Revealed Preference Heterogeneity

(Job Market Paper)

Abi Adams*

University of Oxford

Institute for Fiscal Studies

Abstract

Evaluating the merits of alternative tax and incomes policies often requires knowledge of their impact on consumer demand and welfare. Attempts to quantify these impacts are complicated; aggregate demand responses depend on the population distribution of preferences for commodities, yet consumer preferences go unobserved and economic theory places few restrictions on their form and distribution. These complications become especially acute when individuals are making choices over many goods because transitivity must be imposed for rationality of demand predictions and simultaneity must be addressed. In this paper, I develop a revealed preference methodology to bound demand responses and welfare effects in the presence of unobserved preference heterogeneity for many-good demand systems. I first derive the revealed preference restrictions that are implied by a simple random utility model, and then develop these inequalities into a linear programming problem that allows for the recovery of the model's underlying structural functions. I show how the feasible set of this linear programme can be used to construct virtual prices that enable one to conduct positive and normative analysis for heterogeneous agents. The utility of this approach is demonstrated through an application to household scanner data in which multidimensional preference parameters are recovered, and individual demands and the distribution of demands are predicted for hypothesised price changes.

*I would like to thank Richard Blundell, Steve Bond, Martin Browning, Laurens Cherchye, Ian Crawford, Bram De Rock, Stefan Hoderlein, Jeremias Prassl, Collin Raymond, Frederick Vermeulen and seminar participants at cemmap, Institute for Fiscal Studies and Oxford for useful discussion and comments. I gratefully acknowledge financial support from the European Research Council (ERC) under ERC-2009-AdG grant agreement number 429529. Data supplied by TNS UK Ltd. The use of TNS UK Ltd. data in this work does not imply the endorsement of TNS UK Ltd. in relation to the interpretation or analysis of the data. All errors and omissions are my own. Contact Information: Department of Economics, University of Oxford and Institute for Fiscal Studies, London. Email: `abi.adams@economics.ox.ac.uk`

1 Introduction

The existence of pervasive preference heterogeneity is uncontroversial in consumer theory. Otherwise identical consumers respond in myriad ways to the same economic incentives. Given this empirical fact, accurate prediction of the aggregate distribution of consumer demand, and the unique calculation of welfare responses, requires knowledge of the distribution of preference heterogeneity, and the nature of its interaction with economic constraints. However, preferences cannot be directly observed and economic theory has little to say regarding their specific functional form and population distribution. This creates significant challenges for policy makers when evaluating the merits of tax and incomes policies, and for firms when predicting changes in market demand.

There exist parametric and nonparametric methods that allow one to incorporate unobserved preference heterogeneity into applied demand analysis. However, the assumptions made by parametric specifications are ad hoc and place strong restrictions on individual preferences, whilst many nonparametric methods require that observables have a continuous density or that one observes distributions of demand at a finite number of budgets (i.e. one observes many different individuals making choices in the same price-income environment). Further, it is often the case that fully general nonparametric methods are so flexible that they do not constrain the location of demand responses at new budgets of interest, reducing their utility for policy analysis.

My goal in this paper is to develop a nonparametric method that allows for the recovery of informative bounds on behavioural responses and welfare effects when one only has access to finite (repeated) cross-section data in which consumers purchase many goods. I do not assume that sample prices have a continuous density nor that one observes many different consumers making choices at the same budget. Many applied data sets fit this assumed data environment; high-dimensional cross section data is typically sparse and one often observes individuals making choices subject to different market prices with different incomes. In this paper, I work with a simple, yet flexible, random utility model that allows me to incorporate unobserved preference heterogeneity in a tractable way. This model cannot be rejected for the data set used in my empirical application and yields informative bounds on individual demand responses and welfare effects with hypothesised price changes.

In this introduction, I first outline the challenges posed by unobserved preference heterogeneity for applied work, before situating my proposed revealed preference solution within the existing literature.

The Challenge To appreciate the problems caused by variation in consumer preferences, it is useful to first consider the circumstances under which this heterogeneity could be dealt with in a straightforward manner. Unobserved preference heterogeneity would not be such a thorny issue for applied demand theorists if they had access to long panels on individual consumption data. In this instance, standard revealed preference and integrabilty theory could be applied to recover the (fixed) preferences generating the observable choices of that individual using the time series variation in prices and income. However, consumer panel data is as life under the state of nature: poor and short.[1] Many data sets do not have a sufficiently long panel dimension, nor sufficient price variation, to identify preferences nor generate tight bounds on demand responses.

Econometricians typically deal with the paucity of data at the individual level by pooling observations across consumers and estimating a statistical demand model. This strategy requires that one make assumptions regarding preference heterogeneity in the population. These assumptions, in effect, enable the applied demand theorist to assume away the fact that she operates in a low information environment. For example, if one were to assume that all consumers had identical preferences, it would be as if they had access to much more 'individual' level data with which to predict a consumer demands; data on consumer i's past choice behaviour would be fully informative for predicting consumer j's demands at new budgets.

Although mainstream econometric models are not so restrictive as to assume homogeneous preferences, they do often require that unobserved preference heterogeneity can be captured by an additive error term that is appended to a deterministic relationship which is assumed to hold between observables. That is,

$$\mathbf{q}_i = \mathbf{d}(\mathbf{p}_i, x_i, \mathbf{z}_i) + \boldsymbol{\epsilon}_i \tag{1}$$

i.e. consumer i's demand is generally modeled as a systematic function of prices, $\mathbf{p}$, income, x, and observable characteristics, $\mathbf{z}$, plus an individual-specific component, $\boldsymbol{\epsilon}_i$, that supposedly captures unobserved preference heterogeneity. As Barten (1977) comments, these preference "disturbances are usually tacked on to demand equations as a kind of afterthought."

The desirability of the standard approach is diminished with the recognition that it places very strong assumptions upon the underlying structure of individual preferences. For $\mathbf{d}(\mathbf{p}_i, x_i, \mathbf{z}_i)$ to be integrable and rational, the covariance between unobserved preference parameters and income effects must satisfy certain strong assumptions (see Lewbel, 2001). When these assumptions are violated, model errors cannot be interpreted as unobserved preference heterogeneity terms. Further, even if such assumptions are satisfied, $\boldsymbol{\epsilon}$ will be functionally dependent on prices and incomes except in a handful of special cases (Brown and Walker (1989), Lewbel (2001)). Estimation and inference with this in mind becomes difficult without parametric specifications for the distribution of unobserved preference

[1]See Thomas Hobbes (1651), *Leviathan*, Chapter XIII: Of the Natural Condition of Mankind As Concerning their Felicity and Misery.

parameters and/or the utility function.

Proceeding in a manner that does not unduly restrict the functional form of individual preferences, and does not require additional assumptions on the distribution of unobserved preference parameters, requires the analysis of random preference terms that are non-additive in the demand function, i.e. $\mathbf{d}(\mathbf{p}, x, \mathbf{z}, \boldsymbol{\epsilon})$. Further, for multidimensional demand systems, where $\boldsymbol{\epsilon} \in \mathbb{R}^K$ with $K > 2$, not just non-additivity of unobservables but simultaneity must be addressed; in cases where a different unobserved preference parameter is associated with every good, the demand function for each commodity will typically depend on the value of every unobserved parameter. In this instance, the demand for any given commodity cannot be treated independently of the unobservables associated with other goods. Furthermore, when $K > 2$ transitivity must be imposed for rationality of the demand system.[2]

Current Solutions There is a large literature that addresses the identification and estimation of demand models with nonseparable unobserved preference heterogeneity. For example, random coefficient models typically specify a reduced form specification between demands and other observable characteristics in which the coefficients on observables are random with a non-degenerate distribution (see, for example, Hoderlein, Klemelae and Mammen (2010)). However, such models are typically rather restrictive regarding the permissible class of utility functions generating choice behaviour. Although, more recently, Lewbel and Pendakur (2013) show the nonparametric identification of a random coefficient model with multiple goods in which random coefficients can be interpreted as Barten scales.

Where unobserved heterogeneity is restricted to be unidimensional, more flexible specifications for the demand function are more easily dealt with. For example, Blundell, Kristensen and Matzkin (2014) and Blundell, Horowitz and Parey (2010) use quantile methods to estimate heterogeneous demand functions subject to restrictions from economic theory. Hoderlein and Vanhems (2010) address the identification and estimation of welfare measures when demands depend on a single unobservable. Although these methods are flexible in the contexts in which they are applied, extending quantile-based approaches to a multivariate setting is non-trivial due to the lack of an objective basis for ordering multivariate observations. In this paper, I will tackle multidimensional heterogeneity and must, therefore, take a different approach to this strand of the literature.

This paper is also related to the literature on the identification of nonparametric simultaneous equation models with nonadditive errors. Matzkin (2008) derived general identification results for non-additive simultaneous equation models, following the results of Benkard and Berry (2004) that showed that earlier results of Brown (1983) and Roehrig (1988) do not necessarily guarantee identification.

[2]With only two goods, transitivity has no empirical content and imposing the weak axiom is sufficient for rationality (Rose, 1958).

Matzkin (2010) develops a new approach for the estimation of nonparametric, nonadditive models with simultaneity that is closely related to her proofs of identification. My theoretical framework will be tied to an earlier paper in this literature by Brown and Matzkin (1998), who developed a nonparametric Closest Empirical Distribution method, following Manski (1983), which does not require a parametric specification for the utility function nor unobserved heterogeneity.

Finite Data In contrast to the work above, I here address the recovery of multidimensional unobserved heterogeneity when only a finite number of budgets are observed. This data environment complicates the application of methods that rely on one having sufficient data to identify the derivatives of the functions of interest. I will here deal with the finite, raw demand data itself and use techniques of finite mathematics to recover unobserved preference parameters that are able to rationalise observed behaviour. The methodology follows in the revealed preference tradition of Samuelson (1938, 1948), Afriat (1967) and Varian (1982). Optimising behaviour within a theoretical framework provided by economic theory yields restrictions on choice behaviour. The revealed preference literature derives these restrictions and uses them to test whether an economic model is capable of explaining the variation in the data, and also to set identify features of the structural functions of interest.

Revealed preference restrictions have been derived for general random utility models that can be applied to rationalise the behaviour of cross section distributions of demand with unobserved preference heterogeneity (see Block and Marschak (1960), McFadden and Richter (1991) and McFadden (2005)). The Axiom of Revealed Stochastic Preference (ARSP) represents the set of inequalities that must be satisfied given individual utility maximisation in the presence of unrestricted, unobserved preference heterogeneity. Hoderlein and Stoye (2013) show how to test the weak axiom with a heterogeneous population. In a recent paper, Kitamura and Stoye (2014) provide the tools to bring ARSP to the data nonparametrically (adding the requirement of transitivity to Hoderlein and Stoye's (2013) approach), developing efficient algorithms for encoding the choice data and generating the ARSP inequalities. The output of their procedures can be used to bound expected demands and the distribution of demand at a new budget of interest. Hausman and Newey (2014) show how average surplus can be bounded in a framework with multivariate unobserved heterogeneity using bounds on income effects.

However, the fully general approach of Kitamura and Stoye (2014) typically yields bounds that are "uncomfortably wide" (Kitamura and Stoye, p. 27) and the procedure is computationally very demanding.[3] Further, current revealed preference approaches are designed for application to distributions of demand on the same budget. In many contexts, such as for the household scanner data that I will

[3]The wide bounds could perhaps be a partial product of the fact that they are only able to test ARSP on data from eight time periods at a time.

use in Section 6, significant price and budget heterogeneity are evident in the sample complicating the application of these methods.

My Contribution In this paper, I attempt to find a middle ground between a highly restrictive parametric approach and the fully general approach of Kitamura and Stoye (2014). My goal is to develop a nonparametric method for recovering the sample distribution of unobserved preference heterogeneity and for making informative predictions for demands at new budgets of interest, which remains flexible enough to capture the rich preference heterogeneity evident in consumer microdata. This method will be tailored to reflect the fact that the applied demand theorist faces a finite data environment; I will not assume that sample prices and income have a continuous density nor, in contrast to past revealed preference works, will I assume that one observes distributions of demand at a finite number of price-income regimes. Rather, I will tackle recovery when one observes a finite number of consumers each of whom may face a different market price and income environment.

I take a simple random utility model as an organising theoretical framework and develop a revealed preference approach for testing and recovering the model's underlying structural functions. Although restrictive, I cannot reject this theoretical framework in my empirical application and find that it yields informative predictions for demand responses at new budgets of interest. Further, the method is simple to implement using standard linear programming techniques.

To proceed, I first derive the revealed preference restrictions on observables that are implied by optimising behaviour within the framework. If one can find a solution to these restrictions, then it is possible to find specifications of the model that could have generated observed decision making behaviour. Following Varian (1982), revealed preference approaches to demand prediction identify the "support set" of demands at new budgets of interest. These are the demand bundles that are consistent with the theoretical model plus observed behaviour. Straightforward application of Varian's (1982) results is not possible in our context because this method requires observed behaviour to have been generated by the maximisation of the same utility function. To develop a method for prediction and welfare analysis, I appeal to the literature on virtual budgets, first introduced by Rothbarth (1941) and applied by Neary and Roberts (1980) to develop the theory of choice behaviour under rationing. Using feasible solutions to the revealed preference restrictions, I transform observed price-expenditure combinations into preference-adjusted virtual budgets. The set of virtual budgets for a particular individual are those combinations of prices and incomes that would have caused that individual to choose the set of sample demands that are observed. Once prices are adjusted for differences in preferences in this way, they can be used to bound demand responses and welfare effects at new budgets of interest using revealed preference methods.

Some complications arise from the fact that, without imposition of additional identification conditions, unobserved preference parameters are not uniquely identified by revealed preference arguments. Thus, virtual prices are not uniquely identified. To deal with this, I impose weak restrictions on the utility function and require independence of preference heterogeneity and budget parameters such that there is a unique mapping between observables and unobservables in the population. I then adapt the approach of Brown and Matzkin (1998) to employ Manski's (1983) minimum distance from independence estimation technique to recover a rationalising sample distribution of preference heterogeneity and show that the estimator is consistent for the true distribution of unobserved preference heterogeneity.

I end by demonstrating the utility of the approach using an illustrative empirical application to cross section household scanner data. This data records individual-level quantity and product characteristics and is ideal for the application of revealed-preference type methods. I find that the behaviour of the whole cross section, and of demographic groups, cannot be rationalised without allowing for unobserved preference heterogeneity. I thus recover multidimensional unobserved heterogeneity within cells of observable characteristics and predict demands for various draws from the distribution of unobserved preference heterogeneity for hypothesised price changes. These bounds are shown to be relatively tight and I find that the barycenter of individual support sets closely approximates observed demands. The method thus has predictive value.

Structure This paper proceeds as follows. Section 2 outlines the random utility model that forms the basis of my theoretical framework and Section 3 derives its associated revealed preference restrictions. These restrictions are developed further in Section 4, where I address the prediction of demands and welfare effects for particular draws from the distribution of unobserved heterogeneity. Section 5 derives the conditions under which there is a unique mapping between observables and unobservables in the population, and how these restrictions can be used to refine the solutions to the revealed preference inequalities and allow for virtual prices to be recovered. Finally, Section 6 demonstrates the practical usefulness of the proposed methodology using an empirical application to household scanner data. Section 7 concludes.

2 The Model

The economic theory lying behind classical demand analysis is extremely simple: a consumer chooses the bundle of goods that maximises her utility from the set of all bundles that she can afford. Assuming a linear budget constraint, the consumer's optimisation problem can be formally expressed as:

$$\max_{\mathbf{q}} U(\mathbf{q}, \mathbf{z}, \boldsymbol{\epsilon}) \tag{2}$$

subject to

$$\mathbf{p}'\mathbf{q} \leq x \tag{3}$$

where: $U(\cdot, \mathbf{z}, \boldsymbol{\epsilon})$ is a utility function representing the preferences of the consumer over consumption bundles $\mathbf{q} \in \mathbb{R}^K_+$; $\mathbf{z} \in \mathbb{R}^L$ and $\boldsymbol{\epsilon} \in \mathbb{R}^S$ represent, respectively, vectors of observable and unobservable characteristics that index a consumer's preferences; and $\mathbf{p} \in \mathbf{R}^K_{++}$ and $x > 0$ give the price vector and income that determine the budget constraint that the consumer faces. Prices are normalised such that $p^K = 1$.

Under this framework, variation in choice behaviour across consumers can be explained by differences in the budget sets that different individuals face and by differences in the objective functions that individuals maximise. If preferences were homogeneous across consumers, or if preferences only varied according to observable characteristics, i.e. $U(\mathbf{q}, \mathbf{z}, \boldsymbol{\epsilon}) = U(\mathbf{q}, \mathbf{z})$, then there would only be a single choice on each budget for each demographic cell. Standard results from integrability theory and revealed preference theory could then be applied to recover consumer preferences from observed demands.

In this paper, I will be concerned with recovering preferences when choice *is* dependent upon unobserved preference heterogeneity. I work with a special case of the above model, employing a random utility framework analogous to that analysed by Brown and Matzkin (1998) and McFadden and Fosgerau (2012). This model guarantees invertibility of the demand function in the unobserved preference parameters. Specifically, I restrict the functional form of the utility function to:

$$U(\mathbf{q}, \boldsymbol{\epsilon}) = u(\mathbf{q}) + \sum_{k=1}^{K-1} \epsilon^k q^k \tag{4}$$

$$= u(\mathbf{q}) + \boldsymbol{\epsilon}'\mathbf{q}^{\neg K} \tag{5}$$

where $\boldsymbol{\epsilon} \in \mathbb{R}^{K-1}$ represents unobserved preference heterogeneity and $E(\epsilon^k) = 0$. The distribution of $\boldsymbol{\epsilon}$ across consumers is characterised by an absolutely continuous distribution function $F_\epsilon : \mathbb{R}^{K-1} \rightarrow$

$\mathbb{R}$. Given the dependence of preferences, and thus of consumer demand, upon $\boldsymbol{\epsilon}$, the distribution F_ϵ generates a distribution of demand at each budget set $(\mathbf{p}, x)$. It is further assumed that $U(\mathbf{q}, \boldsymbol{\epsilon})$ is continuous and smooth in the sense of Debreu (Debreu, 1972).

The dependence of consumer preferences upon observable heterogeneity, $\mathbf{z}$, is here suppressed largely for notational ease. Introducing a dependence between preferences and a set of discrete random variables $\mathbf{z}$ is easily done by dividing consumers into cells based upon observable characteristics. The base utility function and distribution of $\boldsymbol{\epsilon}$ can then assumed to be common to each cell (and thus dependent on observables). This strategy will be adopted in the later empirical application.

Given the theoretical framework, unobservable preference heterogeneity manifests itself in shifts to the marginal utilities of the commodities. From the first order conditions, it is the case that

$$\frac{u_k(\mathbf{q})}{u_K(\mathbf{q})} + \frac{\epsilon^k}{u_K(\mathbf{q})} = p^k \tag{6}$$

$$MRS^k(\mathbf{q}) + \frac{\epsilon^k}{u_K(\mathbf{q})} = p^k \tag{7}$$

where $u_k(\mathbf{q}) = \partial u(\mathbf{q})/\partial q_k$. From equation 6, the model generates unobserved preference heterogeneity that is separable in the marginal rate of substitution; across consumers, there exists a common 'base' component to the MRS of the k^{th} good, $MRS^k(\mathbf{q})$, and a separable individual-specific MRS component that is dependent upon a consumer's particular draw of ϵ^k. Figure 1 shows the impact of unobserved preference heterogeneity graphically, depicting a simple 2-good example. The indifference curve passing through $\mathbf{q} = [4, 4]$ is shown for the following three cases:

1. Base utility function: $u(\mathbf{q}) = 0.5\log(q^1) + 0.5\log(q^2)$.

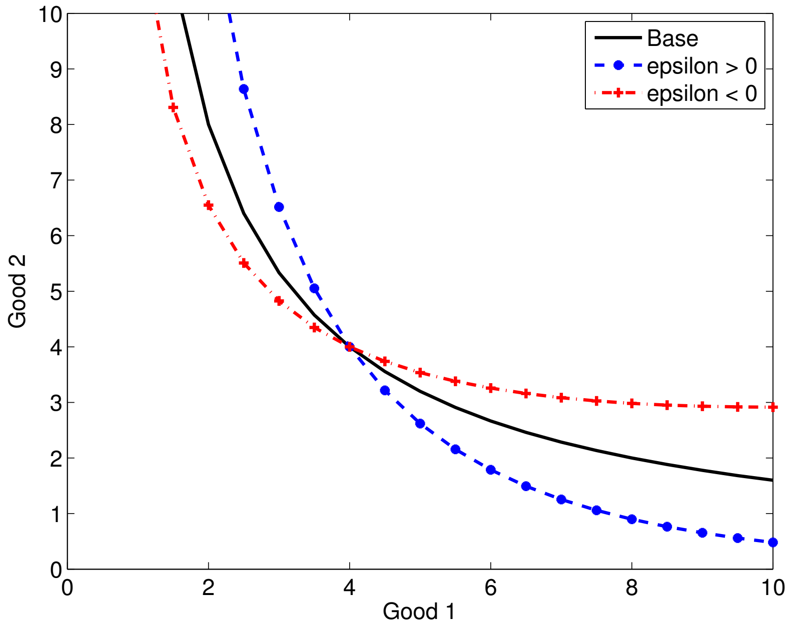
2. Positive perturbation: $u(\mathbf{q}) + \epsilon^1 q^1 = 0.5\log(q^1) + 0.5\log(q^2) + 0.1q^1$.

3. Negative perturbation: $u(\mathbf{q}) - \epsilon^1 q^1 = 0.5\log(q^1) + 0.5\log(q^2) - 0.1q^1$.

The indifference curve associated with $\epsilon = 0.1$ has a steeper gradient through $\mathbf{q} = [4, 4]$ as any reduction in q^1 would have to be compensated with a greater amount of q^2 for the individual to remain on the same indifference curve.

One should note that this specification for the utility function does *not* imply that unobserved heterogeneity is separable in the consumer demand function. From the first order conditions of the model, it is clear that $q^k = f(p^1, ..., p^{K-1}, x, \epsilon^1, ..., \epsilon^{K-1})$ in the absence of separability assumptions or restrictions on income effects. However, it is true that the framework remains rather restrictive regarding the functional structure of unobserved heterogeneity. In a sense, this specification can be considered one step less restrictive than a model that only incorporates observable heterogeneity. If

Figure 1: MRS Perturbation



one did not allow for any interaction between $\mathbf{q}$ and $\boldsymbol{\epsilon}$, i.e. $U(\mathbf{q}, \boldsymbol{\epsilon}) = u(\mathbf{q}) + \boldsymbol{\epsilon}$, then the existence of such heterogeneity would not be behaviourally meaningful.

There clearly exist more flexible specifications for unobserved heterogeneity but these come at the price of wider bounds for demands at new budgets of interest. Although, my framework is restrictive, I cannot reject this model for the data set used in my empirical illustration. I therefore leave the extensions of results in this paper to alternative functional forms for later work.[4]

3 Revealed Preference Restrictions

As a consequence of maximising behaviour within the theoretical framework, a consumer's choice behaviour, and the behaviour of a sample of consumers, will satisfy certain inequality restrictions. Later in this paper, I will show how these restrictions can be used to bound demands and welfare effects.

Imagine that we observe a random sample of the choice behaviour of $N < \infty$ consumers, located in separate geographic markets and thus facing different price regimes, $\{\mathbf{p}_i, \mathbf{q}_i\}_{i=1,\ldots,N}$. Sample prices

[4] A more general version of the random utility model employed in this paper that is able to encompass utility with heterogeneous curvature whilst retaining global invertibility is:

$$U(\mathbf{q}, \boldsymbol{\epsilon}) = u(\mathbf{q}) + v(\mathbf{q}^{\neg K})' \boldsymbol{\epsilon} \tag{8}$$

This functional form is not amenable to the empirical strategy that I employ later in the paper. Specifically, if $v(\cdot)$ is unknown, the necessary and sufficient rationalisation conditions that form the basis of our revealed preference approach to identification, are non-linear in unknowns and cannot be implemented using linear programming techniques. If, however, one is willing to assume a specific function form for $v(\mathbf{q}^{\neg K})$, it is possible to proceed with few amendments to the approach.

do not have a continuous density. In line with the theoretical framework, each consumer is associated with a fixed $(K-1)$-vector $\boldsymbol{\epsilon}^i$ drawn from F_ϵ that is known to them. If a consumer i chooses $\mathbf{q}$ to maximise $u(\mathbf{q}) + \boldsymbol{\epsilon}_i'\mathbf{q}^{\neg K}$, then their behaviour will satisfy (R1) of the following Rationalisation Inequalities. Further, to ensure monotonicity of the utility function for all permissible $\boldsymbol{\epsilon}$, (R2) must hold; intuitively, (R2) places a bound on the support of F_ϵ.

Definition. (Rationalisation Inequalities) Consumer i's choice behaviour is consistent with the maximisation of the utility function $U(\mathbf{q}, \boldsymbol{\epsilon}^i) = u(\mathbf{q}) + \boldsymbol{\epsilon}_i'\mathbf{q}^{\neg K}$ if their choices satisfy the following inequality constraint:

$$u(\mathbf{q}) - u(\mathbf{q}_i) > \boldsymbol{\epsilon}_i'(\mathbf{q}^{\neg K} - \mathbf{q}_i^{\neg K}) \tag{R1}$$

for all $\mathbf{q}$ such that $\mathbf{p}_i'\mathbf{q} \leq \mathbf{p}_i'\mathbf{q}_i$. Monotonicity of $U(\mathbf{q}, \boldsymbol{\epsilon})$ at all recovered $\boldsymbol{\epsilon}$ requires:

$$u_k(\mathbf{q}) + \underline{\epsilon}^k > 0 \tag{R2}$$

for all $i = \{1, ..., N\}$ and $k = \{1, ..., K-1\}$, where

$$\underline{\epsilon}^k = \min_i \epsilon_i^k$$

The satisfaction of the Rationalisation Inequalities cannot be directly tested because $u(\mathbf{q})$ and $\boldsymbol{\epsilon}_i$ are not observed. Further, as each individual is observed only once, a revealed preference test based only on the observation $\{p_i, q_i\}$ is meaningless. Yet, there exist testable revealed preference inequalities defined upon the choices of the cross section. Looking to the random sample of consumers, if there exists a non-empty solution set to the inequalities defined by Theorem 1, then the behaviour of the cross-section can be rationalised by our theoretical framework.[5] Theorem 1 is akin to the equivalence result originally derived by Afriat (1967) for the utility maximisation model with a static, deterministic utility function; imposing $\boldsymbol{\epsilon}_i = \mathbf{0}$ for all $i = 1, ..., N$ returns the standard Afriat inequalities.

[5]This result is the multi-good extension of Adams et al (2014).

Theorem 1. If one can find sets $\{u_i\}_{i=1,...,N}$, $\{\boldsymbol{\epsilon}_i\}_{i=1,...,N}$ and $\{\lambda_i\}_{i=1,...,N}$ with $u_i \in \mathbb{R}$, $\lambda_i \in \mathbb{R}_{++}$ and $\boldsymbol{\epsilon}_i \in \mathbb{R}^{K-1}$, such that:

$$u_i - u_j < \lambda_j \mathbf{p}_j'(\mathbf{q}_i - \mathbf{q}_j) - \boldsymbol{\epsilon}_j'(\mathbf{q}_i^{\neg K} - \mathbf{q}_j^{\neg K}) \tag{A1}$$

$$\epsilon_i^k < \lambda_i p_i^k \tag{A2}$$

$$\epsilon_i^k - \epsilon_j^k < \lambda_i p_i^k \tag{A3}$$

$$\frac{1}{N}\sum_{i=1}^{N} \epsilon_i^k = 0 \tag{A4}$$

for all $i, j = \{1, ..., N\}$, and, for all $k = \{1, ..., K-1\}$, then a random sample of observed choice behaviour $\{\mathbf{p}_i, \mathbf{q}_i\}_{i=1,...,N}$ is consistent with the maximisation of $u(\mathbf{q}) + \boldsymbol{\epsilon}'\mathbf{q}^{\neg K}$.

Proof. See Appendix A.

The unknowns that define the revealed preference inequalities of Theorem 1 have natural interpretations. The numbers $\{u_i\}_{i=1,...,N}$ and $\{\lambda_i\}_{i=1,...,N}$ can be interpreted respectively as measures of the utility level that is dictated by the base utility function, $u_i = u(\mathbf{q}_i)$, and of the marginal utility of income at observed demands. Given the normalisation $p^K = 1$ and the restriction that $\epsilon^K = 0$, it is the case that $\lambda_i = u_K(\mathbf{q})$. ϵ_i^k is to be interpreted as the marginal utility perturbation to good-k relative to that dictated by base utility for consumer i. (A1) follows from the strict concavity of $U(\mathbf{q}, \boldsymbol{\epsilon})$ and optimising behaviour, while (A2) and (A3) impose strict monotonicity of the utility function given all recovered $\boldsymbol{\epsilon}$. (A4) imposes that the sample average of each dimension of unobserved heterogeneity is zero, in accordance with our model.

Each feasible solution to Theorem 1 can be used to construct a rationalising sample distribution function for $\boldsymbol{\epsilon}$ and a base utility function. Let a feasible solution to Theorem 1 be referred to as $\{\hat{u}_i, \hat{\boldsymbol{\epsilon}}_i, \hat{\lambda}_i\}_{i=1,...,N}$. The empirical distribution function of $\boldsymbol{\epsilon}$ associated with this set can be constructed as:

$$\hat{F}_\epsilon(\boldsymbol{\epsilon}) = \frac{1}{N}\sum_{i=1,...,N}^{N} 1\left[\hat{\boldsymbol{\epsilon}}_i \leq \boldsymbol{\epsilon}\right]$$

The proof of Theorem 1 is constructive and provides a method for building a candidate base utility function from the solution set to the revealed preference inequalities (see Appendix A). One specification for a rationalising base utility function is:

$$\hat{u}(\mathbf{q}) = \min_i \{\phi_i(\mathbf{q})\} \tag{9}$$

where for each $i = 1, .., N$, ϕ_i is defined as:

$$\phi_i(\mathbf{q}) \equiv \hat{u}_i + \hat{\lambda}_i \mathbf{p}_i'(\mathbf{q} - \mathbf{q}_i) - \hat{\boldsymbol{\epsilon}}_i'(\mathbf{q}^{\neg K} - \mathbf{q}_i^{\neg K}) - \delta g(\mathbf{q} - \mathbf{q}_i) \tag{10}$$

where g is defined as:

$$g(\mathbf{q}) = \sqrt{(q^1)^2 + ... + (q^K)^2 + T} - \sqrt{T} \tag{11}$$

with $T > 0$. Since g is strictly convex, ϕ_i is strictly concave.

4 Positive and Normative Analysis using Virtual Budgets

Just as one can test the model using data across consumers, it is also possible to conduct positive and normative analysis for particular individuals (i.e. draws of ϵ from the joint distribution of unobserved heterogeneity) in the sample using the observed choice behaviour of the full cross section. In this section, I address how demand responses and welfare effects can be bounded at new budgets of interest given the theoretical framework. To introduce the methodology, it is assumed that the draw of $\boldsymbol{\epsilon}$ characterising each consumer is known. I return to address the recovery of $\boldsymbol{\epsilon}$ in Section 5.

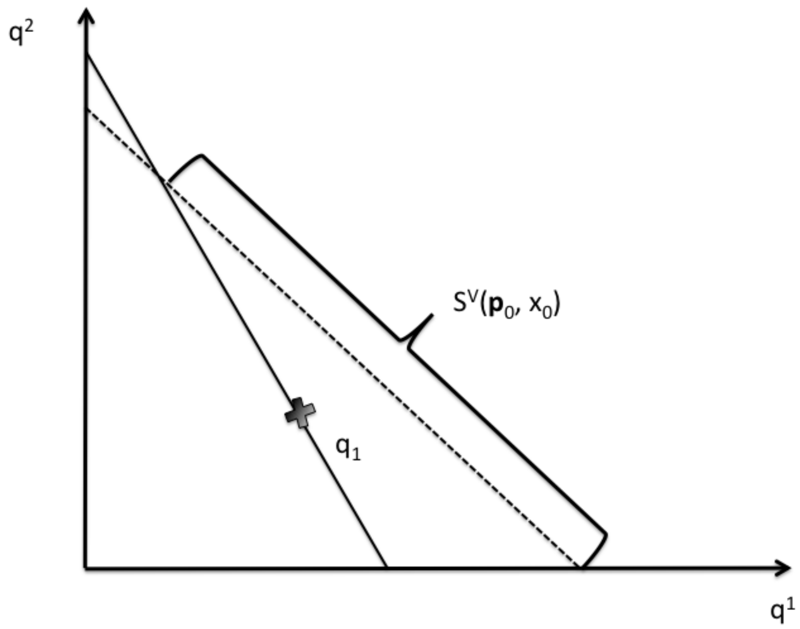
4.1 Demand Prediction

The revealed preference approach to demand prediction recovers the demands at a new budget of interest that are consistent with one's theoretical model given previously observed quantities generated by that model. With finite data, this approach typically identifies a *set* of potential demand responses at a new budget $\{\mathbf{p}_0, x_0\}$.

4.1.1 Traditional approach

Varian (1982) provides a thorough account of how to extrapolate demand behaviour to new budgets of interest for a static utility function, $u(\mathbf{q})$. Conditional on a panel of an individual consumer's consumption behaviour, $\{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,...,T}$, the demand response at $\{\mathbf{p}_0, x_0\}$ will be an element of the

Figure 2: Varian Bound



support set, $S(\mathbf{p}_0, x_0)^V$:

$$S(\mathbf{p}_0, x_0)^V = \left\{ \mathbf{q}_0 \quad : \quad \begin{array}{l} \mathbf{p}_0'\mathbf{q}_0 = x_0 \\ \mathbf{q}_0 \geq \mathbf{0} \\ \{\mathbf{p}_t, \mathbf{q}_t\}_{t=0,\ldots,T} \text{ satisfies SARP} \end{array} \right. \tag{12}$$

where SARP refers to the "Strong Axiom of Revealed Preference".

Definition. Strong Axiom of Revealed Preference (SARP)

$$\mathbf{q}_t \mathbb{R} \mathbf{q}_s \qquad \text{implies} \qquad \mathbf{p}_s'\mathbf{q}_s < \mathbf{p}_s'\mathbf{q}_t$$

where $\mathbb{R}$ represents the revealed preference relation.

We cannot directly apply Varian's methodology because, rather than analysing choices generated by the maximisation of a single utility function, we observe choices consistent with the maximisation of the sample distribution of preferences. We could apply the Varian approach to the data (i.e. the observation) that we have on a particular individual but, as shown in Figure 2, this is unlikely to yield informative bounds. In Figure 2, the observation on consumer 1, $\mathbf{q}_1$, does not tightly constrain responses at $\{\mathbf{p}_0, x_0\}$. Of course, if we had panel data on a consumer, and thus on choice behaviour generated by a particular value of $\boldsymbol{\epsilon}$ (i.e. data generated by the same utility function $u(\mathbf{q}) + \boldsymbol{\epsilon}'\mathbf{q}^{\neg K}$), $\{\mathbf{p}_t, \mathbf{q}_{t,\boldsymbol{\epsilon}}\}_{t=1,\ldots,T}$, then the support set associated with $\boldsymbol{\epsilon}$ could be defined as:

$$S(\mathbf{p}_0, x_0, \boldsymbol{\epsilon})^V = \left\{ \mathbf{q}_{0,\epsilon} \quad : \quad \begin{array}{l} \mathbf{p}_0'\mathbf{q}_{0,\epsilon} = x_0 \\ \mathbf{q}_{0,\epsilon} \geq \mathbf{0} \\ \{\mathbf{p}_t, \mathbf{q}_{t,\epsilon}\}_{t=0,\ldots,T} \text{ satisfies SARP} \end{array} \right. \tag{13}$$

and applied researchers could proceed as in Varian (1982). Blundell, Kristensen and Matzkin (BKM, 2014) are able to apply this methodology to construct revealed preference support sets because, in their two-good setting with the assumptions that they employ, consumers characterised by the same draw of unobserved heterogeneity will occupy the same quantile at each budget. Thus, the BKM support set can be constructed using quantile demands. Extending their approach to the multivariate setting is non-trivial due to the lack of an objective basis for ordering multivariate observations, and thus the difficulty of extending the notion of quantiles to a multidimensional setting.

4.1.2 Using Virtual Budgets

In this paper, I assume that only cross section data is available and that the dimensionality of the demand system does not facilitate the application of BKM (2014)- a different approach is required if we are to gain informative bounds on demand behaviour. I here address how the choices of a cross section can be used to bound the demand responses of a particular individual once observed prices are transformed into virtual prices.

The concept of a virtual budget was first suggested by Rothbarth (1945) and applied by Neary and Roberts (1980) to develop the theory of choice behaviour under rationing. In our setting, the $\boldsymbol{\epsilon}_0$-virtual budget of consumer i, $\{\tilde{\mathbf{p}}_{i,\epsilon_0}, \tilde{x}_{i,\epsilon_0}\}$, is that which induces consumer i to demand the same bundle as the consumer with the utility function $U(\mathbf{q}, \boldsymbol{\epsilon}_0) = u(\mathbf{q}) + \boldsymbol{\epsilon}_0'\mathbf{q}$ when facing market prices $\mathbf{p}$ with income x.

Looking to the first order conditions for consumer i, we have that at their observed demand $\{\mathbf{p}_i, \mathbf{q}_i\}$:[6]

$$u_k(\mathbf{q}_i) + \epsilon_i^k \quad = \quad \lambda_i p_i^k \tag{14}$$

[6] Interior solutions are guaranteed by smoothness of $U(\mathbf{q}, \boldsymbol{\epsilon})$. If corner solutions are admitted, the revealed preference restrictions are left unchanged but virtual prices are not uniquely determined at $q^k = 0$. One could then choose to work with the lowest rationalising virtual price vector.

Adding ϵ_0^k to each side and rearranging gives

$$u_k(\mathbf{q}_i) + \epsilon_0^k = \lambda_i \left(p_i^k + \frac{\epsilon_0^k - \epsilon_i^k}{\lambda_i} \right) \tag{15}$$

$$= \lambda_i \tilde{p}_{i,\epsilon_0}^k \tag{16}$$

Restrictions (A2) and (A3) of Theorem 1 imposes that $\tilde{p}_{i,\epsilon_0}^k > 0$ for any $\epsilon_0^k \geq \min_i \epsilon_i^k$.

The structure of the demand function therefore respects:

$$\mathbf{d}(\tilde{\mathbf{p}}_{i,\epsilon_0}, \tilde{x}_{i,\epsilon_0}, \boldsymbol{\epsilon}_0) = \mathbf{q}_i \tag{17}$$

where

$$\tilde{p}_{i,\epsilon_0}^k = p_i^k + \frac{(\epsilon_0^k - \epsilon_i^k)}{\lambda_i} \tag{18}$$

$$\tilde{x}_{i,\epsilon_0} = \tilde{\mathbf{p}}_{i,\epsilon_0}' \mathbf{q}_i \tag{19}$$

Responses for each individual in the cross section, or for any draw from the joint distribution of unobserved heterogeneity, can be bounded by constructing the support set using these virtual budgets rather than observed price data. Any demand response at $\{\mathbf{p}_0, x_0\}$ must be consistent with the known structure of the demand function. Thus, the "virtual price support set" at $\{\mathbf{p}_0, x_0\}$ for the individual with fixed unobserved heterogeneity $\boldsymbol{\epsilon}$ is given as:

$$S(\mathbf{p}_0, x_0, \boldsymbol{\epsilon}_0)^{VP} = \left\{ \mathbf{q}_{0,\epsilon_0} : \begin{array}{l} \mathbf{p}_0' \mathbf{q}_{0,\epsilon_0} = x_0 \\ \mathbf{q}_{0,\epsilon_0} \geq \mathbf{0} \\ \{\mathbf{p}_0; \mathbf{q}_{0,\epsilon_0}\} \cup \{\tilde{\mathbf{p}}_{i,\epsilon_0}; \mathbf{q}_i\}_{i=1,\ldots,N} \\ \quad \text{satisfies SARP} \end{array} \right. \tag{20}$$

Figure 3 demonstrates the method graphically, drawing upon the insights of Blundell, Browning and Crawford (2008). Figure 3 (a) shows the indifference curves associated with two different draws of ϵ that go through the demand bundles $[2, 5]$ and $[5, 2.5]$. The indifference curves correspond to the utility functions:

1. Blue preferences ($\epsilon_H^1 = 0.1$): $U(\mathbf{q}, \epsilon_H) = 0.5 \log(q^1) + 0.5 \log(q^2) + 0.1 q^1$

2. Red preferences ($\epsilon_L^1 = -0.1$): $U(\mathbf{q}, \epsilon_L) = 0.5 \log(q^1) + 0.5 \log(q^2) - 0.1 q^1$

Panel (b) displays the virtual budgets supporting the quantity bundles for the two different draws of ϵ.

The virtual relative prices for the 'good-1 loving' consumer (blue) are higher than those for ϵ_L. This is because, faced with the same virtual budget, the consumer endowed with ϵ_H will always choose to consume more of good-1 than the consumer endowed with ϵ_L.

Figure 3 (c) gives the income expansion paths going through the two demand bundles for the two draws of ϵ at the virtual prices which support these choices. Blundell, Browning and Crawford (2008) show that the intersection of the income expansion paths with the new budget of interest define the best bounds on demands (for $K = 2$) given the information available.[7] Thus, the support set associated with ϵ_L is shown by $S(\mathbf{p}_0, x_0, \epsilon_L)$ and the support set associated with ϵ_H is defined analogously. Different draws of ϵ are thus associated with different predicted sets.

4.2 Welfare Analysis

Similar arguments can be made to bound the welfare effects of price and income changes. Rather than apply revealed preference techniques to observed price-quantity combinations, one is able to use virtual price-quantity combinations to bound welfare metrics for hypothetical price and income changes for each individual (i.e. draw from the joint distribution of heterogeneity) in a cross section. For example, the compensating variation of a price change from $\mathbf{p}_i$ to $\mathbf{p}_0$ for the individual with $\boldsymbol{\epsilon} = \boldsymbol{\epsilon}_i$ is defined as:

$$CV = e(\mathbf{p}_0, \mathbf{q}_0; \boldsymbol{\epsilon}_i) - e(\mathbf{p}_0, \mathbf{q}_i; \boldsymbol{\epsilon}_i) \tag{21}$$

$$= x_0 - e(\mathbf{p}_0, \mathbf{q}_i; \boldsymbol{\epsilon}_i) \tag{22}$$

where $e(\mathbf{p}, \mathbf{q}; \boldsymbol{\epsilon})$ is the expenditure required to attain the utility associated with bundle $\mathbf{q}$ given prices $\mathbf{p}$ and preferences indexed by $\boldsymbol{\epsilon}$.

With panel data on an individual consumer, Varian (1982) provides algorithms for computing upper and lower bounds on $e(\mathbf{p}, \mathbf{q}; \boldsymbol{\epsilon})$. I adapt these algorithms to accept virtual prices rather than observed prices as inputs. This allows unique, theory consistent welfare metrics to be computed for each individual in the cross section given $\boldsymbol{\epsilon}_0$. For example, an upper bound on $e(\mathbf{p}_0, \mathbf{q}_i; \boldsymbol{\epsilon}_i)$ is calculated as follows:

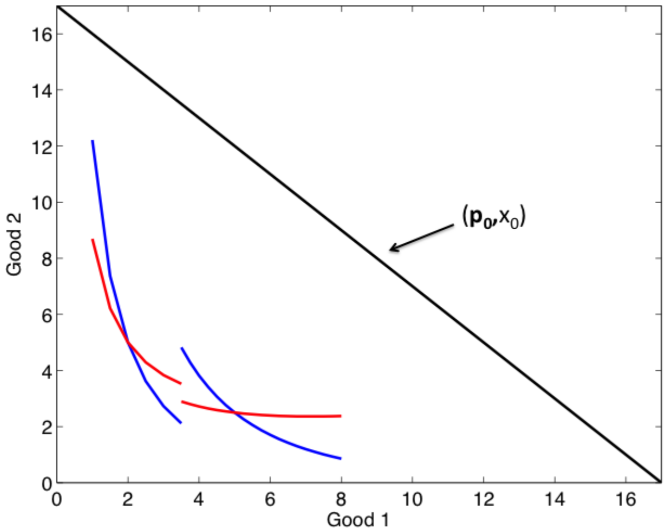
$$e^+(\mathbf{p}_0, \mathbf{q}_i; \boldsymbol{\epsilon}_i) = \min_{\mathbf{q}_j} \mathbf{p}_0' \mathbf{q}_j \tag{23}$$

$$\text{such that:} \quad \mathbf{q}_j \widetilde{\mathbb{R}}_{\epsilon_0} \mathbf{q}_i \tag{24}$$

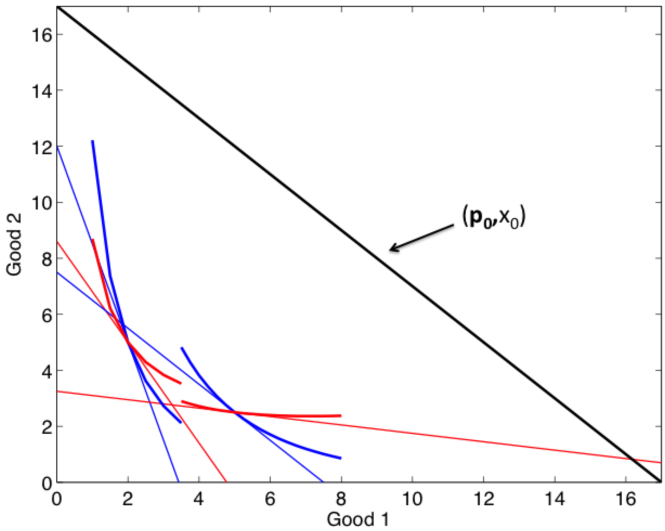
where a bundle $\mathbf{q}_j$ is revealed preferred to $\mathbf{q}_i$ for preference $\boldsymbol{\epsilon}_i$, $\mathbf{q}_j \widetilde{\mathbb{R}}_{\epsilon_i} \mathbf{q}_i$, if for any sequence of observations

[7] For $K > 2$, transitivity can be exploited to refine the support set further (Blundell et al, 2014).

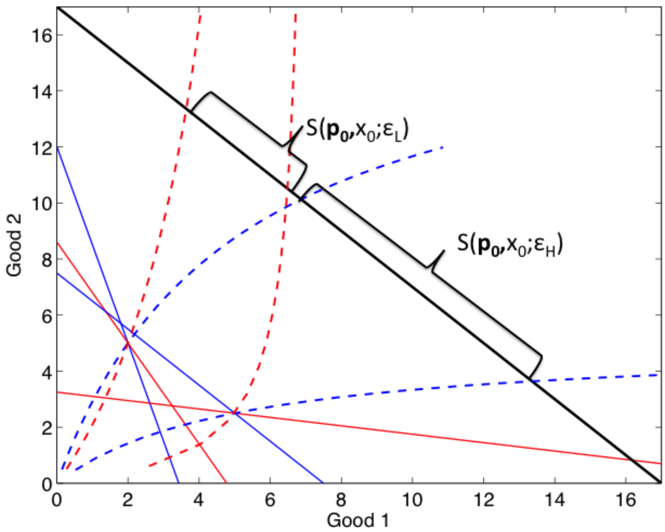
Figure 3: Bounding Demands



(a) Past Choices & Indifference Curves



(b) Virtual Budgets



(b) Engel Curves & Support Sets

j, k, l, m:

$$\tilde{\mathbf{p}}'_{j,\epsilon_i}\mathbf{q}_j \geq \tilde{\mathbf{p}}'_{j,\epsilon_i}\mathbf{q}_k, \tilde{\mathbf{p}}'_{k,\epsilon_i}\mathbf{q}_k \geq \tilde{\mathbf{p}}'_{k,\epsilon_i}\mathbf{q}_l, ..., \tilde{\mathbf{p}}'_{m,\epsilon_i}\mathbf{q}_m \geq \tilde{\mathbf{p}}'_{m,\epsilon_i}\mathbf{q}_i \tag{25}$$

Thus, where revealed preference relations are defined with respect to virtual prices rather than observed prices.

5 Identification Conditions

In Section 4, it was assumed that $\{\boldsymbol{\epsilon}_i, \lambda_i\}_{i=1,...,N}$ were known to the econometrician, allowing unique virtual prices to be constructed for each individual in the data set. If one imposes weak restrictions on the utility function and requires independence of $\boldsymbol{\epsilon}$ and budget parameters, then there is indeed a unique mapping between observables and the unobservables of the model given the population joint distribution of observables.[8] By then adapting the minimum distance from independence estimation technique of Manski (1983), I am able to recover a sample distribution of heterogeneity and values of the utility function at observed quantities that are strongly consistent for the true functions and can be used to construct virtual prices as required by Section 4.[9]

5.1 Nonparametric Identification

The question of identification concerns the mapping between a model's observable and unobservable features. A feature of the model is said to be identified if there exist no alternative specifications of the model that are 'observationally equivalent' to the true feature (see Matzkin (2007) for a comprehensive treatment of the topic). Identification of $F_{\boldsymbol{\epsilon}}$ requires that a unique $\boldsymbol{\epsilon}$ can be identified with each demand bundle given $\mathbf{p}$ and x.

As the model currently stands, F_ϵ is not identified nor does it have bounded support. This means that even if one had access to the population joint distribution of quantities, prices and income, there would not be a unique rationalising base utility function and F_ϵ, nor would it be possible to bound ϵ_i^k. For notational ease, let ξ represent a $(K-1)N$ vector of stacked $\boldsymbol{\epsilon}_i$ vectors and let Ξ give the set of all ξ that are able to rationalise the data set $\{\mathbf{p}_i, \mathbf{q}_i\}_{i=1,...,N}$ (i.e. for which one can find u_i and λ_i such that Theorem 1 is satisfied). Without further restrictions on preferences, if Ξ is nonempty, then it is unbounded.

[8]This is a similar result to Corollary 1 of Brown and Matzkin (1998), although proved without recourse to the results of Brown (1983) and Roehrig (1988).

[9]Note that this is still not sufficient to uniquely identify demands, or the properties of the demand function, at new budget regimes.

Theorem 2. If a non-empty feasible set to the inequalities of Theorem 1 exists, then it is unbounded.

Proof. See Appendix A.

Imposing the set of conditions specified by Theorem 3 is sufficient for there to exist a unique specification of F_ϵ that is consistent with the population distribution of data. This is related to Corollary 1 of Brown and Matzkin (1998) but is proved without a reliance on the results of Brown (1983) and Roherig (1988).[10] The distribution of budgets supporting a particular quantity bundle can be used to identify F_ϵ given $E(\boldsymbol{\epsilon}) = \mathbf{0}$, independence of taste and budget parameters and the restriction of the marginal utility of the K^{th} good to a known, bounded function. Given these assumptions, one is able to disentangle the influence of preference heterogeneity from the marginal utility of income, λ, in producing variation in the budgets that support a particular quantity bundle.

Theorem 3. Identification of $(F^\star_\epsilon, \nabla_{\mathbf{q}} u^\star(\mathbf{q})$

Assume that the vector $(\mathbf{p}, x)$ has a continuous Lebesgue density and that the joint distribution of observables $F_{Q,P,X}$ is identified. Let $\bar{\mathbf{q}} \in Q$ and $\alpha \in R$ be given. Suppose that W is a set of smooth utility functions $u : Q \to R$ such that $\forall u \in W$, $u(\bar{\mathbf{q}}) = \alpha$. Let Ω denote the set of functions $\nabla_{\mathbf{q}} u(\mathbf{q})$ where $u \in W$. Denote by Γ the set of absolutely continuous distribution functions of vectors $(\epsilon^1, ..., \epsilon^{K-1})$ that satisfy $E(\boldsymbol{\epsilon}) = \mathbf{0}$. Then, $(\nabla_{\mathbf{q}} u(\mathbf{q}), F(\boldsymbol{\epsilon}))$ is identified in $(\Omega \times \Gamma)$ if the following conditions are imposed:

1. Independence of $\boldsymbol{\epsilon}$ and $(\mathbf{p}, x)$.
2. $u_K(\mathbf{q}) = f(\mathbf{q})$, with $f(\mathbf{q}) > 0$ and $f(\mathbf{q}) < \infty$.

Proof. See Appendix A.

Restricting the base utility function to be quasilinear in the K^{th} good is a simple way of restricting the marginal utility of the K^{th} good, i.e.

$$U(\mathbf{q}, \boldsymbol{\epsilon}) = v(\mathbf{q}^{\neg K}) + q^K + \sum_{k=1}^{K-1} \epsilon^k q^k \tag{26}$$

which gives $u_K(\mathbf{q}) = 1$. That this assumption can be easily imposed on demand predictions, is one benefit of the specification. However, more complicated functions for $u_K(\mathbf{q})$ can be dealt with provided

[10] Theorem 3, and by extension Corollary 1, of Brown and Matzkin (1998) relied on earlier results of Brown (1983) and Roherig (1988) that were later shown not to be sufficient to guarantee identification (Benkard and Berry, 2004).

that they are *known*.

5.2 Imposing Identification Restrictions

Imposing the identification conditions restricts the set of feasible solutions to the revealed preference restrictions and enables easy construction of virtual prices to facilitate positive and normative analysis. To impose the restrictions of Theorem 3 on the base utility function, the revealed preference inequalities of Theorem 1 are modified to:

$$u_i - u_j < f(\mathbf{q}_j)\mathbf{p}_j'(\mathbf{q}_i - \mathbf{q}_j) - \boldsymbol{\epsilon}_j'(\mathbf{q}_i^{\neg K} - \mathbf{q}_j^{\neg K}) \quad (27)$$

$$\epsilon_i^k < f(\mathbf{q}_i)p_i^k \quad (28)$$

$$\epsilon_i^k - \epsilon_j^k < f(\mathbf{q}_i)p_i^k \quad (29)$$

$$\frac{1}{N}\sum_{i=1}^{N}\epsilon_i^k = 0 \quad (30)$$

with $u_1 = \alpha$, and where $u_K(\mathbf{q}) = f(\mathbf{q})$. Note that these inequalities remain linear in unknowns and are thus easily implemented using standard linear programming techniques.

Independence of $\boldsymbol{\epsilon}$ and $(\mathbf{p}, x)$ is imposed using a nonparametric version of Manski's (1983) Closest Empirical Distribution method. For reasons of practical tractability, I adapt Brown and Matzkin's (1998) estimator to minimise the supremum norm between the joint distribution of $\boldsymbol{\epsilon}$ and $(\mathbf{p}, x)$ and the multiplication of the marginals, rather than, as they suggest, minimise the bounded Lipschitz metric between the distributions.

Formally, the estimates $\hat{\xi} = \{\hat{\boldsymbol{\epsilon}}_i\}_{i=1,..,N}$ and $\{\hat{u}_i\}_{i=1,..,N}$ are selected as the solution to the following optimisation problem:

$$\min_{u,\epsilon} \sup |F_{\epsilon,X,N}(\boldsymbol{\epsilon}, \mathbf{X}) - F_{\epsilon,N}(\boldsymbol{\epsilon})F_{X,N}(\mathbf{X})| \quad (31)$$

subject to:

$$u_i - u_j < f(\mathbf{q}_j)\mathbf{p}_j'(\mathbf{q}_i - \mathbf{q}_j) - \boldsymbol{\epsilon}_j'(\mathbf{q}_i^{\neg K} - \mathbf{q}_j^{\neg K}) \quad (32)$$

$$\epsilon_i^k < f(\mathbf{q}_i)p_i^k \quad (33)$$

$$\epsilon_i^k - \epsilon_j^k < f(\mathbf{q}_i)p_i^k \quad (34)$$

$$\frac{1}{N}\sum_{i=1}^{N}\epsilon_i^k = 0 \quad (35)$$

with $u(\mathbf{q}_1) = \alpha$ and where

$$F_{\epsilon,X,N}(\boldsymbol{\epsilon}, \mathbf{X}) = \frac{1}{N}\sum_{i=1}^{N} 1\left[\boldsymbol{\epsilon}_i^k \leq \boldsymbol{\epsilon}, \mathbf{X}_i \leq \mathbf{X}\right] \tag{36}$$

$$F_{\epsilon,N}(\boldsymbol{\epsilon}) = \frac{1}{N}\sum_{i=1}^{N} 1\left[\boldsymbol{\epsilon}_i^k \leq \boldsymbol{\epsilon}\right] \tag{37}$$

$$F_{X,N}(\mathbf{X}) = \frac{1}{N}\sum_{i=1}^{N} 1\left[\mathbf{X}_i \leq \mathbf{X}\right] \tag{38}$$

The estimator of F_ϵ, $\hat{F}_{\epsilon,N}$ is then:

$$\hat{F}_{\epsilon,N}(\boldsymbol{\epsilon}) = \frac{1}{N}\sum_{i=1}^{N} 1\left[\hat{\boldsymbol{\epsilon}}_i^k \leq \boldsymbol{\epsilon}\right]$$

With the assumptions that F_ϵ possesses absolutely continuous marginal distributions, and that $u(\mathbf{q})$ and its derivatives up to the second order are equicontinuous and uniformly bounded, then $\hat{F}_{\epsilon,N}(\boldsymbol{\epsilon})$ is a strongly consistent estimator for F_ϵ. $\{\hat{\boldsymbol{\epsilon}}_i\}_{i=1,..,N}$ and $f(\mathbf{q})$ can then be used to construct the virtual prices required to bound demands and welfare effects using the methodology outlined in Section 4.[11]

Theorem 4. Strong Consistency of $\hat{F}_{\epsilon,N}(\boldsymbol{\epsilon})$ for F_ϵ
Let the conditions for identification hold, that $u(\mathbf{q})$ and its derivatives up to the second order are equicontinuous and uniformly bounded and that F_ϵ possesses absolutely continuous marginal distributions. Then $\hat{F}_{\epsilon,N}(\boldsymbol{\epsilon})$ is a strongly consistent estimator for F_ϵ.

Proof. See Appendix A.

It is important to note that while the distribution of unobserved preference heterogeneity and the derivatives of the base utility function are identified in the population, demands (and thus features of the demand function, e.g. price elasticities), are not uniquely recovered in a finite data setting. When one only observes a finite number of demands and budget environments, features of the base utility function are incompletely recovered. Therefore, when applied in practise, one continues to recover *sets* of demands at new budgets of interest.

[11] In the demand prediction procedure, one will now additionally impose the restriction that $u_K(\mathbf{q}) = f(\mathbf{q})$.

Summary In the preceding sections, a revealed preference methodology was developed that allows unobserved preference parameters to be recovered and used to predict demand and welfare effects at new budgets of interest. The set of linear inequalities implied by the theoretical framework was derived, and used to construct virtual prices that can be used for positive and normative analysis. I have shown how identifying restrictions and minimum distance for independence estimation techniques can be integrated into the framework to recover a sample rationalising $F_{\epsilon,N}$. Imposing absolute continuity on the marginal distributions of F_ϵ and requiring that $u(\mathbf{q})$ and its derivatives up to the second order are equicontinuous and uniformly bounded gives us that the estimator $F_{\epsilon,N}$ is consistent for F_ϵ.

6 Empirical Illustration

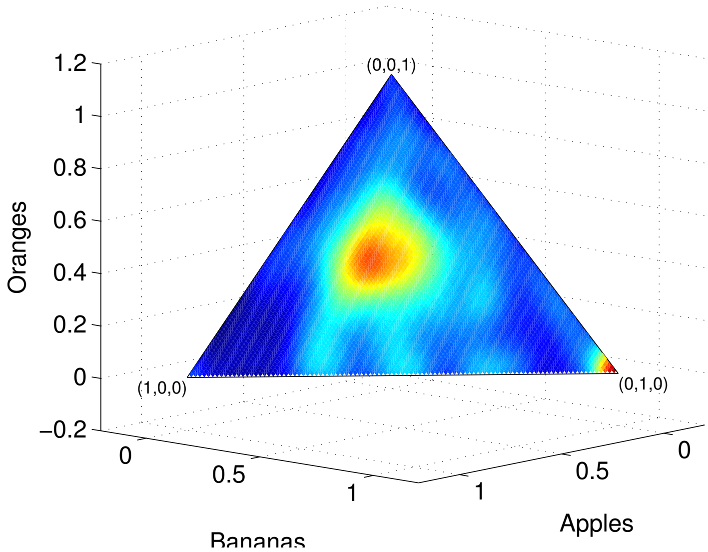
I now demonstrate the utility of the approach via an empirical application to consumer microdata. The data is drawn from the U.K. Kantar Worldpanel. The Worldpanel is one of the largest surveys of consumer behaviour in the world and contains information on domestic food and drink purchases. Participating households are issued with a barcode reader, with which they record the purchases of all barcoded products that are bought into the home. Therefore, all household scannable 'fast-moving consumer goods' are recorded.[12] Leicester (2012) estimates that approximately 20% of all household expenditures are covered by this data source.

The aim of this section is to demonstrate the workings of the methodology in a simple setting in which cross-sectional heterogeneity is a salient feature of the data. I focus on modelling consumer demand for fruit, considering choice over apples, bananas and oranges. This application is complicated enough to require multidimensional heterogeneity but remains simple enough for results to be easily graphically displayed. Although the assumption that these goods constitute a separable subset of the main utility function is restrictive, I find that a necessary condition for $u(apples, bananas, oranges)$ to form an additively separable subset cannot be rejected for the majority of the sample when a panel of household purchases are considered; there exists a well-behaved utility function defined just on consumer purchases of these goods for many households with stable demographic and employment characteristics.[13]

I analyse purchases carried out in the summer of 2011 and aggregate information to this level, partly to ease the computational burden of estimation and to allow for fruit to be treated as a non-durable and non-storable good.[14] Quantities are given by the (observed) kilograms of fruit purchased, while a price index is constructed from the corresponding unit price (total expenditure on a particular good

[12]Purchases from all retailers and online purchases are covered.
[13]Details available from the author on request.
[14]Intertemporal separability of the utility function is required by our empirical application.

Figure 4: Cross Section Budget Share Heterogeneity



divided by the weight of good purchased). The low level of aggregation of our individual commodities means that variation in cross-sectional prices are principally explained by differences in demand and supply across markets, rather than variation in unobserved product qualities and characteristics.[15] For those not consuming a particular good, I assume that they face the average price faced by consumers in their geographic region and social class that is observed in the data.[16]

6.1 Sample heterogeneity

Figure 4 gives the cross-section distribution of budget shares for apples, bananas and oranges that is observed in our sample, estimated using multivariate kernel methods.[17] The bottom left vertex of the simplex represents the choice to spend all of one's budget on apples, the bottom right vertex gives the choice to spend all of ones budget on bananas and the top vertex gives the choice to spend all of one's budget on oranges. Consumers are typically located around the barycenter of the budget share simplex, although there is a group of individuals who spend the majority of their fruit budget on bananas.

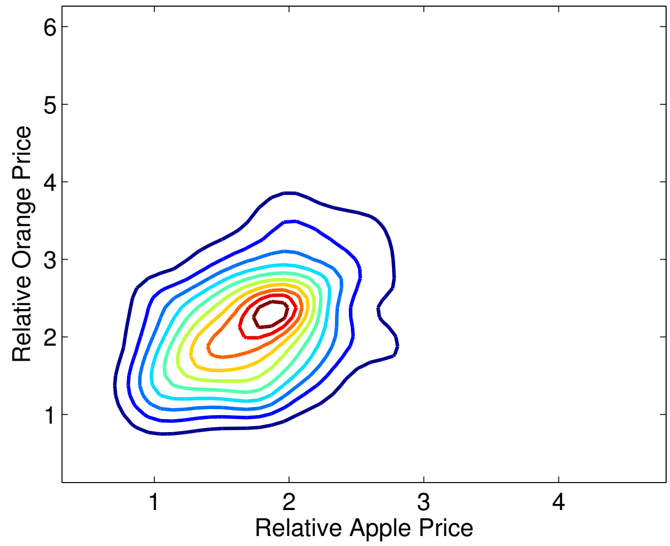
The goal of this section is to rationalise the cross-section variation in budget shares by appeal to variation in prices and total expenditure (differences in constraints) and to unobserved preference heterogeneity (differences in objective functions). Panels (a) and (b) of Figure 5 give the cross-section distribution of the relative price of apples and bananas and of total expenditure. The density functions were again estimated using kernel techniques. The x-axis of panel (a) gives the relative price of apples

[15] See Deaton (1988) for a discussion of the "unit-price" problem.

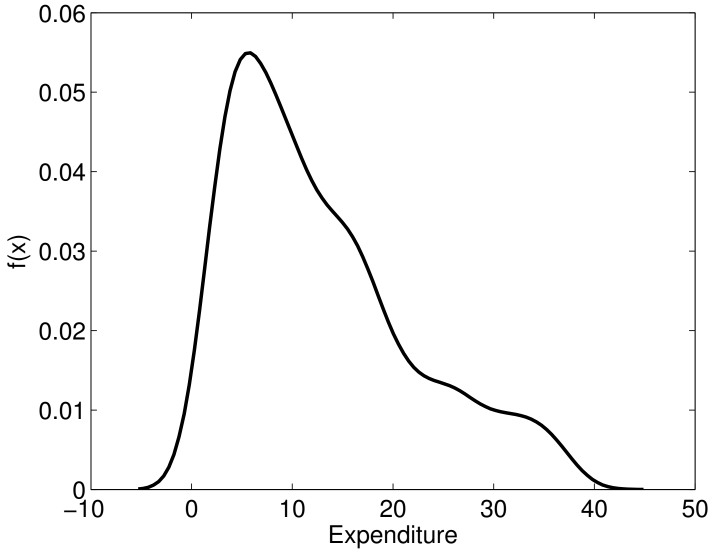
[16] A household's social class is declared as one of the NRS social grades AB, C1, C2, D, E, with a classification lower in the alphabet designating a higher social class.

[17] Multiplicative Gaussian kernel functions and rule-of-thumb bandwidths, as defined in Bowman and Azzalini (1997), were employed for nonparametric bivariate density estimation.

Figure 5: Price & Expenditure Heterogeneity



(a) Relative Prices



(b) Expenditure

and the y-axis gives the relative price of oranges. Bananas are cheaper than both apples and oranges, and one observes a positive correlation between the relative price of apples and oranges. There is also a right skew to the fruit expenditure distribution, as shown in panel (b).

Observed cross-section choice variation cannot be rationalised by variation in constraints alone. The full sample of choices violate the Generalised Axiom of Revealed Preference implying that there does not exist a single utility function that could have generated the data set.[18] Nor is just allowing for preference heterogeneity along observable dimensions sufficient. I partition the sample into observable cells defined on family structure, the presence of children and the education level of the household head.

[18]GARP is a necessary condition for SARP to hold.

Figure 6 gives the empirical cumulative distribution of group sizes resulting. This illustrates my earlier comment about the sparsity of high-dimensional data. Despite the rather broad partitioning of the data, the median number of observations per non-empty cell is only 42 and 90% of cells include fewer than 100 observations- the presumption of finite data is thus well-founded in this context.

Figure 6: Demographic Cell Size

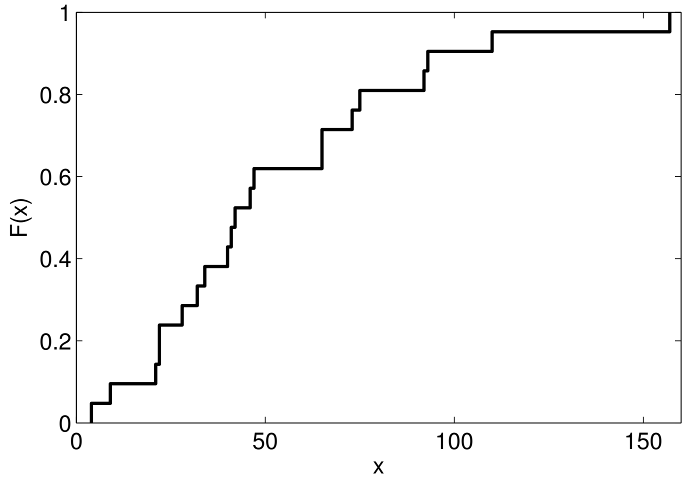


Table 1 gives the rationality results for each demographic cell. Every group with a group size of greater than 10 fails GARP. The Afriat Efficiency Index (Afriat (1967, 1972), Varian (1990, 1991)), a, which gives a measure of the size of a revealed preference violation, suggests that the deviations from a single utility function are large for many groups. The lower is a, the more that the budget constraint must be relaxed for the rationality restrictions to be satisfied. The mean Afriat Efficiency Index is 0.7656 implying that, on average, budget constraints must be relaxed by $\sim 25\%$ to achieve consistency with revealed preference. This suggests severe violations; Varian (1991), for example, suggested that one reject the hypothesis of the maximisation of a single utility function if $a < 0.95$.

6.2 Revealed Preference Heterogeneity

The failure of GARP within observable demographic cells implies that an appeal to unobserved heterogeneity is required to rationalise the behaviour of the cross-section. My framework is sufficient to capture the heterogeneity observed in this data set. All demographic cells can be rationalised by the theoretical framework; a non-empty solution set to Theorem 1 is returned for each group. For my recovery exercises, I restrict preferences to be quasilinear in oranges such that $u_K(\mathbf{q}) = 1$.

Table 2 gives the preference parameters associated with the minimum distance from independence (MDI) solution. The variance of preference parameters for apples within demographic groups is greater than that of bananas. However, there is no systematic covariance in tastes for apples and bananas

Table 1: Rationality by Demographic Cell

Family Type	Children	Education	N	GARP	Afriat Efficiency (a)
Single	No	Low	9	1	1.0000
		Mid	34	0	0.6787
		High	40	0	0.2469
Couple	No	Low	22	0	0.9413
		Mid	28	0	0.9046
		High	22	0	0.8530
Couple	Yes	Low	4	1	1.0000
		Mid	42	0	0.6587
		High	41	0	0.9324
Mulit-Adult	No	Low	93	0	0.7890
		Mid	157	0	0.6608
		High	110	0	0.5655
Multi-Adult	Yes	Low	21	0	0.9933
		Mid	73	0	0.5100
		High	65	0	0.7467
Pensioner Single	No	Low	65	0	0.6531
		Mid	46	0	0.7766
		High	32	0	0.8352
Pensioner Couple	No	Low	92	0	0.8248
		Mid	75	0	0.7011
		High	47	0	0.8066

evident; the sign varies across groups and the magnitude is typically small. Figure 7 shows the contours of the joint distribution of preference parameters for apples (ϵ^1) and bananas (ϵ^1) for multi-adult households who are mid-educated without children. For reference, the same contours of the bivariate normal with the maximum likelihood mean and covariance matrix. Recovered preference parameters are not normally distributed; individuals are more likely to have a strong preference for one good and a strong dislike of the other, than is reflected by the normal distribution.

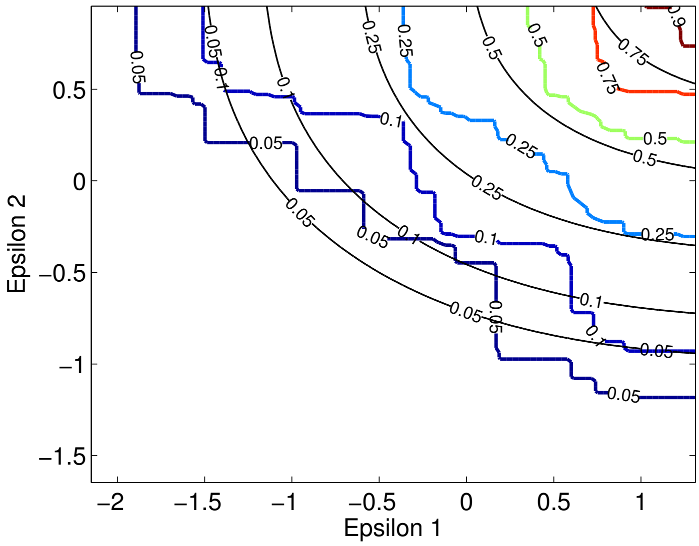
6.3 Demand Prediction

I now turn to the prediction of demand responses for the base utility function and for various draws from the distribution of unobserved heterogeneity. I begin by recovering the demands consistent with the base utility function, i.e. $u(\mathbf{q})$, of each demographic group at the mean price vector. This involves bounding demands with the set of virtual prices, $\tilde{\mathbf{p}}_i = \mathbf{p}_i - \hat{\boldsymbol{\epsilon}}$ for $i = \{1, ..., N\}$, with the restriction that $u_K(\mathbf{q}) = 1$.

Figure 8 gives the budget-share simplex at this new economic environment for a random selection of demographic groups.[19] Each coloured convex set corresponds to the rational demands that are

[19]The groups pictured are: single, no children, high education (blue); couple, children, low education (green); multi-adult, children, mid education (yellow); pensioner couple, low education (red).

Figure 7: Joint Distribution of Preference Parameters



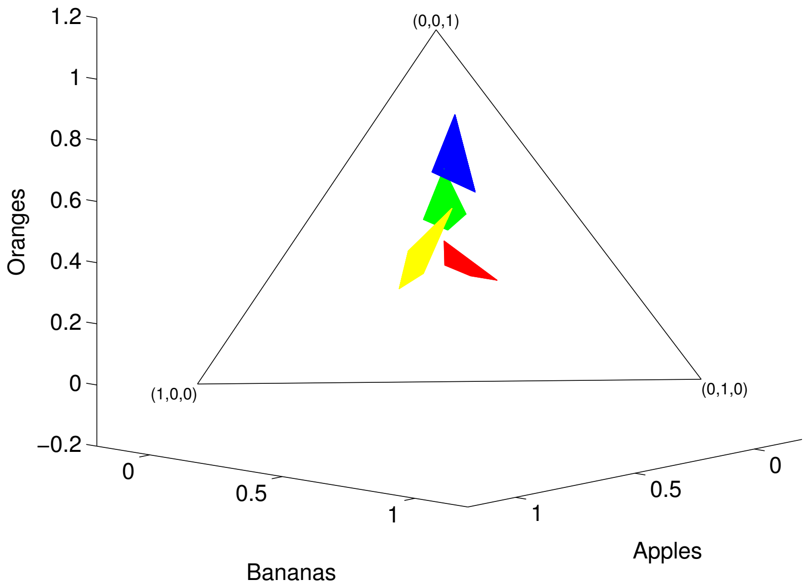
consistent with our theoretical model plus the distribution of unobserved heterogeneity recovered for that particular group. There is heterogeneity in the location and cardinality of the predicted sets, which is a product of preference heterogeneity across groups and variation in the intersections virtual budgets and the new budget of interest. It is interesting to note that the sets are informative for observed demands; it is not the case that "anything goes" with the method.

One can also recover demands at various conditional quantiles of interest within demographic cells. Conditional on a particular quantile of the marginal distribution of unobserved preference heterogeneity

Table 2: MDI Preference Parameters: Summary Statistics

Family Type	Children	Education	Var(Apples)	Var(Bananas)	Cov(Apples, Bananas)
Single	No	Mid	0.5329	0.5057	-0.0386
Single	No	High	0.7057	0.5177	-0.0340
Pensioner	No	Low	0.5061	0.4955	-0.0209
Pensioner	No	Mid	0.4982	0.5113	0.0184
Pensioner	No	High	0.5559	0.4907	-0.0533
Couple	No	Low	1.0898	0.8842	0.0267
Couple	No	Mid	0.7229	0.6765	0.0077
Couple	No	High	0.5549	0.537	-0.0455
Pen. Couple	No	Low	0.6719	0.5967	0.0249
Pen. Couple	No	Mid	0.7758	0.7511	0.0283
Pen. Couple	No	High	0.4380	0.4952	0.0724
Couple	Yes	Mid	0.8429	0.7277	0.0184
Couple	Yes	High	0.5221	0.5059	0.0460
Other	No	Low	0.5883	0.5146	0.0453
Other	Yes	Low	0.4655	0.4388	-0.0185
Other	Yes	Mid	0.5797	0.5097	-0.0145
Other	Yes	High	0.6925	0.6234	-0.1060

Figure 8: Differences in Base Utility Predictions of Demographic Groups at the Average Budget



for one good, $\bar{\tau}^k$, I recover sets of demands consistent with various conditional quantiles, τ, of the distribution of unobserved heterogeneity of the other good, i.e.

$$\epsilon^k_{0,\tau} = Q(\tau|\epsilon^{\neg k} = \bar{\tau}^{\neg k}) \tag{39}$$

Figure 9 shows various support sets at the average budget for, for example, the single pensioner, mid education group. Each coloured set gives a support set consistent with the median taste for apples and a different conditional quantile of the taste for bananas distribution, from 0.1 to 0.9 (labelled). Conditional on being endowed with the median taste for apples, a very low taste for bananas (blue set) is associated with a support set in the apex of the budget share simplex indicating that the overwhelming majority of one's budget is devoted to oranges, and few apples and bananas being consumed. Conversely a very high taste for bananas (red set), is associated with a large portion of one's budget being spent on that good, with smaller shares devoted to oranges and apples.

It is also possible to recover the demands for particular individuals in the sample. To recover the support set for individual j, one bounds demand responses at $\{\mathbf{p}_j, x_j\}$ given j's virtual prices, i.e. sample prices $\{\mathbf{p}_i\}_{i=1,..,N}$ are modified to $\{\tilde{\mathbf{p}}_{i,\epsilon_j}\}_{i=1,...,N}$ where $\tilde{\mathbf{p}}_{i,\epsilon_j} = \mathbf{p}_i - \hat{\boldsymbol{\epsilon}}_i + \hat{\boldsymbol{\epsilon}}_j$ for all individuals $i \neq j$ with the restriction that $u_K(\mathbf{q}) = 1$. Figure 10 shows the support sets associated with four individuals' observed demands, calculated using all quantities and virtual prices *other* than that individual's observed demand (otherwise the method would perfectly predict that individual's demand behaviour given the requirement of strict concavity of the utility function). I refer to these as 'leave-one-out' predictions. An individual's *observed* budget share is also plotted as a black point. As dictated by the methodology, all observed demands are elements of the predicted support sets (as, by definition, the past demands of individuals $i \neq j$ and the demand of individual j must satisfy rationality with the appropriate virtual prices). The cardinality of these sets varies because, as budgets and tastes vary across individuals, the intersections of virtual budgets and the new budgets of interest do not always lead to particularly informative support sets for an individual.

Using individual budget predictions, one is able to compare the distributions of 'leave-one-out' demand predictions and the actual distribution of demands. To do this, I select the barycenter of the uniform distribution over each individual's support set and compare the distributions of budget shares obtained in this manner to the observed budget shares. Intuitively, the barycenter represents the middle of the support set. This point is the optimal choice of demand within the support set if one wishes to minimise the squared loss of an incorrect forecast that is constrained by rationality. Further, this point can be justified by an appeal to the Principle of Maximum Entropy (Jaynes, 1957a; Jaynes, 1957b), which dictates that one select the object of interest consistent with the "maximally uncertain" distribution over outcomes, whilst respecting the knowledge that one has to hand. As Jaynes put it, of all distributions consistent with the data to hand, we should choose the one that is "maximally non-committal with regard to the missing information" (1957a, p.623). This corresponds to maximisation of the Shannon entropy measure subject to the constraints laid down by the data and the theory, i.e. the constraint that the solution must lie within the support set.

Figure 9: Demand at Various Quantiles of Distribution of Banana Tastes

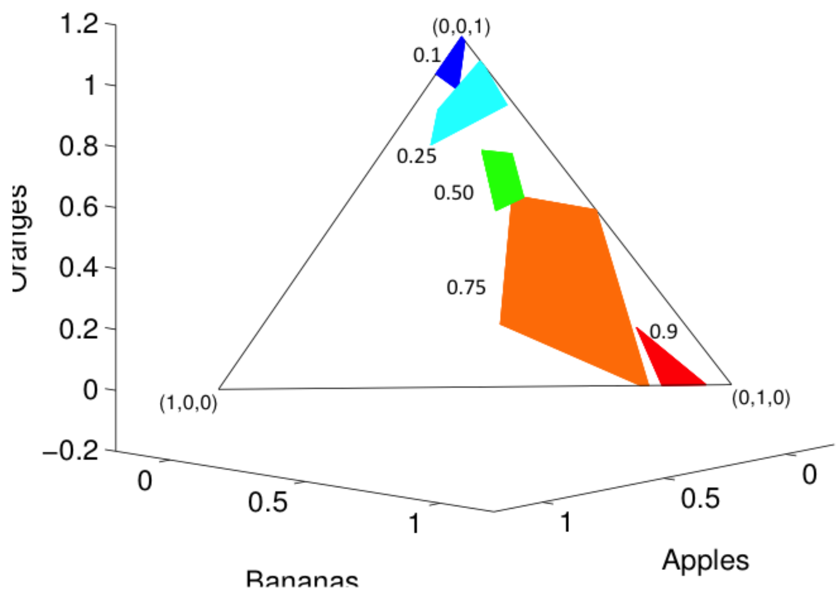
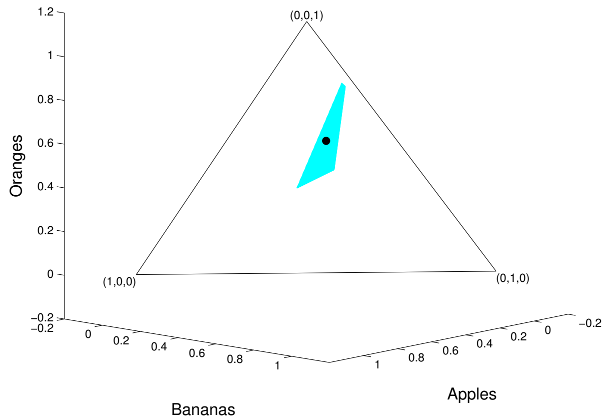
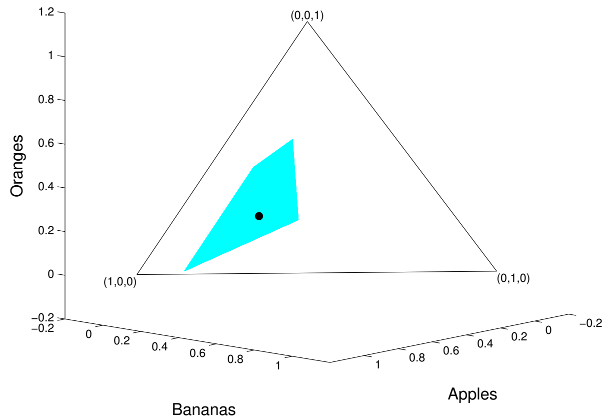


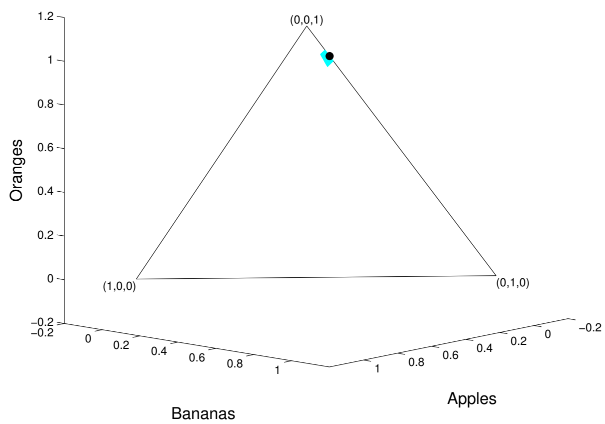
Figure 10: Support Set Predictions and Observed Demands: Single Pensioner, Mid Education



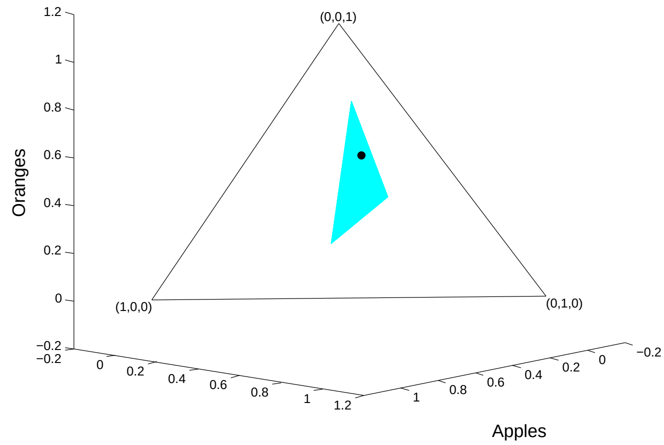
(a) Individual 4



(b) Individual 13



(c) Individual 37



(d) Individual 55

Figure 11 shows the observed and barycenter demand distributions for the full cross section. Kernel density estimates and histograms are both provided for ease of interpretation as most discrepancies occur at the edges of the support of budget shares. Visual inspection suggests that estimated and observed budget share distributions are similar, although the barycenter estimates are less able to return observed demands near corners. This is to be expected given that at least one vertex of an individual's support set will be at the interior of the budget set, and thus the barycenter will always be a strictly interior solution.

Table 3 gives the Mean Absolute Deviation (MAD) of recovered from observed budget shares as a measure of how close the barycenter of the revealed preference support sets are to observed budget shares for apples and bananas. These are reported for each demographic group and for the full cross section. They are calculated as:

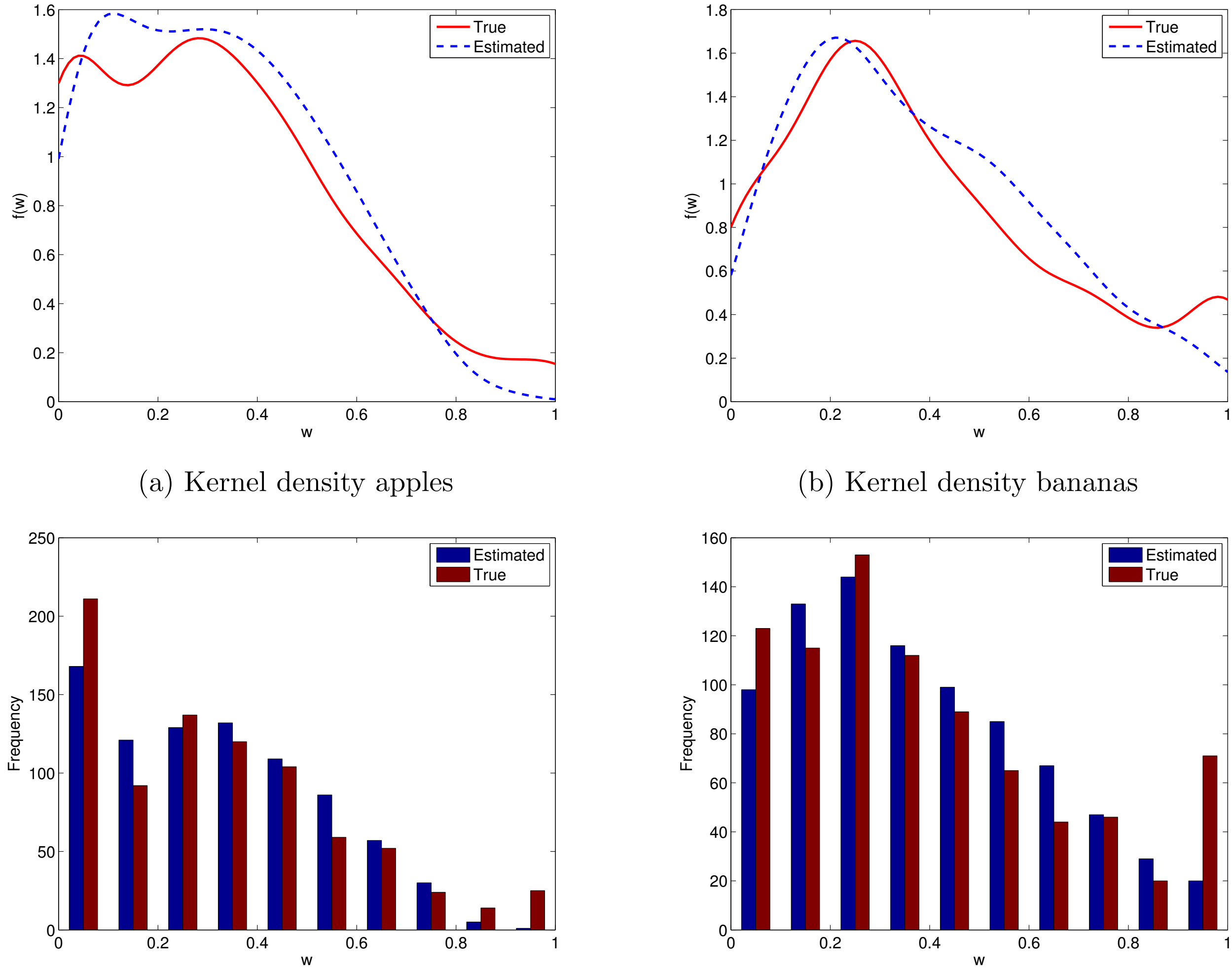
$$MAD = \frac{1}{N_g} \sum_{i=1}^{N_g} |w_i^k - \hat{w}_i^k| \tag{40}$$

Table 3: R^2 and Mean Absolute Deviation

Family Type	Children	Education	MAD Apples	MAD Bananas
Full Sample			0.0774	0.0782
Single	No	Mid	0.1176	0.1167
Single	No	High	0.0998	0.0997
Pensioner	No	Low	0.0845	0.0761
Pensioner	No	Mid	0.1114	0.0900
Pensioner	No	High	0.1056	0.1023
Couple	No	Low	0.1196	0.1384
Couple	No	Mid	0.0949	0.1348
Couple	No	High	0.1158	0.1060
Pen. Couple	No	Low	0.0576	0.0677
Pen. Couple	No	Mid	0.0765	0.0674
Pen. Couple	No	High	0.0707	0.0780
Couple	Yes	Mid	0.0637	0.0856
Couple	Yes	High	0.0776	0.0686
Other	No	Low	0.0566	0.0538
Other	Yes	Low	0.0832	0.0708
Other	Yes	Mid	0.0599	0.0594
Other	Yes	High	0.0527	0.0602

where N_g is the number of individuals in the demographic group/sample in question, w_i^k is the observed budget share of good k of individual i and $\hat{w}_i^k$ is the barycenter of the support set for that individual. The average discrepancy is low for the literature highlighting that calculated budget shares closely approximate observed budget shares providing strong evidence of the utility of my approach.

Figure 11: Calculated and Observed Cross Section Demands



(a) Kernel density apples

(b) Kernel density bananas

(c) Histogram apples

(d) Histogram bananas

7 Conclusion

In this paper I have developed a revealed preference approach for predicting demand responses and calculating welfare effects when preferences differ across consumers. Many empirical applications devote their attention to modelling the relationship between demands and other observed variables, leaving an additively separable error term to capture unobserved preference heterogeneity. This specification places strong restrictions on preferences, which many studies find are invalid. I relax many of these assumptions to develop a methodology that allows for a richer class of underlying individual utility functions. I assume that preference heterogeneity manifests itself in shifts to the marginal utility of commodities, which can then enter non-additively in the demand function. Further, unlike much of the current literature, I allow for multidimensional unobserved heterogeneity and therefore tackle demand systems composed of many goods. This is rare in the literature given the problems of simultaneity and the requirement of transitivity that accompany the move to a many-good demand system.

I derive the linear revealed preference restrictions associated with my theoretical framework and show how the feasible set to these inequalities can be used to construct virtual prices that enable demand and welfare effects to be bounded at new budgets of interest. To refine the set of solutions, I impose a set of identifying restrictions and apply the Closest Empirical Distribution estimation method of Manski (1983) to recover a rationalising sample distribution of unobserved preference parameters. I am able to show that this estimator is strongly consistent for the true distribution of unobserved heterogeneity.

I demonstrate the utility of the approach through an illustrative application to household scanner data. I find that cross-section choices, and the choices within demographic groups, cannot be rationalised by the maximisation of a single utility function. The violations of rationality are large in magnitude, suggesting that they are not a product of measurement error nor marginal optimisation error. Thus, an appeal to unobserved preference heterogeneity is necessary. I recover the unobserved preference parameters that rationalise the choices of each demographic group and bound demand predictions for changes in the prices of commodities. This serves to demonstrate two advantages of my approach. First, this method leads to informative and accurate predictions. By comparing the barycenter of individual support sets to their observed demands, I show that budget shares recovered by my methodology closely approximate observed budget shares, suggesing that the accuracy of the approach dominates many alternatives in the literature. Second, I am able to carry out this analysis on a demand system with many goods. This will allow me to address richer applications than those relying on quantile demands to address unobserved preference heterogeneity given the difficulty of extending quantiles to a multidimensional setting.

In my theoretical framework unobserved heterogeneity manifests itself in shifts to the marginal utility of commodities. This assumption is restrictive but cannot be rejected for my data set. An important extension to my approach would allow individual utility functions to have heterogeneous curvatures. This extension is non-trivial as the revealed preference inequalities characterising such a model are non-linear in unknowns, making them difficult to implement. However, I am exploring the use of convex optimisation techniques in the hope of providing a tractable procedure for this model. Further, the empirical application in this paper is intentionally simple; its aim being to demonstrate the techniques on non-simulated data. I look forward to applying these techniques on a wider range of data on consumer spending and labour supply, to provide interesting policy insights. Such research will have a similar focus to recent work by Manski (2014), who uses revealed preference techniques to examine the robustness of income tax policy evaluations, and Klein and Tartari (2014), who bound labour supply responses to a randomised welfare experiment using revealed preference methods to identify counterfactual choices.

Finally, developing a framework for inference is an important topic for future work, as is an allowance for relaxing independence of ϵ and budget parameters. Early work in these directions is promising; revealed preference restrictions bound unobserved preference parameters even if prices and expenditure are a function of ϵ. Regarding a framework for inference, I hope to examine the connections between Brown and Wegkamp (2002) and Pakes and Pollard (1989) to a setting where the parameters of interest are functions, and to determine the consistency of sub-sampling estimates of the sampling distribution.

References

[1] Abi Adams, Richard Blundell, Martin Browning and Ian Crawford (2014), "Prices versus Preferences: Rationalising Tobacco Consumption", mimeo.

[2] Sydney Afriat (1967),"The Construction of Utility Functions from Expenditure Data", *International Economic Review*, 8(1), 67-77.

[3] Sydney Afriat (1972), "Efficiency Estimation of Production Functions", *International Economic Review*, 13(3), 568-598.

[4] Sydney Afriat (1977), *The Price Index*, London: Cambridge University Press.

[5] Adelchi Azzalini and Adrian Bowman (1997), *Applied Smoothing Techniques for Data Analysis*, New York: Oxford University Press.

[6] Walter Beckert and Richard Blundell (2008), "Heterogeneity and the Non-Paramteric Analysis of Consumer Choice: Conditions for Invertibility", *The Review of Economic Studies*, 75, 1069-1080.

[7] V. E. Benes (1965), *Mathematical Theory of Connecting Networks and Telephone Traffic* New York: Academic Press.

[8] C. Lanier Benkard and Steve Berry (2004), "On the Nonparametric Identification of Nonlinear Simultaneous Equations Models: Comment on B. Brown (1983) and Roehrig (1988), Cowles Foundation Discussion Paper 1482.

[9] H. D. Block and Jacob Marschak (1960), "Random Orderings and Stochastic Theories of Responses", in *Controbutions to Probability and Statistics* edited by I Olkin, S. Ghurye, H. Hoeffding, H. Madow and h. Mann. California: Stanford University Press.

[10] Richard Blundell, Martin Browning and Ian Crawford (2003), "Nonparametric Engel Curves and Revealed Preference", *Econometrica*, 71(1), 205-240.

[11] Richard Blundell, Martin Browning and Ian Crawford (2008),"Best Nonparametric Bounds on Demand Responses", *Econometrica*, 76(6), 1227-1262.

[12] Richard Blundell, Martin Browning, Laurens Cherchye, Ian Crawford, Bram De Rock and Frederic Vermeulen (2012), "Sharp for SARP: Nonparametric Bounds on the Behavioural and Welfare Effects of Price Changes", *IFS Working Paper W12/14*, The Institute for Fiscal Studies.

[13] Richard Blundell, Joel Horowitz and Matthias Parey (2010), "Measuring the Price Responsiveness of Gasoline Demand: Economic Shape Restrictions and Nonparametric Demand Estimation", *Quantitative Economics*, 3(1), 29-51.

[14] Richard Blundell, Dennis Kristensen and Rosa Matzkin (2012), "Bounding Quantile Demand Functions using Revealed Preference Inequalities", cemmap working paper CWP21/11.

[15] Bryan Brown and Mary Walker (1989), "The Random Utility Hypothesis and Inference in Demand Systems", *Econometrica*, 57(4), 815-829.

[16] Donald Brown and Rosa Matzkin (1998), "Estimation of Nonparametric Functions in Simultaneous Equations Models, with an Application to Consumer Demand", *Cowles Foundation Discussion Papers 1175*, Cowles Foundation, Yale University.

[17] Bryan Brown (1983), "The Identification Problem in Systems Nonlinear in the Variables, *Econometrica*, 51, 175-196.

[18] Martin Browning (1989), "A Nonparametric Test of the Life-Cycle Rational Expectations Hypothesis", *International Economic Review*, 30(4) 979-992.

[19] Adrian Bowman and Adelchi Azzalini (1997), *Applied Smoothing Techniques for Data Analysis: the Kernel Approach with S-Plus Illustrations*, Oxford: Oxford University Press.

[20] Laurens Cherchye, Bram De Rock and Frederic Vermeulen (2007), "The Collective Model of Household Consumption: A Nonparametric Characterization", *Econometrica*, 75(2), 553-574.

[21] Gerard Debreu (1972), "Smooth Preferences", *Econometrica*, 40(4), 603-615.

[22] Jerry Hausman and Whitney Newey (2014), "individual Heterogeneity and Average Welfare ", Mimeo.

[23] Stefan Hoderlein, Jussi Klemelae and Enno Mammen (2010), "Analyzing the Random Coefficient Model Nonparametrically", *Econometric Theory*, 26, 804-837.

[24] Stefan Hoderlein and Anna Vanhems (2010), "Welfare Analysis using Nonseparable Models", *cemmap Working Paper*.

[25] Stefan Hoderlein and Jorg Stoye (2013), "Revealed Preferences in a Heterogeneous Population", *Review of Economics and Statistics*, 96(2), 197-213.

[26] Edwin Jaynes (1957a), "Information Theory and Statistical Mechanics I", *Physical Review*, 106(4), 620-630.

[27] Edwin Jaynes (1957b), "Information Theory and Statistical Mechanics II", *Physical Review*, 108(2), 171-190.

[28] Yuichi Kitamura and Jorg Stoye (2013), "Nonparametric Analysis of Random Utility Models: Testing", cemmap working paper CWP36/13.

[29] Patrick Kline and Melissa Tartari (2014), "Labor Supply Responses to a Randomized Welfare Experiment: A Revealed Preference Approach", mimeo.

[30] Andrew Leicester (2012), "How Might In-Home Scanner Technology be Used in Budget Surveys?", *IFS Working Papers W12/01*, The Institute for Fiscal Studies.

[31] Andrew Leicester and Zoe Oldfield (2009), "Using Scanner Technology to Collect Expenditure Data", *Fiscal Studies*, 30(3-4), p309-337.

[32] Arthur Lewbel (2001), "Demand Systems With and Without Errors", *American Economic Review*, 91(3), 611-618.

[33] Arthur Lewbel and Krishna Pendakur (2014), "Generalized Random Coefficients with Equivalence Scale Applications", Mimeo.

[34] Rosa Matzkin (2003), "Nonparametric Estimation of Nonadditive Random Functions", *Econometrica*, 71(5), 1339-1375.

[35] Rosa Matzkin (2008), "Identification in Nonparametric Simultaneous Equations Models", *Econometrica*, 76(5), 945-978.

[36] Rosa Matzkin (2007), "Nonparametric Identification", in *Handbook of Econometrics*, Volume 6B, edited by James Heckman and E. Leamer, Amsterdam: North-Holland. 5307-5368.

[37] Rosa Matzkin (2007), "Heterogeneous Choice", in *Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress of the Econometric Society* edited by Richard Blundell, Whitney Newey and Torsten Persson, Cambridge: Cambridge University Press.

[38] Rosa Matzkin (2010), "Estimation of Nonparametric Models with Simultaneity", Mimeo.

[39] Rosa Matzkin and Marcel Richter (1991), "Testing Strictly Concave Rationality", *Journal of Economic Theory*, 53, 287-303.

[40] Charles Manski (1983), "Closest Empirical Distribution Estimation", *Econometrica*, 51(2), 305-320.

[41] Charles Manski (2014), "Identification of income-leisure preference and evaluation of income tax policy", *Quantitative Economics*, 5, 145-174.

[42] Daniel McFadden and Mogens Fosgerau (2012), "A Theory of the Perturbed Consumer with General Budgets", *NBER Working Papers 17953*, National Bureau of Economic Research, Inc.

[43] Daniel McFadden and Marcel Richter (1991), "Stochastic Rational and Revealed Stochastic Preference", in *Preferences, Uncertainty and Rationality* edited by J. Chipman, D. McFadden and K Richter, Boulder: Westview Press.

[44] Daniel McFadden (2005), "Revealed Stochastic Preference: A Synthesis", *Economic Theory*, 26, 245-264.

[45] Peter Neary and Kevin Roberts (1980), "The Theory of Household Behaviour under Rationing", *European Economic Review*, 13(1), 25-42.

[46] Ranga Rao (1962), "Relations between Weak and Uniform Convergence of Measures with Applications", *The Annals of Mathematical Statistics*, 33(2), 659-680.

[47] Charles Roehrig (1988), "Conditions for Identification in Nonparametric and Parametric Models", *Econometrica*, 56(2), 433-447.

[48] Erwin Rothbarth (1941), "The Measurement of Changes in Real Income under Conditions of Rationing", *Review of Economic studies*, 8(2), 100-107.

[49] Paul Samuelson (1938), "A Note on the Pure Theory of Consumer Behavior", *Economica*, 5(17), 61-71.

[50] Paul Samuelson (1948), "Consumption Theory in Terms of Revealed Preference", *Economica*, 15(60), 243-253.

[51] Hal Varian (1982), "The Nonparametric Approach to Demand Analysis", *Econometrica*, 50(4), 945-973.

[52] Hal Varian (1983), "Non-Parametric Tests of Consumer Behaviour", *Review of Economic Studies*, 50(1), 99-110.

[53] Hal Varian (1990), "Goodness-of-fit in Optimizing Models", *Journal of Econometrics*, 46(1-2), 125-140.

8 Appendix A: Proofs

Theorem 1. If one can find sets $\{u_i\}_{i=1,...,N}$, $\{\boldsymbol{\epsilon}_i\}_{i=1,...,N}$ and $\{\lambda_i\}_{i=1,...,N}$ with $u_i \in \mathbb{R}$, $\lambda_i \in \mathbb{R}_{++}$ and $\boldsymbol{\epsilon}_i \in \mathbb{R}^{K-1}$, such that:

$$u_i - u_j < \lambda_j \mathbf{p}_j'(\mathbf{q}_i - \mathbf{q}_j) - \boldsymbol{\epsilon}_j'(\mathbf{q}_i^{\neg K} - \mathbf{q}_j^{\neg K}) \quad \text{(A1)}$$

$$\epsilon_i^k < \lambda_i p_i^k \quad \text{(A2)}$$

$$\epsilon_i^k - \epsilon_j^k < \lambda_i p_i^k \quad \text{(A3)}$$

$$\frac{1}{N}\sum_{i=1}^{N} \epsilon_i^k = 0 \quad \text{(A4)}$$

for all $i, j = \{1, ..., N\}$, and, for all $k = \{1, ..., K-1\}$, then a random sample of observed choice behaviour $\{\mathbf{p}_i, \mathbf{q}_i\}_{i=1,...,N}$ is consistent with the maximisation of $u(\mathbf{q}) + \boldsymbol{\epsilon}'\mathbf{q}^{\neg K}$.

This proof concerns the sufficiency of conditions (A1) to (A4) for rationalisation by the theoretical framework. I use analgous methods to Varian (1982) and Matzkin and Richter (1991) to prove this sufficiency result.

Following Matzkin and Richter (1991), since there are only a finite number of revealed preference inequalities, there exists a $\delta > 0$ such that:

$$u_i \;<\; u_j + \lambda_j \mathbf{p}_j'(\mathbf{q}_i - \mathbf{q}_j) - \boldsymbol{\epsilon}_j'(\mathbf{q}_i^{\neg K} - \mathbf{q}_j^{\neg K}) - \delta \quad (41)$$

for $i \neq j$. Let $T > 0$ and define a function $g : \mathbb{R}^K \to \mathbb{R}^1$ as:

$$g(\mathbf{q}) = \sqrt{(q^1)^2 + ... + (q^K)^2 + T} - \sqrt{T} \quad (42)$$

Then $g(\mathbf{q}) > 0$ for all $\mathbf{q} > 0$, g is strictly convex and differentiable and

$$\frac{\partial g(\mathbf{q})}{\partial q^k} < 1 \quad (43)$$

For each $i = 1, .., N$, define $\phi_i : \mathbb{R}^K \to \mathbb{R}^1$ as:

$$\phi_i(\mathbf{q}) \equiv u_i + \lambda_i \mathbf{p}_i'(\mathbf{q} - \mathbf{q}_i) - \boldsymbol{\epsilon}_i'(\mathbf{q}^{\neg K} - \mathbf{q}_i^{\neg K}) - \delta g(\mathbf{q} - \mathbf{q}_i) \quad (44)$$

Since g is strictly convex, ϕ_i is strictly concave.

Define the base utility function $u : \mathbb{R}^K \to \mathbb{R}^1$ as:

$$u(\mathbf{q}) = \min_i \{\phi_i(\mathbf{q})\} \tag{45}$$

As the minimum of finitely many strictly concave functions, $u(\mathbf{q})$ is strictly concave.

As $u(\mathbf{q})$ is defined as the minimum of finitely many ϕ_i functions, to prove monotonicity of $u(\mathbf{q})$ it is sufficient to show that each ϕ_i has a strictly positive partial derivative.

$$\frac{\partial \phi_i(\mathbf{q})}{\partial q^k} = \lambda_i p_i^k - \epsilon_i^k - \delta g_k(\mathbf{q} - \mathbf{q}_i) \tag{46}$$

Given the construction of g

$$\frac{\partial \phi_i(\mathbf{q})}{\partial q^k} > \lambda_i p_i^k - \epsilon_i^k - \delta \tag{47}$$

Condition (A2) imposes that

$$\lambda_i p_i^k - \epsilon_i^k > 0 \tag{48}$$

Since there are finitely many indices $i = \{1, ..., N\}$, one can pick a δ small enough such that Equation 47 is positive for all $k = 1, ..., K$.

Strict monotonicty of the function $U(\mathbf{q}, \boldsymbol{\epsilon}_i)$ for all $\mathbf{q}$ and $\boldsymbol{\epsilon}_i$ with $i = \{1, ..., N\}$ is guaranteed by satisfaction of (A3). Defining $U(\mathbf{q}, \epsilon_j) = u(\mathbf{q}) + \boldsymbol{\epsilon}_j' \mathbf{q}^{\neg K}$, then monotonicity of $U(\mathbf{q}, \epsilon_j)$ for $j = 1, .., N$ is guaranteed if for any ϕ_i, U has a strictly positive partial derivative:

$$\begin{aligned} \frac{\partial \phi_i(\mathbf{q})}{\partial q^k} + \epsilon_j^k &= \lambda_i p_i^k - \epsilon_i^k + \epsilon_j^k - \delta g_k(\mathbf{q} - \mathbf{q}_i) && (49) \\ &> \lambda_i p_i^k - \epsilon_i^k + \epsilon_j^k - \delta && (50) \end{aligned}$$

Condition (A3) imposes that

$$\lambda_i p_i^k - \epsilon_i^k + \epsilon_j^k > 0 \tag{51}$$

Since there are finitely many indices $i = \{1, ..., N\}$, one can pick a δ small enough such that Equation 50 is positive for all $k = 1, ..., K$ and $i = \{1, ..., N\}$.

I now show that (R1) of the Rationalisation Inequality is satisfied. Consider an arbitrary bundle, $\mathbf{q}$, that is feasible given the budget of individual i:

$$\mathbf{p}_i' \mathbf{q} \leq \mathbf{p}_i' \mathbf{q}_i \tag{52}$$

Given the specification for the base individual utility function:

$$\begin{aligned} U(\mathbf{q}, \boldsymbol{\epsilon}_i) &= \min_j \{\phi_j(\mathbf{q})\} + \boldsymbol{\epsilon}_i' \mathbf{q}^{\neg K} && (53) \\ &= \min_j \{u_j + \lambda_j \mathbf{p}_j'(\mathbf{q} - \mathbf{q}_j) - \boldsymbol{\epsilon}_j'(\mathbf{q}^{\neg K} - \mathbf{q}_j^{\neg K}) - \delta g(\mathbf{q} - \mathbf{q}_j)\} + \boldsymbol{\epsilon}_i' \mathbf{q}^{\neg K} && (54) \\ &\leq u_i + \lambda_i \mathbf{p}_i'(\mathbf{q} - \mathbf{q}_i) - \boldsymbol{\epsilon}_i'(\mathbf{q}^{\neg K} - \mathbf{q}_i^{\neg K}) - \delta g(\mathbf{q} - \mathbf{q}_i) + \boldsymbol{\epsilon}_i' \mathbf{q}^{\neg K} && (55) \\ &\leq U(\mathbf{q}_i, \boldsymbol{\epsilon}_i) + \lambda_i \mathbf{p}_i'(\mathbf{q} - \mathbf{q}_i) && (56) \end{aligned}$$

Noting that $\mathbf{q}$ is feasible and thus,

$$\begin{aligned} \mathbf{p}_i' \mathbf{q} &\leq \mathbf{p}_i' \mathbf{q}_i && (57) \\ \mathbf{p}_i'(\mathbf{q} - \mathbf{q}_i) &\leq 0 && (58) \end{aligned}$$

and that $g(\mathbf{q}) > 0$, it is the case that:

$$U(\mathbf{q}, \boldsymbol{\epsilon}_i) < U(\mathbf{q}_i, \boldsymbol{\epsilon}_i) \quad (59)$$

In words, any other feasible bundle given $U(\mathbf{q}, \boldsymbol{\epsilon}_i)$ yields lower utility than $\mathbf{q}_i$ and thus (R1) is satisfied.

Rationalisation Inequality (R2) requires:

$$u_k(\mathbf{q}) + \underline{\epsilon}^k > 0 \quad (60)$$

for $i = \{1, ..., N\}$ and $k = \{1, ..., K-1\}$. This is guaranteed by the strict monotonicty of the recovered utility function.

$$\begin{aligned} u_k(\mathbf{q}) + \underline{\epsilon}^k &= \frac{\partial \phi_i(\mathbf{q})}{\partial q^k} + \underline{\epsilon}^k && (61) \\ &= \lambda_i p_i^k - \epsilon_i^k + \underline{\epsilon}^k - \delta g_k(\mathbf{q} - \mathbf{q}_i) && (62) \\ &> \lambda_i p_i^k - \epsilon_i^k + \underline{\epsilon}^k - \delta && (63) \end{aligned}$$

Condition (A3) imposes that

$$\lambda_i p_i^k - \epsilon_i^k + \underline{\epsilon}^k > 0 \quad (64)$$

as $\underline{\epsilon}^k = \min_i \epsilon_i^k$. Then, given there are finitely many indices $i = \{1, ..., N\}$, one can pick a δ small enough such that Equation 63 is positive for all $k = 1, ..., K$ and $i = \{1, ..., N\}$.

This completes the proof.

I also comment on the necessity of conditions (A1), (A2) and (A3). From the first order conditions:

$$\nabla_{\mathbf{q}} u(\mathbf{q}_i) + \boldsymbol{\epsilon}_i = \lambda_i \mathbf{p}_i \tag{65}$$

Concavity of $U(\mathbf{q}, \boldsymbol{\epsilon})$ implies:

$$U(\mathbf{q}_i, \boldsymbol{\epsilon}_j) + U(\mathbf{q}_j, \boldsymbol{\epsilon}_j) \leq \nabla_{\mathbf{q}} U(\mathbf{q}_j, \boldsymbol{\epsilon}_j)'(\mathbf{q}_i - \mathbf{q}_j) \tag{66}$$

$$u(\mathbf{q}_i) - u(\mathbf{q}_j) \leq \nabla_{\mathbf{q}} U(\mathbf{q}_j, \boldsymbol{\epsilon}_j)'(\mathbf{q}_i - \mathbf{q}_j) - \boldsymbol{\epsilon}_j'(\mathbf{q}_i^{\neg K} - \mathbf{q}_j^{\neg K}) \tag{67}$$

Substituting the first order conditions into the concavity condition and rearranging gives:

$$u(\mathbf{q}_i) - u(\mathbf{q}_j) \leq \lambda_j \mathbf{p}_j'(\mathbf{q}_i - \mathbf{q}_j) - \boldsymbol{\epsilon}_j'(\mathbf{q}_i^{\neg K} - \mathbf{q}_j^{\neg K})$$

Letting $u_i = u(\mathbf{q}_i)$, returns (A1).

(A2) and (A3) are required for the base utility function to be strictly increasing in $\mathbf{q}$ for any $\boldsymbol{\epsilon}$ such that $F_\epsilon(\boldsymbol{\epsilon}) > 0$. If $\epsilon_i^k \leq \lambda_i p_i^k$, then $\nabla_{\mathbf{q}^k} u(\mathbf{q}_i) < 0$, violating montonicity. (A2) is thus necessary. (A3) is necessary to impose montonicity at $\mathbf{q}_i$ given $\boldsymbol{\epsilon}_j$ with $j \neq i$.

Theorem 2. If a non-empty feasible set to the inequalities of Theorem 1 exists, then it is unbounded.

Imagine that a feasible solution to Theorem 1 exists. Then it is the case that the following inequalities are satisfied.

$$u_i - u_j < \lambda_j \mathbf{p}_j'(\mathbf{q}_i - \mathbf{q}_j) - \boldsymbol{\epsilon}_j'(\mathbf{q}_i^{\neg K} - \mathbf{q}_j^{\neg K}) \tag{68}$$

Thus, it is also true that for any $\delta > 0$.

$$\delta u_i - \delta u_j < \delta \lambda_j \mathbf{p}_j'(\mathbf{q}_i - \mathbf{q}_j) - \delta \boldsymbol{\epsilon}_j'(\mathbf{q}_i^{\neg K} - \mathbf{q}_j^{\neg K}) \tag{69}$$

$$\delta \epsilon_i^k < \delta \lambda_i p_i^k \tag{70}$$

$$\delta \epsilon_i^k - \delta \epsilon_j^k < \delta \lambda_i p_i^k \tag{71}$$

$$\delta \frac{1}{N} \sum_{i=1}^{N} \epsilon_i^k = 0 \tag{72}$$

Therefore, the feasible set to Theorem 1 is unbounded.

Theorem 3. Identification of $(F_\epsilon^\star, \nabla_{\mathbf{q}} u^\star(\mathbf{q})$

Assume that the vector $(\mathbf{p}, x)$ has a continuous Lebesgue density and that the joint distribution of observables $F_{Q,P,X}$ is identified. Let $\bar{\mathbf{q}} \in Q$ and $\alpha \in R$ be given. Suppose that W is a set of smooth utility functions $u : Q \to R$ such that $\forall u \in W$, $u(\bar{\mathbf{q}}) = \alpha$. Let Ω denote the set of functions $\nabla_{\mathbf{q}} u(\mathbf{q})$ where $u \in W$. Denote by Γ the set of absolutely continuous distribution functions of vectors $(\epsilon^1, ..., \epsilon^{K-1})$ that satisfy $E(\boldsymbol{\epsilon}) = \mathbf{0}$. Then, $(\nabla_{\mathbf{q}} u(\mathbf{q}), F(\boldsymbol{\epsilon}))$ is identified in $(\Omega \times \Gamma)$ if the following conditions are imposed:

1. Independence of $\boldsymbol{\epsilon}$ and $(\mathbf{p}, x)$.
2. $u_K(\mathbf{q}) = f(\mathbf{q})$, with $f(\mathbf{q}) > 0$ and $f(\mathbf{q}) < \infty$.

Intuitively, given the population joint distribution of observables one knows the set of budgets that support any given $\mathbf{q}$. Given that, by assumption, $E(\epsilon_k) = 0$ plus independence of taste and budget parameters and a restriction on $u_K(\mathbf{q})$, one is able to decompose the difference in budgets that support $\mathbf{q}$ into a component caused by the marginal utility of income at $\mathbf{q}$ and a component attributable to preference heterogeneity. One thus recovers the distribution $F_\epsilon^\star$ from the conditional density $f_{\mathbf{p}|Q=\mathbf{q}}$.

Given that $U(\mathbf{q}, \boldsymbol{\epsilon})$ is strictly monotone and smooth in the sense of Debreu, for each $\boldsymbol{\epsilon}$ and a given $\mathbf{q}$, there exists a unique price vector, $\mathbf{p}_{\epsilon,q}$, and expenditure $x_{\epsilon,q} = \mathbf{p}'_{\epsilon,q}\mathbf{q}$, such that:

$$\mathbf{q} = d(\mathbf{p}_{\epsilon,q}, x_{\epsilon,q}, \boldsymbol{\epsilon}) \tag{73}$$

with $\mathbf{p}_{\epsilon',q} = \mathbf{p}_{\epsilon'',q}$ and $x_{\epsilon',q} = x_{\epsilon'',q}$ if and only if $\boldsymbol{\epsilon}' = \boldsymbol{\epsilon}''$.

From the first order conditions of the model we have that:

$$\frac{\nabla_{\mathbf{q}} u(\mathbf{q})}{u_K(\mathbf{q})} + \frac{\boldsymbol{\epsilon}}{u_K(\mathbf{q})} = \mathbf{p}_{\epsilon,q} \tag{74}$$

The expectation of prices supporting some bundle $\mathbf{q}$ is:

$$E(\mathbf{p}|Q = \mathbf{q}) = E\left(\frac{\nabla_{\mathbf{q}} u(\mathbf{q})}{u_K(\mathbf{q})} + \frac{\boldsymbol{\epsilon}}{u_K(\mathbf{q})}|Q = \mathbf{q}\right) \tag{75}$$

$$= \frac{\nabla_{\mathbf{q}} u(\mathbf{q})}{u_K(\mathbf{q})} + \frac{E(\boldsymbol{\epsilon}|Q = \mathbf{q})}{u_K(\mathbf{q})} \tag{76}$$

For good k it is the case that:

$$E(p^k|Q=\mathbf{q}) = \frac{u_k(\mathbf{q})}{u_K(\mathbf{q})} + \frac{1}{u_K(\mathbf{q})}\int_{\frac{-u_k(\mathbf{q})}{u_K(\mathbf{q})}}^{\infty} \epsilon^k f(\epsilon^k|Q=\mathbf{q})\mathrm{d}\epsilon^k \tag{77}$$

Given identification of the joint density $F_{Q,P,X}$, independence of $\boldsymbol{\epsilon}$ and $(\mathbf{p},x)$, and that strict monotonicity implies $\epsilon^k > \frac{-u_k(\mathbf{q})}{u_K(\mathbf{q})}$, the price supporting $\mathbf{q}$ for each $\boldsymbol{\epsilon}$ is observed and thus:

$$\begin{aligned}
E(p^k|Q=\mathbf{q}) &= \frac{u_k(\mathbf{q})}{u_K(\mathbf{q})} + \frac{1}{u_K(\mathbf{q})}\int_{\frac{-u_k(\mathbf{q})}{u_K(\mathbf{q})}}^{\infty} \epsilon^k f(\epsilon^k|Q=\mathbf{q})\mathrm{d}\epsilon^k & (78)\\
&= \frac{u_k(\mathbf{q})}{u_K(\mathbf{q})} + \frac{1}{u_K(\mathbf{q})}\int_{\frac{-u_k(\mathbf{q})}{u_K(\mathbf{q})}}^{\infty} \epsilon^k f(\epsilon^k)\mathrm{d}\epsilon^k & (79)\\
&= \frac{u_k(\mathbf{q})}{u_K(\mathbf{q})} & (80)
\end{aligned}$$

Given that the expectation of all prices supporting $\mathbf{q}$ gives the marginal rate of substitution of the base utility function at $\mathbf{q}$, then the heterogeneity vector characterising some individual that demands $\mathbf{q}$ at prices $\mathbf{p}_{\epsilon',q}$ is given as

$$\begin{aligned}
\mathbf{p}_{\epsilon',q} - E(\mathbf{p}|Q=\mathbf{q}) &= \frac{\boldsymbol{\epsilon}'}{u_K(\mathbf{q})} & (81)\\
\boldsymbol{\epsilon}' &= u_K(\mathbf{q})\left(\mathbf{p}_{\epsilon',q} - E(\mathbf{p}|Q=\mathbf{q})\right) & (82)
\end{aligned}$$

With the restriction of the marginal utility of the K^{th} good to a known and bounded function, $u_K(\mathbf{q}) = f(\mathbf{q})$, then the distribution of heterogeneity parameters can be identified from the distribution of prices supporting any given quantitiy bundle:

$$F_\epsilon(\boldsymbol{\epsilon}) = Pr\left(f(\mathbf{q})(\mathbf{p}_{\epsilon,q} - E(\mathbf{p}|Q=\mathbf{q})) \leq \boldsymbol{\epsilon}\right) \tag{83}$$

From above, the gradient of the base utility function, $\nabla_{\mathbf{q}} u(\mathbf{q})$, can then be recovered as:

$$\begin{aligned}
\nabla_{\mathbf{q}} u(\mathbf{q}) &= u_K(\mathbf{q})E(\mathbf{p}|Q=\mathbf{q}) & (84)\\
&= f(\mathbf{q})E(\mathbf{p}|Q=\mathbf{q}) & (85)
\end{aligned}$$

Theorem 4. Strong Consistency of $\hat{F}_{\epsilon,N}(\boldsymbol{\epsilon})$ for F_ϵ

Let the conditions for identification hold, that $u(\mathbf{q})$ and its derivatives up to the second order are equicontinuous and uniformly bounded and that F_ϵ possesses absolutely continuous marginal distributions. Then $\hat{F}_{\epsilon,N}(\boldsymbol{\epsilon})$ is a strongly consistent estimator for F_ϵ.

In addition to the identification conditions of Theorem 3, the set of additional assumptions required to establish the strong consistency of $F_{\epsilon,N}(\boldsymbol{\epsilon})$ are:

1. $u(\mathbf{q})$ and its derivatives up to the second order be equicontinuous and uniformly bounded. This implies that W, the set of candidate utility functions satisfying all conditions will be compact with respect to the metric d:

$$d(u(\mathbf{q}), u'(\mathbf{q})) = \sup |u(\mathbf{q}) - u'(\mathbf{q})| + \sup \max_{k \in \{1,\ldots,K\}} |u_k(\mathbf{q}) - u'_k(\mathbf{q})| + \sup \max_{k,j \in \{1,\ldots,K\}} |u_{kj}(\mathbf{q}) - u'_{kj}(\mathbf{q})| \tag{86}$$

2. The marginal distributions of F_ϵ are absolutely continuous.

The proof then follows straightforwardly from Theorem 2 of Brown and Matzkin (1998) and Manski (1983) except with convergence with respect to the supremum norm rather than the Prohorov metric. When it is assumed that F_ϵ possesses absolutely continuous marginal distributions, weak convergence of cumulative distribution functions is equivalent to convergence with respect to the supremum norm (Rao, 1962), giving the desired result.