Stochastic Demand and Revealed Preference*

Richard Blundell[†] Dennis Kristensen[‡] Rosa Matzkin[§]

June 2009,
This Version: August 2010

Preliminary Draft

Abstract

This paper develops new techniques for the estimation and testing of stochastic consumer demand models. Particular attention is given to nonseparable unoberserved heterogeneity. The objective is to elicit demand responses from consumer expenditure survey data. Revealed preference inequality restrictions from consumer optimisation conditions are utilized to improve on the nonparametric estimation and testing of demand responses. Bounds on demand reponses to price changes are estimated nonparametrically, and their asymptotic properties are derived. We also devise a test for rationality of demand behavior. An empirical application using data from the U.K. Family Expenditure Survey illustrates the usefulness of the methods.

JEL: C20, D12

Keywords: microeconometrics, consumer behaviour, revealed preference, bounds, testing rationality.

*We would like to thank Ian Crawford and Roger Koenker for sharing Matlab code with us and helping with the implementation. We also wish to thank participants at the Latin-American Meetings in Buenos Aires, the CAM workshop at University of Copenhagen, the Malinvaud Seminar, and seminars at UC San Diego, Universitat Autònoma de Barcelona, and Toulouse School of Economics for helpful comments and suggestions. The research is part of the program of research of the ESRC Centre for the Microeconomic Analysis of Public Policy at IFS. Funding from the ESRC, grant number R000239865 is gratefully acknowleged. Kristensen acknowledges support from the NSF (grant no. SES-0961596). Material from the FES made available by the ONS through the ESRC Data Archive has been used by permission of the controller of HMSO. Neither the ONS nor the ESRC Data Archive bear responsibility for the analysis or the interpretation of the data reported here. The usual disclaimer applies.

[†]Department of Economcis, UCL, and Institute for Fiscal Studies, email: r.blundell@ucl.ac.uk
[‡]Department of Economics, Columbia University and CREATES. email: dk2313@columbia.edu
[§]Department of Economics, UCLA. email: matzkin@econ.ucla.edu.

1 Introduction

This paper develops new nonparametric techniques for the estimation and prediction of consumer demand responses. The objectives are two-fold: First, to impose minimum restrictions on how unobserved heterogeneity enters the demand function of the individual consumer. Second, to utilize restrictions arriving from choice theory to improve demand estimation and prediction.

Our analysis takes place in the commonly occurring empirical setting where only a relatively small number of market prices are observed, but within each of those markets the demand responses of a large number of consumers are reported. In this setting, it is not possible to point identify the predicted demand response to a new, unobserved price. We will instead use restrictions derived from revealed preference theory to establish bounds on the demand responses: if consumers behave according to the axioms of revealed preference their vector of demands at each relative price will satisfy certain well known inequalities (see Afriat, 1973 and Varian, 1982). If, for any individual, these inequalities are violated then that consumer can be deemed to have failed to behave according to the optimisation rules of revealed preference.

Blundell, Browning and Crawford (2003, 2008) extend the analysis of Varian (1982) and obtain 'expansion path based bounds' (E-bounds) which are the tightest possible bounds, given the data and the inequality restrictions implied from choice theory. This paper develops this line of research and uses revealed preference restrictions to obtain improved *quantile* demand estimates at observed prices, and to establish bounds on the predicted *distribution* of demands at new prices.

A key ingredient of the analysis we conduct is the Engel curve which describes the *expansion path* for demand as total expenditure changes. We first develop nonparametric sieve estimators of the stochastic demand functions at any given set of observed prices, these are Engel curves or expansion paths. We then show how revealed preference restrictions can be imposed across expansion paths.

The modelling and estimation of the Engel curve relationship has a long history. Working (1943) and Leser (1963) suggested standard parametric regression models where budget shares are linear functions of log total budget; the so-called Piglog-specification. This simple linear model has since then been generalised in various ways since empirical studies suggested that higher order logarithmic expenditure terms are required for certain expenditure share equations, see e.g. Hausman, Newey and Powell (1995), Lewbel (1991), Banks, Blundell and Lewbel (1997). An obvious way to detect the presence of such higher order terms is non- and semiparametric methods which have been widely used in the econometric analysis of Engel curves, see for example Blundell, Chen and Kristensen (2007), Blundell and Duncan (1998) and Härdle and Jerison (1988).

In the majority of empirical studies it has been standard to impose an additive error structure. This type of specification allows relatively straightforward identification and estimation of the structural parameters. Additive heterogeneity on the other hand imposes very strong assumptions on the class of underlying utility functions, see e.g. Lewbel (2001), and as such is in general inconsistent with economic theory; see also Beckert (2007). In response to this problem with existing empirical consumer demand models, we allow for non-additive heterogeneity.[1] Our approach imposes only weak assumptions on the specification of unobserved heterogeneity in nonparametric demand functions. We derive a constrained nonparametric quantile estimator for the non-additive demand model, and derive its asymptotic properties.

An early treatment of nonparametric identification of non-additive models is Brown (1983) whose results were extended in Roehrig (1988). Brown and Matzkin (1998, 2003) and Matzkin (2003a,b) building on their work propose estimators. A number of other papers have addressed identification and estimation: Chesher (2001, 2002a, 2002b, 2003) consider quantile-driven identification; Ma and Koenker (2003) make use of his results to construct parametric estimators; Imbens and Newey (2009) and Chernozhukov, Imbens and Newey (2003) establish estimators which allow for endogeneity. Our approach draws on this literature and our unconstrained estimator is similar to the one developed in Imbens and Newey (2009).

The second part of the paper is concerned with *predicting the distribution* of consumer demand responses to a new relative price level that has not been previously observed. Given the limited price variation it is not possible to nonparametrically point identify demand at new, previously unobserved, relative prices. However, the revealed preference restrictions allow us to establish bounds on the demands for the nonseparable heterogeneity case that is the focus here. We show that the demand functions estimated at observed prices, can be used to recover these bounds nonparametrically. The estimation problem falls within the framework of partially identified models (see e.g. Manski, 1993). We employ the techniques developed in, amongst others, Chernozhukov, Hong and Tamer (2003) to establish the properties of the nonparametric demand bounds estimators.

The estimation and prediction strategy outlined above heavily exploits revealed preference inequalities. In the third part of the paper, we develop a nonparametric test for this rationality assumption. In this testing problem the testable restrictions take the form of a set of inequality constraints. As such the testing problem is similar to one of estimation and testing when the parameter is on the boundary of the parameter space as analyzed in, for example, Andrews (1999, 2001) and Andrews and Guggenberger (2009). By importing the techniques developed there, we derive the asymptotic properties of the test statistic.

Our empirical analysis is based on data from the British Family Expenditure Survey (FES) where the relative price variation occurs over time, and samples of consumers, each of a particular household type, are observed at specific points in time in particular regional locations. We estimate bounds on demand functions and test the revealed preference inequality restrictions.[2]

An important feature of the estimation of Engel curve is the possible presence of endogeneity in the total expenditure variable. In a parametric framework this can be dealt with using standard IV-techniques. In recent years, a range of different methods have been proposed to deal with this problem in a nonparametric setting. The two main approaches proposed in the literature is

[1] See Lewbel and Pendakur (2009) for one of the few parametric specification that allows non-additive interaction.

[2] We note that other papers have combined nonparametric techniques and economic theory to estimate and test demand systems; see, for example, Haag, Hoderlein and Pendakur (2009), Hoderlein (2008), Hoderlein and Stoye (2009), Lewbel (1995).

nonparametric IV (Ai and Chen (2003), Hall and Horowitz (2003), Darolles, Florens and Renault (2002)) and control functions (Newey, Powell and Vella (1998)). Both these methods have been applied in the empirical analysis of Engel curves (Blundell, Chen and Kristensen (2003) and Blundell, Duncan and Pendakur (1998) respectively). We briefly discuss how our estimators and tests can be extended to handle endogeneity of explanatory variables by using recent results in Chen and Pouzo (2009), Chernozhukov, Imbens and Newey (2007) and Imbens and Newey (2009).

The remainder of the paper is organized as follows: In Section 2, we set up the basic econometric framework. In Section 3-4, we develop estimators of the demand functions at observed prices. The estimation of demand bounds is considered in Section 5, while a test for rationality is developed in Section 6. In Sections 7 and 8, we discuss the implementation of the estimators and how to compute confidence bands. We briefly discuss how to allow for endogenous explanatory variables in Section 9. Section 10 contains an empirical application on British household data. We conclude in Section 11. All proofs have been relegated to the Appendix.

2 Heterogeneous Consumers and Market Prices

2.1 Quantile Demand Functions

Consumer demand depends on market prices, individual income and individual heterogeneity. Suppose we have observed consumers in T separate markets, where T is finite. In what follows we will assume these refer to observations over time but they could equally well refer to geographically separated markets. Let $\mathbf{p}(t)$ be the set of prices for the goods that all consumers face at time $t = 1, ..., T$. At each time point t, we draw a new random sample of $n \geq 1$ consumers.

For each consumer, we observe his or her demands and income level (and potentially some other individual characteristics such as age, education etc.). Let $\mathbf{q}_i(t)$ and $x_i(t)$ be consumer i's ($i = 1, ..., n$) vector of demand and income level at time t ($t = 1, ..., T$). We stress that the data $\{\mathbf{p}(t), \mathbf{q}_i(t), x_i(t)\}$, for $i = 1, ..., n$ and $t = 1, ..., T$, is not a panel data set since we do not observe the same consumer over time. Individual heterogeneity in observed and unobserved characteristics implies that, for any *given* market prices $\mathbf{p}(t)$ and for consumers with income x, there will be a *distribution* of demands. Changes in x map out a distribution of expansion paths.

We focus on the situation where there are only two goods in the economy such that $\mathbf{q}(t) = (q_1(t), q_2(t))' \in \mathbb{R}^2_+$ and $\mathbf{p}(t) = (p_1(t), p_2(t))' \in \mathbb{R}^2_+$.[3] The demand for good 1 is assumed to arrive from the following demand function,

$$q_1(t) = d_1(x(t), \mathbf{p}(t), \varepsilon(t)),$$

where $\varepsilon(t) \in \mathbb{R}$ is an individual specific heterogeneity term that reflects unobserved heterogeneity in preferences and characteristics.[4] To ensure that the budget constraint is met, the demand for

[3] By the end of the paper, we discuss extensions to general, multidimensional goods markets.

[4] The demand function could potentially depend on other observable characteristics besides income, but to keep the notation at a reasonable level we suppress such dependence in the following. If additionally explanatory variables are present, all the following assumptions, arguments and statements are implicitly made conditionally on those.

good two must satisfy:

$$q_2(t) = d_2(x(t), \mathbf{p}(t), \varepsilon(t)) := \frac{x(t) - p_1(t) d_1(x(t), \mathbf{p}(t), \varepsilon(t))}{p_2(t)}. \tag{1}$$

We collect the two demand functions in $\mathbf{d} = (d_1, d_2)$. The demand function $\mathbf{d}$ should be thought of as the solution to an underlying utility maximization problem that the individual consumer solves.

We here consider the often occurring situation where the time span T over which we have observed consumers and prices is small (in the empirical application $T = 6$). In this setting, we are not able to identify the mapping $\mathbf{p} \mapsto \mathbf{d}(x, \mathbf{p}, \varepsilon)$. To emphasize this, we will in the following write

$$\mathbf{d}(x(t), t, \varepsilon(t)) := \mathbf{d}(x(t), \mathbf{p}(t), \varepsilon(t)).$$

So we have a sequence of T demand functions, $\{\mathbf{d}(x, t, \varepsilon)\}_{t=1}^{T}$.

The unobserved heterogeneity $\varepsilon(t)$ is assumed (or normalized) to follow a uniform distribution, $\varepsilon(t) \sim U[0,1]$ and to be independent of $x(t)$.[5] This combined with the assumption that d_1 is invertible in $\varepsilon(t)$ implies that $d_1(x, t, \tau)$ is identified as the τth quantile of $q_1(t) | x(t) = x$ (Matzkin, 2003; Imbens and Newey, 2009):

$$d_1(x, t, \tau) = F^{-1}_{q_1(t)|x(t)=x}(\tau), \quad \tau \in [0,1]. \tag{2}$$

These are the *quantile expansion paths* that describe the way demand changes with income x for any given market t and for any given consumer ε. These are the quantile respresentations of the Engel curves.

2.2 e-Bounds on Demands at New Prices

Consider a particular consumer characterized by some $\tau \in [0,1]$ and income x, with associated sequence of demand functions $\mathbf{d}(x, t, \tau)$, $t = 1, ..., T$. Suppose that the consumer faces a given new price $\mathbf{p}_0$ at an income level x_0. The consumer's new budget set is

$$\mathcal{B}_{\mathbf{p}_0, x_0} = \left\{\mathbf{q} \in \mathbb{R}^2_+ | \mathbf{p}_0' \mathbf{q} = x_0\right\},$$

is compact and convex.

Suppose we observe a set of demands $\{\mathbf{q}_1, \mathbf{q}_2, ...\mathbf{q}_T\}$ which record the choices made by a consumer ($\boldsymbol{\varepsilon}$) when faced by the set of prices $\{\mathbf{p}_1, \mathbf{p}_2, ...\mathbf{p}_T\}$,how do we find the support set for a new price vector $\mathbf{p}_0$ with total outlay x_0? All demands $\mathbf{d}(\mathbf{p}, x, \boldsymbol{\varepsilon})$ are conditioned on unobserved heterogeneity $\boldsymbol{\varepsilon}$.Varian (1986) established that the best *support set* $S^V(\mathbf{p}_0, x_0, \boldsymbol{\varepsilon})$ for $\mathbf{d}(x_0, \mathbf{p}_0, \boldsymbol{\varepsilon})$ is given by:

$$\left\{ \mathbf{q}_0 : \begin{array}{c} \mathbf{p}_0' \mathbf{q}_0 = x_0, \ \mathbf{q}_0 \geq \mathbf{0} \text{ and} \\ \{\mathbf{p(t)}, \mathbf{q(t)}\}_{t=0...T} \text{ satisfies RP} \end{array} \right\}$$

[5] The independence assumption can be relaxed as discussed in Section 9.

Given the expansion paths $\{\mathbf{p(t)}, \mathbf{q}(x(t), \boldsymbol{\varepsilon})\}_{t=1,..T}$, Blundell, Browning and Crawford (2008) define intersection demands $\mathbf{q}_t(\overline{x}(t), \boldsymbol{\varepsilon})$ by $\mathbf{p}_0'\mathbf{q}_t(\overline{x}(t), \boldsymbol{\varepsilon}) = \mathbf{x}_0$. The set of points that are consistent with observed expansion paths and utility maximisation is given by the *support set*:

$$\mathcal{S}_{\mathbf{p}_0,x_0} = \left\{ \mathbf{q} \in \mathcal{B}_{\mathbf{p}_0,x_0} | \mathbf{p}(t)' \mathbf{q} \geq \mathbf{p}(t)' \mathbf{d}(\bar{x}(t), t, \tau), \quad t = 1, ..., T \right\},$$

where $\{\bar{x}(t) : t = 1, ..., T\}$ is a sequence of intersection incomes defined as the solution to

$$\mathbf{p}_0' \mathbf{d}(\bar{x}(t), t, \tau) = x_0, \quad t = 1, ..., T.$$

Note that $\mathcal{S}_{\mathbf{p}_0,x_0}$ is the support set of a given consumer described by τ and as such depends on its value; we here suppress this dependence. This support set uses the quantile expansion paths and RP defines *e-bounds* on demand responses.

Using Blundell, Browning and Crawford (2008), we can establish that if the set of T demands satisfy the Revealed Preference inequalities then the support set is non-empty and convex. The support set can also be shown to define the best bounds on demands $\mathbf{q}_0$ given the observed data and the RP inequalities. These *e-bounds* therefore make maximal use of the heterogeneous expansion paths and the basic nonparametric choice theory in predicting in a new situation. In other words, there do not exist alternative bounds (derived from the same data) which are tighter than the *e-bounds*. It is important to note that the support sets for demand responses are **local** to each point in the distribution of income $\mathbf{x}$ and unobserved heterogeneity $\boldsymbol{\varepsilon}$. This allows for the **distribution of demand responses** to vary across the income distribution in a unrestricted way.

For convenience we rewrite the support set as

$$\mathcal{S}_{\mathbf{p}_0,x_0} = \{ \mathbf{q} \in \mathcal{B}_{\mathbf{p}_0,x_0} | \bar{\mathbf{x}} - \mathbf{P}\mathbf{q} \leq 0 \},$$

where $\mathbf{P}$ is the matrix containing the observed prices and $\bar{\mathbf{x}}$ is the vector of intersection income levels,

$$\mathbf{P} = [\mathbf{p}(1), \cdots, \mathbf{p}(T)]' \in \mathbb{R}_+^{T \times 2}, \quad \bar{\mathbf{x}} = (\bar{x}(1), ..., \bar{x}(T))' \in \mathbb{R}_+^T.$$

We have here utilized that by definition $\mathbf{p}(t)' \mathbf{d}(\bar{x}(t), t, \tau) = \bar{x}(t)$. $\mathcal{S}_{\mathbf{p}_0,x_0}$ is the identified *set* of demand responses for $\mathbf{p}_0, x_0, \boldsymbol{\varepsilon}$, and can be written as a set of moment inequalities.

A central objective of this paper is to provide an estimator for this support set and to investigate its properties. As an initial ingredient for this estimator we first develop sieve estimators of the sequence of demand functions $\mathbf{d}(x, t, \varepsilon) = (d_1(x, t, \varepsilon), d_2(x, t, \varepsilon))$, $t = 1, ..., T$.

3 Unrestricted Sieve Estimator

As a starting point, we assume that for all $t = 1, ..., T$ and all $\tau \in [0, 1]$, the function $x \mapsto d_1(x, t, \tau)$ is situated in some known function space $\mathcal{D}_1$ which is equipped with some (pseudo-)norm $\|\cdot\|$.[6] We specify the precise form of $\mathcal{D}_1$ and $\|\cdot\|$ below. Given the function space $\mathcal{D}_1$, we choose sieve spaces

[6] The function space could without problems be allowed to depend on time, t, and the errors, τ. For notational simplicity, we maintain that the function space is the same across different values of (t, τ).

$\mathcal{D}_{1,n}$ that are finite-dimensional subsets of $\mathcal{D}$. In particular, we will assume that for any function $d_1 \in \mathcal{D}_1$, there exists a sequence $\pi_n d_1 \in \mathcal{D}_{1,n}$ such that $\|\pi_n d_1 - d_1\| \to 0$ as $n \to \infty$. Most standard choices of the function space $\mathcal{D}_1$ can be written on the form

$$\mathcal{D}_1 = \left\{ d_1 : d_1 (x, t, \tau) = \sum_{k \in \mathcal{K}} \pi_k (t, \tau) B_k (x), \quad \pi (t, \tau) \in \mathbb{R}^{|\mathcal{K}|} \right\},$$

for known (basis) functions $\{B_k\}_{k \in \mathcal{K}}$, and some (infinite-dimensional) index set $\mathcal{K}$; see Chen (2007, Section 2.3) for some standard specifications. A natural choice for sieve is then

$$\mathcal{D}_{1,n} = \left\{ d_{n,1} : d_{n,1} (x, t, \tau) = \sum_{k \in \mathcal{K}_n} \pi_k (t, \tau) B_k (x), \quad \pi (t, \tau) \in \mathbb{R}^{|\mathcal{K}_n|} \right\}, \tag{3}$$

for some sequence of (finite-dimensional) sets $\mathcal{K}_n \subseteq \mathcal{K}$. Finally, we define the space of vector functions,

$$\mathcal{D} = \left\{ \mathbf{d} = (d_1, d_2) : d_1 (x, t, \tau) \in \mathcal{D}_1, \quad d_2 (t, x, \tau) := \frac{x - p_1 (t) d_1 (x, t, \tau)}{p_2 (t)} \right\},$$

and with the associated sieve space $\mathcal{D}_n$ obtained by replacing $\mathcal{D}_1$ by $\mathcal{D}_{1,n}$ in the definition of $\mathcal{D}$.

Given the function space $\mathcal{D}$ and its associated sieve, we can construct a sieve estimator of the function $\mathbf{d} (\cdot, t, \tau)$. Given that $d_1 (x, t, \tau)$ is identified as a conditional quantile for any given value of x, c.f. eq. (2), we may employ standard quantile regression techniques to obtain the estimator: Let

$$\rho_\tau (z) = (\mathbb{I} \{z < 0\} - \tau) z, \quad \tau \in [0, 1],$$

be the standard check function used in quantile estimation (see Koenker and Bassett, 1978). We then propose to estimate $\mathbf{d} (x, t, \tau)$ by

$$\hat{\mathbf{d}} (\cdot, t, \tau) = \arg \min_{\mathbf{d}_n \in \mathcal{D}_n} \frac{1}{n} \sum_{i=1}^{n(t)} \rho_\tau (q_{1,i} (t) - d_{n,1} (x_i (t), t, \tau)), \tag{4}$$

for any $t = 1, ..., T$ and $\tau \in [0, 1]$.

The above estimator can be computed using standard numerical methods for linear quantile regressions when the sieve space is on the form in Eq. (3): Define $\underline{B} (x) = \{B_k (x) : k \in \mathcal{K}_n\} \in \mathbb{R}^{|\mathcal{K}_n|}$ as the collection of basis functions spanning the sieve $\mathcal{D}_{1,n}$. Then the sieve estimator is given by $\hat{d}_1 (x, t, \tau) = \hat{\pi} (t, \tau)' \underline{B} (x) = \sum_{k \in \mathcal{K}_n} \hat{\pi}_k (t, \tau) B_k (x)$, where

$$\hat{\pi} (t, \tau) = \arg \min_{\pi \in \mathbb{R}^{|\mathcal{K}_n|}} \frac{1}{n} \sum_{i=1}^{n} \rho_\tau \left(q_{1,i} (t) - \pi' \underline{B} (x_i (t)) \right), \quad \tau \in [0, 1]. \tag{5}$$

That is, the estimator $\hat{\pi} (t, \tau)$ is simply the solution to a standard linear quantile regression problem. Finally, the estimator of the demand function for the "residual" good is given by

$$\hat{d}_2 (x, t, \tau) = \frac{x - p_1 (t) \hat{d}_1 (x, t, \tau)}{p_2 (t)}. \tag{6}$$

To develop an asymptotic theory of the proposed sieve estimator, the following assumptions are imposed on the model:

A.1 Income $x(t)$ has bounded support, $x(t) \in \mathcal{X} = [a, b]$ for $-\infty < a < b < +\infty$, and is independent of $\varepsilon \sim U[0, 1]$.

A.2 The demand function $d_1(x, t, \varepsilon)$ is invertible in ε and is continuously differentiable in (x, ε).

These are fairly standard assumptions in the literature on nonparametric quantile estimation. It would be desirable to weaken the restriction of bounded support, but the cost would be more complicated conditions and proof so we maintain (A.1) (see e.g. Chen, Blundell and Kristensen, 2007 for results with unbounded support). The independence assumption rules out endogenous income; in Section 9, we argue how this can be allowed for by adopting nonparametric IV or control function approaches. We refer to Beckert (2007) and Beckert, and Blundell (2008) for more primitive conditions in terms of the underlying utility-maximization problem for (A.2) to hold.

We restrict our attention to the case where either B-splines or polynomial splines are used to construct the sieve space $\mathcal{D}_{1,n}$. For an introduction to these, we refer to Chen (2007, Section 2.3). All of the following results remain true with other linear sieve spaces after suitable modifications of the conditions. We introduce the following L_2- and sup-norms which will be used to state our convergence rate results:

$$||\mathbf{d}||_2 = \sqrt{E[\|\mathbf{d}(x, t, \tau)\|]}, \quad ||\mathbf{d}||_\infty = \sup_{x \in \mathcal{X}} \|\mathbf{d}(x, t, \tau)\|.$$

The function space $\mathcal{D}_1$ is then restricted to satisfy:

A.3 The function $d_1(\cdot, t, \tau) \in \mathcal{D}_1$, where $\mathcal{D}_1 = \mathcal{W}_2^m([a, b])$ and $\mathcal{W}_2^m([a, b])$ is the Sobolev space of all functions on $[a, b]$ with L_2-integrable derivatives up to order $m \geq 0$.

We now have the following result:

Theorem 1 *Assume that (A.1)-(A.3) hold. Then for any $t = 1, ..., T$ and $\tau \in [0, 1]$:*

$$||\hat{\mathbf{d}}(\cdot, t, \tau) - \mathbf{d}_0(\cdot, t, \tau)||_2 = O_P(\sqrt{k_n/n}) + O_P\left(k_n^{-m}\right),$$

while

$$||\hat{\mathbf{d}}(\cdot, t, \tau) - \mathbf{d}_0(\cdot, t, \tau)||_\infty = O_P(k_n/\sqrt{n}) + O_P\left(k_n^{-m}\right).$$

In particular, with $k_n = O\left(n^{1/(2m+1)}\right)$,

$$||\hat{\mathbf{d}}(\cdot, t, \tau) - \mathbf{d}_0(\cdot, t, \tau)||_2 = O_P\left(n^{-m/(2m+1)}\right),$$

while, with $k_n = O\left(n^{1/(2m+2)}\right)$,

$$||\hat{\mathbf{d}}(\cdot, t, \tau) - \mathbf{d}_0(\cdot, t, \tau)||_\infty = O_P(n^{-m/(2m+2)}).$$

We here state rate results both in the L_2- and sup-norm, and note that while we obtain optimal rates in the L_2-norm this is not the case in the sup-norm. This is a general problem for sieve estimators; see e.g. Newey (1997, Theorem 1) and Belloni et al (2010, Lemma 2.1. and Remark

2.1). However, the rate result in the sup-norm proves helpful when developing the asymptotic properties of the constrained demand function estimator and the demand bound estimator.

To establish the asymptotic distribution of our sieve estimator, we employ the results of Belloni et al (2010) who give general conditions for limiting distributions of sieve estimators. To state the asymptotic distribution, we need some additional notation: Define the covariance matrix

$$W_n(\tau) = \tau(1-\tau) E\left[f_\tau(0|x) \underline{B}_{k_n}(x) \underline{B}_{k_n}(x)'\right]^{-1} E\left[\underline{B}_{k_n}(x) \underline{B}_{k_n}(x)'\right] E\left[f_\tau(0|x) \underline{B}_{k_n}(x) \underline{B}_{k_n}(x)'\right]^{-1}, \tag{7}$$

where $f_\tau(e|x)$ is the conditional distribution of $e(\tau) := d_1(x, \varepsilon) - d_1(x, \tau)$. Note that $W_n(\tau)$ are the asymptotic variance of the estimated coefficients $\hat{\pi}_{k_n}(\tau)$ in the quantile regression model $q_{1,i}(t) = \pi_{k_n}(\tau)' \underline{B}_{k_n}(x) + e(\tau)$, c.f. Koenker and Bassett (1978). We are then able to state the following asymptotic normality result:

Theorem 2 *Assume that (A.1)-(A.3) hold; the smallest eigenvalue of* $E\left[\underline{B}_{k_n}(x) \underline{B}_{k_n}(x)'\right]$ *is bounded and bounded away from zero;* $k_n^4/n = O(1)$ *and* $k_n^{3m-1/2}/n = O(1)$. *Then for any* $x(t) \in \mathcal{X}$, $t = 1, ..., T$, *and* $\tau \in [0, 1]$,

$$\sqrt{n}\Sigma_n^{-1/2}(x, \tau) \begin{pmatrix} \hat{d}_1(x(1), 1, \tau) - d_{0,1}(x(1), 1, \tau) \\ \vdots \\ \hat{d}_1(x(T), T, \tau) - d_{0,1}(x(T), T, \tau) \end{pmatrix} \to^d N(0, I),$$

where $\Sigma_n(x, \tau) = \operatorname{diag}\{\Sigma_n(x(1), 1, \tau), ..., \Sigma_n(x(T), T, \tau)\} \in \mathbb{R}^{T \times T}$ *with*

$$\Sigma_n(x(t), t, \tau) = \underline{B}_{k_n}(x(t))' W_n(\tau) \underline{B}_{k_n}(x(t)) \in \mathbb{R}.$$

An attractive feature of the above result is that for inferential purposes we can treat the sieve estimator as simply a parametric estimator: As already noted, $W_n(\tau)$ is identical to the asymptotic covariance matrix of the estimated coefficients in a quantile regression setting. We then simply have to pre- and postmultiply this by $\underline{B}_{k_n}(x(t))$ to obtain the covariance matrix of the demand function itself.

A consistent estimator of the covariance matrix $\Omega_n(x, \tau)$ can be obtained by replacing $W_n(\tau)$ in the above expression by

$$\hat{W}_n(\tau) = \tau(1-\tau) \hat{E}_n\left[\hat{f}_\tau(0|x) \underline{B}_{k_n}(x) \underline{B}_{k_n}(x)'\right]^{-1} \hat{E}_n\left[\underline{B}_{k_n}(x) \underline{B}_{k_n}(x)'\right] E_n\left[\hat{f}_\tau(0|x) \underline{B}_{k_n}(x) \underline{B}_{k_n}(x)'\right]^{-1},$$

where $\hat{E}_n[g(x)] := n^{-1}\sum_{i=1}^n g(x_i)$ and $\hat{f}_\tau(0|x)$ is a kernel estimator of the conditional density. This estimator is on the same form as the one proposed in Powell (1986) for linear quantile regressions.

A similar distributional result holds for the demand function of the second good, except that the covariance matrix $\Sigma_n(x(t), t, \tau)$ has to be multiplied by $-p_1(t)/p_2(t)$; this follows by the delta method and eq. (6). We also note that the joint distribution of $(\hat{d}_1(x(t), t, \tau), \hat{d}_2(x(t), t, \tau))$ is degenerate due to the budget constraint, c.f. eq. (1).

The above weak convergence result only holds pointwise. As discussed in the following sections, uniform weak convergence results would be useful in order to extend some of our estimators and test statistics. We conjecture that these could be obtained by utilizing the results of Belloni and Chernozhukov (2010).

4 Revealed Preference (RP) Restricted Sieve Estimator

We now wish to impose revealed preferences (RP) restrictions on the function $\mathbf{d}$ in the estimation. Blundell, Browning and Crawford (2003, 2008) provide a more detailed introduction to the relevant revealed preference inequality restrictions and their empirical implications.

Consider a particular income expansion path $\{x(t)\}$ given by

$$x(t) = \mathbf{p}(t)' \mathbf{d}(x(t+1), t+1, \tau).$$

For any consumer chracterised by $\tau \in [0,1]$, the RP inequality restrictions imply:

$$\mathbf{p}(t)' \mathbf{d}(x(t), t, \tau) = x(t) \leq \mathbf{p}(t)' \mathbf{d}(x(s), s, \tau), \ \ s < t, \ t = 2, ..., T. \tag{8}$$

If the demand functions $\mathbf{d}(x, t, \tau)$, $t = 1, ..., T$, satisfy these inequalities for any given income level $x(T)$ we say that "$\mathbf{d}$ *satisfies RP*".

A RP-restricted sieve estimator is easily obtained in principle: First observe that the unrestricted estimator of $\{\mathbf{d}(\cdot, t, \tau)\}_{t=1}^{T}$ of the previous section can be expressed as the solution to the following joint estimation problem.

$$\{\hat{\mathbf{d}}(\cdot, t, \tau)\}_{t=1}^{T} = \arg \min_{\{\mathbf{d}_n(\cdot,t,\tau)\}_{t=1}^{T} \in \mathcal{D}_n^T} \frac{1}{n} \sum_{t=1}^{T} \sum_{i=1}^{n} \rho_\tau (q_{1,i}(t) - d_{1,n}(t, x_i(t))), \quad \tau \in [0,1],$$

where $\mathcal{D}_n^T = \otimes_{t=1}^{T} \mathcal{D}_n$ and $\mathcal{D}_n$ is defined in the previous section. Since there are no restrictions across goods and t, the above definition of $\{\hat{\mathbf{d}}(\cdot, t, \tau)\}_{t=1}^{T}$ is equivalent to the unrestricted estimators in eqs. (4) and (6). In order to impose the RP restrictions, we simply define the constrained function set as

$$\mathcal{D}_C^T := \mathcal{D}^T \cap \{\mathbf{d}(\cdot, \cdot, \tau) \ \text{ satisfies RP}\}, \tag{9}$$

and similarly the constrained sieve as

$$\mathcal{D}_{C,n}^T := \mathcal{D}_n^T \cap \{\mathbf{d}_n(\cdot, \cdot, \tau) \ \text{ satisfies RP}\}.$$

We define the constrained estimator by:

$$\{\hat{\mathbf{d}}_C(\cdot, t, \tau)\}_{t=1}^{T} = \arg \min_{\{\mathbf{d}_n(\cdot,t,\tau)\}_{t=1}^{T} \in \mathcal{D}_{C,n}^T} \frac{1}{n} \sum_{t=1}^{T} \sum_{i=1}^{n} \rho_\tau (q_{1,i}(t) - d_{1,n}(t, x_i(t))), \quad \tau \in [0,1]. \tag{10}$$

Since the RP inequalities impose restrictions across both goods ($l = 1, 2$) and time ($t = 1, ..., T$), the above estimation problem can no longer be split up into individual subproblems as in the unconstrained case.

The proposed estimator shares some similarities with the ones considered in, for example, Berestenau (2004), Gallant and Golub (1984), Mammen and Thomas-Agnan (1999) and Yatchew and Bos (1997) who also consider constrained sieve estimators. However, they focus on least-squares regression while ours is a least-absolute deviation estimator, and they furthermore restrict themselves to linear constraints. There are some results for estimation of monotone quantiles and other linear constraints, see Chernozhukov et al (2006), Koenker and Ng (2005) and Wright (1984), but again their constraints are simpler to analyze and implement. These two issues, a non-smooth criterion function and non-linear constraints, complicate the analysis and implementation of our estimator, and we cannot readily import results from the existing literature.

In order to derive the convergence rate of the constrained sieve estimator, we employ the same proof strategy as found elsewhere in the literature on nonparametric estimation under shape constraints, see e.g. Birke and Dette (2007), Mammen (1991), Mukerjee (1988): We first demonstrate that as $n \to \infty$, the unrestricted estimator, $\hat{\mathbf{d}}$, satisfies RP almost surely. This implies that $\{\hat{\mathbf{d}}(\cdot,t,\tau)\}_{t=1}^{T} \in \mathcal{D}_{C,n}^{T}$ with probability approaching one (w.p.a.1) which in turn means that $\hat{\mathbf{d}} = \hat{\mathbf{d}}_C$ w.p.a.1, since $\hat{\mathbf{d}}_C$ solves a constrained version of the minimization problem that $\hat{\mathbf{d}}$ is a solution to. We are now able to conclude that $\hat{\mathbf{d}}_C$ is asymptotically equivalent $\hat{\mathbf{d}}$, and all the asymptotic properties of $\hat{\mathbf{d}}$ are inherited by $\hat{\mathbf{d}}_C$.

For the above argument to go through, we need to slightly change the definition of the constrained estimator though. We introduce the following generalized version of RP: We say that "$\mathbf{d}$ satisfies RP(ϵ)" for some ("small") $\epsilon \geq 0$ if for any income expansion path,

$$x(t) \leq \mathbf{p}(t)' \mathbf{d}(x(s), s, \tau) + \epsilon, \ \ s < t, \ t = 2, ..., T.$$

The definition of RP(ϵ) is akin to Afriat (1973) who suggests a similar modification of (GA)RP to allow for waste ("partial efficiency"). We then define the constrained function space and its associated sieve as:

$$\mathcal{D}_C^T(\epsilon) = \mathcal{D}^T \cap \{\mathbf{d}(\cdot,\cdot,\tau) \text{ satisfies RP}(\epsilon)\},$$
$$\mathcal{D}_{C,n}^T(\epsilon) = \mathcal{D}_n^T \cap \{\mathbf{d}_n(\cdot,\cdot,\tau) \text{ satisfies RP}(\epsilon)\}.$$

We note that the constrained function space $\mathcal{D}_C^T$ as defined in eq. (9) satisfies $\mathcal{D}_C^T = \mathcal{D}_C^T(0)$. Moreover, it should be clear that $\mathcal{D}_C^T(0) \subset \mathcal{D}_C^T(\epsilon)$ and $\mathcal{D}_{C.n}^T(0) \subset \mathcal{D}_{C,n}^T(\epsilon)$ since RP(ϵ), $\epsilon > 0$, imposes weaker restrictions on the demand functions.

We now re-define our RP constrained estimators to solve the same optimization problem as before, but now the optimization takes place over $\mathcal{D}_{C,n}(\epsilon)$. We let $\hat{\mathbf{d}}_C^{\epsilon}$ denote this estimator, and note that $\hat{\mathbf{d}}_C^0 = \hat{\mathbf{d}}_C$, where $\hat{\mathbf{d}}_C$ is given in Eq. (10). The assumption that $\{\mathbf{d}_0(\cdot,t,\tau)\}_{t=1}^T \in \mathcal{D}_C^T(0)$ implies that $\{\hat{\mathbf{d}}(\cdot,t,\tau)\}_{t=1}^T \in \mathcal{D}_{C,n}^T(\epsilon)$ w.p.a.1. Since $\hat{\mathbf{d}}_C^{\epsilon}$ is a constrained version of $\hat{\mathbf{d}}$, this implies that $\hat{\mathbf{d}}_C^{\epsilon} = \hat{\mathbf{d}}$ w.p.a.1. Similar assumptions and proof strategies have been employed in Birke and Dette (2007) [Mammen (1991)]: They assume that the function being estimated is *strictly* convex [monotone], such that the unconstrained estimator is convex [monotone] w.p.a.1. Since $\mathcal{D}_{C,n}^T(0) \subset \mathcal{D}_{C,n}^T(\epsilon)$, our new estimator will in general be less precise than the one defined as the optimizer over $\mathcal{D}_{C,n}^T(0)$, but for small values of $\epsilon > 0$ the difference should be negligible.

Theorem 3 *Assume that (A.1)-(A.3) hold, and that* $\mathbf{d}_0 \in \mathcal{D}_C^T(0)$. *Then for any* $\epsilon > 0$:

$$||\hat{\mathbf{d}}_C^{\epsilon}(\cdot,t,\tau) - \mathbf{d}_0(\cdot,t,\tau)||_{\infty} = O_P(k_n/\sqrt{n}) + O_P\left(k_n^{-m}\right),$$

for $t = 1, ..., T$. *Moreover, under the conditions in Theorem 2, the restricted estimator has the same asymptotic distribution as the unrestricted estimator given in the same theorem.*

Again, the rate result is not minimax optimal, which is a consequence of that our proof relies on sup-norm convergence of the unrestricted estimator. We conjecture that the restricted estimator will exhibit optimal convergence in the L_2-norm, but have not been able to show this. In terms of convergence rate in the sup-norm, we are not able to show that the additional constraints arising from the RP restrictions lead to any improvements. This is similar to other results in the literature on constrained nonparametric estimation. Kiefer (1982) and Berestenau (2004) establish optimal nonparametric rates in the case of constrained densities and regression functions respectively when the constraints are not binding. In both cases, the optimal rate is the same as for the unconstrained one. However, as demonstrated both analytically and through simulations in Mammen (1991) for monotone restrictions (see also Berestenau, 2004 for simulation results for other restrictions), there may be significant finite-sample gains.

We conjecture that the above result will not in general hold for the estimator $\hat{\mathbf{d}}_C$ defined as the minimizer over $\mathcal{D}_{C,n}^T(0)$. In this case the GARP constraints would be binding, and we can no longer ensure that the unconstrained estimator is situated in the interior of the constrained function space. This in turn means that the unconstrained and constrained estimator most likely are not asymptotically equivalent and very different techniques have to be used to analyze the constrained estimator. In particular, the rate of convergence and/or asymptotic distribution of it would most likely be non-standard. This is, for example, demonstrated in Andrews (1999,2001), Anevski and Hössjer (2006) and Wright (1981) who give results for inequality-constrained parametric and nonparametric problems respectively.

Finally, we note that the above theorem is not specific to our particular quantile sieve estimator. One can by inspection easily see that the arguments employed in our proof can be carried over without any modifications to show that for any unconstrained demand function estimator, the corresponding RP-constrained estimator will be asymptotically equivalent.

5 Estimation of Demand Bounds

Once an estimator of the demand function has been obtained, either unrestricted or restricted, we can proceed to estimate the associated demand bounds. We will here utilize the machinery developed in Chernozhukov, Hong and Tamer (2007), henceforth CHT, and use their results to develop the asymptotic theory of the proposed demand bound estimators.

We have earlier Considered a particular consumer characterized by some $\tau \in [0,1]$ with associated sequence of demand functions $\mathbf{d}(x,t,\tau)$, $t = 1,...,T$. The consumer's budget set associated with new prices $\mathbf{p}_0$ was given by the compact and convex set

$$\mathcal{B}_{\mathbf{p}_0,x_0} = \left\{\mathbf{q} \in \mathbb{R}_+^2 | \mathbf{p}_0' \mathbf{q} = x_0\right\},$$

The demand support set can then be represented as follows:

$$\mathcal{S}_{\mathbf{p}_0,x_0} = \left\{\mathbf{q} \in \mathcal{B}_{\mathbf{p}_0,x_0} | \mathbf{p}(t)' \mathbf{q} \geq \mathbf{p}(t)' \mathbf{d}(\bar{x}(t), t, \tau), \quad t = 1,...,T\right\},$$

where $\{\bar{x}(t) : t = 1, ..., T\}$ is a sequence of intersection incomes defined as the solution to

$$\mathbf{p}_0' \mathbf{d}(\bar{x}(t), t, \tau) = x_0, \quad t = 1, ..., T.$$

A natural estimator of the support set would be to simply substitute the estimated intersection incomes for the unknown ones. Defining the estimated income levels $\hat{\mathbf{x}} = (\hat{x}(1), ..., \hat{x}(T))$ as the solutions to

$$\mathbf{p}_0' \hat{\mathbf{d}}_C(\hat{x}(t), t, \tau) = x_0, \quad t = 1, ..., T,$$

a natural support set estimator would appear to be $\hat{\mathcal{S}}_{\mathbf{p}_0, x_0} = \{\mathbf{q} \in \mathcal{B}_{\mathbf{p}_0, x_0} | \hat{\mathbf{x}} - \mathbf{P}\mathbf{q} \leq 0\}$. However, in order to do inference, in particular obtaining a valid confidence set for $\mathcal{S}_{\mathbf{p}_0, x_0}$, we need to modify this estimator.

First, as stated in Theorems 2-3, the sieve estimators of the demand functions may exhibit different convergence rates over time and income levels. As demonstrated in Appendix B, the estimated intersection income levels, $\hat{x}(t)$, $t = 1, ..., T$, inherit this property,

$$\sqrt{n} W_n^{1/2} (\hat{\mathbf{x}} - \bar{\mathbf{x}}) \to^d N(0, I_T),$$

where I_T denotes the T-dimensional identity matrix, and W_n is a diagonal matrix,

$$W_n = \text{diag}\{w_n(1), ..., w_n(T)\},$$

with positive entries given by

$$w_n(t) = \left[\frac{\mathbf{p}_0' \frac{\partial \mathbf{d}(\bar{x}(t), t)}{\partial x}}{p_{0,1} - \frac{p_{0,2}}{p_2(t)} p_1(t)} \right]^2 \Sigma_n^{-1}(\bar{x}(t), t),$$

where $\Sigma_n(x, t, \tau)$ is the variance component of $\hat{d}_1(x, t, \tau)$ in Theorem 2. Due to the heterogenous normalizations across $t = 1, ..., T$, as described by the weighting matrix W_n, the T inequality constraints that make up the support set are potentially estimated with different rates. Unless this is taken into account in the estimation, we are not able to construct confidence sets. We therefore introduce a sample objective function $Q_n(\mathbf{q})$ that contain normalized versions of the estimated demand bounds,

$$Q_n(\mathbf{q}) = \left\| \hat{W}_n^{1/2} [\hat{\mathbf{x}} - \mathbf{P}\mathbf{q}] \right\|_+^2,$$

where $\|y\|_+ = \|\max\{y, 0\}\|$ for any vector y, and $\hat{W}_n = \text{diag}\{\hat{w}_n(1), ..., \hat{w}_n(T)\}$ being a consistent estimator of W_n. In comparison to the naive estimator suggested earlier, we now normalize $\hat{\mathbf{x}} - \mathbf{P}\mathbf{q}$ with $W_n^{1/2}$. If we could have shown that the intersection incomes converged with same rate (for example, if we could show that $\Sigma_n(x, t) = r_n \Sigma(x, t)$ for some sequence r_n) this normalization would not be required. Given that $\hat{\mathbf{x}}$ in addition is a consistent estimator of $\bar{\mathbf{x}}$, it is straightforward to verify that "the limit" of $Q_n(\mathbf{q})$ is given by

$$\bar{Q}_n(\mathbf{q}) = \left\| W_n^{1/2} [\bar{\mathbf{x}} - \mathbf{P}\mathbf{q}] \right\|_+^2.$$

An important point here is that even though the nonstochastic function $\bar{Q}_n(\mathbf{q})$ depends on $n \geq 1$ through W_n, it still gives a precise characterization of the support set $\mathcal{S}_{\mathbf{p}_0,x_0}$ for any given $n \geq 1$:

$$\bar{Q}_n(\mathbf{q}) = 0 \Leftrightarrow W_n^{1/2}[\bar{\mathbf{x}} - \mathbf{P}\mathbf{q}] \leq 0 \Leftrightarrow \bar{\mathbf{x}} - \mathbf{P}\mathbf{q} \leq 0 \Leftrightarrow \mathbf{q} \in \mathcal{S}_{\mathbf{p}_0,x_0},$$

where the second equivalence follows from the fact that W_n is a diagonal matrix with positive elements.

In addition to the normalizing weights, we also introduce a slackness variable to control for boundary issues. Let $c_n \geq 0$ be some positive sequence, which will be further restricted in the following. We then define our support set estimator as

$$\hat{\mathcal{S}}_{\mathbf{p}_0,x_0}(c_n) = \left\{\mathbf{q} \in \mathcal{B}_{\mathbf{p}_0,x_0} \left| \sqrt{n} Q_n(\mathbf{q}) \leq c_n \right.\right\}.$$

The resulting support set estimator is given as the demand levels that lie within a given contour level c_n of the sample objective function $Q_n(\mathbf{q})$.

It is worth noting that the above formulation of the support set and its estimator in terms of $\bar{Q}_n(\mathbf{q})$ and $Q_n(\mathbf{q})$ is very close to the general formulation of set estimators defined through moment inequalities used in CHT. However, in their setting the limiting objective function, in our case $\bar{Q}_n(\mathbf{q})$, is not allowed to depend on n, so we cannot directly apply their results. However, their proof strategy fortunately carries over to our case without much additional work. This is similar to the extension of standard proofs of consistency and rate results in the point identified case to allow for a sequence of limiting objective functions; see e.g. White (1994).

In order to analyze the set estimator we impose the following conditions on the demand functions and observed prices which together defines the support set:

A.4 $x(t) \mapsto \mathbf{d}(x(t), t, \tau)$ is strictly monotonically increasing, $t = 1, ..., T$.

A.5 The matrix $\mathbf{P} = [\mathbf{p}(1), \cdots, \mathbf{p}(T)]' \in \mathbb{R}_+^{T \times 2}$ has rank 2.

The monotonicity requirement in Condition (A.4) ensures that the intersection income path $\{\bar{x}(t)\}$ is uniquely defined, and is a standard requirement in consumer demand theory. Condition (A.5) states that the observed prices have exhibited sufficient variation so we can distinguish between different demands. In particular, we need at least to have observed at least two prices and furthermore that at least two of these prices cannot be expressed as linear combinations of others.

To state rate results for our support set estimator, we introduce the so-called Hausmann norm which is given by:

$$d_{\mathrm{H}}(\mathcal{A}_1, \mathcal{A}_2) = \max\left\{\sup_{y \in \mathcal{A}_1} \rho(y, \mathcal{A}_2), \sup_{y \in \mathcal{A}_2} \rho(y, \mathcal{A}_1)\right\}, \quad \rho(y, \mathcal{A}) = \inf_{x \in \mathcal{A}} \|x - y\|,$$

for any two sets $\mathcal{A}_1, \mathcal{A}_2$. The following theorem now follows as a straight foward implication of the more general result found in Appendix B:

Theorem 4 *Suppose that (A.1)-(A.5) hold and* $\left\|\hat{W}_n - W_n\right\| \to^P 0$*. Then for any sequence* $c_n \propto \log(n)$*,*

$$d_{\mathrm{H}}(\hat{\mathcal{S}}_{\mathbf{p}_0,x_0}(c_n), \mathcal{S}_{\mathbf{p}_0,x_0}) = O_P(k_n\sqrt{\log(n)/n}) + O_P\left(\log(n)\, k_n^{-m}\right)).$$

If furthermore, the smallest eigenvalue of $E\left[\underline{B}_{k_n}(x)\,\underline{B}_{k_n}(x)'\right]$ *is bounded and bounded away from zero;* $k_n^4/n = O(1)$ *and* $k_n^{3m-1/2}/n = O(1)$*, then:*

$$P(\mathcal{S}_{\mathbf{p}_0,x_0} \subseteq \hat{\mathcal{S}}_{\mathbf{p}_0,x_0}(\hat{c}_n)) \to 1-\alpha,$$

where $\hat{c}_n = \hat{q}_{1-\alpha} + O_P(\log(n))$ *and* $\hat{q}_{1-\alpha}$ *is an estimator of the* $(1-\alpha)$*th quantile of* $\mathcal{C}_{\mathbf{p}_0,x_0}$ *given by*

$$\mathcal{C}_{\mathbf{p}_0,x_0} := \sup_{\mathbf{q}\in\mathcal{S}_{\mathbf{p}_0,x_0}} \|Z + \xi(\mathbf{q})\|_+^2.$$

Here, $Z \sim N(0, I_T)$ *while* $\xi(\mathbf{q}) = (\xi_1(\mathbf{q}), ..., \xi_T(\mathbf{q}))'$ *is given by*

$$\xi_t(\mathbf{q}) = \begin{cases} -\infty, & \mathbf{p}(t)'\mathbf{q} > \mathbf{p}(t)'\mathbf{d}(x(t),t) \\ 0, & \mathbf{p}(t)'\mathbf{q} = \mathbf{p}(t)'\mathbf{d}(x(t),t) \end{cases}, \quad t = 1, ..., T.$$

The first part of the theorem shows that the support set estimator inherits the sup-norm convergence rate of the underlying demand function estimator. The second part shows how a valid confidence set can be constructed for the demand bounds, and is akin to the result found in, for example, CHT's Theorem 5.2. In order to employ the above result in practice, we need to be able to obtain estimators of the quantiles of the random variable $\mathcal{C}_{\mathbf{p}_0,x_0}$ defined in the theorem. The distribution of $\mathcal{C}_{\mathbf{p}_0,x_0}$ is non-standard and cannot be written on closed form, so evaluation of its quantiles has to be done either through simulations or resampling methods. In the latter case, one can either use the modified bootstrap procedures developed in Bugni (2009,2010) and Andrews and Soares (2010) or the subsampling procedure of CHT. We discuss in more detail in Section 8 how the bootstrap procedure of Bugni (2009,2010) can be implemented in our setting.

[The above theorem does not utilize that potentially our estimator has the degeneracy property discussed in, for example, CHT, Section 3.2 and 4.2. We need to add some results for when this is the case since this will improve on the results stated in the above theorem.]

6 Testing for Rationality

In the previous two sections, we have developed estimators of the demand responses under revealed preferences constraints. It is of interest to test whether the consumers in the data set indeed do satisfy these restrictions: First, from an economic point of view it is highly relevant to test the axioms underlying standard choice theory. Second, from an econometric point of view, we wish to test whether the imposed constraints are actually satisfied in data.

We here develop a test for whether the consumers satisfy the revealed preference aximom; that is, are they rational? A natural way of testing this hypothesis would be to compare the unrestricted and restricted demand function estimates, and rejecting if they are "too different" from each other.

Unfortunately, since we have only been able to develop the asymptotic properties of the constrained estimator under the hypothesis that none of the inequalities are binding, the unrestricted and restricted estimators are asymptotically equivalent under the null. Thus, any reasonable test comparing the two estimates would have a degenerate distribution under the null. Instead, we take the same approach as in Blundell et al (2008) and develop a minimum-distance statistic based on the unrestricted estimator alone that has a non-degenerate asymptotic distribution.

For given set of prices $\mathbf{p}_0 = (p_{0,1}, p_{0,2})'$ and income level x_0, define the associated intersection demands,

$$\bar{\mathbf{q}}(t) := \mathbf{d}(\bar{x}(t), t, \tau), \tag{11}$$

where $\{\bar{x}(t) : t = 1, ..., T\}$ are the intersection income levels solving

$$\mathbf{p}_0'\mathbf{d}(\bar{x}(t), t, \tau) = x_0, \quad t = 1, ..., T. \tag{12}$$

We collect the sequence of intersection demands in $\bar{\mathbf{q}} = (\bar{\mathbf{q}}(1), ..., \bar{\mathbf{q}}(T))$. A necessary condition for the demand function to be consistent with rational behaviour is that $\bar{\mathbf{q}}$ lies in the set of demand sequences that satisfy GARP which we denote $\mathbb{S}_{\mathbf{p}_0,x_0}$. As demonstrated in Blundell et al (2008), this set can be written as:

$$\mathbb{S}_{\mathbf{p}_0,x_0} = \left\{ \mathbf{q} \in \mathcal{B}^T_{\mathbf{p}_0,x_0} \middle| \exists V > 0, \lambda \geq 1 : V(t) - V(s) \geq \lambda(t)\, \mathbf{p}(t)' (\mathbf{q}(s) - \mathbf{q}(t)), \quad 1 \leq s, t \leq T \right\}, \tag{13}$$

where $\mathcal{B}_{\mathbf{p}_0,x_0}$ is the budget set defined in the previous section. So for a given $\mathbf{q} \in \mathbb{S}_{\mathbf{p}_0,x_0}$, there exists corresponding vectors $V(\mathbf{q}) = (V(\mathbf{q}, 1), ..., V(\mathbf{q}, T))'$ and $\lambda(\mathbf{q}) = (\lambda(\mathbf{q}, 1), ..., \lambda(\mathbf{q}, T))'$ such that

$$V(\mathbf{q}, t) - V(\mathbf{q}, s) \geq \lambda(\mathbf{q}, t) \mathbf{p}(t)' (\mathbf{q}(s) - \mathbf{q}(t)), \quad 1 \leq s, t \leq T.$$

One can interpret $V(\mathbf{q}, t)$ and $\lambda(\mathbf{q}, t)$ as the utility level and marginal utility corresponding to the demand vector $\mathbf{q}(t)$. The existence of such rationalizes the demand vector. The null hypothesis of interest is therefore that the sequence of intersection demands generated by the demand functions $\mathbf{d}(\cdot, t, \tau)$, $1 \leq t \leq T$, lies in this set:

$$H_0 : \bar{\mathbf{q}} \in \mathbb{S}_{\mathbf{p}_0,x_0}.$$

To test H_0, we first obtain estimates of the intersection demands $\bar{\mathbf{q}}(t)$, $1 \leq t \leq T$. Given the *unrestricted* demand function estimator $\hat{\mathbf{d}}$, these can be estimated by

$$\hat{\mathbf{q}}(t) = \hat{\mathbf{q}}(\mathbf{p}_0, x_0, t, \tau) := \hat{\mathbf{d}}(\hat{x}(t), t, \tau) \in \mathbb{R}^2, \tag{14}$$

where $\{\hat{x}(t)\}$ are the estimated intersection income levels solving

$$\mathbf{p}_0'\hat{\mathbf{d}}(\hat{x}(t), t, \tau) = x_0, \quad t = 1, ..., T. \tag{15}$$

A natural way to test H_0 is then to examine how far away the estimated intersection demands are from $\mathbb{S}_{\mathbf{p}_0,x_0}$ through some distance measure. A natural choice is

$$\rho_n(\hat{\mathbf{q}}, \mathbb{S}_{\mathbf{p}_0,x_0}) := \inf_{\mathbf{q} \in \mathbb{S}_{\mathbf{p}_0,x_0}} \|\hat{\mathbf{q}} - \mathbf{q}\|^2_{\hat{W}_n^{\text{test}}},$$

where $\hat{W}_n^{\text{test}} = (\hat{W}_n^{\text{test}}(1), ..., \hat{W}_n^{\text{test}}(T))$, is a sequence of positive weighting matrices and $\|\cdot\|_{\hat{W}_n^{\text{test}}}$ is a weighted Euclidean norm,

$$\|\hat{\mathbf{q}} - \mathbf{q}\|^2_{\hat{W}_n^{\text{test}}} = \sum_{t=1}^{T} (\hat{\mathbf{q}}(t) - \mathbf{q}(t))' \hat{W}_n^{\text{test}}(t) (\hat{\mathbf{q}}(t) - \mathbf{q}(t)).$$

The weighting matrices will be specified in the following and is chosen so the test statistic has a well-defined limiting distribution. The intuition behind the test statistic is that if indeed the consumer is rational, then $\hat{\mathbf{q}}$ in the limit will be situated in $\mathbb{S}_{\mathbf{p}_0,x_0}$ and as such $\rho(\hat{\mathbf{q}},\mathbb{S}_{\mathbf{p}_0,x_0}) \to^P 0$. Conversely, if the consumer is irrational, then $\lim_P \rho(\hat{\mathbf{q}},\mathbb{S}_{\mathbf{p}_0,x_0}) \neq 0$. Thus, we will in general accept (reject) H_0 if $\rho(\hat{\mathbf{q}},\mathbb{S}_{\mathbf{p}_0,x_0})$ is small (large). We note that since $\mathbb{S}_{\mathbf{p}_0,x_0}$ is a compact, convex set and $\mathbf{q}_1 \mapsto \|\hat{\mathbf{q}} - \mathbf{q}\|^2_{\hat{W}_{\text{test}}}$ is a convex function, there exists a unique projection of $\hat{\mathbf{q}}$ onto $\mathbb{S}_{\mathbf{p}_0,x_0}$,

$$\hat{\mathbf{q}}^* = \arg \min_{\mathbf{q} \in \mathbb{S}_{\mathbf{p}_0,x_0}} \|\hat{\mathbf{q}} - \mathbf{q}\|^2_{\hat{W}_n^{\text{test}}},$$

and we have $\rho_n(\hat{\mathbf{q}},\mathbb{S}_{\mathbf{p}_0,x_0}) = \|\hat{\mathbf{q}} - \hat{\mathbf{q}}^*\|^2_{\hat{W}_n^{\text{test}}}$.

The asymptotics of $\rho_n(\hat{\mathbf{q}},\mathbb{S}_{\mathbf{p}_0,x_0})$ under the null hypothesis of rationality H_0 are non-standard due to the fact that under the null $\bar{\mathbf{q}} =^P \lim_{n\to\infty} \hat{\mathbf{q}}$ may be situated on the boundary of $\mathbb{S}_{\mathbf{p}_0,x_0}$. Thus, the problem falls within the framework of Andrews (1999, 2001) who consider estimation and testing of a parameter on the boundary of a (restricted) parameter space. We employ his general results to derive the asymptotic distribution of $d(\hat{\mathbf{q}},\mathbb{S}_{\mathbf{p}_0,x_0})$. In the Appendix, we demonstrate that there exists well-defined versions of the mappings $V^*(\mathbf{q}) = (V^*(\mathbf{q},1), ..., V^*(\mathbf{q},1))'$ and $\lambda^*(\mathbf{q}) = (\lambda^*(\mathbf{q},1), ..., \lambda^*(\mathbf{q},T))'$ taking a given demand vector into corresponding utility levels and marginal utilities which are differentiable w.r.t. $\mathbf{q}$. Introducing the vector function $\mathbf{C}_1(\mathbf{q}) = (\mathbf{C}_1(\mathbf{q})', \mathbf{C}_1(\mathbf{q})', \mathbf{C}_1(\mathbf{q})')'$ where $\mathbf{C}_1(\mathbf{q}) = \{C_1(\mathbf{q},s,t)\}_{1\leq s,t\leq T} \in \mathbb{R}^{2T}$, $\mathbf{C}_2(\mathbf{q}) = \{C_2(\mathbf{q},t)\}_{1\leq t\leq T} \in \mathbb{R}^T$ and $\mathbf{C}_3(\mathbf{q}) = \{C_3(\mathbf{q},t)\}_{1\leq t\leq T} \in \mathbb{R}^T$ and

$$\begin{aligned} C_1(\mathbf{q},s,t) &:= V^*(\mathbf{q},s) - V^*(\mathbf{q},t) + \lambda^*(\mathbf{q},t)\mathbf{p}(t)'(\mathbf{q}(s) - \mathbf{q}(t)), \\ C_2(\mathbf{q},t) &:= \mathbf{p}_0\mathbf{q}(t) - x_0 \\ C_3(\mathbf{q},t) &:= -\mathbf{q}(t), \end{aligned} \tag{16}$$

we can rewrite the set of intersection demands that satisfy GARP as

$$\mathbb{S}_{\mathbf{p}_0,x_0} = \{\mathbf{q} \in \mathbb{R}^{2T} | \mathbf{C}(\mathbf{q}) \leq 0\},$$

and define the associated cone $\Lambda_{\mathbf{p}_0,x_0}$ by

$$\Lambda_{\mathbf{p}_0,x_0} = \left\{ v \in \mathbb{R}^T : \frac{\partial \mathbf{C}(\bar{\mathbf{q}})}{\partial \mathbf{q}} (A \otimes I_T) v \leq 0 \right\}, \tag{17}$$

where $A = [1, p_{0,2}/p_{0,1}]'$. The asymptotic distribution of the test statistic will then be governed by the shape of this cone. Note that while $\mathbb{S}_{\mathbf{p}_0,x_0}$ lies in $\mathbb{R}^{2T}$, the cone $\Lambda_{\mathbf{p}_0,x_0}$ governing the distribution of the test statistic is situated in the lower-dimensional space $\mathbb{R}^T$. This is due to the fact that due to the budget constraint $\mathbb{S}_{\mathbf{p}_0,x_0}$ is a T-dimensional subspace in $\mathbb{R}^{2T}$. As a consequence, even though the null hypothesis involves $2T$-dimensional restrictions on $\mathbf{q}$, the limiting distribution of the test statistic only involves a T-dimensional distribution as shown below.

As was the case for the support set estimator, care has to be taken in the choice of the weighting matrix defining the test statistic $\rho_n\left(\hat{\mathbf{q}},\mathbb{S}_{\mathbf{p}_0,x_0}\right)$ in order for it to have a well-defined limiting distribution. We here choose the sequence $\hat{W}_n^{\text{test}}(t)$ as a consistent estimator of

$$W_n^{\text{test}}(t)=\frac{n\Sigma_n\left(\bar{x}(t),t\right)}{\|A\|^2\|\Psi(t)\|}AA'\in\mathbb{R}^{(L+1)\times(L+1)}, \tag{18}$$

where

$$\Psi(t):=1+\frac{d_1\left(\bar{x}(t),t\right)}{\partial x(t)}\left[\mathbf{p}_0'\frac{\partial\mathbf{d}(\bar{x}(t),t)}{\partial x}\right]^{-1}\left[p_{0,1}-\frac{p_{0,2}}{p_L(t)}p_1(t)\right],$$

for $1\le t\le T$. The following result then holds:

Theorem 5 *Assume that (A.1)-(A.5) hold,* $\left\|\hat{\Sigma}_n(x,\tau)-\Sigma_n(x,\tau)\right\|\to^P 0$*, and* $\hat{\Psi}(t)\to^P\Psi(t)$ *for* $1\le t\le T$*. Then*

$$\rho_n\left(\hat{\mathbf{q}},\mathbb{S}_{\mathbf{p}_0,x_0}\right)\to^d\rho\left(Z,\Lambda_{\mathbf{p}_0,x_0}\right):=\inf_{\lambda\in\Lambda_{\mathbf{p}_0,x_0}}\|\lambda-Z\|^2,$$

where $\Lambda_{\mathbf{p}_0,x_0}$ *is given in Eq. (17) and* $Z\sim N(0,I_T)$*.*

The distributions of the estimator and the test statistic are non-standard. Andrews, (1999, Theorem 5) shows that the asymptotic distribution $\lambda^*=\arg\inf_{\lambda\in\Lambda_{\mathbf{p}_0,x_0}}\|\lambda-Z\|^2$ can be written as a linear projection of Z. Alternatively, it may be simulated. It may also be of interest to draw inference regarding the constrained set of demands, $\hat{\mathbf{q}}^*$. Applying Andrews (1999, Theorem 3) we obtain that $\sqrt{r_n}(\hat{\mathbf{q}}^*-\bar{\mathbf{q}}^*)\to^d\lambda^*$ where $\bar{\mathbf{q}}^*:=\arg\min_{\mathbf{q}\in\mathbb{S}_{\mathbf{p}_0,x_0}}\|\bar{\mathbf{q}}-\mathbf{q}\|_\Lambda^2$.

The proposed test only examines rationality for a particular income level, x_0, and set of prices, $\mathbf{p}_0$. A stronger test should examine rationality uniformly over incomes and prices; a natural test statistic would be $\sup_{\mathbf{p}_0,x_0}d\left(\hat{\mathbf{q}},\mathbb{S}_{\mathbf{p}_0,x_0}\right)$ where the sup is taken over some compact set. The asymptotic analysis of this test statistic is outside the scope of this study and is instead left for future research.

7 Practical Implementation

In this section, we discuss in further detail how the unconstrained and constrained estimators can be implemented. In the following, we again suppress the dependence on τ since this is kept fixed throughout. Let $\mathbf{d}_n\in\mathcal{D}_n^T$ be some given function in the sieve space. This function can be represented by its corresponding set of basis function coefficients, $\pi=\left[\pi(1)',...,\pi(T)'\right]'\in\mathbb{R}^{|\mathcal{K}_n|T}$ be a given set of parameter values.

Also, choose (a large number of) M income "termination" values, $x_m^*(T)$, $m=1,...,M$. The latter will be used to generated income paths. The idea is that as $M\to\infty$, we cover all possible income paths in the limit.

We then first compute M SMP paths $\{x_m^*(t)\}$, $m=1,...,M$:

$$x_m^*(t)=\mathbf{p}(t)'\,\mathbf{d}_n(x_m^*(t+1),t+1), \tag{19}$$

where

$$d_{1,n}(x,t) = \pi(t)' B(x), \quad d_{2,n}(x,t) = \frac{x - p_1(t)\, d_{1,n}(x,t)}{p_2(t)}.$$

Note that $x^*(t)$ implicitly depends on π. For any of these paths, say, $\{x^*(t)\}$, we can rewrite the restriction in Eq. (19) as:

$$a(s,t,\pi)\,\pi(s) \leq b(s,t,\pi), \quad s < t,$$

where

$$\begin{aligned} a(s,t,\pi) &= \left\{ \frac{p_2(t)}{p_2(s)} p_1(s) - p_1(t) \right\} B(x^*(s))' \in \mathbb{R}^{|\mathcal{K}_n|}, \\ b(s,t,\pi) &= \frac{p_2(t)}{p_2(s)} x^*(s) - x^*(t) \in \mathbb{R}, \end{aligned} \tag{20}$$

for $s < t$. For the given set of M income paths, we collect these inequalities and write them on matrix form,

$$A(\pi)\,\pi \leq b(\pi),$$

where

$$A(\pi) = \left[O_{1\times(s-1)|\mathcal{K}_n|}, a_m(s,t,\pi), O_{1\times(T-s)|\mathcal{K}_n|}, \right]_{m=1,\dots,M,s<t}, \quad b = \left[b_m(s,t,\pi) \right]_{m=1,\dots,M,s<t}.$$

Here, $a_m(s,t,\pi)$ and $b_m(s,t,\pi)$ denote the coefficients in Eq. (20) associated with the income path $\{x_m^*(t)\}$, $m = 1, ..., M$. For example, with $T = 3$ and $M = 1$, we have

$$\begin{bmatrix} a(1,2,\pi) & O_{1\times|\mathcal{K}_n|} & O_{1\times|\mathcal{K}_n|} \\ a(1,3,\pi) & O_{1\times|\mathcal{K}_n|} & O_{1\times|\mathcal{K}_n|} \\ O_{1\times|\mathcal{K}_n|} & a(2,3,\pi) & O_{1\times|\mathcal{K}_n|} \end{bmatrix} \begin{bmatrix} \pi(1) \\ \pi(2) \\ \pi(3) \end{bmatrix} \leq \begin{bmatrix} b(1,2,\pi) \\ b(1,3,\pi) \\ b(2,3,\pi) \end{bmatrix}.$$

We now see that the original estimation problem is an inequality constrained quantile estimation problem:

$$\hat{\pi}_C = \arg\min_{\pi} \frac{1}{n} \sum_{t=1}^{T} \sum_{i=1}^{n} \rho_\tau \left(q_{1,i}(t) - \pi(t)' W_i(t) \right) \qquad \text{s.t. } A(\pi)\,\pi \leq b(\pi). \tag{21}$$

Unfortunately, $A(\pi)$ and $b(\pi)$ both depend on π; otherwise, the estimator would be a simple linearly constrained quantile estimator as discussed in Koenker and Ng (2005).

In some cases, it may be simpler to compute an least-squares projection estimator instead:

$$\hat{\pi}_C = \min_{\pi} \|\hat{\pi} - \pi\| \text{ s.t. } A(\pi)\,\pi \leq b(\pi), \tag{22}$$

where $\hat{\pi}$ is the unconstrained estimator of the coefficients given in Eq. (5).

8 Bootstrap Inference

We here propose to employ bootstrap procedures to compute confidence regions and critical values of the estimators and tests developed in the previous sections. Since the estimators and test statistics are non-standard in the sense that they suffer from boundary problems (binding constraints) and/or the population parameters are not point identified, standard bootstrap procedures will not be valid. Instead, we base our proposed bootstrap procedures on the ideas developed in Bugni (2009,2010) in the context of moment inequalities. Both estimation problems in Sections 5 and 6 can be expressed as a set of moment inequalities and as such fit into the framework of Bugni (2009,2010).

As an alternative to the proposed bootstrap procedure, one can employ "plug-in" methods, where nuisance parameters appearing in the relevant asymptotic distributions are estimated from the sample such that one can directly evaluate quantiles from the "estimated" asymptotic distribution (using simulations). These can then be used to obtain confidence regions for the population parameters of interest.

Below, we demonstrate that each of our estimators and test statistics fit into the framework of Bugni (2009,2010), and describe how his bootstrap procedure can be used in our setting. First, we briefly summarize his procedure: Suppose we are given a set of "moments" , $m(\theta) \in \mathbb{R}^q$, that defines the set of parameters through equality constraints,

$$\Theta_I := \{\theta \in \Theta : m(\theta) \leq 0\}.$$

This set may be a singleton such that we have point identification. We have at our disposal a sample estimator of $m(\theta)$, say $\hat{m}(\theta) \in \mathbb{R}^q$, such that $\{\sqrt{n}(\hat{m}(\theta) - m(\theta)) : \theta \in \Theta\}$ has a well-defined, tight weak limit. This is then used to define the estimator of Θ_I by

$$\hat{\Theta}_I(c_n) := \left\{\theta \in \Theta : \|\hat{m}(\theta)\|_+^2 \leq c_n/n\right\},$$

where c_n is a slack variable satisfying $c_n/n \to 0$ and $\log\log(n)/c_n \to 0$. Bugni (20010) then proposes the following bootstrap procedure given our estimator $\hat{\Theta}_I(c_n)$:

1. For $b = 1, ..., B$: Draw a bootstrap sample with replacement from the data and compute the moment estimator based on the bootstrap sample, $m_b^*(\theta)$.

2. Compute

$$e_{b,i}^*(\theta) := \sqrt{n}(\hat{m}_{b,i}^*(\theta) - \hat{m}_i(\theta)) \times \mathbb{I}\left\{|\hat{m}_i(\theta)| \leq c_n/\sqrt{n}\right\}, \quad , i = 1, ..., q,$$

   and

$$\Gamma_b^* := \sup_{\theta \in \hat{\Theta}_I(c_n)} \|e_b^*(\theta)\|_+^2.$$

The empirical $(1-\alpha)$ quantile of $\{\Gamma_b^* : b = 1, ..., B\}$, $\hat{c}_{1-\alpha}$, is then used to estimate the $(1-\alpha)$ quantile of the statistic $\sup_{\theta \in \Theta_I} G(\hat{m}(\theta))$. Moreover, the $(1-\alpha)$ confidence set of Θ_I is estimated by

$$\hat{\Theta}_I(1-\alpha) = \left\{\theta \in \Theta : \|\hat{m}(\theta)\|_+ \leq \hat{c}_{1-\alpha}/n\right\}.$$

Bootstrapping RP-Restricted Demand Estimates: We here wish to bootstrap the constrained version of the demand function, $\mathbf{d}(x,t,\tau)$ for a given value of (x,t,τ). We focus on the restricted least-squares estimator given in eq. (22). Let $\pi(t,\tau)$ be the set of coefficients corresponding to a given demand function $\mathbf{d}(x)$ situated in the function space $\mathcal{D}$. With $\hat{d}_1(x,t,\tau)$ denoting the unrestricted estimator

$$\hat{m}_1(d_1) = \mathbf{d}(x,t,\tau) - \hat{\mathbf{d}}(x,t,\tau), \quad \hat{m}_2(\mathbf{d}) = \hat{\mathbf{d}}(x,t,\tau) - \mathbf{d}(x,t,\tau), \quad \hat{m}_3(\mathbf{d} = A(\pi(\mathbf{d}))\pi(\mathbf{d}) - b(\pi(\mathbf{d})),$$

where $A(\cdot)$ and $b(\cdot)$ were defined in Section 7. We now have that identified set is the singleton $\mathcal{D}_I = \{\mathbf{d}_0(x,t,\tau)\}$ and its estimator is the singleton $\hat{\mathcal{D}}_I(\epsilon) = \{\hat{\mathbf{d}}_C^{\epsilon}(x,t,\tau)\}$. We define the confidence set $\hat{\mathcal{D}}_I(c_n) = \{\pi : \left\|\Sigma_n^{-1/2}(x,t,\tau)\hat{m}(\mathbf{d})\right\|_+^2 \leq c_n/n\}$, and propose to use the following bootstrap procedure to obtain a consistent estimator of c_n.

1. Draw a bootstrap sample with replacement from the data,

$$\{Z_i^*(t) = (y_i^*(t), x_i^*(t)) : i = 1,...,n\},$$

   for $1 \leq t \leq T$. Compute the unrestricted demand function estimator based on the bootstrap sample, say $\hat{\mathbf{d}}^*$.

2. With $\hat{m}^*(\mathbf{d})$ being the moments with $\hat{\mathbf{d}}$ replaced by $\hat{\mathbf{d}}^*$,

$$e_{b,i}^*(d) := \sqrt{n}\Sigma_n^{-1/2}(x,t,\tau)(m_i^*(\mathbf{d}) - m_i(\mathbf{d})) \times \mathbb{I}\left\{\left|\Sigma_n^{-1/2}\hat{m}_i(\mathbf{d})\right| \leq c_n/\sqrt{r_n}\right\}, \quad , t = 1,...,T,$$

   and

$$\Gamma_b^* := \sup_{d \in \hat{\mathcal{D}}_I(c_n)} G(e_b^*(d)).$$

We then compute the confidence region of the demand function by

$$\hat{\mathcal{D}}(1-\alpha) = \{d \in \mathcal{D}_n : G(\hat{m}(\mathbf{d})) \leq \hat{c}_{1-\alpha}/\sqrt{r_n}\}.$$

Bootstrapping Demand Bounds: Define

$$m(\mathbf{q}) = \mathbf{P}\mathbf{q} - a \in \mathbb{R}^T, \quad \hat{m}(y) = \mathbf{P}\mathbf{q} - \hat{a} \in \mathbb{R}^T,$$

where $\mathbf{P}$ is given in Assumption C.2, while a and $\hat{a}$ are given as

$$a(t) := \mathbf{p}(t)'\mathbf{d}(x(t),t), \quad \hat{a}(t) := \mathbf{p}(t)'\hat{\mathbf{d}}(\hat{x}(t),t), \quad t = 1,...,T.$$

Finally, note that $\mathcal{B}$ is the "parameter space." We now have that $\mathcal{S} = \Theta_I$ and $\hat{\mathcal{S}}(c) = \hat{\Theta}_I$, and we use the following bootstrap procedure to obtain a confidence region for the demand bounds:

1. For $b = 1, ..., B$: Draw a bootstrap sample with replacement from the data,

$$\left\{Z^*_{t,i}(b) : i = 1, ..., n,\ t = 1, ..., T\right\}.$$

   Compute the demand function estimator based on the bootstrap sample, $\hat{\mathbf{d}}^*_b(x(t), t)$, and the income path associated with $(\mathbf{p}_0, x_0)$, $\{\hat{x}^*_b(t)\}$,

$$\mathbf{p}_0'\hat{\mathbf{d}}^*_b(\hat{x}^*_b(t), t) = x_0, \quad t = 1, ..., T.$$

2. With $\hat{m}^*_{b,t}(\mathbf{q}) = \mathbf{p}(t)'\mathbf{q} - \mathbf{p}(t)'\hat{\mathbf{d}}^*_b(\hat{x}^*_b(t), t)$, compute

$$e^*_{b,t}(\mathbf{q}) := \sqrt{r_n}(\hat{m}^*_{b,t}(\mathbf{q}) - \hat{m}_t(\mathbf{q})) \times \mathbb{I}\{|\hat{m}_t(\mathbf{q})| \leq c_n/\sqrt{r_n}\}, \quad , t = 1, ..., T,$$

   and

$$\Gamma^*_b := \sup_{\mathbf{q}\in\hat{\mathcal{S}}(c_n)} G(e^*_b(\mathbf{q})).$$

We then compute the confidence region of the bounds by

$$\hat{\mathcal{S}}(1-\alpha) = \{\mathbf{q} \in \mathcal{B} : G(\hat{m}(\mathbf{q})) \leq \hat{c}_{1-\alpha}/\sqrt{r_n}\}.$$

Bootstrapping RP Test: To translate the "moment" equalities used to test for rationality into moment inequalities, we define $\hat{m}(\underline{\mathbf{q}}) = \left(\hat{m}_1(\underline{\mathbf{q}}), \hat{m}_2(\underline{\mathbf{q}}), \hat{m}_3(\underline{\mathbf{q}})\right) \in \mathbb{R}^{4T}$, where $\hat{m}_1(\underline{\mathbf{q}}) = \underline{\mathbf{q}} - \hat{\mathbf{d}}(\mathbf{p}_0, x_0) \in \mathbb{R}^{2T}$, $\hat{m}_2(\underline{\mathbf{q}}) = \hat{\mathbf{d}}(\mathbf{p}_0, x_0) - \underline{\mathbf{q}} \in \mathbb{R}^{2T}$, and $\hat{m}_3(\underline{\mathbf{q}}) = B(\underline{\mathbf{q}})$, where $B(\underline{\mathbf{q}})$ is defined in Section 7. Furthermore, we restrict the weighting matrix in the test statistic, W_n, to be diagonal, $W_n = diag\{w_t\}$. We then introduce the following set of RP-restricted demands:

$$\hat{\mathbb{S}}_{\text{RP}}(c_n) := \left\{\underline{\mathbf{q}} \in \mathbb{S} : G(\hat{m}(\underline{\mathbf{q}})) \leq c_n/\sqrt{r_n}\right\},$$

and note that $\hat{\mathbb{S}}_{\text{RP}}(0) = \{\underline{\hat{\mathbf{q}}}^*\}$ while $MD_n(\underline{\mathbf{q}}|\mathbf{p}_0, x_0) := \sup_{y\in\hat{\mathbb{S}}_{\text{RP}}(0)} G(\hat{m}(\underline{\mathbf{q}}))$. Our bootstrap procedure now proceeds as follows:

1. For $b = 1, ..., B$: Draw a bootstrap sample with replacement from the data,

$$\left\{Z^*_{t,i}(b) : i = 1, ..., n,\ t = 1, ..., T\right\}.$$

   Compute the demand function estimator based on the bootstrap sample,

$$\hat{\mathbf{d}}^*_b(\mathbf{p}_0, x_0) = (\hat{\mathbf{d}}^*_b(\hat{x}^*_b(1), 1), ..., \hat{\mathbf{d}}^*_b(\hat{x}^*_b(T), T)) \in \mathbb{R}^{2T}$$

   where $\hat{\mathbf{d}}^*_b$ is the unrestricted demand function estimator, and $\{\hat{x}^*_b(t)\}$ solves

$$\mathbf{p}_0'\hat{\mathbf{d}}^*_b(\hat{x}^*_b(t), t) = x_0, \quad t = 1, ..., T.$$

2. With $\hat{m}_b^*(\underline{\mathbf{q}}) = \left(\hat{m}_{b,1}^*(\underline{\mathbf{q}}), \hat{m}_{b,2}^*(\underline{\mathbf{q}})\right)$, where $\hat{m}_{b,1}^*(\underline{\mathbf{q}}) = \underline{\mathbf{q}} - \hat{\mathbf{d}}_b^*(\mathbf{p}_0, x_0)$, $\hat{m}_{b,2}^*(\underline{\mathbf{q}}) = \hat{\mathbf{d}}_b^*(\mathbf{p}_0, x_0) - \underline{\mathbf{q}}$ and $\hat{m}_{b,3}^*(\underline{\mathbf{q}}) = B(\underline{\mathbf{q}})$, compute

$$e_{b,t}^*(\theta) := \sqrt{r_n}(\hat{m}_{b,t}^*(\theta) - \hat{m}_t(\theta)) \times \mathbb{I}\{|\hat{m}_t(\theta)| \le c_n/\sqrt{r_n}\}, \quad , t = 1, ..., 2T,$$

and

$$\Gamma_b^* := \sup_{\underline{\mathbf{q}} \in \hat{\mathbb{S}}_{\mathrm{RP}}(c_n)} G(e_b^*(\underline{\mathbf{q}})).$$

The $(1-\alpha)$ critical value of $MD_n(\underline{\mathbf{q}}|\mathbf{p}_0, x_0)$ is now estimated by the $(1-\alpha)$ quantile of $\{\Gamma_b^* : b = 1, ..., B\}$, $\hat{c}_{1-\alpha}$.

We unfortunately have no theoretical justification for the above two bootstrap procedure when the demand function estimators are nonparametric. The theoretical results in Bugni (2010) are restricted to parametric estimation problems; in particular, the moments are based on i.i.d. averages that converge with $\sqrt{n}$-rate and are unbiased. In contrast, our "moments" are based on nonparametric estimators that converge with slower rates and are biased (in finite samples). As such, we can unfortunately not directly appeal to his theoretical results showing the validity of the above bootstrap procedures. It is outside the scope of this paper to demonstrate that the bootstrap procedures proposed here are in fact valid. It should be noted though that if we treat the sieve estimator as a parametric estimator (keeping $|\mathcal{K}_n|$ fixed), then Bugni's results do apply.

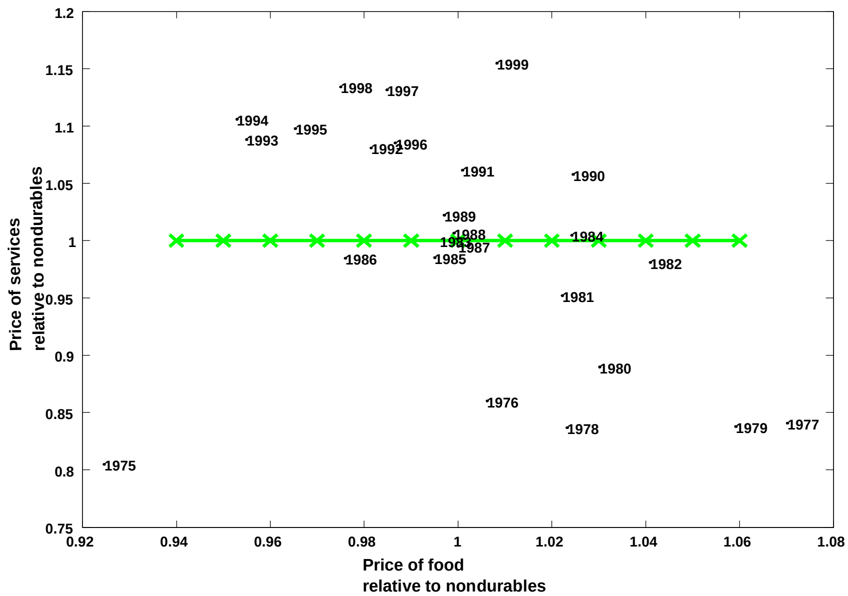
The finite-sample biases of nonparametric estimators can lead to inferior performance of bootstrap procedures if the same smoothing parameter is used for the estimator based on the actual sample, and those based on the bootstrap samples, see Hall (1992, Section 4.5). As in Blundell et al (2007), we handle this issue by computing the demand function estimators in Step 2 of the bootstrap procedures with $|\mathcal{K}_n|$ (and/or the penalization term) chosen slightly larger (smaller) than for the estimator based on the actual sample.

9 Endogenous Income

Suppose that income, $x(t)$, is endogenous such that the independence assumption made in (A.2) fails. The proposed sieve quantile estimator will in this case be inconsistent. One can instead then employ the IV quantile estimators developed in Chen and Pouzo (2009) and Chernozhukov, Imbens and Newey (2007). Alternatively, the control function approach taken in Imbens and Newey (2009) can be used.

With the assumptions and results of either of these three papers replacing our assumptions (A.1)-(A.3) and our Theorem 1, the remaining results of ours as stated in Theorems 2-5 remain valid since these follow from the properties of the unconstrained estimator. Thus, all the results stated in Theorems 2-5 go through except that the convergence rates and asymptotic distributions have to be modified to adjust for the use of another unrestricted estimator.

Figure 1: Relative prices in the FES: 1975 to 1999



10 Empirical Application

In our application we apply the methodology for constructing demand bounds under revealed preference restrictions to data from the British Family Expenditure Survey. The data set contains expenditure data and prices from British households. We use the same sample selection as in Blundell et al (2008) and we refer to that paper for a more detailed description. The distribution of relative prices over the central period of the data is give in Figure 1. It shows periods of quite dense relative prices, in the 1980s for example, and periods of sparse relative prices as in the 1970s.

We choose food as our primary good, and then group the other goods together in this application. The basic distribution of the Engel curve data are described in Figures 2 and 3.

10.1 The Quantile Sieve Estimates of Expansion Paths

In the estimation, we use polynomial splines

$$d_{1,n}(x,t,\tau) = \pi\left(t,\tau\right)' B_{K_n}(y)' = \sum_{j=0}^{q_n} \pi_j\left(t,\tau\right) x^j + \sum_{k=1}^{r_n} \pi_{q_n+k}\left(t,\tau\right)\left(x-\nu_k\left(t\right)\right)_+^{q_n}, \quad K_n = q_n + r_n + 1, \tag{23}$$

where $q_n \geq 1$ is the order of the polynomial and ν_k, $k = 1, ..., r_n$, are the knots. For a given choice of r_n, we place the knots according to the sample quantiles of $x_i\left(t\right)$, $i = 1, ..., n$, i.e., $\nu_k\left(t\right)$ was chosen as the estimated $k/\left(r_n+1\right)$-th quantile of $x\left(t\right)$.

In the implementation of the quantile sieve estimator, a small penalization term was added to the objective function to robustify the estimators (see Blundell, Chen and Kristensen, 2007 for a

Figure 2: The Engel Curve Distribution

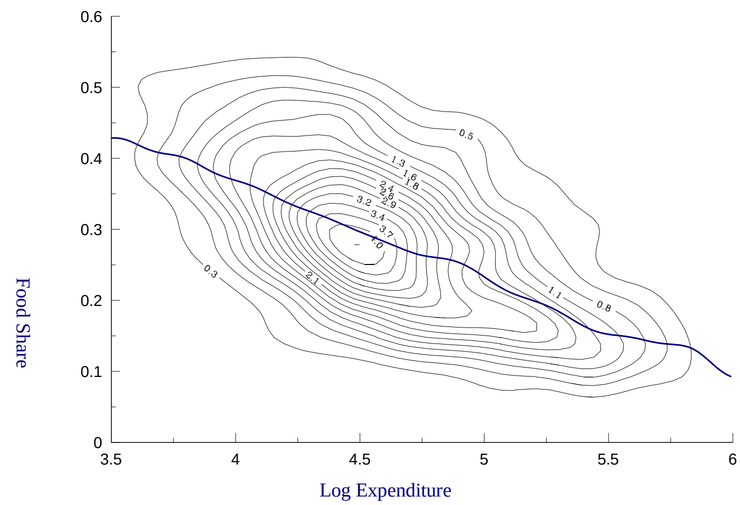
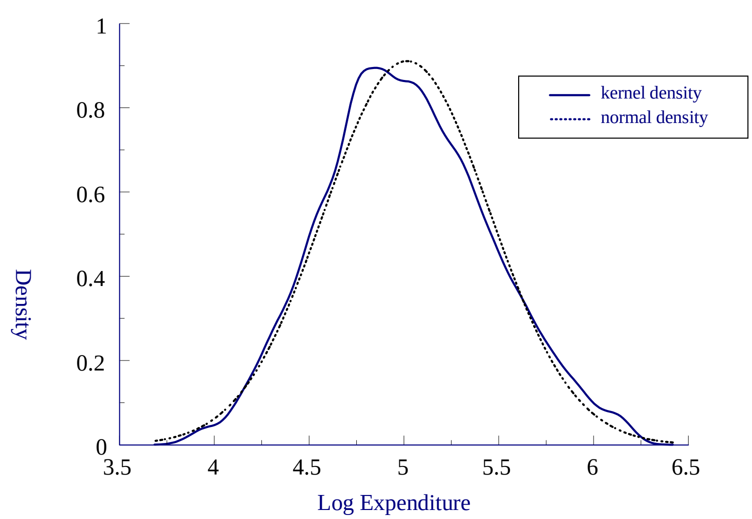


Figure 3: The Density of Log Consumption



similar approach). That is,

$$\hat{\pi}(t,\tau) = \arg \min_{\pi \in \mathbb{R}^{|\mathcal{K}_n|}} \frac{1}{n} \sum_{i=1}^{n} \rho_\tau \left(q_{1,i}(t) - \pi' W_i(t) \right) + \lambda Q(\pi), \quad \tau \in [0,1], \tag{24}$$

where

$$W_i(t) = \left(1, x_i(t), ..., x_i(t)^q, (x - \nu_1(t))_+^{q_n}, ..., (x - \nu_{r_n}(t))_+^{q_n}\right)',$$

and $\lambda Q(\pi)$ is an L_1-penalty term. Here, $Q(\pi)$ is the total variation of $\partial d_{n,1}(x) / (\partial x)$,

$$Q(\pi) = \int_a^b \left| \pi' \frac{\partial^2 B(x)}{\partial x^2} \right| dx \in \mathbb{R}_+,$$

while $\lambda > 0$ is the penalization weight that controls the smoothness of the resulting estimator.

By following the arguments of Koenker, Ng and Portnoy (1994), the above estimation problem can be formulated as a linear programming problem.[7] The restricted estimator is computed by solving the least-squares problem in Eq. (22). We opted for this estimator instead of the quantile estimator proposed in Eq. (21) since numerically we found it easier to solve the constrained least-squares problem.

In our application, we focus on FES data for the six year period 1983-1988. As in Blundell at al (2008) we use a group of demographically homogeneous households and estimate conditional quantile spline expansion paths using a 3rd order polynomial ($q_n = 3$) with $r_n = 5$ knots. The RP restrictions are imposed at 100 x-points over the empirical support x (log total expenditure on non-durables and services).

Each household is defined by a point in the distribution of log income and unobserved heterogeneity (x, ε). For households at the median of the income (total expenditure) distribution the unrestricted $\tau-$quantile expansion paths for food share are given in Figure 4.

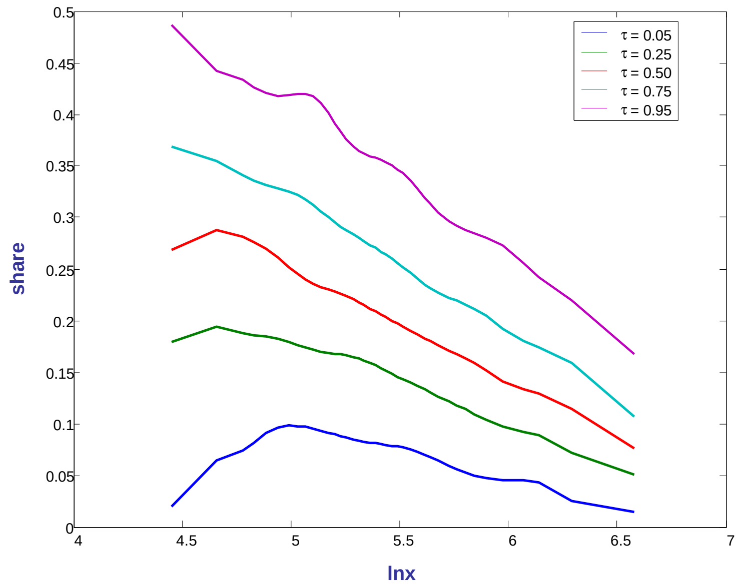
10.2 Estimated Demand Bounds

The key parameter of interest in this study is the consumer response at some new relative price $\mathbf{p}_0$ and income x or at some sequence of relative prices. The later defines the demand curve for $(x, \boldsymbol{\varepsilon})$. The estimated (e-)bounds (support sets) using the revealed preference inequalities and our FES data are given in Figures 5-8.

The Figures display the bounds on demand responses across the two dimensions of individual heterogeneity - income and unobserved heterogeneity. For a given income we can look at demand bounds for consumers with stronger or weaker preferences for food. Figure 5 shows the bounds on demands at the median income for the 50th percentile ($\tau = .5$) of the unobserved taste distribution. Notice that where the relative prices are quite dense the bounds are correspodngly narrow. Figure 6 contrasts this for a consumer at the 25% percentile of the heterogeneity distribution - a consumer with much weaker taste for food. At all points demands are much lower and the price response is somewhat less steep. Figure 7 considers a consumer with a strong taste for food - at the 75th

[7] The computation of the unrestricted estimator was done using Matlab code kindly provided by Roger Koenker.

Figure 4: $\tau-$Quantile Expansion Paths for Food Shares



percentile of the taste distribution. Demand shifts up at all points. The bounds remain quite narrow where the relative prices are dense. Finally, to illustrate the power of this approach, Figure 8 considers a higher income consumer but with median taste for food.

11 Summary and Conclusions

This paper has developed new techniques for the estimation and testing of stochastic consumer demand models. For general non-additive stochastic demand functions, we have demonstrated how revealed preference inequality restrictions from consumer optimisation conditions can be utilized to improve on the nonparametric estimation and testing of demand responses. We have shown how bounds on demand responses to price changes can be estimated nonparametrically, and derive their asymptotic properties utilizing recent results on estimation and testing of parameters characterized by moment inequalities. We also devise a test for rationality of the consumers.

An empirical application using data from the British Family Expenditure Survey has illustrated the usefulness of the methods. New insights have been provided about the price responsiveness of demand, especially across different income groups.

Figure 5: Quantile (RP-Rest) e-Bounds on Demand (Median InC) $\tau = .50$

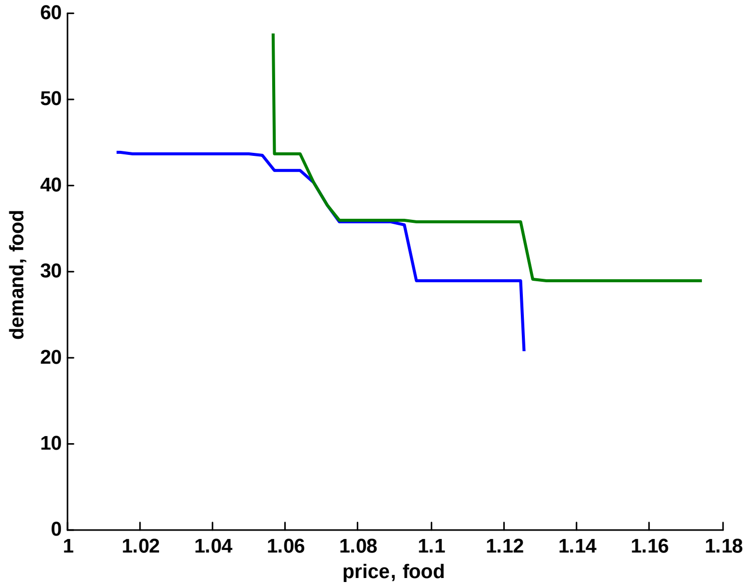


Figure 6: Quantile (RP-Rest) e-Bounds on Demand (Median Inc) $\tau = .25$

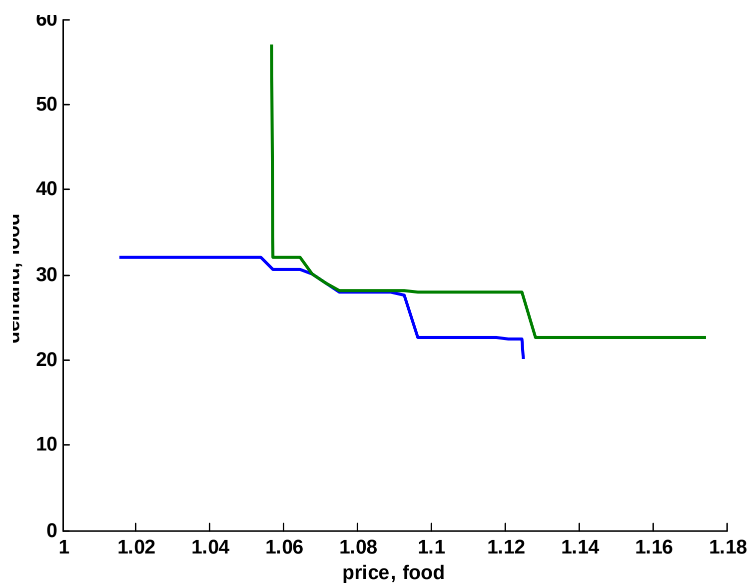


Figure 7: Quantile (RP-Rest) e-Bounds on Demand (Median Inc) $\tau = .75$

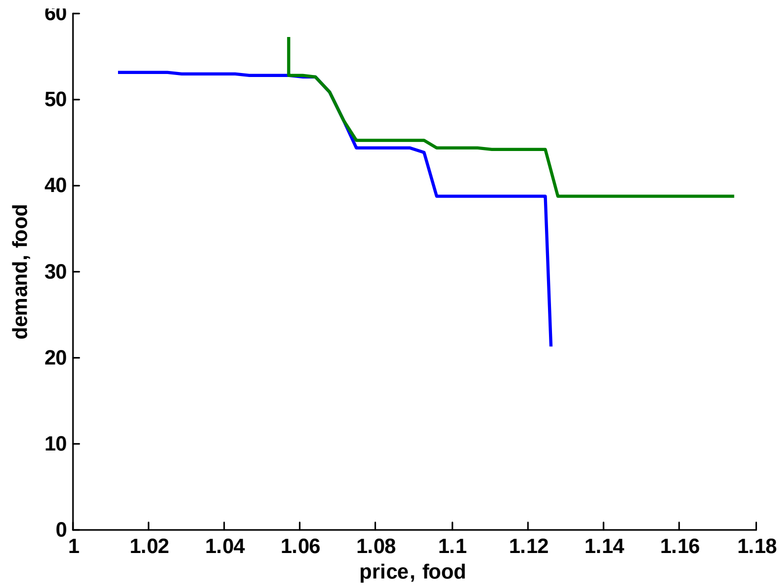
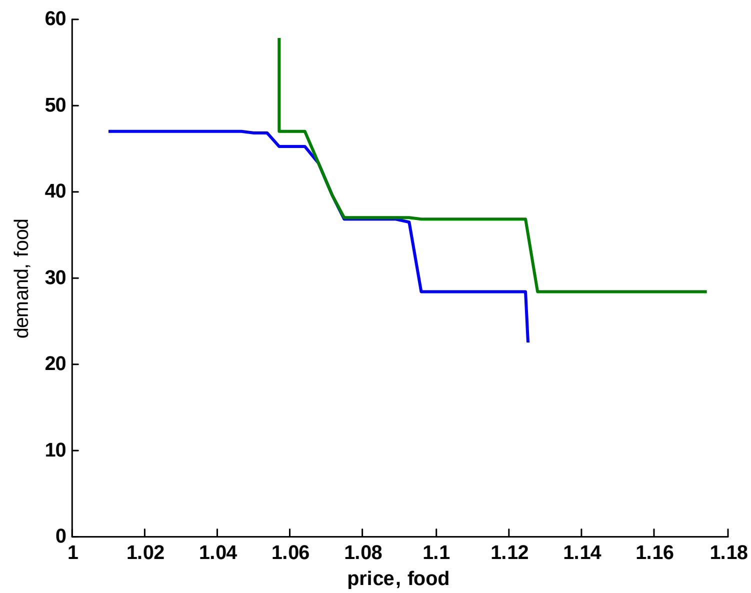


Figure 8: Quantile (RP-Rest) e-Bounds on Demand (75% Inc)



References

Afriat, S.N. (1973): "On a System of Inequalities in Demand Analysis: An Extension of the Classical Method," *International Economic Review* 14, 460-472.

Anevski, D. and O. Hössjer (2006): "A General Asymptotic Scheme for Inference Under Order Restrictions," *Annals of Statistics* 34, 1874-1930.

Andrews, D.W.K. (1997): "Estimation When a Parameter Is on a Boundary: Theory and Applications," Cowles Foundation Discussion Paper No. 1153.

Andrews, D.W.K. (1999): "Estimation When a Parameter Is on a Boundary," *Econometrica* 67, 1341-1383.

Andrews, D.W.K. (2001): "Testing When a Parameter Is on the Boundary of the Maintained Hypothesis," *Econometrica* 69, 683-734.

Andrews, D.W.K. and P. Guggenberger (2009): "Validity of Subsampling and 'Plug-in Asymptotic' Inference for Parameters Defined by Moment Inequalities," *Econometric Theory* 25, 669-709.

Andrews, D.W.K. and G. Soares (2010): "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica* 78, 119–157.

Beckert, W. (2007): "Specification and Identification of Stochastic Demand Models," *Econometric Reviews* 26, 669-683.

Beckert, W. and R. Blundell (2008): "Heterogeneity and the nonparametric analysis of consumer choice: conditions for invertibility," *Review of Economic Studies* 75, 1069-1080.

Belloni, A., X. Chen, V. Chernozhukov and Z. Liao (2010): "On Limiting Distributions of Possibly Unbounded Functionals of Linear Sieve M-Estimators," Manuscript, Dep. of Economics, Yale University.

Belloni, A. and V. Chernozhukov (2010): "Conditional Quantile Processes under Increasing Dimension," Manuscript, Dep. of Economics, MIT.

Beresteanu, A. (2006): "Nonparametric Estimation of Regression Functions under Restrictions on Partial Derivatives," working paper, Duke University.

Birke, M. and H. Dette (2007): "Estimating a Convex Function in Nonparametric Regression," *Scandinavian Journal of Statistics* 34, 384-404.

Blundell, R., M. Browning and I. Crawford (2003): "Nonparametric Engel Curves and Revealed Preference," *Econometrica* 71, 205-240.

Blundell, R., M. Browning and I. Crawford (2008): "Best Nonparametric Bounds on Demand Responses," *Econometrica* 76, 1227-1262.

Blundell, R., X. Chen and D. Kristensen (2007): "Semi-Nonparametric IV Estimation of Shape-Invariant Engel Curves," *Econometrica* 75, 1613-1669.

Brown, B.W. (1983): "The Identification Problem in Systems Nonlinear in the Variables," *Econometrica*, 51, 175-196.

Bugni, F. (2009): "Bootstrap Inference in Partially Identified Models Defined by Moment Inequalities: Coverage of the Points of the Identified Set," manuscript, Department of Economics, Duke University.

Bugni, F. (2010): "Bootstrap Inference in Partially Identified Models Defined by Moment Inequalities: Coverage of the Identified Set," *Econometrica* 78, 735–753.

Chen, X. (2007): "Large Sample Sieve Estimation of Semi-Nonparametric Models," in *Handbook of Econometrics*, Vol. 6B (eds. J. Heckman and E. Leamer), 5549-5632. Amsterdam: North-Holland.

Chen, X. and D. Pouzo (2007): "Estimation of Nonparametric Conditional Moment Models with Possibly Nonsmooth Moments," Cowles Foundation Discussion Papers, No. 1650.

Chen, X. and X. Shen (1998): "Sieve Extremum Estimates for Weakly Dependent Data," *Econometrica* 66, 289-314.

Chernozhukov, V., I. Fernandez-Val and A. Galichon (2006): "Improving Point and Interval Estimators of Monotone Functions by Rearrangement," *Biometrika* 96, 559–575.

Chernozhukov, V., G. Imbens and W.K. Newey (2007): "Instrumental Variable Estimation of Nonseparable Models," *Journal of Econometrics* 139, Pages 4-14

Chernozhukov, V., H. Hong and E. Tamer (2007): "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica* 75, 1243 - 1284.

Chesher, A. (2003): "Identification in Nonseparable Models," *Econometrica* 71, 1405-1441.

Gallant, A. R. & Golub, G. H. (1984): "Imposing Curvature Restrictions on Flexible Functional Forms," *Journal of Econometrics* 26, 295-321.

Haag, B.R., S. Hoderlein and K. Pendakur (2009): "Testing and Imposing Slutsky Symmetry in Nonparametric Demand Systems," *Journal of Econometrics* 153, 33-50.

Hall, P. (1992): *The Bootstrap and Edgeworth Expansion*. New York: Springer Verlag.

Hall, P. and J. Horowitz (2005): "Nonparametric Methods for Inference in the Presence of Instrumental Variables," *Annals of Statistics* 33, 2904-2929.

Hausman, J.A., W.K. Newey & J.L. Powell (1995): "Nonlinear Errors in Variables: Estimation of Some Engel Curves," *Journal of Econometrics* 65, 205-233.

Hoderlein, S. (2008): "How Many Consumers are Rational?" Manuscript, Dep. of Economics, Brown University.

Hoderlein, S. and J. Stoye (2009) "Revealed Preferences in a Heterogeneous Population," Manuscript, Dep. of Economics, Brown University.

Imbens, G. and W.K. Newey (2009): "Identification and Estimation of Triangular Simultaneous Equations Models without Additivity," *Econometrica* 77, 1481-1512.

Kiefer, J. (1982): "Optimum Rates for Non-parametric Density and Regression Estimates under Order Restrictions," in *Statistics and Probability: Essays in Honor of C. R. Rao* (eds. G. Kallianpur , P.R. Krishnaiah & J.K. Ghosh), 419-428. Amsterdam: North-Holland.

Klepper, S. and E.E. Learner (1984): "Consistent Sets of Estimates for Regressions with Errors in All Variables," *Econometrica* 52, 163-184.

Koenker, R. and G. Bassett, Jr. (1978): "Regression Quantiles," *Econometrica* 46, 33-50.

Koenker, R. and P. Ng (2005): "Inequality Constrained Quantile Regression," Sankhyā 67, 418-440.

Koenker, R., P. Ng and S. Portnoy (1994): "Quantile smoothing splines," *Biometrika* 81, 673-680.

Leser, C. E. V. (1963): "Forms of Engel Functions," *Econometrica*, 31, 694-703.

Lewbel, A. (1991): "The Rank of Demand Systems: Theory and Nonparametric Estimation," *Econometrica*, 59, 711-730.

Lewbel, A. (1995): "Consistent nonparametric hypothesis tests with an application to Slutsky symmetry," *Journal of Econometrics* 67, 379-401.

Lewbel, A. (2001): "Demand Systems With and Without Errors," *American Economic Review* 91, 611-618.

Lewbel, A. and K. Pendakur (2009): "Tricks With Hicks: The EASI Demand System," *American Economic Review* 99, 827-863.

Mammen, E. (1991): "Estimating a Smooth Monotone Regression Function," *Annals of Statistics* 19, 724-740.

Mammen, E., J.S. Marron, B.A. Turlach and M.P. Wand (2001): "A General Projection Framework for Constrained Smoothing," *Statistical Science* 16, 232–248.

Mammen, E. and C. Thomas-Agnan (1999): "Smoothing Splines and Shape Restrictions," *Scandinavian Journal of Statistics* 26, 239-252.

Matzkin, R.L. (2003): "Nonparametric Estimation of Nonadditive Random Functions," *Econometrica*, 71, 1339-1375.

Matzkin, R.L. (2005): "Identification in Nonparametric Simultaneous Equations," *Econometrica* 76, 945 - 978

Mukerjee, H. (1988): "Monotone Nonparametric Regression," *Annals of Statistic*s 16, 741-750.

Newey, W.K. (1997) Convergence Rates and Asymptotic Normality for Series Estimators," *Journal of Econometrics* 79, 147-168.

Newey, W.K. and D. McFadden (1994): "Large Sample Estimation and Hypothesis Testing," in *Handbook of Econometrics*, Vol. 4 (Eds. R.F. Engle and D.L. McFadden). Elsevier, Amsterdam.

Newey, W.K. & J. Powell (2003): "Instrumental Variables Estimation for Nonparametric Models," *Econometrica*, 71, 1565-1578.

Powell, J.L (1986): "Censored Regression Quantiles," *Journal of Econometrics* 32, 143-155.

Roehrig, C.S. (1988) "Conditions for Identification in Nonparametric and Parametric Models," *Econometrica*, 56, 433-447.

Van de Geer, S. (2000): *Empirical Processes in M-estimation.* Cambridge University Press.

Van der Vaart, A. and J. Wellner (1996): *Weak Convergence and Empirical Processes: With Applications to Statistics.* Springer-Verlag, New York.

White, H. (1994) *Estimation, Inference and Specification Analysis.* Econometric Society Monographs, Vol. 22. Cambridge University Press, Cambridge.

Working, H. (1943): "Statistical Laws of Family Expenditure," *Journal of the American Statistical Association*, 38, 43-56.

Wright, F.T. (1981): "The Asymptotic Behavior of Monotone Regression Estimates," *Annals of Statististics* 9, 443-448.

Wright, F.T. (1984): "The Asymptotic Behavior of Monotone Percentile Regression Estimates," *Canadian Journal of Statistics* 12, 229-236.

Yatchew, A. and Bos, L. (1997): "Nonparametric Regression and Testing in Economic Models," *Journal of Quantitative Economics* 13, 81{131.

A Proofs of Section 2 and 3

Proof of Theorem 1. We suppress the dependence on t since it is kept fixed in the following. Write the first demand equation as a quantile regression,

$$q_1 = d_1(x,\tau) + e(\tau), \tag{25}$$

where e is defined as the generalized residual, $e(\tau) := d_1(x,\varepsilon) - d_1(x,\tau)$. This formulation of the model for corresponds to the quantile regression considered in Chen (2007, Section 3.2.2). We then wish to verify the conditions stated there. First, we show that the distribution of $e(\tau)|x$ is described by a density $f_{t,\tau}(e|x)$ that satisfies

$$0 < \inf_{x\in\mathcal{X}} f_\tau(e|x) \le \sup_{x\in\mathcal{X}} f_\tau(e|x) < \infty, \tag{26}$$

$$\sup_{x\in\mathcal{X}} |f_\tau(e|x) - f_\tau(0|x)| \to 0, \quad |e| \to 0. \tag{27}$$

To see this, we first note that due to the independence between x and ε, the invertibility and differentiability of $\varepsilon \mapsto d_1(x,\varepsilon)$, and that $\varepsilon \sim U[0,1]$,

$$f_\tau(e|x) = \mathbb{I}\left\{0 \le d_1^{-1}(x, e - d(x,\tau)) \le 1\right\} \left|\frac{\partial d_1^{-1}(x, e - d_1(x,\tau))}{\partial e}\right|.$$

From this expression it is easily seen that Eq. (26) holds since $d_1(x,t,\varepsilon)$ and its derivative w.r.t. ε are continuous in x and $\mathcal{X}$ is compact. Eq. (27) clearly holds pointwise due to the continuity of $\varepsilon \mapsto d_1(x,\varepsilon)$. This can be extended to uniform convergence since $\sup_{x\in\mathcal{X},e\in[0,1]} f_\tau(e|x) < \infty$.

Combining the above results with the arguments given in the Proof of Chen (2007, Proposition 3.4), we now conclude that Chen (2007, Theorem 3.2) applies such that

$$||\hat{\mathbf{d}}(\cdot,t,\tau) - \mathbf{d}_0(\cdot,t,\tau)||_2 = O_P\left(\max\left\{\delta_n, \|\pi_n d_{1,0}(\cdot,t,\tau) - d_1(\cdot,t,\tau)\|_\infty\right\}\right)$$

where

$$\delta_n = \inf_{\delta\in(0,1)}\left\{\frac{1}{\sqrt{n}\delta^2}\int_{b\delta^2}^{\delta}\sqrt{H_{[]}(w,\mathcal{F}_n,\|\cdot\|)}dw \le \text{const.}\right\},$$

and $\pi_n d_{1,0}$ is an element in $\mathcal{D}_{1,n}$. Here, $H_{[]}(w,\mathcal{F}_n(\delta),\|\cdot\|_2) = \log\left(N_{[]}(w,\mathcal{F}_n(\delta),\|\cdot\|_2)\right)$ with $N_{[]}(w,\mathcal{F}_n(\delta),\|\cdot\|_2)$ denoting the so-called L_2-covering numbers with bracketing of the function class $\mathcal{F}_n(\delta)$, see Van der Vaart and Wellner (1996) and van de Geer (2000) for the precise definitions. To complete the proof, we appeal to Chen and Shen (1998, p. 311) to obtain that in the case of splines $\delta_n = O(\sqrt{k_n/n})$ and $\|\pi_n d_{1,0}(\cdot,t,\tau) - d_1(\cdot,t,\tau)\|_\infty = O(k_n^{-m})$.

The convergence rate result in the sup-norm is a direct consequence of Lemma 2.1 and Remark 2.1 in Belloni et al, 2010). ■

Proof of Theorem 2. First note that since data is independent over the time, it is sufficient to derive the marginal distributions of $\hat{d}_1(x(t),t,\tau)$, $t = 1,...,T$. This will follow from Belloni et al

(2010, Corollary 6.1) if their Conditions 6.1-6.2 hold under our assumptions. First, their Condition 6.1 is shown to hold in the Proof of Theorem 1, while their Condition 6.2 holds since

$$\left|f_\tau\left(e_1|x\right)-f_\tau\left(e_2|x\right)\right| \leq C\left|\frac{\partial d_1^{-1}\left(x, e_1-d_1(x, \tau)\right)}{\partial e}-\frac{\partial d_1^{-1}\left(x, e_2-d_1(x, \tau)\right)}{\partial e}\right| \leq C\left|e_1-e_2\right|,$$

where we have used that d_1 is continuously differentiable. ■

Proof of Theorem 3. Let $r_n=k_n / \sqrt{n}+k_n^{-m}$ denote the uniform rate of the unrestricted estimator, and let $x_0(t)$ be a given income expansion path generated from $\mathbf{d}$. We first note that the expansion path based on the unconstrained demand function satisfies

$$\hat{x}(T-1)-x(T-1)=\mathbf{p}(T-1)^{\prime}\left[\hat{\mathbf{d}}(x(T), t, \tau)-\mathbf{d}(x(T), t, \tau)\right]=O_P\left(r_n\right).$$

By recursion, we easily extend this to $\max_{t=1, \ldots, T}|\hat{x}(t)-x(t)|=O_P\left(r_n\right)$. It therefore follows that

$$\begin{aligned}
\hat{x}(t)-\mathbf{p}(t)^{\prime} \hat{\mathbf{d}}(\hat{x}(s), s, \tau) & =\{\hat{x}(t)-x(t)\}+\mathbf{p}(t)^{\prime}\left\{\mathbf{d}_0\left(x_0(s), s, \tau\right)-\hat{\mathbf{d}}(\hat{x}(s), s, \tau)\right\} \\
& \quad +x(t)-\mathbf{p}(t)^{\prime} \mathbf{d}_0\left(x_0(s), s, \tau\right) \\
& =O_P\left(r_n\right),
\end{aligned}$$

where we used that $x(t)=\mathbf{p}(t)^{\prime} \mathbf{d}_0(x(s), s, \tau)$. Thus, $\hat{x}(t)-\mathbf{p}(t)^{\prime} \hat{\mathbf{d}}(\hat{x}(s), s, \tau) \leq \epsilon$ with probability approaching one (w.p.a.1) as $r_n \rightarrow 0$. This proves that $\hat{\mathbf{d}} \in \mathcal{D}_{C, n}^T(\epsilon)$ w.p.a.1 such that $\hat{\mathbf{d}}_C^\epsilon=\hat{\mathbf{d}}$ w.p.a.1 as $r_n \rightarrow 0$. Since the restricted and unrestricted estimators are asymptotically equivalent, they must share convergence rates and asymptotic distributions. ■

B Proof of Theorem 4

We here prove a more general version of Theorem 4 since we believe this has independent interest. In particular, the general result takes as input any set of demand function estimators and derive the asymptotic properties of the corresponding bounds. The result is stated in such a fashion that it allows for both fully parametric, semi- and nonparametric first-step estimators and for any number of goods.

We consider a consumer with income x_0 who faces prices $\mathbf{p}_0=\left(p_{0,1}, \ldots, p_{0, L+1}\right)^{\prime}$ for the $L+1$ goods in the economy. The consumer's budget set is then given as:

$$\mathcal{B}_{\mathbf{p}_0, x_0}=\left\{\mathbf{q} \in \mathbb{R}_{+}^{L+1} | \mathbf{p}_0^{\prime} \mathbf{q}=x_0\right\},$$

which is compact and convex. Suppose that we have observed T prices, $\mathbf{p}(1), \ldots, \mathbf{p}(T)$, $\mathbf{p}(t)=\left(p_1(t), \ldots, p_{L+1}(t)\right)^{\prime}$, and let $\mathbf{d}(x(t), t)=\left(\mathbf{d}_1(x(t), t), \ldots, \mathbf{d}_{L+1}(x(t), t)\right)^{\prime}$, $t=1, \ldots, T$, denote the consumer's corresponding demand functions. Since the demand function has to satisfy $\mathbf{p}(t)^{\prime} \mathbf{d}(x, t)=x$, the demand for the $(L+1)$th good is simply given as

$$\mathbf{d}_{L+1}(x, t)=\frac{x-\mathbf{p}_{1: L}(t)^{\prime} \mathbf{d}_{1: L}(x, t)}{p_L(t)}. \tag{28}$$

The closure of the consumer's so-called demand support set can be represented as follows:

$$\mathcal{S}_{\mathbf{p}_0,x_0} = \left\{ \mathbf{q} \in \mathcal{B}_{\mathbf{p}_0,x_0} | \mathbf{p}(t)' \mathbf{q} \geq \bar{x}(t), \quad t = 1, ..., T \right\},$$

where $\{\bar{x}(t) : t = 1, ..., T\}$ is the intersection income path solving

$$\mathbf{p}_0' \mathbf{d}(\bar{x}(t), t, \tau) = x_0, \quad t = 1, ..., T.$$

For later use, note that, using the identity in eq. (28), the left hand side of the above equation can be rewritten as

$$\begin{aligned} \mathbf{p}_0' \mathbf{d}(x(t), t, \tau) &= \mathbf{p}_{0,1:L}' \mathbf{d}_{1:L}(x(t), t) + \mathbf{p}_{0,L+1} \frac{x(t) - \mathbf{p}_{1:L}(t)' \mathbf{d}_{1:L}(x(t), t)}{\mathbf{p}_L(t)} \\ &= \left[\mathbf{p}_{0,1:L} - \frac{p_{0,L+1}}{p_L(t)} \mathbf{p}_{1:L}(t) \right]' \mathbf{d}_{1:L}(x(t), t) + \frac{p_{0,L+1}}{p_L(t)} x(t) \end{aligned}$$

Also note that we can rewrite the support set as

$$\mathcal{S}_{\mathbf{p}_0,x_0} = \{ \mathbf{q} \in \mathcal{B}_{\mathbf{p}_0,x_0} | \bar{\mathbf{x}} - \mathbf{P}\mathbf{q} \leq \mathbf{0} \},$$

where $\mathbf{P}$ is the matrix containing the observed prices and $\bar{\mathbf{x}}$ is the vector of intersection income levels,

$$\mathbf{P} = [\mathbf{p}(1), \cdots, \mathbf{p}(T)]' \in \mathbb{R}_+^{T \times (L+1)}, \quad \bar{\mathbf{x}} = (\bar{x}(1), ..., \bar{x}(T))' \in \mathbb{R}_+^T.$$

Suppose that we have available estimators of these, $\hat{\mathbf{d}}(x(t), t)$. Again, the $(L+1)$th component of the estimator is restricted to satisfy eq. (28). We then in the following develop a support set estimator and analyze its theoretical properties. In order to provide a formal analysis, we impose the following regularity conditions:

C.1 $x(t) \mapsto \mathbf{d}(x(t), t)$ is montonically increasing and continuously differentiable.

C.2 The estimators $\hat{\mathbf{d}}_{1:L}(x, 1), ..., \hat{\mathbf{d}}_{1:L}(x, T)$ are mutually independent over time, and there exists sequences of nonsingular matrices $\Omega_n(x, t) \in \mathbb{R}^{L \times L}$ such that

$$\sup_{x \in \mathcal{X}} \left\| \Omega_n^{1/2}(x, t) (\hat{\mathbf{d}}_{1:L}(x, t) - \mathbf{d}_{1:L}(x, t)) \right\| = O_P(1/\sqrt{r_n})$$

for some sequence r_n.

C.3 At the intersection income levels,

$$\sqrt{r_n} \Omega_n^{1/2}(\bar{x}(t), t) (\hat{\mathbf{d}}_{1:L}(\bar{x}(t), t) - \mathbf{d}_{1:L}(\bar{x}(t), t)) \rightarrow^d N(0, V(\bar{x}(t), t)),$$

for some positive definite matrix $V(x(t), t) \in \mathbb{R}^{L \times L}$.

C.4 The estimator is differentiable and satisfies $\sup_{x \in \mathcal{X}} \left\| \partial \hat{\mathbf{d}}_{1:L}(x, t) / (\partial x) - \partial \mathbf{d}_{1:L}(x, t) / (\partial x) \right\| = o_P(1)$.

The monotonicity requirement in Condition (C.1) ensures that the intersection income path $\{\bar{x}(t)\}$ is uniquely defined and is a standard requirement in consumer demand theory. The differentiability condition in conjunction with (C.4) allow us to use standard delta method arguments to derive the asymptotic distribution of the intersection income levels.

Condition (C.2) introduces two sequences, a matrix $\Omega_n(x,t)$ and a scalar r_n. The condition states that once the demand estimator has been normalized by $\Omega_n^{1/2}(x,t)$ it converges with rate $\sqrt{r_n}$. (C.3) is a further strengthening and states that the estimator when normalized by $\sqrt{r_n}\Omega_n^{1/2}(x,t)$ converges towards a normal distribution. We have formulated (C.2)-(C.3) to cover as many potential estimators as possible. For parametric estimators, (C.2)-(C.3) will in general hold with $r_n = n$ and $\Omega_n(x,t) = I_L$. With nonparametric estimators, one may potentially choose $\Omega_n(x,t)$ and r_n in (C.2) and (C.3) differently: Most nonparametric estimators depend on a smoothing parameter (such as a bandwidth or number of basis functions) that can be chosen differently depending on whether a rate result is sought (as in (C.2)) or asymptotic distributional results (as in (C.3)). In particular, for the sieve quantile estimator, to obtain rate results we will choose $\Omega_n(x,t) = I$ and $r_n = O(k_n/\sqrt{n}) + O(k_n^{-m})$ with no restrictions on the sequence k_n; to obtain distributional results, we will choose $\Omega_n(x,t) = \Sigma_n^{-1}(x,t)$ as the inverse of the sequence of variance matrices given in Theorem 2 and $r_n = n$ in which case (C.3) holds under the restrictions on k_n imposed in Theorem 2

The following lemma states the properties of the estimated income paths under (C.1)-(C.4):

Lemma 6 *Assume that (C.1)-(C.2) hold. Then*

$$|\hat{x}(t) - \bar{x}(t)| = O_P\left(1/\sqrt{\|r_n\Omega_n(x(t),t)\|}\right).$$

If in addition (C.3)-(C.4) hold then,

$$\sqrt{r_n w_n(t)}\,(\hat{x}(t) - \bar{x}(t)) \to^d N(0,1),$$

where

$$w_n(t) := \left\| \left[\mathbf{p}_0' \frac{\partial \mathbf{d}(\bar{x}(t),t)}{\partial x}\right]^{-1} \left[\mathbf{p}_{0,1:L} - \frac{p_{0,L+1}}{p_L(t)}\mathbf{p}_{1:L}(t)\right]' \Omega_n^{-1/2}(x,t)\, V^{1/2} \right\|^{-2} > 0.$$

Proof. We treat the estimation of $\bar{x}(t)$ as a GMM estimation problem: Define

$$\hat{G}(x,t) = \mathbf{p}_0'\hat{\mathbf{d}}(x,t) - x_0 = \left[\mathbf{p}_{0,1:L} - \frac{p_{0,L+1}}{p_L(t)}\mathbf{p}_{1:L}(t)\right]' \hat{\mathbf{d}}_{1:L}(x,t) + \frac{p_{0,L+1}}{p_L(t)}x - x_0$$

and

$$G(x,t) = \mathbf{p}_0'\mathbf{d}(x,t) - x_0 = \left[\mathbf{p}_{0,1:L} - \frac{p_{0,L+1}}{p_L(t)}\mathbf{p}_{1:L}(t)\right]' \mathbf{d}_{1:L}(x,t) + \frac{p_{0,L+1}}{p_L(t)}x - x_0.$$

We then have that the estimated and true intersection incomes satisfy $\hat{x}(t) = \arg\min_{x\in\mathcal{X}} \hat{G}^2(x,t)$ and $\bar{x}(t) = \arg\min_{x\in\mathcal{X}} G^2(x,t)$ respectively. Given the requirement in (C.1) that the demand function is montonically increasing, $\bar{x}(t)$ is unique. Furthermore, since the demand function is continuous, so is $G(x,t)$. Finally, we note that

$$\begin{aligned}
& \sup_{x\in\mathcal{X}}\left|\hat{G}(x,t)-G(x,t)\right| \\
= & \sup_{x\in\mathcal{X}}\left|\left[\mathbf{p}_{0,1:L}-\frac{\mathbf{p}_{0,L+1}}{\mathbf{p}_L(t)}\mathbf{p}_{1:L}(t)\right]'\Sigma_n^{-1/2}(x,t)\,\Sigma_n^{1/2}(x,t)\left[\hat{\mathbf{d}}_{1:L}(x,t)-\mathbf{d}_{1:L}(x,t)\right]\right| \\
\leq & \left\|\mathbf{p}_{0,1:L}-\frac{\mathbf{p}_{0,L+1}}{\mathbf{p}_L(t)}\mathbf{p}_{1:L}(t)\right\|\left\|\Sigma_n^{-1/2}(x,t)\right\|\sup_{x\in\mathcal{X}}\left\|\Omega_n^{1/2}(x,t)\left(\hat{\mathbf{d}}_{1:L}(x,t)-\mathbf{d}_{1:L}(x,t)\right)\right\| \\
= & O_P\left(1/\sqrt{\|r_n\Omega_n(x,t)\|}\right),
\end{aligned}$$

where the last equality follows from (C.2). It now follows from standard consistency results for extremum estimators (see e.g. Newey and McFadden, 1994, Theorem 2.1) that $\hat{x}(t)\to^P\bar{x}(t)$. To obtain the rate result, we utilize that $\mathbf{d}_{1:L}(x,t)$ is continuously differentiable, c.f. (C.1), which implies that for any x in a small neighbourhood of $\bar{x}(t)$,

$$\begin{aligned}
G(x,t)-G(\bar{x}(t),t) &= \frac{\partial G(\tilde{x}(t),t)}{\partial x}[x-\bar{x}(t)] \\
&= \left\{\mathbf{p}_0'\frac{\partial\mathbf{d}(\tilde{x}(t),t)}{\partial x}+o_P(1)\right\}[x-\bar{x}(t)],
\end{aligned}$$

where $\tilde{x}(t)\in[x,\bar{x}(t)]$ and $\mathbf{p}_0'\partial\mathbf{d}(\tilde{x}(t),t)/(\partial x)\neq 0$. Thus, there exists $\kappa>0$ such that

$$|G(x,t)|=|G(x,t)-G(\bar{x}(t),t)|\geq\kappa|x-\bar{x}(t)|.$$

Given consistency, we therefore have

$$\begin{aligned}
|\hat{x}(t)-\bar{x}(t)| &\leq \kappa|G(\hat{x}(t),t)|\ \text{(wpa 1)} \\
&\leq \kappa\left(\left|G(\hat{x}(t),t)-\hat{G}(\hat{x}(t),t)\right|+\left|\hat{G}(\hat{x}(t),t)\right|\right) \\
&\leq \kappa\left(\left|G(\hat{x}(t),t)-\hat{G}(\hat{x}(t),t)\right|+\left|\hat{G}(x(t),t)\right|\right) \\
&= \kappa\left(\left|G(\hat{x}(t),t)-\hat{G}(\hat{x}(t),t)\right|+\left|\hat{G}(x(t),t)-G(x(t),t)\right|\right) \\
&= O_P\left(1/\sqrt{\|r_n\Omega_n(x(t),t)\|}\right).
\end{aligned}$$

Next, by a first-order Taylor expansion,

$$0=\hat{G}(\hat{x}(t),t)=\hat{G}(\bar{x}(t),t)+\frac{\partial\hat{G}(\tilde{x}(t),t)}{\partial x}(\hat{x}(t)-\bar{x}(t)),$$

where $\tilde{x}(t)\in[\hat{x}(t),\bar{x}(t)]$; in particular, $\tilde{x}(t)\to^P\bar{x}(t)$. This together with (C.4) implies

$$\frac{\partial\hat{G}(\tilde{x}(t),t)}{\partial x}\to^P\frac{\partial G(\bar{x}(t),t)}{\partial x}=\mathbf{p}_0'\frac{\partial\mathbf{d}(\bar{x}(t),t)}{\partial x}>0. \tag{29}$$

Moreover, with $\Delta_n(t) := \hat{\mathbf{d}}_{1:L}(\bar{x}(t), t) - \mathbf{d}_{1:L}(\bar{x}(t), t)$,

$$
\begin{aligned}
& \left[\mathbf{p}_0' \frac{\partial \mathbf{d}(\bar{x}(t), t)}{\partial x}\right]^{-1} \hat{G}(\bar{x}(t), t) \\
= & \left[\mathbf{p}_0' \frac{\partial \mathbf{d}(\bar{x}(t), t)}{\partial x}\right]^{-1} \left[\mathbf{p}_{0,1:L} - \frac{p_{0,L+1}}{p_L(t)} \mathbf{p}_{1:L}(t)\right]' \Delta_n(t) \\
= & \left[\mathbf{p}_0' \frac{\partial \mathbf{d}(\bar{x}(t), t)}{\partial x}\right]^{-1} \left[\mathbf{p}_{0,1:L} - \frac{p_{0,L+1}}{p_L(t)} \mathbf{p}_{1:L}(t)\right]' \Omega_n^{-1/2}(\bar{x}(t), t) V^{1/2} \left\{V^{-1/2} \Omega_n^{1/2}(\bar{x}(t), t) \Delta_n(t)\right\} \\
= & : a_n(t)' \left\{V^{-1/2} \Omega_n^{1/2}(\bar{x}(t), t) \Delta_n(t)\right\},
\end{aligned}
$$

where $V^{-1/2}\sqrt{r_n}\Omega_n^{1/2}(\bar{x}(t), t)\Delta_n(t) \to^d N(0, I_L)$ by (C.3). Next, observe that $w_n(t)$ defined in the lemma satisfies $w_n(t) = \|a_n(t)\|^{-2}$. Thus,

$$
\sqrt{r_n w_n(t)}(\hat{x}(t) - \bar{x}(t)) = \frac{a_n(t)'(1 + o_P(1))}{\|a_n(t)\|} \left\{V^{-1/2}\sqrt{r_n}\Omega_n^{1/2}(x(t), t)\Delta_n(t)\right\} \to^d N(0, 1).
$$

■

In the case where the demand function estimators have a common rate of convergence $\sqrt{r_n}$, $\Omega_n^{1/2}(x, t)$ can be chosen as the identity, and the asymptotic normality result simplifies to

$$
\sqrt{r_n}(\hat{x}(t) - \bar{x}(t)) \to^d N\left(0, w^2(t)\right),
$$

where

$$
w^2(t) := \left[\mathbf{p}_0' \frac{\partial \mathbf{d}(\bar{x}(t), t)}{\partial x}\right]^{-2} \left[\mathbf{p}_{0,1:L} - \frac{p_{0,L+1}}{p_L(t)} \mathbf{p}_{1:L}(t)\right]' V \left[\mathbf{p}_{0,1:L} - \frac{p_{0,L+1}}{p_L(t)} \mathbf{p}_{1:L}(t)\right].
$$

To define the support estimator, we introduce a diagonal weighting matrix given by

$$
W_n = \text{diag}\{w_n(1), ..., w_n(T)\},
$$

and assume that we have a consistent estimator of this,

$$
\hat{W}_n = \text{diag}\{\hat{w}_n(1), ..., \hat{w}_n(T)\}.
$$

Given $\hat{W}_n$, we introduce the following criterion function which is simply an $(L+1)$-dimensional generalization of the one introduced in the main text,

$$
Q_n(\mathbf{q}) = \left\|\hat{W}_n^{1/2}[\hat{\mathbf{x}} - \mathbf{P}\mathbf{q}]\right\|_+^2,
$$

with its limit given by

$$
\bar{Q}_n(\mathbf{q}) = \left\|W_n^{1/2}[\bar{\mathbf{x}} - \mathbf{P}\mathbf{q}]\right\|_+^2.
$$

Note that in the case where the intersection incomes converge with same rate (such that $\Omega_n(x, t)$ can be chosen as the identity matrix) the normalizations $\hat{W}_n^{1/2}$ and $W_n^{1/2}$ are not required. We note that the true support set can be expressed as

$$
\mathcal{S}_{\mathbf{p}_0, x_0} = \left\{\mathbf{q} \in \mathcal{B}_{\mathbf{p}_0, x_0} \left| \bar{Q}_n(\mathbf{q}) = 0\right.\right\}.
$$

This motivates us to define our support set estimator as

$$\hat{\mathcal{S}}_{\mathbf{p}_0,x_0}(c_n) = \{\mathbf{q} \in \mathcal{B}_{\mathbf{p}_0,x_0} \,|r_n Q_n(\mathbf{q}) \leq c_n\},$$

for some contour level c_n that we will choose in the following.

In order to analyze the set estimator we impose the following condition on the observed prices which is a multi-good version of (A.5) in the main text:

C.5 The matrix $\mathbf{P} = [\mathbf{p}(1), \cdots, \mathbf{p}(T)]' \in \mathbb{R}_+^{T\times(L+1)}$ has rank $L+1$.

The following theorem gives rate of convergence of the support set estimator and confidence sets for the unknown support set. Theorem 4 follows as a special case of this general result.

Theorem 7 *Assume that (C.1)-(C.2) and (C.5) hold, and that* $\hat{w}_n(t) = w_n(t) + o_P(1)$. *Then for any sequence* $c_n \propto \log(n)$,

$$d_{\mathrm{H}}(\hat{\mathcal{S}}_{\mathbf{p}_0,x_0}(c_n), \mathcal{S}_{\mathbf{p}_0,x_0}) = O_P(\sqrt{\log(n)/(r_n w_n^*)}),$$

where $w_n^* = \min_{t=1,\ldots,T} w_n(t)$.

If furthermore (C.3)-(C.4) hold, then

$$P(\mathcal{S}_{\mathbf{p}_0,x_0} \subseteq \hat{\mathcal{S}}_{\mathbf{p}_0,x_0}(\hat{c}_n)) \to 1-\alpha,$$

where $\hat{c}_n = \hat{q}_{1-\alpha} + O_P(\log(n))$ *with* $\hat{q}_{1-\alpha}$ *being an estimator of* $(1-\alpha)$*th quantile of* $\mathcal{C}_{\mathbf{p}_0,x_0}$ *given by*

$$\mathcal{C}_{\mathbf{p}_0,x_0} := \sup_{\mathbf{q}\in\mathcal{S}_{\mathbf{p}_0,x_0}} \|Z + \xi(\mathbf{q})\|_+^2.$$

Here, $Z \sim N(0, I_T)$ *while* $\xi(\mathbf{q}) = (\xi_1(\mathbf{q}), ..., \xi_T(\mathbf{q}))'$ *is given by*

$$\xi_t(\mathbf{q}) = \begin{cases} -\infty, & \mathbf{p}(t)'\mathbf{q} > \mathbf{p}(t)'\mathbf{d}(x(t), t) \\ 0, & \mathbf{p}(t)'\mathbf{q} = \mathbf{p}(t)'\mathbf{d}(x(t), t) \end{cases}, \quad t = 1, ..., T.$$

Proof. We follow the same proof strategy as in CHT and first verify that slightly modified versions of their Conditions C.1-C.2 are satisfied with our definitions of $\bar{Q}_n(\mathbf{q})$ and $Q_n(\mathbf{q})$. For convenience, define

$$m_n(\mathbf{q}) := \hat{\mathbf{x}} - \mathbf{P}\mathbf{q}, \quad \bar{m}(\mathbf{q}) := \bar{\mathbf{x}} - \mathbf{P}\mathbf{q}.$$

We then have uniformly in $\mathbf{q} \in \mathcal{B}_{\mathbf{p}_0,x_0}$,

$$\begin{aligned} Q_n(\mathbf{q}) &= \left\| \hat{W}_n^{1/2}\{m_n(\mathbf{q}) - \bar{m}_n(\mathbf{q})\} + \hat{W}_n^{1/2}\bar{m}(\mathbf{q}) \right\|_+^2 \\ &= \left\| \hat{W}_n^{1/2}\{\hat{\mathbf{x}} - \bar{\mathbf{x}}\} + \hat{W}_n^{1/2}\bar{m}(\mathbf{q}) \right\|_+^2, \\ &= \left\| \hat{W}_n^{1/2}\bar{m}(\mathbf{q}) \right\|_+^2 + O_P(1/\sqrt{r_n}) \\ &= \bar{Q}_n(\mathbf{q}) + O_P(1/\sqrt{r_n}), \end{aligned}$$

since $\hat{W}_n^{1/2}\{\hat{\mathbf{x}}-\bar{\mathbf{x}}\}=O_P\left(1/\sqrt{r_n}\right)$. Moreover,

$$\begin{aligned}
r_n Q_n(\mathbf{q}) &= \left\|\sqrt{r_n}\hat{W}_n^{1/2}\{m_n(\mathbf{q})-\bar{m}_n(\mathbf{q})\}+\sqrt{r_n}\hat{W}_n^{1/2}\bar{m}(\mathbf{q})\right\|_+^2 \\
&= \left\|\sqrt{r_n}\hat{W}_n^{1/2}\{\hat{\mathbf{x}}-\bar{\mathbf{x}}\}+\sqrt{r_n}\hat{W}_n^{1/2}\bar{m}(\mathbf{q})\right\|_+^2 \\
&= \frac{\left\|\sqrt{r_n}\hat{W}_n^{1/2}\{\hat{\mathbf{x}}-\bar{\mathbf{x}}\}+\sqrt{r_n}\hat{W}_n^{1/2}\bar{m}(\mathbf{q})\right\|_+^2}{\left\|\sqrt{r_n}\hat{W}_n^{1/2}\bar{m}(\mathbf{q})\right\|_+^2}\left\|\sqrt{r_n}\hat{W}_n^{1/2}\bar{m}(\mathbf{q})\right\|_+^2,
\end{aligned}$$

where $\sqrt{r_n}\hat{W}_n^{1/2}\{\hat{\mathbf{x}}-\bar{\mathbf{x}}\}=O_P(1)$ by Lemma 6, while $\left\|\sqrt{r_n}\hat{W}_n^{1/2}\bar{m}(\mathbf{q})\right\|_+^2\geq r_n w_n^*\lambda_{\min}^2\rho^2(\mathbf{q},\mathcal{S}_{\mathbf{p}_0,x_0})$ by Lemma 8 below. By the same arguments as in CHT, Proof of Theorem 4.2(Step 1), it now follows that $r_nQ_n(\mathbf{q})\geq r_nw_n^*\lambda_{\min}^2\rho^2(\mathbf{q},\mathcal{S}_{\mathbf{p}_0,x_0})/2$ wpa 1.

This shows that Condition C.1-C.2 of CHT hold in our case as well, except that the limiting objective function $\bar{Q}_n(\mathbf{q})$ and the constant $\kappa=\kappa_n=w_n^*\lambda_{\min}^2$ in their Condition C.2 both depend on n. We now proceed as in CHT, Proof of Theorem 3.1 to obtain the claimed rate result.

To show the validity of the proposed confidence set, we verify CHT's Condition C.4: We first note that for any given $\mathbf{q}$,

$$\hat{W}_n^{1/2}m_n(\mathbf{q})=\sqrt{r_n}\hat{W}_n^{1/2}\{\hat{\mathbf{x}}-\bar{\mathbf{x}}\}+\hat{W}_n^{1/2}\bar{m}(\mathbf{q})=_d Z+W_n^{1/2}\bar{m}(\mathbf{q})+o_P(1),$$

where Z is defined in the theorem. Next, for any $\mathbf{q}_1,\mathbf{q}_2$,

$$\left\|\hat{W}_n^{1/2}m_n(\mathbf{q}_1)-\hat{W}_n^{1/2}m(\mathbf{q}_2)\right\|=\left\|\hat{W}_n^{1/2}\mathbf{P}\{\mathbf{q}_1-\mathbf{q}_2\}\right\|\leq c\|\mathbf{q}_1-\mathbf{q}_2\|.$$

This proves that the stochastic process $\mathbf{q}\mapsto\left\{\hat{W}_n^{1/2}m_n(\mathbf{q})-W_n^{1/2}\bar{m}(\mathbf{q})\right\}$ weakly converges on the compact set $\mathcal{B}_{\mathbf{p}_0,x_0}$ towards Z, c.f. Van der Vaart and Wellner (2000, Example 1.5.10). In particular, $\hat{W}_n^{1/2}m_n(\mathbf{q})=_d Z+W_n^{1/2}\bar{m}(\mathbf{q})+o_P(1)$ uniformly in $\mathbf{q}$, which in turn implies that

$$r_nQ_n(\mathbf{q})=\left\|\sqrt{r_n}\hat{W}_n^{1/2}\{\hat{\mathbf{x}}-\bar{\mathbf{x}}\}+\sqrt{r_n}\hat{W}_n^{1/2}\bar{m}(\mathbf{q})\right\|_+^2=_d\left\|Z+\sqrt{r_n}W_n^{1/2}\bar{m}(\mathbf{q})\right\|_+^2+o_P(1),$$

uniformly in $\mathbf{q}$. The random variable $\mathcal{C}_n:=\sup_{\mathbf{q}\in\mathcal{S}_{\mathbf{p}_0,x_0}}r_nQ_n(\mathbf{q})$ therefore satisfies

$$\mathcal{C}_n=_d\sup_{\mathbf{q}\in\mathcal{S}_{\mathbf{p}_0,x_0}}\left\|Z+\sqrt{r_n}W_n^{1/2}\bar{m}(\mathbf{q})\right\|_+^2+o_P(1),$$

where $\sqrt{r_nw_n(t)}\bar{m}_t(\mathbf{q})=0$ for all n if $\bar{m}_t(\mathbf{q})=0$ and $\sqrt{r_nw_n(t)}\bar{m}_t(\mathbf{q})\to-\infty$ if $\bar{m}_t(\mathbf{q})<0$, $t=1,...,T$. Thus,

$$\lim_{n\to\infty}\sup_{\mathbf{q}\in\mathcal{S}_{\mathbf{p}_0,x_0}}\left\|Z+\sqrt{r_n}\hat{W}_n^{1/2}\bar{m}(\mathbf{q})\right\|_+^2=_d\sup_{\mathbf{q}\in\mathcal{S}_{\mathbf{p}_0,x_0}}\|Z+\xi(\mathbf{q})\|_+^2,$$

with $\xi(\mathbf{q})$ defined in the theorem. This proves the second claim. ■

Lemma 8 *Under (C.5), $\left\|W_n^{1/2}\bar{m}_n(\mathbf{q})\right\|_+^2\geq w_n^*\lambda_{\min}^2\rho(\mathbf{q},\mathcal{S}_{\mathbf{p}_0,x_0})$, where $\lambda_{\min}>0$ is the minimum eigenvalue of $\mathbf{PP}'$, and $w_n^*=\min_t w_n(t)$*

Proof. Consider any $\mathbf{q} \in \mathcal{B}_{\mathbf{p}_0,x_0} \backslash \mathcal{S}_{\mathbf{p}_0,x_0}$: Let $\mathbf{q}^* = \arg\min_{\mathbf{q}' \in \mathcal{S}_{\mathbf{p}_0,x_0}} ||\mathbf{q} - \mathbf{q}'||$ be the unique point in $\mathcal{S}_{\mathbf{p}_0,x_0}$ which has minimum distance to $\mathbf{q}$. Let $\delta^* = \mathbf{q} - \mathbf{q}^*$ be the difference such that $\|\delta^*\| = \rho(\mathbf{q}, \mathcal{S}_{\mathbf{p}_0,x_0})$. We can decompose the rows of $(\mathbf{P}, \bar{\mathbf{x}})$ into binding and non-binding constraints respectively of θ^*. Let $\left(\mathbf{P}^{(1)}, \bar{\mathbf{x}}^{(1)}\right)$ and $\left(\mathbf{P}^{(2)}, \bar{\mathbf{x}}^{(2)}\right)$, with $\mathbf{P}^{(1)} = \left[\mathbf{p}_1^{(1)\prime}, ..., \mathbf{p}_{T_1}^{(1)\prime}\right]'$ and $\bar{\mathbf{x}}^{(1)} = (x_1^{(1)}, ..., x_{T_1}^{(1)})'$, denote the set of rows which contain the binding and non-binding constraints respectively. That is, $\bar{m}_n^{(1)}(\mathbf{q}^*) := \mathbf{P}^{(1)}\mathbf{q}^* - \bar{\mathbf{x}}^{(1)} = 0$ while $\bar{m}_n^{(1)}(\mathbf{q}^*) := \mathbf{P}^{(2)}\mathbf{q}^* - \bar{\mathbf{x}}^{(2)} < 0$. Due to Assumption 2, $\lambda_{\min} > 0$ and we have

$$\lambda_{\min} \|\delta^*\| \leq \left\|\mathbf{P}^{(1)}\delta^*\right\| \leq \max_{t=1,...,T_1} |\mathbf{p}_t^{(1)\prime}\delta^*|.$$

Thus, there exists at least one $t_0 \in \{1, ..., T_1\}$ such that either $\mathbf{p}_{t_0}^{(1)\prime}\delta^* \leq -\lambda_{\min} \|\delta^*\|$ or $\lambda_{\min} \|\delta^*\| \leq \mathbf{p}_{t_0}^{(1)\prime}\delta^*$. We then obtain

$$\begin{aligned} \left\|W_n^{1/2}\bar{m}_n(\mathbf{q})\right\|_+^2 &= \sum_{t=1}^{T} w_n(t) \left|\mathbf{p}_t'\mathbf{q} - \bar{x}(t)\right|_+^2 \geq w_n(t_0) \left|\mathbf{p}_{t_0}^{(1)\prime}\mathbf{q} - \bar{x}_{t_0}^{(1)}(t)\right|_+^2 = w_n(t_0) \left|\mathbf{p}_{t_0}^{(1)\prime}\delta^*\right|_+^2 \\ &\geq w_n^*\lambda_{\min}^2 \|\delta^*\|^2 = w_n^*\lambda_{\min}^2\rho^2(\mathbf{q}, \mathcal{S}_{\mathbf{p}_0,x_0}). \end{aligned}$$

■

C Proof of Theorem 5

As with Theorem 4, we here prove a more general result that takes as input any estimate of the demand system. In order to develop the theory, it proves useful to reformulate $\mathbb{S}_{\mathbf{p}_0,x_0}$ in terms of inequality constraints involving differentiable functions. We claim that we can write $\mathbb{S}_{\mathbf{p}_0,x_0}$ as

$$\mathbb{S}_{\mathbf{p}_0,x_0} = \left\{\mathbf{q} \in \mathcal{B}_{\mathbf{p}_0,x_0}^T | V^*(\mathbf{q}, t) - V^*(\mathbf{q}, s) \geq \lambda^*(\mathbf{q}, t)\mathbf{p}(t)'(\mathbf{q}(s) - \mathbf{q}(t)), \quad s, t = 1, ..., T\right\},$$

where $V^* : \mathbb{R}_+^{(L+1)T} \mapsto \mathbb{R}_+^T$ and $\lambda^* : \mathbb{R}_+^{(L+1)T} \mapsto \mathbb{R}_+^T$ are differentiable functions. For $\mathbf{q} \in \mathbb{S}_{\mathbf{p}_0,x_0}$, the mappings can be constructed by, for example, the Varian-Afriat algorithm. More generally, we may define the two mappings as

$$\begin{aligned} (V^*(\mathbf{q}), \lambda^*(\mathbf{q})) \quad &: \quad = \arg\min_{V,\lambda \in \mathbb{R}^T} V'V + \lambda'\lambda \\ \text{s.t. } V(t) - V(s) \quad &\geq \quad \lambda(t)\,\mathbf{p}(t)'(\mathbf{q}(s) - \mathbf{q}(t)), \ \lambda(t) \geq 1, \ V(t) > 0, \quad t = 1, ..., T. \end{aligned}$$

This convex optimization problem has a unique solution for any given value of $\mathbf{q} \in \mathcal{B}_{\mathbf{p}_0,x_0}^T$, and clearly the mapping $\mathbf{q} \mapsto (V^*(\mathbf{q}), \lambda^*(\mathbf{q}))$ is continuously differentiable. We can then write the constrained set as $\mathbb{S}_{\mathbf{p}_0,x_0} = \{\mathbf{q} \in \mathbb{R}^{(L+1)T} | \mathbf{C}(\mathbf{q}) \leq 0\}$, where $\mathbf{C}(\mathbf{q})$ is the obvious multi-good extension of the function $\mathbf{C}(\mathbf{q})$ defined in eq. (16). We note that $\partial\mathbf{C}(\mathbf{q})/(\partial\mathbf{q}) = \{\partial C(\mathbf{q}, s, t)/(\partial\mathbf{q})\}_{1 \leq s,t \leq T}$ is given by

$$\frac{\partial C(\mathbf{q}, s, t)}{\partial \mathbf{q}(u)} := \frac{\partial V^*(\mathbf{q}, s)}{\partial \mathbf{q}(u)} - \frac{\partial V^*(\mathbf{q}, t)}{\partial \mathbf{q}(u)} + \frac{\partial \lambda^*(\mathbf{q}, t)}{\partial \mathbf{q}(u)}\mathbf{p}(t)'\Delta(s, t, u),$$

and $\Delta(s, t, u) = 1$ if $u = s$, $\Delta(s, t, u) = -1$ if $u = t$, and zero otherwise.

Now, we define an analyze a general estimator of $\mathbb{S}_{\mathbf{p}_0,x_0}$: Let $\hat{\mathbf{d}}(x,t) = (\hat{d}_1(x,t), ..., \hat{d}_{L+1}(x,t))'$ be the general demand function estimators considered in the previous section, and $\hat{\mathbf{q}} = (\hat{\mathbf{q}}(1), ..., \hat{\mathbf{q}}(T))$, $\hat{\mathbf{q}}(t) \in \mathbb{R}^{L+1}$ be the estimated intersection demands computed as in eqs. (14)-(15). The test statistic takes the same form, $\rho_n(\hat{\mathbf{q}}, \mathbb{S}_{\mathbf{p}_0,x_0}) := \inf_{\mathbf{q}\in\mathbb{S}_{\mathbf{p}_0,x_0}} \|\hat{\mathbf{q}} - \mathbf{q}\|^2_{\hat{W}_n^{\text{test}}}$ except that now the weights are $(L+1)\times(L+1)$ matrices. Since the estimator $\hat{\mathbf{q}}(t) = (\hat{q}_1(t), ..., \hat{q}_{L+1}(t))'$ has a degenerate distribution due to the budget constraint, $\hat{q}_{L+1}(t) = (x_0 - \mathbf{p}_{0,1:L}(t)'\hat{\mathbf{q}}_{1:L}(t))/p_{0,L}(t)$, we reparameterize the test statistic to avoid this degeneracy: For any vector $\mathbf{q} = (\mathbf{q}(1), ..., \mathbf{q}(T))$, with $\mathbf{q}(t) = (q_1(t), ..., q_{L+1}(t))$, define $\theta = (\theta(1), ..., \theta(T))$ as

$$\theta(t) := \mathbf{q}_{1:L}(t) \in \mathbb{R}^L.$$

In particular, we let $\hat{\theta}(t)$ and $\theta_0(t)$ denote the first L components of the estimated and true intersection demand, $\hat{\mathbf{q}}_{1:L}(t)$ and $\bar{\mathbf{q}}_{1:L}(t)$ respectively. We now rephrase the test statistic in this new parameterization. To define the GARP-restricted part of the parameter space, introduce the affine mapping $\mathbf{q}(\theta(t))$ defined as

$$\mathbf{q}(\theta(t)) := A\theta(t) + b,$$

where

$$A = \begin{bmatrix} I_L \\ \mathbf{p}'_{0,1:L}/p_{0,L+1} \end{bmatrix} \in \mathbb{R}^{(L+1)\times L}, \quad b = \begin{bmatrix} \mathbf{0}_L \\ x_0/p_{0,L+1} \end{bmatrix} \in \mathbb{R}^{L+1}.$$

We can then write the θ-parameterization of the set $\mathbb{S}_{\mathbf{p}_0,x_0}$ as

$$\Theta_0 := \left\{\theta \in \mathbb{R}^{LT} : \mathbf{C}(\mathbf{q}(\theta)) \leq 0 \right\}. \tag{30}$$

We introduce the unrestricted parameter space $\Theta \subset \mathbb{R}^{LT}$ as some compact set containing Θ_0 and with θ_0 being an interior point. We define the objective function $\ell_n(\theta)$ corresponding to the distance measure as

$$\ell_n(\theta) := -\sum_{t=1}^{T} (\hat{\mathbf{q}} - \mathbf{q}(\theta(t)))' \hat{W}_n^{\text{test}}(t) (\hat{\mathbf{q}} - \mathbf{q}(\theta(t))),$$

and let $\hat{\theta} = (\hat{\theta}(1), ..., \hat{\theta}(T))$ denote the unrestricted extremum estimator

$$\hat{\theta} = \arg\max_{\theta\in\Theta} \ell(\theta),$$

which obviously is given as $\hat{\theta}(t) = \hat{\mathbf{q}}_{1:L}(t)$ (for n large enough). Similarly, the restricted extremum estimator,

$$\hat{\theta}_0 = \arg\max_{\theta\in\Theta_0} \ell_n(\theta)$$

equals the first L components of the projection of the intersection demand onto the test statistic, $\hat{\theta}_0(t) = \hat{\mathbf{q}}^*_{1:L}(t)$. Since $\ell_n(\hat{\theta}) = 0$, our test statistic can be written as

$$\rho_n(\hat{\mathbf{q}}, \mathbb{S}_{\mathbf{p}_0,x_0}) = -\ell_n(\hat{\theta}_0) = \left\{\ell_n(\hat{\theta}) - \ell_n(\theta_0)\right\} - \left\{\ell_n(\hat{\theta}_0) - \ell_n(\theta_0)\right\}. \tag{31}$$

The parameter θ_0 lies in the interior of Θ by construction, so the first term, $\ell_n(\hat{\theta}) - \ell_n(\theta_0)$, can be analyzed using standard techniques. However, θ_0 may be on boundary of the restricted parameter space Θ_0, and so the second term will have a non-standard asymptotic distribution. In order to derive its limit, we employ the general results of Andrews (2001); henceforth A01. To this end, we verify his Assumptions 1-6 for the restricted estimator.

First, we analyze the objective function $\ell_n(\theta)$. Given that $\hat{\mathbf{q}}(t) = A\hat{\theta}(t) + b$, we can rewrite $\ell_n(\theta)$ as

$$\ell_n(\theta) = -\sum_{t=1}^{T} (\hat{\theta}(t) - \theta(t))' A' \hat{W}_n^{\text{test}}(t) A(\hat{\theta}(t) - \theta(t)),$$

such that the first and second order derivatives are given by

$$\frac{\partial \ell_n(\theta)}{\partial \theta(t)} = 2A'\hat{W}_n^{\text{test}}(t) A(\hat{\theta}(t)) - \theta(t)), \quad \frac{\partial^2 \ell_n(\theta)}{\partial \theta(s) \partial \theta(t)'} = \begin{cases} -2A'\hat{W}_n^{\text{test}}(t) A, & s = t \\ 0 & s \neq t \end{cases},$$

while the third derivatives are zero. Since the derivatives depend on estimated intersection demands, we take a closer look at those: First note that with $\Delta_n(t) := \hat{\mathbf{d}}_{1:L}(\bar{x}(t), t) - \mathbf{d}_{1:L}(\bar{x}(t), t)$,

$$\begin{aligned} \hat{\theta}(t) - \theta_0(t) &= \hat{\mathbf{d}}_{1:L}(\hat{x}(t), t) - \mathbf{d}_{1:L}(\bar{x}(t), t) \\ &= \left\{ \hat{\mathbf{d}}_{1:L}(\hat{x}(t), t) - \hat{\mathbf{d}}_{1:L}(\bar{x}(t), t) \right\} + \Delta_n(t) \\ &= \left[\frac{\mathbf{d}_{1:L}(\bar{x}(t), t)}{\partial x(t)} + o_P(1) \right] \{\hat{x}(t) - \bar{x}(t)\} + \Delta_n(t), \end{aligned}$$

where, from the proof of Lemma 6,

$$\hat{x}(t) - \bar{x}(t) \simeq \left[\mathbf{p}_0' \frac{\partial \mathbf{d}(\bar{x}(t), t)}{\partial x} \right]^{-1} \left[\mathbf{p}_{0,1:L} - \frac{p_{0,L+1}}{p_L(t)} \mathbf{p}_{1:L}(t) \right]' \Delta_n(t).$$

In total,

$$\hat{\theta}(t)) - \theta_0(t) \simeq \Psi(t) \Delta_n(t),$$

$$\Psi(t) := I_L + \frac{\mathbf{d}_{1:L}(\bar{x}(t), t)}{\partial x(t)} \left[\mathbf{p}_0' \frac{\partial \mathbf{d}(\bar{x}(t), t)}{\partial x} \right]^{-1} \left[\mathbf{p}_{0,1:L} - \frac{p_{0,L+1}}{p_L(t)} \mathbf{p}_{1:L}(t) \right]'.$$

Utilizing the asymptotic expansion of $\hat{\theta}(t) - \theta_0(t)$, we now see that with the weights chosen as consistent estimators of

$$W_n^{\text{test}}(t) = A(A'A)^{-1} B_n(t) (A'A)^{-1} A' \in \mathbb{R}^{(L+1)\times(L+1)},$$

where

$$B_n(t) := \Psi(t) \left(\Psi(t)' \Psi(t) \right)^{-1} r_n \Omega_n(\bar{x}(t), t) \left(\Psi(t)' \Psi(t) \right)^{-1} \Psi(t)' \in \mathbb{R}^{L\times L},$$

we obtain

$$\begin{aligned} \frac{1}{2} B_n^{-1/2}(t) \frac{\partial \ell_n(\theta_0)}{\partial \theta(t)} &\simeq B_n^{1/2}(t) \Psi(t) \Delta_n(t) = \Psi(t) \left(\Psi(t)' \Psi(t) \right)^{-1} \left\{ \sqrt{r_n} \Omega_n^{1/2}(\bar{x}(t), t) \Delta_n(t) \right\} \\ &\to {}^d Z(t) \overset{\text{def}}{\sim} N\left(0, \Psi(t) \left(\Psi(t)' \Psi(t) \right)^{-1} V \left(\Psi(t)' \Psi(t) \right)^{-1} \Psi(t)' \right). \end{aligned}$$

and

$$\frac{1}{2}B_n^{-1/2}(t)\frac{\partial^2 \ell_n(\theta_0)}{\partial\theta(t)\,\partial\theta(t)'}B_n^{-1/2}(t) = -B_n^{-1/2}(t)\,B_n(t)\,B_n^{-1/2}(t) = -I_L.$$

These results imply A01's Assumptions 1, 2* and 3 which in turn imply his Assumption 4.

Finally, we verify A01's assumption 5 by showing that cone Λ defined in eq. (17) locally approximates $B_n^{1/2}(\Theta_0 - \theta_0)$. In order to do so, we impose the following additional restriction on the sequence $\Omega_n(x,t)$:

C.6 The sequence $\Omega_n(x,t)$ is well-approximated by $\sigma_n^2 I_L$ for some scalar sequence $\sigma_n^2 \to \infty$ in the sense that $\underline{c}\sigma_n^2 I_L \leq \Omega_n(x,t) \leq \bar{c}\sigma_n^2 I_L$ for some constants $0 < \underline{c} \leq \bar{c} < \infty$.

Condition (C.6) is satisfied for most standard estimators, including the sieve quantile estimators proposed here. Under this condition,

$$\bar{b}r_n\sigma_n^2 I_L \leq B_n(t) \leq \bar{b}r_n\sigma_n^2 I_L \tag{32}$$

for constants $0 < \underline{b} \leq \bar{b} < \infty$. Using these bounds on $B_n(t)$ together with the fact that Θ_0 can be expressed through a set of smooth inequality constraints, c.f. eq. (30), we obtain by the same arguments as in Andrews (1997, Proof of Lemma 2) that $B_n^{1/2}(\Theta - \theta_0)$ is locally approximated by the cone Λ given in eq. (17).

It now follows by A01, Theorem 1, that $\ell_n(\hat{\theta}_0) - \ell_n(\theta_0) \to^d Z'Z - \inf_{\lambda\in\Lambda}\|\lambda - Z\|^2$, while by standard results $\ell_n(\hat{\theta}) - \ell_n(\theta_0) \to^d Z'Z$, where $Z = (Z(1),...,Z(T))$. Substituting those limits into eq. (31), the following result is obtained:

Theorem 9 *Assume that (C.1)-(C.6) hold, and*

$$\hat{W}_n^{\text{test}}(t) = A\left(A'A\right)^{-1}\hat{B}_n(t)\left(A'A\right)^{-1}A' \in \mathbb{R}^{(L+1)\times(L+1)},$$

where, with $\hat{\Psi} \to^P \Psi$ *and* $||\hat{\Omega}_n(\hat{x}(t),t) - \Omega(\bar{x}(t),t)|| \to^P 0$,

$$\hat{B}_n(t) := \hat{\Psi}(t)\left(\hat{\Psi}(t)'\hat{\Psi}(t)\right)^{-1} r_n\hat{\Omega}_n(\hat{x}(t),t)\left(\hat{\Psi}(t)'\hat{\Psi}(t)\right)^{-1}\hat{\Psi}(t)' \in \mathbb{R}^{L\times L},$$

Then, with $Z(t) \sim N\left(0, \Psi(t)\left(\Psi(t)'\Psi(t)\right)^{-1}V\left(\Psi(t)'\Psi(t)\right)^{-1}\Psi(t)'\right)$,

$$\rho_n\left(\hat{\mathbf{q}},\mathbb{S}_{\mathbf{p}_0,x_0}\right) \to^d \inf_{\lambda\in\Lambda}\|\lambda - Z\|^2.$$

Theorem 5 is a consequence of the above result: First, note that with the choice of $\hat{W}_n^{\text{test}}(t)$ given in eq. (18), the limit is $Z \sim N(0, I_T)$. Second, (C.6) is satisfied by the sieve estimator since $0 < \underline{f} \leq f_\tau(0|x) \leq \bar{f}$ which in turn implies that $W_n(\tau)$ defined in eq. (7) satisfies

$$\tau(1-\tau)\underline{f}^2 E\left[\underline{B}_{k_n}(x)\underline{B}_{k_n}(x)'\right]^{-1} \leq W_n(\tau) \leq \tau(1-\tau)\bar{f}^2 E\left[\underline{B}_{k_n}(x)\underline{B}_{k_n}(x)'\right]^{-1},$$

which in turn implies that, given the assumptions on the eigenvalues of $E\left[\underline{B}_{k_n}(x)\underline{B}_{k_n}(x)'\right]$, $\underline{c}k_n \leq \Sigma_n(x(t),t,\tau) \leq \bar{c}k_n$ for suitably constants $0 < \underline{c} \leq \bar{c} < \infty$.