Institute for Fiscal Studies

Demand analysis with partially observed prices

IFS Working Paper W15/16

Ian Crawford
Matthew Polisson

1965 - 2015
50
E·S·R·C
50 YEARS SHAPING SOCIETY

The Institute for Fiscal Studies (IFS) is an independent research institute whose remit is to carry out rigorous economic research into public policy and to disseminate the findings of this research. IFS receives generous support from the Economic and Social Research Council, in particular via the ESRC Centre for the Microeconomic Analysis of Public Policy (CPP). The content of our working papers is the work of their authors and does not necessarily represent the views of IFS research staff or affiliates.

Demand Analysis with Partially Observed Prices

Ian Crawford and Matthew Polisson*

June 23, 2015

Abstract: In empirical demand, industrial organization, and labor economics, prices are often unobserved or unobservable since they may only be recorded when an agent transacts. In the absence of any additional information, this partial observability of prices is known to lead to a number of identification problems. However, in this paper, we show that theory-consistent demand analysis remains feasible in the presence of partially observed prices, and hence partially observed implied budget sets, even if we are agnostic about the nature of the missing prices. Our revealed preference approach is empirically meaningful and easy to implement. We illustrate using simple examples.

Keywords: demand, missing prices, partial identification, revealed preference

JEL classification numbers: D11, D12

* Crawford (ian.crawford@economics.ox.ac.uk): Department of Economics, University of Oxford, Manor Road, Oxford, OX1 3UQ, United Kingdom; Nuffield College, New Road, Oxford, OX1 1NF, United Kingdom; Institute for Fiscal Studies, 7 Ridgmount Street, London, WC1E 7AE, United Kingdom; Polisson (matthew.polisson@le.ac.uk): Department of Economics, University of Leicester, University Road, Leicester, LE1 7RH, United Kingdom; Institute for Fiscal Studies, 7 Ridgmount Street, London, WC1E 7AE, United Kingdom. Part of this research was carried out while Matthew Polisson was visiting Caltech in 2015, and he would like to thank Caltech for their hospitality and support. We acknowledge financial assistance from the ESRC Centre for the Microeconomic Analysis of Public Policy (No. RES-544-28-5001). We are grateful to Abi Adams, Jerry Hausman, and conference participants at the 2014 Workshop on Nonparametric Demand (IFS and CeMMAP) for helpful contributions.

1. Introduction

It is not uncommon for a data set to contain incomplete observations. The particular, yet important, case that we study in this paper arises due to the fact that prices are typically only observed when a transaction occurs, e.g., the wage of a worker is only observed when that individual is employed, or the price of a good is only observed when a consumer makes a purchase. This means that an agent's implied choice set is ill-defined from the perspective of the econometrician. The combination of missing prices and zero purchases is an important feature of many data sets and presents numerous challenges to empirical work. These problems are typically magnified when the data involved are high dimensional, e.g., when an agent is choosing on both the extensive and intensive margins among many goods at once.

Manski (2003) studies a number of identification problems that relate to data of this type, e.g., how to estimate the joint distribution of prices, or a feature of that distribution such as its mean. A number of other important papers (e.g., Heckman (1979), Deaton and Irish (1984), Keen (1986), Atkinson, Gomulka, and Stern (1990), and Meghir and Robin (1992)) study the problem of recovering consumer preferences in the presence of data with this particular feature. The principal difficulty with such data is that while, as the above studies show, zero purchases can arise for several reasons,[1] and the correct procedure for dealing with the resulting partial observability of prices generally depends upon that reason,[2] the econometrician may not be able to distinguish between them.

In this paper, we have two main objectives. First, we emphasize the importance of the partial observability of prices, and hence implied budget sets,[3] in observational consumer panel data, a feature made more obviously prominent as the dimensions of the data grow large. While much of the current econometric work in consumer demand[4] tends to assume access to an *idealized* data set, we argue that it is important to recognize the prominent

[1] Zero purchases might be the result of choices (e.g., corners), constraints (e.g., temporary local unavailability due to the fact that a good has not yet been introduced or has been discontinued), or measurement (e.g., infrequencies).

[2] For example, infrequencies may demand one approach, choices another, and rationing yet another.

[3] Notice that 'true' budget sets are also typically partially observed in observational consumer panel data since we usually observe expenditure rather than income. This is a longstanding issue in empirical consumer demand and industrial organization that is normally resolved by invoking a separability argument.

[4] The literature on the estimation of consumer demand has deep roots in economics, dating back to the earliest linear models proposed by Klein and Rubin (1947–1948), Samuelson (1947–1948), Geary (1950), and Stone (1954), through to the more flexible parametric approaches of Deaton and Muellbaeur (1980a,

features of *real* observational consumer panel data, and therefore to give the problem of partially observed prices first-order attention. Second, we present a novel result, which is elementary, illuminating, and constructive, and which serves as an agnostic point of departure for thinking about the problem of partial price observability. We claim that it is *not*, in fact, necessary to know the exact nature of a zero purchase and corresponding missing price in order to make meaningful empirical progress when prices are partially observed. We show that it is possible to carry out both positive and normative economic analysis of demand and consumer behavior even in the presence of partially observed prices. This is a result of importance to many applied researchers.

2. Data Setting

Consider a finite set of repeated observations on an individual consumer. Suppose that there are K goods, each indexed by $k \in \{1, 2, \ldots, K\}$, and T observations, each indexed by $t \in \{1, 2, \ldots, T\}$. Let $x_k^t \in \mathbb{R}_+$ denote the consumer's demand for good k at observation t, and let her corresponding consumption bundle be given by $x^t = (x_1^t, x_2^t, \ldots, x_K^t) \geq 0$. We denote the price of good k at observation t by $p_k^t \in \mathbb{R}_{++}$ and the corresponding price vector by $p^t = (p_1^t, p_2^t, \ldots, p_K^t) \gg 0$. In an observational setting, a data set is therefore given by

$$\mathcal{D} = \left\{ \left(p_k^t \,|\, x_k^t > 0,\ x_k^t \right) \right\}_{k=1,2,\ldots,K}^{t=1,2,\ldots,T},$$

that is, we observe all of the demands, but prices are only observed conditional on a non-zero demand. The data set alone contains no information about $p_k^t \,|\, x_k^t = 0$.

1980b), Jorgenson, Lau, and Stoker (1982), Banks, Blundell, and Lewbel (1997), and Lewbel and Pendakur (2009)—this list is far from exhaustive. More recently, the econometric literature has tended towards a semi- or non-parametric approach to consumer demand, which has spawned a large set of papers establishing necessary and/or sufficient conditions for identification, consistent estimation, and statistical inference in a variety of settings, e.g., Lewbel (1991), Hausman and Newey (1995), Lewbel (1995), Newey, Powell, and Vella (1999), Chesher (2003), Matzkin (2003, 2008), Newey and Powell (2003), Altonji and Matzkin (2005), Blundell, Chen, and Kristensen (2007), Chernozhukov, Imbens, and Newey (2007), Beckert and Blundell (2008), Haag, Hoderlein, and Pendakur (2009), and Imbens and Newey (2009). A further stream has introduced new notions of partial identification, e.g., Manski (2003, 2007), Imbens and Manski (2004), Chernozhukov, Hong, and Tamer (2007), and Chernozhukov, Lee, and Rosen (2013). A related literature on stochastic discrete choice demand beginning with Block and Marschak (1960), Brown and Walker (1989), and McFadden and Richter (1991) has spawned a large body of applied work in both empirical consumer demand and industrial organization (e.g., Berry, Levinsohn, and Pakes (1995)). More recently, several papers have attempted to marry the demand and revealed preference approaches, e.g., Blundell, Browning, and Crawford (2003, 2007, 2008), Hoderlein (2011), Kitamura and Stoye (2013), Blundell, Kristensen, and Matzkin (2014), and Hoderlein and Stoye (2014, 2015).

As an example, with 6 goods and 4 observations, the schematics of a data set for an individual consumer might look something like this:

$$X = \begin{pmatrix} x_1^1 & 0 & 0 & x_1^4 \\ 0 & x_2^2 & x_2^3 & 0 \\ x_3^1 & 0 & x_3^3 & x_3^4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & x_5^3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad P = \begin{pmatrix} p_1^1 & \cdot & \cdot & p_1^4 \\ \cdot & p_2^2 & p_2^3 & \cdot \\ p_3^1 & \cdot & p_3^3 & p_3^4 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & p_5^3 & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}.$$

The goods are arranged in rows, and the observations in columns. In this example, the consumer purchases good 3 frequently, goods 1 and 2 only occasionally, good 5 rarely, and goods 4 and 6 never at all. The majority (two thirds) of the price data are missing.

To give an example of a relatively new source of data that is potentially extremely valuable but which suffers from missing prices, consider electronically gathered consumer panel data (known sometimes as scanner data). The increased availability of these data has made it possible to carry out a wide range of new empirical work on consumer demand and industrial organization, particularly involving highly differentiated and disaggregated goods, often down to the stock-keeping unit (UPC) barcode. The key features of such data sets are that (i) the number of products is typically very large, (ii) there are many instances of zero demands, and (iii) prices are only recorded when a consumer makes a purchase. As a consequence, while quantities and expenditures are indeed fully observed for every item (where both are equal to zero when a product is unpurchased), prices are only partially observed. The price data for an individual consumer are, in fact, very likely to be extremely sparse. As an example, consider a typical consumer drawn from the Kantar Worldpanel who was observed to have purchased 2,901 different products over the course of 207 days. This amounts to 600,507 product/day observations. However, most products were purchased rarely (60% only once, and 81% three times or fewer), and as a result, 591,073 (or 98.4%) of the corresponding prices were unobserved or unobservable.[5]

One strategy to counteract missing prices for an individual consumer has been to construct a local average of observed prices. Typically, this is an average of prices for similar

[5] We thank Abi Adams for providing the figures used in this example.

products, bought by similar consumers, in similar locations, on similar dates. There is of course a natural *empirical* tradeoff between the number of cells and the number of observations within each cell, and the sparsity of observed prices in highly disaggregated consumer panel data makes this aspect of the curse of dimensionality even more pronounced. However, if the empirical obstacles to price imputation and aggregation are demanding, the *theoretical* problems associated with such procedures are at least as formidable, since it is not typically feasible to identify the nature of a zero purchase and its corresponding missing price.[6]

In the next section, we show that empirical demand analysis in the form of heterogeneous revealed preference is possible even when prices are only partially observed. Furthermore, meaningful empirical content obtains even under minimal assumptions over preferences.

3. Revealed Preference

Suppose that the true data-generating process for the data described in the previous section involves the maximization of a stable preference subject to a sequence of linear budget and rationing constraints with exogenous prices and incomes.[7,8] At every observation t, the consumer's optimization problem is therefore given by

$$\max_{x \in \mathbb{R}_+^K} u(x) \text{ subject to } p^t \cdot x \leq e^t \text{ and } x_k^t = 0 \text{ for any } k \in \mathcal{K}^t,$$

where the utility function u is increasing, concave, and continuous, and where $e^t > 0$ is the consumer's income at observation t, and $\mathcal{K}^t$ the index set of goods which are unavailable to the consumer at observation t. Notice that neither e^t nor $\mathcal{K}^t$ are available to the econometrician, i.e., it is impossible for the analyst to test whether the consumer exhausts her income, and whether a zero purchase is a choice or due to rationing.

[6] See, e.g., Meghir and Robin (1992), pp. 54–55.

[7] In this paper, we restrict our attention to the static model. It would be straightforward to extend these ideas to an intertemporal setting using the approach of Browning (1989). It would also be possible to introduce intertemporal nonseparabilities following Crawford (2010) and Demuynck and Verriest (2013). This would allow for the possibility of observed or unobserved stocks of durables, the presence of which might also influence a consumer's decision to transact. We reserve these and other extensions as topics of future research in order not to obscure the main insights.

[8] In many of the data sets where zero purchases and missing prices might present a problem, it is also the case that many goods are only available in discrete amounts. Polisson and Quah (2013) discuss the nonparametric revealed preference approach in a discrete consumption space. The availability of some goods in continuous amounts or an outside good (typically money) available in continuous amounts is enough to ensure that discreteness adds no important complications to the arguments set out below.

For every good k and at every observation t, the consumer's first order conditions are then given according to

$$\begin{aligned} u_k(x^t) &= \lambda^t p_k^t \;\text{ for all }\; x_k^t > 0, \\ u_k(x^t) &\leq \lambda^t p_k^t \;\text{ for all }\; x_k^t = 0,\; k \notin \mathcal{K}^t, \\ u_k(x^t) &= \lambda^t \left(p_k^t + \frac{\mu_k^t}{\lambda^t} \right) \;\text{ for all }\; x_k^t = 0,\; k \in \mathcal{K}^t, \end{aligned}$$

where $u_k(x^t)$ denotes the partial derivative of the utility function u with respect to good k evaluated at the consumption bundle x^t,[9] and where the multiplier $\lambda^t \in \mathbb{R}_{++}$ is the marginal utility of income at observation t, and the multiplier $\mu_k^t \in \mathbb{R}_{++}$ the marginal (utility) cost of rationing good k at observation t. The demands generated under these circumstances can be expressed in terms of unrationed demands by allowing for choice over the entire product space, but replacing the observed market prices with a vector of 'support' prices.[10] The support prices are such that an unrationed choice problem would generate exactly the same demands as those which were generated under rationing. Monotonicity, concavity, and continuity of the consumer's utility function are sufficient to guarantee the existence of a set of strictly positive support prices consistent with any set of demands.[11] The support prices themselves are a mixture of various prices: for purchased goods, they are identical to observed prices; for goods that are available but not purchased, they are equal to reservation prices; and for goods that are not available, they are equal to 'virtual' prices (the lowest prices consistent with zero demands in the absence of any rationing constraints). Denoting the support price of good k at observation t by π_k^t, we have

$$\begin{aligned} \pi_k^t &= p_k^t \;\text{ for all }\; x_k^t > 0, \\ \pi_k^t &= \frac{u_k(x^t)}{\lambda^t} \;\text{ for all }\; x_k^t = 0,\; k \notin \mathcal{K}^t, \\ \pi_k^t &= p_k^t + \frac{\mu_k^t}{\lambda^t} \;\text{ for all }\; x_k^t = 0,\; k \in \mathcal{K}^t. \end{aligned}$$

[9] Note that we use differentiability of the utility function here to develop a simple argument for the purposes of building intuition; we do not appeal to differentiability in order to establish any formal results.

[10] See Hicks (1940), Rothbarth (1941), Neary and Roberts (1980), and Hausman (1997) for seminal treatments of consumer behavior under rationing and the economic valuation of new goods. See also Varian (1983) and Fleissig and Whitney (2011) for a revealed preference approach to rationing, and relatedly, Demuynck and Seel (2014) for a revealed preference approach to limited consideration.

[11] See Neary and Roberts (1980), pp. 27–29, for the formal result.

Using these support prices, the observed demand at observation t is simply the solution to the following unrationed constrained optimization problem:

$$\max_{x \in \mathbb{R}^K_+} u(x) \text{ subject to } \pi^t \cdot x \leq e^t.$$

In data of this kind, when the demand for a good is zero, the corresponding price is either *unobserved* (when the zero purchase is the result of a choice) or *unobservable* (when the zero purchase is due to unavailability). There is often no obvious way for a researcher to identify which type of zero purchase obtains, unless of course she has recourse to some outside identifying information.[12]

What restrictions, if any, does economic theory imply about consumer behavior in such circumstances when prices (and hence implied budget sets) are only partially observed? The following definition sets out formally what is required in order to *rationalize* the data set $\mathcal{D} = \{(p^t_k \,|\, x^t_k > 0,\ x^t_k)\}^{t=1,2,\ldots,T}_{k=1,2,\ldots,K}$.

Definition 1. *A utility function* $u : \mathbb{R}^K_+ \to \mathbb{R}$ rationalizes *the data set* $\mathcal{D}$ *if there exist support prices* $\pi^t \in \mathbb{R}^K_{++}$ *(with* $\pi^t_k = p^t_k$ *for any* $x^t_k > 0$*) such that, at every observation* $t = 1, 2, \ldots, T$, $u(x^t) \geq u(x)$ *for any* $x \in \{x \in \mathbb{R}^K_+ : \pi^t \cdot x \leq \pi^t \cdot x^t\}$.

The above definition states that in order to rationalize the observed behavior, there must exist a utility function *and* corresponding reservation and virtual prices such that the observed choices are indeed maximizing. Our main result is below.

Proposition 1. *The following statements are equivalent:*

(1) The data set $\mathcal{D}$ *is rationalizable by a nonsatiated utility function* $u : \mathbb{R}^K_+ \to \mathbb{R}$.

(2) The data set $\mathcal{D}$ *is rationalizable by a utility function* $u : \mathbb{R}^K_+ \to \mathbb{R}$*, which is increasing, concave, and continuous.*

(3) Given the data set $\mathcal{D}$*, at every observation* $t = 1, 2, \ldots, T$*, there exist numbers* $u^t \in \mathbb{R}$ *and* $\lambda^t \in \mathbb{R}_{++}$*, and vectors* $\rho^t \in \mathbb{R}^K_{++}$*, such that*

$$u^{t'} \leq u^t + \rho^t \cdot (x^{t'} - x^t) \text{ for all } t, t' = 1, 2, \ldots, T,$$

[12] One such source of information would clearly be to observe another consumer making a purchase of the same product in the same location at the same time. With products disaggregated down to the level of UPC barcodes, such an occurrence is highly unlikely in most data sets due to the curse of dimensionality.

$$\rho_k^t = \lambda^t p_k^t \ \ (\textit{for any} \ \ x_k^t > 0) \ \ \textit{for all} \ \ k = 1, 2, \ldots, K, \ t = 1, 2, \ldots, T.$$

Proof. The proof is given in the Appendix. □

Our main proposition establishes a set of necessary and sufficient conditions for maximizing behavior in the presence of partially observed prices. The following remarks help to situate the result:

(i) The equivalence between statements (1) and (2) implies that strong monotonicity, concavity, and continuity of the utility function are without loss of generality. As in the classical setting when prices are fully observed, these additional properties of the utility function (beyond nonsatiation) are untestable in a finite data setting, i.e., we get them for free.

(ii) Statement (3) reveals that the problem is linear, and therefore easily solvable, using computationally efficient algorithms.

(iii) The support prices themselves are a mixture of observed prices, reservation prices, and virtual prices. Notice that reservation prices and virtual prices can be constructed from the (not necessarily unique) solution to the set of inequalities in statement (3). Further note that in the absence of any identifying information about the nature of a zero purchase, we are unable to empirically distinguish between reservation prices and virtual prices.[13] Nonetheless, as Proposition 1 indicates, and as the examples in the next section illustrate, it is not necessary to draw such a distinction.

(iv) Afriat's (1967) Theorem obtains when prices are fully observed.

(v) Products that are never purchased can be excluded from the empirical analysis entirely since they provide no further restrictions on the data. This potentially eases the empirical implementation.

(vi) If the price of a good that is purchased is never observed, then the restrictions on the data are vacuously satisfied and any choice behavior is rationalizable. This is not a

[13] The support price of good k at observation t can be constructed according to $\pi_k^t = \rho_k^t/\lambda^t$. Notice that $\pi_k^t = p_k^t$ when prices are observed, and that we can construct $\pi_k^t = u_k(x^t)/\lambda^t$ or $\pi_k^t = p_k^t + \mu_k^t/\lambda^t$ when prices are unobserved or unobservable.

circumstance that deserves much emphasis here, primarily since the data structure of interest precludes it, but nonetheless it is worth noting that Proposition 1 delivers the same result as Theorem 1 in Varian (1988) under these circumstances.

(vii) Proposition 1 can be used to make demand predictions and conduct welfare analysis.

On this last point, we can elaborate. Given the data set $\mathcal{D} = \{(p_k^t \,|\, x_k^t > 0,\ x_k^t)\}_{k=1,2,\ldots,K}^{t=1,2,\ldots,T}$, for some hypothetical normalized price vector $p^0 \in \mathbb{R}_{++}^K$, we can define the set of consumption bundles which are rationalizable at this price vector according to

$$\mathcal{S}(p^0 \,|\, \mathcal{D}) = \{x^0 \in \mathbb{R}_+^K \ :\ p^0 \cdot x^0 = 1,\ \mathcal{D} \cup \{(p^0, x^0)\} \text{ is rationalizable}\},$$

i.e., the set of demand predictions at a hypothetical budget must be consistent with the observed data. The following proposition is important in empirical work.

Proposition 2. *Given any data set $\mathcal{D}$ which satisfies the conditions in Proposition 1, the support set $\mathcal{S}(p^0 \,|\, \mathcal{D})$ is convex.*

Proof. The proof is given in the Appendix. □

Convexity of the support set is an important property both for describing bounds on demand responses and also making welfare comparisons (see, e.g., Blundell, Browning, and Crawford (2008) and Blundell *et al.* (2014)). It is therefore useful that this property is preserved even under partially observed prices. Notice that if all prices were observed and we did have access to the full data set $\mathcal{O} = \{(p_k^t, x_k^t)\}_{k=1,2,\ldots,K}^{t=1,2,\ldots,T}$, the zeros would be known to be corner colutions and the standard Varian (1982) support set $\mathcal{S}(p^0 \,|\, \mathcal{O})$ obtains. Further notice that in general $\mathcal{S}(p^0 \,|\, \mathcal{O}) \subseteq \mathcal{S}(p^0 \,|\, \mathcal{D})$, and therefore that the coverage probability of the 'true' support set by the support set available when prices are only partially observed is equal to one.

To summarize, economic theory provides empirically meaningful restrictions on observables even when prices are only partially observed (i.e., only when the consumer transacts). These restrictions allow choice behavior to be examined for consistency with utility maximization without the need to impute missing prices (typically thought to be either reservation or virtual prices). Subject to these conditions being satisfied, procedures are available to provide bounds on demand forecasts and welfare measures.

4. Examples

4.1 Falsifiability

As an example, with 3 goods and 5 observations, the data on an individual consumer might be represented by

$$X = \begin{pmatrix} 5 & 0 & 4 & 2 & 1 \\ 2 & 4 & 0 & 5 & 3 \\ 0 & 0 & 4 & 0 & 1 \end{pmatrix}, \qquad P = \begin{pmatrix} 5 & \cdot & 3 & 4 & 2 \\ 4 & 5 & \cdot & 4 & 1 \\ \cdot & \cdot & 3 & \cdot & 2 \end{pmatrix},$$

with goods arranged in rows and observations in columns. This is an example of a data set which is *not* rationalizable, i.e., there do not exist any prices that support the observed consumption choices as having arisen from the maximization of a nonsatiated preference. In this simple case, it is relatively easy to see why. If we focus on a subset of the data containing only observations 1 and 4, then we have

$$X' = \begin{pmatrix} 5 & 2 \\ 2 & 5 \\ 0 & 0 \end{pmatrix}, \qquad P' = \begin{pmatrix} 5 & 4 \\ 4 & 4 \\ \cdot & \cdot \end{pmatrix}.$$

Since products which are never purchased can be excluded, we are left with

$$X'' = \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}, \qquad P'' = \begin{pmatrix} 5 & 4 \\ 4 & 4 \end{pmatrix},$$

which clearly violates any notion of rationality in the sense that $(2,5)$ is purchased when $(5,2)$ is precisely affordable, but $(5,2)$ is purchased when $(2,5)$ is more than affordable.

4.2 Rationalizability

As another example, again with 3 goods and 5 observations, the data on an individual consumer might be represented by

$$X = \begin{pmatrix} 1 & 1 & 1 & 0 & 4 \\ 0 & 2 & 0 & 1 & 0 \\ 1 & 2 & 1 & 4 & 4 \end{pmatrix}, \qquad P = \begin{pmatrix} 3 & 2 & 5 & \cdot & 5 \\ \cdot & 2 & \cdot & 2 & \cdot \\ 4 & 4 & 1 & 4 & 4 \end{pmatrix}.$$

This data set is, in fact, rationalizable. Given two sets of potential support prices

$$\Pi = \begin{pmatrix} 3 & 2 & 5 & 4 & 5 \\ 4 & 2 & 4 & 2 & 4 \\ 4 & 4 & 1 & 4 & 4 \end{pmatrix}, \qquad \Pi' = \begin{pmatrix} 3 & 2 & 5 & 5 & 5 \\ 4 & 2 & 2 & 2 & 3 \\ 4 & 4 & 1 & 4 & 4 \end{pmatrix},$$

the set of demands X is rationalizable by Π but not Π', i.e., the data are rationalizable in the sense that there exist some prices supporting the maintained hypothesis of utility maximization, but it is also possible to find prices which refute this hypothesis. What this example illustrates is that rationalizability based on some procedure for imputing the missing prices, and any ensuing counterfactual and welfare analysis, is sensitive to the choice of imputation. Within the framework developed in this paper, we allow the researcher to be completely agnostic about any unobserved or unobservable prices, which reduces the probability of model misspecfication due to price imputation and aggregation.

4.3 Demand Predictions

A first question one might ask of a data set which is rationalizable is how to make demand predictions at hypothetical or previously unobserved budgets. Estimating demand functions and bounding demand responses have been longstanding positive economic objectives in the empirical demand and revealed preference literatures. Essentially, one can appeal to the structure of the model of utility maximization in order to construct counterfactuals of interest. The support sets defined in the previous section allow us to compute bounds on these counterfactuals straightforwardly.

Recall the example from the previous subsection. Suppose we are interested in bounding a demand response at the means of the observed prices $p^* = (3.75, 2, 3.4)$ and expenditure $e^* = 16.2$. We are therefore interested in the support sets given by $\mathcal{S}(p^0 \mid \{(P, X)\})$ and $\mathcal{S}(p^0 \mid \{(\Pi, X)\})$, where the hypothetical price vector $p^0 = p^*/e^*$ has been normalized. The support sets can be depicted as budget shares on the unit simplex as shown in Figure 1.[14] The union of the lighter and darker shaded regions corresponds to $\mathcal{S}(p^0 \mid \{(P, X)\})$, and the

[14] Orientation: if a consumer were to devote her entire budget to good 3, then her demand would lie in the top corner of the simplex; if her demands were such that she had equal budget shares, this would be represented by a point at the center; if she decided not to purchase good 1, then her demands would be represented by a point somewhere on the edge connecting the top and the bottom-right corners.

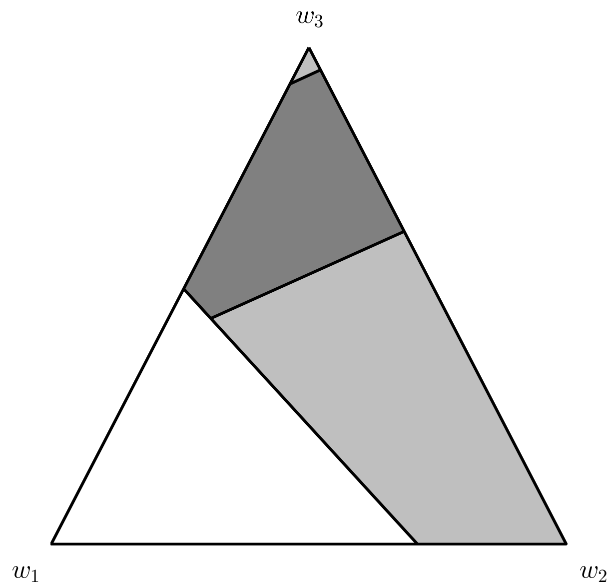

Figure 1: Support Sets on the Budget Simplex

darker shaded region to $\mathcal{S}(p^0 \mid \{(\Pi, X)\})$. The set-valued predictions are 'sharp' in the sense that they exhaust the empirical content of both the theory (maximizing behavior) and the data (observed prices and quantities).

4.4 Welfare Analysis

The bounds on the compensating and equivalent variation associated with the change in prices and income from $\{(p^2, e^2)\}$ to $\{(p^*, e^*)\}$ are given by

$$CV = c(p^*, u^*) - c(p^*, u^2) \in [1.65, 9.05],$$

$$EV = c(p^2, u^*) - c(p^2, u^2) \in [0.005, 4.906].$$

5. Conclusions

Consumer panels with very finely disaggregated products are a relatively new and potentially very rich source of data for applied work in consumer demand and empirical industrial organization. These data sets also suffer from the pervasive problem of partially observed prices. This paper has shown that economic theory continues to provide meaningful restrictions, even nonparametrically, in the presence of partially observed prices. We have defined a set of necessary and sufficient conditions for theoretical consistency, and further demonstrated how they might be used to make counterfactual demand predictions and to perform welfare analysis. Many important challenges remain—partially observable prices weaken the restrictions of economic theory, thereby making it easier for a data set to appear consistent with maximizing behavior and widening the bounds on demand forecasts and welfare measures compared to the fully observable case. Finding ways to improve these bounds, perhaps by combining the conditions outlined in this paper with the nonparametric statistical methods used by Blundell *et al.* (2003, 2008, 2015), remains an important task.

Appendix

A.1 Preliminaries

Let $\mathcal{O} = \{(p^t_k, x^t_k)\}^{t=1,2,\ldots,T}_{k=1,2,\ldots,K}$ be a set of observations drawn from a consumer. Each observation consists of a price vector $p^t = (p^t_1, p^t_2, \ldots, p^t_K) \gg 0$ and a corresponding consumption bundle $x^t = (x^t_1, x^t_2, \ldots, x^t_K) \geq 0$. Given the data set $\mathcal{O}$, we say that (1) x^t is *directly revealed preferred* to x^s ($x^t \succeq^* x^s$) if $p^t \cdot x^s \leq p^t \cdot x^t$, (2) x^t is *strictly directly revealed preferred* to x^s ($x^t \succ^* x^s$) if $p^t \cdot x^s < p^t \cdot x^t$, and (3) x^t is *revealed preferred* to x^s ($x^t \succeq x^s$) if $x^t \succeq^* x^i$, $x^i \succeq^* x^j$, … , $x^k \succeq^* x^l$, $x^l \succeq^* x^s$. The data set $\mathcal{O}$ obeys the Generalized Axiom of Revealed Preference (GARP) so long as $x^t \succeq x^s \implies x^s \not\succ^* x^t$.

Now we restrict our attention to the data set $\mathcal{D} = \{(p^t_k \,|\, x^t_k > 0,\ x^t_k)\}^{t=1,2,\ldots,T}_{k=1,2,\ldots,K}$ and the notion of rationalizability in Definition 1.

Theorem 1. *The following statements are equivalent:*

(1) The data set $\mathcal{D}$ is rationalizable by a nonsatiated utility function $u : \mathbb{R}^K_+ \to \mathbb{R}$.

(2) Given the data set $\mathcal{D}$, at every observation $t = 1, 2, \ldots, T$, there exist support prices $\pi^t \in \mathbb{R}^K_{++}$ (with $\pi^t_k = p^t_k$ for any $x^t_k > 0$), such that $\{(\pi^t_k, x^t_k)\}^{t=1,2,\ldots,T}_{k=1,2,\ldots,K}$ obeys GARP.

(3) Given the data set $\mathcal{D}$, at every observation $t = 1, 2, \ldots, T$, there exist support prices $\pi^t \in \mathbb{R}^K_{++}$ (with $\pi^t_k = p^t_k$ for any $x^t_k > 0$), and numbers $u_t \in \mathbb{R}$ and $\lambda_t \in \mathbb{R}_{++}$, such that

$$u^{t'} \leq u^t + \lambda^t \pi^t \cdot (x^{t'} - x^t) \ \text{ for all } \ t, t' = 1, 2, \ldots, T.$$

(4) The data set $\mathcal{D}$ is rationalizable by a utility function $u : \mathbb{R}^K_+ \to \mathbb{R}$, which is increasing, concave, and continuous.

Proof. See Afriat (1967), Diewert (1973), and Varian (1982). $\square$

It is easy to see how Afriat's Theorem might be adapted to account for the partial observability of prices, i.e., to establish necessary and sufficient conditions on the data set $\mathcal{D} = \{(p^t_k \,|\, x^t_k > 0,\ x^t_k)\}^{t=1,2,\ldots,T}_{k=1,2,\ldots,K}$. All of the usual results obtain, i.e., the costlessness of assuming monotonicity, concavity, and continuity over and above nonsatiation, and the equivalence between checking a no-cycling condition on the data and finding a solution to a set of inequalities constructed from the data. However, in their current forms, the conditions in statements (2) and (3) are not implementable; in statement (2), GARP is defined over a partially observed price vector, and the inequalities in statement (3) are nonlinear, which is a computationally hard problem. Proposition 1 remedies this by establishing a further equivalence.

A.2 Proof of Proposition 1

Proof. Necessity: Given the data set $\mathcal{D}$, at every observation $t = 1, 2, \ldots, T$, suppose that there exist support prices $\pi^t \in \mathbb{R}^K_{++}$ (with $\pi^t_k = p^t_k$ for any $x^t_k > 0$), and numbers $u^t \in \mathbb{R}$ and $\lambda^t \in \mathbb{R}_{++}$, such that

$$u^{t'} \leq u^t + \lambda^t \pi^t \cdot (x^{t'} - x^t) \ \text{ for all } \ t,\, t' = 1, 2, \ldots, T.$$

Let $\rho^t = \lambda^t \pi^t$ for all $t = 1, 2, \ldots, T$. Notice that $\rho^t \in \mathbb{R}^K_{++}$ and that $\rho^t_k = \lambda^t p^t_k$ for any $x^t_k > 0$.

Sufficiency: Given the data set $\mathcal{D}$, at every observation $t = 1, 2, \ldots, T$, suppose that there exist numbers $u^t \in \mathbb{R}$ and $\lambda^t \in \mathbb{R}_{++}$, and vectors $\rho^t \in \mathbb{R}^K_{++}$, such that

$$u^{t'} \leq u^t + \rho^t \cdot (x^{t'} - x^t) \ \text{ for all } \ t,\, t' = 1, 2, \ldots, T,$$

$$\rho^t_k = \lambda^t p^t_k \ \text{ (for any } \ x^t_k > 0) \ \text{ for all } \ k = 1, 2, \ldots, K,\ t = 1, 2, \ldots, T.$$

This implies that, at every observation $t = 1, 2, \ldots, T$, there must also exist numbers $u^t \in \mathbb{R}$, $\lambda^t \in \mathbb{R}_{++}$, and $\rho^t_k \in \mathbb{R}_{++}$ (for any $x^t_k = 0$), such that

$$u^{t'} \leq u^t + \lambda^t \sum_{x^t_k > 0} p^t_k (x^{t'}_k - x^t_k) + \sum_{x^t_k = 0} \rho^t_k (x^{t'}_k - x^t_k) \ \text{ for all } \ t,\, t' = 1, 2, \ldots, T.$$

For all $k = 1, 2, \dots, K$, $t = 1, 2, \dots, T$, let $\pi_k^t = p_k^t$ for any $x_k^t > 0$ and $\pi_k^t = \rho_k^t / \lambda^t$ for any $x_k^t = 0$. Notice that $\pi^t \in \mathbb{R}_{++}^K$. $\square$

A.3 Proof of Proposition 2

Proof. Given the data set $\mathcal{D}$, at every observation $t = 1, 2, \dots, T$, suppose that there exist numbers $u^t \in \mathbb{R}$ and $\lambda^t \in \mathbb{R}_{++}$, and vectors $\rho^t \in \mathbb{R}_{++}^K$, such that

$$u^{t'} \leq u^t + \rho^t \cdot (x^{t'} - x^t) \text{ for all } t, t' = 1, 2, \dots, T,$$

$$\rho_k^t = \lambda^t p_k^t \text{ (for any } x_k^t > 0) \text{ for all } k = 1, 2, \dots, K, \ t = 1, 2, \dots, T.$$

Let $v^t = u^t$, $\mu^t = \lambda^t$, and $\eta^t = \rho^t$ for all $t = 1, 2, \dots, T$. Given the data set $\mathcal{D}$, for some hypothetical price vector $p^0 \in \mathbb{R}_{++}^K$, choose any $x^0, y^0 \in \mathcal{S}(p^0 \,|\, \mathcal{D})$. (Notice that $\mathcal{S}(p^0 \,|\, \mathcal{D})$ is non-empty. Since there exist support prices $\pi^t \in \mathbb{R}_{++}^K$ (with $\pi_k^t = p_k^t$ for any $x_k^t > 0$) at every observation $t = 1, 2, \dots, T$, such that $\{(\pi_k^t, x_k^t)\}_{k=1,2,\dots,K}^{t=1,2,\dots,T}$ obeys GARP, there is a convex preference which rationalizes $\mathcal{D}$. In fact, an increasing, concave, and continuous utility function can be constructed from $\{(\pi_k^t, x_k^t)\}_{k=1,2,\dots,K}^{t=1,2,\dots,T}$. Maximizing this function by choosing $x^0 \in \mathbb{R}_+^K$ subject to $p^0 \cdot x^0 = 1$ implies that $\mathcal{S}(p^0 \,|\, \mathcal{D})$ is always non-empty.) First define u^0 and v^0 according to

$$u^0 = \min_t \{u^t + \rho^t \cdot (x^0 - x^t)\},$$

$$v^0 = \min_t \{v^t + \eta^t \cdot (y^0 - x^t)\},$$

next define λ^0 and μ^0 according to

$$\lambda^0 = \max \{1, \max_t \{(u^t - u^0)/p^0 \cdot (x^t - x^0) : p^0 \cdot (x^t - x^0) \neq 0\}\},$$

$$\mu^0 = \max \{1, \max_t \{(v^t - v^0)/p^0 \cdot (x^t - y^0) : p^0 \cdot (x^t - y^0) \neq 0\}\},$$

and lastly, define ρ^0 and η^0 according to

$$\rho^0 = \lambda^0 p^0,$$

$$\eta^0 = \mu^0 p^0.$$

Notice that $u^0, v^0 \in \mathbb{R}$, $\lambda^0, \mu^0 \in \mathbb{R}_{++}$, and $\rho^0, \eta^0 \in \mathbb{R}_{++}^K$. For all $t = 0, 1, \dots, T$, let $w^t = \alpha u^t + (1 - \alpha) v^t$, $\gamma^t = \alpha \lambda^t + (1 - \alpha) \mu^t$, and $\sigma^t = \alpha \rho^t + (1 - \alpha) \eta^t$ for some $\alpha \in [0, 1]$.

Notice that $u^t = v^t = w^t$, $\lambda^t = \mu^t = \gamma^t$, and $\rho^t = \eta^t = \sigma^t$ for all $t = 1, 2, \ldots, T$. Therefore, at every observation $t = 1, 2, \ldots, T$, there must exist numbers $w^t \in \mathbb{R}$ and $\gamma^t \in \mathbb{R}_{++}$, and vectors $\sigma^t \in \mathbb{R}^K_{++}$, such that

$$w^{t'} \leq w^t + \sigma^t \cdot (x^{t'} - x^t) \ \text{ for all } \ t, t' = 1, 2, \ldots, T,$$

$$\sigma^t_k = \gamma^t p^t_k \ \text{ (for any } \ x^t_k > 0) \ \text{ for all } \ k = 1, 2, \ldots, K, \ t = 1, 2, \ldots, T.$$

Consider two remaining sets of inequalities. In the first set of inequalities, there exist numbers u^t, $v^t \in \mathbb{R}$ for all $t = 0, 1, \ldots, T$, and vectors ρ^t, $\eta^t \in \mathbb{R}^K_{++}$ for all $t = 1, 2, \ldots, T$, such that

$$u^0 \leq u^t + \rho^t \cdot (x^0 - x^t) \ \text{ for all } \ t = 1, 2, \ldots, T,$$

$$v^0 \leq v^t + \eta^t \cdot (y^0 - x^t) \ \text{ for all } \ t = 1, 2, \ldots, T.$$

This is guaranteed by the definitions of u^0 and v^0. For some $\alpha \in [0, 1]$, taking a convex combination of the above inequalities, there exist numbers $w^t \in \mathbb{R}$ for all $t = 0, 1, \ldots, T$, and vectors $\sigma^t \in \mathbb{R}^K_{++}$ for all $t = 1, 2, \ldots, T$, such that

$$w^0 \leq w^t + \sigma^t \cdot ((\alpha x^0 + (1 - \alpha) y^0) - x^t) \ \text{ for all } \ t = 1, 2, \ldots, T.$$

In the second set of inequalities, there exist numbers u^t, $v^t \in \mathbb{R}$ for all $t = 0, 1, \ldots, T$, and numbers λ^0, $\mu^0 \in \mathbb{R}_{++}$, such that

$$u^t \leq u^0 + \lambda^0 p^0 \cdot (x^t - x^0) \ \text{ for all } \ t = 1, 2, \ldots, T,$$

$$v^t \leq v^0 + \mu^0 p^0 \cdot (x^t - y^0) \ \text{ for all } \ t = 1, 2, \ldots, T.$$

This is guaranteed by the definitions of λ^0 and μ^0. Since $\rho^0 = \lambda^0 p^0$, $\eta^0 = \mu^0 p^0$, and $p^0 \cdot x^0 = p^0 \cdot y^0 = 1$, there exist numbers u^t, $v^t \in \mathbb{R}$ for all $t = 0, 1, \ldots, T$, numbers λ^0, $\mu^0 \in \mathbb{R}_{++}$, and vectors ρ^0, $\eta^0 \in \mathbb{R}^K_{++}$, such that

$$u^t \leq u^0 + \rho^0 \cdot x^t - \lambda^0 \ \text{ for all } \ t = 1, 2, \ldots, T,$$

$$v^t \leq v^0 + \eta^0 \cdot x^t - \mu^0 \ \text{ for all } \ t = 1, 2, \ldots, T.$$

For some $\alpha \in [0, 1]$, taking a convex combination of the above inequalities, since $p^0 \cdot (\alpha x^0 + (1 - \alpha) y^0) = 1$, there exist numbers $w^t \in \mathbb{R}$ for all $t = 0, 1, \ldots, T$, a number $\gamma^0 \in \mathbb{R}_{++}$, and a vector $\sigma^0 \in \mathbb{R}^K_{++}$, such that

$$w^t \leq w^0 + \sigma^0 \cdot x^t - \gamma^0 p^0 \cdot (\alpha x^0 + (1 - \alpha) y^0) \ \text{ for all } \ t = 1, 2, \ldots, T.$$

Since $\rho^0 = \lambda^0 p^0$ and $\eta^0 = \mu^0 p^0$, then $\sigma^0 = \gamma^0 p^0$, and there exist numbers $w^t \in \mathbb{R}$ and vectors $\sigma^t \in \mathbb{R}^K_{++}$ for all $t = 0, 1, \ldots, T$, and a number $\gamma^0 \in \mathbb{R}_{++}$, such that

$$w^t \leq w^0 + \sigma^0 \cdot (x^t - (\alpha x^0 + (1-\alpha) y^0)) \text{ for all } t = 1, 2, \ldots, T,$$

with $\sigma^0 = \gamma^0 p^0$. Therefore, for some $\alpha \in [0,1]$, the consumption bundle $z^0 = \alpha x^0 + (1-\alpha) y^0$, a convex combination of x^0 and y^0, is also in the support set $\mathcal{S}(p^0 \,|\, \mathcal{D})$. $\square$

References

ALTONJI, J., and R. MATZKIN (2005): "Cross Section and Panel Data Estimators for Nonseparable Models with Endogenous Regressors," *Econometrica*, 73(4), 1053–1103.

AFRIAT, S. N. (1967): "The Construction of Utility Functions from Expenditure Data", *International Economic Review*, 8(1), 67–77.

ATKINSON, A. B., J. GOMULKA, and N. H. STERN (1990): "Spending on Alcohol: Evidence from the Family Expenditure Survey 1970–1983," *Economic Journal*, 100(402), 808–827.

BANKS, J., R. BLUNDELL, and A. LEWBEL (1997): "Quadratic Engel Curves and Consumer Demand," *Review of Economics and Statistics*, 79(4), 527–539.

BECKERT, W., and R. BLUNDELL (2008): "Heterogeneity and the Non-Parametric Analysis of Consumer Choice: Conditions for Invertibility," *Review of Economic Studies*, 75(4), 1069–1080.

BERRY, S., J. LEVINSOHN, and A. PAKES (1995): "Automobile Prices in Market Equilbrium," *Econometrica*, 63(4), 841-890.

BLOCK, H., and J. MARSCHAK (1960): "Random Orderings and Stochastic Theories of Response," in I. OKLIN, ed., *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling* (Stanford, CA: Stanford University Press).

BLUNDELL, R., D. KRISTENSEN, and R. MATZKIN (2014): "Bounding Quantile Demand Functions Using Revealed Preference Inequalities," *Journal of Econometrics*, 179(2), 112–127.

BLUNDELL, R., M. BROWNING, L. CHERCHYE, I. CRAWFORD, B. DE ROCK, and F. VERMEULEN (2015): "Sharp for SARP: Nonparametric Bounds on Counterfactual Demands," *American Economic Journal: Microeconomics*, 7(1), 43–60.

BLUNDELL, R., M. BROWNING, and I. CRAWFORD (2003): "Nonparametric Engel Curves and Revealed Preference," *Econometrica*, 71(1), 205–240.

BLUNDELL, R., M. BROWNING, and I. CRAWFORD (2008): "Best Nonparametric Bounds on Demand Responses," *Econometrica*, 76(6), 1227–1262.

BLUNDELL, R., M. BROWNING, and I. CRAWFORD (2007): "Improving Revealed Preference Bounds on Demand Responses," *International Economic Review*, 48(4), 1227–1244.

BLUNDELL, R., X. CHEN, and D. KRISTENSEN (2007): "Semi-Nonparametric IV Estimation of Shape-Invariant Engel Curves," *Econometrica*, 75(6), 1613–1669.

BROWN, B. W., and M. B. WALKER (1989): "The Random Utility Hypothesis and Inference in Demand Systems," *Econometrica*, 57(4), 815–829.

BROWNING, M. (1989): "A Nonparametric Test of the Life-Cycle Rational Expectations Hypothesis," *International Economic Review*, 30(4), 979–992.

CHERNOZHUKOV, V., G. W. IMBENS, and W. K. NEWEY (2007): "Instrumental Variable Estimation of Nonseparable Models," *Journal of Econometrics*, 139(1), 4–14.

CHERNOZHUKOV, V., H. HONG, and E. TAMER (2007): "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica*, 75(5), 1243–1284.

CHERNOZHUKOV, V., S. LEE, and A. M. ROSEN (2013): "Intersection Bounds: Estimation and Inference," *Econometrica*, 81(2), 667–737.

CHESHER, A. (2003): "Identification in Nonseparable Models," *Econometrica*, 71(5), 1405–1441.

CRAWFORD, I. (2010): "Habits Revealed", *Review of Economic Studies*, 77(4), 1382–1402.

DEATON, A., and J. MUELLBAUER (1980a): "An Almost Ideal Demand System," *American Economic Review*, 70(3), 312–336.

DEATON, A., and J. MUELLBAEUR (1980b): *Economics and Consumer Behavior* (Cambridge, Cambridge University Press).

DEATON, A., and M. IRISH (1984): "Statistical Models for Zero Expenditures in Household Budgets," *Journal of Public Economics*, 23(1–2), 59–80.

DEMUYNCK , T., and C. SEEL (2014): "Revealed Preference with Limited Consideration," mimeo.

DEMUYNCK, T., and E. VERRIEST (2013): "I'll Never Forget My First Cigarette: A Revealed Preference Analysis of the 'Habits As Durables' Model," *International Economic Review*, 54(2), 717–738.

DIEWERT, W. E. (1973): "Afriat and Revealed Preference Theory," *Review of Economic Studies*, 40(3), 419–425.

FLEISSIG, A. R., and G. WHITNEY (2011): "A Revealed Preference Test of Rationing," *Economic Letters*, 113(3), 234-236.

GEARY, R. C. (1950): "A Note on 'A Constant-Utility Index of the Cost of Living'," *Review of Economic Studies*, 18 (1), 65–66.

HAAG, B. R., S. HODERLEIN, and K. PENDAKUR (2009): "Testing and Imposing Slutsky Symmetry in Nonparametric Demand Systems," *Journal of Econometrics*, 153(1), 33–50.

HAUSMAN, J. A. (1997): "Valuation of New Goods under Perfect and Imperfect Competition," in *The Economics of New Goods*, ed. T. BRESNAHAN and R. GORDON. Chicago, University of Chicago Press.

HAUSMAN, J. A., and W. K. NEWEY (1995): "Nonparametric Estimation of Exact Consumer Surplus and Deadweight Loss," *Econometrica*, 63(6), 1445–1476.

HECKMAN, J. T. (1979): “Sample Selection Bias as a Specification Error,” *Econometrica*, 47(1), 153–161.

HICKS, J. R. (1940), “The Valuation of the Social Income,” *Economica*, 7(26), 105–124.

IMBENS, G. W., and W. K. NEWEY (2009): “Identification and Estimation of Triangular Simultaneous Equations Models without Additivity,” *Econometrica*, 77(5), 1481–1512.

HODERLEIN, S. (2011): “How Many Consumers are Rational?”, *Journal of Econometrics*, 164(2), 294–309.

HODERLEIN, S., and J. STOYE (2014): “Revealed Preferences in a Heterogeneous Population,” *Review of Economic Statistics*, 197–213.

HODERLEIN, S., and J. STOYE (2015): “Testing Stochastic Rationality and Predicting Stochastic Demand: The Case of Two Goods,” *Economic Theory Bulletin*, forthcoming.

HOUTHAKKER, H. S., and J. TOBIN (1952): “Estimates of the Free Demand for Rationed Foodstuffs”, *Economic Journal*, 62(245), 103–118.

IMBENS, G. W., and C. F. MANSKI (2004): “Confidence Intervals for Partially Identified Parameters,” *Econometrica*, 72(6), 1845–1857.

JORGENSON, D. W., L. J. LAU, and T. M. STOKER (1982): “The Transcendental Logarithmic Model of Aggregate Consumer Behavior,” in R. BASMANN and G. RHODES, eds., *Advances in Econometrics*, Volume 1 (Greenwich, CT: JAI Press).

KEEN, M. (1986): “Zero Expenditures and the Estimation of Engel Curves,” *Journal of Applied Econometrics*, 1(3), 277–286.

KITAMURA, Y., and J. STOYE (2013): “Nonparametric Analysis of Random Utility Models: Testing,” CeMMAP Working Paper CWP36/13.

KLEIN, L. R., and H. RUBIN (1947–1948): “A Constant-Utility Index of the Cost of Living,” *Review of Economic Studies*, 15 (2), 84–87.

LEWBEL, A. (1991): “The Rank of Demand Systems: Theory and Nonparametric Estimation,” *Econometrica*, 59(3), 711–730.

LEWBEL, A. (1995): “Consistent Nonparametric Hypothesis Tests with an Application to Slutsky Symmetry,” *Journal of Econometrics*, 67(2), 379–401.

LEWBEL, A., and K. PENDAKUR (2009): “Tricks with Hicks: The EASI Demand System,” *American Economic Review*, 99(3), 827–863.

MANSKI, C. F. (2003): *Partial Identification of Probability Distributions.* New York, Springer-Verlag.

MANSKI, C. F. (2007): “Partial Identification of Counterfactual Probabilities,” *International Economic Review*, 48(4), 1393–1410.

MATZKIN, R. L. (2003): “Nonparametric Estimation of Nonadditive Random Functions,” *Econometrica*, 71(5), 1339–1375.

MATZKIN, R. L. (2008): “Identification in Nonparametric Simultaneous Equations Models,” *Econometrica*, 76(5), 945–978.

McFADDEN, D., and M. RICHTER (1991): "Stochastic Rationality and Revealed Stochastic Preference," in J. CHIPMAN, D. McFADDEN, and M. RICHTER, eds., *Preferences, Uncertainty and Optimality: Essays in Honour of Leo Hurwicz* (Boulder, CO: Westview Press).

MEGHIR, C., and J.-M. ROBIN (1992): "Frequency of Purchase and the Estimation of Demand Systems," *Journal of Econometrics*, 53(1–3), 53–85.

NEARY, J. P., and K. W. S. ROBERTS (1980): "The Theory of Household Behaviour under Rationing," *European Economic Review*, 13(1), 25–42.

NEWEY, W. K., and J. L. POWELL (2003): "Instrumental Variables Estimation for Nonparametric Models," *Econometrica*, 71(5), 1565–1578.

NEWEY, W. K., J. L. POWELL, and F. VELLA (1999): "Nonparametric Estimation of Triangular Simultaneous Equations Models," *Econometrica*, 67(3),565–604.

POLISSON, M., and J. K.-H. QUAH (2013): "Revealed Preference in a Discrete Consumption Space," *American Economic Journal: Microeconomics*, 5(1), 28–34.

ROTHBARTH, E. (1941): "The Measurement of Changes in Real Income under Conditions of Rationing," *Review of Economic Studies*, 8(2), 100–107.

SAMUELSON, P. A. (1947–1948): "Some Implications of 'Linearity'," *Review of Economic Studies*, 15 (2), 88–90.

STONE, R. (1954): "Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand," *Economic Journal*, 64 (255), 511–527.

VARIAN, H. R. (1982): "The Nonparametric Approach to Demand Analysis," *Econometrica*, 50(4), 945–973.

VARIAN, H. R. (1983): "Non-Parametric Tests of Consumer Behaviour," *Review of Economic Studies*, 50(1), 99–110.

VARIAN, H. R. (1988): "Revealed Preference with a Subset of Goods," *Journal of Economic Theory*, 46(1), 179–185.