UNIVERSITY OF LEICESTER

School of Business
Economics Division

Demand Analysis with Partially Observed Prices

Ian Crawford, University of Oxford

Matthew Polisson, University of Leicester

Working Paper No. 15/12

Demand Analysis with Partially Observed Prices

Ian Crawford and Matthew Polisson*

December 14, 2016

Abstract: In empirical demand, industrial organization, and labor economics, prices are often unobserved since they may only be recorded when an agent transacts. 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 and UC Berkeley, and he would like to thank these institutions 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, Krishna Pendakur, and conference participants at the 2014 Workshop on Nonparametric Demand (IFS and CeMMAP) for helpful contributions.

1. Introduction

It is extremely common for consumer microdata to contain incomplete observations on prices since prices are typically only observed when a transaction occurs.[1] The combination of missing prices and zero demands is often a feature of many observational data sets and presents numerous challenges to empirical work. These problems are usually amplified when the data involved are high dimensional, and as a consequence, when a large number of zero demands are observed simultaneously.

To give a motivating example from a relatively new source of data that is potentially extremely valuable but which suffers from missing prices, consider electronically gathered consumer panel data (sometimes known 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 (SKU) barcode. The key features of such data sets are that (i) the number of products is typically very large, (ii) many goods are only available in discrete amounts, (iii) there are many instances of zero demands, and (iv) prices are only recorded when a consumer makes a purchase. As a consequence, while quantities and expenditures are fully observed for every item (where both are equal to zero when a product is not purchased), 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,[2] 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. The problem of zero purchases and partially observed prices is therefore extremely prominent in these new rich scanner data sets.

The fundamental problem stemming from partially observed prices is that, as an observer, one is never able to fully infer the consumer's choice set, i.e., the analyst has incomplete information about alternatives which an individual might have selected but did not purchase.

[1] As an example, the wage of a worker is only observed when that individual is employed.

[2] The Kantar Worldpanel is a representative rolling panel of British households who scan all home food purchases, much like Nielsen home scanner data in the US.

This makes it extremely difficult to infer the decision maker's preferences from observed choice behavior.[3] It would appear that strong economic assumptions plus advanced econometric machinery would be necessary in order to make any progress when zero demands and partially observed prices are a significant feature of the data. The objective of this paper is to show that this is not necessarily the case.

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 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. Perhaps somewhat surprisingly, we are able to do so using classical revealed preference arguments in consumer choice. At first glance, it would seem that the revealed preference methods of Samuelson (1938, 1948), Afriat (1967), Diewert (1973), and Varian (1982) could have no great purchase on a problem with partially observed prices, since the arguments themselves are normally constructed from fully observed price and consumption vectors. However, we show this not to be the case. We also provide a set of examples which demonstrate how to apply our results in order to test microdata for consistency with the maximization hypothesis, to make demand predictions out-of-sample, and to perform welfare analysis.

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 $x^t = (x_1^t, x_2^t, \ldots, x_K^t) \geq 0$. Note that while we assume, for expositional simplicity, that the consumption set is classical, i.e., there are a finite number of infinitely divisible goods that can be consumed in continuous amounts, the framework and the results to follow can accommodate a coarse granularity or discreteness

[3] Manski (2003) studies a number of statistical identification problems that relate to partial 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 important papers (e.g., Heckman (1979), Wales and Woodland (1983), Deaton and Irish (1984), Keen (1986), Lee and Pitt (1986), Atkinson, Gomulka, and Stern (1990), and Meghir and Robin (1992)) study the problems of estimating demand and/or recovering consumer preferences in the presence of data which feature missing prices.

in the consumption space without any difficulty or significant modification whatsoever.[4] 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 then 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 contains no information about $p_k^t \,|\, x_k^t = 0$.

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.

3. Treatments for Missing Prices

In labor economics, where missing prices are the wages of workers who are not in employment, the conventional assumption is often that labor force participation does not occur

[4] For data sets in which zero demands and partially observed prices are a significant feature, it is also the case that many goods are only available in discrete amounts. To accommodate this without complicating the exposition, we make use of the results in Polisson and Quah (2013), which state that if a consumer's preference over a set of discrete goods is weakly separable from at least one infinitely divisible outside good, then as long as overall utility is strictly increasing in such an outside good, discreteness adds no complications to the usual analysis, i.e., price and demand observations of the discrete goods obey the Generalized Axiom of Revealed Preference (GARP) if and only if the consumer is maximizing utility. In our setting, an infinitely divisible outside good is simply any money that remains if a consumer fails to exhaust her budget on discrete consumption. Therefore, since GARP is necessary and sufficient for utility maximization both in continuous and discrete consumption spaces (subject to the preceding argument), and since the use of money as an outside option is uncontroversial in the current empirical practice surrounding discrete choice models, we make use of this result in order to simplify the exposition and make the heuristic assumption that $x_k^t \in \mathbb{R}_+$.

at random, nor is it the case that work is unavailable, but rather that the wage offer is lower than the individual's reservation wage. The estimation of labor supply models therefore needs to take into account choices on both extensive and intensive margins with this choice-driven selection problem in mind.

In the context of demand estimation, missing prices usually occur across many goods simultaneously. Zero purchases might arise for several reasons; it may be, as in the labor supply case, that the offer price is above the individual's reservation price; it may be that the product is unavailable; it may be a consequence of a form of measurement error typically described as the 'infrequency of purchase' problem. Since infrequency of purchase is usually assumed to be a measurement error issue, it is commonly dealt with by adopting an instrumental variables approach.[5] Zero demands that are a consequence of rationing have, to our knowledge, only been studied in the context of a single good. In such a case, the demand for this good is estimated using any non-zero observations and the support price inferred via some form of extrapolation.[6] Zero purchases which are a consequence of the offer price being above the individual's reservation price can be dealt with using the methods in Wales and Woodland (1983) or Lee and Pitt (1986).

However, it would seem that the latter two approaches to this particular selection problem are rarely, if ever, applied in empirical practice. Wales and Woodland (1983) make use of the direct utility function and the Kuhn-Tucker conditions of the corresponding utility maximization problem to construct a likelihood function that is composed of contributions from the densities for every possible consumption pattern. For data sets with K goods, N of which are consumed, there are $\sum_{N=1}^{K} \frac{K!}{N!(K-N)!}$ possible consumption patterns and corresponding densities to incorporate. The Lee and Pitt (1986) approach exploits the dual representation of preferences (and therefore can make use of a richer set of available parametric demand systems) and virtual prices that can be estimated as part of the problem. It is important to note that both approaches depend heavily on functional form assumptions over preferences. More problematically, however, both methods also seem to be extremely difficult to apply in empirical practice due to the curse of dimensionality—in the former case, the analysis is complicated by the number of possible consumption patterns; in the

[5] See Keen (1986) and Atkinson, Gomulka, and Stern (1990).
[6] See Hausman, Leonard, and Zona (1994) and Hausman (1997).

latter case, the estimation of many virtual prices for goods not purchased simultaneously leads to a complex censoring problem, requiring the solution of a large number of multivariate probability integrals when evaluating the likelihood function. As a result of these difficulties, rigorous approaches to the analysis of corner solutions do not appear to have been taken up with much enthusiasm by applied researchers. Instead, practitioners tend to favor more ad hoc strategies. Broadly speaking, there are two types: aggregation and/or imputation.

The simple aggregation approach is to combine purchases across goods and observations, thereby reducing the dimensions of the demand system, until zero demands are no longer a significant problem. The difficulty with this approach is that the prices needed to form a proper (by this, we mean theory-consistent) price index for the group are not observed. Instead, an aggregate unit value (group total expenditure divided by group total quantity) is used as an implicit price index, but this leads to another problem—the 'price' of a group constructed in this way depends on the consumption pattern within that group, and thus apparent price variation in the group price index is partly a function of preference heterogeneity, i.e., two consumers facing identical prices but who have difference tastes would be measured as having different price indices at the group level. Deaton (1987, 1988) and Crawford, Laisney, and Preston (2003) discuss some consequences of this problem for demand system estimation. This approach also obscures the fine detail on spending patterns, which is one of the features that makes scanner data so attractive in the first place.

The imputation approach is to say that for an individual consumer who is missing a price for a given product/location/time, we are free to search within the broader data set for another consumer observed to have purchased the same product, in the same location, at the same time, and then to impute, i.e., replace, the missing price with this transaction price that has been observed for that other consumer. In high dimensional consumer panel data, another consequence of the curse of dimensionality is that accurately defined product/location/time cells are very likely to be empty,[7] so the definitions of these cells tend to be non-local in practice.[8] Even if this is not the case, and even if the imputation is accurate,

[7] Note that national/regional pricing schemes in supermarkets may facilitate price matching across a wider range of locations.

[8] Consider, for example, 1,000 observations scattered uniform-randomly in a 10-dimensional unit hypercube $[0, 1]^{10}$. The probability of an observation occurring in an neighborhood described by a hypercube with edges of length 0.2 is 0.2^{10}, so the expected number of observations in such a neighborhood is 0.0001024;

simply imputing a missing price in this way and then estimating a regular demand system as normal is a dubious manner in which to proceed. At their heart, demand systems identify preferences because relative prices are equal to the local marginal rates of substitution; sufficient variation in relative prices and income therefore allows the econometrician to trace out the shape of indifference curve maps. But this is only true for interior solutions as Wales and Woodford (1983) and Lee and Pitt (1986) make clear. At corners, relative prices only reveal the local rates of substitution if the consumer is just indifferent between purchasing and not purchasing, and the econometrician simply cannot assume this.

Therefore, as a summary and as argued by Meghir and Robin (1992),[9] the appropriate solution to missing prices in consumption data generally depends on the reason for having observed zero demands. Different reasons require different solutions by-and-large, and although different conventional assumptions are made across these contexts, it is often in reality difficult to know why a consumer did not make a purchase. Somewhat more optimistically, we show in the next section 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.

4. Revealed Preference

We first appeal to some familiar arguments in order to build intuition before establishing the formal results. Suppose that the true data-generating process for the data described in the previous section involves the maximization of a stable utility function subject to a sequence of linear budgets (each determined by exogenously given prices and income) and rationing constraints.[10] At every observation t, the consumer's constrained optimization

one would need approaching 10 million observations before one could expect to see a single observation in such a neighborhood.

[9] See Meghir and Robin (1992), pp. 54–55.

[10] In this paper, we restrict our attention to a static model. It would be straightforward to extend the ideas developed here 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.

problem is therefore given by

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

where the utility function u is increasing, concave, and continuous, and where $e^t \in \mathbb{R}_{++}$ 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 demand is a choice or due to rationing.

If we further assume differentiability of the utility function,[11] then for every good k and at every observation t, the consumer's first-order conditions are given according to

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

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, and where the multiplier $\lambda^t \in \mathbb{R}_{++}$ is the marginal utility of income at observation t, and the multiplier $\mu^t_k \in \mathbb{R}_{++}$ 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.[12] 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.[13] 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

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

[12] 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.

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

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 $\pi_k^t \in \mathbb{R}_{++}$, we have

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

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.

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

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* $\pi_k^t = p_k^t$ *for any* $x_k^t > 0$*) 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 simply states that in order to rationalize the observed behavior, there must exist a utility function *and* corresponding support 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*

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

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

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 within the revealed preference literature, and more broadly, as a point of departure for treating missing prices in high dimensional consumer panel data.

(i) The equivalence between statements (1) and (2) implies that 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 potentially a mixture of observed prices, reservation prices, and virtual prices. Notice that reservation prices and virtual prices can actually 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 demand, we are unable to empirically identify reservation prices and virtual prices.[14] Nonetheless, as Proposition 1 indicates, and as the examples in the next section illustrate, it is not necessary to draw such a distinction.

[14] The support price of good k at observation t can be recovered according to $\pi^t_k = \rho^t_k / \lambda^t$. Notice that $\pi^t_k = p^t_k$ when prices are observed, and that we can construct $\pi^t_k = u_k(x^t)/\lambda^t$ or $\pi^t_k = p^t_k + \mu^t_k / \lambda^t$ when prices are unobserved or unobservable.

(iv) Products that are never purchased can be excluded from the rationalizability analysis entirely since they provide no further restrictions on the data.[15]

(v) In many of the data sets where zero purchases and their corresponding missing prices might arise, it is also the case that many goods are only available in discrete amounts. As we noted earlier in the paper, Polisson and Quah (2013) develop a revealed preference approach in a discrete consumption space. In this classical setting, the availability of some goods in continuous amounts or an unobserved outside good (typically money) available in continuous amounts is enough to ensure that discreteness adds no important complications to the usual arguments; the same applies in a setting in which prices are only partially observed.

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

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

(viii) Blow, Browning, and Crawford (2008) provide a revealed preference analysis of the characteristics model, a framework which was initially put forward by Gorman (1956) and Lancaster (1966). In this context, Blow, Browning, and Crawford (2008) note that when allowing for missing prices, any unobserved price can be set arbitrarily high, which reduces the number of conditions on the data that need to be checked. In the special case where each good is a characteristic, conditional (L) of Theorem 2 in Blow, Browning, and Crawford (2008) then coincides with statement (2) of Proposition 1. What is important to note here is that Proposition 1 is a distinctly more general result since it establishes the *complete* analog to the original Afriat (1967) theorem (minus GARP) allowing for missing prices, i.e., an equivalence between weaker and stronger notions of rationalizability, meaning that monotonicity, concavity, and continuity of

[15] This claim may no longer hold if we are interested in making demand predictions or conducting welfare analysis at hypothetical budgets. See the following section.

the utility function can be assumed without cost in such environment. While this is not particularly difficult to show, it is important to highlight. As a consequence of these flexible features, we are able to carry out the full extent of demand analysis.

(ix) Proposition 1 can be used to make demand predictions and conduct welfare analysis following an approach first developed by Varian (1982), and later extended by Blundell, Browning, and Crawford (2003, 2007, 2008) and Blundell *et al.* (2015).

To elaborate on this last point, 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 extremely important property both for describing bounds on demand responses and making welfare comparisons (see, e.g., Blundell, Browning, and Crawford (2003, 2007, 2008) and Blundell *et al.* (2015)). It is therefore useful and significant that this property is preserved in a situation where prices are only partially observed. Notice that if all prices are fully observed and we have access to the complete data set $\mathcal{O} = \{(p_k^t, x_k^t)\}_{k=1,2,\ldots,K}^{t=1,2,\ldots,T}$, i.e, any zero demands are known to be corner solutions, then 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., observed only when the consumer transacts. These restrictions allow choice behavior to be examined for consistency with utility maximization without imposing heavy economic assumptions and/or importing complicated econometric machinery, and without aggregating/imputing missing prices. Subject to these conditions being satisfied, tractable procedures are available to provide bounds on demand forecasts and welfare measures.

5. Examples

In this section, we provide a number of simple examples to illustrate how our main results can be used to carry out all of the principal tasks of empirical work in consumer demand, i.e., testing utility maximization, predicting new demands, and performing welfare analysis, even in the presence of partially observed prices. We have confined ourselves in all examples to low dimensional problems, i.e., relatively few products and observations, and a large proportion of missing prices/zero demands for a number of important reasons.

Firstly, if observable choice behavior is not rationalizable, then a low dimensional environment makes it easier to see what, in the data themselves, is giving rise to such a result (as we shall see). With many goods and observations, this is likely to become much more opaque. Secondly, the most important differences, both theoretically and computationally, between an environment with K as opposed to $K + 1$ goods arise when moving from two to three goods. This is because the Weak Axiom of Revealed Preference (WARP) is a necessary and sufficient condition for rationalizability with only two goods, and WARP does not involve the transitivity of preferences. Any violation of WARP via an intransitivity in a two-good environment is necessarily also a direct violation of WARP, and so transitivity adds no empirical content in a two-good environment. In a three-good environment it adds a very great deal (see, e.g., Kihlstrom, Mas-Colell, and Sonnenschein (1976)). Therefore, in terms of the empirical demands of consumer theory as well as its computational complexity, moving from three to four goods, or from three to forty goods, adds little of substance. Thirdly, the choice of relatively few observations in each example is again purposeful, since it is well known that the ability of revealed preference techniques to detect violations is weakly increasing in the number of observations (see, e.g., Blundell, Browning, and Crawford (2003, 2007, 2008)). In an environment with missing data and few observations, it might be easy to assume that revealed preference techniques would have little or no ability to restrict consumer behavior. Our choice of examples is designed to show that this is not the case.

We begin by illustrating how our results can be used to determine whether a data set with partially observed prices can be rationalized by utility maximization, and we consider a data set with 3 goods and 5 observations on a single consumer, 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.

In this data set, we have only a few observations, and one third of the prices are missing. Yet 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 utility function. And it is relatively easy to see why.

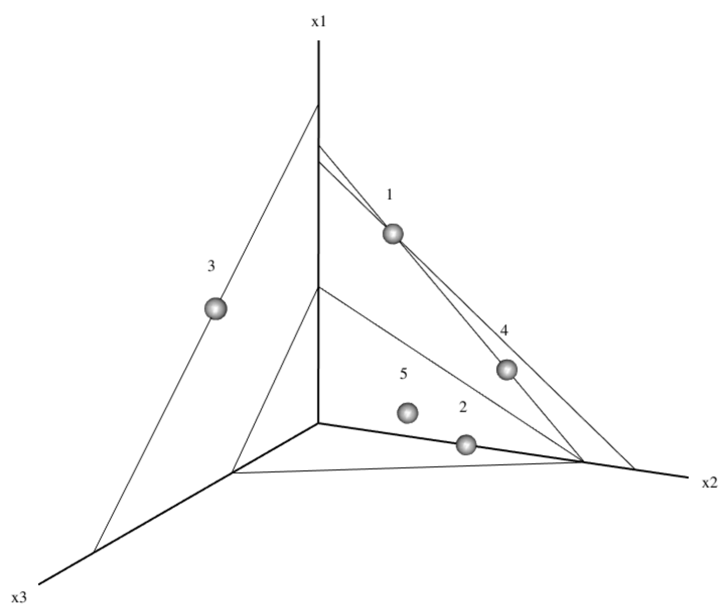

Figure 1: Nonrationalizable Choices

Figure 1 depicts the data. The budgets for all but observation 5 are partially observed. The observations of interest are 1 and 4, in which good 3 has zero demand but in which there are positive demands for goods 1 and 2. If we focus on these, then we have

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

Products which are never purchased can be excluded since they add no restrictions on consumer behavior, and so we are simply 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 exactly affordable, but $(5,2)$ is purchased when $(2,5)$ costs strictly less. Another way to understand this violation is that whatever were the prices of the third good at these two observations, the demands are both zero, and therefore the third good is irrelevant when assessing whether or not one bundle was affordable when the other was chosen. This is illustrated plainly in Figure 1—observations 1 and 4 clearly violate the revealed preference restrictions regardless of the price of the third good.

As an example of a data set which is rationalizable, consider the following:

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

Here we have as many goods and observations as in the first example, which was not rationalizable, and one less missing price. In such a situation, one might expect it to be easier to reject utility maximization since we have access to more information on prices, but the data are in fact rationalizable. To illustrate that this is not a vacuous achievement, i.e., an 'anything goes' result, note that we are able to distinguish between support prices which are theory-consistent versus theory-inconsistent. Consider the price matrices

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

The set of demands given above is rationalizable by Π (which is therefore a valid set of support prices) but not Π' (which is invalid).

What this example also illustrates is that our conditions can assist researchers in imputing prices. It is straightforward to assess whether imputed prices are theory-consistent, and if not, then it is possible for a researcher to exploit our conditions in order to choose an alternative but theory-consistent set of imputed prices that are close (in some suitable metric) to the initial imputation. Essentially this would amount to a minimum distance perturbation to the imputed prices such that they lie in the set of valid support prices.

5.2 Demand Predictions

The next question one might ask of a data set that is rationalizable is how to exploit the economic theory in order 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 (see, e.g., Blundell, Browning, and Crawford (2003, 2007, 2008) and Blundell *et al.* (2015)). Essentially, one can appeal to the structure of the model of utility maximization in order to construct any conceivable counterfactuals of interest.

Recall our earlier example which was shown to be rationalizable. Suppose that we are interested in bounding demands 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. This is simply the set of forecast demands such that the observed data and the forecast data satisfy our conditions when pooled together.

With three goods, support sets can be conveniently depicted in two dimensions as budget shares (w_k) on the unit simplex as shown in Figure 2.[16] The shaded areas (both light and dark) together represent the set of forecast budget shares consistent with our conditions and the data, i.e. $\mathcal{S}(p^0 \mid \{(P, X)\})$, the set of demands that we are able to predict using our five

[16] The orientation is as follows: 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.

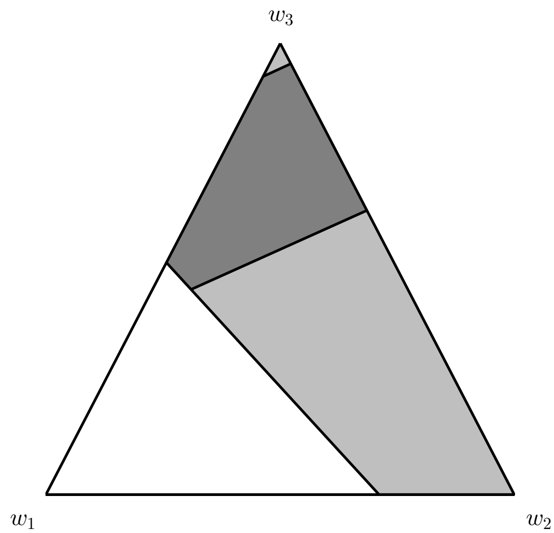

Figure 2: Support Sets on the Budget Simplex

partial observations. The set is large relative to the outcome space, but recall that we only have five observations on three goods, and many of the price observations are missing.

Of course, we could make tighter predictions if we actually knew the prices or had reliable and theory-consisted imputed values. To illustrate the cost of the missing information on the precision of out-of-sample predictions, we have drawn a darker shaded region using the imputed prices Π from the example in section 5.1 above. This corresponds to the the support set $\mathcal{S}(p^0 \,|\, \{(\Pi, X)\})$, which is computed as in Varian (1982) using the theory-consistent Π. In both cases the set-valued predictions are 'sharp' in the sense that they exhaust the empirical content of both the theory (maximizing behavior) and the data, and the benefit in terms of precision of 'knowing' true prices is clearly illustrated by the manifest shrinking of the support set. Note that any set of imputed prices where the imputation is done correctly, i.e., in a way that is consistent with both theory and existing data, necessarily results in a support set of demand predictions which is a subset of that which can be achieved when prices are missing.

5.3 Welfare Analysis

Given a data set which satisfies the conditions that we require for basic rationalizability, and a forecast of demands for budgets which have not previously been observed, the next question typically involves some notion of welfare analysis. The typical welfare question asks what is the welfare cost/benefit to the consumer of facing some observed set of prices compared to some counterfactual alternative. For example, this counterfactual might be the result of a tax or benefit change.

Revealed preference methods allow us to compute bounds on money metric welfare measures, such as compensating variation (CV) and equivalent variation (EV), by putting bounds on the relevant reference indifference curves and then establishing the minimum cost of reaching these indifference curves using either initial or counterfactual prices. This is described in the context of fully observed prices and known demands in Varian (1982). Blundell *et al.* (2015) show how to adapt this method to accommodate that fact that $\{(p^*, e^*)\}$ is a counterfactual budget, rather than an observed budget, which implies that the demand forecast is bounded rather than known. In our missing prices example which was already shown to be rationalizable, using the results from this paper, it is straightforward to bound the CV and EV associated with a change in prices and income from, e.g., $\{(p^2, e^2)\}$ to $\{(p^*, e^*)\}$, by simply modifying the approach of Varian (1982) and Blundell *et al* (2015). The welfare bounds in this instance are given by

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

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

These bounds both show that the consumer's welfare is higher under the counterfactual regime than the initial regime, i.e., if the new regime were the result of a tax reform, then this individual would be better off by a monetary equivalent value which lies within these bounds. The CV and EV bounds do overlap, but it is clearly possible for them to differ since preferences may well not be homothetic or quasilinear. Revealed preference tests of homotheticity and quasilinearity both exist (see Varian (1983) and Brown and Calsamiglia (2007)). These restrictions could therefore be appended to our results in order to investigate partially observed prices in these special cases.

6. Conclusions

Consumer panel data 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. But these data sets and many others also suffer from the pervasive problem of partially observed prices. At present, one might argue that most applied research treats this problem in a manner that is almost certain to be ad hoc, and in many instances, altogether unsystematic. 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 demonstrated how they might be used constructively to make counterfactual demand predictions and to perform welfare analysis.

Appendix

A.1 Preliminaries

Let $\mathcal{O} = \{(p_k^t, x_k^t)\}_{k=1,2,\ldots,K}^{t=1,2,\ldots,T}$ be a set of observations drawn from a consumer. Each observation consists of a price vector $p^t = (p_1^t, p_2^t, \ldots, p_K^t) \gg 0$ and a corresponding consumption bundle $x^t = (x_1^t, x_2^t, \ldots, x_K^t) \geq 0$. 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 $\mathcal{D} = \{(p_k^t \,|\, x_k^t > 0,\ x_k^t)\}_{k=1,2,\ldots,K}^{t=1,2,\ldots,T}$ 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 $\pi_k^t = p_k^t$ for any $x_k^t > 0$), such that $\{(\pi_k^t, x_k^t)\}_{k=1,2,\ldots,K}^{t=1,2,\ldots,T}$ 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 $\pi_k^t = p_k^t$ for any $x_k^t > 0$), and numbers $u^t \in \mathbb{R}$ and $\lambda^t \in \mathbb{R}_{++}$, such that

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

(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 $k = 1, 2, \ldots, K,\ t = 1, 2, \ldots, T$, 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 $t = 1, 2, \ldots, T$, 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

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

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

Let $v^t = u^t$, $\mu^t = \lambda^t$, and $\eta^t = \rho^t$ for all $t = 1, 2, \ldots, T$. 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 $t = 1, 2, \ldots, T$, such that $\{(\pi_k^t, x_k^t)\}_{k=1,2,\ldots,K}^{t=1,2,\ldots,T}$ obeys GARP, there is a convex preference which rationalizes $\mathcal{D}$. In fact, an increasing, concave, and continuous utility function can be constructed from $\{(\pi_k^t, x_k^t)\}_{k=1,2,\ldots,K}^{t=1,2,\ldots,T}$. 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, $v^0 \in \mathbb{R}$, λ^0, $\mu^0 \in \mathbb{R}_{++}$, and ρ^0, $\eta^0 \in \mathbb{R}_{++}^K$. For all $t = 0, 1, \ldots, T$, 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 $\mathcal{S}(p^0 \mid \mathcal{D})$. $\square$

References

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.

BLOW L., M. BROWNING, and I. CRAWFORD (2008): "Revealed Preference Analysis of Characteristics Models," *Review of Economic Studies*, 75(2), 371–389.

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 (2007): "Improving Revealed Preference Bounds on Demand Responses," *International Economic Review*, 48(4), 1227–1244.

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

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.

BROWN, D. J., and C. CALSAMIGLIA (2007): "The Nonparametric Approach to Applied Welfare Analysis," *Economic Theory*, 31(1), 183–188.

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

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

CRAWFORD, I., F. LAISNEY, and I. PRESTON (2003): "Estimation of Household Demand Systems with Theoretically Compatible Engel Curves and Unit Value Specifications," *Journal of Econometrics*, 114(2), 221–241.

DEATON, A. (1987): "Estimation of Own- and Cross-Price Elasticities from Household Survey Data," *Journal of Econometrics*, 36(1–2), 7–30.

DEATON, A. (1988): "Quality, Quantity, and Spatial Variation of Price," *American Economic Review*, 78(3), 418–430.

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," unpublished paper.

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.

GORMAN, W. M. (1980): "A Possible Procedure for Analysing Quality Differentials in the Egg Market," *Review of Economic Studies*, 47(5), 843–856. [First circulated in 1956].

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

HAUSMAN, J., G. LEONARD, and J. D. ZONA (1994): "Competitive Analysis with Differentiated Products," *Annales D'Économie et de Statistique*, 34, 159–180.

HECKMAN, J. J. (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.

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

KIHLSTROM, R., A. MAS-COLELL, and H. SONNENSCHEIN (1976): "The Demand Theory of the Weak Axiom of Revealed Preference," *Econometrica*, 44(5), 971–978.

LANCASTER, K. J. (1966): "A New Approach to Consumer Theory," *Journal of Political Economy*, 74(2), 132–157.

LEE, L.-F., and M. M. PITT (1986): "Microeconometric Demand System with Binding Nonnegativity Constraints: The Dual Approach," *Econometrica*, 54(5), 1237–1242.

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

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.

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. (1938): “A Note on the Pure Theory of Consumer’s Behaviour,” *Economica*, 5(17), 61–71.

SAMUELSON, P. A. (1948): “Consumption Theory in Terms of Revealed Preference,” *Economica*, 15(60), 243–253.

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.

WALES, T. J., and A. D. WOODLAND (1983): “Estimation of Consumer Demand Systems with Binding Non-Negativity Constraints,” *Journal of Econometrics*, 21(3), 263–285.