Journal of Econometrics 179 (2014) 112–127

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Bounding quantile demand functions using revealed preference inequalities

Richard Blundell [a,b], Dennis Kristensen [a,b,c,*], Rosa Matzkin [d]

[a] Department of Economics, UCL, United Kingdom
[b] Institute for Fiscal Studies, United Kingdom
[c] CREATES, University of Aarhus, United Kingdom
[d] Department of Economics, UCLA, United States

ARTICLE INFO

Article history:
Received 27 March 2012
Received in revised form
15 October 2013
Accepted 1 January 2014
Available online 30 January 2014

JEL classification:
C20
D12

Keywords:
Consumer behavior
Revealed preference
Bounds
Quantile regression

ABSTRACT

This paper develops a new approach to 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 counterfactual demand. A nonparametric estimator for these bounds is developed and asymptotic properties are derived. An empirical application using data from the UK Family Expenditure Survey illustrates the usefulness of the methods.

1. Introduction

This paper develops a new approach to the estimation and prediction of individual 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 individual consumer demand. We propose both unconstrained and RP constrained nonparametric estimators for individual demand functions with nonadditive 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 Muellbauer (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 are not typically additive in the unobserved tastes. The identification and estimation of consumer demand models that are consistent with unobserved taste variation therefore require analyzing demand models with nonadditive random terms.

An early treatment of identification of semiparametric nonadditive models is Brown (1983) whose results were extended to nonparametric models in Roehrig (1988). Building on their work, Matzkin (2003) derives nonparametric identification and quantile-driven estimation in one equation non-additive models, and Matzkin (2008) derives nonparametric identification in simultaneous equations non-additive models. A number of authors have addressed identification and estimation in triangular models. Among these, Chesher (2003, 2007) considers quantile-driven identification while Chernozhukov et al. (2007a) and Imbens and Newey (2009) develop quantile-based nonparametric estimators. Our approach draws on this literature.

Our proposed procedure incorporates nonadditive methods and inequality restrictions derived from economic theory. 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 (1938), Richter (1966), Houthakker

* Correspondence to: Department of Economics, University College London, Gower Street, London WC1E 6BT, United Kingdom. Tel.: +44 0 20 7679 5846.
E-mail addresses: r.blundell@ucl.ac.uk (R. Blundell), d.kristensen@ucl.ac.uk (D. Kristensen), matzkin@econ.ucla.edu (R. Matzkin).

http://dx.doi.org/10.1016/j.jeconom.2014.01.005
0304-4076/

(1950), Afriat (1967) and Varian (1982). Our analysis follows Varian (1982), where the inequalities developed in Afriat (1973) are used to characterize bounds on individual demand responses to new prices. We extend the RP approach of Afriat and Varian to the case where demand observations are from repeated cross sectional data. This requires additional restrictions to connect identical preferences across budgets. First of all, we have to assume that the unobserved preferences remain stable over time. Under this assumption, Blundell et al. (2008) connect the *average* consumer across incomes and prices, and develop bounds on the demand of this consumer under new prices.

In the first part of this paper, we provide a general theory for inference on counterfactual demand bounds using RP inequalities. We take as starting point the availability of an estimator of the demand functions at observed prices. We take no stand on the underlying identification scheme and the precise nature of this estimator, and only require it to satisfy weak regularity conditions, including consistency and pointwise asymptotic normality. The conditions allow for the demand function estimator (and thereby the bounds) to be nonparametric, semiparametric or parametric and so should cover all relevant scenarios. Under these high-level conditions, we show the corresponding estimated bounds are consistent and derive tools for constructing confidence sets.

In the second part of the paper, we consider a particular demand function estimator based on quantiles of the unobserved component entering consumers' preferences. We connect consumers across budgets by mapping each of them into a quantile of the heterogeneity distribution. Formally, we assume that the demand of each consumer can be described by a function of income (and potentially other observed characteristics) together with an unobserved component capturing tastes and other individual-specific unobserved characteristics. Assuming that the demand function is invertible w.r.t. unobserved component, a particular point in the (conditional) distribution of demand corresponds to a unique value of the unobserved taste. In this setting, 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.

We then develop specific nonparametric conditional quantile-type estimators of demand, and show that the general theory of the first part of the paper applies to these estimators. We focus on the case of two goods and a scalar unobserved component with the idea being that when a demand function depends monotonically on only one unobservable random term, the function can be identified from the conditional distribution of demand, given prices and income. This identification and estimation scheme is straightforward to extend to the case of multiple goods if one maintains the assumption that demand for each good is a function of a scalar error component.

Another relevant extension of the proposed quantile estimator would be to allow for multiple unobservables entering the demand for each good. However, identification and estimation of demand functions in this setting requires in general methods for simultaneous equations, which are usually more demanding in terms of assumptions and estimation methods than the class of models considered here (see Matzkin, 2008). Assume, as we do in most of the paper, that the unobservables are distributed independently of prices and income. When each demand function depends on a vector of unobservable random terms, the system of demand functions cannot in general be identified, at each of the quantiles of the marginal distributions of unobservables, from only the conditional distribution of the vector of demands, given prices and income, even when the system of demand functions is invertible in the vector of unobservables (see Benkard and Berry, 2006) and Example 3 in (Matzkin, 2007). Further restrictions are needed. One could, for example, consider representing the system of demands as a triangular system of equations, and estimating the equations sequentially using conditional quantile methods. However, the set of simultaneous equations that are observationally equivalent to triangular systems possess very restrictive properties (see Blundell and Matzkin, forthcoming; Blundell et al., 2013a).

The problem of estimating counterfactual demand using RP inequalities falls within the framework of partially identified models (see e.g. Manski, 1993). We employ the techniques developed in, amongst others, Chernozhukov et al. (2007b) to establish the properties of the demand bounds estimators. Our aim here is to develop bounds on the quantiles of predicted (counterfactual) demands, while we do not directly address testing the revealed preference restrictions. There is a long history of studies that have combined nonparametric techniques to test restrictions from consumer theory; see Lewbel (1995), Haag et al. (2009) and Blundell, Horowitz and Parey (2012) and references therein. These methods are not directly applicable to the revealed preference inequalities in the quantile demand framework we consider here. More recently, Hoderlein and Stoye (forthcoming), Hoderlein and Stoye (2013) and Kitamura and Stoye (2012) have developed an attractive alternative approach to testing that employs stochastic revealed preference inequalities (McFadden and Richter, 1991; McFadden, 2005). Their method focuses on the behavior of partitions of observed budgets that are consistent with the existence of a distribution of preferences generating the observed distribution of demand. They thus require weaker conditions on unobserved heterogeneity.

The remainder of the paper is organized as follows: In Section 2, we set up our framework for modeling heterogeneous consumer choice. A general theory for estimation of demand function bounds is developed in Section 3. In Section 4 we propose sieve estimators for the quantile Engel curves in a two-good economy. In Section 5 we discuss the implementation of the estimator and examine how to compute confidence sets. We then 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 in Section 6. Section 7 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 and lemmas have been relegated to Appendices A and B respectively.

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 $\mathbf{p}(t) \in \mathbb{R}_+^{L+1}$ be the set of prices for the $L+1$, $L \geq 1$, goods that all consumers face at time $t = 1, \ldots, T$. At each time point t, we draw a new random sample of $n \geq 1$ consumers. For each consumer, we observe his or her demands and income level (and potentially some other individual characteristics such as age, education etc., which we suppress in this discussion).

Let $\mathbf{q}_i(t) \in \mathbb{R}_+^{L+1}$ and $x_i(t) \in \mathbb{R}_+$ be consumer i's ($i = 1, \ldots, n$) vector of demand and income level at time t ($t = 1, \ldots, T$). We stress that the data $\{\mathbf{p}(t), \mathbf{q}_i(t), x_i(t)\}$, for $i = 1, \ldots, n$ and $t = 1, \ldots, T$, 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.

The demand $\mathbf{q}(t) = (q_1(t), \ldots, q_{L+1}(t))'$ for a given consumer is assumed to satisfy

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

for some vector function $\mathbf{d}(x(t), \mathbf{p}(t), \varepsilon) = (\mathbf{d}_1(x(t), \mathbf{p}(t), \varepsilon), \ldots, \mathbf{d}_L(x(t), \mathbf{p}(t), \varepsilon))'$ where ε is a *time-invariant* individual specific heterogeneity term that reflects unobserved heterogeneity in preferences and characteristics.[1] In this and the following section, the dimension of ε is left unrestricted and may even be infinite-dimensional. When we develop the quantile estimators of $\mathbf{d}$ in Section 4, we will however impose further restrictions. To ensure that the budget constraint is met, the demand for good $L+1$ must satisfy:

$$q_{L+1}(t) = d_{L+1}(x(t), \mathbf{p}(t), \varepsilon) := \frac{x(t) - \mathbf{p}_{1:L}(t)' \mathbf{d}_{1:L}(x(t), \mathbf{p}(t), \varepsilon)}{p_L(t)}, \qquad (1)$$

where $\mathbf{p}_{1:L}(t)$ and $\mathbf{d}_{1:L}$ are the first L elements of $\mathbf{p}(t)$ and $\mathbf{d}$ respectively. The demand function $\mathbf{d}$ should be thought of as the solution to an underlying utility maximization problem over the subset of goods 1 through $L+1$.

We consider the 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 $\mathbf{p} \mapsto \mathbf{d}(x, \mathbf{p}, \varepsilon)$. On the other hand, as we shall see, it is possible to identify the function $(x, \varepsilon) \mapsto \mathbf{d}(x, \mathbf{p}(t), \varepsilon)$ at each of the observed prices under suitable regularity conditions. 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 Engel curves, $\{\mathbf{d}(x, t, \varepsilon)\}_{t=1}^T$. One consequence of this partial identification is that we cannot point identify demand responses to a new price, say $\mathbf{p}_0 \neq \mathbf{p}(t)$, $t = 1, \ldots, T$. Instead we propose to use revealed preference (RP) constraints involving $\{\mathbf{d}(x, t, \varepsilon)\}_{t=1}^T$ to construct bounds for such counterfactual demands.

2.2. Bounds on quantile demand functions

Consider a particular consumer characterized by some ε with associated demand function $\mathbf{d}(x, \mathbf{p}, \varepsilon)$. Suppose that the consumer faces a given new price $\mathbf{p}_0$ at an income level x_0. Without full knowledge of $\mathbf{d}(x, \mathbf{p}, \varepsilon)$, what can we learn about the demand for this consumer, $\mathbf{q}_0 = \mathbf{d}(x_0, \mathbf{p}_0, \varepsilon)$?

Suppose that for a given sequence of prices $\{\mathbf{p}(t)\}_{t=1}^T$ we have observed the consumer's demand responses $\{\mathbf{q}_\varepsilon(t) = \mathbf{d}(x(t), \mathbf{p}(t), \varepsilon)\}_{t=1}^T$, but not the underlying demand function. In this situation, using results of Afriat (1967), Varian (1982) derived bounds on the values that the counterfactual demand $\mathbf{q}_0$ can take thereby leading to a support set for $\mathbf{q}_0$. These bounds were based on the assumption that the consumer is rational and so satisfies the generalized axiom of preferences (GARP): For any given chain, $\mathbf{q}_\varepsilon(t_1), \mathbf{q}_\varepsilon(t_2), \ldots, \mathbf{q}_\varepsilon(t_N)$, for some of $N \geq 1$, satisfying $\mathbf{p}(t_k)' \mathbf{q}_\varepsilon(t_k) \geq \mathbf{p}(t_k)' \mathbf{q}_\varepsilon(t_{k+1})$, $k = 1, \ldots, N$, it must hold that $\mathbf{p}(t_N)' \mathbf{q}_\varepsilon(t_1) \geq \mathbf{p}(t_N)' \mathbf{q}_\varepsilon(t_N)$.

Alternatively, one can employ the Strong Axiom of Revealed Preference (SARP) developed by Samuelson (1938), Houthakker (1950) and Richter (1966). SARP is a strengthening of GARP and requires that for the same chain defined above, the strict inequality $\mathbf{p}(t_N)' \mathbf{q}_\varepsilon(t_1) > \mathbf{p}(t_N)' \mathbf{q}_\varepsilon(t_N)$ must hold. SARP characterizes finite sets of demand data generated by strictly convex and strictly monotone preferences, c.f. Matzkin and Richter (1991). GARP, on the other hand, is consistent with demand data generated by weakly convex preferences, which may generate non-unique demands. Mas-Colell (1978) established the following connection between GARP and SARP: Under a boundary condition, if the true preferences generating demand are strictly convex and monotone and the resulting demand functions are income Lipschitzian, the sequence of preferences constructed from a finite number of observations using GARP will have a unique limit as the data becomes dense. The limit are the unique preferences consistent with SARP, which satisfy the RP conditions with strict inequality. Since, however, the upper-contour sets generated by such RP conditions are open sets, while the upper contour sets implied by GARP are closed sets, we use the latter, as Varian, for constructing support sets.

The support set developed by Varian only uses the information contained in $\{\mathbf{p}(t), \mathbf{q}_\varepsilon(t)\}_{t=1}^T$. If we in fact have access to the sequence of Engel curves $\{\mathbf{d}(x, t, \varepsilon)\}_{t=1}^T$, Varian's bounds can be tightened as demonstrated by Blundell et al. (2008): Define intersection demands as $\bar{\mathbf{q}}_\varepsilon(t) = \mathbf{d}(\bar{x}_\varepsilon(t), t, \varepsilon)$, where $\bar{x}_\varepsilon(t)$ solves

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

A tighter support set that is consistent with observed expansion paths *and* utility maximization is then given as

$$\mathcal{S}_{\mathbf{p}_0, x_0, \varepsilon} = \left\{ \mathbf{q} \in \mathcal{B}_{\mathbf{p}_0, x_0} | \mathbf{p}(t)' \mathbf{q} \geq \mathbf{p}(t)' \bar{\mathbf{q}}_\varepsilon(t), 1 \leq t \leq T \right\} = \left\{ \mathbf{q} \in \mathcal{B}_{\mathbf{p}_0, x_0} | \bar{\mathbf{x}}_\varepsilon - \boldsymbol{P}\boldsymbol{q} \leq 0 \right\}, \qquad (2)$$

where

$$\mathcal{B}_{\mathbf{p}_0, x_0} = \left\{ \mathbf{q} \in \mathbb{R}_+^{L+1} | \mathbf{p}_0' \mathbf{q} = x_0 \right\}. \qquad (3)$$

Here, the second equality in Eq. (2) used that, by definition, $\mathbf{p}(t)' \bar{\mathbf{q}}_\varepsilon(t) = \bar{x}_\varepsilon(t)$ and that the T inequality constraints can be written on matrix form with

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

$$\bar{\mathbf{x}}_\varepsilon = (\bar{x}_\varepsilon(1), \ldots, \bar{x}_\varepsilon(T))' \in \mathbb{R}_+^T.$$

This 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.

From Blundell et al. (2008), we know that the support set $\mathcal{S}_{\mathbf{p}_0, x_0, \varepsilon}$ is non-empty and convex. Moreover, in the case of two goods, $\mathcal{S}_{\mathbf{p}_0, x_0, \varepsilon}$ defines bounds on $\mathbf{q}_0$ that are sharp given $\{\mathbf{d}(x, t, \varepsilon)\}_{t=1}^T$ and the RP inequalities since it makes maximal use of the heterogeneous expansion paths and the basic nonparametric choice theory.[2] In other words, there do not exist alternative bounds (derived from the same data) which are tighter. 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. 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

[1] 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] The bounds described are not necessarily sharp in a general economy with more than two goods since they do not utilize all constraints implied by rationality.

track this consumer across different price regimes. If a given consumer's ε is not time-varying, this set of demand functions will provide a full characterization of his behavior 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, $\varepsilon_1, \ldots, \varepsilon_T$, knowledge of $\mathbf{d}(x, t, \varepsilon)$, $t = 1, \ldots, T$, does not provide information of this particular consumer's behavior 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 $(x, \varepsilon) \mapsto \mathbf{d}(x, t, \varepsilon)$, $t = 1, \ldots, T$, 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, it is well-known that prices exhibit strong time series dependence (see Lewbel and Ng, 2005)) which will 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 et al. (2011) could be combined with the methods developed here to construct bounds for this more general case.

3. Estimation of demand bounds

A central objective of this paper is to develop inferential tools for the support set $\mathcal{S}_{\mathbf{p}_0,x_0,\varepsilon}$ when the demand functions are unknown to the researcher. In this section, we provide a general framework for estimation and inference of $\mathcal{S}_{\mathbf{p}_0,x_0,\varepsilon}$ given some initial estimators of the sequence of demand functions $\{\mathbf{d}(x, t, \varepsilon)\}_{t=1}^T$. In particular, under fairly general regularity conditions on the demand function estimators, $\{\hat{\mathbf{d}}(x, t, \varepsilon)\}_{t=1}^T$, we provide an asymptotic analysis of the corresponding support set estimator. The analysis will utilize the machinery developed in Chernozhukov et al. (2007b), henceforth CHT, who develop a general framework for the analysis of set estimators.

We take as given the availability of a sequence of demand function estimator, $\hat{\mathbf{d}}(x, t, \varepsilon)$ for $t = 1, \ldots, T$, whose $(L+1)$th component is restricted to satisfy Eq. (1). The general theory of this section takes no stand on the identification scheme that has been used in order to identify $\mathbf{d}(x, t, \varepsilon)$, and the precise nature of the estimator $\hat{\mathbf{d}}(x, t, \varepsilon)$. We will only require that the estimator is consistent and pointwise asymptotically normally distributed. In particular, we allow for both non-, semi- and fully parametric estimators, and so should cover most relevant scenarios. In the subsequent sections, we propose a specific nonparametric quantile estimator of $\mathbf{d}(x, t, \varepsilon)$ for the case of two goods ($L = 1$) that falls within the general setting of this section. But other estimators could be used.

To illustrate the generality of our results, we here provide some examples of potential demand estimators that are covered by the theory below: One could, for example, entertain the additive model of Blundell et al. (2008), where it is assumed that $\mathbf{d}_{1:L}(x, t, \varepsilon) = \bar{\mathbf{d}}_{1:L}(x, t) + \Sigma(x, t)\varepsilon$ for some functions $\bar{\mathbf{d}}_{1:L}(x, t) \in \mathbb{R}^L$ and $\Sigma(x, t) \in \mathbb{R}^{L\times L}$. Assuming that $\varepsilon \in \mathbb{R}^L$ satisfies $E[\varepsilon|x] = 0$, Blundell et al. (2008) estimate $\bar{\mathbf{d}}_{1:L}(x, t)$ and $\Sigma(x, t)$ using nonparametric kernel regression. Blundell et al. (2007) allow for x to be endogenous and develop nonparametric sieve IV estimators of the model under the assumption that $E[\varepsilon|w] = 0$ where w is a set of instruments. As another example, Imbens and Newey (2009) consider a class of triangular models where ε is allowed to enter non-additively but in a restricted manner and derive quantile-type estimators of $\mathbf{d}$. Finally, Matzkin (2003, 2008) derives identification results for other classes of non-additive models. One could then develop estimators of $\mathbf{d}$ based on her identification schemes which in turn could be used to compute estimated bounds for counterfactual demand.

We impose the following regularity conditions on the population demand and its estimator:

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

C.2 The estimators $\hat{\mathbf{d}}_{1:L}(x, 1, \varepsilon), \ldots, \hat{\mathbf{d}}_{1:L}(x, T, \varepsilon)$ are mutually independent over time, and there exists a sequence of nonsingular matrices $\Omega_n(x, t, \varepsilon) \in \mathbb{R}^{L\times L}$ such that
$$\sup_{x\in\mathcal{X}} \left\| \Omega_n^{1/2}(x, t, \varepsilon)\left(\hat{\mathbf{d}}_{1:L}(x, t, \varepsilon) - \mathbf{d}_{1:L}(x, t, \varepsilon)\right)\right\| = O_P\left(1/\sqrt{r_n}\right)$$
for some sequence r_n.

C.3 At the intersection income levels,
$$\sqrt{r_n}\Omega_n^{1/2}(\bar{x}(t), t, \varepsilon)\left(\hat{\mathbf{d}}_{1:L}(\bar{x}(t), t, \varepsilon) - \mathbf{d}_{1:L}(\bar{x}(t), t, \varepsilon)\right) \to^d N(0, V(\bar{x}(t), t, \varepsilon)),$$
for some positive definite matrix $V(x(t), t, \varepsilon) \in \mathbb{R}^{L\times L}$.

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

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

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

Condition (C.2) introduces two sequences, a matrix $\Omega_n(x, t, \varepsilon)$ and a scalar r_n. The condition states that once the demand estimator has been normalized by $\Omega_n^{1/2}(x, t, \varepsilon)$ it converges with rate $\sqrt{r_n}$. (C.3) is a further strengthening and states that the estimator when normalized by $\sqrt{r_n}\Omega_n^{1/2}(x, t, \varepsilon)$ converges towards a normal distribution. We have formulated (C.2)–(C.3) to cover as many potential estimators as possible. For parametric estimators, (C.2)–(C.3) will in general hold with $r_n = n$ and $\Omega_n(x, t) = I_L$. With nonparametric estimators, one may potentially choose $\Omega_n(x, t, \varepsilon)$ 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 developed in the subsequent section, to obtain rate results we will choose $\Omega_n(x, t, \varepsilon) = I_L$ and $r_n = O(k_n/\sqrt{n}) + O\left(k_n^{-m}\right)$

where k_n is number of sieve terms and m is the degree of smoothness of $x \mapsto \mathbf{d}(x, t, \varepsilon)$; to obtain distributional results, we will choose $\Omega_n(x, t, \varepsilon)$ as the inverse of the sequence of variance matrices of the estimator and $r_n = n$ in which case (C.3) holds under suitable restrictions on k_n.

Condition (C.4) will hold in great generality, while (C.5) requires that the observed prices have exhibited sufficient variation so we can distinguish between different demands. For the latter to hold, it is necessary that $T \geq L + 1$ and that at least $L + 1$ of these prices cannot be expressed as linear combinations of others.

Given an estimator $\hat{\mathbf{d}}(x, t, \varepsilon)$, it is natural to estimate the support set by simply substituting the estimated intersection incomes for the unknown ones. Defining the estimated intersection income levels $\hat{\mathbf{x}}_{\boldsymbol{\varepsilon}} = (\hat{x}_{\boldsymbol{\varepsilon}}(1), \ldots, \hat{x}_{\boldsymbol{\varepsilon}}(T))$ as the solutions to

$$\mathbf{p}_0' \hat{\mathbf{d}}(\hat{x}_{\boldsymbol{\varepsilon}}(t), t, \tau) = x_0, \quad t = 1, \ldots, T, \tag{4}$$

a natural support set estimator would appear to be $\hat{\mathcal{S}}_{\mathbf{p}_0, x_0, \tau} = \left\{\mathbf{q} \in \mathcal{B}_{\mathbf{p}_0, x_0} | \hat{\mathbf{x}}_{\boldsymbol{\tau}} - \mathbf{Pq} \leq 0\right\}$. 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. Conditions (C.2)–(C.4) allow the estimators of the demand functions to exhibit different convergence rates across time and income levels. As demonstrated in Lemma 4, the estimated intersection income levels, $\hat{\mathbf{x}}_{\boldsymbol{\varepsilon}} = \left\{\hat{x}_{\varepsilon}(t)\right\}_{t=1}^{T}$, inherit this property,

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

where I_T denotes the T-dimensional identity matrix, and $W_n(\varepsilon)$ is a diagonal matrix,

$$W_n(\varepsilon) = \operatorname{diag}\{w_n(1, \varepsilon), \ldots, w_n(T, \varepsilon)\},$$

with positive entries $w_n(t, \varepsilon)$ in the diagonal whose expressions can be found in Eq. (19). Due to the heterogeneous normalizations across $t = 1, \ldots, T$, as described by the weighting matrix $W_n(\varepsilon)$, 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, $Q_{n,\varepsilon}(\mathbf{q}) = \|\hat{W}_n^{1/2}(\varepsilon)\left[\hat{\mathbf{x}}_{\varepsilon} - \mathbf{Pq}\right]\|_+^2$, where $\|\mathbf{x}\|_+ = \|\max\{\mathbf{x}, 0\}\|$ for any vector $\mathbf{x}$, and $\hat{W}_n(\varepsilon) = \operatorname{diag}\left\{\hat{w}_n(1, \varepsilon), \ldots, \hat{w}_n(T, \varepsilon)\right\}$ is a consistent estimator of $W_n(\varepsilon)$. In comparison to the naive estimator suggested earlier, we now normalize $\hat{\mathbf{x}}_{\tau} - \mathbf{Pq}$ with $W_n^{1/2}$. In the case where the intersection incomes converge with same rate, this normalization would not be required since $\Omega_n^{1/2}(x, t, \varepsilon)$ can be chosen as the identity in this case.

Given that $\hat{\mathbf{x}}_{\varepsilon}$ is a consistent estimator of $\bar{\mathbf{x}}_{\varepsilon}$, it is straightforward to verify that $\sup_{\mathbf{q} \in \mathcal{B}_{\mathbf{p}_0, x_0}} |Q_n(\mathbf{q}|\varepsilon) - \bar{Q}_n(\mathbf{q}|\varepsilon)| \to^P 0$ (see the proof of Theorem 1 below), where $\bar{Q}_n(\mathbf{q}|\varepsilon) = \|W_n^{1/2}[\bar{\mathbf{x}}_{\varepsilon} - \mathbf{Pq}]\|_+^2$ is the non-stochastic version of $Q_n(\mathbf{q}|\varepsilon)$. Note that even though $\bar{Q}_n(\mathbf{q}|\varepsilon)$ is a sequence of functions (due to the presence of W_n), it still gives a precise characterization of the support set $\mathcal{S}_{\mathbf{p}_0, x_0, \varepsilon}$ for any given $n \geq 1$:

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

where the second equivalence follows from the fact that W_n is a diagonal matrix with positive elements. We then introduce the following set containing the demand levels that lie within a given contour level c of the sample objective function $Q_n(\mathbf{q}|\varepsilon)$,

$$\hat{\mathcal{S}}_{\mathbf{p}_0, x_0, \varepsilon}(c) = \left\{\mathbf{q} \in \mathcal{B}_{\mathbf{p}_0, x_0} | Q_n(\mathbf{q}|\varepsilon) \leq c\right\}. \tag{5}$$

In particular, our support set estimator is $\hat{\mathcal{S}}_{\mathbf{p}_0, x_0, \varepsilon}(0)$.

It is worth noting that the above formulation of the support set and its estimator in terms of $\bar{Q}_n(\mathbf{q}|\varepsilon)$ and $Q_n(\mathbf{q}|\varepsilon)$ is very close to the general formulation of set estimators defined through moment inequalities used in CHT. However, in their setting the limiting objective function, in our case $\bar{Q}_n(\mathbf{q}|\varepsilon)$, is not allowed to depend on n, so we cannot directly apply their results. Their proof strategy can fortunately be generalized without much additional work to apply to our case. 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. Finally, note that CHT introduce additional nuisance parameters in the form of a sequence of slack variables in the definition of their general estimator. Fortunately, our estimation problem satisfies the degeneracy property discussed in, for example, CHT, Sections 3.2 and 4.2, and so we can avoid using slack variables in the estimation. It is also worth noting that the estimation problem falls within the framework of Shi and Shum (2012) who consider plug-in estimators of identified set, except that our first-stage estimator, $\hat{\mathbf{x}}_{\varepsilon}$, is a nonparametric estimator; see also Kline and Tamer (2013).

We consider convergence of the estimated support set in terms of the Hausdorff distance,

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

$$\rho(y, \mathcal{A}) = \inf_{x \in \mathcal{A}} \|x - y\|,$$

for any two sets $\mathcal{A}_1, \mathcal{A}_2$.

Theorem 1. *Assume that* (C.1)–(C.2) *and* (C.5) *hold, and that* $\hat{w}_n(t) = w_n(t) + o_P(1)$. *Then* $d_{\mathrm{H}}(\hat{\mathcal{S}}_{\mathbf{p}_0, x_0}(0), \mathcal{S}_{\mathbf{p}_0, x_0}) = O_P(\sqrt{1/(r_n w_n^*)})$ *where* $w_n^* = \min_{t=1,\ldots,T} w_n(t)$.

If furthermore (C.3)–(C.4) *hold, then* $P(\mathcal{S}_{\mathbf{p}_0, x_0, \varepsilon} \subseteq \hat{\mathcal{S}}_{\mathbf{p}_0, x_0, \varepsilon}(\hat{q}_{1-\alpha})) \to 1 - \alpha$, *where* $\hat{q}_{1-\alpha}$ *is an estimator of* $(1-\alpha)$*th quantile of* $\mathcal{C}_{\mathbf{p}_0, x_0} := \sup_{\mathbf{q} \in \mathcal{S}_{\mathbf{p}_0, x_0}} \|Z + \xi(\mathbf{q})\|_+^2$. *Here,* $Z \sim N(0, I_T)$ *while* $\xi(\mathbf{q}) = (\xi_1(\mathbf{q}), \ldots, \xi_T(\mathbf{q}))'$ *is given by*

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

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 $C_{\mathbf{p}_0, x_0}$ which is the limiting distribution of $\sup_{\mathbf{q} \in \mathcal{S}_{\mathbf{p}_0, x_0, \varepsilon}} n\left\{Q_n(\mathbf{q}|\tau) - \bar{Q}_n(\mathbf{q}|\tau)\right\}$. 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 $C_{\mathbf{p}_0, x_0}$ 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 $\bar{x}_{\varepsilon}(t), t = 1, \ldots, T$, 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 second part of the theorem to construct a confidence set, quantiles of the random variable $C_{\mathbf{p}_0, x_0}$ defined in the theorem has to be computed. The distribution of $C_{\mathbf{p}_0, x_0}$ is non-standard and cannot be written in closed form. So its quantiles need to be evaluated using simulations (CHT) or resampling methods ((Bugni, 2010); CHT). Alternatively, given that our estimator falls within the framework of Shi and Shum (2012), a plug-in method can be employed: Let $CI_t(1-\alpha) \subseteq \mathbb{R}^L$ be an asymptotically valid confidence set for $\bar{\mathbf{x}}_{\boldsymbol{\varepsilon}}$ given the estimator $\hat{\mathbf{x}}_{\varepsilon}$,

$P\left(\bar{\mathbf{x}}_{\boldsymbol{\varepsilon}} \in CI_t\left(1-\alpha\right)\right) \to 1-\alpha$. This can be constructed using standard methods. One choice is to rely on an asymptotic approximation,

$$CI_t\left(1-\alpha\right) = \left\{\bar{\mathbf{x}}_{\boldsymbol{\varepsilon}} \in \mathbb{R}^T : \left(\bar{\mathbf{x}}_{\boldsymbol{\varepsilon}} - \hat{\mathbf{x}}_{\varepsilon}\right)' r_n^{-1} \hat{W}_{n,t}^{-1}\left(\varepsilon\right) \times \left(\bar{\mathbf{x}}_{\boldsymbol{\varepsilon}} - \hat{\mathbf{x}}_{\varepsilon}\right) \le \chi_T^2\left(1-\alpha\right)\right\},$$

for some estimator $\hat{W}_{n,t}\left(\varepsilon\right)$ of $W_{n,t}\left(\varepsilon\right)$. Another way of constructing $CI_t\left(1-\alpha\right)$ is by using standard bootstrap methods. It now easily follows that

$$\tilde{\mathcal{S}}_{\mathbf{p}_0,x_0,\varepsilon}\left(1-\alpha\right) := \left\{\mathbf{q} \in \mathbb{R}^L : \|\hat{W}_n^{1/2}\left(\varepsilon\right)\left[\mathbf{x}_{\varepsilon} - \mathbf{Pq}\right]\|_+^2 \le 0, \bar{\mathbf{x}}_{\boldsymbol{\varepsilon}} \in CI_t\left(1-\alpha\right)\right\}$$

is a consistent confidence set of $\mathcal{S}_{\mathbf{p}_0,x_0,\varepsilon}$, $P(\mathcal{S}_{\mathbf{p}_0,x_0,\varepsilon} \subseteq \tilde{\mathcal{S}}_{\mathbf{p}_0,x_0,\varepsilon}(1-\alpha)) \to 1-\alpha$.

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 Section 5 of CHT.

4. Quantile sieve estimation of demand

The previous section developed a general theory for support set estimation taking as input some estimator of the demand functions. In this section, we propose a specific estimator that allows for unobserved heterogeneity to enter in a fairly unrestricted manner. This estimator requires us to restrict our attention to the case of two goods ($L = 1$) and assume that ε is univariate or, more generally, to assume that $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_L)'$ and $d_l\left(x, \varepsilon\right) = d_l\left(x, \varepsilon_l\right)$ only depends on ε_l, $l = 1, \ldots, L$. Since the proposed estimation theory for $L > 1$ is a straightforward generalization of the case $L = 1$, we assume in the following that $L = 1$.

In order for $\mathbf{d}\left(x, t, \varepsilon\right) = \left(d_1\left(x, t, \varepsilon\right), d_2\left(x, t, \varepsilon\right)\right)'$ 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, $\varepsilon \sim U\left[0, 1\right]$. This in particular implies that the distribution of ε cannot change over time. We will furthermore assume ε to be independent of $x\left(t\right)$.[3]

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 $d_1\left(x, t, \tau\right)$, $\tau \in [0, 1]$, is identified as the τth quantile of $q_1\left(t\right)|x\left(t\right) = x$ (Matzkin, 2003; Imbens and Newey, 2009), $d_1\left(x, t, \tau\right) = F^{-1}_{q_1(t)|x(t)=x}\left(\tau\right)$, $\tau \in [0, 1]$. 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, 2009). However, this would not permit the application of the methods developed in this paper as demands relating to individual consumers are not directly identified. Importantly, Hoderlein and Stoye (2013) argue that in a two-good setting the assumption of ε being a scalar is vacuous for the computation of demand bounds.

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 $t = 1, \ldots, T$ and all $\tau \in [0, 1]$, the function $x \mapsto d_1\left(x, t, \tau\right)$ is situated in some known function space $\mathcal{D}_1$ which is equipped with some (pseudo-)norm $\|\cdot\|$.[4] 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 $n \to \infty$. Assuming that $\mathcal{D}_1$ is spanned by a set of known (basis) functions $\{B_k\}_{k\in\mathcal{K}}$ (see Chen (2007), Section 2.3 for examples), we focus on linear sieves,

$$\mathcal{D}_{n,1} = \left\{d_{n,1} : d_{n,1}\left(x, t, \tau\right) = \sum_{k\in\mathcal{K}_n} \pi_k\left(t, \tau\right) B_k\left(x\right), \pi\left(t, \tau\right) \in \mathbb{R}^{|\mathcal{K}_n|}\right\}, \quad (6)$$

for some sequence of (finite-dimensional) sets $\mathcal{K}_n \subseteq \mathcal{K}$. Finally, we define the space of vector functions, $\mathcal{D} = \{\mathbf{d} : d_1 \in \mathcal{D}_1, d_2 = (x - p_1 d_1)/p_2\}$, with the corresponding sieve space $\mathcal{D}_n$ obtained by replacing $\mathcal{D}_1$ by $\mathcal{D}_{n,1}$ in the definition of $\mathcal{D}$.

Given that $d_1\left(x, t, \tau\right)$ is identified as a conditional quantile for any given value of x, we may employ standard quantile regression techniques to obtain the estimator: Let $\rho_\tau\left(z\right) = \left(\tau - \mathbb{I}\left\{z < 0\right\}\right) z$, $\tau \in [0, 1]$, be the standard check function used in quantile estimation (see Koenker and Bassett, 1978). We then propose the following estimator:

$$\hat{\mathbf{d}}\left(\cdot, t, \tau\right) = \arg\min_{\mathbf{d}_n \in \mathcal{D}_n} \frac{1}{n} \sum_{i=1}^{n} \rho_\tau\left(q_{1,i}\left(t\right) - d_{n,1}\left(x_i\left(t\right), t, \tau\right)\right), \quad t = 1, \ldots, T, \quad (7)$$

In practice, given that the sieve is linear, this estimator takes the form of a linear quantile regression.

The above estimator does not utilize that, if indeed the consumers are rational, the demand function has to satisfy RP restrictions. Since the unconstrained estimator is consistent, it will asymptotically satisfy the RP restrictions. However, in finite samples, there is no reason why the estimator should satisfy these restrictions and so imposing these restrictions is expected to reduce estimation errors. Consider a given consumer characterized by $\tau \in [0, 1]$, and construct the following particular income expansion path $\left\{\tilde{x}_\tau\left(t\right)\right\}$ recursively by

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

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

$$\tilde{x}_\tau\left(t\right) \le \mathbf{p}\left(t\right)' \mathbf{d}\left(\tilde{x}\left(s\right), s, \tau\right), \quad s < t, \ t = 1, \ldots, T. \quad (8)$$

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

[3] The independence assumption can be relaxed as discussed in Section 7.

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

A RP-restricted sieve estimator is now easily obtained in principle: First observe that the unrestricted estimator of $\{\mathbf{d}(\cdot,t,\tau)\}_{t=1}^{T}$ developed above can be thought of as the solution to a joint estimation problem over the function space $\mathcal{D}_n^T = \otimes_{t=1}^{T}\mathcal{D}_n$ with $\mathcal{D}_n$ defined above. Since there are no restrictions across the T time periods, the joint estimation problem can be decomposed into T separate estimation problems as given in Eq. (7). In order to impose the RP restrictions, we define the constrained function set for the T demand functions as

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

and similarly the constrained sieve as $\mathcal{D}_{C,n}^T := \mathcal{D}_n^T \cap \{\mathbf{d}_n(\cdot,\cdot,\tau)$ satisfies RP$\}$. The constrained estimator is then defined as

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

Note that since the RP inequalities impose restrictions across time ($t = 1,\dots,T$), the above estimation problem can no longer be split into individual subproblems as in the unconstrained case. The above estimator, however, 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). So a computationally attractive alternative is to update ("rearrange") the initial unconstrained estimator using least-squares:

$$\{\tilde{\mathbf{d}}_C(\cdot,t,\tau)\}_{t=1}^{T} = \arg\min_{\mathbf{d}_n(\cdot,\cdot)\in\mathcal{D}_{C,n}^T} \frac{1}{n} \times \sum_{t=1}^{T}\sum_{i=1}^{n}\left(\hat{d}_1(t,x_i(t)) - d_{n,1}(t,x_i(t))\right)^2. \quad (11)$$

In particular, $\{\tilde{\mathbf{d}}_C(\cdot,t,\tau)\}_{t=1}^{T}$ can be computed using standard numerical optimization algorithms. At the same time, under conditions discussed below, the two estimators, $\hat{\mathbf{d}}_C$ and $\tilde{\mathbf{d}}_C$, are asymptotically equivalent, and so we will in the following use $\hat{\mathbf{d}}_C$ to denote either of the two.

As explained in Section 2.3, by the results in Mas-Colell (1978, Theorem 4), if we are willing to assume that the demand function satisfies a boundary condition and is income-Lipschitzian, the RP constraints will be satisfied with strict inequality in the limit. Hence, as argued below, the shape constraints will not be binding and so the asymptotic properties of the constrained estimator will be the same as that for the unconstrained. On the other hand, if the constraints are binding, the constrained estimator $\hat{\mathbf{d}}_C$ is expected to have a non-standard asymptotic distribution; see, for example, Wright (1981), Andrews (1999), Anevski and Hössjer (2006) who give results for inequality-constrained parametric and nonparametric problems respectively. Ideally, we would like to analyze the properties of $\hat{\mathbf{d}}_C$ also in the case of binding constraints using similar techniques, but this proves technically very demanding. This is due to the fact that the RP constraints are global and cannot, as in the case of monotonicity or positivity constraints, be formulated as simple, pointwise inequality constraints.

Instead, in order to be robust towards binding constraints, we adapt an idea used elsewhere in the literature on nonparametric estimation under shape constraints where we remove the binding constraints through the introduction of a certain level of slack, see e.g. Birke and Dette (2007), Mammen (1991), Mukerjee (1988). This is done here by introducing the following generalized version of RP: We say that "$\mathbf{d}$ satisfies RP(e)" for some constant $e \geq 0$ if for any income expansion path,

$$e\tilde{x}_\tau(t) \leq \mathbf{p}(t)'\mathbf{d}\left(\tilde{x}_\tau(s),s,\tau\right), \quad s<t, t=2,\dots,T.$$

The definition of RP(e) is akin to Afriat (1973) who suggests a similar modification of (GA)RP to allow for waste ("partial efficiency"). We can interpret e as Afriat's "efficiency parameter": With $e = 1$, no waste is allowed for; as e decreases, the more waste we allow for; with $e = 0$, 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 $\mathcal{D}_C^T(e) = \mathcal{D}^T \cap \{\mathbf{d}(\cdot,\cdot,\tau)$ satisfies RP$(e)\}$ and $\mathcal{D}_{C,n}^T(e) = \mathcal{D}_n^T \cap \{\mathbf{d}_n(\cdot,\cdot,\tau)$ satisfies RP$(e)\}$. We note that the constrained function space $\mathcal{D}_C^T$ as defined in Eq. (9) satisfies $\mathcal{D}_C^T = \mathcal{D}_C^T(1)$. Moreover, it should be clear that $\mathcal{D}_C^T(\bar{e}) \subseteq \mathcal{D}_C^T(e)$ for $0 \leq e \leq \bar{e} \leq 1$ since RP(e) imposes weaker restrictions on the demand functions compared to RP($\bar{e}$). One could in principle make the waste parameter $e = e_n \to 1$ at a suitable rate. This would however complicate the analysis and so we treat e as a fixed, user-chosen constant in the following.

We now re-define our RP constrained estimators to solve the same optimization problem as before, but now the optimization takes place over $\mathcal{D}_{C,n}(e)$ for some given choice of e. Let $\hat{\mathbf{d}}_C^e$ denote this estimator, and note that $\hat{\mathbf{d}}_C^1 = \hat{\mathbf{d}}_C$, where $\hat{\mathbf{d}}_C$ is given in Eq. (10). Suppose that $\{\mathbf{d}(\cdot,t,\tau)\}_{t=1}^{T} \in \mathcal{D}_C^T(\bar{e})$ for some $\bar{e} \leq 1$; this implies that the unconstrained estimator satisfies $\{\hat{\mathbf{d}}(\cdot,t,\tau)\}_{t=1}^{T} \in \mathcal{D}_{C,n}^T(e)$ w.p.a.1, for any $e < \bar{e}$. Since $\hat{\mathbf{d}}_C^e$ is a constrained version of $\hat{\mathbf{d}}$, this implies that $\hat{\mathbf{d}}_C^e = \hat{\mathbf{d}}$ w.p.a.1. We may therefore conclude that $\hat{\mathbf{d}}_C^e$ is asymptotically equivalent $\hat{\mathbf{d}}$, and all the asymptotic properties of $\hat{\mathbf{d}}$ are inherited by $\hat{\mathbf{d}}_C^e$.

Suppose the results of Mas-Colell (1978, Theorem 4) apply, such that no constraints are binding for the population demand functions. If the support of $x(t)$ is compact, it then holds that $\{\mathbf{d}(\cdot,t,\tau)\}_{t=1}^{T} \in \mathcal{D}_C^T(\bar{e})$ for some $\bar{e} > 1$. In this case, the above arguments go through with $e = 1 < \bar{e}$, and so the constrained estimator $\hat{\mathbf{d}}_C = \hat{\mathbf{d}}_C^1$ will be asymptotically equivalent to $\hat{\mathbf{d}}$.

Additional constraints could be imposed on the estimator to further improve its finite-sample performance. For example, we could use the quantile-rearrangement method of Chernozhukov et al. (2009) to ensure that no crossing is taking place across the quantile range.

To analyze the two quantile sieve estimators, 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, $\|\mathbf{d}(\cdot,t,\tau)\|_2 = \sqrt{E\left[\|\mathbf{d}(x,t,\tau)\|^2\right]}$ and $\|\mathbf{d}(\cdot,t,\tau)\|_\infty = \sup_{x\in\mathcal{X}}\|\mathbf{d}(x,t,\tau)\|$. The following assumptions are then 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 $\varepsilon \sim U[0,1]$, $1 \leq t \leq T$.
A.2 The demand function $d_1(x,t,\varepsilon)$ is strictly increasing in ε, $1 \leq t \leq T$.
A.3 The function $d_1(\cdot,t,\tau) \in \mathcal{D}_1$, where $\mathcal{D}_1 = \mathcal{W}_2^m([a,b])$ and $\mathcal{W}_2^m([a,b])$ is the Sobolev space of all functions on $[a,b]$ with L_2-integrable derivatives up to order $m \geq 0$, $1 \leq t \leq T$.

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 assumptions and proofs so we maintain (A.1) for simplicity (see e.g. Chen, Blundell and Kristensen, 2007 for results with unbounded support). The independence assumption rules out endogenous income; in Section 7, we explain 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.

Finally, to state the asymptotic distribution, we need some additional notation: Define the sequence of covariance matrices

$$V_n(\tau) = \tau(1-\tau) H_n^{-1}(t,\tau)\,\Omega_n(t,\tau)\,H_n^{-1}(t,\tau), \tag{12}$$

$$\Omega_n(t,\tau) = E\left[\underline{B}_{k_n}(x(t))\,\underline{B}_{k_n}(x(t))'\right],$$
$$H_n(t,\tau) = E\left[f(0|t,x(t),\tau)\,\underline{B}_{k_n}(x(t))\,\underline{B}_{k_n}(x(t))'\right]. \tag{13}$$

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

$$f(e|t,x,\tau) = f_{q_1(t)|x(t)}(\kappa + d_1(x,t,\tau)|x), \tag{14}$$

where $f_{q_1(t)|x(t)}(\cdot|x)$ is the conditional density of $q_1(t)$ given $x(t)$.

Theorem 2. *Assume that* (A.1)–(A.3) *hold. Then the unconstrained estimator satisfies the following: for any* $1 \le t \le T$ *and* $\tau \in [0,1]$*:*

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

while

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

If, in addition, 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*

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

where $I_T \in \mathbb{R}^{T\times T}$ *denotes the identity matrix, and* $\Sigma_n(x,\tau) = \text{diag}\left\{\{\Sigma_n(x(1),1,\tau)\}_{t=1}^T\right\} \in \mathbb{R}^{T\times T}$ *with* $\Sigma_n(x(t),t,\tau) = \underline{B}_{k_n}(x(t))'\,V_n(t,\tau)\,\underline{B}_{k_n}(x(t))$*.*

Suppose that $\mathbf{d} \in \mathcal{D}_C^T(\bar{e})$ *for some* $\bar{e} > 0$*. Then the constrained estimator* $\hat{\mathbf{d}}_C^e(\cdot,t,\tau)$ *with* $0 \le e < \bar{e}$ *has the same asymptotic properties as* $\hat{\mathbf{d}}$*. In particular, if the results of Mas-Colell (1978) apply, then* $\hat{\mathbf{d}}_C = \hat{\mathbf{d}}_C^1$ *is asymptotically equivalent to* $\hat{\mathbf{d}}$*.*

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 et al. (2010, Lemma 2.1. and Remark 2.1). The asymptotic independence of the estimators across time is due to the fact that a new sample of consumers are drawn at each time period. 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 et al. (2011), and so could potentially be used to examine uniform convergence of the resulting bounds.

A consistent estimator of the covariance matrix $\Sigma_n(x,\tau)$ can be obtained by replacing expectations with sample averages in the definition of $\Omega_n(t,\tau)$ and $H_n(t,\tau)$ in Eq. (13) and $f(0|t,x,\tau)$ by $\hat{f}_{q_1(t)|x(t)}(\hat{d}_1(x,t,\tau)|x)$ with $\hat{f}_{q_1(t)|x(t)}(q|x)$ being, for example, a kernel estimator of the conditional density.

The first-order asymptotic properties of the constrained estimator are identical to those of the unconstrained one. This is similar to other results in the literature on constrained nonparametric estimation. For example, 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. We conjecture that the constrained estimator is higher-order efficient relative to the unconstrained one, so that in finite samples it provides a more precise estimate. Simulation results reported in the working paper version of this paper (Blundell et al., 2011) support this conjecture. However, a formal proof of it seems very daunting,[5] and so the verification of this is left for future research.

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.

Given the results in Theorem 2, we can now verify the general conditions (C.1)–(C.4) to obtain the following results for the estimated bounds based on the proposed sieve quantile estimator:

Theorem 3. *Suppose that* (A.1)–(A.3) *hold,* $x \mapsto d_1(x,t,\varepsilon)$ *is strictly increasing, and* $\|\hat{W}_n - W_n\| \to^P 0$*. Then* (C.1)–(C.4) *hold for the quantile estimators. In particular,*

$$d_H(\hat{\mathcal{S}}_{\mathbf{p}_0,x_0,\tau}(0), \mathcal{S}_{\mathbf{p}_0,x_0,\tau}) = O_P(k_n\sqrt{n}) + O_P\left(k_n^{-m}\right).$$

If furthermore, the eigenvalues of $E\left[\underline{B}_{k_n}(x(t))\,\underline{B}_{k_n}(x(t))'\right]$, $t = 1,\ldots,T$*, 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* $P(\mathcal{S}_{\mathbf{p}_0,x_0,\tau} \subseteq \hat{\mathcal{S}}_{\mathbf{p}_0,x_0,\tau}(\hat{q}_{1-\alpha})) \to 1-\alpha$ *where* $\hat{q}_{1-\alpha}$ *is an estimator of the* $(1-\alpha)$*th quantile of* $C_{\mathbf{p}_0,x_0}$ *defined in Theorem 1.*

Note that instead of using $\hat{\mathcal{S}}_{\mathbf{p}_0,x_0,\tau}(\hat{q}_{1-\alpha})$ as a confidence set, one could alternatively use the plug-in version discussed after Theorem 1.

5. Computation of estimators

The implementation of the estimated demand functions and bounds may be computationally challenging. We here propose relative simple numerical algorithms for their computation that are not too demanding in the sense that we were able to implement them in Matlab on a standard desktop.

In the computation of the constrained demand function estimator, we have to check whether the RP constraints are satisfied for a given candidate estimator. However, it is not numerically feasible to check that a given candidate 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) $M \ge 1$ income "termination" values, $\tilde{x}_i(T)$, $i = 1,\ldots,M$. The latter will be used to generated income paths. By choosing M sufficiently large, we hope to cover most of the possible income paths. For a given member of the unconstrained 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 $\{\tilde{x}_i(t)\}$, $\tilde{x}_i(t) = \mathbf{p}(t)'\,\mathbf{d}_n(\tilde{x}_i(t+1), t+1)$ for $i = 1,\ldots,M$. For each of these paths, which implicitly depends on the sieve coefficients to be estimated, $\pi = \{\pi(t)\}_{t=1}^T$, we check whether Eq. (8) holds. By defining

$$a_i(s,t,\pi) = \left\{\frac{p_2(t)}{p_2(s)}p_1(s) - p_1(t)\right\}\underline{B}_{k_n}(\tilde{x}_i(s))' \in \mathbb{R}^{k_n}, \tag{15}$$

$$b_i(s,t,\pi) = \frac{p_2(t)}{p_2(s)}\tilde{x}_i(s) - \tilde{x}_i(t) \in \mathbb{R},$$

[5] For example, Chernozhukov et al. (2009) have to spend considerable effort to show such results for a much simpler constrained estimator.

for $s < t$, the RP constraints can be written more conveniently in matrix form as $A_M(\pi)\pi \le b_M(\pi)$, where,

$$A_M(\pi) = \left[O_{1\times(s-1)k_n}, a_i(s,t,\pi), O_{1\times(T-s)k_n},\right]_{i=1,\ldots,M,s<t},$$
$$b_M(\pi) = [b_i(s,t,\pi)]_{i=1,\ldots,M,s<t},$$

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 in Eq. (11) should then be well-approximated by

$$\hat{\pi}_C = \arg\min_{\pi} \frac{1}{n}\sum_{t=1}^{T}\sum_{i=1}^{n}\left(\hat{d}_1(t, x_i(t)) - \pi(t)'\underline{B}_{k_n}(x_i(t))\right)^2$$
$$\text{s.t.}\ A_M(\pi)\pi \le b_M(\pi). \quad (16)$$

For a fixed grid size M, the above implementation will only deliver an approximate version of the Eq. (11) with the approximate estimator being less precise relative to the exact, but infeasible one. However, as the grid size increases, the approximation errors will vanish. In the empirical analysis, we tried out different grid sizes and found little sensitivity of the estimator to the choice of this for $M = 50 - 100$. In the reported results, we choose the grid as the empirical 1–99 percentiles of x so that $M = 99$.

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.

Once an estimator of $\mathbf{d}$, either constrained or unconstrained, has been obtained, the computation of bounds proceeds in two steps: For a given set of prices, $\mathbf{p}_0$, and income level, x_0, we first solve Eq. (4) w.r.t. $\hat{x}_\varepsilon(t)$ using a numerical equation solver for $t = 1, \ldots, T$. Next, approximate estimators of the bounds are found as solution to a linear programming problems: Given some cut-off level $\hat{c}_n \ge 0$, 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:

$$\tilde{q}_{\text{up},1} = \arg\max_{\mathbf{q}\in\mathbb{R}^2} q_1 \ \text{s.t.}\ \hat{\mathbf{A}}\mathbf{q} \le \hat{\mathbf{b}} \quad\text{and}\quad p_0\mathbf{q} = x_0,$$
$$\tilde{q}_{\text{low},1} = \arg\min_{\mathbf{q}\in\mathbb{R}^2} q_1 \ \text{s.t.}\ \hat{\mathbf{A}}\mathbf{q} \le \hat{\mathbf{b}} \quad\text{and}\quad p_0\mathbf{q} = x_0.$$

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:

$$\tilde{q}_{\text{up},1} = \arg\max_{\mathbf{q}\in\mathcal{B}_{\mathbf{p}_0,x_0}} q_1 \ \text{s.t.}\ nQ_n(\mathbf{q}) \le \hat{c}_n,$$
$$\tilde{q}_{\text{low},1} = \arg\min_{\mathbf{q}\in\mathcal{B}_{\mathbf{p}_0,x_0}} q_1 \ \text{s.t.}\ nQ_n(\mathbf{q}) \le \hat{c}_n,$$

where the optimization algorithm is started at $\tilde{q}_{\text{up},1}$ and $\tilde{q}_{\text{low},1}$ respectively. Note that the theory allows us to set $\hat{c}_n = 0$, but in practice we found that the linear programming solver was unable to find a solution with $\hat{c}_n = 0$. To resolve this issue, we then gradually increased $\hat{c}_n$ until the solver found a solution. In general, this procedure worked well, but some numerical irregularities were found as discussed below in the presentation of the empirical results.

Confidence regions for these demand bounds can be obtained by choosing the cut-off level $\hat{c}_n$ as $\hat{c}_n = \hat{q}_{1-\alpha}$, 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 1. This can be computed by simulations. We first rewrite $C_{\mathbf{p}_0,x_0,\tau}$: Letting $\bar{T}_b = \max_{\mathbf{q}\in\mathcal{S}_{\mathbf{p}_0,x_0,\tau}}\sum_{t=1}^{T}\xi_\tau(t,\mathbf{q})$, $\xi_t(t,\mathbf{q}) := \mathbb{I}\left\{\bar{x}_\tau(t) = \mathbf{p}(t)'\mathbf{q}\right\}$, 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 $\hat{T}_b = \max_{\mathbf{q}\in\hat{\mathcal{S}}_{\mathbf{p}_0,x_0,\tau}}\sum_{t=1}^{T}\hat{\xi}_\tau(t,\mathbf{q})$, $\hat{\xi}_\tau(t,\mathbf{q}) = \mathbb{I}\left\{\hat{x}_\tau(t) \ge \mathbf{p}(t)'\mathbf{q} - a_n\right\}$ with $a_n \propto \sqrt{\log(n)/n}$, we propose to compute approximate quantiles by simulating from $\hat{C}_{\mathbf{p}_0,x_0,\tau} = \sum_{t=1}^{\hat{T}_b}\max\{Z(t),0\}^2$.

In the working paper version (Blundell et al., 2011), we provided a simulation study showing that at observed prices, our demand estimators perform 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.

6. Empirical application

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) which is a repeated cross-section survey consisting of around 7000 British households in each year containing expenditure data and prices. The same data set was used in Blundell et al. (2008) to construct demand bounds under the assumption that the demand function was additive, $d_l(x,\varepsilon_l) = d_l(x) + \varepsilon_l$. This restriction implies that unobserved heterogeneity only affects demands in terms of location shifts. Here, we allow demand to be non-additive in x and ε and thereby for richer interaction between the two, thereby providing new insights into price responsiveness of demand across the distribution of unobserved tastes.

6.1. Data

We select as data the subsample of couples with children who own a car providing us with between 1421 and 1906 observations per year. In our application, we focus on FES data for the eight year period 1983–1990. We use total spending on nondurables to define our total expenditure, x, defined as total expenditure on non-durables and services. 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 Fig. 5.1. Similar distributions are found for the other years in the data set.

6.2. The sieve estimates of quantile expansion paths

In the estimation, we implement the sieve estimator along the lines described in Section 5. 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 the left panel of Fig. 5.2 together with 95% confidence intervals that have been computed using the asymptotically normal approximation of Theorem 2.

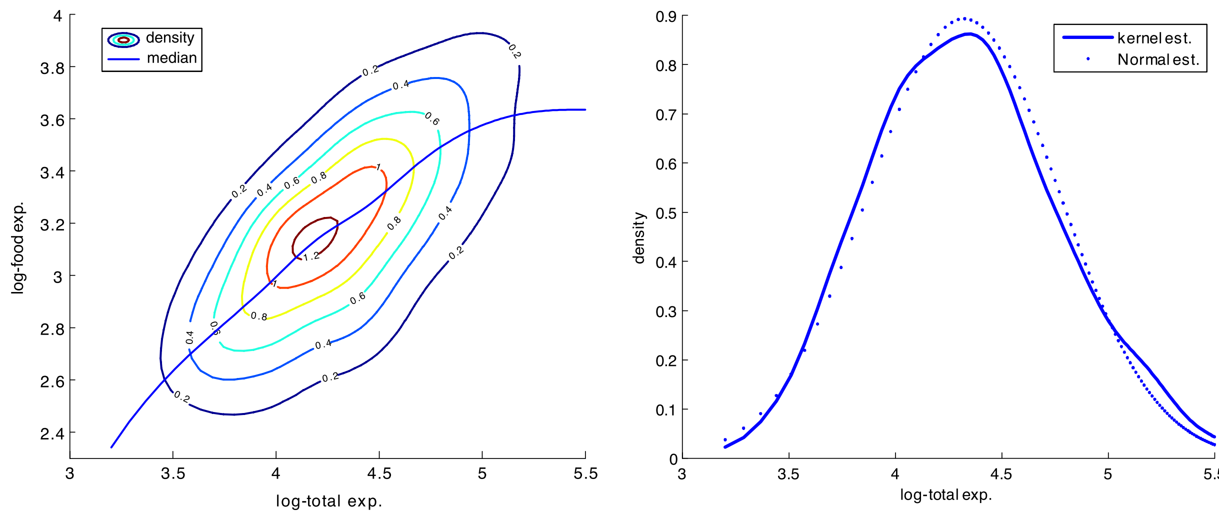

Fig. 5.1. Engel curve distribution and distribution of total expenditure, 1985.

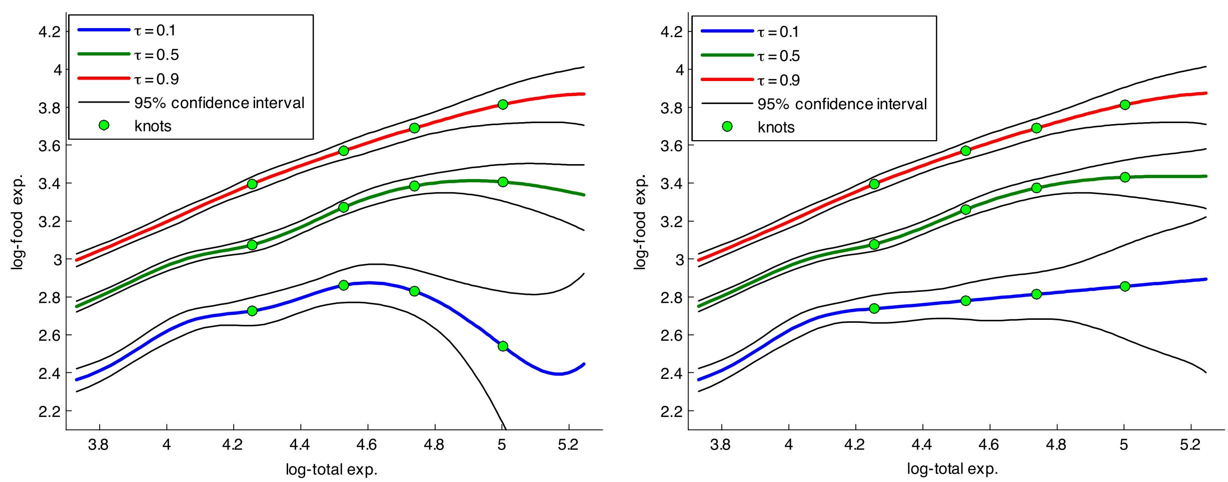

Fig. 5.2. Unconstrained and constrained demand function estimates, 1983.

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 reflecting 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 with $e = 0.99$ thereby allowing for 1% waste. The constrained quantile Engel curve estimates for $t = 1983$ can be found in right panel of Fig. 5.2. 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 of the unconstrained estimator are removed. However, the overall shapes remain quite similar. The 95% confidence intervals are also here computed using the asymptotically normal approximation of Theorem 2 assuming that the RP constraints are non-binding. If in fact the constraints are binding, the confidence intervals will in general be distorted and so should be interpreted with care.

6.3. *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 Fig. 5.3, 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 are relatively more narrow within the range of observed prices ($0.94 \leq p_{1,t} \leq 1.01$) in our sample, while for prices far away from observed prices the bounds widen and become less informative. The kinks in the bounds appear at observed prices since these impose additional restrictions on the demand. Note, however, that the plotted bounds are smooth and do not take the form of step functions. This appears to be a side effect of our implementation of the bound estimators as presented in Section 6 where a positive cut-off point in the computation of the bounds, $\hat{c}_n > 0$, functions as a type of smoothing.

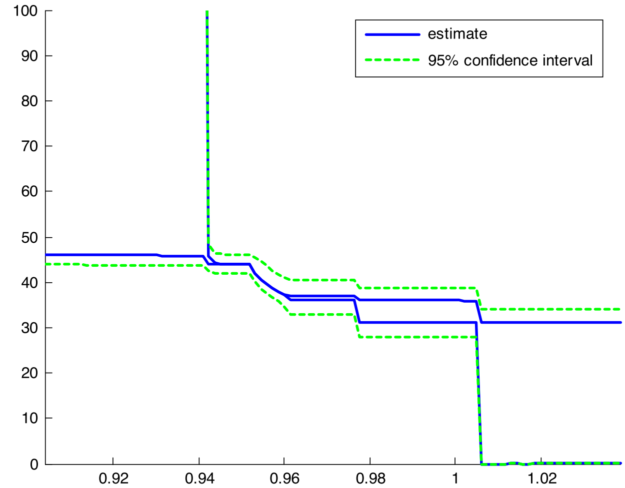

Fig. 5.3. Estimated demand bounds and 95% confidence sets at median income, $\tau = 0.5, T = 8$.

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. The top-left panel of Fig. 5.4 shows the estimated confidence sets for the bounds on the quantile demand function at the median income for the 10th percentile ($\tau = 0.1$) of the unobserved taste distribution. Notice that where the relative prices are quite dense the bounds are correspondingly narrow. The top-right panel of Fig. 5.4 contrasts this for a consumer at the 50% ($\tau = 0.5$) 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. The lower-left panel of Fig. 5.4 considers a consumer with an even stronger taste for food—at the 90th percentile ($\tau = 0.9$) of the taste distribution. Demand shifts further up at all points. The bounds remain quite narrow where the relative prices are dense. At a few price levels and chosen values of τ, the demand bounds grow wider as we increase T; for example, compare the demand bounds for $T = 6$ and 8 in the top-left panel when price of food is between 1 and 1.02. This was caused by the aforementioned problems with a finding a solution to the linear programming problem for larger values of T which lead to us having to increase the cut-off point $\hat{c}_n$.

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 Fig. 5.5, which presents 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 bounds are shown in the left panel

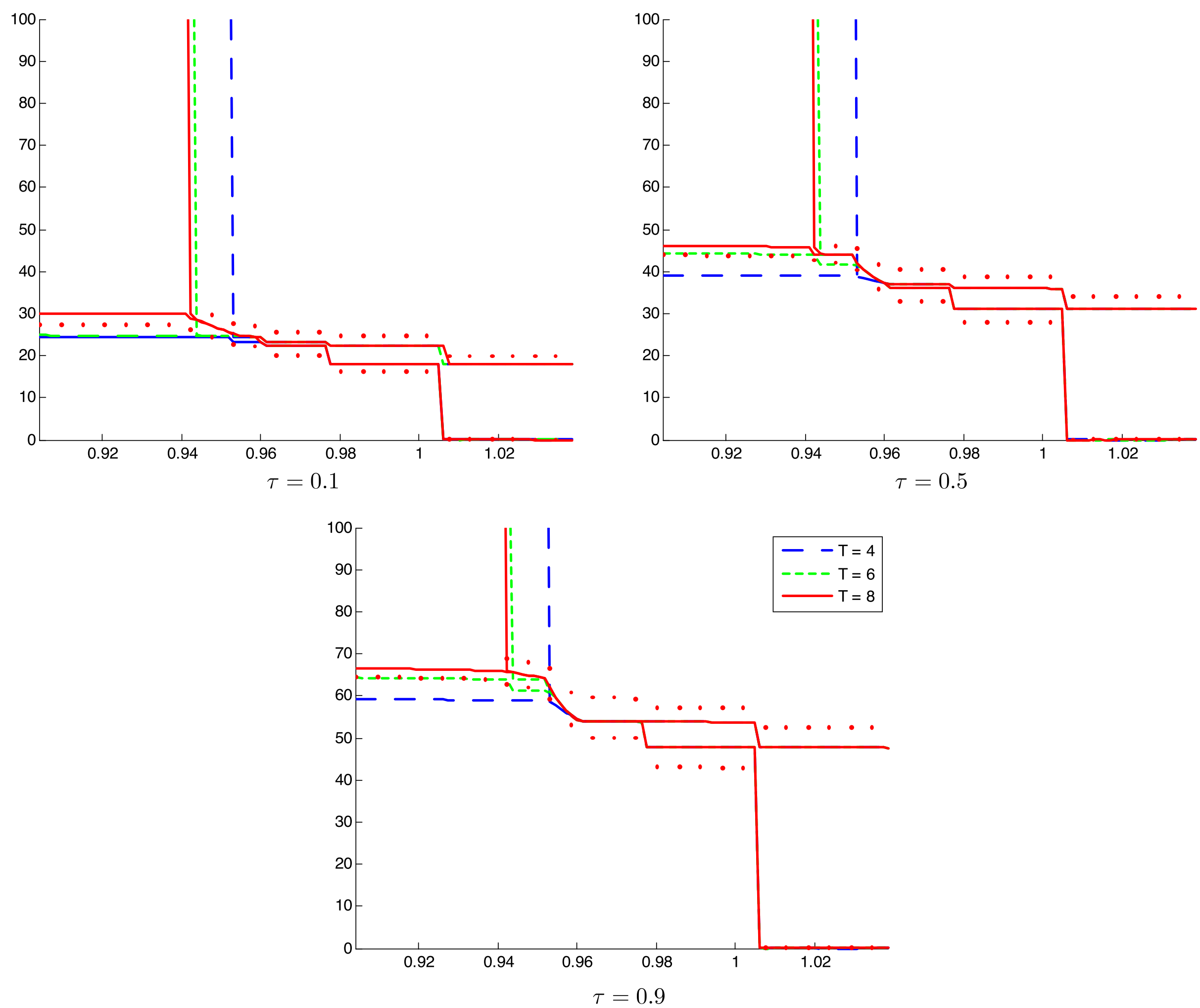

Fig. 5.4. Bounds at median income for $\tau = 0.1$, 0.5 and 0.9.

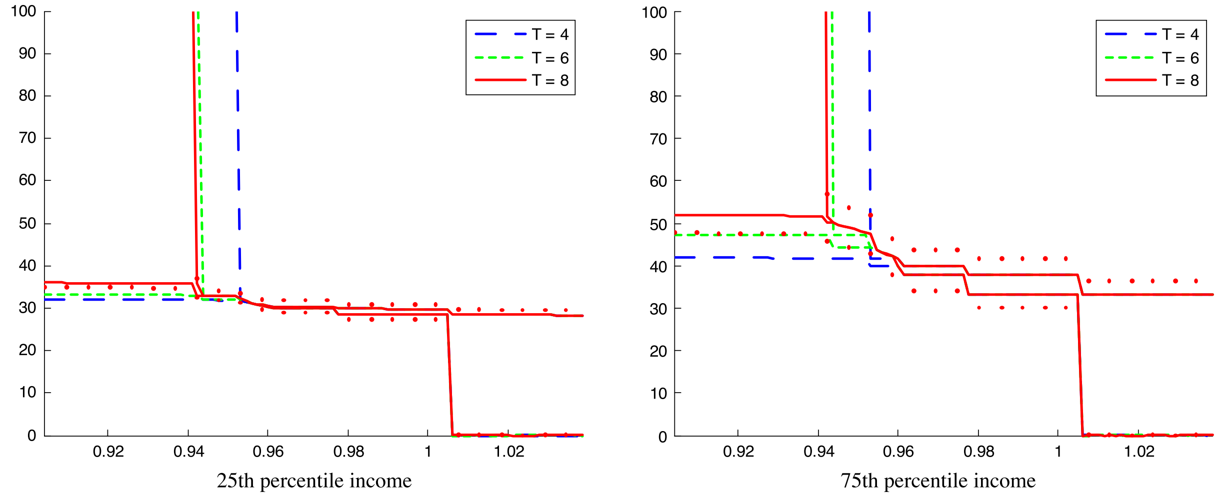

Fig. 5.5. Bounds at $\tau = 0.5$ and 25th and 75th percentile income.

of Fig. 5.5. 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 the right panel of Fig. 5.5. Comparing the two plots, we see that the overall shape remains the same, but that demand bounds are compressed as income levels are decreased.

7. 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 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. In recent years, nonparametric estimation methods for additive regression models under endogeneity have been proposed; see, for example, Newey et al. (1999), Newey and Powell (2003) and Hall and Horowitz (2005). These have been applied in the empirical analysis of Engel curves with additive errors (Blundell et al., 1998, 2007). These nonparametric techniques have recently been generalized to the case of quantile models; see, e.g., Chernozhukov et al. (2007a), Imbens and Newey (2009) and Chen and Pouzo (2012). With the assumptions and results of either of these three papers replacing our assumptions (A.1)–(A.3) and our Theorem 2, the general results for estimated bounds as given in Theorem 1 will still go through.

The specific estimators developed in the paper assumes that the demand function for good l takes the form $d_l\,(x, \varepsilon_l)$ where ε_l is a scalar. When the number of goods $L > 2$, this is a rather restrictive assumption which ideally should be weakened to allow for multiple unobserved components entering each demand function. Building on Matzkin (2008), Blundell et al. (2013b) provide new identification and estimation methods for consumer demand models that can be used to estimate nonparametric systems of demand functions where each function is nonadditive on a vector of unobservable random terms.

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, amongst others, Self and Liang (1987) and Wolak (1991). An attractive alternative approach is to work directly with the revealed preference inequalities, avoiding estimating quantile demands and relaxing the restrictions on unobserved heterogeneity see, in particular, the recent work of Hoderlein and Stoye (forthcoming, 2013) and Kitamura and Stoye (2012). We leave the extension of these results to nonparametric quantile estimation for future research.

Acknowledgments

A previous version of this paper was circulated under the title *Stochastic Demand and Revealed Preference*. We would like to thank the editor, associate editor and three referees for their comments. We are grateful to Ian Crawford and Roger Koenker for sharing Matlab code with us and helping with the implementation. We also wish to thank Xiaohong Chen, Yong Hyeon Yang and participants at the Malinvaud Seminar, the ES World Congress in Shanghai, the Cowles–Cemmap workshop in microeconometrics, and seminars at Brown University, UC San Diego, Universitat Autònoma de Barcelona, University of Rochester, Tilburg University, and Toulouse School of Economics for helpful comments and suggestions. The research is part of the program of research of the

ESRC Centre for the Microeconomic Analysis of Public Policy at IFS. Funding from the ESRC, grant number R000239865 is gratefully acknowledged. Kristensen gratefully acknowledges financial support from the European Research Council through a starting grant (No 312474), and the Danish Research Council through its funding of CREATES. Matzkin acknowledges support from NSF (grants BCS-0852261, SES-1062090, and SES-0833058). Material from the FES made available by the ONS through the ESRC Data Archive has been used by permission of the controller of HMSO. Neither the ONS nor the ESRC Data Archive bear responsibility for the analysis or the interpretation of the data reported here. The usual disclaimer applies.

Appendix A. Proofs

Proof of Theorem 1. Since ε is kept fixed throughout, we suppress any dependence on this in the following. We follow the same proof strategy as in CHT and first verify that slightly modified versions of their Conditions C.1–C.3 are satisfied with our definitions of $\bar{Q}_n(\mathbf{q})$ and $Q_n(\mathbf{q})$. For convenience, define $m_n(\mathbf{q}|\varepsilon) := \hat{\mathbf{x}} - \mathbf{Pq}$ and $\bar{m}(\mathbf{q}) := \bar{\mathbf{x}} - \mathbf{Pq}$. We then have uniformly in $\mathbf{q} \in \mathcal{B}_{\mathbf{p}_0,x_0}$,

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

since $\hat{W}_n^{1/2}\left\{\hat{\mathbf{x}} - \bar{\mathbf{x}}\right\} = O_P\left(1/\sqrt{r_n}\right)$ by Lemma 4. Moreover,

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

where $\left\| \sqrt{r_n}\hat{W}_n^{1/2}\bar{m}(\mathbf{q}) \right\|_+^2 \geq r_n w_n^* C^2 \rho^2\left(\mathbf{q}, \mathcal{S}_{\mathbf{p}_0,x_0}\right)$ by Lemma 5. By the same arguments as in CHT, Proof of Theorem 4.2 (Step 1), it now follows that $r_nQ_n(\mathbf{q}) \geq r_n w_n^* C^2 \rho^2\left(\mathbf{q}, \mathcal{S}_{\mathbf{p}_0,x_0}\right)/2$ 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 $\bar{Q}_n(\mathbf{q})$ and the constant $\kappa = \kappa_n = w_n^* C^2$ in their Condition C.2 both depend on n. Finally, the verification that their Condition C.3 is satisfied follows by the same arguments as in Shi and Shum (2012). We now proceed as in CHT, Proof of Theorem 3.2 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 $\mathbf{q}$,

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

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

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

where $c_n \to^P c < \infty$. This proves that the stochastic process $\mathbf{q} \mapsto \left\{\hat{W}_n^{1/2}m_n(\mathbf{q}) - W_n^{1/2}\bar{m}(\mathbf{q})\right\}$ weakly converges on the compact set $\mathcal{B}_{\mathbf{p}_0,x_0}$ towards Z, c.f. Van der Vaart and Wellner (1996, Example 1.5.10). In particular, $\hat{W}_n^{1/2}m_n(\mathbf{q}) = Z_n + W_n^{1/2}\bar{m}(\mathbf{q}) + o_P(1)$ uniformly in $\mathbf{q}$, which in turn implies that, by Slutsky's theorem,

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

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

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

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

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

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

Proof of Theorem 2. First consider the unconstrained estimator: We write the first demand equation as a quantile regression, $q_1(t) = d_1(x, t, \tau) + e(t, \tau)$, where $e(t, \tau) := d_1(x, t, \varepsilon) - d_1(x, t, \tau)$. 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. (14). We claim that

$$
0 < \inf_{x\in\mathcal{X}} f(0|x, t, \tau) \leq \sup_{x\in\mathcal{X}} f(0|x, t, \tau) < \infty, \tag{17}
$$

$$
\sup_{x\in\mathcal{X}} |f(e|x, t, \tau) - f(0|x, t, \tau)| \to 0, \quad |e| \to 0. \tag{18}
$$

From the definition of expression it is easily seen that Eq. (17) holds since $d_1(x, t, \varepsilon)$ and its derivative w.r.t. ε are continuous in x and $\mathcal{X}$ is compact. Eq. (18) 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

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

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 L_2-covering numbers with bracketing of the function class $\mathcal{F}_n(\delta)$, see Van der Vaart and Wellner (1996) for the precise definitions. To complete the proof, we note that in the case of splines $\delta_n = O(\sqrt{k_n/n})$ and $\left\|\pi_n d_{1,0}(\cdot, t, \tau) - d_1(\cdot, t, \tau)\right\|_2 = O\left(k_n^{-m}\right)$. The convergence rate result in the sup-norm is a direct consequence of Lemma 2.1 and Remark 2.1 in Chen et al. (2010).

To derive the asymptotic distribution of $\{\hat{d}_1(x(t), t, \tau)\}_{t=1}^T$, first note that since data is independent over the time, it is sufficient to derive the marginal distributions. We do this by verifying Conditions 6.1–6.2 of Chen et al. (2010, Corollary 6.1).

Their Condition 6.1 is shown to hold above, while their Condition 6.2(i) holds since

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

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, $\xi_0(k_n) := \sup_{x\in\mathcal{X}}\left\|\underline{B}_{k_n}(x)\right\| \leq c\sqrt{k_n}$. Thus, their Condition 6.2(iv) becomes $\xi_0^2(k_n)k_n^3/n \simeq k_n^4/n = O(1)$ and $\xi_0(k_n)k_n^{-3m}n = k_n^{-3m+1/2}n = O(1)$. Finally, the condition, (ii) of their Corollary 6.1 becomes $nk_n^{-2m-1} = O(1)$.

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

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

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

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

Thus, since $\tilde{x}_\tau(t) \leq \mathbf{p}(t)'\mathbf{d}\left(\tilde{x}_\tau(s),s,\tau\right)$, we have $e\widehat{\tilde{x}}_\tau(t) \leq \mathbf{p}(t)'\hat{\mathbf{d}}(\widehat{\tilde{x}}_\tau(s),s,\tau)$ with probability approaching one (w.p.a.1) as $r_n \to 0$. This proves that $\hat{\mathbf{d}} \in \mathcal{D}_{C,n}^T(e)$ w.p.a.1 such that $\hat{\mathbf{d}}_C^e = \hat{\mathbf{d}}$ 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. ■

Appendix B. Lemmas

Lemma 4. *Assume that* (C.1)–(C.2) *hold. Then*

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

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

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

where

$$\begin{aligned}w_n(t,\varepsilon) := &\left\|\left[\mathbf{p}_0'\frac{\partial\mathbf{d}(\bar{x}(t),t,\varepsilon)}{\partial x}\right]^{-1}\left[\mathbf{p}_{0,1:L}-\frac{p_{0,L+1}}{p_L(t)}\mathbf{p}_{1:L}(t)\right]'\right.\\ &\left.\times\Omega_n^{-1/2}(\bar{x}(t),t,\varepsilon)V^{1/2}(\bar{x}(t),t,\varepsilon)\right\|^{-2} > 0. \quad (19)\end{aligned}$$

Proof. Since ε is kept fixed throughout, we suppress any dependence on this in the following. We treat the estimation of $\bar{x}(t)$ as a GMM estimation problem: Define

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

and

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

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

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

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

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

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

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

Given consistency, we therefore have

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

Next, by a first-order Taylor expansion,

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

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

$$\frac{\partial\hat{G}\left(\tilde{x}(t),t\right)}{\partial x}\to^P\frac{\partial G(\bar{x}(t),t)}{\partial x} = \mathbf{p}_0'\frac{\partial\mathbf{d}(\bar{x}(t),t)}{\partial x} > 0. \quad (20)$$

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

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

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 $w_n(t) = w_n(t, \varepsilon)$ defined in the lemma satisfies $w_n(t) = \|a_n(t)\|^{-2}$. Thus,

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

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

Proof. The inequality is trivial for $\mathbf{q} \in \mathcal{S}_{\mathbf{p}_0,x_0}$. Consider any $\mathbf{q} \in \mathcal{B}_{\mathbf{p}_0,x_0} \setminus \mathcal{S}_{\mathbf{p}_0,x_0}$: Let $\mathbf{q}^* = \arg\min_{\mathbf{q}' \in \mathcal{S}_{\mathbf{p}_0,x_0}} \|\mathbf{q} - \mathbf{q}'\|$ be the unique point in $\mathcal{S}_{\mathbf{p}_0,x_0}$ which has minimum distance to $\mathbf{q}$. Let $\delta^* = \mathbf{q}^* - \mathbf{q}$ be the difference such that $\|\delta^*\| = \rho\left(\mathbf{q}, \mathcal{S}_{\mathbf{p}_0,x_0}\right)$. We can decompose the rows of $(\mathbf{P}, \bar{\mathbf{x}})$ 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 $\mathbf{P}^{(1)} = \left[\mathbf{p}^{(1)}(1), \ldots, \mathbf{p}^{(1)}(T_1)\right]' \in \mathbb{R}^{T_1 \times (L+1)}$ and $\bar{\mathbf{x}}^{(1)} = (x^{(1)}(1), \ldots, x^{(1)}(T_1))' \in \mathbb{R}^{T_1}$ for some $T_1 \leq L + 1$, denote the set of rows which contain the binding and non-binding constraints respectively. That is, $\bar{m}^{(1)}(\mathbf{q}^*) := \bar{\mathbf{x}}^{(1)} - \mathbf{P}^{(1)}\mathbf{q}^* = 0$ while $\bar{m}^{(2)}(\mathbf{q}^*) := \bar{\mathbf{x}}^{(2)} - \mathbf{P}^{(2)}\mathbf{q}^* < 0$. 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 $c_1 > 0$,

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

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

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

References

Afriat, S., 1967. The construction of a utility function from demand data. Internat. Econom. Rev. 8, 67–77.
Afriat, S.N., 1973. On a system of inequalities in demand analysis: an extension of the classical method. Internat. Econom. Rev. 14, 460–472.
Andrews, D.W.K., 1999. Estimation when a parameter is on a boundary. Econometrica 67, 1341–1383.
Anevski, D., Hössjer, O., 2006. A general asymptotic scheme for inference under order restrictions. Ann. Statist. 34, 1874–1930.
Beckert, W., 2007. Specification and identification of stochastic demand models. Econom. Rev. 26, 669–683.
Beckert, W., Blundell, R., 2008. Heterogeneity and the nonparametric analysis of consumer choice: conditions for invertibility. Rev. Econom. Stud. 75, 1069–1080.
Belloni, A., Chernozhukov, V., Fernandez-Val, I., 2011. Conditional quantile processes under increasing dimension. Cemmap working paper CWP19/11.
Benkard, L., Berry, S., 2006. On the nonparametric identification of nonlinear simultaneous equations models: comment on Brown (1983) and Roehrig (1988). Econometrica 74, 1429–1440.
Birke, M., Dette, H., 2007. Estimating a convex function in nonparametric regression. Scand. J. Stat. 34, 384–404.
Blundell, R., Browning, M., Crawford, I., 2008. Best nonparametric bounds on demand responses. Econometrica 76, 1227–1262.
Blundell, R., Chen, X., Kristensen, D., 2007. Semi-nonparametric IV estimation of shape-invariant engel curves. Econometrica 75, 1613–1669.
Blundell, R., Duncan, A., Pendakur, K., 1998. Semiparametric estimation and consumer demand. J. Appl. Econom. 13, 435–461.
Blundell, R., Kristensen, D., Matzkin, R., 2011. Bounding quantile demand functions using revealed preference inequalities. Cemmap working paper CWP21/11.
Blundell, R., Kristensen, D., Matzkin, R., 2013a. Control functions and simultaneous equations methods. Amer. Econ. Rev. Pap. Proc. 103, 563–569.
Blundell, R., Kristensen, D., Matzkin, R., 2013b. Consumer demand with unobserved heterogeneity. Mimeo, UCLA.
Blundell, R., Matzkin, R., 2013. Control functions in nonseparable simultaneous equations models. Quant. Econ. (forthcoming).
Brown, B.W., 1983. The identification problem in systems nonlinear in the variables. Econometrica 51, 175–196.
Brown, B.W, Walker, M.B., 1989. The random utility hypothesis and inference in demand systems. Econometrica 57, 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: Heckman, J., Leamer, E. (Eds.), Handbook of Econometrics, vol. 6B. North-Holland, Amsterdam.
Chen, X., Chernozhukov, V., Liao, Z., 2010. On limiting distributions of possibly unbounded functionals of linear sieve M-estimators. Manuscript, Dep. of Economics, Yale University.
Chen, X., Pouzo, D., 2012. Estimation of nonparametric conditional moment models with possibly nonsmooth moments. Econometrica 80, 277–322.
Cherchye, L., De Rock B, B., Vermeulen, F., 2011. The revealed preference approach to collective consumption behavior: testing and sharing rule recovery. Rev. Econom. Stud. 78, 176–198.
Chernozhukov, V., Fernandez-Val, I., Galichon, A., 2009. Improving point and interval estimators of monotone functions by rearrangement. Biometrika 96, 559–575.
Chernozhukov, V., Imbens, G., Newey, W.K., 2007a. Instrumental variable estimation of nonseparable models. J. Econometrics 139, 4–14.
Chernozhukov, V., Hong, H., Tamer, E., 2007b. 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. In: Blundell, R., Persson, T., Newey, W. (Eds.), Advances in Economics and Econometrics, Theory and Applications, 9th World Congress. Cambridge University Press, Cambridge.
Deaton, A., Muellbauer, J., 1980. Economics and Consumer Behavior. Cambridge University Press.
Haag, B.R., Hoderlein, S., Pendakur, K., 2009. Testing and imposing Slutsky symmetry in nonparametric demand systems. J. Econometrics 153, 33–50.
Hall, P., Horowitz, J., 2005. Nonparametric methods for inference in the presence of instrumental variables. Ann. Statist. 33, 2904–2929.
Hoderlein, S., Mammen, E., 2007. Identification of marginal effects in nonseparable models without monotonicity. Econometrica 75, 1513–1518.
Hoderlein, S., Mammen, E., 2009. Identification and estimation of local average derivatives in non-separable models without monotonicity. Econom. J. 12, 1–25.
Hoderlein, S., Stoye, J., 2013. Revealed preferences in a heterogeneous population. Rev. Econ. Stat. (forthcoming).
Hoderlein, S., Stoye, J., 2013. Testing stochastic rationality and predicting stochastic demand: the case of two goods. Manuscript, Dep. of Economics, Brown University.
Houthakker, H.S., 1950. Revealed preference and the utility function. Economica 17, 159–174.
Imbens, G., Newey, W.K., 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: Kallianpur, G., Krishnaiah, P.R., Ghosh, J.K. (Eds.), Statistics and Probability: Essays in Honor of C.R. Rao. North-Holland, Amsterdam, pp. 419–428.
Kitamura, Y., Stoye, J., 2012. Nonparametric analysis of random utility models. Manuscript, Dep. of Economics, Yale University.
Kline, B., Tamer, E., 2013. Default Bayesian inference in a class of partially identified models. Manuscript, Northwestern University.
Koenker, R., Bassett Jr., G., 1978. Regression quantiles. Econometrica 46, 33–50.
Koenker, R., Ng, P., 2005. Inequality constrained quantile regression. Sankhyā 67, 418–440.
Lewbel, A., 1995. Consistent nonparametric hypothesis tests with an application to Slutsky symmetry. J. Econometrics 67, 379–401.
Lewbel, A., 2001. Demand systems with and without errors. Amer. Econ. Rev. 91, 611–618.
Lewbel, A., Ng, S., 2005. Demand systems with nonstationary prices. Rev. Econ. Stat. 87, 479–494.

Mammen, E., 1991. Estimating a smooth monotone regression function. Ann. Statist. 19, 724–740.

Manski, C., 1993. Identification problems in the social sciences. Sociol. Methodol. 23, 1–56.

Mas-Colell, A., 1978. On revealed preference analysis. Rev. Econom. Stud. 45, 121–131.

Matzkin, R.L., 2003. Nonparametric estimation of nonadditive random functions. Econometrica 71, 1339–1375.

Matzkin, R.L., 2007. Heterogeneous choice. In: Blundell, R., Newey, W., Persson, T. (Eds.), Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress of the Econometric Society. Cambridge University Press, Cambridge.

Matzkin, R.L., 2008. Identification in nonparametric simultaneous equations. Econometrica 76, 945–978.

Matzkin, R.L., Richter, M.K., 1991. Testing strictly concave rationality. J. Econom. Theory 53, 287–303.

McElroy, M.B., 1987. Additive general error models for production, cost, and derived demand or share systems. J. Political Econom. 95, 737–757.

McFadden, D.L., 2005. Revealed stochastic preference: a synthesis. Econom. Theory 26, 245–264.

McFadden, D., Richter, K., 1991. Stochastic rationality and revealed stochastic preference. In: Chipman, J., McFadden, D., Richter, K. (Eds.), Preferences, Uncertainty, and Rationality. Westview Press, pp. 161–186.

Mukerjee, H., 1988. Monotone nonparametric regression. Ann. Statist. 16, 741–750.

Newey, W.K., 1997. Convergence rates and asymptotic normality for series estimators. J. Econometrics 79, 147–168.

Newey, W.K., Powell, J., 2003. Instrumental variables estimation for nonparametric models. Econometrica 71, 1565–1578.

Newey, W.K., Powell, J., Vella, F., 1999. Nonparametric estimation of triangular simultaneous equations models. Econometrica 67, 565–603.

Richter, M.K., 1966. Revealed preference theory. Econometrica 34, 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 (5), 61–71.

Shi, X., Shum, M., 2012. Simple two-stage inference for a class of partially identified models. Manuscript University of Wisconsin.

Self, S.G., Liang, K., 1987. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. J. Amer. Statist. Assoc. 82, 605–610.

Van der Vaart, A., Wellner, J., 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.

Wolak, F.A., 1991. The local nature of hypothesis tests involving nonlinear inequality constraints. Econometrica 59, 981–995.

Wright, F.T., 1981. The asymptotic behavior of monotone regression estimates. Ann. Statist. 9, 443–448.