Testing Stochastic Rationality and Predicting Stochastic Demand: The Case of Two Goods

Stefan Hoderlein*
Boston College

Jörg Stoye†
Cornell University

October 3, 2014

Abstract

This paper precisely delineates the empirical content of consumer rationality in the following setting: Data are from a repeated cross-section; unobserved heterogeneity is completely unrestricted; however, there are only two goods. Simple closed-form expressions determine whether (population level) data are consistent with these assumptions. Bounds on counterfactual distributions of demand, and parameters thereof, follow.

A striking finding is that any rationalizable collection of cross-sectional distributions can be rationalized by pretending that the ordering of individual consumers on the budget line is maintained across budgets. Hence, this seemingly strong assumption does not tighten the bounds.

Keywords: revealed preference, stochastic demand, random utility models, integrability of demand.

JEL codes: C14, D11.

*Stefan Hoderlein, Department of Economics, Boston College, 140 Commonwealth Avenue, Chestnut Hill, MA 02467, USA, email: stefan_hoderlein@yahoo.com. First version, December 2011.

†Jörg Stoye, Department of Economics, Cornell University, Uris Hall, Ithaca, NY 14850, stoye@cornell.edu. Stoye gratefully acknowledges support through NSF grant SES-1260980. We thank Chuck Manski and two anonymous referees for helpful comments but assume responsibility for any and all errors.

1 Introduction

This paper provides simple closed-form results that precisely characterize the empirical content of random utility models of demand when there are two goods. The leading application is demand analysis with repeated cross-sections. Thus, we assume that a researcher observes the distributions of demand generated by some population of consumers on a (finite or infinite) set of budgets. We focus on the assumption that every member of the population is rational in the sense of maximizing a utility function and that utility increases in consumption ("more is better"). We refrain from any other restriction on unobserved heterogeneity. Under these assumptions, we resolve the following questions:

- Can we decide whether the data could have been generated by the postulated model? We will henceforth call this the *rationalizability* question.

- Using the data and only maintaining rationality and "more is better," how can we bound the distribution of demand on an as yet unobserved budget? This will be called the *prediction* question. Prediction will in general be set-valued, i.e. we provide bounds on the distribution of interest.

A practically useful aspect of this paper is that we answer these questions through closed-form expressions. This allows researchers to exactly determine the bite of rationality assumptions in their data without an explicit – restricted or unrestricted – estimation of demand systems. Prediction can be reduced to evaluating upper and lower bounding c.d.f.'s that are provided in closed form. Because these bounding c.d.f.'s exhaust the implications of rationality for the cross-sectional distribution of demand, tight bounds on many other parameters follow immediately (e.g., through Stoye (2010)). The explicit display of these nonparametric conditions reveals that they are rather weak, especially if one wants to extrapolate beyond the support of observed prices. This is a reminder that tests of *rationalizability* may not be very powerful tests of *rationality* (in the sense that many irrational populations will generate rationalizable data), and that one might expect efficiency gains from restricted estimation to be limited. Some of these issues surfaced in the literature discussed below.

This paper's main theoretical novelty is the following observation. Consider imposing *rank invariance* with respect to consumption of good 2. That is, assume that if consumer i consumes more of good 2 than consumer j on some budget, she does so on all budgets. This is a seemingly very strong restriction. It effectively recovers panel data by force of assumption: The γ-th quantile of demand for good 2 is now attributable to the same consumer throughout, and one can test the rationality of that consumer using existing insights from demand theory. Yet we find that for our purposes, imposing rank invariance is without loss of generality: A sequence of cross-sectional distributions is rationalizable iff the behavior of all consumers in the fictitious panel generated by rank invariance is. By the same token, imposing rank invariance does not affect predictive bounds. This observation is reminiscent of a well-known finding for nonstochastic demand data, which are rationalizable by a nonstochastic utility function iff they are also rationalizable by a strictly concave one (Afriat (1967), Varian (1982)).

Our insights are specific to the two-good case. However, this case is prominent in empirical analysis; for example, consider estimation of demand functions for one good (and therefore implicitly for that good versus a numeraire) and analysis of labor supply. Recent papers that nonparametrically analyze demand with unobserved heterogeneity in a two-good setting include Blundell, Horowitz, and Parey (2012, 2013), Blundell, Kristensen, and Matzkin (2014), Hausman and Newey (2013), and Manski (2014). Analysis of the two-good case is frequently motivated by separability assumptions; in the conclusion, we will briefly comment on the interpretation of our results in that context.

While we point out some links to current research on statistical inference, we do not directly contribute to the literature on testing rationality in finite samples. See Hoderlein and Stoye (2014), Manski (2007), and in particular Kitamura and Stoye (2013) for relevant work.

2 Related Literature

The rationalization and prediction questions motivated a classic literature. For individual demand, the empirical content of rationality is embodied in the "Afriat inequalities" (Afriat (1973)); see also Varian (1982) especially for prediction. The extension to stochastic demand generated by a heterogeneous population was provided by McFadden and Richter ((1991); see also Block and Marschak (1960), Falmagne (1978), Barberà and Pattanaik (1986), and McFadden (2005)), but it is not immediate how to compute their criterion.

Nonparametric demand analysis using repeated cross-sections is also the subject of a large, recent literature. One important approach is to extract one or more individual demand systems and then use standard revealed preference analysis. This has the considerable advantage of leveraging classic results, though at the price of restricting heterogeneity.

The most prominent method within this approach is to investigate rationalizability and extrapolation of expected demand. This was pioneered by Blundell, Browning, and Crawford (2003, 2008, BBC henceforth), who nonparametrically estimate Engel curves and then impose the weak axiom of revealed preference (WARP). The approach works in two- and higher-dimensional commodity spaces, though only the recent extension to the strong axiom of revealed preference (SARP) by Blundell, Browning, Cherchye, Crawford, de Rock, and Vermeulen (2014) exhausts the implications of rationalizability in the latter case. Blundell, Horowitz, and Parey (2012) estimate stochastic demand for one good, and hence effectively for one good versus a numeraire. Average demand is presumed to be a formally valid demand function; indeed, constrained estimation of this function imposing the Slutsky conditions is the main contribution. See Haag, Hoderlein, and Pendakur (2009) for a related exercise in higher dimensions.

An important alternative is to examine quantiles of demand rather than its expectation. This was recently pursued by Blundell, Kristensen, and Matzkin (2014, BKM henceforth) as well as Blundell, Horowitz, and Parey (2013, BHP henceforth). Both restrict attention to two goods, at least in their empirical implementations; BKM assume finitely many budgets, whereas BHP consider continuous prices. Both motivate the analysis by the rank invariance assumption that is central to the present paper.

A third group of papers uses the same assumptions that we do, that is, they give up on recovering individual demand systems. Hoderlein and Stoye (2014, HS henceforth) bound the fraction of a population who violate WARP. The analysis is conducted for both two and more goods, with a qualitative change occurring (and the analysis becoming much more intricate) in the latter case. The present paper builds on the two-good analysis in HS. Kitamura and Stoye (2013, KS henceforth) develop and implement a finite sample test of rationalizability when there are many goods. They also consider prediction, though they do not provide confidence regions for predicted demand. From a technical point of view, the scope of their analysis extends ours, with the caveat that they restrict attention to continuous distributions and finite sets of budgets. However, their approach is computationally demanding, whereas we provide closed-form results that are readily computed even for large sets of budgets, as well as additional insights that do not extend to the more general case.

In general, the aggregation or invertibility assumptions used by the three groups of papers are related as follows. Stochastic rationalizability implies integrability of expected demand only under specific conditions laid out by Lewbel (2001). The reverse implication does not hold at all. Invertibility of demand, e.g. through rank invariance, and rationalizability of the resulting demand systems are jointly (much) stronger than stochastic rationalizability. However, we discover that this is not true for repeated cross-sections with two goods, where rationalizability of quantile demand and stochastic rationalizability are observationally equivalent.

We conclude this section by mentioning some other related papers. Hoderlein (2011) shows that certain features of rationalizable individual (nonstochastic) demand, like adding up and standard properties of the Slutsky matrix, are inherited by average demand under very weak conditions. He develops and implements a corresponding test. The catch is that the test is of a one-sided implication of rationalizability, so that its power is fundamentally limited. Manski (2007) analyzes rationalizability and prediction for finite (and in practice, small) choice sets, and Manski (2014) applies this analysis to a labor-leisure trade-off, i.e. a setting with two commodities. The environment considered, and questions asked, are similar to ours, but Manski analyzes discrete budgets, and budgets that are induced by tax rates might furthermore have kinks. Both of these features mean that Manski's (2007, 2014) and our findings are not directly comparable.

3 The Case of Two Budgets

Consider a population of agents, all of whom face the same sequence of budgets in $\mathbb{R}^2_+$. Income is normalized to 1 and dropped from notation, so that budgets are characterized by (normalized) price vectors $\mathbf{p} = (p^1, p^2)$. We assume that "more is better," hence consumers exhaust their budgets. We also assume that individual choice from a given budget is nonstochastic. Individual demand is, then, given by the function $q(\mathbf{p}, \mathbf{a}) : \mathbb{R}^2_+ \times \mathcal{A} \to \mathbb{R}_+$, where $q(\mathbf{p}, \mathbf{a}) \in [0, 1/p^2]$ is quantity demanded of good 2, quantity demanded of good 1 is implied through budget exhaustion, and the unobservable covariate $\mathbf{a} \in \mathcal{A}$ parameterizes heterogeneity. From the researcher's point of view, $\mathbf{a}$ is an unobserved random variable with unknown but fixed cross-sectional distribution, and hence $q(\mathbf{p}, \mathbf{a})$ is a random variable supported on

$[0, 1/p^2]$. The researcher knows, or can estimate, the distribution of $q_t = q(\mathbf{p}_t, \mathbf{a})$ for each $\mathbf{p}_t$ in a sequence of price regimes $\{\mathbf{p}_t\}_{t \in \mathcal{T}}$. We take the index set $\mathcal{T}$ and the sequence $\{\mathbf{p}_t\}_{t \in \mathcal{T}}$ to be nonstochastic or conditioned upon. The interpretation of t as time periods is for concreteness but is not essential. The distribution of q_t may be continuous, discrete, or mixed discrete-continuous. Let F_t denote its c.d.f. and define $F_t^-(z) = \Pr(q_t < z)$; of course, F_t and F_t^- coincide if the former is continuous. An observable covariate is omitted for simplicity but could easily be introduced; see numerous of the aforecited papers. Similarly, we assume that the income process is exogenous. Note that, while exogeneity of income and nonsatiation are substantive assumptions, normalization of income to 1 does not lose any generality.

The nonstochastic collection of demands $\{q(\mathbf{p}_t, \mathbf{a})\}_{t \in \mathcal{T}}$ induced by a given, fixed $\mathbf{a}$ is rationalizable if there exists a utility function $u_{\mathbf{a}} : \mathbb{R}_+^2 \to \mathbb{R}$ such that $q(\mathbf{p}_t, \mathbf{a})$ maximizes $u_{\mathbf{a}}$ on $\{\mathbf{x} \in \mathbb{R}_+^2 : \mathbf{p}_t \mathbf{x} \leq 1\}$ for all t. This obtains iff $\{q(\mathbf{p}_t, \mathbf{a})\}_{t \in \mathcal{T}}$ fulfils the weak axiom of revealed preference (WARP), i.e. iff there is no pair of time indices $(s, t) \in \mathcal{T} \times \mathcal{T}$ such that $p_s^1 \left(1 - p_s^2 q(\mathbf{p}_t, \mathbf{a})\right)/p_s^1 + p_s^2 q(\mathbf{p}_t, \mathbf{a}) \leq 1$ and $p_t^1 \left(1 - p_t^2 q(\mathbf{p}_s, \mathbf{a})\right)/p_t^1 + p_t^2 q(\mathbf{p}_s, \mathbf{a}) \leq 1$, with at least one inequality strict. Furthermore, recall Afriat's (1967) finding that if nonstochastic demand is rationalizable, then it is rationalizable by a strictly convex utility function. Hence, strict convexity of $u_{\mathbf{a}}$ can be imposed without further loss of generality, in which case $q(\mathbf{p}_t, \mathbf{a}) = \arg\max_{x \in [0, 1/p_t^2]} u_{\mathbf{a}}(\left(1 - p^2 x\right)/p^1, x)$.

A collection of cross-sectional distributions $\{F_t\}_{t \in \mathcal{T}}$ is stochastically rationalizable if there exists a distribution G over strictly convex utility functions u such that

$$F_t(z) = \Pr(\arg \max_{x \in [0, 1/p_t^2]} u(\left(1 - p^2 x\right)/p^1, x) \leq z),$$

where the probability is evaluated with respect to G. By choosing G to be the distribution of $u_{\mathbf{a}}$ induced by the distribution of $\mathbf{a}$, one can see that $\{F_t\}_{t \in \mathcal{T}}$ is stochastically rationalizable if $\{q(\mathbf{p}_t, \mathbf{a})\}_{t \in \mathcal{T}}$ is rationalizable for each $\mathbf{a}$. The converse does not hold: A population consisting of irrational (in the sense of violating WARP) agents may induce a rationalizable collection of cross-sectional demand distributions.[1]

The following assumption will play a prominent role.[2]

[1] The distribution G defines a Random Utility Model; thus, $\{F_t\}_{t \in \mathcal{T}}$ is stochastically rationalizable iff it can arise as population distribution from some Random Utility Model. In principle, the model could also be interpreted as describing single-agent stochastic choice, but our assumptions are geared to demand analysis with repeated cross-sections.

[2] The name "rank invariance" is inspired by Chernozhukov and Hansen (2005), who impose the assumption

Rank Invariance

$\mathcal{A} = [0,1]$ *(possibly after normalization) and* $q(\mathbf{p}, \mathbf{a})$ *is nondecreasing in* $\mathbf{a}$ for every $\mathbf{p}$.

For continuously distributed demand, rank invariance implies that across all observed budgets and for any $\gamma \in [0,1]$, the γ-quantile of the demand distribution is attributable to the same consumer. After estimating these quantiles for different budgets, this allows one to leverage single-consumer demand theory, as is pursued by BHP and BKM.

Rank invariance does have "microfoundations" in the sense that with two goods, it follows from restrictions on structural parameters which more generally would allow one to recover the underlying individual demand systems; see Beckert and Blundell (2008) and references therein. Nonetheless, it is intuitively restrictive, is consistent with random coefficient models only in knife-edge cases, and is contradicted in lab experiments.[3]

Initially consider one pair of budgets, thus $\mathcal{T} = \{s, t\}$. In this case, stochastic rationalizability can be characterized as follows.

Proposition 1 *Fix any price vectors* $(\mathbf{p}_s, \mathbf{p}_t)$*; w.l.o.g., assume* $p_s^2 \leq p_t^2$*. Let*

$$q_{st}^* = \left(p_t^1 - p_s^1\right) / \left(p_t^1 p_s^2 - p_s^1 p_t^2\right),$$

the quantity at which the corresponding budget lines intersect. Then the following statements are mutually equivalent:

(i) *The cross-sectional distributions of demand on budgets* $\mathbf{p}_s$ *and* $\mathbf{p}_t$ *can be stochastically rationalized.*

(ii) *The cross-sectional distributions of demand on budgets* $\mathbf{p}_s$ *and* $\mathbf{p}_t$ *can be stochastically rationalized by assuming rank invariance.*

(iii)

$$\Pr\left(q_s \leq q_{st}^*\right) + \Pr\left(q_t \geq q_{st}^*\right) - \min\left\{\Pr\left(q_s = q_{st}^*\right), \Pr\left(q_t = q_{st}^*\right)\right\} \leq 1,$$

in a treatment effects context. Hoderlein and Stoye (2014) call the same assumption "quantile constancy."

[3] Andreoni and Miller (2002) observe repeated choices by the same consumers in a sequence of linear budget sets with two goods. Rank invariance is frequently violated; one violation is visually apparent from comparing panels c and d of their Figure 2. We are not aware of direct tests of rank invariance outide the lab. This would require consumer panel data, which are rare; see Christensen (2014) for references and a discussion. That said, the heterogeneity of income effects estimated by Hoderlein, Klemelä, and Mammen (2010) and Christensen (2014) would preclude rank invariance in two-good versions of their models.

which with continuous demand is equivalent to

$$F_s(q_{st}^*) \leq F_t(q_{st}^*).$$

If the budgets do not intersect, all conditions hold vacuously (the last one because the intersection of budgets is not in the positive quadrant, thus the inequalities necessarily hold with equality), illustrating that consumer rationality is not falsifiable in this case.

Proof. Individual demand $(q(\mathbf{p}_s, \mathbf{a}), q(\mathbf{p}_t, \mathbf{a}))$ violates WARP iff $q(\mathbf{p}_s, \mathbf{a}) \leq q_{st}^*$ and $q(\mathbf{p}_t, \mathbf{a}) \geq q_{st}^*$ but not $q(\mathbf{p}_s, \mathbf{a}) = q(\mathbf{p}_t, \mathbf{a}) = q_{st}^*$. Thus demand is stochastically rationalizable if the joint population distribution of (q_s, q_t) fulfils $\Pr(q_s \leq q_{st}^*, q_t \geq q_{st}^*) - \Pr(q_s = q_t = q_{st}^*) = 0$. Now, $\Pr(q_s \leq q_{st}^*, q_t \geq q_{st}^*) \geq \max\{\Pr(q_s \leq q_{st}^*) + \Pr(q_t \geq q_{st}^*) - 1, 0\}$ by the Fréchet-Hoeffding lower bound, whereas $\Pr(q_s = q_t = q_{st}^*) \leq \min\{\Pr(q_s = q_{st}^*), \Pr(q_t = q_{st}^*)\}$ by the Fréchet-Hoeffding upper bound, thus the probability in question is bounded below by $\max\{\Pr(q_s \leq q_{st}^*) + \Pr(q_t \geq q_{st}^*) - \min\{\Pr(q_s = q_{st}^*), \Pr(q_t = q_{st}^*)\} - 1, 0\}$ and can be zero only if condition (iii) holds. This yields (i)$\Rightarrow$(iii). Next, condition (iii) implies that $\Pr(q_s < q_{st}^*) + \Pr(q_t \geq q_{st}^*) \leq 1$, hence that $\Pr(q_s < q_{st}^*) \leq \Pr(q_t < q_{st}^*)$. If one were to assume rank invariance, it would follow that any consumer with $q_s < q_{st}^*$ also has $q_t < q_{st}^*$. Similarly, every consumer with $q_t > q_{st}^*$ would have $q_s > q_{st}^*$. Hence, every consumer would fulfil WARP. It follows that if condition (iii) holds, then one can rationalize observed demand by imposing rank invariance. This establishes (iii)$\Rightarrow$(ii). (ii)$\Rightarrow$(i) is obvious. ∎

The simplified version of condition (iii) for continuous demand has a visual interpretation: As a budget plane is rotated, the probability mass on the part of the plane that is rotated toward the origin must shrink. We emphasize that (i)$\Leftrightarrow$(iii) was known in the literature.[4] This paper's main novelty is to recognize (i)$\Leftrightarrow$(ii), i.e. that the marginal distributions of choices on two budgets are jointly rationalizable iff they are rationalizable by imposing rank invariance, and to spell out some implications of this fact.

An important application is counterfactual prediction. Given the distribution F_t of q_t and assuming rationality, a candidate distribution F_s of q_s is rationalizable with or without rank invariance iff (F_s, F_t) jointly fulfil (iii). This insight and the resulting bounds are summarized in the following corollary.

[4] See McFadden and Richter (1991), Matzkin (2006, for the continuous case), and HS. We provide an independent proof because it is very short and instructive. (iii)$\Rightarrow$(i) can also be proved by combining results in Bandyopadhyay, Dasgupta, and Pattanaik (2002, 2004).

Corollary 1 *Fix any price vectors* $(\mathbf{p}_s, \mathbf{p}_t)$ *and let the distribution of* q_t *be known. If one assumes rationalizability of demand, then:*

(i) *If* $p_t^2 > p_s^2$*, then* $F_s^-(q_{st}^*) \leq F_t^-(q_{st}^*)$ *and* $F_s(q_{st}^*) \leq F_t(q_{st}^*)$*.*

(ii) *If* $p_t^2 < p_s^2$*, then* $F_s^-(q_{st}^*) \geq F_t^-(q_{st}^*)$ *and* $F_s(q_{st}^*) \geq F_t(q_{st}^*)$*.*

(iii) F_s *is not otherwise restricted, nor are there cross-restrictions across values of* q *(other than* F_s *being a c.d.f.).*

All of (i)-(iii) remain true if one also imposes rank invariance.

Remark 1 *The set of distributions consistent with these bounds is not closed under weak convergence, and the implied bounds on many parameters are open. To see this, consider minimization of* F_s *(and hence maximization of the distribution of* q *in the sense of first-order dominance) in case (ii). The solution is approximated by a point mass at* $q = 1/p_s^2$ *and a second point mass that approaches* q_{st}^* *from below. The limit of this sequence violates WARP. Among other things, this implies that the identified set for expected demand is open at the upper bound.*

For many purposes, it would seem natural to replace the bounds with their closure; indeed, this will typically make no difference in estimation. This closure is succinctly characterized as the set of distributions whose c.d.f. F *fulfils* $\underline{F}_s(q) \leq F(q) \leq \overline{F}_s(q)$*, where*

$$\underline{F}_s(q) = \begin{cases} F_t(q_{st}^*) & \text{if } p_t^2 < p_s^2 \text{ and } q \geq q_{st}^* \\ 1 & \text{if } q = 1/p_s^2 \\ 0 & \text{otherwise.} \end{cases}$$

$$\overline{F}_s(q) = \begin{cases} F_t^-(q_{st}^*) & \text{if } p_t^2 > p_s^2 \text{ and } q < q_{st}^* \\ 1 & \text{otherwise.} \end{cases}$$

This is the "bounds on c.d.f.'s" scenario considered at length in Stoye (2010). Hence, results in Stoye (2010) now yield (the closure of) sharp, joint bounds on many combinations of parameters of F_s*, e.g. its expectation and variance or median and interquartile range.*

Again, a remarkable aspect of the finding is that rank invariance, even though it is substantively so strong as to be incredible in many applications, does not restrict the distribution of demand on a counterfactual budget. For a visual illustration, consider figure 1. Say that observed data were a generated by a population of size 4 who chose pairs of con-

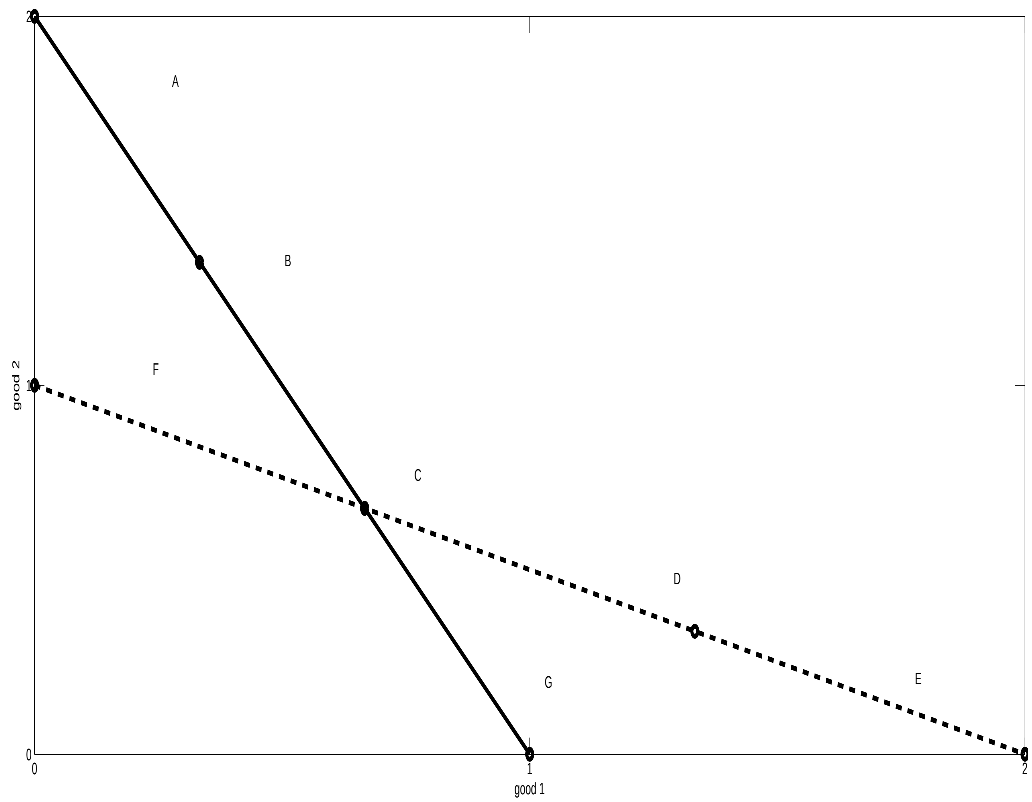

Figure 1: An illustration of Lemma 1.

sumption bundles $\{(A,F),(B,D),(C,C),(G,E)\}$, respectively.[5] Observe that all consumers fulfil WARP but that the population violates rank invariance.

The researcher observed only the implied cross-sectional distributions, which are uniform on $\{A,B,C,G\}$ and $\{F,C,D,E\}$, respectively. Having been generated by rational consumers, one would hope that these cross-sectional distributions are rationalizable. Indeed, condition (iii) from Proposition 1 holds with slack. Proposition 1 then asserts that the cross-sectional distributions can be rationalized by imposing rank invariance, i.e. by assuming that consumption pairs $\{(A,F),(B,C),(C,D),(G,E)\}$ are attributable to the same consumers. While the assumption is false in the example, the rationalization is indeed successful – all of these consumption pairs fulfil WARP. The punch line of Proposition 1 is that this example generalizes.

The point is not that rank invariance is a weak restriction on the underlying population; to the contrary, it is unlikely to hold. However, the assumptions of stochastic rationalizability and "stochastic rationalizability plus rank invariance" are observationally equivalent for repeated cross-sections with two goods. If repeated observations existed for at least some consumers, rank invariance would have bite. For example, if it were known that the consumer who picked B from the first budget picked C from the second one, then rank invariance would imply that A and F were also picked by the same consumer; but, as the example illustrates, this was not implied by rationalizability alone.

We further discuss the importance of these findings after generalizing them. We conclude this section by pointing out that Proposition 1 and Corollary 2 could be easily implemented with finite samples. For example, the hypothesis of rationalizability applied to continuously distributed demand is just

$$H_0 : \Pr\left(q_s < q^*_{st}\right) + \Pr\left(q_t > q^*_{st}\right) \leq 1,$$

a standard one-sided testing problem.

4 Extension to Many Budgets

A main reason that these results are interesting is that with two goods and when budgets are exhausted, WARP implies the strong axiom of revealed preference (SARP) and, therefore,

[5] For clarity, we here abstract from any measurement error or any other "ε_{it}" term that would induce violations of rank invariance in observed data even if it were a true feature of the distribution of preferences.

characterizes rationalizability of individual demand (Rose (1958)).[6] Hence, if there are two goods but many budgets, the rationalizability and prediction questions are resolved by applying Proposition 1 to every pair of budgets. Formally, consider a collection $(\mathbf{p}_t, F_t)_{t\in\mathcal{T}}$ of price vectors and corresponding demand distributions, where the index set $\mathcal{T}$ is initially assumed finite but can be arbitrarily large. Then we have:

Proposition 2 *Fix any finite collection $\{\mathbf{p}_t, F_t\}_{t\in\mathcal{T}}$. The following statements are equivalent:*

(i) *The cross-sectional distributions of demand on budgets $\{\mathbf{p}_t\}_{t\in\mathcal{T}}$ can be stochastically rationalized.*

(ii) *The cross-sectional distributions of demand on budgets $\{\mathbf{p}_t\}_{t\in\mathcal{T}}$ can be stochastically rationalized by assuming rank invariance.*

(iii) *Condition (iii) from Proposition 1 applies to every $(\mathbf{p}_s, F_s, \mathbf{p}_t, F_t)$ with $(s,t) \in \mathcal{T}\times\mathcal{T}$.*

Proof. Suppose condition (iii) is violated for one (s,t), then some consumer must violate WARP and hence SARP. Thus, (i)⇒(iii). Next, suppose the condition holds for all (s,t). Rank invariance would then be a coherent assumption in the sense that it implies a well-defined joint distribution of demand on all budgets. Furthermore, it would imply that, by (iii)⇒(ii) from Proposition 1, all consumers fulfil WARP and hence SARP. Thus, (iii)⇒(ii). (ii)⇒(i) is again obvious. ■

Some book-keeping leads to the following corollary.

Corollary 2 *Fix the collection $\{\mathbf{p}_t, F_t\}_{t\in\mathcal{T}}$, but assume that F_s corresponding to some budget of interest $\mathbf{p}_s$ is unknown. Assuming rationalizability only, F_s can be bounded as follows:*

$$\max\left\{\max_{t\in\mathcal{T}:p_t^2<p_s^2,q_{st}^*<q} F_t(q_{st}^*), \max_{t\in\mathcal{T}:p_t^2<p_s^2,q_{st}^*=q} F_t^-(q_{st}^*)\right\} \le F_s^-(q) \le \min_{t\in\mathcal{T}:p_t^2>p_s^2,q_{st}^*\ge q} F_t^-(q_{st}^*)$$

$$\max_{t\in\mathcal{T}:p_t^2<p_s^2,q_{st}^*=q} F_t(q_{st}^*) \le F_s(q) \le \min\left\{\min_{t\in\mathcal{T}:p_t^2>p_s^2,q_{st}^*>q} F_t^-(q_{st}^*), \min_{t\in\mathcal{T}:p_t^2>p_s^2,q_{st}^*=q} F_t(q_{st}^*)\right\},$$

with the understanding that max*-operators corresponding to empty sets are ignored. These bounds simplify to*

$$\max_{t\in\mathcal{T}:p_t^2<p_s^2,q_{st}^*\le q} F_t(q_{st}^*) \le F_s^-(q) \text{ and } F_s(q) \le \min_{t\in\mathcal{T}:p_t^2>p_s^2,q_{st}^*\ge q} F_t(q_{st}^*)$$

[6] Recall that SARP excludes revealed preference cyces of arbitrary length. In this paper's notation, call $q(\mathbf{p}_s, \mathbf{a})$ revealed preferred to $q(\mathbf{p}_t, \mathbf{a})$ if $p_s^1\left(1 - p_s^2 q(\mathbf{p}_t, \mathbf{a})\right)/p_s^1 + p_s^2 q(\mathbf{p}_t, \mathbf{a}) \le 1$, then there must not exist any finite length cycle of revealed preference in which at least one of the inequalities is strict.

if all observed distributions are continuous. They are sharp uniformly in q, that is, any c.d.f. that fulfils them is a possible distribution of demand.

All of these statements are true with and without imposing rank invariance.

Remark 2 *As before, the bounds are open (under weak convergence) but their closure can be characterized in terms of sandwich c.d.f.'s, which now are*

$$\underline{F}_s(q) = \begin{cases} \max_{t\in\mathcal{T}:p_t^2<p_s^2,q_{st}^*\leq q} F_t(q_{st}^*) & \text{if the maximum is over a nonempty set} \\ & \text{and } q\in[0,1/p_s^2), \\ 1 & \text{if } q=1/p_s^2, \\ 0 & \text{otherwise.} \end{cases}$$

$$\overline{F}_s(q) = \begin{cases} \min_{t\in\mathcal{T}:p_t^2>p_s^2,q_{st}^*>q} F_t^-(q_{st}^*) & \text{if the minimum is over a nonempty set} \\ & \text{and } q\in[0,1/p_s^2), \\ 1 & \text{otherwise.} \end{cases}$$

The closure of bounds on numerous parameters of F_s follows just as in remark 1.

Remark 3 *To compute joint bounds on cross-sectional distributions on a set of budgets $\{\mathbf{p}_{s_1},\mathbf{p}_{s_2},\ldots\}$, one can in principle proceed as follows. Use corollary 4 to bound F_{s_1}. For each candidate distribution F_{s_1} within the identified set, compute the identified set for F_{s_2} taking $\{\mathbf{p}_t,F_t\}_{t\in\mathcal{T}}\cup\{\mathbf{p}_{s_1},F_{s_1}\}$ as known. Iterate. The resulting bounds are not to be confused with bounds on the joint distribution. However, they imply best possible joint bounds across budgets on all parameters that only depend on cross-sectional distributions, e.g. on expected demand. These joint bounds are not in general the Cartesian product of the individual bounds.*

While computational cost of fully characterizing these joint bounds will escalate quickly, two observations are of interest. First, it is easy to verify whether a given tuple of candidate distributions is in the identified set. More importantly, some additional book-keeping reveals that $\left\{\underline{F}_{s_1},\underline{F}_{s_2}\right\}$ and $\left\{\overline{F}_{s_1}.\overline{F}_{s_2}\right\}$ – but not, in general, $\left\{\underline{F}_{s_1}.\overline{F}_{s_2}\right\}$ – can be approximated by mutually consistent c.d.f.'s. By Proposition 3, this consistency extends to larger collections of counterfactual budgets. Thus, for any nonsingleton set of counterfactual budgets and any parameter that respects first-order dominance, all lower as well as all upper bounds on the parameter – though not all combinations thereof – can be achieved simultaneously across counterfactual budgets. For example, the joint identification region for expected demand across several counterfactual budgets contains its own meet and join.

Remark 4 *The data are rationalizable iff prediction bounds are nonempty for all counterfactual budgets. To economize on case distinctions, we verify this in the continuous case. Here, the bound on $F_s(q)$ will be empty if the lower bound exceeds the upper one, i.e. if there exist t, v s.t. $p_t^2 < p_s^2 < p_v^2$, $q_{st}^* \leq q_{sv}^*$, and $F_t(q_{st}^*) > F_v(q_{sv}^*)$. In this case, comparison of budgets s and v would reveal a violation of the condition from Proposition 1(iii) because it would follow that $q_{sv}^* \geq q_{tv}^* \geq q_{st}^*$, thus $F_t(q_{tv}^*) \geq F_t(q_{st}^*)$ and $F_v(q_{tv}^*) \leq F_v(q_{sv}^*)$, thus $F_t(q_{tv}^*) > F_v(q_{tv}^*)$. Thus, if the data are rationalizable, then the bounds are nonempty. On the other hand, if data are not rationalizable, this fact must be apparent from at least one pair of price vectors $(\mathbf{p}_t, \mathbf{p}_v)$. Let $\mathbf{p}_s$ be a proper mixture of $\mathbf{p}_t$ and $\mathbf{p}_v$, then prediction bounds for F_s are empty. Note, however, that even if data are not rationalizable, bounds may be nonempty for some specific choices of $\mathbf{p}_s$.*

Remark 5 *Proposition 3 and corollary 4 extend to infinite index sets $\mathcal{T}$ under mild regularity conditions, e.g. McFadden's (2005, Corollary 5.3.2) smoothness assumptions on demand. The reason is that under these conditions, McFadden (2005) shows that rationalizability is implied by his Axiom of Revealed Stochastic Preference (ARSP), yet any instance of ARSP applies to a finite selection from $\mathcal{T}$ (though it must hold for all such selections), and WARP for individual demand implies SARP for individual demand, hence ARSP, on every such selection.*

Remark 6 *For this remark only, assume that the data are consistent with rationality and that the focus is on extrapolation to cross-sectional distributions on one or more counterfactual budgets. In this case, it is w.l.o.g. to assume rank invariance even if one has panel data, i.e. if consumers' identities are tracked across budgets, irrespective of whether rank invariance holds in these data. To see this, recall that predictive bounds on F_s will collect the corresponding marginals of all rationalizable joint distributions of demand on budgets $\{\mathbf{p}_t\}_{t \in \mathcal{T}} \cup \{\mathbf{p}_s\}$. Any such joint distribution will imply a rationalizable sequence of marginal distributions. By Proposition 3, that same sequence is then also rationalizable through rank invariance. Thus, for the specific purpose outlined in this remark, the panel dimension of the data can be discarded in favor of imposing rank invariance.*[7]

[7] This remark applies only to nonparametric revealed preference analysis along the lines of this paper, e.g. nonparametric extrapolation from data in Andreoni and Miller (2002). If one uses panel data to estimate more structured models, the panel dimension might certainly be useful. Also, a given panel data set might reveal violations of WARP. If one is willing to attribute them to data issues but wants to pose a well-

Going back to Figure 1, the additional insight needed to generate Proposition 3 is twofold. First, WARP implies SARP, so all pairwise comparisons of budgets jointly suffice to check rationalizability. Second, the rank invariance assumption is consistent across budgets in the following sense: If we "link" the distributions on budget 1 and 2 through rank invariance and do the same for budgets 2 and 3, then we implicitly defined a link between budgets 1 and 3 that also corresponds to rank invariance. Hence, the fictitious panel generated through imposing rank invariance is well-defined. A similar consistency does not, of course, obtain for arbitrary pairwise copulas.

The conditions from Proposition 3(iii) are readily verifiable even when $\mathcal{T}$ is large. This is in stark contrast to checking ARSP more generally. The currently most powerful computational approach to the latter was provided by KS. Using advanced numerical solvers, KS are able to analyze up to $T = 8$ budgets when there are three goods and could analyze substantially more budgets when there are only two goods. The complexity of their problem increases faster than exponentially with T, however. In contrast, verifying the condition of Proposition 3(iii) requires at most $T(T-1)/2$ checks of the inequality in Proposition 1(iii). Regarding the strength of these bounds, note that $\underline{F}_s(q)$ and $\overline{F}_s(q)$ are potentially close to each other, and bounds therefore potentially tight, as long as the maxima and minima in their definitions are over nonempty sets. This corresponds to extrapolation to budgets that interpolate between observed ones. As one leaves the support of observed budgets, the bounds become very wide. A similar issue is encountered in BBC's prediction exercise and seems to be a fundamental limitation of nonparametric demand extrapolation based on weak rationality restrictions. Interestingly, an apparently substantial addition of structure (i.e., rank invariance) does not affect this finding.

Regarding finite sample inference, the simultaneous testing of rationality on many budgets raises subtle issues. In general, it would amount to specification testing of a moment inequalities model, an area of current research. Here are some ideas. First, with finite index set $\mathcal{T}$, the null of rationalizability is the conjunction of finitely many inequality hypotheses. Second, in the continuous case, this null is equivalent to

$$H_0 : \max_{t \in \mathcal{T}} \left\{ \Pr\left(q_s < q^*_{st}\right) + \Pr\left(q_t > q^*_{st}\right) \right\} \leq 1,$$

defined extrapolation problem, one will have to "iron them out" in some manner; else, there will not exist a rationalizable joint distribution of choices over all budgets. Our remark then applies to the ironed out data.

which can be tested using the tools developed in Chernozhukov, Lee, and Rosen's (2013) work on intersection bounds. Similarly, the bounds on F_s from corollary 4 are intersection bounds. Finally, asymptotic results in KS apply to this paper's setting, though they do not exploit the special structure of the two-good case.

5 Conclusion

We focus in this conclusion on some extensions and on the methodological relevance of our finding, in particular the observation that if rationalizability is assumed, then rank invariance does not further constrain cross-sectional distributions. This finding is reminiscent of the well-known observation (Varian (1982)) that if rationalizability of individual demand (in the sense of SARP) is assumed, then assuming convex preferences has no further empirical content. The findings arise from very different features of the problem and are easily combined. Thus, if one assumes rationalizability of demand, then one may as well assume rationalizability through a set of convex utility functions that are parameterized by $\mathbf{a} \in [0, 1]$ and where revealed taste for good 2 increases in $\mathbf{a}$.

An immediate implication is that none of the identification of counterfactual demand in BKM is owed to their seemingly restrictive rank invariance assumption, including if one were to apply their methods to panel data. This makes their method widely applicable. At the same time, for the two-good space of their empirical implementation, it suggests an alternative approach because it implies that up to estimation uncertainty, their bounds recover the closed-form characterizations reported above. One could use Chernozhukov, Lee, and Rosen's (2013) work on intersection bounds to directly test these characterizations and estimate the predictive bounds, sidestepping the construction of a fictitious "γ-quantile" consumer. Also, BKM focus on bounding quantiles of demand. While bounds on other features of demand are logically implied, it would arguably be more straightforward to bound expected demand, say, by directly computing the objects $(\underline{F}_s, \overline{F}_s)$ defined earlier.

We also note that the finding can be turned around: By assuming that demand is monotonic in unobservables, one will never implicitly assume rationalizability when there is none, nor assume it away when it actually obtains. Given that rank invariance assumptions are generally considered strong, and given that more structured models of demand which allow for heterogeneity typically violate it, this arguably illustrates that rationalizability is a rather weak restriction on data. This assessment is corroborated by inspection of the bounds and ties in with other findings in the literature, e.g. the observation that imposing negative semidefiniteness in nonparametric demand estimation does not affect the estimator's rate of convergence (Haag, Hoderlein, and Pendakur (2009)).

Analysis of the two-good case is frequently, and sometimes implicitly, motivated by separability conditions on utility. In this case, assumptions well beyond rationality must be invoked to motivate the setting, and this may seem to undercut the motivation for testing a bare-bones rationality assumption within the setting. However, there is a growing literature on demand estimation using minimal assumptions beyond separability. Researchers interested in this line of work might want to learn the implications of assuming that agents act as if the choice between two goods could be separated from their other decision problems, but without any restriction on their preferences beyond this. The present paper provides just that analysis.

One might also wonder about generalizability of results to settings where there are two commodities, but budget sets are not fully described by linear budget constraints. Under mild conditions outlined in Polisson and Quah (2013), our findings generalize to the case where consumers are constrained to a discrete subset $\mathcal{X}$ of $\mathbb{R}^2_+$, but can pick any point in this subset that is below a given budget plane, i.e. prices $\mathbf{p}$ induce budget $\{\mathbf{x} \in \mathcal{X} : \mathbf{px} \leq 1\}$. This follows from combining Polisson and Quah's (2013) insights with the observation that, as revealed by close inspection of Rose (1958), WARP continues to imply SARP in this setting. Rose's (1958) finding, and therefore our results, do not extend to arbitrary collections of discrete choice sets, even if choice objects are elements of $\mathbb{R}^2$, nor if budgets can be kinked, as is relevant in analysis of labor supply.

We conclude with some remarks about the case of more than two goods. Our finding about rank invariance does not seem to generalize. To begin, it is unclear how to define rank invariance in higher dimensions: Budget hyperplanes now have two or more dimensions, so that there is no obvious ranking of consumers on them. Imposing rank invariance for one commodity is arbitrary, and imposing it for all commodities implies knife-edge conditions on cross-sectional distributions. However, we are extremely skeptical that any adaptation of rank invariance could be implied by, or even generally consistent with, stochastic rationalizability. For example, consider price vectors $\mathbf{p}_s = (.05, .03, .02)$ and $\mathbf{p}_t = (.02, .1, .03)$ and cross-sectional distributions of demand that are uniform over $\{(10, 0, 25), (11, 15, 0)\}$ in period s and $\{(40,2,0),(2,0,32)\}$ in period t. Stochastic rationalizability obtains because if all consumers who choose $(10,0,25)$ also choose $(40,2,0)$, then each individual demand is rationalizable. However, rank invariance for any one good (or for all three, which is logically possible here due to careful rigging of the example) implies that consumers who choose $(10,0,25)$ also choose $(2,0,32)$ and then violate WARP. Less stringent adaptations of rank invariance, like association or positive quadrant dependence,[8] will still require positive correlation between choice of $(10,0,25)$ and $(2,0,32)$ and, therefore, contradict stochastic rationalizability.

That said, our closed-form expressions could still be used to test a necessary condition for rationalizability and to construct outer (non-tight) bounds on counterfactual demand. These results are theoretically less satisfying than those in KS, but are much easier to implement. Cosaert and Demuynck (2014) and Kawaguchi (2012) recently made intermediate proposals, both in terms of strength of conditions and of computational expense; recall also that Hoderlein (2011) provided other restrictions that are (under weak assumptions) necessary but not sufficient for rationalizability. Comparing the performance of these approaches in empirical examples would be a fruitful area for future research.

References

[1] Afriat, S., 1967. The Construction of a Utility Function from Demand Data, *International Economic Review*, 8, 67-77.

[2] —, 1973. On a System of Inequalities in Demand Analysis: An Extension of the Classical Method. *International Economic Review*, 14, 460-472.

[3] Andreoni, J., and J. Miller, 2002. Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism. *Econometrica*, 70, 737-753.

[4] Bandyopadhyay, T., I. Dasgupta, and P.K. Pattanaik, 2002. Demand Aggregation and the Weak Axiom of Stochastic Revealed Preference. *Journal of Economic Theory*, 107, 483-489.

[5] —, 2004. A General Revealed Preference Theorem for Stochastic Demand Behavior. *Economic Theory*, 23, 589-599.

[8] See Hoderlein and Stoye (2014) for definitions of these concepts.

[6] Barberà, S., and P.K. Pattanaik, 1986. Falmagne and the Rationalizability of Stochastic Choices in Terms of Random Orderings. *Econometrica*, 54, 707-715.

[7] Beckert, W. and R. Blundell, 2008. Heterogeneity and the Non-parametric Analysis of Consumer Choice: Conditions for Invertibility. *Review of Economic Studies*, 75, 1069-1080.

[8] Block, H. and J. Marschak, 1960. Random Orderings and Stochastic Theories of Response. In I. Oklin et al. (eds.), *Contributions to Probability and Statistics*. Stanford University Press.

[9] Blundell, R., M. Browning, L. Cherchye, I. Crawford, B. de Rock, and F. Vermeulen, 2014. Sharp for SARP: Nonparametric Bounds on the Behavioural and Welfare Effects of Price Changes. *American Economic Journal: Microeconomics*, forthcoming.

[10] Blundell, R., M. Browning, and I. Crawford, 2003. Nonparametric Engel Curves and Revealed Preference. *Econometrica*, 71, 205-240.

[11] —, 2008. Best Nonparametric Bounds on Demand Responses. *Econometrica*, 76, 1227-1262.

[12] Blundell, R., J. Horowitz, and M. Parey, 2012. Measuring the Price Responsiveness of Gasoline Demand: Economic Shape Restrictions and Nonparametric Demand Estimation. *Quantitative Economics*, 3, 29-51.

[13] —, 2013. Nonparametric Estimation of a Heterogeneous Demand Function under the Slutsky Inequality Restriction. CeMMAP working paper 54/13.

[14] Blundell, R., D. Kristensen, and R.L. Matzkin. 2014. Bounding Quantile Demand Functions Using Revealed Preference Inequalities. *Journal of Econometrics*, 179, 112–127.

[15] Chernozhukov, V., and C. Hansen, 2005. An IV Model of Quantile Treatment Effects. *Econometrica*, 73, 245-261.

[16] Chernozhukov, V., S. Lee, and A.M. Rosen, 2013. Intersection Bounds: Estimation and Inference. *Econometrica*, 81, 667–737.

[17] Christensen, M., 2014. Heterogeneity in Consumer Demands and the Income Effect: Evidence from Panel Data. *Scandinavian Journal of Economics*, 116, 335–355.

[18] Cosaert, S. and T. Demuynck, 2014. Nonparametric Welfare and Demand Analysis with Unobserved Individual Heterogeneity. Discussion Paper, Maastricht University.

[19] Epstein, L.G. and A.J. Yatchew, 1985. Non-parametric Hypothesis Testing Procedures and Applications to Demand Analysis. *Journal of Econometrics*, 30, 149-169.

[20] Falmagne, J.-C., 1978. A Representation Theorem for Finite Random Scale Systems. *Journal of Mathematical Psychology*, 18, 52-72.

[21] Haag, B.R., S. Hoderlein, and K. Pendakur, 2009. Testing and Imposing Slutsky Symmetry in Nonparametric Demand Systems. *Journal of Econometrics*, 153, 33-50.

[22] Hausman, J. and W. Newey, 2013. Individual Heterogeneity and Average Welfare. CeMMAP working paper 34/13.

[23] Hoderlein, S., 2011. How Many Consumers are Rational? *Journal of Econometrics*, 164, 294–309.

[24] Hoderlein, S., J. Klemelä, and E. Mammen, 2010. Analyzing the Random Coefficient Model Nonparametrically, *Econometric Theory*, 26, 804-837.

[25] Hoderlein, S., and J. Stoye, 2014. Revealed Preferences in a Heterogeneous Population. *Review of Economics and Statistics*, 96, 197–213.

[26] Kawaguchi, K., 2012. Nonparametric Test of Axioms of Revealed Stochastic Preference. Discussion paper, London School of Economics.

[27] Kitamura, Y., and J. Stoye, 2013. Nonparametric Analysis of Random Utility Models: Testing. CeMMAP working paper 36/13.

[28] Lewbel, A., 2001. Demand Systems with and without Errors. *American Economic Review*, 91, 611-618.

[29] Manski, C.F., 2007. Partial Identification of Counterfactual Probabilities. *International Economic Review*, 48, 1393–1410.

[30] —, 2014. Identification of Income-Leisure Preferences and Evaluation of Income Tax Policy. *Quantitative Economics*, 5, 145–174.

[31] Matzkin, R., 2006. Heterogeneous Choice. In R. Blundell, W. Newey, and T. Persson (eds.), *Advances in Econometrics: Proceedings of the 9th World Congress*. Cambridge University Press.

[32] McFadden, D., 2005. Revealed Stochastic Preference: A Synthesis. *Economic Theory*, 26, 245-264.

[33] 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.* Westview Press.

[34] Polisson, M., and J.K.-H. Quah, 2013. Revealed Preference in a Discrete Consumption Space, *American Economic Journal: Microeconomics*, 5, 28-34.

[35] Rose, H., 1958. Consistency of Preferences: The Two-Commodity Case. *Review of Economic Studies*, 25, 124-125.

[36] Stoye, J., 2010. Partial Identification of Spread Parameters. *Quantitative Economics*, 1, 323-357.

[37] Varian, H., 1982. The Nonparametric Approach to Demand Analysis. *Econometrica*, 50, 945-974.