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