IFN Working Paper No. 1324, 2020

Income Elasticities Without Parameters

Per Hjertstrand

Research Institute of Industrial Economics
P.O. Box 55665
SE-102 15 Stockholm, Sweden
info@ifn.se
www.ifn.se

Income Elasticities Without Parameters

Per Hjertstrand*

Research Institute of Industrial Economics (IFN)

Stockholm, Sweden

March 2020

Abstract

This paper proposes a simple non-parametric framework to calculate income elasticities from a data set of observed prices and consumed quantities without having to estimate any parameters. The framework can be applied when the price-quantity data satisfies a revealed preference axiom called the strong version of the strong axiom of revealed preference (SSARP). The framework is used to calculate income elasticities for food and non-alcoholic drinks from a rich panel of household expenditures. For this category, it is found that households are rather homogeneous in their demand responses.

1 Introduction

Knowing the magnitude of income elasticities across the income distribution has significant implications in several areas of economics. For example, in macroeconomics, policymakers that are responsible for maintaining price stability (e.g., central banks) require reliable quantitative estimates of money demand. If money demand is stable, the income elasticity equals the rate of money growth that is consistent with long-run price stability (See e.g., Mulligen and Sala-i-Martin, 1992). In empirical microeconomics, quantifying income elasticities for different categories of goods is important when making welfare comparisons within or between countries (Almås, 2012). More specifically, estimating income elasticities for food is important when analyzing patterns in cost of living indices since changes in the tariff structure and food import quotas has led to large increases in relative food prices in some countries (Blundell et al., 2008). As such, the income elasticity is often used as a key parameter when debating the level (or existence) of indirect tax rates (e.g., value added taxes).

This paper proposes a simple non-parametric framework to calculate income elasticities from a flexible demand system without having to estimate any parameters. The method is based on revealed preference theory, and consequently, assumes only knowledge of data on prices and consumed quantities of the goods.

The standard approach to calculate income elasticities in empirical consumption analysis is to postulate a parametric functional form for the indirect utility (or cost) function, and then estimate the corresponding demand functions using observed price and quantity data. Income elasticities are calculated directly from the estimated demand functions. However, this procedure is only satisfactory when

*I am grateful to Lanny Zrill for comments on an earlier draft. Financial support from Jan Wallander och Tom Hedelius stiftelse and Marianne och Marcus Wallenberg stiftelse is gratefully acknowledged. E-mail: Per.Hjertstrand@ifn.se.

the assumed parametric form is a good approximation of the correct one. Otherwise, elasticity estimates may be severely biased and even inconsistent. In contrast, the framework proposed in this paper makes no parametric assumptions for the utility function, and income elasticities can be calculated whenever the price-quantity data satisfies the strong version of the strong axiom of revealed preference (SSARP), an axiom only moderately more restrictive than the well-known strong axiom of revealed preference (SARP).

An important element of the proposed framework is that the income elasticities are calculated from a non-parametric utility function that, by construction, satisfies all theoretical regularity restrictions (i.e., continuity, strict monotonicity and concavity). Thus, this non-parametric utility function is always consistent with economic theory (provided that SSARP holds). In contrast, estimated parametric demand systems rarely satisfy theoretical regularity restrictions, and it is generally considered difficult to impose such conditions on parametric demand systems (Diewert and Wales, 1985; Barnett and Serletis, 2008).[1]

Another important element of the proposed framework is that, since no parameters are estimated, the method avoids any problems associated with degrees of freedom. Hence, the method can be meaningfully applied to datasets with any number of observations. This is especially advantageous when dealing with household panel data, where consumption expenditures for each household are typically observed over a few time periods, since the method can be applied to data from every individual household. Thus, because income elasticities are calculated separately for every household, our framework escapes any preference homogeneity assumptions across households and allows the researcher to optimally exploit the structure of such data. In contrast, this is impossible using parametric demand systems, since this approach requires more observations than free parameters (i.e., the demand system does not have enough degrees of freedom to be applicable for each individual household). Hence, estimating income elasticities in panel household data using parametric demand systems require some form of pooling, and consequently, rest on preference homogeneity assumptions across households (e.g., based on observable demographic factors).

Directly related to the previously described advantage of avoiding problems with degrees of freedom, our method can also be applied to data with any number of goods, and therefore to very disaggregated data. This is not possible for parametric demand systems, since the number of independent variables in the system grows polynomially in the number of goods. Omission of relevant prices (i.e., independent variables) creates an omitted-variables problem, which yields biased estimates. Thus, using the parametric approach require some kind of aggregation scheme, often in the form of separability assumptions (See e.g., Deaton and Muellbauer, 1980). In contrast, since the proposed non-parametric framework is applicable to any number of goods, no separability assumptions are needed.

The proposed framework is based on a revealed preference characterization of differentiable utility maximization developed by Chiappori and Rochet (1987). Specifically, they show that a data set of prices and quantities can be rationalized by a continuous, strictly increasing, strongly concave and infinitely differentiable utility function if and only if the data satisfies SSARP. This utility function is very flexible since it contains 2T parameters (which can be interpreted as utility and marginal utility values of the utility function), where T denotes the number of observations in the data.[2] In this paper, I simply propose

[1] In particular, parametric demand systems generally lose their flexibility when theoretical regularity conditions are imposed globally or even at every observation. Moreover, Barnett (2002) note that theoretical regularity is often equated with concavity and argues that imposing monotonicity on the utility function is equally as important, which adds to the complexity of estimating demand systems that are restricted to be consistent with all aspects of economic theory.

[2] Chiappori and Rochet's (1987) rationalizing utility function is based on Afriat's (1967) utility construction for the generalized axiom of revealed preference (GARP) which also requires 2T parameters.

to calculate income elasticities from Chiappori and Rochet's rationalizing utility function. Remarkably, although the utility function consists of 2T parameters, all these cancel out in the expression for the income elasticity, and I show that this elasticity is a simple function of the observed price-quantity data. Thus, there is no need to solve for any of the 2T parameters in Chiappori and Rochet's rationalizing utility function in order to calculate the income elasticities.[3]

Being based on a revealed preference theory, the proposed framework is deterministic in the sense that it does not contain any stochastic element. This implies that the framework neither allows any deviation from perfect utility maximizing behavior nor that the data is measured with errors, resulting in violations of revealed preference (SSARP). However, observed price-quantity data is rarely perfectly consistent with revealed preference restrictions. To deal with this, I propose two methods that make the framework applicable in situations when revealed preference is violated. In the first method, I suggest calculating the income elasticities from the largest subset of the data satisfying revealed preference. Appendix A outlines an efficient procedure to calculate the largest subset based on solving a mixed-integer linear programming problem. In the second method, I suggest to calculate the minimal adjustment of expenditure such that the data satisfies revealed preference. The income elasticities are then calculated taking the adjustment into account.

I use the framework to calculate income elasticities for food and non-alcoholic drinks at home in a rich panel of Spanish household consumption expenditures. The expenditures for each household are recorded in five to eight consecutive quarters, making the proposed methods ideal for analyzing heterogeneity in demand responses between households, since income elasticities are calculated for every individual household. The average income elasticity across all households is approximately 0.2. Looking at the entire distribution reveals that 75% of all households have an elasticity greater than 0.14 but only 25% of them have an elasticity larger than 0.21. Thus, based on these results, I find that households are rather homogeneous in terms of their demand responses for the category food and non-alcoholic drinks at home.

This paper is organized as follows. The next section provides the required revealed preference theory to derive the income elasticity. Section 3 gives the expression for the income elasticity. Section 4 outlines the two methods to deal with data violating revealed preference. Section 5 contains the application and Section 6 concludes.

2 Revealed preference

Suppose a consumer chooses from K goods and assets, which are indexed by $\mathbb{K} = \{1, ..., K\}$. The goods and assets are observed in a *finite* number of time periods, which are indexed by $\mathbb{T} = \{1, ..., T\}$. Let $\mathbf{x}_t = (x_{1t}, ..., x_{Kt}) \in \mathbb{R}^K_+$ denote the observed quantity-vector at time $t \in \mathbb{T}$ with corresponding price-

[3] The income elasticities proposed here are point elasticities (i.e., calculated at a specific point on the demand curve). Chavas and Cox (1997) propose calculating arc elasticities (i.e., over a range of the demand function) from two representations of preferences that bound the family of utility functions that rationalizes the data under the generalized axiom of revealed preference (GARP). Their approach requires sizeable changes in prices or income, since both representations of preferences under GARP are not everywhere differentiable (they consider demand responses from 20% changes in total expenditure). Blundell et al. (2008) suggest a method to obtain non-parametric bounds on predicted consumer responses to price changes, also in the form of arc elasticities. They consider a dataset that satisfies revealed preference and estimate potential responses to a new combination of prices and incomes that would not violate revealed preference given the observed data.

vector $\mathbf{p}_t = (p_{1t}, ..., p_{Kt}) \in \mathbb{R}^K_{++}$.[4] Throughout the paper, I refer to the list $\mathbb{D} = \{\mathbf{p}_t, \mathbf{x}_t\}_{t\in\mathbb{T}}$ as 'the data'.

Consider the following definitions of the concept of revealed preferences (Varian, 1982). We say that $\mathbf{x}_t$ is directly revealed preferred to $\mathbf{x}_s$ written $\mathbf{x}_t R^D \mathbf{x}_s$ if $\mathbf{p}_t\mathbf{x}_t \geq \mathbf{p}_t\mathbf{x}_s$. This means that the bundle $\mathbf{x}_t$ is directly revealed preferred over the bundle $\mathbf{x}_s$ if $\mathbf{x}_t$ was chosen when $\mathbf{x}_s$ also was affordable at prices $\mathbf{p}_t$. $\mathbf{x}_t$ is revealed preferred to $\mathbf{x}_s$ written $\mathbf{x}_t R \mathbf{x}_s$ if there exists a sequence of observations $(t, u, v, ..., w, s) \in \mathbb{T}$ such that $\mathbf{x}_t R^D \mathbf{x}_u$, $\mathbf{x}_u R^D \mathbf{x}_v$, ..., $\mathbf{x}_w R^D \mathbf{x}_s$. Thus, the relation R exploits transitivity of preferences.

Definition 1 *Consider a data set $\mathbb{D} = \{\mathbf{p}_t, \mathbf{x}_t\}_{t\in\mathbb{T}}$:*

- *$\mathbb{D}$ satisfies the Strong Axiom of Revealed Preference (SARP) if $\mathbf{x}_t R \mathbf{x}_s$ and $\mathbf{x}_t \neq \mathbf{x}_s$ implies $\mathbf{p}_s\mathbf{x}_s < \mathbf{p}_s\mathbf{x}_t$.*
- *$\mathbb{D}$ satisfies the Strong version of the Strong Axiom of Revealed Preference (SSARP) if $\mathbb{D}$ satisfies SARP and if $\mathbf{p}_t \neq \mathbf{p}_s$ implies $\mathbf{x}_t \neq \mathbf{x}_s$ for all $s, t \in \mathbb{T}$.*

SARP is well-known and states that it cannot be that the bundle $\mathbf{x}_t$ is preferred over the distinct bundle $\mathbf{x}_s$, when at the same time the cost of $\mathbf{x}_s$ is strictly less than the cost of $\mathbf{x}_t$ at prices $\mathbf{p}_s$. Chiappori and Rochet (1987) introduced SSARP, which says that the data, in addition to satisfying SARP, cannot contain situations when the same quantity bundle is purchased at two different price vectors.

Definition 2 *Consider a data set $\mathbb{D} = \{\mathbf{p}_t, \mathbf{x}_t\}_{t\in\mathbb{T}}$ and a utility function $u : \mathbb{R}^K_+ \longmapsto \mathbb{R}$. For any $\mathbf{x} \in \mathbb{R}^K_+$ and all $t \in \mathbb{T}$ such that $\mathbf{p}_t\mathbf{x}_t \geq \mathbf{p}_t\mathbf{x}$,*

- *the data $\mathbb{D}$ is strictly rationalized by u if $u(\mathbf{x}_t) > u(\mathbf{x})$ whenever $\mathbf{x}_t \neq \mathbf{x}$,*
- *the data $\mathbb{D}$ is strongly rationalized by u if $u(\mathbf{x}_t) > u(\mathbf{x})$.*

Matzkin and Richter (1991) provided a revealed characterization of SARP by showing that there exists a continuous, strictly increasing and strictly concave utility function that strictly rationalizes $\mathbb{D}$ if and only if $\mathbb{D}$ satisfies SARP. However, SARP does not ensure that there exists a differentiable utility function that (strictly or strongly) rationalizes the data (See e.g., Figure 1 in Chiappori and Rochet, 1987). In contrast, empirical demand analysis is primarily concerned with calculation of point elasticities and welfare comparisons, and consequently, is usually based on differentiable (indirect) utility (or cost) functions. As a reaction to this, Chiappori and Rochet (1987) showed that adding to SARP the condition that the same bundle cannot be purchased at two distinct price vectors ensures that there also exists an *infinitely differentiable* utility function in the set of rationalizing utility functions.

Theorem 1 (Chiappori and Rochet, 1987) *Consider a data set $\mathbb{D} = \{\mathbf{p}_t, \mathbf{x}_t\}_{t\in\mathbb{T}}$. The following conditions are equivalent:*

- *the data $\mathbb{D}$ can be strongly rationalized by a strictly increasing, strongly concave and infinitely differentiable utility function $u(\mathbf{x})$, whose Hessian matrix satisfies:*

$$\nabla^2_{\mathbf{xx}} u(\mathbf{x}_t) = -\varepsilon I^K, \tag{1}$$

for all $t \in \mathbb{T}$, where $\varepsilon > 0$ is a (sufficiently) small scalar, and I^K is the $K \times K$ identity matrix.

[4] The following notation is used: The inner product of two vectors $(x, y) \in \mathbb{R}^K$ is defined as $xy = \sum_{l=1}^K x_l y_l$. For all $(x, y) \in \mathbb{R}^K$, $x \geqq y$ if $x_i \geq y_i$ for all $i = 1, \dots, K$; $x \geq y$ if $x \geqq y$ and $x \neq y$; and $x > y$ if $x_i > y_i$ for all $i = 1, \dots, K$. We denote $\mathbb{R}^L_+ = \{x \in \mathbb{R}^L : x \geqq (0, \dots, 0)\}$ and $\mathbb{R}^L_{++} = \{x \in \mathbb{R}^L : x > (0, \dots, 0)\}$. The gradient vector of a differentiable function $f(\mathbf{x})$ is denoted $\nabla_{\mathbf{x}} f(\mathbf{x})$. The Hessian matrix (i.e., matrix of second-order partial derivatives) of a twice-differentiable function $f(\mathbf{x})$ is denoted $\nabla^2_{\mathbf{xx}} f(\mathbf{x})$.

- *the data $\mathbb{D}$ satisfies SSARP.*

This result states that the data $\mathbb{D}$ satisfies SSARP *if and only if* there exists an infinitely differentiable and well-behaved (i.e., strictly increasing and strongly concave) utility function that (strongly) rationalizes $\mathbb{D}$. Interestingly, Theorem 1 shows that the Hessian matrix of this utility function takes a very simple form since it is diagonal with all non-zero (diagonal) entries equal to $-\varepsilon < 0$.[5] Although there may exist other utility functions that rationalizes the data, it is important to note that the utility function in Theorem 1 does not rely on any parametric assumptions and is very flexible since it consists of twice as many parameters as there are observations.

3 Income elasticity

This section shows that the income elasticities derived from the rationalizing utility function in Theorem 1 are simple functions of the data $\mathbb{D} = \{\mathbf{p}_t, \mathbf{x}_t\}_{t\in\mathbb{T}}$. Define total expenditure as $m = \mathbf{px}$, and let $\mathbf{h}(\mathbf{p}, m) = (h_1(\mathbf{p}, m), ..., h_K(\mathbf{p}, m))$ denote the vector of (Marshallian) demand functions at prices $\mathbf{p}$ and expenditure $m = \mathbf{px}$. Also, define $\nabla_m \mathbf{h}(\mathbf{p}, m) = (\partial h_1(\mathbf{p}, m)/\partial m, ..., \partial h_K(\mathbf{p}, m)/\partial m)$ as the gradient of the demand function with respect to expenditure $m = \mathbf{px}$. Using comparative statics $\nabla_m \mathbf{h}(\mathbf{p}, m)$ can be calculated from the Hessian of the utility function as:[6]

$$\nabla_m \mathbf{h}(\mathbf{p}, m) = \frac{\left(\nabla^2_{\mathbf{xx}} u(\mathbf{x})\right)^{-1} \mathbf{p}}{\mathbf{p}\left(\nabla^2_{\mathbf{xx}} u(\mathbf{x})\right)^{-1} \mathbf{p}}. \tag{2}$$

Let the income elasticity of good $k \in \mathbb{K}$ be defined as:

$$E_k(\mathbf{p}, m) = \frac{\partial h_k(\mathbf{p}, m)}{\partial m} \frac{m}{x_k}.$$

Using the Hessian (1) to calculate the income elasticity for the rationalizing utility function in Theorem 1, evaluated at observation $t \in \mathbb{T}$, yields:

$$E_{kt}(\mathbf{p}_t, m_t) = \frac{p^2_{kt}}{\mathbf{p}_t\mathbf{p}_t} \frac{1}{w_{kt}}, \tag{3}$$

where $m_t = \mathbf{p}_t\mathbf{x}_t$ is total expenditure at observation $t \in \mathbb{T}$, and

$$w_{kt} = \frac{p_{kt}x_{kt}}{m_t},$$

denotes the budget share for good $k \in \mathbb{K}$ at observation $t \in \mathbb{T}$.

The income elasticity, E_{kt}, is given by the product between the reciprocal of the budget share of the k^{th} good, $1/w_{kt}$, and the quadratic price of the k^{th} good relative to the sum of quadratic prices, $p^2_{kt}/\mathbf{p}_t\mathbf{p}_t = p^2_{kt}/\sum_{j\in\mathbb{K}} p^2_{jt}$. Hence, E_{kt} is independent of any unknown (nuisance) parameters, and consequently, can be calculated directly from the observed data $\mathbb{D} = \{\mathbf{p}_t, \mathbf{x}_t\}_{t\in\mathbb{T}}$. Note, however, that the expressions (3) are local in the sense that they only hold at the observed data points.

[5] The Hessian (1) is not explicitly stated in Chiappori and Rochet (1987) but can be derived from Lemma 1 and the last two equations on page 690 in their paper.

[6] See, for example, Shone (1975, Ch. 4.2, p.82-91).

4 The income elasticity when revealed preference is violated

A data set may violate SSARP, in which case there cannot exist any utility function that (strongly) rationalizes the data. In this section, I outline two simple ways of calculating the income elasticities when the data violates SSARP.

In the first method, called *Method A*, I propose to simply delete the observations that are causing the violation of SSARP, and calculate the income elasticities from the remaining set of observations that satisfies SSARP. Appendix A describes a computationally efficient method based on solving a mixed integer linear programming problem to calculate the maximal subset of the data consistent with SSARP. In a second step, the income elasticities are calculated at every observation in this set from (3). Although this method is simple and straightforward, it has the disadvantage that the maximal subset of consistent observations may not be unique. In other words, there may be two (or more) distinct maximal subsets of the same size that satisfies SSARP, and gives different income elasticities, but are impossible to discriminate between.

The second method, called *Method B*, consists of applying the Afriat efficiency index (AEI) to calculate the minimal adjustment of expenditure such that the data satisfies a weaker form of SSARP, called SSARP(e), and then use these adjustments when calculating the income elasticities. Let $e \in (0,1]$ be a scalar, and define for all $s,t \in \mathbb{T}$ the relation $R^D(e)$ as $\mathbf{x}_t R^D(e)\,\mathbf{x}_s$ if $e\mathbf{p}_t\mathbf{x}_t \geq \mathbf{p}_t\mathbf{x}_s$, or equivalently (See Appendix B):

$$\sum_{k=1}^{K} w_{kt}^{e} \geq \frac{\mathbf{p}_t\mathbf{x}_s}{\mathbf{p}_t\mathbf{x}_t}, \tag{4}$$

where $w_{kt}^e = e \times w_{kt}$ for all goods $k \in \mathbb{K}$ and observations $t \in \mathbb{T}$. Let $R(e)$ be the transitive closure of $R^D(e)$. Analogous to Definition 1, a data set $\mathbb{D} = \{\mathbf{p}_t, \mathbf{x}_t\}_{t\in\mathbb{T}}$ satisfies SARP(e) if $\mathbf{x}_t R(e)\,\mathbf{x}_s$ and $\mathbf{x}_t \neq \mathbf{x}_s$ implies $\sum_{k=1}^{K} w_{ks}^e < \frac{\mathbf{p}_s\mathbf{x}_t}{\mathbf{p}_s\mathbf{x}_s}$ (or equivalently $e\mathbf{p}_s\mathbf{x}_s < \mathbf{p}_s\mathbf{x}_t$), and satisfies SSARP($e$) if: *(i)* $\mathbb{D}$ satisfies SARP(e); and *(ii)* if $\mathbf{p}_t \neq \mathbf{p}_s$ implies $\mathbf{x}_t \neq \mathbf{x}_s$ for all $s,t \in \mathbb{T}$. The AEI is defined as the maximal value of e that satisfies SSARP(e).[7] I propose the following adjusted version of the income elasticity:

$$E_{kt}^{\mathrm{AEI}}(\mathbf{p}_t, m_t) = \frac{p_{kt}^2}{\mathbf{p}_t\mathbf{p}_t}\frac{1}{w_{kt}^{\mathrm{AEI}}},$$

where $w_{kt}^{\mathrm{AEI}} = \mathrm{AEI}\times w_{kt}$. The advantage of this method is that it is very easy to calculate the AEI using a binary search algorithm; Appendix B provides the details. However, one disadvantage is that the method is only applicable for a "sufficiently large" AEI, but what constitutes "sufficiently large" is subjective. Varian (1990) suggests a lower bound of 0.95, but points out that the level should depend on the problem at hand, that is, the number of observations, the power of the test, and the model under consideration. Moreover, the AEI only accounts for the worst violation and neglects all other violations. Consequently, a single large violation can make the AEI arbitrarily small even if there are no other violations (Dean and Martin, 2016).

5 Do income elasticities differ between households?

This section applies the proposed framework to a large micro panel of households' consumption expenditures. The purpose of this application is to answer the question how much income elasticities differ

[7] The AEI (also known as the critical cost efficiency index, CCEI) was introduced by Afriat (1972) and Varian (1990), and is a measure of wasted income: if a consumer has an AEI of $e < 1$, then he could have obtained the same level of utility by spending only a fraction of e of what he actually spent to obtain this level.

between heterogeneous households.

Data. I use data from the Spanish Continuous Family Expenditure Survey (Encuesta Continua de Presupuestos Familiares, abbreviated ECPF). This panel is a quarterly budget survey, ranging from 1985-1997, that interviews Spanish households for up to a maximum of eight consecutive quarters on their consumption expenditures. This data was obtained from Crawford (2010) and has previously been analyzed in, among others, Browning and Collado (2001) and Cherchye et al. (2015). From this data, I use a subsample of couples with and without children, where the husband is employed full-time and the wife is outside the labor force. Moreover, I assume that durable goods are weakly separable from nondurables and therefore only consider consumption expenditures on nondurable consumption categories.[8] The price data are national price indices for the corresponding expenditure categories. Overall, there is data on 21,866 observations for 3,134 households.

I focus on consumption expenditures for the category "food and non-alcoholic drinks at home", and calculate income elasticities for this category for every individual household.[9] By analyzing each household separately, I allow for maximal unrestricted heterogeneity between households since there is no need to specify how the heterogeneity enters in the model.[10] As such, the analysis even captures nonobservable heterogeneity between households, and we avoid any (debatable) preference homogeneity assumptions across households that share similar demographic factors. I follow the standard convention in the demand literature and present mean income elasticities across all time periods.

Results. From the total 3,134 households, I exclude 14 households who did not record any purchases of "food and non-alcoholic drinks at home", leaving us with data on 3,120 households (21,772 observations). Out of these 3,120 households, the consumption data for 2,916 (93.46%) households (20,266 observations) satisfy SSARP. Hence, the data for 204 (6.54%) households (1,506 observations) violated SSARP.[11] I begin looking at those households who pass SSARP. The graph on the left-hand side in Figure 1 presents the kernel density of the income elasticities (red solid line).[12] The distribution is unimodal and marginally positively skewed. It has a mode around 0.2, implying that the elasticities are clustered around this point. The right tail vanish quickly and shows that the income elasticities very rarely exceeds 0.5. The graph on the right-hand side in Figure 1 plots the kernel density of the budget shares (red solid line). This figure shows that the shares are rather dispersed, with a mode around 0.5. Thus, households are heterogeneous in how much they spend on "food and non-alcoholic drinks at home".

The first panel in Table 1 reports summary statistics of the budget share, the quadratic price of "food and non-alcoholic drinks at home" relative to the sum of all quadratic prices, $p^2_{\text{FOOD}} / \sum_{j \in \mathbb{K}} p^2_j$ (called 'Price share'), and the income elasticity. We see that 50% (Median) of the 2,916 households that satisfy

[8] The nondurables are aggregated into the following 15 consumption categories: (i) food and non-alcoholic drinks at home, (ii) alcohol, (iii) tobacco, (iv) energy at home, (v) services at home, (vi) non-durables at home, (vii) non-durable medicines, (viii) medical services, (ix) transportation, (x) petrol, (xi) leisure, (xii) personal services, (xiii) personal non-durables, (xiv) restaurants and bars, and (xv) travelling.

[9] The analysis is based on a unitary model of household consumption. Note, however, that a recent literature has shown that household behavior can potentially be better explained by a collective household model (e.g., Cherchye et al. 2007).

[10] The standard (parametric) approach in the literature is to pool the household data and perform the analysis conditional on demographic factors to control for heterogeneity between households. However, this approach implicitly assumes that most of the hetorogenity can be explained by the (observable) demographic factors.

[11] The data for the same 2,916 households also satisfy SARP.

[12] Implemented using the default settings in the R package `density` (i.e., with a gaussian kernel).

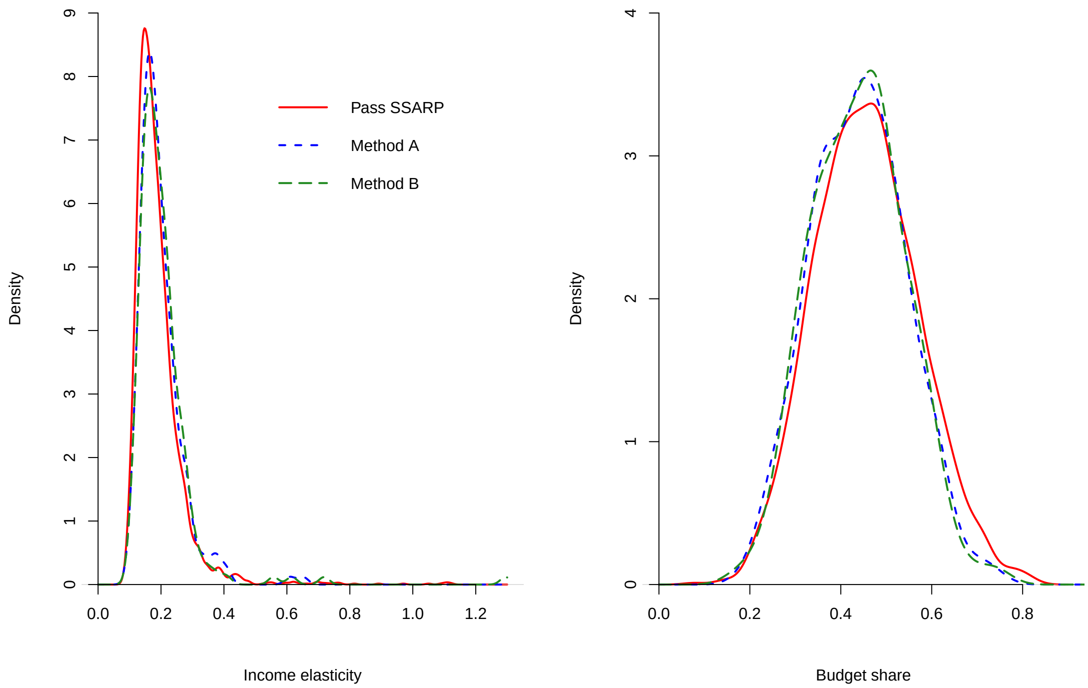

Figure 1: Kernel densities of the income elasticities (left) and budget shares (right).

SSARP have an income elasticity of 0.1719 or lower and that the elasticity is equal or less than 0.2136 for 75% (3rd quartile) of these households. In general, the results in the first panel in Table 1 mainly reveal two things. First, the relative price of "food and non-alcoholic drinks at home" is fairly constant across households (its standard deviation is very low). Second, the income elasticities do not seem to vary much across households, but when it does, the variation is almost entirely due to differences in the fraction of expenditure that is spent on "food and non-alcoholic drinks at home", as indicated by the budget shares. In our model, the income elasticity is inversely related to the budget share, implying that the elasticity increases as the fraction of expenditure spent on "food and non-alcoholic drinks at home" decreases (the correlation between the income elasticity and the budget share is -0.3645).[13]

Consider next the 204 households that violated SSARP. The graph on the left-hand side in Figure 1 presents the kernel density of the income elasticities calculated using Method A (blue dashed line) and Method B (green long-dashed line). As seen from this figure, the distributions are very close to the distribution of the elasticities for households that satisfied SSARP. This can also be seen from the kernel densities of the budget shares in the graph on the right-hand side in Figure 1. The second and third panels in Table 1 reports summary statistics calculated using the two different methods. The second panel reports the results from Method A. The HMI reported in the first row gives summary statistics across all households of the maximal fraction of observations consistent with SSARP. These numbers indicate that households are only making minor violations of SSARP. As seen from Table 1, the income elasticities are very similar to those households that satisfy SSARP, which confirms what we see from

[13]Note that the budget share is also inversely related to the income elasticity in parametric demand models such as the (quadratic) almost ideal demand system and the translog demand system.

Table 1: Summary statistics across all households

	Households	Mean	Std.	Min	1st quartile	Median	3rd quartile	Max
				Households that satisfy SSARP				
Budget share	2,916	0.4570	0.1145	0.0683	0.3771	0.4539	0.5337	0.8347
'Price share'	2,916	0.0735	0.0083	0.0596	0.0624	0.0763	0.0807	0.0827
Income elasticity	2,916	0.1973	0.1853	0.0853	0.1434	0.1719	0.2136	5.2822
				Households that violate SSARP: Method A				
HMI	204	0.8572	0.0333	0.6667	0.8571	0.8750	0.8750	0.8750
Budget share	204	0.4397	0.1036	0.1910	0.3629	0.4419	0.5126	0.7378
'Price share'	204	0.0745	0.0080	0.0592	0.0676	0.0787	0.0810	0.0827
Income elasticity	204	0.1985	0.0724	0.1073	0.1538	0.1826	0.2253	0.6510
				Households that violate SSARP: Method B				
AEI	204	0.9963	0.0048	0.9698	0.9949	0.9980	0.9993	1.0000
Budget share	204	0.4395	0.1027	0.1682	0.3607	0.4372	0.5061	0.7525
'Price share'	204	0.0746	0.0080	0.0596	0.0672	0.0787	0.0808	0.0825
Income elasticity	204	0.2056	0.1076	0.1070	0.1548	0.1867	0.2269	1.2945

the kernel densities in Figure 1. Consider next the results from Method B in the third panel. The AEI is high across the 204 households, which supports the results from the HMI indicating that households are only making minor violations of SSARP. Overall, the income elasticities are very similar to those obtained using Method A.

Are the estimates of the income elasticity reasonable? A partial answer can be given by comparing the results to previous studies that estimates the income elasticity for food and non-alcoholic beverages at home. Also using the ECPF data, Christensen (2014) estimates a parametric demand model and finds income elasticities in the range of $0.6 - 0.7$. Seale et al. (2003) estimates income elasticities of food and beverages for low, middle and high income countries at the macro level and finds an income elasticity of 0.44 for Spain, and a mean elasticity (across all high income countries) of 0.34. Using U.K. household data in repeated cross-sections, Blundell, Pashardes and Weber (1993) find estimates of the income elasticity to be in the range $0.5 - 0.6$. Using more recent household data covering the period 1998-2000 from the U.K. National Food Survey (NFS), Lechene (2001) estimates the income elasticity for "food and beverages at home" to be 0.2. Also using U.K. household data from repeated cross-sections, Paluch et al. (2012) applies a nonparametric framework to estimate micro and macro income elasticities, and finds elasticities for food between 0.16 and 0.2. Based on the results from these studies, the income elasticities seem reasonable, and match especially well with the latter studies on the demand for food and non-alcoholic beverages in the U.K.

6 Conclusion

This paper outlines a framework to calculate income elasticities from a flexible non-parametric utility function. By construction, this utility function satisfies theoretical regularity conditions, and the framework can be applied to any number of goods and to any number of observations. The income elasticities are simple functions of the price-quantity data, and as such, are point identified from the data. In other words, the income elasticities do not depend on any parameters that need to be estimated.

Usually, income elasticities are presented together with various types of price elasticities in empirical demand analysis. In the current framework, Marshallian and Hicksian price elasticities are functions of parameters that need to be calculated (i.e., utility and marginal utility indices in Chiappori and Rochet's (1987) rationalizing utility function). These parameters are not point identified, implying that price elasticities are not point identified. However, under additional restrictions, for instance by imposing that the utility function satisfies the law of demand at every observation, it is possible to show that price elasticities are partially (set) identified. I leave further analysis of such price elasticities for future work.

References

Afriat, S.N. (1967). The construction of a utility function from expenditure data. *International Economic Review* **8**, p. 67-77.

Almås, I. (2012). International income inequality: Measuring PPP bias by estimating Engel curves for food. *American Economic Review* **102**, p. 1093-1117.

Barnett, W.A. and A. Serletis (2008). Consumer preferences and demand systems. *Journal of Econometrics* **147**, p. 210-224.

Barnett, W.A. (2002). Tastes and technology: Curvature is not sufficient for regularity. *Journal of Econometrics* **108**, p. 199-202.

Blundell, R., M. Browning and I. Crawford (2008). Best nonparametric bounds on demand responses. *Econometrica* **76**, p. 1227–1262.

Blundell, R., P. Pashardes and G. Weber (1993). What do we learn about consumer demand patterns from micro data? *American Economic Review* **83**, p. 570-597.

Browning, M. and M.D. Collado (2001). The response of expenditures to anticipated income changes: panel data estimates. *American Economic Review* **91**, p. 681–692.

Chavas, J.-P. and T. Cox (1997). On nonparametric demand analysis. *European Economic Review* **41**, p. 75-95.

Cherchye, L., T. Demuynck, B. De Rock, and P. Hjertstrand (2015). Revealed preference tests for weak separability: An integer programming approach. *Journal of Econometrics* **186**, p. 129-141.

Cherchye, L., B. De Rock and F. Vermeulen (2007). The collective model of household consumption: a nonparametric characterization. *Econometrica* **75**, p. 553–574.

Chiappori, P.-A. and J.-C. Rochet (1987). Revealed preferences and differentiable demand. *Econometrica* **55**, p. 687-691.

Christensen, M. (2014). Heterogeneity in consumer demands and the income effect: Evidence from panel data. *The Scandinavian Journal of Economics* **116**, p. 335-355.

Crawford, I. (2010). Habits revealed. *Review Economic Studies* **77**, p. 1382–1402.

Dean, M. and D. Martin (2016). Measuring rationality with the minimum cost of revealed preference violations. *Review of Economics and Statistics* **98**. p. 524-534.

Deaton, A. and J. Muellbauer (1980). *Economics and Consumer Behavior*. Cambridge University Press.

Diewert, W.E. (1973). Afriat and revealed preference theory. *Review of Economic Studies* **40**, p. 419-425.

Diewert, W.E. and T.J. Wales (1985). Flexible functional forms and global curvature conditions. *Econometrica* **55**, p. 43-68.

Heufer, J. and P. Hjertstrand (2015). Consistent subsets: Computationally feasible methods to compute the Houtman-Maks-Index. *Economics Letters* **128**, p. 87-89.

Houtman, M. and J. Maks (1985). Determining all maximal data subsets consistent with revealed preference. *Kwantitative Methoden* **19**, p. 89–104.

Lechene, V. (2001). Income and price elasticities of demand for food. Section 5 of the Annual report on food expenditure, consumption and nutrient intakes, Report of the National Food Survey Committee, London, The Stationery Office.

Matzkin, R. and M. Richter (1991). Testing strictly concave utility. *Journal of Economic Theory* **53**, p. 287-303

Mulligan, C.B. and X. Sala-i-Martin (1992). U.S. money demand: Surprising cross-sectional estimates. *Brookings Papers on Economic Activity* **2**, p. 285-343.

Paluch, M., A. Kneip, and W. Hildenbrand. (2012). Individual versus aggregate income elasticities for heterogeneous populations. *Journal of Applied Econometrics* **27**, p. 847-869

Seale Jr, J., A. Regmi, and J. Bernstein (2003). International evidence on food consumption patterns. Technical Bulletin Number 1904, Electronic Report from the Economic Research Service, United States Department of Agriculture.

Shone, R. (1975). *Microeconomics: A modern treatment*. Academic Press, New York.

Varian, H.R. (1982). The nonparametric approach to demand analysis. *Econometrica* **50**, p. 945-972.

Appendix

A Procedure to find the maximal subset of observations satisfying SSARP

This appendix describes a computationally efficient procedure to calculate the largest subset of the data satisfying SSARP. Chiappori and Rochet (1987) showed that SSARP is equivalent to that there exist numbers u_t and $\lambda_t > 0$ for all $t \in \mathbb{T}$ such that the data $\mathbb{D} = \{\mathbf{p}_t, \mathbf{x}_t\}_{t\in\mathbb{T}}$ satisfies the following inequalities (for all $s, t \in \mathbb{T}$):

$$u_s - u_t - \lambda_t (\mathbf{p}_t\mathbf{x}_s - \mathbf{p}_t\mathbf{x}_t) \leq -\varepsilon\Psi, \tag{A1}$$

where $\varepsilon > 0$ is a scalar and Ψ is defined as $\Psi = \max_{s,t\in\mathbb{T}} \left\{ \frac{1}{2} \sum_{k\in\mathbb{K}} (x_{kt} - x_{ks})^2 \right\}$. These inequalities can be weakened by introducing a binary $T-$dimensional vector $\mathbf{h} = (h_1, ..., h_T) \in \{0,1\}^T$ and multiply total expenditure in each time period with the corresponding period t binary variable h_t. Replacing expenditure $\mathbf{p}_t\mathbf{x}_t$ with the product $h_t\mathbf{p}_t\mathbf{x}_t$ for every $t \in \mathbb{T}$ in (A1) gives the weaker set of inequalities:

$$u_s - u_t - \lambda_t (\mathbf{p}_t\mathbf{x}_s - h_t\mathbf{p}_t\mathbf{x}_t) \leq -\varepsilon\Psi. \tag{A2}$$

The Houtman–Maks index (HMI) is defined as the maximal fraction of non-zero elements in the binary vector $\mathbf{h}$ such that the inequalities (A2) hold. One way to calculate the HMI is to find the $\mathbf{h}$ closest to the unit vector in the L1-norm, which consists of solving the nonlinear problem:

$$\text{HMI} = \arg \min_{\{u_t,\lambda_t,h_t\}_{t\in\mathbb{T}}} \frac{1}{T} \sum_{t\in\mathbb{T}} (1 - h_t) \quad \text{s.t.} \quad u_s - u_t - \lambda_t (\mathbf{p}_t\mathbf{x}_s - h_t\mathbf{p}_t\mathbf{x}_t) \leq -\varepsilon\Psi, \tag{A3}$$

for all $s, t \in \mathbb{T}$. Let $\widehat{\mathbf{h}} = \left(\widehat{h}_1, ..., \widehat{h}_T\right)$ be the optimal solution from this problem. The largest subset of the data $\mathbb{D}$ consistent with SSARP (in the L1-norm) is given by set of observations for which $\widehat{h}_t = 1$.

However, the problem (A3) is of limited empirical use since the constraints are nonlinear in the term $\lambda_t h_t \mathbf{p}_t \mathbf{x}_t$. A more convenient implementation can be obtained by noticing that the term $\lambda_t h_t$ is the product of a binary variable and a continuous variable and that the latter without loss of generality can be bounded as $0 < L \leq \lambda_t \leq U$ with $L \neq U$. By defining the variable $y_t = \lambda_t h_t$, it is easy to show that the product is equivalent to the following set of linear inequalities:

$$\begin{aligned} &L h_t \leq y_t \\ &y_t \leq U h_t \\ &L (1 - h_t) \leq \lambda_t - y_t \\ &\lambda_t - y_t \leq U (1 - h_t) \\ &0 \leq y_t \leq U. \end{aligned} \tag{A4}$$

This allow us to linearize the non-linear term in (A2) with the linear equalities (A4), and reformulate the problem (A3) as a mixed integer linear programming (MILP) problem since the elements in $\mathbf{h}$ only takes binary values. Specifically, I propose to calculate the HMI (in the L1-norm) by solving the following

MILP problem (for a given $\varepsilon > 0$):

$$\begin{aligned}
\text{HMI} \quad = \quad & \arg \min_{\{u_t, \lambda_t, h_t, y_t\}_{t \in \mathbb{T}}} \frac{1}{T} \sum_{t \in \mathbb{T}} (1 - h_t) \quad \text{s.t.} \qquad \text{(A5)} \\
& u_s - u_t - \lambda_t \mathbf{p}_t \mathbf{x}_s + y_t \mathbf{p}_t \mathbf{x}_t \leq -\varepsilon \Psi \\
& L h_t - y_t \leq 0 \\
& y_t - U h_t \leq 0 \\
& y_t - \lambda_t - L h_t \leq -L \\
& \lambda_t - y_t + U h_t \leq U \\
& L \leq \lambda_t \leq U \\
& h_t \in \{0, 1\} \\
& 0 \leq y_t \leq U.
\end{aligned}$$

This problem gives an exact and global solution (because every local solution to a MILP problem is a global solution), and there exist efficient algorithms for solving such problems in practice (e.g. branch and bound, and cutting plane).

Finally, some remarks on the origins and generalizations of this procedure. Houtman and Maks (1985) originally proposed the HMI as a measure of goodness-of-fit, or more precisely, as a measure of how close a data set $\mathbb{D} = \{\mathbf{p}_t, \mathbf{x}_t\}_{t \in \mathbb{T}}$ comes to satisfying revealed preference. Heufer and Hjertstrand (2015) propose a computationally efficient procedure to calculate the HMI for the generalized axiom of revealed preference (GARP). GARP is the standard model in empirical applications of consumer rationality, and its significance is summarized by the well-known Afriat's theorem, which states that a data set $\mathbb{D} = \{\mathbf{p}_t, \mathbf{x}_t\}_{t \in \mathbb{T}}$ can be rationalized by a continuous, strictly increasing and concave utility function if and only if $\mathbb{D}$ satisfies GARP, or equivalently, if and only if $\mathbb{D}$ satisfies the so called Afriat inequalities (Varian, 1982). A close inspection of the inequalities (A1) show that these reduces to the Afriat inequalities when imposing $\varepsilon = 0$. Hence, by setting $\varepsilon = 0$, the problem (A5) can be used to calculate the HMI for GARP. Thus, in relation to Heufer and Hjertstrand (2015), the problem (A5) (with $\varepsilon = 0$) provides an equivalent but alternative method to calculate the HMI for GARP.

B Calculating the income elasticities using the AEI

Derivation of Eq. (4). Recall that $\mathbf{x}_t R^D(e)\, \mathbf{x}_s$ if $e\mathbf{p}_t\mathbf{x}_t \geq \mathbf{p}_t\mathbf{x}_t$ and that w_{kt} denotes the budget share for good $k \in \mathbb{K}$ at observation $t \in \mathbb{T}$. We have:

$$\begin{aligned}
& e\mathbf{p}_t\mathbf{x}_t \geq \mathbf{p}_t\mathbf{x}_s \\
\iff \quad & e \geq \frac{\mathbf{p}_t\mathbf{x}_s}{\mathbf{p}_t\mathbf{x}_t} \\
\iff \quad & e \sum_{k=1}^{K} w_{kt} \geq \frac{\mathbf{p}_t\mathbf{x}_s}{\mathbf{p}_t\mathbf{x}_t} \quad [\text{since } \sum_{k=1}^{K} w_{kt} = 1] \\
\iff \quad & \sum_{k=1}^{K} e w_{kt} \geq \frac{\mathbf{p}_t\mathbf{x}_s}{\mathbf{p}_t\mathbf{x}_t} \\
\iff \quad & \sum_{k=1}^{K} w_{kt}^{e} \geq \frac{\mathbf{p}_t\mathbf{x}_s}{\mathbf{p}_t\mathbf{x}_t},
\end{aligned}$$

where $w_{kt}^{e} = e \times w_{kt}$ for all $k \in \mathbb{K}$ and $t \in \mathbb{T}$.

Binary search algorithm to calculate the AEI. Consider the relation $R^D(e)$. If $R^D(e)$ holds for some e then it also holds for some $e' < e$. This monotonicity condition implies that the maximal e such that SSARP(e) holds can be calculated using a simple binary search algorithm. Let F_l denote an initial feasible lower bound of e, i.e., a number e chosen sufficiently small such that the data $\mathbb{D} = \{\mathbf{p}_t, \mathbf{x}_t\}_{t\in\mathbb{T}}$ satisfies SSARP(e). Without loss of generality we can set $F_l = 0$. Analogously, let F_u denote an infeasible upper bound, e.g. $F_u = 1$. The binary search consists of the following steps. Start with the midpoint $(F_u + F_l)/2$ and check whether SSARP$\left(e^{(1)}\right)$ hold with $e^{(1)} = (F_u + F_l)/2$. If an optimal solution is found, then set $F_l = e^{(1)}$ (keeping F_u fixed) and check whether $e^{(2)} = (F_u + F_l)/2$ satisfy SSARP$\left(e^{(2)}\right)$. On the other hand, if no optimal solution is found, set $F_u = e^{(1)}$ (keeping F_l fixed) and check whether $e^{(2)} = (F_u + F_l)/2$ satisfy SSARP$\left(e^{(2)}\right)$. At each iteration of the binary search, the range $[F_l, F_u]$, which contains the solution, is halved. As such, the width of the interval decreases exponentially in the number of iterations. A possible termination criterion of the binary search is when $(F_u - F_l)/F_l \leq \psi$, for some very small positive number ψ (in the application I set $\psi = 10^{-12}$).