Bounding Quantile Demand Functions using Revealed Preference Inequalities[*]

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

June 2009
This Version: February 2012

Abstract

This paper develops a new technique for the estimation of consumer demand models with unobserved heterogeneity subject to revealed preference inequality restrictions. Particular attention is given to nonseparable heterogeneity. The inequality restrictions are used to identify bounds on quantile demand functions. A nonparametric estimator for these bounds is developed and asymptotic properties are derived. An empirical application using data from the U.K. Family Expenditure Survey illustrates the usefulness of the methods by deriving bounds and confidence sets for estimated quantile demand functions.

JEL: C20, D12

Keywords: consumer behaviour, revealed preference, bounds, quantile regression.

1 Introduction

This paper develops a new nonparametric approach to the estimation and prediction of consumer demand responses for heterogeneous consumers. The objectives are two-fold: First, to utilize inequality restrictions arriving from revealed preference (RP) theory to improve demand estimation and prediction. Second, to relax restrictions on unobserved heterogeneity in consumer demand functions. We propose both unconstrained and RP constrained nonparametric estimators for demand functions with non-additive unobserved tastes, and derive their asymptotic properties.

Estimation of consumer demand models, and of the utility functions generating consumer demand, have attracted attention since a long time ago (see, for example, Deaton and Muelbauer (1980) and the references therein.) However, within these models, allowing for unobserved taste variation has succeeded only in very specific cases (e.g., McElroy (1987)). As Brown and Walker (1989) and Lewbel (2001) have shown, demand functions generated from random utility functions do not typically generate demand function where the unobserved tastes are additive. The identification and estimation of consumer demand models that are consistent with unobserved taste variation require analyzing demand models with nonadditive random terms.

An early treatment of identification of non-additive models is Brown (1983) whose results were extended to nonparametric models in Roehrig (1988). Building on their work, Matzkin (2003, 2008, 2010) derives general identification results for non-additive models. A number of other authors have addressed specific issues in identification and estimation: For example, Chesher (2003, 2007) considers quantile-driven identification with Ma and Koenker (2003) making use of his results to construct parametric estimators. Imbens and Newey (2009) and Chernozhukov, Imbens and Newey (2007) also develop quantile-based estimators which allow for endogeneity. Our approach draws on this literature. Our unrestricted estimator is similar to the one developed in Imbens and Newey (2009).

Our proposed procedure incorporates into nonadditive methods shape restrictions derived from economic theory. The shape restrictions allow us to deal with the common empirical situation where only a relatively small number of market prices are observed. If each consumer is choosing demand by maximizing his or her preferences, demand of such consumer will satisfy the well known axioms of RP of Samuelson (1948), Houthakker (1950), Richter (1966), Afriat (1967) and Varian (1982). Our analysis follows Varian (1982), where the inequalitites developed in Afriat (1973)

are used to characterize bounds on individual demand responses to new prices. As in Blundell, Browning and Crawford (2003, 2008), we extend the RP approach of Afriat and Varian to the case where demand observations are from cross sectional data. This requires additional restrictions to connect identical preferences across budgets. Blundell, Browning and Crawford (2003, 2008)'s method connected the *average* consumer across incomes and prices, and developed bounds on the demand of this consumer under new prices. In this paper, we connect consumers across budgets by mapping each of them into a quantile, and develop bounds on the demand of each quantile. When preferences of each consumer are separable into a monotone increasing function of some goods and an unobserved taste, each quantile corresponds to a unique value of the unobserved taste. Under such assumption, our method connects across budgets consumers with identical unobserved tastes. Other methods of connecting consumers with the same unobserved taste across budgets are, of course, possible.

A key ingredient in our analysis is the Engel curve for heterogeneous consumers. This describes, for each consumer, the *expansion path* for demand as total expenditure changes. The modelling and estimation of the Engel curve relationship has a long history. For example, Working (1943) and Leser (1963) suggested parametric regression models where budget shares are linear functions of log total budget; the so-called Piglog-specification. This simple specification has subsequently 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). A natural way to allow for more flexible specifications is through non- and semiparametric methods which have been widely used in the econometric analysis of Engel curves; see for example Blundell, Chen and Kristensen (2007) and Blundell, Duncan and Pendakur (1998).

These studies largely concern *average* demands, that is they effectively assume an additive error structure and consequently impose strong assumptions on the class of underlying utility functions, see e.g. Lewbel (2001). As mentioned above, we allow for non-additive heterogeneity[1]. Under a monotonicity (or invertibility) restriction we show that the expansion path for each consumer is identified by the conditional quantile. We first develop a nonparametric unconstrained conditional quantile estimator. We then use RP inequalities to derive sharp bounds on quantile demand functions. The results in this paper refer to the case of two goods. The extension to multiple goods

---

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

and multiple tastes is non-trivial and is left to future work.

RP inequality restrictions only allow us to establish bounds on quantile predicted demands. The estimation problem therefore 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 quantile demand bounds estimators.

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 under the RP inequality restrictions and show the practical usefulness of our approach.[2]

The remainder of the paper is organized as follows: In Section 2, we set up our framework for modelling heterogeneous consumer choice. In Section 3 we develop unrestricted sieve estimator for the quantile Engel curves. Section 4 extends this to the RP constrained case. The estimation of demand function bounds is then developed in Section 5. In Section 6 we discuss the implementation of the estimator and examine how to compute confidence sets. Section 7 contains a simulation study. In section 8 we apply our approach to household expenditure data and estimate bounds on the quantile functions of predicted demands for food for a sample of British households. Section 9 concludes and also points to some relevant extensions. In particular, we discuss how our estimator can be extended to handle endogeneity of explanatory variables by using the recent results on nonparametric estimation of quantile models under endogeneity. We also examine possible routes to testing for rationality. All proofs have been relegated to the Appendix.

2 Heterogeneous Consumers and Market Prices

2.1 Quantile Expansion Paths

Consumer demand depends on market prices, individual income and individual heterogeneity. Suppose we observe consumers in $T \geq 1$ separate markets, where T is finite. In what follows we will assume these refer to time periods 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

---

[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 and Stoye (2009), Lewbel (1995).

observe his or her demands and income level (and potentially some other individual characteristics such as age, education etc., which we suppress in this discussion).

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. Rather, it is a repeated cross-section where, for each new price, a new cross section of consumers is drawn from the population. 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 two good case such that $\mathbf{q}(t) = (q_1(t), q_2(t))' \in \mathbb{R}_+^2$ and $\mathbf{p}(t) = (p_1(t), p_2(t))' \in \mathbb{R}_+^2$. The demand for good 1 is given by:

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

where ε is a time-invariant individual specific heterogeneity term that reflects unobserved heterogeneity in preferences and characteristics.[3] To ensure that the budget constraint is met, the demand for good two must satisfy:

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

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 over the subset of goods 1 and 2.

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 $T \leq 8$). In this setting, we are not able to identify the mapping $\mathbf{p} \mapsto \mathbf{d}(x, \mathbf{p}, \varepsilon)$. We will show that it is possible to identify the function $(x, \varepsilon) \mapsto \mathbf{d}(x, \mathbf{p}(t), \varepsilon)$ at each of the observed prices. To emphasize this, we will in the following write

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

So we have a sequence of T demand functions, $\{\mathbf{d}(x, t, \varepsilon)\}_{t=1}^T$. One consequence of this partial identification is that we cannot point identify demand responses to a new price, say $\mathbf{p}_0 \neq \mathbf{p}(t)$, $t = 1, ..., T$. Instead we propose to use revealed preference (RP) constraints involving $\{\mathbf{d}(x, t, \varepsilon)\}_{t=1}^T$ to construct so-called e-bounds for such counterfactual demands.

---

[3] 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.

2.2   e-Bounds on Quantile Demands Functions

Consider a particular consumer characterized by some $\varepsilon \in [0,1]$ and income x, with associated sequence of demand functions $\mathbf{d}\,(x,t,\varepsilon)$, $t = 1,...,T$. 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

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

which 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 given consumer characterized by a particular value of ε when faced by the set of prices $\{\mathbf{p}_1, \mathbf{p}_2, ...\mathbf{p}_T\}$. All demands are generated by $\mathbf{d}\,(x,\mathbf{p},\tau)$ conditioned on the specified unobserved heterogeneity ε. How do we find the support set of the demand for this consumer when he faces a new price vector $\mathbf{p}_0$ with total outlay x_0? Varian (1983) established that, under the generalized axiom of RP, the demand $\mathbf{d}\,(x_0,\mathbf{p}_0,\varepsilon)$ will be situated in the *support set* $\mathcal{S}_{\mathbf{p}_0,x_0,\tau}^V$ given by:

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

see Varian (1983) for the precise definition of the RP restrictions. This set can be tightened by introducing so-called expansion paths $\{\mathbf{p(t)}, \mathbf{q}\,(x(t),\tau)\}_{t=1,..T}$ as done in Blundell, Browning and Crawford (2008) (BBC08 in the following): Define intersection demands $\mathbf{q}\,(\bar{x}_\tau(t),t,\tau) = \mathbf{d}\,(\bar{x}_\tau\,(t)\,,\mathbf{p}\,(t)\,,\tau)$ where $\{\bar{x}_\varepsilon\,(t) : t = 1,...,T\}$ is a sequence of intersection incomes defined as the solution to

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

The set of points that are consistent with observed expansion paths *and* utility maximization is given by the *support set*:

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

It is the identified *set* of demand responses for any prices $\mathbf{p}_0$, incomes x_0 and heterogeneity ε. In particular, the support set defines bounds on possible quantile demand responses. We follow BBC08 and refer to these as *e-bounds*.

Using BBC08, we can establish that if the set of T demands satisfy the RP inequalities then the support set is non-empty and convex. In the case of two goods, the support set can also be shown

to define bounds on demands $\mathbf{q}_0$ that are sharp given the observed data and the RP inequalities. These *e-bounds* in this case 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*. In particular, it will in general give tighter bounds compared to Varian's version. 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 ε. This allows for the **distribution of demand responses** to vary across the income distribution in a unrestricted way.

For convenience, utilizing that by definition $\mathbf{p}(t)' \mathbf{d}(\bar{x}_\varepsilon(t), t, \varepsilon) = \bar{x}_\varepsilon(t)$, we rewrite the support set $\mathcal{S}_{\mathbf{p}_0, x_0, \varepsilon}$ in terms of a set of linear "moment" inequalities:

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

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

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

Some comments regarding the underlying assumptions used to establish the above bounds are in order:

First, a key assumption for the above analysis to be valid for a given consumer is that his unobserved component, ε, is time-invariant. This allows us to use the repeated cross-sectional data to track this consumer across different price regimes. In particular, under regularity conditions stated below, we can identify $d(x, \mathbf{p}(t), \varepsilon)$, $t = 1, ..., T$, from data. If a given consumer's ε is not time-varying, this set of demand functions will provide a full characterisation of his behaviour across the T price regimes. This in turn allows us to construct bounds for counterfactual demands for the consumer. On the other hand, if a consumer's ε is time varying, say, $\varepsilon_1, ..., \varepsilon_T$, the knowledge about $d(x, \mathbf{p}(t), \varepsilon)$, $t = 1, ..., T$, does not provide information of this particular consumer's behaviour over time unless we are given information about the particular sequence of ε's. In particular, the above bounds are not valid for this consumer.

Second, the above bounds analysis for counterfactual demand is motivated by the empirically relevant situation where only little price variation is available (small T). A different approach to statistical inference about counterfactual demand in our setting would be to develop estimators that, as $n, T \to \infty$, allows identification of demand responses to prices as well, $(x, \mathbf{p}, \varepsilon) \mapsto \mathbf{d}(x, \mathbf{p}, \varepsilon)$.

This would allow one to compute point estimates of $\mathbf{d}(x, \mathbf{p}_0, \varepsilon)$ which would be consistent for any value of $\mathbf{p}_0$ as $n, T \to \infty$. Moreover, the asymptotic distribution of the estimator as $n, T \to \infty$ could be used to construct confidence bands for the counterfactual demand; in particular, these bands would take into account the finite-sample variation of $\mathbf{p}(t)$. The outlined approach is an alternative to ours where we only establish estimators of $(x, \varepsilon) \mapsto \mathbf{d}(x, t, \varepsilon)$, $t = 1, .., T$, and conduct statistical inference for fixed T and $n \to \infty$. However, for small T, the confidence bands obtained from the alternative approach will in general be quite imprecise - in particular in a nonparametric setting - since they rely on asymptotic approximations, and so we expect that our procedure provides a more robust set of confidence bands for counterfactual demands. Moreover, prices are well-known to exhibit strong time series dependence (see Lewbel and Ng, 2005) which will lead to lead to further deterioration of nonparametric estimators in finite samples.

Finally, we would like to point out that our analysis focuses on economic agents whose demand decisions - given ε - are fully described by their income and the prices they face. In case of households with cohabiting couples, this assumption may be violated. While it is outside the scope of this paper to provide an analysis of collective demand decisions, we conjecture that recent results on RP of collective consumption as in Cherchye, De Rock and Vermeulen (2011) could be combined with the methods developed here to construct bounds for this more general case.

3   An Unrestricted Sieve Estimator

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

In order for $\mathbf{d}(x, t, \varepsilon)$ to be nonparametrically identified, additional constraints have to be imposed on the function and the random variables (x, ε). First, the distribution of unobserved heterogeneity ε is in general not identified from data, and so will be to assumed (or normalized) to be univariate and to follow a uniform distribution, $\varepsilon \sim U[0, 1]$. We will furthermore assume ε to be independent of $x(t)$.[4]

Next, we assume that d_1 is invertible in ε. Sufficient conditions for this to hold in demand models can be found in Matzkin (2003) and Beckert and Blundell (2008). This combined with the above restrictions on ε implies that $d_1(x, t, \tau)$, $\tau \in [0, 1]$, is identified as the τth quantile of

---

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

$q_1(t) | x(t) = x$ (Matzkin, 2003; Imbens and Newey, 2009):

$$d_1(x, t, \tau) = F^{-1}_{q_1(t)|x(t)=x}(\tau), \quad \tau \in [0, 1].\tag{3}$$

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, that is, quantile representations of Engel curves. Based on the above characterization of d_1, we will in the following develop nonparametric quantile estimators of the function.

The assumptions of a univariate and uniformly distributed ε and invertibility of d_1 are restrictive, but it is not possible to weaken those in our general setting without loosing identification of d_1 and thereby consistency of our quantile demand function estimator. Consistent estimators of marginal effects and average derivatives of non-additive models that are robust to deviations from the above assumptions are provided in Hoderlein and Mammen (2007, 2008). However, this would not permit the application of the methods developed in this paper as demands relating to individual consumers are not directly identified.

Given the above identification result, we proceed to develop a sieve quantile estimator of d_1. 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 $\|\cdot\|$.[5] We specify the precise form of $\mathcal{D}_1$ and $\|\cdot\|$ below. Given the function space $\mathcal{D}_1$, we choose sieve spaces $\mathcal{D}_{n,1}$ 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 $\pi_n d_1 \in \mathcal{D}_{n,1}$ 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

$$\mathcal{D}_{n,1} = \left\{ d_{n,1} : d_{n,1}(x, t, \tau) = \sum_{k \in \mathcal{K}_n} \pi_k(t, \tau) B_k(x), \quad \pi(t, \tau) \in \mathbb{R}^{|\mathcal{K}_n|} \right\},\tag{4}$$

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

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

---

[5] The function space could without problems be allowed to change over time, $t = 1, ..., T$. For notational simplicity, we maintain that the function space is the same across time.

with associated sieve space $\mathcal{D}_n$ obtained by replacing $\mathcal{D}_1$ by $\mathcal{D}_{n,1}$ 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. (3), we may employ standard quantile regression techniques to obtain the estimator: Let

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

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

$$\hat{\mathbf{d}}(\cdot, t, \tau) = \arg \min_{\mathbf{d}_n \in \mathcal{D}_n} \frac{1}{n} \sum_{i=1}^n \rho_\tau(q_{1,i}(t) - d_{n,1}(x_i(t), t, \tau)), \tag{5}$$

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. (4): Define $\underline{B}_{k_n}(x) = \{B_k(x) : k \in \mathcal{K}_n\} \in \mathbb{R}^{k_n}$, where $k_n = |\mathcal{K}_n|$, as the collection of basis functions spanning the sieve $\mathcal{D}_{n,1}$. Then the sieve estimator is given by $\hat{d}_1(x, t, \tau) = \hat{\pi}(t, \tau)' \underline{B}_{k_n}(x) = \sum_{k \in \mathcal{K}_n} \hat{\pi}_k(t, \tau) B_k(x)$, where

$$\hat{\pi}(t, \tau) = \arg \min_{\pi \in \mathbb{R}^{|\mathcal{K}_n|}} \frac{1}{n} \sum_{i=1}^n \rho_\tau(q_{1,i}(t) - \pi' \underline{B}_{k_n}(x_i(t))), \quad \tau \in [0, 1]. \tag{6}$$

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

$$\hat{d}_2(x, t, \tau) = \frac{x - p_1(t) \hat{d}_1(x, t, \tau)}{p_2(t)}. \tag{7}$$

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 $\varepsilon \sim U[0, 1]$, $1 \le t \le T$.

A.2 The demand function $d_1(x, t, \varepsilon)$ is strictly increasing in ε, $1 \le t \le T$.

The assumption of bounded support is fairly standard in the literature on sieve estimation. It should be possible 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 Matzkin (2003), 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 B-splines are used to construct the sieve space $\mathcal{D}_{n,1}$. For an introduction to these, we refer to Chen (2007, Section 2.3). All of the following results goes through for 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:

$$||\mathbf{d}\left(\cdot,t,\tau\right)||_2 = \sqrt{E\left[||\mathbf{d}\left(x,t,\tau\right)||^2\right]}, \quad ||\mathbf{d}\left(\cdot,t,\tau\right)||_\infty = \sup_{x\in\mathcal{X}}||\mathbf{d}\left(x,t,\tau\right)||.$$

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

A.3 The function $d_1\left(\cdot,t,\tau\right)\in\mathcal{D}_1$, 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 $m\geq 0$, $1\leq t\leq T$.

We now have the following result:

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

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

while

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

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

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

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

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

We here state 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 Chen, Chernozhukov and Liao (2010, Lemma

10

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 Chen 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 sequence of covariance matrices

$$V_n(\tau) = \tau(1-\tau) H_n^{-1}(t,\tau) \Omega_n(t,\tau) H_n^{-1}(t,\tau), \tag{8}$$

with

$$\Omega_n(t,\tau) = E\left[\underline{B}_{k_n}(x(t)) \underline{B}_{k_n}(x(t))'\right], \quad H_n(t,\tau) = E\left[f(0|t,x(t),\tau) \underline{B}_{k_n}(x(t)) \underline{B}_{k_n}(x(t))'\right].$$

Here, $f(0|t,x,\tau)$ denotes the conditional distribution of $e(t,\tau) := q_1(t) - d_1(x(t),t,\tau)$ given $x(t) = x$; this is given by

$$f(e|t,x,\tau) = f_{q_1(t)|x(t)}(e + d_1(x,t,\tau)|x), \tag{9}$$

where $f_{q_1(t)|x(t)}(\cdot|x)$ is the conditional density of $q_1(t)$ given $x(t)$. Note that $V_n(t,\tau)$ takes the same form as the asymptotic variance of the estimated coefficients $\hat{\pi}_{k_n}(\tau)$ in the quantile regression model $q_{1,i}(t) = \pi_{k_n}(\tau)' \underline{B}_{k_n}(x(t)) + e(t,\tau)$ where we treat $\underline{B}_{k_n}(x(t))$ as a set of regressors of fixed dimension, c.f. Powell (1986). We are then able to state the following asymptotic normality result:

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

$$\sqrt{n}\Sigma_n^{-1/2}(x,\tau) \begin{pmatrix} \hat{d}_1(x(1),1,\tau) - d_1(x(1),1,\tau) \\ \vdots \\ \hat{d}_1(x(T),T,\tau) - d_1(x(T),T,\tau) \end{pmatrix} \to^d N(0,I_T),$$

where $I_T \in \mathbb{R}^{T \times T}$ denotes the identity matrix, and $\Sigma_n(x,\tau) = \mathrm{diag}\{\Sigma_n(x(1),1,\tau),....,\Sigma_n(x(T),T,\tau)\} \in \mathbb{R}^{T \times T}$ with

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

The independence of the estimators across time is due to the fact that a new sample of consumers are drawn at each time period.

An attractive feature of the above result is that for a given sample, we can simply treat the sieve estimator as a parametric estimator: As already noted, $V_n(t, \tau)$ in eq. (8) 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 $\Sigma_n(x, \tau)$ can be obtained by replacing $V_n(t, \tau)$ in the above expression by

$$\hat{V}_n(t, \tau) = \tau(1 - \tau)\hat{H}_n^{-1}(t, \tau)\hat{\Omega}_n(t, \tau)\hat{H}_n^{-1}(t, \tau), \tag{10}$$

where

$$\hat{\Omega}_n(t, \tau) = \frac{1}{n}\sum_{i=1}^{n}\underline{B}_{k_n}(x_i(t))\underline{B}_{k_n}(x_i(t))', \quad \hat{H}_n(t, \tau) = \frac{1}{n}\sum_{i=1}^{n}\hat{f}(0|t, x_i(t), \tau)\underline{B}_{k_n}(x_i(t))\underline{B}_{k_n}(x_i(t))'.$$

Here, $\hat{f}(0|t, x, \tau) = \hat{f}_{q_1(t)|x(t)}(\hat{d}_1(x, t, \tau)|x)$ with $\hat{f}_{q_1(t)|x(t)}(q|x)$ being, for example, a kernel estimator of the conditional density. This asymptotic variance 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 $[p_2(t)/p_1(t)]^2$; this follows by the delta method and eq. (7). 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 is only stated in a pointwise version. As discussed in the following sections, uniform weak convergence results would be useful if the goal is to analyze demand bounds across a continuum of consumers (that is, for τ in some interval of $[0, 1]$). These can be obtained from the general results in Belloni, Chernozhukov and Fernandez-Val (2011), and so could potentially be used to examine uniform convergence of the resulting bounds. For simplicity, we here focus on pointwise results.

4   A Revealed Preference (RP) Restricted Sieve Estimator

If the consumer is indeed rational, then the unconstrained estimator will asymptotically satisfy the RP restrictions. However, in finite samples, there is no reason why the estimator should satisfy these restrictions. This motivates us to directly impose RP restrictions in the estimation of demand functions.

Consider a given consumer characterised by $\tau \in [0, 1]$, and construct the following particular income expansion path $\{\tilde{x}_\tau(t)\}$ recursively by

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

where we initialize the sequence at a given "termination" income level $x_\tau(T) \in \mathbb{R}_+$. The weak axiom of RP imply the following set of inequality constraints:

$$\tilde{x}_\tau(t) \leq \mathbf{p}(t)' \mathbf{d}(\tilde{x}(s), s, \tau), \ s < t, \ t = 1, ..., T. \tag{11}$$

If the demand functions $\mathbf{d}(x, t, \tau)$, $t = 1, ..., T$, satisfy these inequalities for any given income level $x_\tau(T)$, we say that "$\mathbf{d}(\cdot, \cdot, \tau)$ *satisfies RP*". Note that these constraints are invariant to the particular ordering of prices; any arbitrary ordering of prices will impose the same constraints on the overall set of demand functions.

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

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

for $\tau \in [0, 1]$, 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 the T time periods, the above definition of $\{\hat{\mathbf{d}}(\cdot, t, \tau)\}_{t=1}^{T}$ is equivalent to the unrestricted estimators in eqs. (5) and (7).

In order to impose the RP restrictions, we define the constrained function set as

$$\mathcal{D}_C^{T} := \mathcal{D}^{T} \cap \{\mathbf{d}(\cdot, \cdot, \tau) \ \text{satisfies RP}\}, \tag{12}$$

and similarly the constrained sieve as

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

The constrained estimator is then obtained as:

$$\{\hat{\mathbf{d}}_C(\cdot, t, \tau)\}_{t=1}^{T} = \arg\min_{\{\mathbf{d}_n(\cdot, t, \tau)\}_{t=1}^{T} \in \mathcal{D}_{C,n}^{T}} \frac{1}{n} \sum_{t=1}^{T} \sum_{i=1}^{n} \rho_\tau(q_{1,i}(t) - d_{n,1}(t, x_i(t))). \tag{13}$$

Note that since the RP inequalities impose restrictions across 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, 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 quantile estimator, and they furthermore restrict themselves to linear constraints. There are some results for estimation of monotone quantiles and other linear constraints, see Chernozhukov, Fernandez-Val and Galichon (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 constant $0 \leq \epsilon \leq 1$ if for any income expansion path,

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

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 can interpret ϵ as Afriat's so-called "efficiency parameter": With $\epsilon = 1$, no waste is allowed for; as ϵ decreases, the more waste we allow for; with $\epsilon = 0$, any sequence of demand functions is rationalizable. With this generalized version of GARP, we then define the corresponding 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. (12) satisfies $\mathcal{D}_{C}^{T} = \mathcal{D}_{C}^{T}(1)$.

Moreover, it should be clear that $\mathcal{D}_C^T(\bar{\epsilon}) \subseteq \mathcal{D}_C^T(\epsilon)$ for $0 \le \epsilon \le \bar{\epsilon} \le 1$ since $\mathrm{RP}(\epsilon)$ imposes weaker restrictions on the demand functions compared to $\mathrm{RP}(\bar{\epsilon})$

We now re-define our RP constrained estimators to solve the same optimization problem as before, but now the optimization takes place over $\mathcal{D}_{C,n}(\epsilon)$ for some given choice of ϵ. We let $\hat{\mathbf{d}}_C^\epsilon$ denote this estimator, and note that $\hat{\mathbf{d}}_C^1 = \hat{\mathbf{d}}_C$, where $\hat{\mathbf{d}}_C$ is given in eq. (13). Suppose now that $\{\mathbf{d}(\cdot, t, \tau)\}_{t=1}^T \in \mathcal{D}_C^T(\bar{\epsilon})$ for some $\bar{\epsilon} > \epsilon$; this implies that the unconstrained estimator satisfies $\{\hat{\mathbf{d}}(\cdot, t, \tau)\}_{t=1}^T \in \mathcal{D}_{C,n}^T(\epsilon)$ 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 $\mathcal{D}_C^T(\bar{\epsilon}) \subseteq \mathcal{D}_C^T(\epsilon)$, our new estimator will in general be less precise than the one defined as the optimizer over $\mathcal{D}_C^T(\bar{\epsilon})$, but if the difference $\bar{\epsilon} - \epsilon > 0$ is not too big, the additional estimation error should be negligible.

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

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

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.

The convergence rate in the sup-norm is identical to the one for the unconstrained estimator and as such is not minimax optimal. On the other hand, the constrained estimator does exhibit optimal convergence rate in the L_2-norm; we have left this result out to save space. 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) establishes 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, there may be significant finite-sample gains.

We conjecture that the above distributional result will not in general hold for the estimator $\hat{\mathbf{d}}_C$ defined as the minimizer over $\mathcal{D}_{C,n}^T(1)$ (where no waste is allowed). 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 first-order equivalent and very different techniques have to be used to analyze the constrained estimator. In particular, the asymptotic distribution of the constrained estimator would most likely be non-standard. This is, for example, demonstrated in Andrews (1999), Anevski and Hössjer (2006) and Wright (1981) who give results for inequality-constrained parametric and nonparametric problems respectively.

Finally, we note that the proof technique used to obtain 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 adapted to show that for any unconstrained demand function estimator, the corresponding RP-constrained estimator will be asymptotically equivalent when allowing for waste.

5 Estimation of Bounds on Quantile Demands

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, 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 $\mathbf{p}_0$ is given by the compact and convex set $\mathcal{B}_{\mathbf{p}_0, x_0}$ as given in eq. (2). Recall from Section 2 that the demand support set can then be represented as $\mathcal{S}_{\mathbf{p}_0, x_0, \tau} = \{\mathbf{q} \in \mathcal{B}_{\mathbf{p}_0, x_0} | \bar{\mathbf{x}}_\tau - \mathbf{P}\mathbf{q} \leq 0\}$, where $\bar{\mathbf{x}}_\tau := \{\bar{x}_\tau(t) : t = 1, ..., T\}$ is a sequence of intersection incomes defined as the solution to $\mathbf{p}_0' \mathbf{d}(\bar{x}_\tau(t), t, \tau) = x_0$, $t = 1, ..., T$.

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 $\hat{\mathbf{x}}_\tau = (\hat{x}_\tau(1), ..., \hat{x}_\tau(T))$ as the solutions to

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

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

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

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

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

where $\mathbf{d}_x(x, t, \tau) = \partial \mathbf{d}(x, t, \tau) / (\partial x)$ and $\Sigma_n(x, t, \tau)$ is the variance of $\hat{d}_1(x, t, \tau)$ as given 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. This has to be taken into account in order to construct valid confidence sets. We therefore introduce a sample objective function $Q_n(\mathbf{q})$ that contain normalized versions of the estimated demand bounds:

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

where $\|\mathbf{x}\|_+ = \|\max\{\mathbf{x}, 0\}\|$ for any vector $\mathbf{x}$, and $\hat{W}_n = \text{diag}\{\hat{w}_n(1), ..., \hat{w}_n(T)\}$ is a consistent estimator of W_n. In comparison to the naive estimator suggested earlier, we now normalize $\hat{\mathbf{x}}_\tau - \mathbf{P}\mathbf{q}$ 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 $\hat{\mathbf{x}}_\tau$ in addition is a consistent estimator of $\bar{\mathbf{x}}$, it is straightforward to verify that $\sup_{\mathbf{q} \in \mathcal{B}_{\mathbf{p}_0, x_0}} |Q_n(\mathbf{q}|\tau) - \bar{Q}_n(\mathbf{q}|\tau)| \rightarrow^P 0$ (see the Appendix), where $\bar{Q}_n(\mathbf{q}|\tau)$ is the non-stochastic version of $Q_n(\mathbf{q}|\tau)$ given by

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

An important point here is that even though $\bar{Q}_n(\mathbf{q}|\tau)$ is a sequence of functions (due to the presence of W_n), it still gives a precise characterization of the support set $\mathcal{S}_{\mathbf{p}_0, x_0}$ for any given $n \geq 1$:

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

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

$$\hat{\mathcal{S}}_{\mathbf{p}_0,x_0,\tau}\left(c_n\right) = \left\{ \mathbf{q} \in \mathcal{B}_{\mathbf{p}_0,x_0} \left| nQ_n\left(\mathbf{q}|\tau\right) \leq c_n \right. \right\}. \tag{14}$$

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 $Q_n\left(\mathbf{q}|\tau\right)$.

It is worth noting that the above formulation of the support set and its estimator in terms of $\bar{Q}_n\left(\mathbf{q}|\tau\right)$ and $Q_n\left(\mathbf{q}|\tau\right)$ 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 $\bar{Q}_n\left(\mathbf{q}|\tau\right)$, 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 $\mathbf{d}\left(x,t,\tau\right)$ is strictly increasing in x, $t = 1,...,T$.

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

The monotonicity requirement in Condition (A.4) ensures that the intersection income path $\left\{\bar{x}\left(t\right)\right\}$ 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 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 Hausdorff norm which is given by:

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

for any two sets $\mathcal{A}_1, \mathcal{A}_2$. In Appendix B, we derive the asymptotic properties of the estimated support set for general demand function estimators allowing for multiple goods and other types of estimators than sieves. The following theorem follows as a straightforward implication of this more general result:

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

$$d_{\mathrm{H}}(\hat{\mathcal{S}}_{\mathbf{p_0},x_0,\tau}(c_n), \mathcal{S}_{\mathbf{p_0},x_0,\tau}) = O_P(k_n \sqrt{\log(n)/n}) + O_P\left(\log(n) k_n^{-m}\right).$$

If furthermore, the eigenvalues of $E\left[\underline{B}_{k_n}(x(t))\underline{B}_{k_n}(x(t))'\right]$, $t = 1,...T$, *are bounded and bounded away from zero; ;* $k_n^4/n = O(1)$, $nk_n^{-3m+1/2} = O(1)$ *and* $nk_n^{-2m-1} = o(1)$, *then:*

$$P(\mathcal{S}_{\mathbf{p_0},x_0,\tau} \subseteq \hat{\mathcal{S}}_{\mathbf{p_0},x_0,\tau}(\hat{c}_n)) \to 1 - \alpha,$$

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* $C_{\mathbf{p_0},x_0,\tau}$ *given by*

$$C_{\mathbf{p_0},x_0,\tau} := \sup_{\mathbf{q}\in\mathcal{S}_{\mathbf{p_0},x_0,\tau}} \|Z + \xi_\tau(\mathbf{q})\|_+^2.$$

Here, $Z \sim N(0, I_T)$ *while* $\xi_\tau(\mathbf{q}) = (\xi_\tau(1,\mathbf{q}),...,\xi_\tau(T,\mathbf{q}))'$ *is given by*

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

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. The critical values are based on quantiles of $C_{\mathbf{p_0},x_0,\tau}$ which is the limiting distribution of $\sup_{\mathbf{q}\in\mathcal{S}_{\mathbf{p_0},x_0}} n\left\{Q_n(\mathbf{q}|\tau) - \bar{Q}_n(\mathbf{q}|\tau)\right\}$. Thus, the confidence set is constructed by inversion of the statistic defining the set estimator. As can be seen from the theorem, the distribution of $C_{\mathbf{p_0},x_0,\tau}$ depends on T-dimensional vectors Z and $\xi(\mathbf{q})$. The former is simply the limiting joint distribution of the (appropriately normalized) estimates of the intersection incomes $\bar{x}(t)$, $t = 1,...,T$, while the latter keeps track of which of the constraints are binding (in the population) with only the binding ones influencing the distribution.

In order to employ the above result in practice, we need to be able to obtain estimators of the quantiles of the random variable $C_{\mathbf{p_0},x_0,\tau}$ defined in the theorem. The distribution of $C_{\mathbf{p_0},x_0,\tau}$ is non-standard and cannot be written on closed form, so evaluation of its quantiles has to be done either through simulations (CHT) or resampling methods such as modified bootstrap (Bugni, 2010; Andrews and Soares, 2010) or subsampling (CHT).

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. If the degeneracy property should hold, we can choose $c_n = 0$ in the first part, and $\hat{c}_n = \hat{q}_{1-\alpha}$ in the second part.

Finally, we note that we have here constructed confidence bounds for the identified support set. One may instead be interested in constructing confidence bounds for the unidentified demand point. This can be done by using the results in CHT, Section 5.

6 Practical Implementation

In this section, we discuss in further detail how the demand function estimators and support set estimators can be implemented.

6.1 Computation of Constrained Estimator

In the following, we suppress the dependence on τ for notational convenience since this is kept fixed throughout.

For numerical ease, we propose a slightly different implementation of the RP constrained estimator compared to the one analyzed in Section 4. The reason for this is that the original constrained estimator requires solving a quantile regression problem with nonlinear constraints which is not easily implemented in standard software packages (in particular, the objective function is non-differentiable which makes standard search algorithms unreliable). Instead, we reformulate the constrained estimator as the solution to a constrained least-squares problem that standard numerical algorithms can handle: Given the unconstrained estimator $\hat{d}_1(t, x)$, we propose to estimate the constrained version as

$$\{\hat{\mathbf{d}}_C(\cdot, t)\}_{t=1}^T = \arg\min_{\mathbf{d}_n(\cdot, \cdot) \in \mathcal{D}_{C,n}^T} \frac{1}{n} \sum_{t=1}^T \sum_{i=1}^n \left(\hat{d}_1(t, x_i(t)) - d_{n,1}(t, x_i(t)) \right)^2. \tag{15}$$

It is not numerically feasible to check that a given candidate estimator satisfies the RP constraints across all potential income expansion paths of which there exists a continuum. Instead, we only check the RP constraints on a discrete grid as follows: First, choose (a large number of) M income "termination" values, $\tilde{x}_m(T)$, $m = 1, ..., M$. The latter will be used to generated income paths. For sufficiently large M, we hope to cover most of the possible income paths. For a given member of the constrained sieve, say $\{\mathbf{d}_n(x, t)\}_{t=1}^T$, where $d_{n,1}(x, t) = \pi(t)' \underline{B}_{k_n}(x)$, we then check whether it satisfies RP across this grid: Compute M SMP paths $\{\tilde{x}_m(t)\}$, $m = 1, ..., M$:

$$\tilde{x}_m(t) = \mathbf{p}(t)' \mathbf{d}_n(\tilde{x}_m(t+1), t+1), \tag{16}$$

For any of these paths, say, $\{\tilde{x}_m(t)\}$, we check whether eq. (11) holds. By defining

$$
\begin{aligned}
a_m(s,t,\pi) &= \left\{ \frac{p_2(t)}{p_2(s)} p_1(s) - p_1(t) \right\} \underline{B}_{k_n}(\tilde{x}_m(s))' \in \mathbb{R}^{k_n}, \\
b_m(s,t,\pi) &= \frac{p_2(t)}{p_2(s)} \tilde{x}_m(s) - \tilde{x}_m(t) \in \mathbb{R},
\end{aligned}
\tag{17}
$$

for $s < t$, the RP constraints can be written more conveniently on matrix form as $A(\pi)\pi \leq b(\pi)$, where

$$
A(\pi) = \left[O_{1\times(s-1)k_n}, a_m(s,t,\pi), O_{1\times(T-s)k_n}, \right]_{m=1,\ldots,M,s<t}, \quad b = [b_m(s,t,\pi)]_{m=1,\ldots,M,s<t},
$$

and $O_{p\times q}$ denotes the $(p \times q)$-dimensional matrix of zeros. This highlights that the constraints are nonlinear in π; if the constraints instead were linear, the constrained estimator could simply be implemented as discussed in Koenker and Ng (2005). Our original least-squares problem should then be well-approximated by

$$
\hat{\pi}_C = \arg\min_{\pi} \frac{1}{n} \sum_{t=1}^{T} \sum_{i=1}^{n} \left(\hat{d}_1(t, x_i(t)) - \pi(t)' \underline{B}_{k_n}(x_i(t)) \right)^2 \quad \text{s.t. } A(\pi)\pi \leq b(\pi).
\tag{18}
$$

For moderate/large values of T, solving the above optimization problem is still quite a formidable task. For example, with a sieve of dimension $k_n = 8$ and $T = 8$ (as is the case in our empirical application), we have a total of 64 parameters to solve for. Fortunately, this numerical issue can to some extent be bypassed by running the following iterative procedure: To initialize the procedure, note that for $T = 1$ the constrained estimator is equal to the unconstrained one, since in this case no RP constraints exist. Now, given an estimator for T periods worth of constraints, we can solve the constrained estimator for $T + 1$ periods by starting the numerical algorithm at the estimates obtained for T periods together with the unconstrained estimator for period $t = T + 1$. In our experience, this procedure is quite robust and allows numerical solutions to the constrained estimation problem with relatively large number of sieve terms and time periods.

6.2 Demand Bounds

To compute the demand bounds and their confidence sets we proceed in two steps: First, approximate estimators are found as solution to a linear programming problems: Given some cut-off level $\hat{c}_n$, we define $\hat{\mathbf{A}} = \hat{W}_n \mathbf{P} \in \mathbb{R}^{T\times 2}$, $\hat{\mathbf{b}} = \hat{c}_n + \hat{W}_n \hat{\mathbf{x}} \in \mathbb{R}^T$, and then compute:

$$
\begin{aligned}
\tilde{q}_{\text{up},1} &= \arg_{\mathbf{q}\in\mathbb{R}^2} \max q_1 \quad \text{s.t.} \quad \hat{\mathbf{A}}\mathbf{q} \leq \hat{\mathbf{b}} \text{ and } p_0\mathbf{q} = x_0, \\
\tilde{q}_{\text{low},1} &= \arg_{\mathbf{q}\in\mathbb{R}^2} \min q_1 \quad \text{s.t.} \quad \hat{\mathbf{A}}\mathbf{q} \leq \hat{\mathbf{b}} \text{ and } p_0\mathbf{q} = x_0.
\end{aligned}
$$

21

This yields approximate estimates of the upper and lower bounds for demand for good 1. The final estimates are then obtained by solving the following two optimization problems numerically:

$$\tilde{q}_{\text{up},1} = \arg_{\mathbf{q} \in \mathcal{B}_{\mathbf{p}0,x_0}} \max q_1 \quad \text{s.t.} \quad nQ_n(\mathbf{q}) \leq \hat{c}_n,$$

$$\tilde{q}_{\text{low},1} = \arg_{\mathbf{q} \in \mathcal{B}_{\mathbf{p}0,x_0}} \min q_1 \quad \text{s.t.} \quad nQ_n(\mathbf{q}) \leq \hat{c}_n,$$

where the optimization algorithm is started at $\tilde{q}_{\text{up},1}$ and $\tilde{q}_{\text{low},1}$ respectively.

Confidence regions for these demand bounds can be obtained by choosing the cut-off level $\hat{c}_n$ as $\hat{c}_n = \hat{q}_{1-\alpha} + O_P(\log(n))$, where $\hat{q}_{1-\alpha}$ is an estimator of the $(1-\alpha)$th quantile of $C_{\mathbf{p}0,x_0,\tau}$ defined in Theorem 4. This can be computed by simulations. We first rewrite $C_{\mathbf{p}0,x_0,\tau}$: Letting $\bar{T}_b = \max_{\mathbf{q} \in \mathcal{S}_{\mathbf{p}0,x_0,\tau}} \sum_{t=1}^{T} \xi_\tau(t,\mathbf{q})$, $\xi_t(t,\mathbf{q}) := \mathbb{I}\{\bar{x}_\tau(t) = \mathbf{p}(t)'\mathbf{q}\}$, denote the maximum number of binding constraints across all points in $\mathcal{S}_{\mathbf{p}0,x_0,\tau}$, we can write $C_{\mathbf{p}0,x_0,\tau} = \sum_{t=1}^{\bar{T}_b} \max\{Z(t),0\}^2$, where $\{Z(t)\}_{t=1}^{T} \sim N(0,I_T)$. Given a consistent estimator $\hat{T}_b = \max_{\mathbf{q} \in \hat{\mathcal{S}}_{\mathbf{p}0,x_0,\tau}} \sum_{t=1}^{T} \hat{\xi}_\tau(t,\mathbf{q})$, $\hat{\xi}_\tau(t,\mathbf{q}) = \mathbb{I}\{\hat{x}_\tau(t) \geq \mathbf{p}(t)'\mathbf{q} - a_n\}$ with $a_n \propto \sqrt{\log(n)/n}$, we propose to compute approximate quantiles by simulating from $\hat{C}_{\mathbf{p}0,x_0,\tau} = \sum_{t=1}^{\hat{T}_b} \max\{Z(t),0\}^2$.

7 Simulation Study

To investigate the finite-sample performance of our estimators, we conduct a small simulation study where we take as data generating process a simple random coefficient Cobb-Douglas model:

$$q_{1i}(t) = \alpha_i \frac{x_i(t)}{p_1(t)}, \quad q_{2i}(t) = (1-\alpha_i) \frac{x_i(t)}{p_2(t)},$$

where $\alpha_i \sim N(\bar{\alpha}, 0.005^2)$ is the random coefficient. This is a quite simple specification where the generalized axiom of RP (GARP) is satisfied when $\alpha_i \in [0,1]$. However, we will not utilize any of the structure in the Cobb-Douglas model in the implementation of our nonparametric estimators of demand functions and corresponding bounds. As such, we expect that the only change in our simulation results when moving to more complicated, non-linear DGP's of demand would be that additional biases due to these nonlinearities would appear in the estimates.

In line with our empirical application in the next section, we refer to good 1 as 'food'. The prices are chosen as those in the data set used in our empirical application; similarly, $x_i(t)$ is drawn from the empirical distribution of total expenditures of this data set. The data-generating value of $\bar{\alpha}$ is chosen as the mean share of food in the sample. We included $T = 6$ different prices in

the simulation study and used the same number of observations as in the empirical application, $n = 1448$.

In the estimation of $d_1(x, t, \tau) = \Phi^{-1}_{\bar{\alpha}, 0.005}(\tau) \, x / p_1(t)$, where $\Phi^{-1}_{\bar{\alpha}, 0.005}(\tau)$ is the quantile function of α_i, we employ (log-transformed) polynomial splines,

$$d_{n,1}(x, t, \tau) = \pi(t, \tau)' B_{K_n}(\log x)' = \sum_{j=0}^{q_n} \pi_j(t, \tau) \log(x)^j + \sum_{k=1}^{r_n} \pi_{q_n+k}(t, \tau)(\log x - \nu_k(t))_+^{q_n}, \quad (19)$$

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

$$\hat{\pi}(t, \tau) = \arg \min_{\pi \in \mathbb{R}^{|\mathcal{K}_n|}} \frac{1}{n} \sum_{i=1}^{n} \rho_\tau \left(q_{1,i}(t) - \pi' \underline{B}_{k_n}(\log x_i(t)) \right) + \lambda Q(\pi), \quad \tau \in [0, 1], \quad (20)$$

where

$$\underline{B}_{k_n}(x) = \left(1, x, ..., x^{q_n}, (x - \nu_1(t))_+^{q_n}, ..., (x - \nu_{r_n}(t))_+^{q_n} \right)',$$

and $\lambda Q(\pi)$ is an L_1-penalty term. Here, $Q(\pi)$ is the total variation of $\partial^2 d_{n,1}(x) / (\partial x^2)$,

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

while $\lambda > 0$ is the penalization weight that controls the smoothness of the resulting estimator. With $\lambda = \lambda_n \to 0$ sufficiently fast as $n \to \infty$, this will not interfere with the asymptotic properties derived in the previous sections. Throughout, we use a third order $(q = 3)$ spline with 4 knots $(r_n = 4)$ so the total number of basis functions is $k_n = 8$.

While the unconstrained demand estimator can be implemented straightforwardly in standard quantile regression software,[6] the constrained optimization problem was implemented using the recursive algorithm outlined in Section 6. Our results are based on 1000 simulated data sets. We only report results for $t = 4$ and $\tau = 0.5$. The results for other time periods were very similar, while the performance of the constrained estimator relative to the unconstrained one improved further as we moved away from the median.

---

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

In Figure 1 and 2, we report the mean and 95% confidence intervals for the unconstrained and constrained estimators of the expenditure share, $d_1(x, t, \tau) p_1(t) / x(t) = \alpha(\tau)$ where $\alpha(\tau)$ is the τth quantile of $\alpha_i \sim N(\bar{\alpha}, 0.005^2)$. As can be seen from the figures, at the median, both estimators are very precise with hardly any biases and small variances. There are only minor differences between the two estimators; this is in contrast to the empirical application where the differences are more pronounced. The strong performance of and the small differences between the two estimators are probably due to the very simple model; in particular, the true demand model is contained in the sieve that we use.
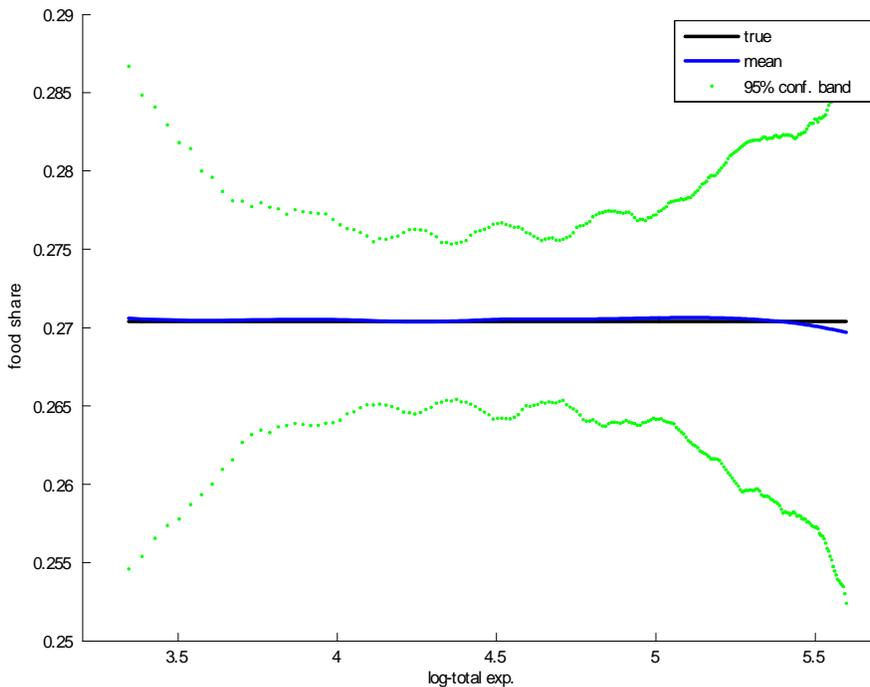

Figure 1: Simulation study, performance of unconstrained estimator at $\tau = 0.5$.

To get a better idea of the ranking of the two estimators, we therefore report pointwise and integrated bias, standard deviation and RMSE in Table 1. From this table, we see that the unconstrained estimator is slightly less biased, but on the other hand exhibits more variance compared to the constrained one; this is particularly evident at values of x in the tail of the distribution of total expenditure. Overall, the constrained estimator has an RMSE that is 17.3% smaller compared to the unrestricted one, and so clearly dominates.

24
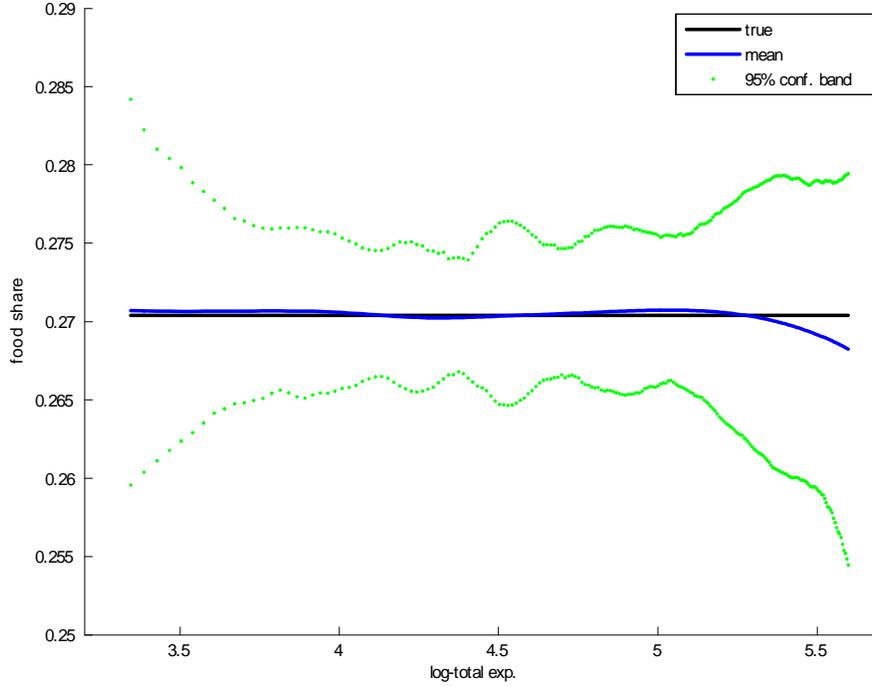
Figure 2: Simulation study, performance of constrained estimator at $\tau = 0.5$.

log (x)	Bias		Standard dev.		RMSE	
	unconstr.	constr.	unconstr.	constr.	unconstr.	constr.
3.3450	0.2028	0.3191	9.8093	7.6865	9.8114	7.6931
4.0518	0.0920	0.1341	3.4781	2.7315	3.4793	2.7348
4.4758	0.1116	-0.0681	3.5716	3.1427	3.5734	3.1434
4.7725	0.1569	0.1717	3.7040	2.9187	3.7073	2.9238
5.0011	0.2256	0.3373	4.0127	2.9841	4.0191	3.0031
5.1870	0.2338	0.2220	5.5910	4.1268	5.5959	4.1328
5.3437	0.0874	-0.2499	7.1645	6.1877	7.1650	6.1927
5.4792	-0.2265	-1.0578	7.6852	7.5625	7.6885	7.6361
5.5984	-0.6958	-2.1580	10.7271	9.7234	10.7496	9.9600
Average	0.0468	-0.1792	5.4468	4.6190	5.4508	4.6473

Table 1: Performance of unconstrained and constrained estimator.

Notes: All numbers have been scaled up by a factor 10^3.

Once the simulated constrained estimators have been computed, we proceed to obtain corresponding demand bounds for new prices, p_0.[7] This allows us to investigate the quality of the

---

[7]The computation of the bounds was done using Matlab code kindly provided by Ian Crawford.

estimated demand bounds in finite samples. In Figure 3, we show the mean and 95% confidence intervals for the estimated demand bounds for good 1 ('food') across a range of $p_{0,1}$, the price of good 1 while we keep $p_{0,2}$ fixed at the price level for good two at time $T = 6$. This is done for a consumer with mean income (x_0 chosen as the population) mean and $\tau = 0.5$. For comparison, we have also plotted the unknown, true bounds for our choices of p_0, x_0 and τ. The estimator suffers from some biases but the 95% confidence interval includes the true bounds and are reasonably tight. Parts of the biases may be due to computational issues since the bounds are here computed by (i) numerically solving for the income expansion path, and (ii) solving a linear programming problem.
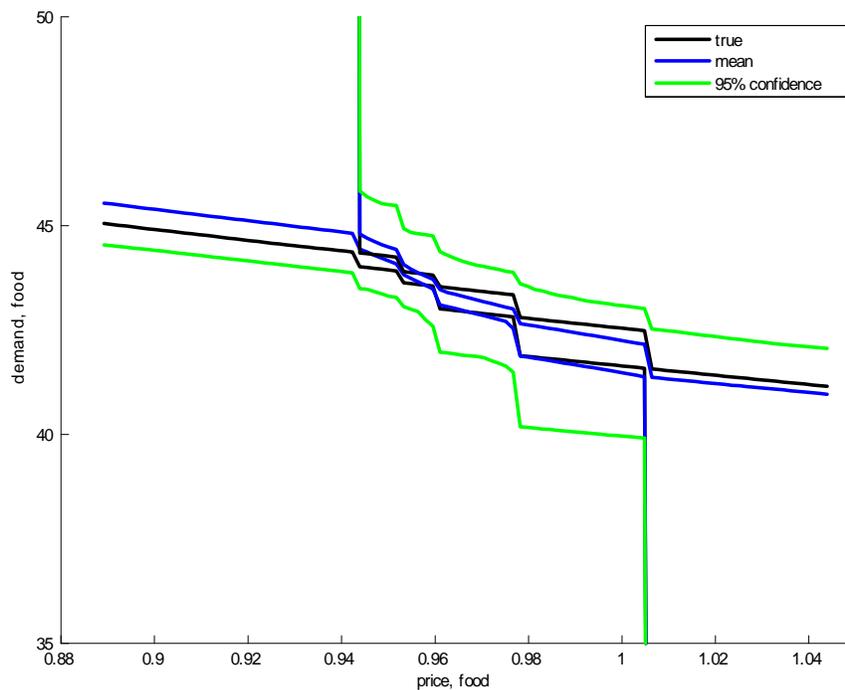
Figure 3: Simulation study, demand bounds at median income and $\tau = 0.5$.

In summary, at observed prices, our demand estimators do very well for the random coefficient Cobb-Douglas models with small biases and variances. Moreover, as expected, the constrained estimator dominates the unconstrained one in terms of MSE. Finally, estimated bounds on (predicted) demands at new prices are somewhat more biased, but still perform satisfactorily.

8 Empirical Application

8.1 Data

In our application we apply the methodology for constructing quantile demand bounds under RP inequality restrictions to data from the British Family Expenditure Survey (FES). The data set contains expenditure data and prices from British households. We use the same sample selection as in BBC08 and we refer to that paper for a more detailed description. We choose food as our primary good, and then group the other goods together in this application. In our application, we focus on FES data for the eight year period 1983-1990. We follow BBC08 and use a group of demographically homogeneous households made up of couples with two children.

The distribution of relative food prices over the central period of the data is give in Figure 4. As a guide to the variation in the expenditure data, the basic distribution of the Engel curve data for the year 1985 are described in Figures 5 and 6. Total expenditure x is defined as total expenditure on non-durables and services. Similar distributions are found for the other years in the data set.
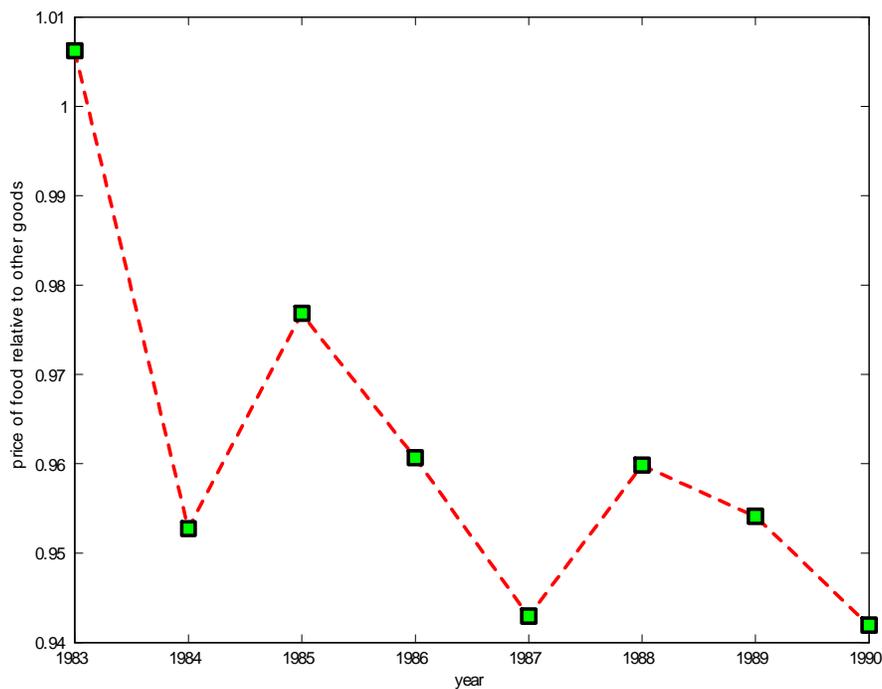

Figure 4: Relative food prices in the FES, 1983 to 1990.
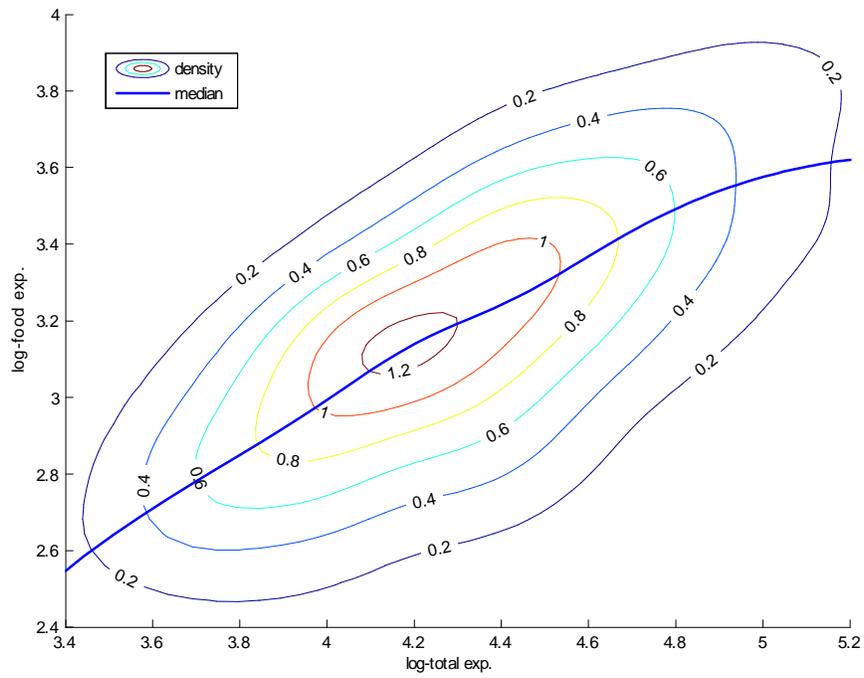
Figure 5: The Engel Curve Distribution, 1985.
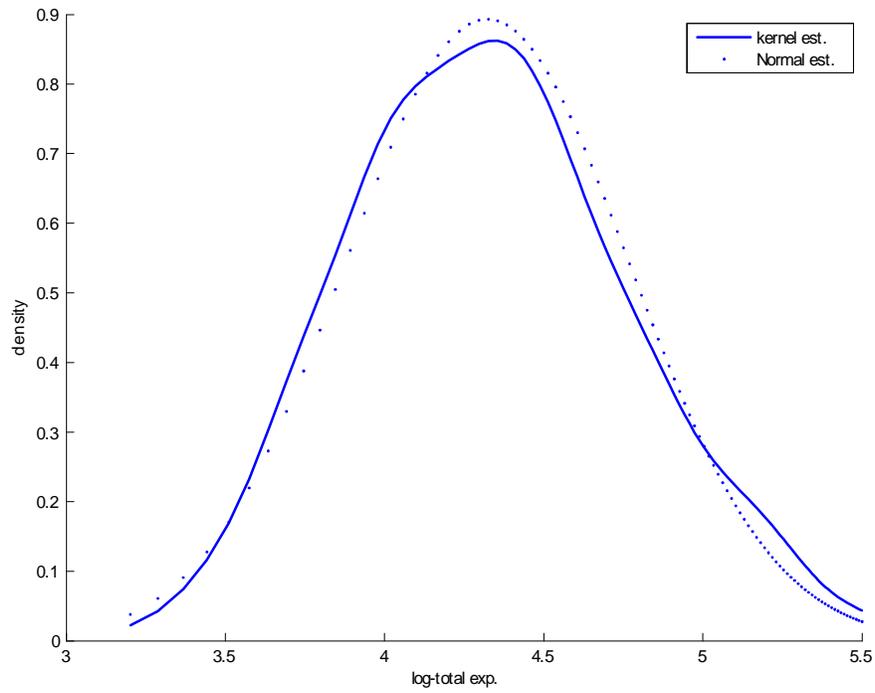

Figure 6: Distribution of total expenditure, 1985.
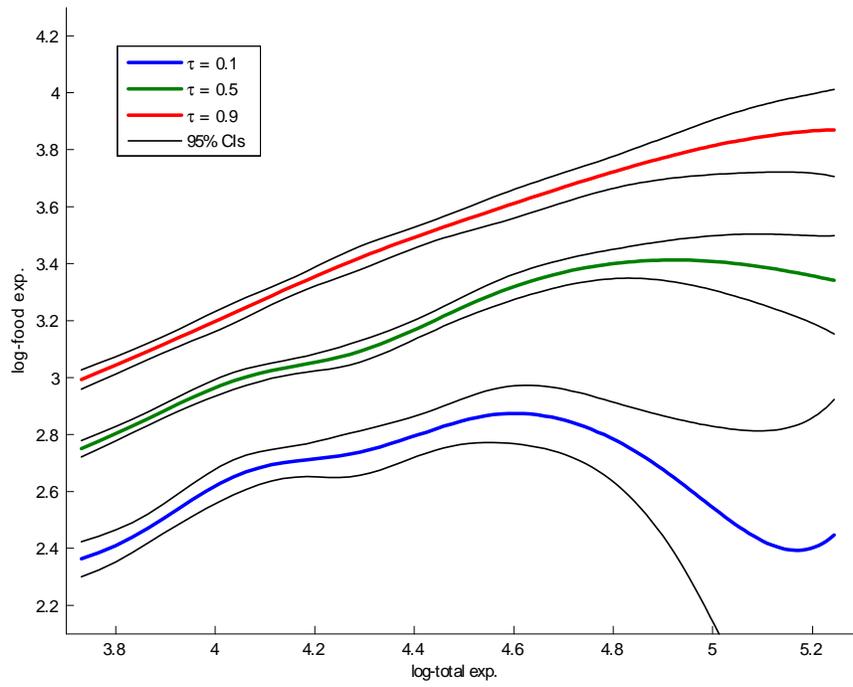
Figure 7: Unconstrained quantile Engel curve estimates, $t = 1983$.
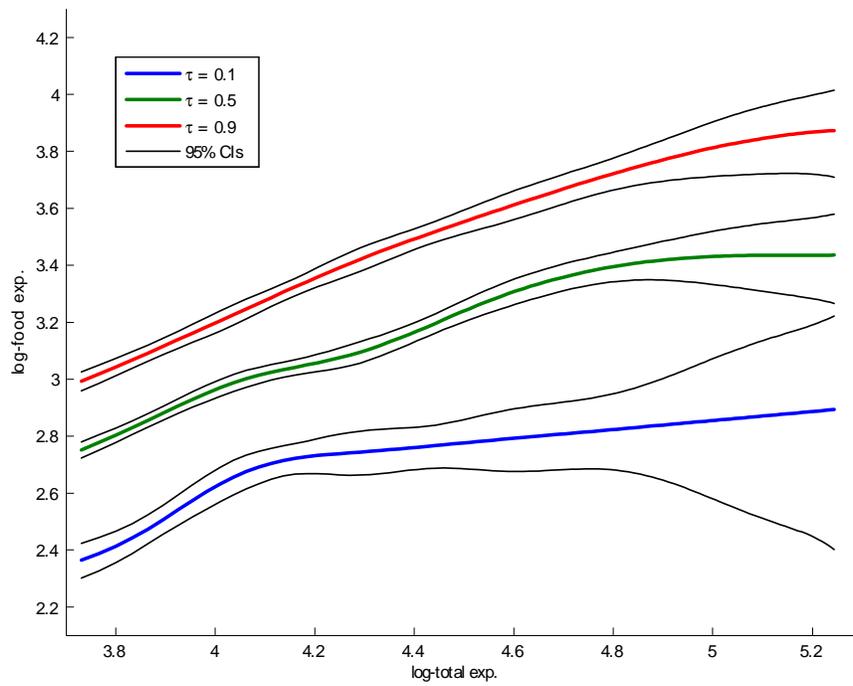
Figure 8: RP Constrained quantile Engel curve estimates, $t = 1983$.

8.2 The Sieve Estimates of Quantile Expansion Paths

In the estimation, we implement the sieve estimator along the lines described in the simulation study. We use a 3rd order polynomial spline ($q_n = 3$) with $r_n = 5$ knots. Each household is defined by a point in the distribution of log income and unobserved heterogeneity ε. As an example, for the year $t = 1983$, the unconstrained expansion paths estimates as a function of x for each of three quantiles ($\tau = 0.1$, 0.5 and 0.90) of the distribution of unobserved heterogeneity are given in Figure 7.

The value of τ in these figures can be interpreted as the taste for food relative to other goods with a higher value of τ reflects stronger preferences for food. We see that the demand functions for the three different types of consumers are similar, but the shape does change as we move across the distribution of unobserved heterogeneity τ. This supports the use of the non-additive demand models that allow for richer interactions between $\log x$ and τ.

Next, we re-estimate the quantile expansion paths (Engel curves) under the RP and monotonicity restrictions. The constrained quantile Engel curve estimates for $t = 1983$ can be found in Figure 8. Comparing the constrained with the unconstrained estimates, imposing monotonicity and RP restrictions tend to remove some of the wiggles found in the unrestricted estimates. The impact of the constraints vary across the different quantiles; for $\tau = 0.90$, the constrained and unconstrained estimators are very close, while substantial shifts in the demand functions happen at $\tau = 0.50$ and $\tau = 0.10$. In particular, the decreases in demand observed at the lower quantiles in Figure 7 are removed. However, the overall shapes remain quite similar.

8.3 Estimated Demand Bounds and Confidence Sets

A key parameter of interest in this study is the distribution of predicted consumer responses for some new relative price $\mathbf{p}_0$ and income x_0. For any x_0, this will allow us to describe the demand curve for a sequence of relative prices. For any price $\mathbf{p}_0$, we estimate bounds (support set) for each quantile demand curve at income x_0 using the RP inequalities. In our FES data we consider bounds on the demand curve at new prices of food while keeping the price of remaining goods fixed at $p_{0,2} = 1$.

We first investigate how precisely the bounds are estimated. In Figure 9, we report the estimated bounds together with the 95% confidence interval across a range of prices for food for a median income consumer. While the estimated bounds are quite narrow, the corresponding confidence

intervals are somewhat larger thus taking into account the sampling uncertainty. We also note that the bounds relatively narrow within the range of observed prices (compare with Figure 4), but for prices far away from observed prices the bounds widen and become less informative.
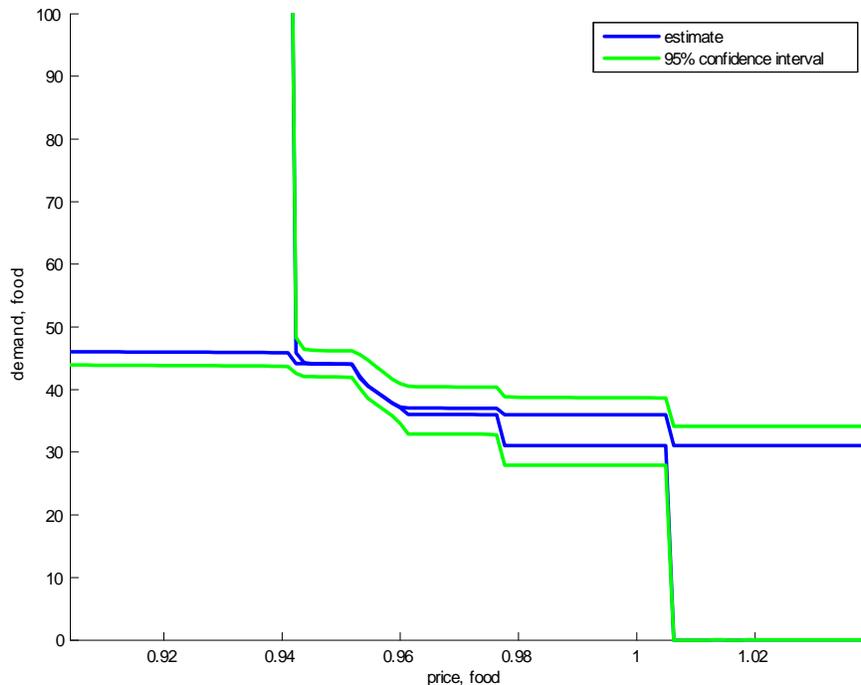
Figure 9: Estimated demand bounds and 95% confidence sets at median income, $\tau = 0.5$, $T = 8$.

Next, we examine how demand responds to changes 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. Each figure contains three sets of bound estimates corresponding to using price information for $T = 4$, 6 and 8 time periods. To avoid too cluttered figures, we only report confidence sets for the bounds for $T = 8$, the *dotted* bounds; the confidence sets for $T = 4$ and 6 are qualitatively the same. Figure 10 shows the estimated confidence sets for the bounds on the quantile demand function at the median income for the 10th percentile ($\tau = .1$) of the unobserved taste distribution. Notice that where the relative prices are quite dense the bounds are correspondingly narrow. Figure 11 contrasts this for a consumer at the 50% ($\tau = .5$) percentile of the heterogeneity distribution - a consumer with stronger taste for food. At all points demands are higher and the price response is somewhat steeper. Figure 12 considers a consumer with an even stronger taste for food - at the 90th percentile ($\tau = .9$) of the

taste distribution. Demand shifts further up at all points. The bounds remain quite narrow where the relative prices are dense.
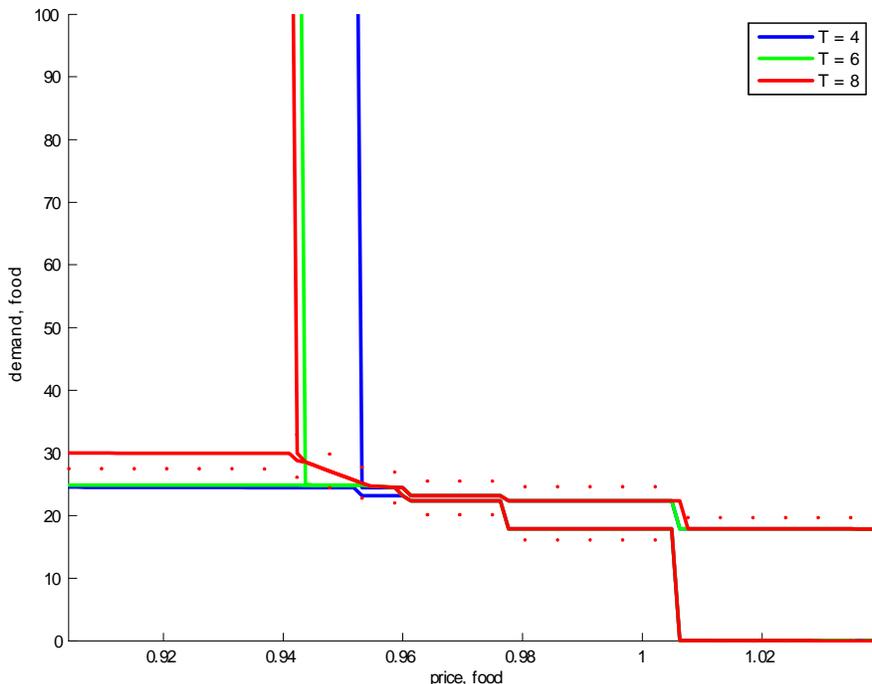
Figure 10: Estimated bounds at median income, $\tau = 0.1$

Finally, we can examine how changes in the total outlay level, x_0, affects the demand bounds. We focus on the median consumer with $\tau = 0.5$. Consider Figure 11, which presents the confidence sets on the demand bounds at median total outlay, as the baseline case. We now decrease the consumers total outlay to the 25th percentile level in the sample; the resulting confidence sets are shown in Figure 13. As expected predicted demand drops uniformly across prices compared to the ones reported for the higher income level (note here that the scale of the y-axis is slightly different from the earlier figures). The sets for the median consumer with outlay x_0 at the 75th percentile of the sample are found in Figure 14. Comparing the two figures, we see that the overall shape remains the same, but that demands bounds are compressed as income levels are decreased.

9   Conclusions and Extensions

This paper has developed a new approach to the estimation of consumer demand models with non-separable unobserved heterogeneity. For general non-additive stochastic demand functions, we
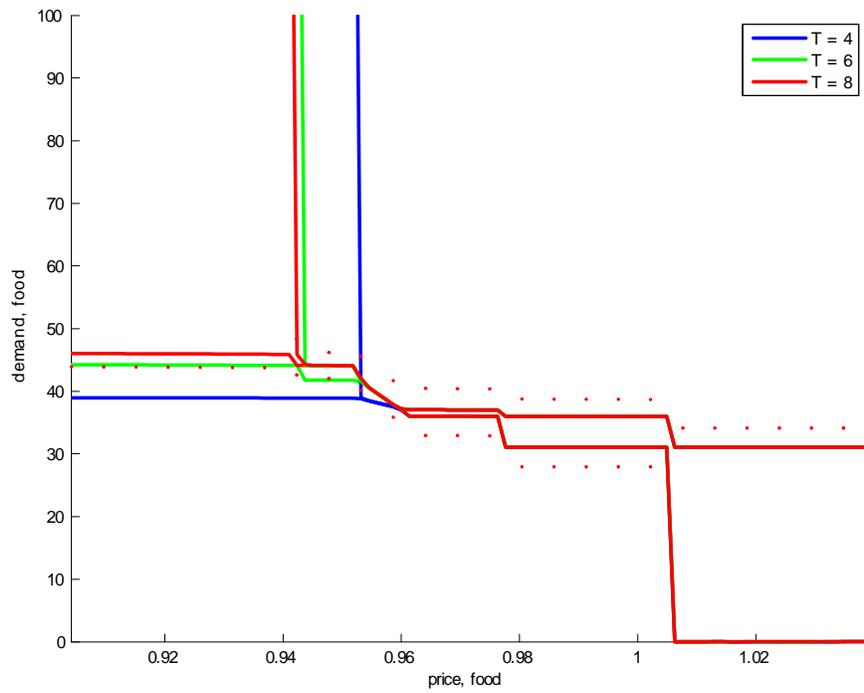
Figure 11: Estimated bounds at median income, $\tau = 0.5$.
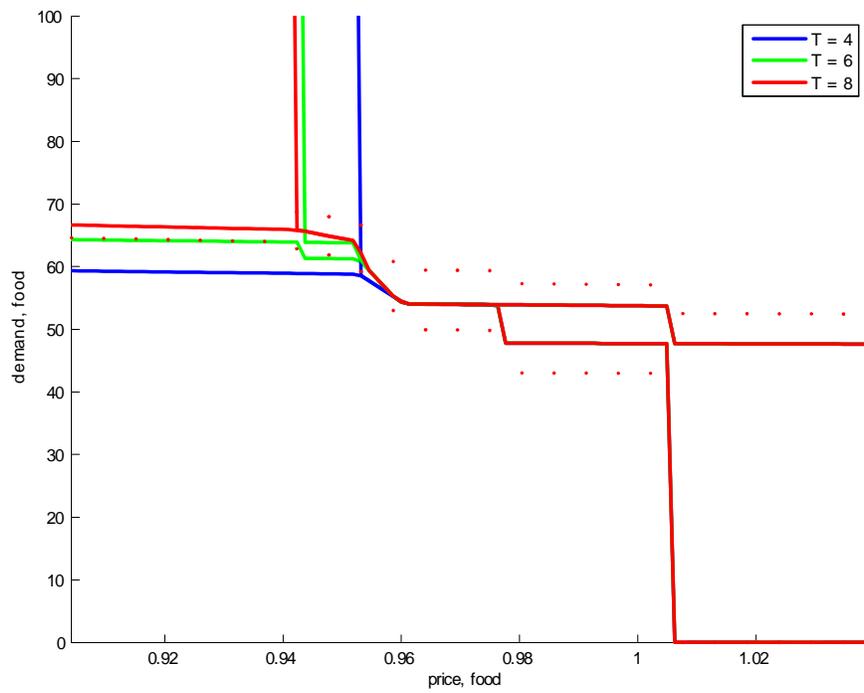
Figure 12: Estimated bounds at median income, $\tau = 0.9$.
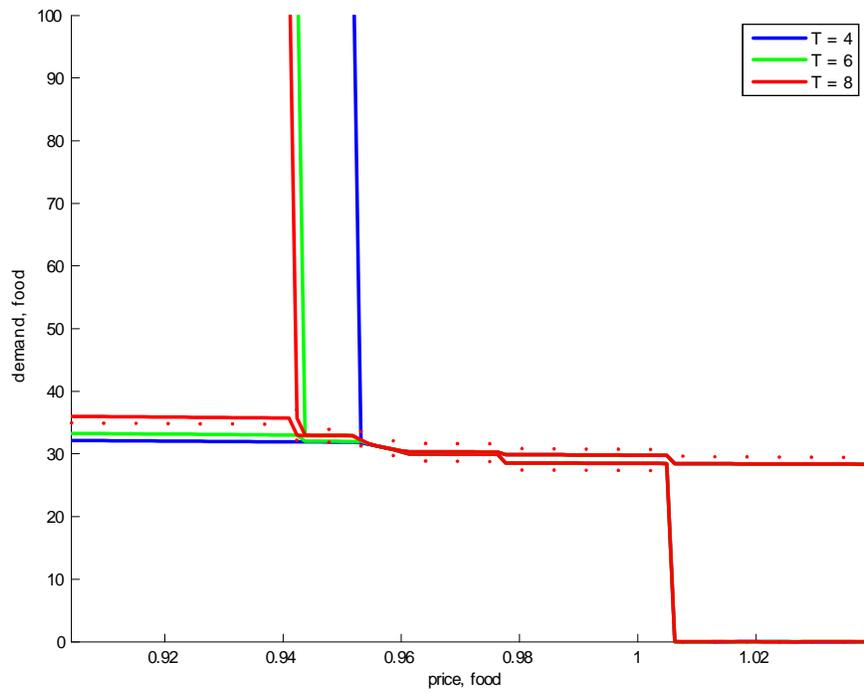
Figure 13: Estimated bounds at 25th percentile income, $\tau = 0.5$.
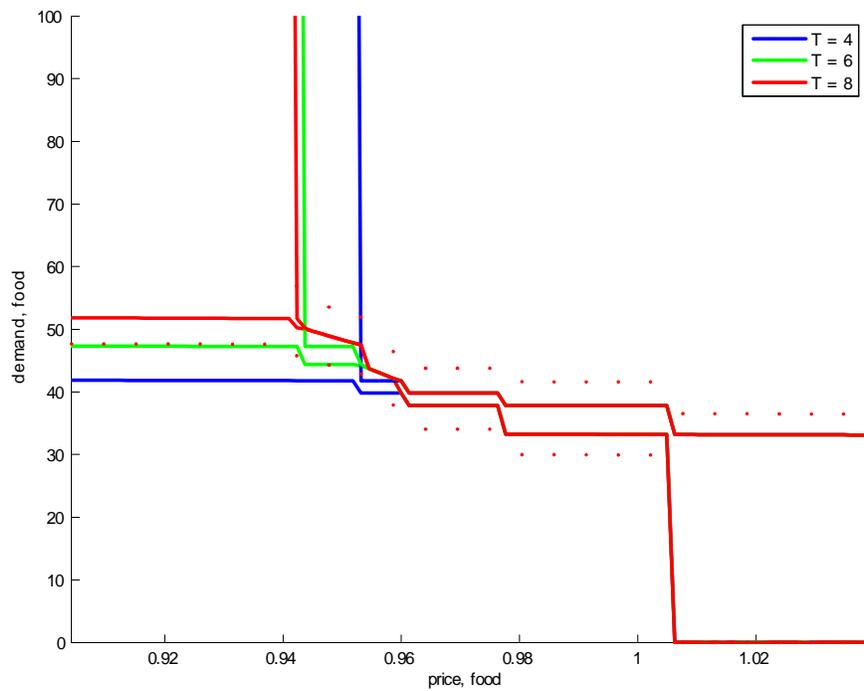
Figure 14: Estimated bounds at 75th percentile income, $\tau = 0.5$.

34

have demonstrated how RP inequality restrictions can be utilized to improve on the nonparametric estimation of demand responses. We have shown how bounds on demand responses to price changes can be estimated, and derive their asymptotic properties using results on the estimation of parameters characterized by moment inequalities.

An empirical application using individual consumer 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 across the distribution of unobserved tastes and different percentiles of the income distribution.

It would be natural to extend our results to allow for endogeneity of the total expenditure variable such that the independence assumption made in (A.2) can be weakened. The proposed sieve quantile estimator will in this case be inconsistent. In a parametric framework this can be dealt with using standard instrumental variables (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 in additive regression models is nonparametric IV (Ai and Chen, 2003; Hall and Horowitz, 2005; Newey and Powell, 2003) and control functions (Newey, Powell and Vella, 1998). Both these methods have been applied in the empirical analysis of Engel curves with additive errors (Blundell, Chen and Kristensen, 2003, and Blundell, Duncan and Pendakur, 1998 respectively). These two approaches have recently been employed in the estimation of quantile models: Chernozhukov, Imbens and Newey (2007) and Chen and Pouzo (2008a,b) develop nonparametric IV methods for quantile models, while Imbens and Newey (2007) consider control function methods. 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.

Finally, it would also 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. A natural way of testing the rationality 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 could take the same approach as in Blundell et al (2008) and develop a minimum-distance statistic based on the unrestricted estimator alone. The hypothesis involves inequality constraints, and so the testing of it falls within the non-standard setting analyzed in Gourieroux et al (1982), Self and Liang (1987) and Wolak (1989,1991). We leave the extension of these results to nonparametric quantile estimation for future research.

References

Ai, C. and X. Chen (2003): "Efficient Estimation of Models with Conditional Moment Restrictions Containing Unknown Functions," *Econometrica*, 71, 1795-1844.

Afriat, S. (1967): "The Construction of a Utility Function from Demand Data," *International Economic Review,* No. 8, pp. 67-77.

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

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

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

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

Banks, J., R. Blundell and A. Lewbel (1997): "Quadratic Engel Curves, Indirect Tax Reform and Welfare Measurement," *Review of Economics and Statistics* LXXIX, 527-539.

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. and V. Chernozhukov and I. Fernandez-Val (2011): "Conditional Quantile Processes under Increasing Dimension," Cemmap working paper CWP19/11.

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.

Blundell, R., A. Duncan and K. Pendakur (1998): "Semiparametric Estimation and Consumer Demand," *Journal of Applied Econometrics* 13, 435-461.

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

Brown, B. W, and M. B. Walker (1989): "The Random Utility Hypothesis and Inference in Demand Systems," *Econometrica*, No. 57, pp. 815-829.

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

Chen, X. and D. Pouzo (2008a): "Estimation of Nonparametric Conditional Moment Models with Possibly Nonsmooth Moments," Cowles Foundation Discussion Papers, No. 1650.

Chen, X. and D. Pouzo (2008b): "Efficient Estimation of Semiparametric Conditional Moment Models with Possibly Nonsmooth Moments," *Journal of Econometrics* 152, 46-60.

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

Cherchye L., B. De Rock B and F. Vermeulen (2011): "The Revealed Preference Approach to Collective Consumption Behavior: Testing and Sharing Rule Recovery," *Review of Economic Studies* 78, 176-198.

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.

Chesher, A. (2007): "Identification of Nonadditive Structural Functions," Chapter 1 of Volume III of *Advances in Economics and Econometrics, Theory and Applications, 9th World Congress* (eds. R. Blundell, T. Persson and W. Newey), Cambridge: Cambridge University Press.

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

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

Gourieroux, C., A. Holly and A. Monfort (1982): "Likelihood Ratio Test, Wald Test and Kuhn-Tucker Test in Linear Models with Inequality Constraints in the Regression Parameters," *Econometrica* 50, 63-80.

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. 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 and J.L. Powell (1995): "Nonlinear Errors in Variables: Estimation of Some Engel Curves," *Journal of Econometrics* 65, 205-233.

Hoderlein, S. and E. Mammen (2008): "Identification and Estimation of Local Average Derivatives in Non-separable Models Without Monotonicity," *Econometrics Journal* 12, 1-25.

Hoderlein, S. and E. Mammen (2007): "Identification of Marginal Effects in Nonseparable Models Without Monotonicity," *Econometrica* 75, 1513-1518.

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

Houthakker, H. S. (1950): "Revealed Preference and the Utility Function," *Economica,* New Series, Vol. 17, No. 66, pp. 159-174.

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.

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.

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.

Lewbel, A. and S. Ng (2005): "Demand Systems with Nonstationary Prices," *Review of Economics and Statistics* 87, 479-494.

Ma, L. and R. Koenker (2006): "Quantile Regression Methods for Recursive Structural Equation Models," *Journal of Econometrics* 134, 471-506.

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

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

Manski, C. (1993): "Identification Problems in the Social Sciences," *Sociological Methodology* 23, 1-56.

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

Matzkin, R.L. (2008): "Identification in Nonparametric Simultaneous Equations," *Econometrica* 76, 945-978

Matzkin, R.L. (2010): "Estimation of Nonparametric Models with Simultaneity," mimeo, UCLA.

McElroy, M. B. (1987): "Additive General Error Models for Production, Cost, and Derived Demand or Share Systems," *Journal of Political Economy,* Vol. 95, No. 4, pp. 737-757.

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.

Richter, M. K. (1966): "Revealed Preference Theory," *Econometrica,* Vol. 34, No. 3, pp. 635-645.

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

Samuelson, P. A. (1938): "A Note on the Pure Theory of Consumer Behavior," *Economica,* No. 5, pp. 61-71.

Self, S.G. and K. Liang (1987): "Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests under Nonstandard Conditions," *Journal of the American Statistical Association* 82, 605-610.

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.

Varian, H.R. (1982): "The Nonparametric Approach to Demand Analysis," *Econometrica* 50, 945-973.

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

Wolak, F.A. (1989): "Local and Global Testing of Linear and Nonlinear Inequality Constraints in Nonlinear Econometric Models, *Econometric Theory* 5, 1-35.

Wolak, F.A. (1991): "The Local Nature of Hypothesis Tests Involving Nonlinear Inequality Constraints," *Econometrica* 59, 981-995.

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

A Proofs of Theorems 1-3

Proof of Theorem 1. We write the first demand equation as a quantile regression,

$$q_1(t) = d_1(x, t, \tau) + e(t, \tau), \qquad (21)$$

where $e(t, \tau)$ is defined as the generalized residual, $e(t, \tau) := d_1(x, t, \varepsilon) - d_1(x, t, \tau)$. This formulation of the model for corresponds to the quantile regression considered in Chen (2007, Section 3.2.2). We then verify the conditions stated there. First, we note that the distribution of $e(\tau)|x$ is described by the density $f(e|x, t, \tau)$ given in eq. (9). We claim that

$$0 < \inf_{x \in \mathcal{X}} f(0|x, t, \tau) \leq \sup_{x \in \mathcal{X}} f(0|x, t, \tau) < \infty, \qquad (22)$$

$$\sup_{x \in \mathcal{X}} |f(e|x, t, \tau) - f(0|x, t, \tau)| \to 0, \quad |e| \to 0. \qquad (23)$$

From the definition of expression it is easily seen that eq. (22) holds since $d_1(x, t, \varepsilon)$ and its derivative w.r.t. ε are continuous in x and $\mathcal{X}$ is compact. eq. (23) clearly holds pointwise due to the continuity of $\varepsilon \mapsto d_1(x, t, \varepsilon)$. This can be extended to uniform convergence since $\sup_{x \in \mathcal{X}, e \in [0,1]} f(e|x, t, \tau) < \infty$.

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

$$||\hat{\mathbf{d}}(\cdot, t, \tau) - \mathbf{d}(\cdot, t, \tau)||_2 = O_P(\max\{\delta_n, \|\pi_n d_1(\cdot, t, \tau) - d_1(\cdot, t, \tau)\|_2\})$$

where

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

and $\pi_n d_1$ is an element in $\mathcal{D}_{n,1}$. Here, $H_{[]}(w, \mathcal{F}_n(\delta), \|\cdot\|_2) = \log\left(N_{[]}(w, \mathcal{F}_n(\delta), \|\cdot\|_2)\right)$ denotes the log of 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 $\|\pi_n d_{1,0}(\cdot, t, \tau) - d_1(\cdot, t, \tau)\|_2 = O(k_n^{-m})$.

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 $\hat{d}_1\left(x\left(t\right),t,\tau\right)$, $t = 1, ..., T$. This will follow from Chen et al (2010, Corollary 6.1) if their Conditions 6.1-6.2 hold under our assumptions. Their Condition 6.1 is shown to hold in the Proof of Theorem 1. Their Condition 6.2(i) holds since

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

where we have used that d_1 is continuously differentiable, while Condition 6.2(iii) holds by assumption. To verify their Conditions 6.2(ii) and (iv), first note that, since we are using splines, $\xi_0\left(k_n\right) := \sup_{x\in\mathcal{X}}\left\|\underline{B}_{k_n}\left(x\right)\right\| \leq c\sqrt{k_n}$. Thus, their Condition 6.2(iv) becomes $\xi_0^2\left(k_n\right)k_n^3/n \simeq k_n^4/n = O\left(1\right)$ and $\xi_0\left(k_n\right)k_n^{-3m}n = k_n^{-3m+1/2}n = O\left(1\right)$. Finally, the condition (ii) of their Corollary 6.1 becomes $nk_n^{-2m-1} = O\left(1\right)$ ∎

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

$$\widehat{\tilde{x}}_\tau\left(T-1\right) - \tilde{x}_\tau\left(T-1\right) = \mathbf{p}\left(T-1\right)'\left[\hat{\mathbf{d}}\left(x_\tau\left(T\right),T,\tau\right) - \mathbf{d}\left(x_\tau\left(T\right),T,\tau\right)\right] = O_P\left(r_n\right).$$

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

$$
\begin{aligned}
&\left\{\widehat{\tilde{x}}_\tau\left(t\right) - \mathbf{p}\left(t\right)'\hat{\mathbf{d}}\left(\widehat{\tilde{x}}_\tau\left(s\right),s,\tau\right)\right\} - \left\{\tilde{x}_\tau\left(t\right) - \mathbf{p}\left(t\right)'\mathbf{d}\left(\tilde{x}_\tau\left(s\right),s,\tau\right)\right\}\\
=\ &\left\{\widehat{\tilde{x}}_\tau\left(t\right) - \tilde{x}_\tau\left(t\right)\right\} + \mathbf{p}\left(t\right)'\left\{\hat{\mathbf{d}}\left(\widehat{\tilde{x}}_\tau\left(s\right),s,\tau\right) - \mathbf{d}\left(\widehat{\tilde{x}}_\tau\left(s\right),s,\tau\right)\right\}\\
&+\mathbf{p}\left(t\right)'\left\{\mathbf{d}\left(\widehat{\tilde{x}}_\tau\left(s\right),s,\tau\right) - \mathbf{d}\left(\tilde{x}_\tau\left(s\right),s,\tau\right)\right\}\\
=\ &O_P\left(r_n\right),
\end{aligned}
$$

Thus, since $\tilde{x}_\tau\left(t\right) \leq \mathbf{p}\left(t\right)'\mathbf{d}\left(\tilde{x}_\tau\left(s\right),s,\tau\right)$, we have $\epsilon\widehat{\tilde{x}}_\tau\left(t\right) \leq \mathbf{p}\left(t\right)'\hat{\mathbf{d}}\left(\widehat{\tilde{x}}_\tau\left(s\right),s,\tau\right)$ with probability approaching one (w.p.a.1) as $r_n \to 0$. This proves that $\hat{\mathbf{d}} \in \mathcal{D}_{C,n}^T\left(\epsilon\right)$ w.p.a.1 such that $\hat{\mathbf{d}}_C^\epsilon = \hat{\mathbf{d}}$ w.p.a.1 as $r_n \to 0$. 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. It should be emphasized though that the e-bounds in a general economy with more than two goods are not necessarily sharp since they do not utilize all constraints implied by rationality.

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

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

which is compact and convex. Suppose that we have observed T prices, $\mathbf{p}(1), ..., \mathbf{p}(T)$, $\mathbf{p}(t) = (p_1(t), ..., p_{L+1}(t))'$, and let $\mathbf{d}(x(t), t) = (\mathbf{d}_1(x(t), t), ..., \mathbf{d}_{L+1}(x(t), t))'$, $t = 1, ..., T$, denote the consumer's corresponding demand functions where we suppress dependence on ε since this is kept fixed. Since the demand function has to satisfy $\mathbf{p}(t)' \mathbf{d}(x, t) = x$, the demand for the $(L+1)$th good is simply given as

$$\mathbf{d}_{L+1}(x, t) = \frac{x - \mathbf{p}_{1:L}(t)' \mathbf{d}_{1:L}(x, t)}{p_L(t)}. \tag{24}$$

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. (24), the left hand side of the above equation can be rewritten as

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

Also note that we can rewrite the support set as

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

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. (24). 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 monotonically 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,

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

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

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

The monotonicity requirement in Condition (C.1) ensures that the intersection income path $\{\bar{x}\left(t\right)\}$ 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 $\Omega_n\left(x,t\right)$ and a scalar r_n. The condition states that once the demand estimator has been normalized by $\Omega_n^{1/2}\left(x,t\right)$ it converges with rate $\sqrt{r_n}$. (C.3) is a further strengthening and states that the estimator when normalized by $\sqrt{r_n}\Omega_n^{1/2}\left(x,t\right)$ 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 $\Omega_n\left(x,t\right)=I_L$. With nonparametric estimators, one may potentially choose $\Omega_n\left(x,t\right)$ 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 $\Omega_n\left(x,t\right)=I$ and $r_n=O(k_n/\sqrt{n})+O\left(k_n^{-m}\right)$ with no restrictions on the sequence k_n; to obtain distributional results, we will choose $\Omega_n\left(x,t\right)=\Sigma_n^{-1}\left(x,t\right)$ 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 5 *Assume that (C.1)-(C.2) hold. Then*

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

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

$$\sqrt{r_nw_n\left(t\right)}\left(\hat{x}\left(t\right)-\bar{x}\left(t\right)\right)\rightarrow^d N\left(0,1\right),$$

where

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

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

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

44

and

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

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

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

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 $\hat{x}\left(t\right) \to^P \bar{x}\left(t\right)$. 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 sufficiently small neighborhood of $\bar{x}\left(t\right)$,

$$G\left(x,t\right) - G\left(\bar{x}\left(t\right),t\right) = \frac{\partial G\left(\tilde{x}\left(t\right),t\right)}{\partial x}\left[x - \bar{x}\left(t\right)\right]$$

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

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

Given consistency, we therefore have

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

Next, by a first-order Taylor expansion,

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

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

$$\frac{\partial\hat{G}\left(\tilde{x}\left(t\right),t\right)}{\partial x} \to^P \frac{\partial G\left(\bar{x}\left(t\right),t\right)}{\partial x} = \mathbf{p}_0'\frac{\partial\mathbf{d}(\bar{x}\left(t\right),t)}{\partial x} > 0. \tag{25}$$

45

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

$$\left[\mathbf{p}_0' \frac{\partial \mathbf{d}(\bar{x}(t),t)}{\partial x}\right]^{-1} \hat{G}(\bar{x}(t),t)$$

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

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

$$= : a_n(t)' \left\{V^{-1/2}\Omega_n^{1/2}(\bar{x}(t),t)\Delta_n(t)\right\},$$

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,

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

∎

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 lemma simplifies to $\sqrt{r_n}(\hat{x}(t) - \bar{x}(t)) \to^d N(0,w^2(t))$, 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, $Q_n(\mathbf{q}) = \left\|\hat{W}_n^{1/2}[\hat{\mathbf{x}} - \mathbf{Pq}]\right\|_+^2$, with its limit given by $\bar{Q}_n(\mathbf{q}) = \left\|W_n^{1/2}[\bar{\mathbf{x}} - \mathbf{Pq}]\right\|_+^2$. 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

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

This motivates us to define our support set estimator as

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

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 6 *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

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

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 $\xi(\mathbf{q}) = (\xi_1(\mathbf{q}), \ldots, \xi_T(\mathbf{q}))'$ is given by

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

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 $\mathbf{q} \in \mathcal{B}_{\mathbf{p}_0, x_0}$,

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

since $\hat{W}_n^{1/2} \{\hat{\mathbf{x}} - \bar{\mathbf{x}}\} = O_P(1/\sqrt{r_n})$ by Lemma 5. 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 $\left\|\sqrt{r_n}\hat{W}_n^{1/2}\bar{m}\left(\mathbf{q}\right)\right\|_+^2 \geq r_n w_n^* C^2 \rho^2\left(\mathbf{q}, \mathcal{S}_{\mathbf{p}_0,x_0}\right)$ by Lemma 7 below. By the same arguments as in CHT, Proof of Theorem 4.2(Step 1), it now follows that $r_n Q_n\left(\mathbf{q}\right) \geq r_n w_n^* C^2 \rho^2\left(\mathbf{q}, \mathcal{S}_{\mathbf{p}_0,x_0}\right)/2$ w.p.a 1. This shows that Condition C.1-C.2 of CHT hold in our case as well, except that the limiting objective function $\bar{Q}_n\left(\mathbf{q}\right)$ and the constant $\kappa = \kappa_n = w_n^* C^2$ 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 $\mathbf{q}$,

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

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

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

where $c_n \to^P c < \infty$. This proves that the stochastic process $\mathbf{q} \mapsto \left\{\hat{W}_n^{1/2} m_n\left(\mathbf{q}\right) - W_n^{1/2}\bar{m}\left(\mathbf{q}\right)\right\}$ 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, $\hat{W}_n^{1/2} m_n\left(\mathbf{q}\right) = Z_n + W_n^{1/2}\bar{m}\left(\mathbf{q}\right) + o_P\left(1\right)$ uniformly in $\mathbf{q}$, which in turn implies that, by Slutsky's theorem,

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

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

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

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

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

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

Lemma 7 *Under (C.5),* $\left\|W_n^{1/2}\bar{m}\left(\mathbf{q}\right)\right\|_+^2 \geq w_n^* C^2 \rho\left(\mathbf{q}, \mathcal{S}_{\mathbf{p}_0,x_0}\right)$ *for some constant* $C < \infty$ *and* $w_n^* = \min_{t=1,...,T} w_n\left(t\right)$.

Proof. The inequality is trivial for $\mathbf{q} \in \mathcal{S}_{\mathbf{p}_0,x_0}$. Consider any $\mathbf{q} \in \mathcal{B}_{\mathbf{p}_0,x_0}\backslash\mathcal{S}_{\mathbf{p}_0,x_0}$: Let $\mathbf{q}^* = \arg\min_{\mathbf{q}'\in\mathcal{S}_{\mathbf{p}_0,x_0}}\left\|\mathbf{q} - \mathbf{q}'\right\|$ 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 $\left\|\delta^*\right\| = \rho\left(\mathbf{q}, \mathcal{S}_{\mathbf{p}_0,x_0}\right)$. We can decompose the rows of $\left(\mathbf{P}, \bar{\mathbf{x}}\right)$ into binding and non-binding constraints respectively of $\mathbf{q}^*$. Let $\left(\mathbf{P}^{(1)}, \bar{\mathbf{x}}^{(1)}\right)$ and $\left(\mathbf{P}^{(2)}, \bar{\mathbf{x}}^{(2)}\right)$, with $\mathbf{P}^{(1)} = \left[\mathbf{p}^{(1)}\left(1\right), ..., \mathbf{p}^{(1)}\left(T_1\right)\right]' \in \mathbb{R}^{T_1\times(L+1)}$ and $\bar{\mathbf{x}}^{(1)} = (x^{(1)}\left(1\right), ..., x^{(1)}\left(T_1\right))' \in \mathbb{R}^{T_1}$ for some $T_1 \leq L + 1$, denote the set of rows which contain the binding and non-binding constraints respectively. That is, $\bar{m}^{(1)}\left(\mathbf{q}^*\right) := \bar{\mathbf{x}}^{(1)} - \mathbf{P}^{(1)}\mathbf{q}^* = 0$ while $\bar{m}^{(2)}\left(\mathbf{q}^*\right) := \bar{\mathbf{x}}^{(2)} - \mathbf{P}^{(2)}\mathbf{q}^* < 0$. The $(T_1 \times T_1)$-matrix

$\mathbf{P}^{(1)}\mathbf{P}^{(1)\prime}$ must necessarily have rank T_1 with its eigenvalues bounded above away from zero. Thus, for some $c_1 > 0$,

$$c_1 \|\delta^*\| \le \left\|\mathbf{P}^{(1)}\delta^*\right\| \le T_1 \max_{t=1,\dots,T_1} |\mathbf{p}^{(1)}(t)' \delta^*|.$$

Moreover, $\mathbf{p}^{(1)}(t)' \delta^* \ge 0$ for all $t \in \{1,\dots,T_1\}$. As a consequence, with $C = c_1/T$, there exists at least one $t_0 \in \{1,\dots,T_1\}$ such that $C\|\delta^*\| \le \mathbf{p}^{(1)}(t_0)' \delta^*$. We then obtain

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

∎