 years.

[Figure 4 about here.]

Distribution of the compensating variation Figure 5 shows the distribution of the compensating variation,

$$x_t - \mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu})$$

Here, x_0 is taken to be the median income in 2000 and x_t is the median income in cross section t . This compensating variation gives the difference between the median income in year t and the minimum income that would be necessary in order to obtain the welfare level at budget (\mathbf{p}_0, x_0) . Values above zero indicate a welfare gain for a household at the median income in year t compared to a household at the median income in year 1994. We see that all quantiles are below zero for the years 1994-1999 and 2001 and quantiles are above zero for the years 2005-2007. Once again, there seems to be quite a lot of heterogeneity present in the population. For many years, the range between the 10th and 90th percentile is around \$400 per year which is substantial.

[Figure 5 about here.]

Figure 6 gives another illustration of the kind of questions that can be answered giving the framework in this paper. The figure gives bounds on the average of the Lasprey-Konüs cost of living for different starting quantiles,

$$\int \frac{\mu(\mathbf{p}_t, \mathbf{p}_0, x_{0,i}, \boldsymbol{\nu})}{x_{0,i}} dF(\boldsymbol{\nu}).$$

Here, $x_{0,i}$ represents the income at the i th quantile of the income distribution in 1994, \mathbf{p}_0 are the prices in 1994 and \mathbf{p}_t is the price vector for 2007. The figure gives an idea of the average price increase (over the heterogeneous population) for households starting at different quantiles of the income distribution in 1994. On average one sees an increase over the quantiles, which means that (on average) the cost of living for households starting at the lower end of the income distribution in 1994 was lower than for household starting at the higher end of the income distribution. In other words, the households that started at the lower end of the income distribution had (on average) a lower increase in the cost of living. Also, notice that the upper bound for the lowest quantile is below the lower bound for the upper quantile. This shows that the average cost of living values are significantly different (although the numbers are not very distinct in absolute terms).

[Figure 6 about here.]

Bounds on the demand functions As a last exercise, let us have a look at the bounds on the demanded consumption shares for counterfactual price regimes. To keep focus, we restrict ourselves to the computation of bounds for the own price effect for the food aggregate. We take as reference year 2001 as this year had a price vector which was closest to the average price vector over all considered periods. We fix the expenditure level at the median level in 2001 and we let the price for food, $p_{2001,j}$ range from 0.94 times the price in 2001 until 1.06 times its price. We keep the prices of all other goods fixed at the 2001 level. Figure 7 presents the results for three quantiles, the 10th (in black) the median (in red) and the 90th (in blue). Again we see a lot of heterogeneity in the demand curves over the population although the price responses look very similar across the three quantiles.

[Figure 7 about here.]

5 Conclusions

In this paper, we used elementary revealed preference techniques together with nonparametric estimation techniques in order to bound the distribution of the money metric utility and the demand functions over a population of heterogeneous households. Our methodology has two attractive features. First of all, the results are entirely nonparametric which means that they are not dependent on any functional form imposed on the underlying utility functions. Second, we impose minimal conditions on the structure of the individual, unobserved, heterogeneity. When we apply our techniques to data from the US consumer expenditure survey, we find that our method generates narrow and informative bounds on the quantiles of the money metric utility function. Our results also demonstrate that individual heterogeneity creates considerable variation in welfare between households in the population (conditional for the same level of expenditure). We also demonstrate how our results can be used to obtain informative bounds on the distribution of the demand functions in counterfactual price-income situations.

There are several avenues for follow up research. First of all, we only briefly touched upon the highly relevant topic of statistical inference. However, given that our data is obtained from a random sample, measurement error and small sample bias influence our bounds, and statistical inference becomes highly relevant. Next, it would be interesting to see how our methodology extends to discrete choice settings. One way to incorporate discrete choices would be consider a setting where individuals make discrete choices in addition to continuous choices. In fact, many of the results from this paper readily extend to this setting. Alternatively one could imagine a setting where all choices are discrete (see Manski (2007) and Sher, il Kim, Fox, and Patrick (2011) for a theoretical account of stochastic revealed preferences recovery in such setting). It would be interesting to look how the methodology developed in this paper transfers to such discrete choice setting. Finally, it would also be interesting to see how other (more strict) stochastic revealed preference axioms that explicitly take into account transitivity may even further improve our bounds.

References

- Afriat, S. N., 1967. The construction of utility functions from expenditure data. *International Economic Review* 8, 67–77.
- Afriat, S. N., 1973. On a system of inequalities in demand analysis: An extension of the classical method. *International Economic Review* 14, 460–472.
- Bandyopadhyay, T., Dasgupta, I., Pattanaik, P. K., 1999. Stochastic revealed preference and the theory of demand. *Journal of Economic Theory* 84, 95–110.
- Bandyopadhyay, T., Dasgupta, I., Pattanaik, P. K., 2002. Demand aggregation and the weak axiom of stochastic revealed preference. *Journal of Economic Theory* 107, 483–489.
- Bandyopadhyay, T., Dasgupta, I., Pattanaik, P. K., 2004. A general revealed preference theorem for stochastic demand behavior. *Economic Theory* 23, 589–599.
- Banks, J., Blundell, R., Lewbel, A., 1997. Quadratic engel curves and consumer demand. *The Review of Economics and Statistics* 79, 527–539.
- Barberá, S., Pattanaik, P. K., 1986. Falmagne and the rationalizability of stochastic choices in terms of random orderings. *Econometrica* 54, 707–715.
- Block, H. D., Marschak, J., 1959. Random orderings and stochastic theories of response. Discussion Paper 66, Cowles Foundation.
- Blundell, R., 2005. Presidential address: How revealing is revealed preference? *Journal of the European Economic Association* 3, 211–235.
- Blundell, R., Browning, M., Cherchye, L., Crawford, I., De Rock, B., Vermeulen, F., 2012. Sharp for SARP: Nonparametric bounds on the behavioral and welfare effects of price changes. *American Economic Journal: Microeconomics* forthcoming (12.11).
- Blundell, R., Browning, M., Crawford, I., 2003. Nonparametric engel curves and revealed preference. *Econometrica* 71, 205–240.
- Blundell, R., Browning, M., Crawford, I., 2007. Improving revealed preference bounds on demand responses. *International Economic Review* 48, 1227–1244.
- Blundell, R., Browning, M., Crawford, I., 2008. Best nonparametric bounds on demand responses. *Econometrica* 76, 1227–1262.
- Blundell, R., Horowitz, J., Parey, M., 2013. Nonparametric estimation of a heterogeneous demand function under the Slutsky inequality restriction. Tech. rep., cemmap.
- Blundell, R., Kristensen, D., Matzkin, R., 2011. Bounding quantile demand functions using revealed preference inequalities. *Journal of Econometrics* forthcoming.

- Bronars, S. G., 1987. The power of nonparametric tests of preference maximization. *Econometrica* 55, 693–698.
- Brown, B. W., Walker, M. B., 1989. The random utility hypothesis and inference in demand systems. *Econometrica* 57, 815–829.
- Cai, Z., 2002. Regression quantiles for time series. *Econometric Theory* 18, 169–192.
- Chernozhukov, V., Hong, H., Tamer, E., 2007. Estimation and confidence regions for parameter sets in econometric models. *Econometrica* 75, 1243–1284.
- Chernozhukov, V., Lee, S., Rosen, A. M., 2013. Intersection bounds: Estimation and inference. *Econometrica* 81, 667–737.
- Christensen, L. R., Jorgenson, D. W., Lau, L. J., 1975. Transcendental logarithmic utility functions. *American Economic Review* 65, 367–383.
- Cohen, M., Falmagne, J.-C., 1990. Random utility representation of binary choice probabilities: A new class of necessary conditions. *Journal of Mathematical Psychology* 34, 88–94.
- Cohen, M. A., 1980. Random utility systems – the infinite case. *Journal of Mathematical Psychology* 22, 1–23.
- Deaton, A., Muellbauer, J., 1980. *Economics and Consumer Behavior*. Cambridge University Press.
- Diewert, W. E., 1973. Afriat and revealed preference theory. *The Review of Economic Studies* 40, 419–425.
- Falmagne, J. C., 1978. A representation theorem for finite random scale systems. *Journal of Mathematical Psychology* 18, 52–72.
- Fishburn, P. C., 1978. Choice probabilities and choice functions. *Journal of Mathematical Psychology* 18, 205–219.
- Fishburn, P. C., 1992. Induced binary probabilities and the linear ordering polytope: A status report. *Mathematical Social Sciences* 23, 67–80.
- Fishburn, P. C., Falmagne, J.-C., 1989. Binary choice probabilities and rankings. *Economics Letters* 31, 113–117.
- Guilbaud, G. T., 1953. Sur une difficulté de la théorie du risque. *Colloques Internationaux du Centre National de Recherche Scientifique, Econometrie* 40, 19–25.
- Hausman, J. A., Newey, W. K., 2013. Individual heterogeneity and average welfare. Tech. rep., cemmap.
- Hoderlein, S., 2011. How many consumers are rational. *Journal of Econometrics* 164, 294–309.

- Hoderlein, S., Stoye, J., 2013. Revealed preference in a heterogeneous population. *Review of Economics and Statistics* forthcoming.
- Hoderlein, S., Vanhems, A., 2013. Estimating the distribution of welfare effects using quantiles. Tech. rep.
- Houthakker, H. S., 1950. Revealed preference and the utility function. *Economica* 17, 159–174.
- Imbens, G. W., Manski, C. F., 2004. Confidence interval for partially identified parameters. *Econometrica* 72, 1845–1857.
- Kawaguchi, K., 2012. Nonparametric test of axioms of revealed stochastic preference. Tech. rep., LSE.
- Kitamura, Y., Stoye, J., 2013. Nonparametric analysis of random utility models. Tech. rep., Cemmap.
- Lewbel, A., 2001. Demand systems with and without errors. *American Economic Review* 91, 611–618.
- Li, Q., Racine, S., 2007. *Nonparametric Econometrics*. Princeton University Press, Princeton.
- Manski, C., 2007. Partial identification of counterfactual choice probabilities. *International Economic Review* 48, 1393–1410.
- Marschak, J., 1960. *Stanford Symposium on Mathematical Methods in the Social Sciences*. Stanford University Press, Ch. Binary Choice Constraints on Random Utility Indicators.
- McElroy, M. B., 1987. Additive general error models for production, cost, and derived demand or share systems. *Journal of Political Economy* 95, 737–757.
- McFadden, D., 1975. Tchebycheff bounds for the space of agent characteristics. *Journal of Mathematical Economics* 2, 225–242.
- McFadden, D., Richter, M. K., 1971. On the extension of a set function on a set of events to a probability on the generated boolean σ -algebra. Tech. rep., University of California, Berkeley.
- McFadden, D. L., 2005. Revealed stochastic preference: a synthesis. *Economic Theory* 26, 245–264.
- Newey, W. K., Powell, J. L., Vella, F., 1999. Nonparametric estimation of triangular simultaneous equations models. *Econometrica* 67, 565–603.
- Pollak, R. A., 1971. The theory of the cost of living index. Research Discussion Paper 11, Bureau of Labor Statistics, Washington DC.
- Rose, H., 1958. Consistency of preference: The two-commodity case. *Review of Economic Studies* 25, 124–125.

- Samuelson, P. A., 1938. A note on the pure theory of consumer's behavior. *Economica* 5, 61–71.
- Sher, I., il Kim, K., Fox, J. T., Patrick, 2011. Partial identification of heterogeneity in preference orderings over discrete choices. Tech. rep., NBER.
- Stoye, J., 2009. More on confidence intervals for partially identified parameters. *Econometrica* 77, 1299–1315.
- Stoye, J., Hoderlein, S., 2013. Testing stochastic rationality and predicting stochastic demand: The case of two goods. Tech. rep.
- Tucker, H., 1967. *A Graduate Course in Probability*. Academic Press, New York.
- Varian, H., 1982. The nonparametric approach to demand analysis. *Econometrica* 50, 945–974.
- Varian, H., 1990. Goodness-of-fit in optimizing models. *Journal of Econometrics* 46, 125–140.

A proof of consistency and asymptotic normality of $\hat{\ell}_t(\pi)$

In this appendix we demonstrate the consistency and asymptotic normality of $\hat{\ell}_t(\pi)$. The proofs are similar to Cai (2002). For ease of notation, we write $\hat{\ell}$ for $\hat{\ell}_t(\pi)$ and ℓ for $\ell_t(\pi)$. We first show that $\hat{\ell} \xrightarrow{P} \ell$.

Consistency Given that $\hat{r}_{t,v}(x, y) \xrightarrow{P} r_{t,v}(x, y)$ and both functions are monotone and bounded in y and bounded, it follows from Tucker (1967, Theorem 1) that,

$$\sup_y |\hat{r}_{t,v}(x, y) - r_{t,v}(x, y)| \xrightarrow{P} 0.$$

For any $\varepsilon > 0$, set,

$$\delta(\varepsilon) = \min\{\pi - r_{t,v}(x, \ell + \varepsilon), r_{t,v}(x, \ell - \varepsilon) - \pi\} > 0.$$

This uses the fact that $r_{t,v}(x, y)$ is strictly decreasing in y at (x, y) .

Lemma 2. *If $|\ell - \hat{\ell}| \geq \varepsilon$ then $|r_{t,v}(x, \hat{\ell}) - \hat{r}_{t,v}(x, \hat{\ell})| \geq \delta(\varepsilon)$.*

Proof. Assume first that $\ell \geq \hat{\ell}$. Then we have that $\hat{\ell} \leq \ell - \varepsilon$. This implies $r_{t,v}(x, \hat{\ell}) \geq r_{t,v}(x, \ell - \varepsilon)$. As such,

$$r_{t,v}(x, \hat{\ell}) - \hat{r}_{t,v}(x, \hat{\ell}) \geq r_{t,v}(x, \ell - \varepsilon) - \pi \geq \delta(\varepsilon).$$

This shows that $|r_{t,v}(x, \hat{\ell}) - \hat{r}_{t,v}(x, \hat{\ell})| \geq \delta(\varepsilon)$.

If $\ell < \hat{\ell}$ we have that $\hat{\ell} > \ell + \varepsilon$ and consequentially, $r_{t,v}(x, \hat{\ell}) \leq r_{t,v}(x, \ell + \varepsilon)$. As such,

$$\hat{r}_{t,v}(x, \hat{\ell}) - r_{t,v}(x, \hat{\ell}) \geq \pi - r_{t,v}(x, \ell + \varepsilon) \geq \delta(\varepsilon).$$

Again, we see that $|r_{t,v}(x, \hat{\ell}) - \hat{r}_{t,v}(x, \hat{\ell})| \geq \delta(\varepsilon)$. As such, we see that for every $\varepsilon > 0$,

$$\Pr\left(|\ell - \hat{\ell}| \geq \varepsilon\right) \leq \Pr\left(|\hat{r}_{t,v}(x, \hat{\ell}) - r_{t,v}(x, \hat{\ell})| \leq \delta(\varepsilon)\right).$$

As the right hand side goes to zero by consistency of $\hat{r}_{t,v}(x, y)$ for $n \rightarrow \infty$, $h \rightarrow 0$ and $nh \rightarrow \infty$, the left hand side also goes to zero. Given this, we see that $\hat{\ell} \xrightarrow{P} \ell$. \square

Asymptotic normality Let us now look at the asymptotic distribution of $\hat{\ell}_n$. We know that $\hat{r}_{t,v}(x, y)$ has the following limiting distribution,

$$\frac{\sqrt{nh}}{V(x, y)^{1/2}}(\hat{r}_{t,v}(x, y) - r_{t,v}(x, y) - B(x, y)) \rightarrow \mathcal{N}(0, 1),$$

Where $B(x, y) = O(h^2)$ is the asymptotic bias and $V(x, y)$ is the asymptotic variance. We first consider two lemma's.

Lemma 3. *If $\varepsilon_n/h = o(1)$, then,*

$$\hat{r}_{t,v}(x, y + \varepsilon_n) - \hat{r}_{t,v}(x, y) = \frac{\partial r_{t,v}(x, y)}{\partial y} \varepsilon_n + o_P(\varepsilon_n) + o_P(h^2) + o_P((nh)^{-1/2}).$$

Proof. The proof involves straightforward but long and tedious computation. As such, it is available upon request. \square

Lemma 4. *Fix a number $v < \infty$ and define $\varepsilon_n = \frac{vV(x, y)^{1/2} - B(x, y)}{\left|\frac{\partial r_{t,v}(x, y)}{\partial y}\right|}$. Then if $h \rightarrow 0$, $nh^3 \rightarrow \infty$ and*

$nh^5 = O(1)$, then,

1. $\varepsilon_n/h = o(1)$.
2. $\sqrt{nh} o(\varepsilon_n) = o(1)$.

Proof. Easy. \square

Now we are ready to demonstrate the following,

$$\sqrt{nh} \frac{\left|\frac{r_{t,v}(x, \ell)}{\partial \ell}\right|}{V(x, \ell)^{1/2}} \left(\hat{\ell}_n - \ell + \frac{B(x, \ell)}{\left|\frac{\partial r_{t,v}(x, \ell)}{\partial \ell}\right|} \right) \rightarrow^d N(0, 1).$$

Proof. Denote by $\Phi(v)$ is the normal cumulative distribution function, then

$$\begin{aligned}
& \Pr \left((nh)^{1/2} \frac{\left| \frac{\partial r_{t,v}(x,\ell)}{\partial \ell} \right|}{V(x,\ell)^{1/2}} \left(\hat{\ell}_n - \ell + \frac{B(x,\ell)}{\left| \frac{\partial r_{t,v}(x,\ell)}{\partial \ell} \right|} \right) \leq v \right), \\
&= \Pr \left(\hat{\ell}_n - \ell \leq \frac{\frac{vV(x,\ell)^{1/2}}{(nh)^{1/2}} - B(x,\ell)}{\left| \frac{\partial r_{t,v}(x,\ell)}{\partial \ell} \right|} \right), \\
&= \Pr \left(\hat{\ell}_n - \ell \leq \varepsilon_n \right), \\
&= \Pr \left(\hat{\ell}_n \leq \varepsilon_n + \ell \right), \\
&= \Pr \left(\hat{r}_{t,v}(x, \hat{\ell}_n) \geq \hat{r}_{t,v}(x, \ell + \varepsilon_n) \right), \\
&= \Pr \left(\hat{r}_{t,v}(x, \hat{\ell}_n) \geq \hat{r}_{t,v}(x, \ell) + \frac{\partial r_{t,v}(x, \ell)}{\partial \ell} \varepsilon_n + o_P(\varepsilon_n) + o_P(h^2) + o_P((nh)^{-1/2}) \right), \\
&= \Pr \left(\hat{r}_{t,v}(x, \hat{\ell}_n) - \hat{r}_{t,v}(x, \ell) \geq - \left| \frac{\partial r_{t,v}(x, \ell)}{\partial \ell} \right| \varepsilon_n + o_P(\varepsilon_n) + o_p(h^2) + o_p((nh)^{-1/2}) \right), \\
&\sim \Pr \left((nh)^{1/2} \frac{1}{V(x, y)^{1/2}} \left(\hat{r}_{t,v}(x, \ell) - \hat{r}_{t,v}(x, \hat{\ell}_n) - B(x, \ell) \right) \geq -v \right) \\
&= \Pr \left((nh)^{-1/2} \frac{1}{V(x, y)^{1/2}} \left(\hat{r}_{t,v}(x, \ell) - r_{t,v}(x, \ell) - B(x, \ell) \right) \geq -v \right) \xrightarrow{d} \Phi(v).
\end{aligned}$$

□

B Construction of aggregates

Food is an aggregate of cereals, bakery products, beef, pork, poultry, seafood, other meat, eggs, milk products, other dairy products, fresh fruit, fresh vegetables, processed fruit, processed vegetables, sweets, fat and oils, non-alcoholic beverages, prepared food, snacks and condiments.

Other non-durables contain expenditures on alcohol consumption, tobacco, clothes (for all household members), footwear, reading material, stationery, school supplies, cleaning products, garden supplies, household textile, non-durable housewares, medical products, personal care products, audio-visual equipment, recreational goods, pet goods and vehicle expenses.

Services include utilities, media bills, repair services, insurance, postal services, gasoline, vehicle expenses (services), public transportation, medical care services, personal care services, recreational services, home services, rental services, membership fees, school fees, other fees, pet services and care services.

Figure 1: Illustration of Lemma 1

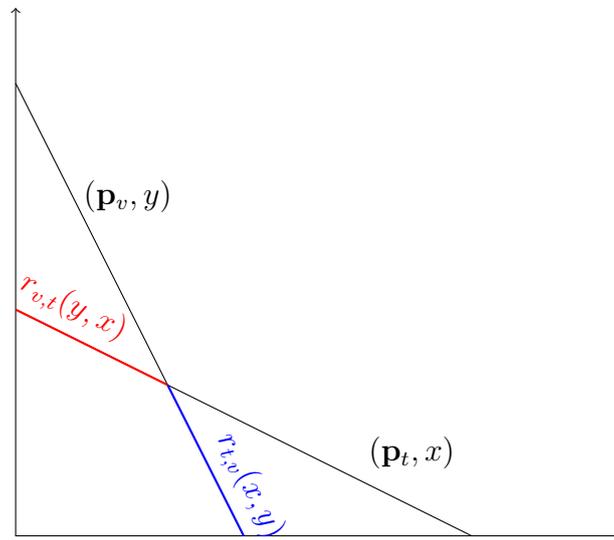


Figure 2: Illustration of the construction of the bounds

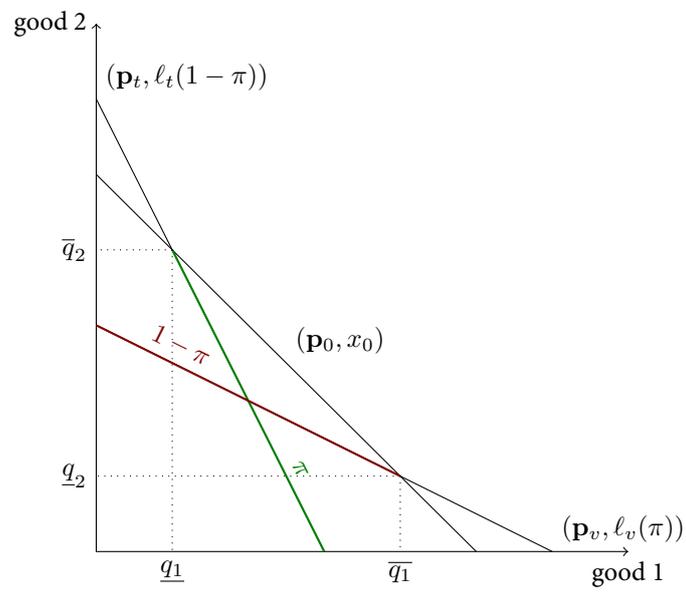


Figure 3: Evolution of average consumption shares

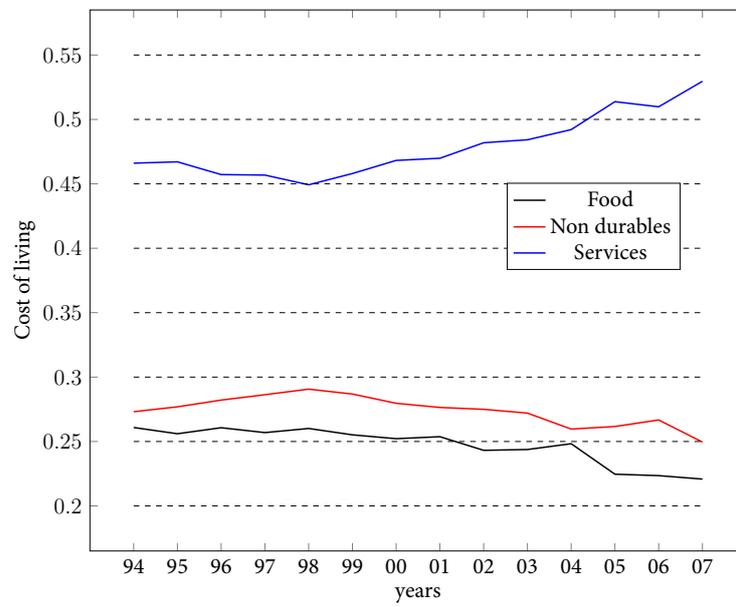


Figure 4: Distribution of the cost of living

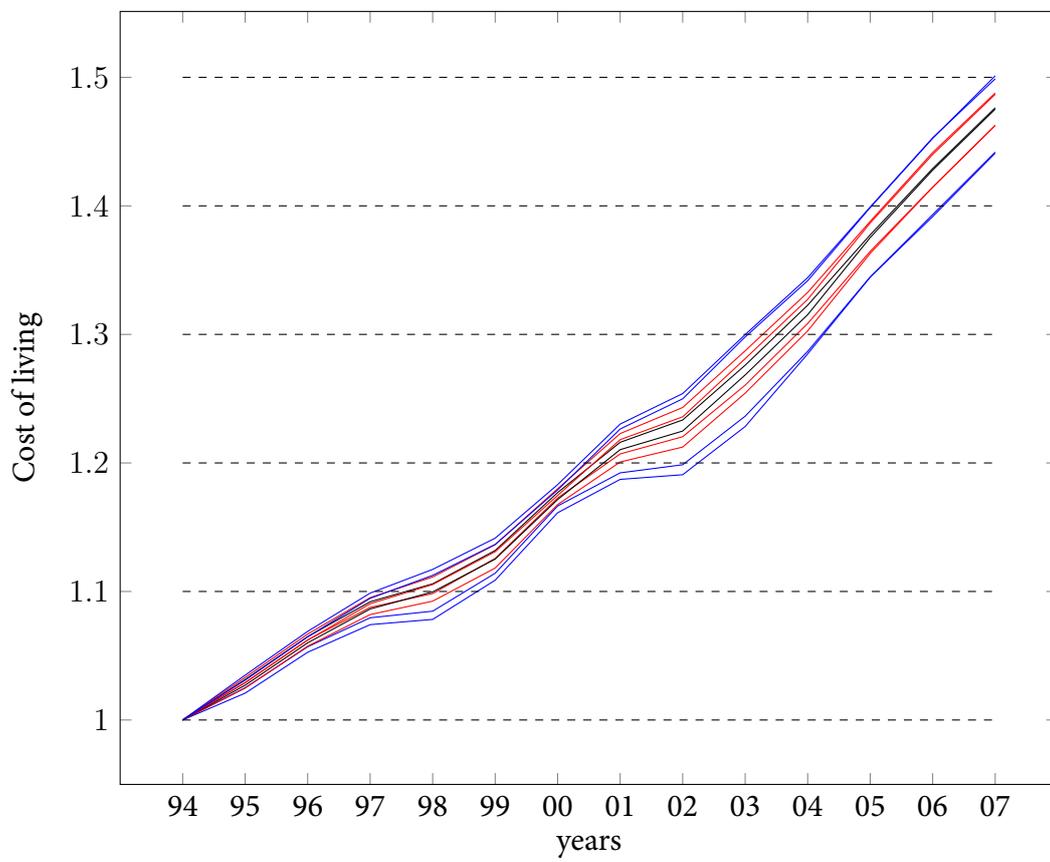


Figure 5: Distribution of compensating variation, baseyear 2000

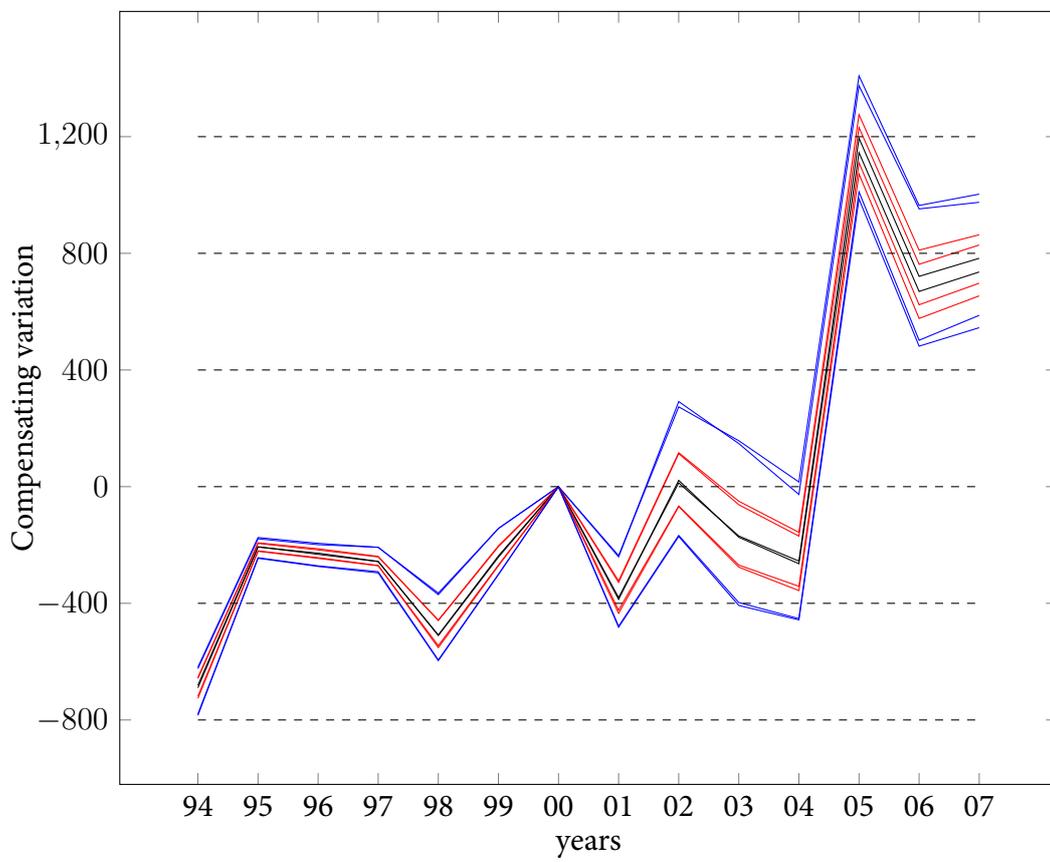


Figure 6: Change in mean cost of living 1994-2007 for different starting quantiles of income

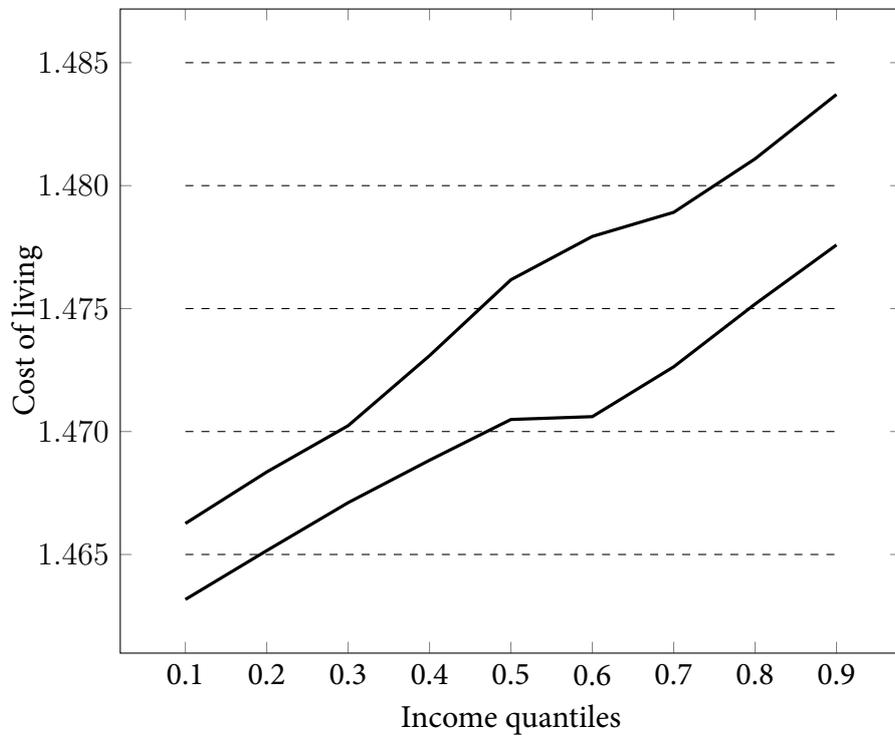


Figure 7: Bounds on the demand shares

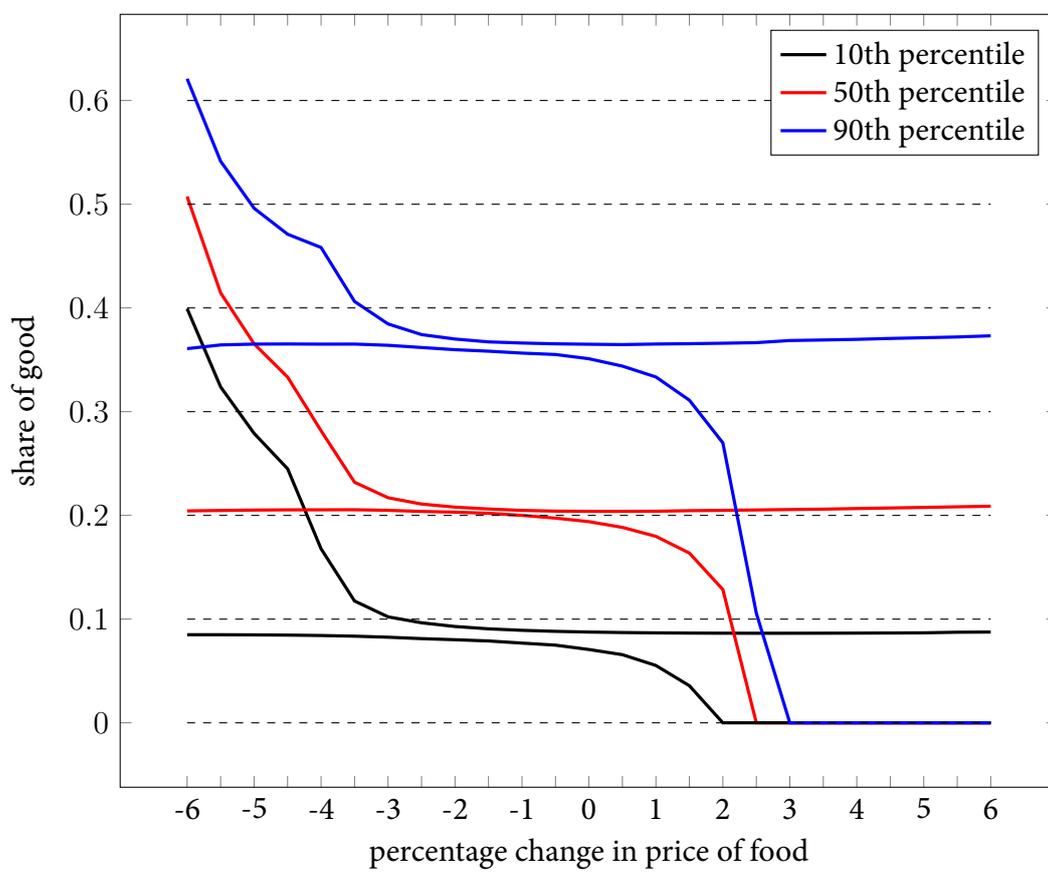


Table 1: Bounds on the mean Laspreyes Konüs cost of living index

year	Price indices			Lerner	Nonparametric Bounds		bounds
	L	P	T		Pollak	BBC	
1994	1.0000	1.0000	1.0000	[1.0000, 1.0000]	[1.0000, 1.0000]	[1.0000, 1.0000]	[1.0000 1.0000]
1995	1.0276	1.0271	1.0269	[1.0086, 1.0357]	[1.0086, 1.0276]	[1.0268, 1.0270]	[1.0257 1.0284]
1996	1.0606	1.0597	1.0594	[1.0358, 1.0708]	[1.0358, 1.0606]	[1.0592, 1.0595]	[1.0575 1.0614]
1997	1.0863	1.0846	1.0845	[1.1019, 1.0972]	[1.1019, 1.0863]	[1.0845, 1.0845]	[1.0820 1.0877]
1998	1.0977	1.0929	1.0937	[1.0327, 1.1236]	[1.0327, 1.0977]	[1.0946, 1.0929]	[1.0889 1.1000]
1999	1.1246	1.1213	1.1216	[1.0709, 1.1470]	[1.0709, 1.1246]	[1.1221, 1.1212]	[1.1173 1.1267]
2000	1.1717	1.1713	1.1713	[1.1886, 1.1652]	[1.1886, 1.1717]	[1.1714, 1.1711]	[1.1691 1.1739]
2001	1.2070	1.2048	1.2051	[1.1456, 1.2437]	[1.1456, 1.2070]	[1.2054, 1.2051]	[1.2003 1.2114]
2002	1.2212	1.2157	1.2171	[1.1301, 1.2742]	[1.1301, 1.2212]	[1.2186, 1.2169]	[1.2103 1.2274]
2003	1.2624	1.2568	1.2585	[1.1659, 1.3263]	[1.1659, 1.2624]	[1.2605, 1.2583]	[1.2528 1.2704]
2004	1.3100	1.3070	1.3076	[1.2243, 1.3679]	[1.2243, 1.3100]	[1.3084, 1.3078]	[1.3020 1.3172]
2005	1.3669	1.3699	1.3676	[1.3115, 1.4247]	[1.3115, 1.3669]	[1.3680, 1.3704]	[1.3659 1.3754]
2006	1.4184	1.4208	1.4193	[1.3400, 1.4839]	[1.3400, 1.4184]	[1.4215, 1.4233]	[1.4166 1.4287]
2007	1.4658	1.4688	1.4684	[1.3966, 1.5276]	[1.3966, 1.4658]	[1.4683, 1.4694]	[1.4649 1.4755]