ONLINE APPENDIX
REVEALED PRICE PREFERENCE: THEORY AND EMPIRICAL ANALYSIS

RAHUL DEB[◊], YUICHI KITAMURA[†], JOHN K.-H. QUAH[‡], AND JÖRG STOYE[⋆]

MARCH 18, 2022

APPENDIX A.1. GAPP AND GARP

In this section, we first state and explain Afriat's Theorem. After that we cover a number of topics on GAPP and GARP and their relationship: augmented utility functions that lead to both properties holding in a data set (Section A.1.2); demand predictions at out-of-sample prices under GAPP and under GARP (Section A.1.3); and on reconciling differing revealed preference relations under GAPP and GARP (Section A.1.4).

A.1.1. *Afriat's Theorem*

Recall that, given a data set $\mathcal{D} = \{(p^t, x^t)\}_{t=1}^T$, a utility function $\widetilde{U} : \mathbb{R}^L_+ \to \mathbb{R}$ is said to rationalize $\mathcal{D}$ if, for all $t \in T$, we have $\widetilde{U}(x^t) \geq \widetilde{U}(x)$ for all $x \in \{x \in \mathbb{R}^L_+ : \ p^t \cdot x \leq p^t \cdot x^t\}$; in other words, x^t is the bundle that maximizes $\widetilde{U}$ among all bundles that cost $p^t \cdot x^t$ or less. Afriat's Theorem characterizes those data sets that can be rationalized in this sense. Below is the formal statement of Afriat's Theorem along with some remarks that relate this theorem to results in the paper.

Afriat's Theorem (Afriat (1967))**.** *Given a data set $\mathcal{D} = \{(p^t, x^t)\}_{t=1}^T$, the following are equivalent:*

(1) $\mathcal{D}$ can be rationalized by a locally nonsatiated utility function.
(2) $\mathcal{D}$ satisfies GARP.
(3) $\mathcal{D}$ can be rationalized by a strictly increasing, continuous, and concave utility function.

REMARK 1. That (1) implies (2) is clear, given the definition of GARP (see Section 2.2 in the main paper). The substantive part of Afriat's Theorem is the claim that (2) implies (3). Standard proofs (see, for instance, Fostel, Scarf, and Todd (2004) or Quah (2014)) work by showing that a consequence of GARP is that there exist numbers ϕ^t and $\lambda^t > 0$ (for all $t \in T$) that solve the so-called *Afriat inequalities*

$$\phi^{t'} \leq \phi^t + \lambda^t p^t \cdot (x^{t'} - x^t) \ \text{ for all } t' \neq t. \tag{A.1}$$

[◊]UNIVERSITY OF TORONTO, [†]YALE UNIVERSITY, [‡]JOHNS HOPKINS UNIVERSITY, [⋆]CORNELL UNIVERSITY.

E-mail addresses: rahul.deb@utoronto.ca, yuichi.kitamura@yale.edu, john.quah@jhu.edu, stoye@cornell.edu.

Once this is established, it is straightforward to show that

$$\widetilde{U}(x) = \min_{t \in T} \left\{ \phi^t + \lambda^t \, p^t \cdot (x - x^t) \right\} \tag{A.2}$$

rationalizes $\mathcal{D}$, with the utility of the observed consumption bundles satisfying $\widetilde{U}(x^t) = \phi^t$. The function $\widetilde{U}$ is the lower envelope of a finite number of strictly increasing affine functions, and so it is strictly increasing, continuous, and concave. A remarkable feature of this theorem is that while GARP follows simply from local nonsatiation of the utility function, it is nonetheless sufficient to guarantee that $\mathcal{D}$ is rationalized by a utility function with significantly stronger properties. Our results Theorem 1 and Theorem 2 share this feature.

REMARK 2. To be precise, GARP guarantees that there is preference $\succsim$ (i.e., a complete, reflexive, and transitive binary relation) on $\mathcal{X}$ that extends the (potentially incomplete) revealed preference relations $\succeq_x^*$ and $\succ_x^*$ in the following sense: if $x^{t'} \succeq_x^* x^t$, then $x^{t'} \succsim x^t$ and if $x^{t'} \succ_x^* x^t$ then $x^{t'} \succ x^t$. One could then proceed to show that, for *any* such preference $\succsim$, there is in turn a utility function $\widetilde{U}$ that rationalizes $\mathcal{D}$ and extends $\succsim$ (from $\mathcal{X}$ to $\mathbb{R}_+^L$) in the sense that $\widetilde{U}(x^{t'}) \geq (>) \widetilde{U}(x^t)$ if $x^{t'} \succsim (\succ) x^t$ (see Quah (2014)). This has implications on the inferences one could draw from the data. If $x^{t'} \not\succeq_x^* x^t$ (or if $x^{t'} \succeq_x^* x^t$ but $x^{t'} \not\succ_x^* x^t$) then it is *always possible* to find a preference extending the revealed preference relations such that $x^t \succ x^{t'}$ (or $x^{t'} \sim x^t$ respectively).[1] Therefore, $x^{t'} \succeq_x^* (\succ_x^*) x^t$ if and only if every locally nonsatiated utility function rationalizing $\mathcal{D}$ has the property that $\widetilde{U}(x^{t'}) \geq (>) \widetilde{U}(x^t)$.

Similarly, we show in Proposition 2 that the revealed price preference relation contains the most detailed information for welfare comparisons in our model.

REMARK 3. A feature of Afriat's Theorem that is less often remarked upon is that in fact $\widetilde{U}$, as given by (A.2), is well-defined, strictly increasing, continuous, and concave on the domain $\mathbb{R}^L$, rather than just the positive orthant $\mathbb{R}_+^L$. Furthermore,

$$x^t \in \underset{\left\{x \in \mathbb{R}^L : \, p^t \cdot x \leq p^t \cdot x^t\right\}}{\operatorname{argmax}} \widetilde{U}(x) \qquad \text{for all } t \in T. \tag{A.3}$$

In other words, $\widetilde{U}$ can be extended beyond the positive orthant and x^t remains optimal under $\widetilde{U}$ in the set $\left\{x \in \mathbb{R}^L : \; p^t \cdot x \leq p^t \cdot x^t\right\}$. (Compare (A.3) with (3).) We utilize this feature when we apply Afriat's Theorem in our proof of Theorem 1.

[1] We use $x^{t'} \sim x^t$ to mean that $x^{t'} \succsim x^t$ and $x^t \succsim x^{t'}$.

A.1.2. *Models that satisfy both GAPP and GARP*

Suppose that a data $\mathcal{D} = \{(p^t, x^t)\}_{t+1}^T$ is collected from a consumer who is maximizing an augmented utility function of the form

$$U(x, -e) = h(\widetilde{U}(x), -e), \tag{A.4}$$

where h is strictly increasing (in both its arguments) and $\widetilde{U} : \mathbb{R}_+^L \to \mathbb{R}$ is strictly increasing. In this case, obviously the data set obeys GAPP, but it must also obey GARP, because if x^t maximizes U then x^t also maximizes $\widetilde{U}$ in the set $\{x \in \mathbb{R}_+^L : p^t \cdot x \le p^t \cdot x^t\}$. Thus GAPP and GARP are not mutually exclusive properties and to say that a data set satisfies one is not to say that it violates the other; depending on the issue being studied, the analyst could exploit GAPP, or GARP, or perhaps even both in conjunction.

An interesting question worth investigating is the characterization of those data sets $\mathcal{D}$ generated by consumers who maximize an augmented utility function of the form (A.4). Such a characterization must involve a property stronger than both GAPP and GARP; indeed, related work that characterizes rationalization by weakly separable preferences in the context of the constrained-maximization model (see Quah (2014)) suggests that rationalization by an augmented utility function of the form (A.4) will involve a property *strictly* stronger than the combination of GAPP and GARP. A special case of (A.4) is, of course, the quasilinear form, where $U(x, -e) = \widetilde{U}(x) - e$. In this case, a full characterization is known and the rationalizing property is sometimes referred to as the *strong law of demand* (see Brown and Calsamiglia (2007)); obviously the strong law of demand implies both GAPP and GARP.

In our analysis of the Progresa data reported in Section 7.1, we find that 2061 out of 2488 households pass GAPP (83%), 2375 households pass GARP (95%), and 35 households (a bit more than 1%) fail both tests. Interestingly, 1983 households (80%) pass both GAPP and GARP, which is suggestive (but not conclusive) evidence that a very large proportion of households from the Progresa data could be rationalized by an augmented utility function of the form (A.4).

A.1.3. *Comparing demand predictions under GAPP and GARP*

Suppose a data set $\mathcal{D} = \{(p^t, x^t)\}_{t=1}^T$ obeys GARP. Then we know from Afriat's Theorem that there is a utility function $\widetilde{U} : \mathbb{R}_+^L \to \mathbb{R}$ for which x^t is constrained optimal, for all t. What could this model tell us about the demand at some price $\hat{p}$ that is not among the observed prices? In this model, the predicted demand also depends on the level of total expenditure on the observed goods. Suppose the expenditure is required to be some $w > 0$; then the predicted demand will be those bundles x with $\hat{p} \cdot x = w$ that are compatible with the model when combined with $\mathcal{D}$. By Afriat's Theorem, this is means that x

is a predicted demand if and only if the following conditions are satisfied: $\hat{p} \cdot x = w$ and the data set $\mathcal{D} \cup \{(\hat{p}, x)\}$ obeys GARP.

Now suppose that $\mathcal{D} = \{(p^t, x^t)\}_{t=1}^T$ also obeys GAPP. Then we know it is also compatible with the augmented utility model and we could ask what the augmented utility model would say about demand at the price $\hat{p}$. This is equivalent to identifying bundles x such that $\mathcal{D} \cup \{(\hat{p}, x)\}$ obeys GAPP. Since $\mathcal{D} \cup \{(\hat{p}, x)\}$ obeys GAPP if and only if $\mathcal{D} \cup \{(\hat{p}, \lambda x)\}$ obeys GAPP for any $\lambda > 0$ (see Section 3.1), we know that *the set of predicted demands at $\hat{p}$ forms a cone*.

Not surprisingly, these two models will typically have different predictions, even at the same expenditure level $w > 0$. To illustrate this, consider the following example.

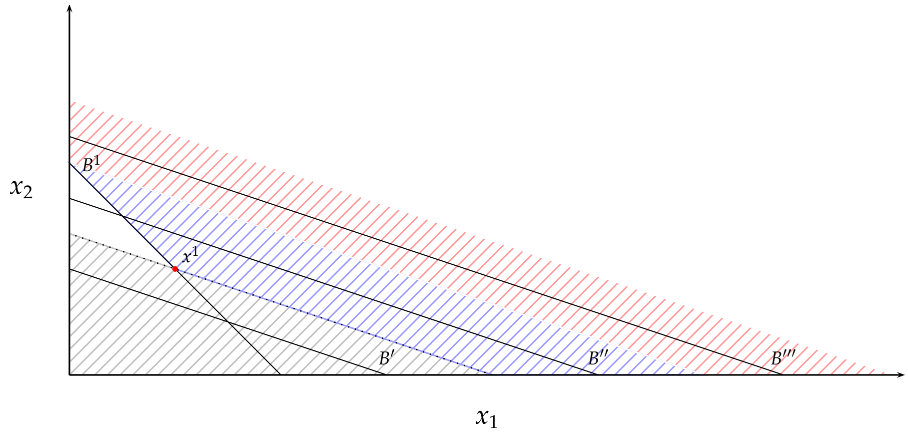
Example A.1. Suppose $\mathcal{D}$ consists of the single observation $p^1 = (1, 1)$ and $x^1 = (1, 1)$. What is the predicted demand at $\hat{p} = (1/4, 3/2)$? We study the predictions under the constrained-optimization model, with and without imposing homotheticity on the utility function, and the augmented utility model.

Consider first the constrained-optimization model. (a) Suppose that $w < \hat{p} \cdot x^1 = 7/4$; the line of points/bundles incurring this level of expenditure is depicted by B' in Figure A.1a. In this case, any bundle with $\hat{p} \cdot x = w$ will *not* be revealed preferred to x^1 and so x can be any bundle in gray shaded area without violating GARP. (b) Now suppose $w \geq \hat{p} \cdot (0, 2) = 3$; the bundles with $\hat{p} \cdot x = w$ is depicted as B''' in Figure A.1a. Then if $x \cdot \hat{p} = w$, we have $x \cdot p^1 > 2$. In other words, x^1 will never be revealed preferred to x. Once again, x can be any bundle in the red shaded area (that extends indefinitely towards the north east) without GARP being violated. (c) Lastly, we turn to the case where $w \in [7/4, 3)$; a line with bundles satisfying this property is B''. Then any bundle satisfying $\hat{p} \cdot x = w$ will be revealed preferred to x^1. So GARP requires that x^1 is *not* revealed preferred to x, that is, $p^1 \cdot x > p^1 \cdot x^1 = 2$ and therefore, all bundles in the blue shaded area will not violate GARP.

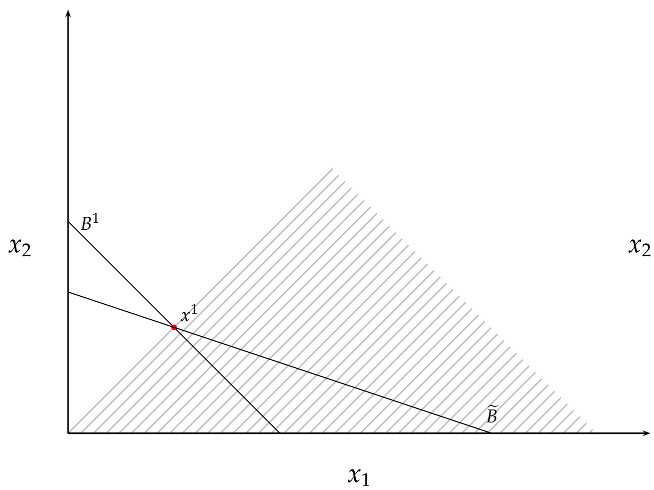
The shaded area in Figure A.1a gives the predicted demands at $\hat{p}$ using GARP.

What happens to the predictions of the constrained-maximization model when the utility function is required to be homothetic? It is well known that homothetic utility functions generate demand that is linear in cones. Therefore, for any $x \in \mathbb{R}^2_+$, the data set $\{(p^1, x^1), (\hat{p}, x)\}$ can be rationalized (in the constrained-maximization sense) by a homothetic utility function if and only if $\{(p^1, x^1), (\hat{p}, \lambda x)\}$ can also be rationalized in this sense, for any $\lambda > 0$. In other words, as in the augmented utility model, the set of predicted demands forms a cone.

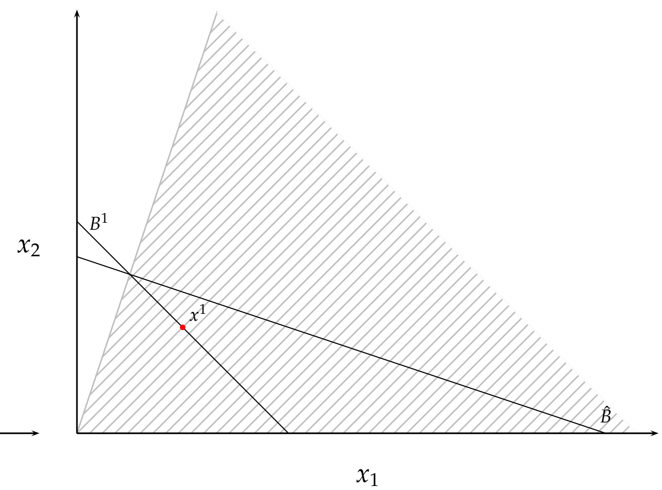
The characterization of data sets that are constrained-optimal according to some homothetic preference is given in Varian (1983), where the precise condition is known as the *homothetic axiom of revealed preference* or HARP, for short. In our simple case, to guarantee



(a) Counterfactuals using GARP



(b) Counterfactuals using HARP



(c) Counterfactuals using GAPP

FIGURE A.1. Counterfactuals with different consumption models

that $\{(p^1, x^1), (\hat{p}, \lambda x)\}$ satisfies HARP, we set $w = \hat{p} \cdot x^1$ and consider the bundles with $\hat{p} \cdot x = w$; the bundles at this expenditure level are depicted by $\widetilde{B}$ in Figure A.1b. At this expenditure level, GARP requires that x satisfies $p^1 \cdot x > p^1 \cdot x^1$ and, for any such x, we have $\{(p^1, x^1), (\hat{p}, \lambda x)\}$ satisfying HARP; in other words, the set of predicted demands is the cone generated by these bundles of x. This cone is depicted by the shaded region in Figure A.1b.

In the case of the augmented utility model, recall that if x satisfies $\hat{p} \cdot x = p^1 \cdot x^1 = 2$, then $\{(p^1, x^1), (\hat{p}, x)\}$ satisfies GAPP if and only if it satisfies GARP (see Proposition 1). The budget line with the property that $\hat{p} \cdot x = 2$ is $\hat{B}$ in Figure A.1c and, in this case, GARP (equivalently, GAPP) requires that $p^1 \cdot x > p^1 \cdot x^1 = 2$. The shaded area gives the predicted demands at $\hat{p}$. Notice that the cone in this case contains the cone in Figure A.1b, which is consistent with the fact that HARP is a stronger property than GAPP. Furthermore, the predicted demands under GAPP is neither a subset nor a superset of that under GARP, which is again unsurprising given that these two properties are not comparable.

A.1.4. *Revealed preferences under GAPP and GARP*

Both GARP and GAPP forbids the existences of strict cycles over revealed preference relations: in the case of GARP, the revealed preference relation is defined over bundles and in the case of GAPP it is defined over prices. It is entirely possible for these revealed preference relations to disagree with each other; this occurrence should not be thought of as strange, nor is it an indication that one model is better of worse compared to the other. The two conclusions apply to different objects and either, or both, of them could be interesting to the analyst.

To be precise, suppose that a data $\mathcal{D} = \{(p^t, x^t)\}_{t=1}^T$ is collected from a consumer who is maximizing an augmented utility function of the form (A.4). Such a data set will obey both GAPP and GARP. It is possible for the price p^t to be strictly revealed preferred to p^s (whether directly or indirectly) and for the bundle x^s to be revealed strictly preferred to x^t. If this occurs, is the agent better off in observation t or in observation s? The fact that p^t is revealed strictly preferred to p^s means that

$$U(x^t, -p^t \cdot x^t) > U(x^s, -p^s \cdot x^s)$$

while the fact that x^s is revealed strictly preferred to x^t means that

$$\widetilde{U}(x^t) < \widetilde{U}(x^s).$$

In other words, the consumer's augmented utility is higher in observation t than in observation s, even though her sub-utility on the observed bundles is lower in observation t; these two phenomena are not mutually exclusive.

Another observation worth making is that it is sometimes possible to conclude that an out-of-sample price $\hat{p}$ is superior to some in-sample price p^{t_1} observed in $\mathcal{D}$, even though one has no inkling what the demand will be at $\hat{p}$. Indeed, $\hat{p}$ is revealed preferred to p^{t_1} whenever $\hat{p} \cdot x^{t_1} \leq p^{t_1} \cdot x^{t_1}$ (and, more generally, this relation between $\hat{p}$ and some other in-sample price p^t can be extended via transitive closure). It follows that at the (unobserved) optimal bundle at $\hat{p}$, which we denote by $\hat{x}$, we must have

$$U(\hat{x}, -\hat{p} \cdot \hat{x}) \geq U(x^{t_1}, -\hat{p} \cdot x^{t_1}) > U(x^{t_1}, -p^{t_1} \cdot x^{t_1}).$$

This is true even though, as we know from Section A.1.3, the predicted demand at $\hat{p}$ under the augmented utility model can be an arbitrarily small or large bundle. On the other hand, without knowing the agent's expenditure level at $\hat{p}$, it is impossible to tell if the sub-utility $\widetilde{U}(\hat{x})$ is greater or lower than $\widetilde{U}(x^{t_1})$. Put another way, while GAPP may allow the observer to rank $\hat{p}$ with p^{t_1}, it is impossible to rank the subutility of the demand bundle at these two observation using GARP, without some information or assumption on the expenditure level at $\hat{p}$.

Example A.2. Suppose $\mathcal{D}$ consists of two observations,

$$\begin{aligned}(p^{t_1}, x^{t_1}) &= ((2,2),(2,2)) \text{ and} \\ (p^{t_2}, x^{t_2}) &= ((1,1),(1,1)).\end{aligned}$$

It is straightforward to check that this data set can be generated by a consumer maximizing

$$U(x, -e) = \widetilde{U}(x) - f(e),$$

for strictly increasing functions $\widetilde{U}$ and f. Clearly, p^{t_2} is revealed preferred to p^{t_1} and x^{t_1} is revealed preferred to x^{t_2}. In this case, the consumer's augmented utility is higher at t_2 compared to t_1, even though her sub-utility on the observed goods is lower at t_2 compared to t_1.

Now suppose the data consists of just the observation (p^{t_1}, x^{t_1}). Obviously, we can still conclude that the consumer prefers $\hat{p} = (1,1)$ to p^{t_1} and derives greater augmented utility from $\hat{p}$ than from p^{t_1}. However, nothing can be said about the consumer's subutility without further information on expenditure. If the expenditure is lower than the expenditure at t_1, which is 8, then the subutility achieved at $\hat{p}$ must be lower than the subutility of x^1 and if the expenditure is higher than 8, then the sub-utility achieve must be lower than that of x^{t_1}.

Appendix A.2. Proof of Proposition 2

(1) We have already shown the 'only if' part of this claim, so we need to show the 'if' part holds. From the proof of Theorem 1, we know that for a large M, it is the case that $p^t \succeq_p p^{t'}$ if and only if $(x^t, M - p^t \cdot x^t) \succeq_x (x^{t'}, M - p^{t'} \cdot x^{t'})$ and hence $p^t \succeq_p^* p^{t'}$ if and only if $(x^t, M - p^t \cdot x^t) \succeq_x^* (x^{t'}, M - p^{t'} \cdot x^{t'})$. If $p^t \not\succeq_p^* p^{t'}$, then $(x^t, M - p^t \cdot x^t) \not\succeq_x^* (x^{t'}, M - p^{t'} \cdot x^{t'})$ and hence there is a utility function $\widetilde{U} : \mathbb{R}_+^{L+1} \to \mathbb{R}$ rationalizing the augmented data set $\widetilde{D}$ such that $\widetilde{U}(x^t, M - p^t \cdot x^t) < \widetilde{U}(x^{t'}, M - p^{t'} \cdot x^{t'})$ (see Remark 2 in Section A.1.1). This in turn implies that the augmented utility function U (as defined by (5)), has the property that $U(x^t, -p^t \cdot x^t) < U(x^{t'}, -p^{t'} \cdot x^{t'})$ or, equivalently, $V(p^t) < V(p^{t'})$.

(2) Given part (1), we need only show that if $p^t \succeq_p^* p^{t'}$ but $p^t \not\succ_p^* p^{t'}$, then there is some augmented utility function U such that $U(x^t, -p^t \cdot x^t) = U(x^{t'}, -p^{t'} \cdot x^{t'})$. To see that this holds, note that if $p^t \succeq_p^* p^{t'}$ but $p^t \not\succ_p^* p^{t'}$, then $(x^t, M - p^t \cdot x^t) \succeq_x^* (x^{t'}, M - p^{t'} \cdot x^{t'})$ but $(x^t, M - p^t \cdot x^t) \not\succ_x^* (x^{t'}, M - p^{t'} \cdot x^{t'})$. In this case there is a utility function $\widetilde{U} : \mathbb{R}_+^{L+1} \to \mathbb{R}$ rationalizing the augmented data set $\widetilde{D}$ such that $\widetilde{U}(x^t, M - p^t \cdot x^t) = \widetilde{U}(x^{t'}, M - p^{t'} \cdot x^{t'})$.

This in turn implies that the augmented utility function U (as defined by (5)) satisfies $U(x^t, -p^t \cdot x^t) = U(x^{t'}, -p^{t'} \cdot x^{t'})$ and so $V(p^t) = V(p^{t'})$. ■

APPENDIX A.3. PRICE INDICES TO DEFLATE NOMINAL EXPENDITURE

In this section, we build on the discussion in Section 3.5. Suppose that, at observation t, the consumer chooses (x^t, y^t) to maximize $\widetilde{U}(x,y)$, subject to $p^t \cdot x^t + q^t \cdot y^t \leq M^t$. We are interested in the conditions under which there is an index k^t, depending on the prices of the outside goods, such that the deflated data $\{(p^t/k^t, x^t)\}_{t=1}^T$ obeys GAPP (and hence can be rationalized as maximizing an augmented utility function). In the main paper, we explained that this holds if prices of the outside goods move up and down proportionately (so there is no change to their prices relative to each other). When relative prices *are* allowed to change, it is still possible to obtain a deflator k^t guaranteeing that $\{(p^t/k^t, x^t)\}_{t=1}^T$ obeys GAPP, but stronger assumptions will have to be imposed on the utility function $\widetilde{U}$. We outline a set of sufficient conditions for this to hold.

Suppose that the outside goods are weakly separable from the observed goods, so the overall utility function has the form $\widetilde{U}(x, \tilde{u}(y))$, where $\tilde{u}(y)$ is the sub-utility of the bundle y of outside goods. Furthermore, we assume that $\tilde{u}$ has an indirect utility $\tilde{v}$ of the following form:

$$\tilde{v}(q,m) = h\left(\frac{m}{f(q)} + b(q), g_1(q), g_2(q), \ldots, g_N(q)\right)$$

where f, b, $g_1, \ldots, g_N$ are all real-valued functions of the prices q of the outside goods, and m is the expenditure devoted to those goods. This formulation covers a number of standard functional forms used in empirical analysis. If $\tilde{v}(q,m) = m/f(q)$ where f is one-homogeneous then the preference it generates is homothetic; if $\tilde{v}(q,m) = (m/f(q)) + b(q)$, where b is zero-homogeneous, then we obtain the Gorman polar form (see Gorman (1961)). Another example is the form

$$\ln \tilde{v}(q,m) = \left\{\left[\frac{\ln m - \ln f(q)}{g_1(q)}\right]^{-1} + g_2(q)\right\}^{-1} \tag{A.5}$$

where g_1 and g_2 are zero-homogeneous functions. If $g_2 \equiv 0$, the form (A.5) generates the Price Invariant Generalized Logarithmic (PIGLOG) demand system (Muellbauer, 1976); if further functional form restrictions are imposed on f and g_1, we obtain the Almost Ideal Demand System (AIDS) of Deaton and Muellbauer (1980). The Quadratic Almost Ideal Demand System (QUAIDS) is a generalization of AIDS that has greater flexibility to model empirically relevant Engel curves (see Banks, Blundell, and Lewbel (1997)); it is a special case of (A.5) with functional form restrictions on f, g_1, and g_2.

We assume that the consumer's total wealth M^t varies with t in such a way that, should the consumer devote all of this wealth to the unobserved goods, then her utility is constant. This captures the idea that the consumer's *real wealth* (as measured by the indirect utility function v) is unchanged across observations. While we permit prices of the unobserved goods to change, we require that they change in such a way that $g_1(q^t)$, $g_2(q^t), \ldots, g_N(q^t)$ remain constant at $\bar{g}_1, \bar{g}_2 \ldots, \bar{g}_N$ (respectively) for all t. Given the form of $\tilde{v}$, this implies that $(M^t/f(q^t)) + b(q^t)$ is constant for all t; let this constant be C. Thus we can think of the consumer as choosing (x, c) to maximize $\widetilde{U}(x, \tilde{v}(c, \bar{g}_1, \bar{g}_2, \ldots, \bar{g}_N))$ subject to $p^t \cdot x + (c - b(q^t))f(q^t) \leq (C - b(q^t))f(q^t)$. This inequality can be written as

$$\frac{p^t \cdot x}{f(q^t)} + c \leq C.$$

It follows that the data set $\{(p^t/f(q^t), x^t)\}_{t=1}^T$ will obey GAPP.

APPENDIX A.4. NONLINEAR PRICING, THE RATIONALITY INDEX, AND RELATED TOPICS

In this section, we formulate and prove a rationalization result that allows for both imperfect rationalization and nonlinear pricing. This result generalizes Theorem 2 and Theorem 1 by allowing for imperfect rationality. We explain how this result is crucial in helping us to calculate the rationality index (introduced in Section 3.4) and other variations on that index that provide a measure of departures from exact rationality. We also use this result to show that the bounds on the compensating and equivalent variations obtained in Section 3.3 are tight.

A.4.1. $\overline{\vartheta}$-*rationalization*

We are in the setting of Section 4. The consumer chooses her consumption from the space $X \subseteq \mathbb{R}^L_+$. A price system is a map $\psi : X \to \mathbb{R}_+$, where $\psi(x)$ is the cost of purchasing $x \in X$. Let $\overline{\vartheta} = (\vartheta^1, \vartheta^2, \ldots, \vartheta^T) \in (0, 1]^T$. An augmented utility function $U : X \times \mathbb{R}_+ \to \mathbb{R}$ provides a $\overline{\vartheta}$-rationalization of a data set $\mathcal{D} = \{(\psi^t, x^t)\}_{t=1}^T$ if, at each observation t,

$$U(x^t, -\psi^t(x^t)) \geq U(x, -(\vartheta^t)^{-1}\psi^t(x)) \text{ for all } x \in X.$$

Note that this definition of $\overline{\vartheta}$-rationalization generalizes the notion introduced in Section 3.4, which can be thought of as the special case where $\vartheta^t = \vartheta^{t'}$ for all t, $t' \in T$. The context here is also more general since we allow for nonlinear pricing (as introduced in Section 4). Obviously, if a data set can be exactly rationalized then it is $\overline{\vartheta}$-rationalized with $\overline{\vartheta} = (1, 1, \ldots, 1)$; note also that if a data set can be $\overline{\vartheta}$-rationalized then it can also be $\overline{\vartheta}'$-rationalized for $\overline{\vartheta}' < \overline{\vartheta}$. A consumer whose observations cannot be exactly rationalized but can be $\overline{\vartheta}$-rationalized for some $\overline{\vartheta} < (1, 1, \ldots, 1)$ exhibits limited rationality in the sense discussed in Section 3.4.

The calculation of the rationality index hinges on our ability to ascertain whether a data set $\mathcal{D}$ has a $\overline{\vartheta}$-rationalization for a given $\overline{\vartheta}$. It is possible to characterize those data sets that can be $\overline{\vartheta}$-rationalized using a modified version of the GAPP test, as we now explain.

Let $\overline{\vartheta} \in (0,1]^T$. Define the relations $\succeq_{p,\overline{\vartheta}}$ and $\succ_{p,\overline{\vartheta}}$ in the following way:

$$\psi^{t'} \succeq_{p,\overline{\vartheta}} \psi^t \textit{ if } \psi^{t'}(x^t) \leq \vartheta^{t'}\psi^t(x^t) \textit{ and } \psi^{t'} \succ_{p,\overline{\vartheta}} \psi^t \textit{ if } \psi^{t'}(x^t) < \vartheta^{t'}\psi^t(x^t).$$

Denote the transitive closure of $\succeq_{p,\overline{\vartheta}}$ by $\succeq^*_{p,\overline{\vartheta}}$. Obviously these definitions generalize the ones given for revealed preference relations over prices provided in Section 4.

The data set $\mathcal{D}$ obeys $\overline{\vartheta}$-*GAPP* if

there do not exist observations $t, t' \in T$ such that $\psi^{t'} \succeq^*_{p,\overline{\vartheta}} \psi^t$ and $\psi^t \succ_{p,\overline{\vartheta}} \psi^{t'}$.

The next result states that $\overline{\vartheta}$-GAPP characterizes $\overline{\vartheta}$-rationalization.

Theorem A.4.1. *A data set $\mathcal{D} = \{(\psi^t, x^t)\}_{t=1}^T$ can be $\overline{\vartheta}$-rationalized by an augmented utility function for some $\overline{\vartheta} \in (0,1]^T$ if and only if it satisfies $\overline{\vartheta}$-GAPP.*

REMARK 1. This theorem states, in particular, that $\mathcal{D} = \{(\psi^t, x^t)\}_{t=1}^T$ can be rationalized by an augmented utility function if and only if it satisfies GAPP, which corresponds to the special case where $\overline{\vartheta} = (1, 1, \ldots, 1)$. So it covers the first claim in Theorem 2 (the part before "Furthermore,…") and also the equivalence of statements (1) and (2) in Theorem 1. For the proof of the second claim in Theorem 2 see the end of this subsection. Unlike the proof we gave of Theorem 1 in the main body of the paper, our proof of Theorem A.4.1 does not appeal to Afriat's Theorem, though it is clearly inspired by it. In particular, we show that $\overline{\vartheta}$-GAPP implies that there is a solution to a system of linear inequalities (see Lemma A.1 below), analogous to the so-called Afriat inequalities usually derived in the proof of Afriat's Theorem and then use those inequalities to explicitly construct a piecewise linear augmented utility function that rationalizes the data.

REMARK 2. Note that checking whether or not $\overline{\vartheta}$-GAPP holds for a given $\overline{\vartheta}$ is computationally undemanding: the relations $\succeq_{p,\overline{\vartheta}}$ and $\succ_{p,\overline{\vartheta}}$ can be easily constructed; once this has been obtained, we can apply Warshall's algorithm to compute the transitive closure of the revealed preference relations and then check for violations of $\overline{\vartheta}$-GAPP.

REMARK 3. Suppose we impose the mild restriction that every bundle that is an observed choice has a strictly positive value under any of the other price observation, that is, $\psi^{t'}(x^t) > 0$ whenever $\psi^{t'} \neq \psi^t$. Then we can choose sufficiently small $\vartheta > 0$ so that $\psi^{t'}(x^t) > \vartheta\psi^t(x^t)$ whenever $\psi^{t'} \neq \psi^t$. If we let $\overline{\vartheta} = (\vartheta, \ldots, \vartheta)$, then $\mathcal{D}$ must obey $\overline{\vartheta}$-GAPP simply because the relation $\succ_{p,\overline{\vartheta}}$ is empty. Thus every data set is $\overline{\vartheta}$-rationalizable for $\overline{\vartheta}$ sufficiently close to zero.

Proof of Theorem A.4.1. Suppose $\mathcal{D}$ can be $\overline{\vartheta}$-rationalized by an augmented utility function for some $\overline{\vartheta} \in (0,1]^T$. In that case, if $\psi^{t'} \succeq_{p,\overline{\vartheta}} \psi^t$, then $\psi^{t'}(x^t) \leq \vartheta^{t'}\psi^t(x^t)$ and so

$$U(x^{t'}, -\psi^{t'}(x^{t'})) \geq U(x^t, -(\vartheta^{t'})^{-1}\psi^{t'}(x^t)) \geq U(x^t, -\psi^t(x^t)), \tag{A.6}$$

where the first inequality follows from the (imperfect) optimality of $x^{t'}$ and the second from the property that U is strictly decreasing in expenditure. It follows that if $\psi^{t'} \succeq^*_{p,\overline{\vartheta}} \psi^t$, then $U(x^{t'}, -\psi^{t'}(x^{t'})) \geq U(x^t, -\psi^t(x^t))$. Similarly, if $\psi^{t'} \succ_{p,\overline{\vartheta}} \psi^t$, then $\psi^{t'}(x^t) < \vartheta^{t'}\psi^t(x^t)$ and we obtain $U(x^{t'}, -\psi^{t'}(x^{t'})) > U(x^t, -\psi^t(x^t))$ since the second inequality in (A.6) will now be strict. It is then clear that we cannot simultaneously have $\psi^{t'} \succeq^*_{p,\overline{\vartheta}} \psi^t$, and $\psi^t \succ_{p,\overline{\vartheta}} \psi^{t'}$, which establishes $\overline{\vartheta}$-GAPP.

Conversely, suppose that $\mathcal{D}$ obeys $\overline{\vartheta}$-GAPP. Then there is a complete preorder $\succsim$ defined on the set $\{p^t\}_{t\in T}$ that extends $\succeq_{p,\overline{\vartheta}}$ and $\succ_{p,\overline{\vartheta}}$ in the sense that such $\psi^{t'} \succsim \psi^t$ if $\psi^{t'} \succeq^*_{p,\overline{\vartheta}} \psi^t$ and $\psi^{t'} \succ \psi^t$ if $\psi^{t'} \succ_{p,\overline{\vartheta}} \psi^t$, where $\succ$ is the asymmetric part of $\succsim$. We first prove the following lemma.

Lemma A.1. *Suppose $\mathcal{D}$ obeys $\overline{\vartheta}$-GAPP and let $\succsim$ be a complete preorder that extends $\succeq_{p,\overline{\vartheta}}$ and $\succ_{p,\overline{\vartheta}}$. Then there are numbers ϕ^t and $\lambda^t > 0$ (for $t = 1, 2, \ldots, T$) with the following properties:*

(a) $\phi^{t'} > \phi^t$ *if* $\psi^{t'} \succ \psi^t$;
(b) $\phi^{t'} = \phi^t$ *if* $\psi^{t'} \sim \psi^t$; *and*
(c) $\phi^{t'} \leq \phi^t + \lambda^t(\psi^t(x^{t'}) - \vartheta^t\psi^{t'}(x^{t'}))$ *for all* $t \neq t'$.

Proof. Let $z^{ij} = \psi^i(x^j) - \vartheta^i\psi^j(x^j)$ for i, $j \in T$. Note that, for $i \neq j$, $z^{ij} < 0$ implies that $\psi^i \succ \psi^j$ and $z^{ij} \leq 0$ implies that $\psi^i \succsim \psi^j$. We shall explicitly construct ϕ^t and $\lambda^t > 0$ that satisfy the required conditions. With no loss of generality, suppose that $\psi^{t+1} \succsim \psi^t$ for $t = 1, 2, \ldots, T-1$.

First, choose ϕ^1 to be any number and λ^1 to be any strictly positive number. Suppose $\psi^2 \succ \psi^1$. Then $\min_{j>1} z^{1j} > 0$, because if $z^{1j'} = \psi^1(x^{j'}) - \vartheta^1\psi^{j'}(x^{j'}) \leq 0$ for some $j' > 1$, then $\psi^1 \succsim \psi^{j'}$, which is a contradiction. So there is ϕ^2 such that

$$\phi^1 < \phi^2 < \min_{j>1}\{\phi^1 + \lambda^1 z^{1j}\}. \tag{A.7}$$

If $\psi^2 \sim \psi^1$ then $\min_{j>1} z^{1j} \geq 0$ because if $z^{1j'} = \psi^1(x^{j'}) - \vartheta\psi^{j'}(x^{j'}) < 0$ for some $j' > 1$, then $\psi^1 \succ \psi^{j'}$, which is a contradiction. Setting $\phi^2 = \phi^1$, we obtain

$$\phi^1 = \phi^2 \leq \min_{j>1}\{\phi^1 + \lambda^1 z^{1j}\}. \tag{A.8}$$

We claim that there is $\lambda^2 > 0$ such that

$$\phi^1 \leq \phi^2 + \lambda^2 z^{21}.$$

Clearly this inequality holds if $z^{21} \geq 0$. If $z^{21} = \psi^2(x^1) - \vartheta^2\psi^1(x^1) < 0$, then $\psi^2 \succ \psi^1$; this implies that $\phi^1 < \phi^2$ and thus the inequality holds for λ^2 sufficiently small.

We now go on to choose ϕ^3 and λ^3. Since $\psi^j \succsim \psi^i$ for all $j > 2$ and $i = 1, 2$, we obtain $z^{ij} \geq 0$. Consider two cases: when $\psi^3 \succ \psi^2 \succsim \psi^1$ and $\psi^3 \sim \psi^2 \succsim \psi^1$. In the former case, both $\min_{j>2} z^{1j} > 0$ and $\min_{j>2} z^{2j} > 0$. Therefore

$$\phi^2 < \min_{j>2}\{\phi^2 + \lambda^2 z^{2j}\}.$$

If $\phi^2 = \phi^1$, obviously we also have

$$\phi^2 < \min_{j>2}\{\phi^1 + \lambda^1 z^{1j}\};$$

this inequality also holds if $\phi^2 > \phi^1$ since in that case (A.7) holds. It follows that we can find ϕ^3 such that

$$\phi^2 < \phi^3 < \min\left\{\min_{j>2}\{\phi^1 + \lambda^1 z^{1j}\}, \min_{j>2}\{\phi^2 + \lambda^2 z^{2j}\}\right\}.$$

We turn to the case where $\psi^3 \sim \psi^2 \succsim \psi^1$. It follows from (A.7) and (A.8) that $\phi^2 \leq \min_{j>2}\{\phi^1 + \lambda^1 z^{2j}\}$. We also know that $z^{2j} \geq 0$ for all $j > 2$. Therefore, we can choose ϕ^3 such that

$$\phi^2 = \phi^3 \leq \min\left\{\min_{j>2}\{\phi^1 + \lambda^1 z^{1j}\}, \min_{j>2}\{\phi^2 + \lambda^2 z^{2j}\}\right\}.$$

Now choose $\lambda^3 > 0$ sufficiently small so that

$$\phi^i \leq \phi^3 + \lambda^3 z^{3i} \text{ for } i = 1, 2.$$

Clearly that this inequality holds for any $\lambda^3 > 0$ if $z^{3i} \geq 0$. If $z^{3i} < 0$ then $\psi^3 \succ \psi^i$, in which case $\phi^3 > \phi^i$ and the inequality will be satisfied for λ^3 sufficiently small.

Repeating this argument, we choose ϕ^t (for $t \leq T - 1$) such that if $\psi^t \succ \psi^{t-1}$ then

$$\phi^{t-1} < \phi^t < \min_{s \leq t-1}\left\{\min_{j>t-1}\{\phi^s + \lambda^s z^{sj}\}\right\} \tag{A.9}$$

and if $\psi^t \sim \psi^{t-1}$ then

$$\phi^{t-1} = \phi^t \leq \min_{s \leq t-1}\left\{\min_{j>t-1}\{\phi^s + \lambda^s z^{sj}\}\right\}; \tag{A.10}$$

and $\lambda^t > 0$ (for $t = 2, 3, \ldots, T$) such that

$$\phi^i \leq \phi^t + \lambda^t z^{ti} \text{ for } i \leq t - 1. \tag{A.11}$$

For a fixed t', (A.9) and (A.10) guarantee that $\phi^{t'} \leq \phi^t + \lambda^t z^{tt'}$ for $t < t'$ while (A.11) guarantees that this inequality holds for $t > t'$. So we have found λ^t and ϕ^t to obey condition (c), while the first two conditions hold by construction. ■

We now return to the proof that (2) implies (3). Let $\succsim$ be a complete preorder that extends $\succeq_{p,\overline{\vartheta}}$ and $\succ_{p,\overline{\vartheta}}$ and let the numbers ϕ^t and $\lambda^t > 0$ (for $t = 1, 2, \ldots, T$) satisfy properties (a) – (c) in Lemma A.1. Define the function $U : X \times \mathbb{R}_- \to \mathbb{R}$ by

$$U(x, -e) = \min_{t \in T} \{\phi^t + \lambda^t(\psi^t(x) - \vartheta^t e)\}. \tag{A.12}$$

This function is an augmented utility function since it is strictly increasing in the last argument. We claim that this function also satisfies the property that, at each $t \in T$,

$$U(x^t, -\psi^t(x^t)) \geq U(x, -(\vartheta^t)^{-1}\psi^t(x)) \text{ for all } x \in X.$$

Indeed, at a given observation s, for any $t \neq s$, we have $\phi^t + \lambda^t(\psi^t(x^s) - \vartheta^t\psi^s(x^s)) \geq \phi^s$ by condition (c); furthermore, $\phi^s + \lambda^s(\psi^s(x^s) - \vartheta^s\psi^s(x^s)) \geq \phi^s$ since $\lambda^s > 0$ and $\vartheta^s \in (0, 1]$. Therefore, $U(x^s, -\psi^s(x^s)) \geq \phi^s$. On the other hand, by the definition of U,

$$U(x, -(\vartheta^s)^{-1}\psi^s(x)) \leq \phi^s + \lambda^s(\psi^s(x) - \psi^s(x)) \leq \phi^s.$$

So $U(x^s, -\psi^s(x^s)) \geq U(x, -\vartheta^{-1}\psi^s(x))$ for all x. ■

The augmented utility function U at the price system ψ induces an indirect utility given by $V(\psi) = \max_{x \in X} U(x, -\psi(x))$. In the case where GAPP holds and exact rationalization is possible, one could also choose the rationalizing utility function U so that its indirect utility V agrees with any ordering over $\{\psi^t\}_{t=1}^T$ that is consistent with the revealed preference relations. (Note that this feature is also present in Afriat's Theorem; see Remark 2 in Section A.1.1.) The following result is used in Section A.5.

Theorem A.4.2. *Suppose the data set $\mathcal{D} = \{(\psi^t, x^t)\}_{t=1}^T$ obeys GAPP and let $\succsim$ be a complete preorder on $\{\psi^t\}_{t=1}^T$ that extends $\succeq_p$ and $\succ_p$. Then there is an augmented utility function U : $X \times \mathbb{R}_- \to \mathbb{R}$ that rationalizes $\mathcal{D}$ such that $V(\psi^{t'}) = V(\psi^t)$ if $\psi^{t'} \sim \psi^t$ and $V(\psi^t) > V(\psi^t)$ if $\psi^{t'} \succ \psi^t$ (where $\sim$ and $\succ$ are the symmetric and asymmetric parts of $\succsim$).*

Proof. From the proof of Theorem A.4.1, we know that $U(x, -e)$ as given by (A.12) (with $\theta^t = 1$ for all t) rationalizes $\mathcal{D}$. We can then conclude that $V(\psi^t) = U(x^t, -\psi(x^t)) = \phi^t$ because $\phi^t \leq \phi^{t'} + \lambda^{t'}(\psi^{t'}(x^t) - \psi^t(x^t))$ from part (c) of Lemma A.1. Finally, V satisfies the required properties because of (a) and (b) in Lemma A.1. ■

We end this subsection with the proof of Theorem 2; this result is obtained as a corollary of Theorem A.4.1.

Proof of Theorem 2. Choosing $\overline{\vartheta} = (1, 1, \ldots, 1)$, Theorem A.4.1 states, in particular, that $\mathcal{D} = \{(\psi^t, x^t)\}_{t=1}^T$ can be rationalized by an augmented utility function if and only if it satisfies GAPP. It remains for us to show that, under assumptions (i), (ii), and (iii), this utility function could be extended to one defined on a closed set Y containing X and that is increasing in x_K. We know from the proof of Theorem 2 that the function $U : X \to \mathbb{R}$ given by

$$U(x, -e) = \min_{t \in T}\{\phi^t + \lambda^t(\psi^t(x) - e)\}.$$

rationalizes the data (see A.12). It suffices to show that each function ψ^t, which is defined on X could be extended to a continuous function on Y that is strictly increasing in x_K, in which case we could correspondingly extend U and the extension would be continuous and strictly increasing in x_K (since $\lambda^T > 0$).

That ψ^t admits such an extension is guaranteed by (i), (ii), and (iii). A quick way of arriving at this conclusion is to appeal to Levin's Theorem, which is a version of Szpilrajn's Theorem for closed preorders (see Nishimura, Ok, and Quah (2017) for a proof of Levin's Theorem). Since ψ^t is continuous, it induces a closed preorder $\succsim'$ on X and therefore also on Y.[2] For $K \subset L$, let $\geq_K$ be the partial order on Y such that, for x' and x in $\mathbb{R}^L$, we have $x' \geq_K x$ if $x'_i \geq x_i$ for all $i \in K$ and $x'_i = x_i$ for $i \notin K$. It is straightforward to check that, for any number M, the set

$$\{x \in Y : \text{there is } \tilde{x} \in X \text{ with } \tilde{x} \geq_K x \text{ and } M \geq \psi^t(\tilde{x})\}$$

is a compact set in Y. (Recall that Y is closed, contains X, and is contained in $\mathbb{R}^L_+$.) Using this property, one could check that $\succsim''$, defined as the transitive closure of $\succsim'$ and $\geq_K$, is also a closed prorder on Y. Levin's Theorem then guarantees that there is a *complete* and closed preorder $\succsim$ on Y that extends $\succsim''$ and has a continuous representation $V : Y \to \mathbb{R}$. In particular, V must be strictly increasing in x_K and satisfies the following property: $V(x') \geq (>)\, V(x)$ if $\psi^t(x') \geq (>)\, \psi^t(x)$, for $x', x \in X$. Furthermore, our assumptions guarantee that that $\{V(x) : x \in X\} \subseteq \mathbb{R}$ is a closed set. These properties guarantee that we could choose a strictly increasing transformation h defined on the range of V, i.e., the set $\{V(x) : x \in Y\}$, so that $h(V(x)) = \psi^t(x)$ for all $x \in X$. Therefore the function $h \circ V : Y \to \mathbb{R}$ is a continuous extension of $\psi^t : X \to \mathbb{R}$ that is strictly increasing in x_K. ■

A.4.2. *Rationality indices and their computation*

Given a data set $\mathcal{D} = \{(\psi^t, x^t)\}_{t=1}^T$, we know that it admits a $(\vartheta, \vartheta, \ldots, \vartheta)$-rationalization for some $\vartheta > 0$ (see Remark 3 following Theorem A.4.1). This guarantees that the *rationality index*, given by

$$\vartheta^* = \sup\{\vartheta \in (0, 1] : \mathcal{D} \text{ has a } (\vartheta, \vartheta, \ldots, \vartheta)\text{-rationalization}\},$$

is well-defined. Note that this definition generalizes the definition provided in Section 3.4 of the main paper, which applies to the linear price environment. A data set that can

[2] A preorder $\succsim'$ defined on a set X is closed if $\{(a, b) \in X \times X : a \succsim' b\}$ is a closed subset of $X \times X$.

be rationalized exactly has a rationality index of 1 and we could use the closeness of ϑ^* to 1 as a measure of the data set's closeness to exactly rationality.

Given the characterization of $\overline{\vartheta}$-rationality stated in Theorem A.4.1, we also have

$$\vartheta^* = \sup\{\vartheta \in (0,1] : \mathcal{D} \text{ satisfies } (\vartheta, \vartheta, \ldots, \vartheta)\text{-GAPP}\}. \tag{A.13}$$

This identity provides us with a practical way of calculating ϑ^*. Indeed, ϑ^* can be obtained through a binary search algorithm that works as follows. We first set the lower and upper bounds on ϑ^* to be $\vartheta^L = 0$ and $\vartheta^H = 1$. We then check (by checking $\overline{\vartheta}$-GAPP) whether the data set passes or fails the test at $\vartheta = (\vartheta^L + \vartheta^H)/2$ (to be precise, at $\overline{\vartheta} = (\vartheta, \vartheta, \ldots, \vartheta)$); if it passes the test, then we update both ϑ^* and its lower bound to $(\vartheta^L + \vartheta^H)/2$; if it fails the test, then we update ϑ^* to ϑ^L and the upper bound on ϑ^* to $(\vartheta^L + \vartheta^H)/2$. We then repeat the procedure, selecting and testing the new midpoint of the updated lower and upper bounds. The algorithm terminates when the lower and upper bounds are sufficiently close.

There are other plausible variations on the rationality index, based on the way one aggregates ϑ^t across observations. Let $F : (0,1]^T \to \mathbb{R}_+$ be any weakly increasing function taking nonnegative values such that $F(1,1,\ldots,1) = 1$. We can then construct a generalized rationality index

$$F^* = \sup\{F(\overline{\vartheta}) : \mathcal{D} \text{ has a } \overline{\vartheta}\text{-rationalization}\}.$$

The rationality index ϑ^* corresponds to the case where F is defined by

$$F(\overline{\vartheta}) = \min\{\vartheta^1, \vartheta^2, \ldots, \vartheta^T\}.$$

As an alternative to this, one could choose

$$F(\overline{\vartheta}) = 1 - \sqrt{(1-\vartheta^1)^2 + (1-\vartheta^2)^2 + \ldots + (1-\vartheta^T)^2},$$

which leads to a measure of rationality based on the sum of square differences from the case of exact rationality (where $\overline{\vartheta} = (1,1,\ldots,1)$).

Computing these generalized rationality indices can be more demanding than computing the (basic) rationality index ϑ^* since in searching for those values of $\overline{\vartheta}$ that $\overline{\vartheta}$-rationalizes the data and maximizes $F(\overline{\vartheta})$, we would not in general be able to confine ourselves to the the case where $\vartheta^t = \vartheta^{t'}$ for all t, t'. In the literature on measuring GARP violations, there are indices, such as the one proposed by Varian (1990), that involve solving a maximization problem with the same mathematical structure. (In that case the problem is to find the best way to break up revealed preference cycles over consumption bundles rather than over price vectors.) Algorithms that have been devised to compute Varian's index (see Halevy, Persitz, and Zrill (2018) and Polisson, Quah, and Renou (2020)) can also be used to compute F^*.

A.4.3. $\overline{\vartheta}$-*GAPP and* $\overline{\vartheta}$-*GARP*

We confine our discussion to the environment where prices are linear, so the data set has the form $\mathcal{D} = \{(p^t, x^t)\}_{t=1}^T$. Let $\overline{\vartheta} \in (0,1]^T$. We say that a utility function $\widetilde{U} : \mathbb{R}^L_+ \to \mathbb{R}$ $\overline{\vartheta}$-*rationalizes* $\mathcal{D}$ *in the sense of Afriat* if $\widetilde{U}(x^t) \geq \widetilde{U}(x)$ for all $x \in B^t_{\overline{\vartheta}}$, where

$$B^t_{\overline{\vartheta}} = \{x \in \mathbb{R}^L_+ : p^t \cdot x \leq \vartheta^t p^t \cdot x^t\}.$$

$\overline{\vartheta}$-rationalization in this sense admits a characterization similar to the one we gave for $\overline{\vartheta}$-rationalization in the augmented utility model.

Define the relations $\succeq_{x,\overline{\vartheta}}$ and $\succ_{x,\overline{\vartheta}}$ on the set $\{x^t\}_{t=1}^T$ in the following way:

$x^{t'} \succeq_{x,\vartheta} x^t$ *if* $p^{t'} \cdot x^t \leq \vartheta^{t'} p^{t'} \cdot x^{t'}$ *and* $x^{t'} \succ_{x,\vartheta} x^t$ *if* $p^{t'} \cdot x^t < \vartheta^{t'} p^{t'} \cdot x^{t'}$

Denote the transitive closure of $\succeq_{x,\overline{\vartheta}}$ by $\succeq^*_{x,\overline{\vartheta}}$. Obviously these definitions generalize the ones given for the revealed preference relations over bundles (see Section 2.2 of the main paper). With these definitions in place, we can also generalize the definition of GARP. We say that the data set $\mathcal{D}$ obeys $\overline{\vartheta}$-*GARP* if

there do not exist observations $t, t' \in T$ such that $x^{t'} \succeq^*_{x,\overline{\vartheta}} x^t$ and $x^t \succ_{x,\overline{\vartheta}} x^{t'}$.

It is straightforward to show that $\overline{\vartheta}$-GARP is necessary for the $\overline{\vartheta}$-rationalization of $\mathcal{D}$ (in the sense of Afriat) by a locally nonsatiated utility function $\widetilde{U} : \mathbb{R}^L_+ \to \mathbb{R}$. It is also known (see Halevy, Persitz, and Zrill (2018)) that $\overline{\vartheta}$-GARP is sufficient to guarantee the $\overline{\vartheta}$-rationalization of $\mathcal{D}$ (in Afriat's sense) by a continuous, strictly increasing and concave utility function $\widetilde{U} : \mathbb{R}^L_+ \to \mathbb{R}$.[3] By definition, the critical cost efficiency index c^* satisfies

$$c^* = \sup\{\vartheta \in (0,1] : \mathcal{D} \text{ has a } (\vartheta, \vartheta, \dots, \vartheta)\text{-rationalization in the sense of Afriat}\}$$

and since $\overline{\vartheta}$-rationalization in Afriat's sense can be characterized by $\overline{\vartheta}$-GARP, we obtain

$$c^* = \sup\{\vartheta \in (0,1] : \mathcal{D} \text{ satisfies } (\vartheta, \vartheta, \dots, \vartheta)\text{-GARP}\}. \tag{A.14}$$

With these observations in place, the proof of Proposition 3 is now straightforward.

Proof of Proposition 3. First we note that there is a generalization to Proposition 1: it is straightforward to check $p^{t'} \succeq_{p,\overline{\vartheta}} p^t$ if and only if $\breve{x}^{t'} \succeq_{x,\overline{\vartheta}} \breve{x}^t$ and $p^{t'} \succ_{p,\overline{\vartheta}} p^t$ if and only if $\breve{x}^{t'} \succ_{x,\overline{\vartheta}} \breve{x}^t$. Thus, $\mathcal{D}$ satisfies $\overline{\vartheta}$-GAPP if and only if $\breve{\mathcal{D}}$ satisfies $\overline{\vartheta}$-GARP. Then it follows

[3]Indeed, we could obtain this result by modifying our proof of Theorem A.4.1. First, $\overline{\vartheta}$-GARP guarantees that there is a complete preorder $\succsim$ on $\{x^t\}_{t=1}^T$ that extends $\succeq_{x,\overline{\vartheta}}$ and $\succ_{x,\overline{\vartheta}}$. Then, by mimicking the proof of Lemma A.1, one could guarantee the existence of numbers ϕ^t and $\lambda^t > 0$ (for $t = 1, 2, \dots, T$) with the following properties: (a) $\phi^{t'} > \phi^t$ if $x^{t'} \succ x^t$; (b) $\phi^{t'} = \phi^t$ if $x^{t'} \sim x^t$; and (c) $\phi^{t'} \leq \phi^t + \lambda^t p^t \cdot (x^{t'} - \vartheta^t x^t)$ for all $t \neq t'$. The utility function $\widetilde{U} : \mathbb{R}^L_+ \to \mathbb{R}$ given by

$$U(x) = \min_{t \in T}\{\phi^t + \lambda^t p^t \cdot (x - \vartheta^t x^t)\}$$

is a continuous, concave, and strictly increasing. It is straightforward to check that property (c) guarantee that $\widetilde{U}$ rationalizes $\mathcal{D}$ in Afriat's sense.

immediately from (A.13) and (A.14) that the critical cost efficiency index of $\breve{\mathcal{D}}$ is equal to the rationality index of $\mathcal{D}$. ∎

A.4.4. *Allowing for variation in product characteristics across observations*

In Section 4.1(3) we considered a model of differentiated goods, where each product is represented by a vector of product characteristics in the space $\mathbb{R}^L_+$. We assumed in that section that the set of available goods, X, is fixed across observations but that assumption is not crucial to our model or test. We now allow the range of products available to the consumer to vary across observations.

The changes we have in mind include the introduction of new products and also changes to characteristics of an existing product. The latter could be a substantive change — for example, a change to the formula for a breakfast cereal — or it could be a change (say) to the amount of money spent on advertising that alters a product's utility (in the broad sense). All these cases could be formally captured by a data set $\mathcal{D} = \{(\psi^t, x^t, X^t)\}_{t=1}^T$, where X^t is the set of products available at observation t, x^t (as usual) is the product chosen, and $\psi^t : X^t \to \mathbb{R}_+$ is the price system as observation t. Notice that the price system at observation t is defined on X^t (the set of available products at observation t). An augmented utility function $U : Y \times \mathbb{R}_- \to \mathbb{R}$, where Y is a subset of $\mathbb{R}^L_+$ containing $\cup_{t \in T} X^t$ rationalizes $\mathcal{D}$ if, at each observation t,

$$U(x^t, -\psi^t(x^t)) \geq U(x, -\psi^t(x)) \text{ for all } x \in X^t;$$

in other words, x^t and its associated expenditure gives greater utility than any other product available at observation t. Sometimes, there is universal agreement that certain product characteristics $K \subset L$ will always make the product more desirable; in this case, we would also like the rationalizing utility function to be increasing in x_K.

Developing a test of whether $\mathcal{D} = \{(\psi^t, x^t, X^t)\}_{t=1}^T$ can be rationalized by an augmented utility function that is increasing in x_K requires a modification of the notion of revealed preference.

We say that $\psi^{t'}$ is *directly revealed preferred* to ψ^t, and denote it by $\psi^{t'} \succeq_{vp} \psi^t$ if $\psi^{t'}(\hat{x}) \leq \psi^t(x^t)$ where $\hat{x} \in X^{t'}$ and $\hat{x} \geq_K x^t$.[4] In other words, $\psi^{t'}$ is directly revealed preferred to ψ^t if there is a product $\hat{x}$ available at t' that is weakly superior to x^t in the dimensions belonging to K, the same in the other dimensions, and costs less than x^t. We say that $\psi^{t'}$ is *directly strictly revealed preferred* to ψ^t, and denote it by $\psi^{t'} \succ_{vp} \psi^t$ if $\psi^{t'}$ is directly revealed preferred to ψ^t and, either $\psi^{t'}(\hat{x}) < \psi^t(x^t)$ or $\hat{x} >_K x^t$. We denote the transitive closure of $\succeq_{vp}$ by $\succeq^*_{vp}$, that is, $\psi^{t'} \succeq^*_{vp} \psi^t$ if there are $t_1, t_2, \ldots, t_N$ in T such that $\psi^{t'} \succeq_{vp} \psi^{t_1}$, $\psi^{t_1} \succeq_{vp} \psi^{t_2}, \ldots, \psi^{t_{N-1}} \succeq_{vp} \psi^{t_N}$, and $\psi^{t_N} \succeq_{vp} \psi^{t'}$; in this case we say that $\psi^{t'}$ is *revealed*

[4]The partial order $\geq_K$ is defined as follows: $x'' \geq_K x'$ if $x''_{-K} = x'_{-K}$ and $x''_K \geq x'_K$.

preferred to ψ^t. If anywhere along this sequence, it is possible to replace $\succeq_{vp}$ with $\succ_{vp}$ then we denote that relation by $\psi^{t'} \succ^*_{vp} \psi^t$ and say that $\psi^{t'}$ is *strictly revealed preferred* to ψ^t.

It is straightforward to check that if $\mathcal{D}$ can be rationalized by an augmented utility function that is strictly increasing in x_K then it obeys *GAPP with respect to* $\succeq^*_{vp}$ *and* $\succ^*_{vp}$, in the following sense:

there do not exist observations $t, t' \in T$ such that $\psi^{t'} \succeq^*_{vp} \psi^t$ and $\psi^t \succ^*_{vp} \psi^{t'}$.

The following theorem asserts that the converse is also true.

Theorem A.4.3. *Let the data set be* $\mathcal{D} = \{(\psi^t, x^t, X^t)\}_{t=1}^T$, *where* X^t *is finite for all* $t \in T$ *and* $\psi^t : X^t \to \mathbb{R}_+$ *is strictly increasing in* x_K, *i.e., if* $x'' >_K x'$ *and both* x'' *and* x' *are in* X^t, *then* $\psi^t(x'') > \psi^t(x')$. *Let* Y *be a closed set in* $\mathbb{R}^L_+$ *containing* $\cup_{t \in T} X^t$.

Then $\mathcal{D}$ *can be rationalized by an augmented utility function* $U : Y \times \mathbb{R}_- \to \mathbb{R}$ *that is strictly increasing in* x_K *if and only if satisfies GAPP with respect to* $\succeq^*_{vp}$ *and* $\succ^*_{vp}$.

Proof. We skip the proof of the necessity of GAPP, which is straightforward, and turn to establishing its sufficiency. Let $X = \cup_{t \in T} X^t$. We claim that we can extend the function $\psi^t : X^t \to \mathbb{R}_+$ to a function $\underline{\psi}^t : X \to \mathbb{R}$ that is increasing in x_K and such that $\underline{\mathcal{D}} = \{(\underline{\psi}^t, x^t)\}_{t=1}^T$ satisfies GAPP (with respect to the revealed preference orders $\succeq^*_p$ and $\succ^*_p$ induced by $\underline{\mathcal{D}}$). Then an application of Theorem 2 will guarantee that $\underline{\mathcal{D}}$, and thus also $\mathcal{D}$, can be rationalized by an augmented utility function $U : Y \times \mathbb{R}_- \to \mathbb{R}$ that is strictly increasing in x_K.

To guarantee that $\underline{\mathcal{D}}$ satisfies GAPP, with respect to $\succeq^*_p$ and $\succ^*_p$, we need to specify $\underline{\psi}^t(x)$, for $x \in X \setminus X^t$, in such a way that $\succeq^*_p = \succeq^*_{vp}$ and $\succ^*_p = \succ^*_{vp}$. Then GAPP holds with respect to $\succeq^*_p$ and $\succ^*_p$ because GAPP holds with respect to $\succeq^*_{vp}$ and $\succ^*_{vp}$. Because X is finite, such an extension ψ^t can be obtained with no technical difficulty. For $x \in X \setminus X^t$, if there is no $x' \in X^t$ such that $x' >_K x$, we choose $\underline{\psi}^t(x) > \max\{\psi^s(x^s) : s \in T\}$, while making sure that $\underline{\psi}^t$ remains increasing in x_K. If there is $x' \in X^t$ such that $x' >_K x$, then choose $\underline{\psi}^t(x)$ to be strictly lower than $\psi^t(x')$, but if $x = x^s$ for some observation s, then choose $\underline{\psi}^t(x) = \underline{\psi}^t(x^s) > \psi^s(x^s)$ if $\psi^t(x') > \psi^s(x^s)$. In this way, we guarantee $\succeq^*_p = \succeq^*_{vp}$ and $\succ^*_p = \succ^*_{vp}$. ■

Appendix A.5. More on compensating variation

Our objective is to prove equation (11) from the body of the paper:

$$\inf(\mu_c) = \max\{m_c^s : m_c^s \text{ satisfies (10) for some } s \in S\} \tag{A.15}$$

where (10) requires $p^{t_2} x^s + m_c^s = p^s x^s$.

Proof. Since S is a finite set, there is $\bar{s} \in S$ that achieves the maximum on the right of (A.15). We have already shown that $\inf(\mu_c) \geq m_c^{\bar{s}}$, so it remains to show that they are equal. We shall do this by producing, for any $\epsilon > 0$, an augmented utility function rationalizing $\mathcal{D}$ for which the compensating variation is smaller than $m_c^{\bar{s}} + \epsilon$.

To this end, let U be any augmented utility function that rationalizes $\mathcal{D} = \{(p^t, x^t)\}_{t=1}^T$; we know that U exists since $\mathcal{D}$ obeys GAPP by assumption. Let $\hat{\psi} : X \to \mathbb{R}_+$ be the nonlinear price system given by $\hat{\psi}(x) = p^{t_2} \cdot x + m_c^{\bar{s}} + \epsilon$ and suppose that $\hat{x} \in \text{argmax}_{x \in X} U(x, -\hat{\psi}(x))$. Now consider the data set $\mathcal{D}' = \mathcal{D} \cup \{(\hat{\psi}, \hat{x})\}$. Obviously this data set can be rationalized (in fact it is rationalized by U). Furthermore, $\hat{\psi} \not\succeq_p p^s$ for any $s \in S$. This is because

$$\hat{\psi}(x^s) = p^{t_2} x^s + m_c^{\bar{s}} + \epsilon > p^{t_2} \cdot x^s + m_c^s = p^s x^s$$

for any $s \in S$. (Recall that, be definition, $m_c^{\bar{s}} \geq m_c^s$ for all $s \in S$.) Thus there is a complete preorder $\succsim$ on $\{p^t\}_{t=1}^T \cup \{\hat{\psi}\}$, completing the revealed preference relations on $\mathcal{D}'$ such that $p^{t_1} \succ \hat{\psi}$. By Theorem A.4.2, there is an augmented utility $\hat{U}$ rationalizing $\mathcal{D}'$ such that its indirect utility $\hat{V}$ satisfies $\hat{V}(p^{t_1}) > \hat{V}(\hat{\psi})$. In other words,

$$\hat{V}(\hat{\psi}) = \max_{x \in X} \hat{U}(x, -p^{t_2} \cdot x - m_c^{\bar{s}} - \epsilon) < \hat{U}(x^{t_1}, -p^{t_1} \cdot x^{t_1}).$$

So for the augmented utility function $\hat{U}$, the compensating variation must be smaller than $m_c^{\bar{s}} + \epsilon$. ■

Our treatment of the compensating and equivalent variations can be easily extended to allow for nonlinear pricing. We give a sketch of the procedure for calculating a bound on the compensating variation and leave the reader to fill in the details; this procedure is completely analogous to the one for linear prices described in Section 3.3

Let U be the consumer's augmented utility function. Suppose that the initial price is ψ^{t_1} and it changes to ψ^{t_2}, leading to a change in consumption from x^{t_1} to x^{t_2}. Then the compensating variation μ_c is, by definition, the variable that solves the equation

$$\max_{x \in \mathbb{R}_+^L} U(x, -\psi^{t_2}(x) - \mu_c) = V(\psi^{t_1}) = U(x^{t_1}, -\psi^{t_1}(x^{t_1})). \tag{A.16}$$

Note that μ_c is unique since U is strictly increasing in the last argument. We could think of μ_c as the lump sum transferred *from* the consumer (if it is positive) or *to* the consumer (if it is negative) after the price change that will make her indifferent between the two situations.

Now suppose a data set $\mathcal{D}$ obeys GAPP and contains the observation (ψ^{t_1}, x^{t_1}). How can we form a lower bound of the compensating variation of a price change from ψ^{t_1} to ψ^{t_2}? (Note that our discussion is valid whether or not ψ^{t_2} is an observed price system in the $\mathcal{D}$.) Formally, we wish to find

$$\inf\{\mu_c : \mu_c \text{ solves (A.16) for some augmented utility function } U \text{ that rationalizes } \mathcal{D}\}.$$

Abusing terminology somewhat, we shall denote this term by $\inf(\mu_c)$.

We now describe how to compute this bound. Let $S \subset T$ be the set of observations such that $s \in S$ if $\psi^s \succeq_p^* \psi^{t_1}$. This set is nonempty since it contains p^{t_1} itself. For each $s \in S$, there is m_c^s such that

$$\psi^{t_2}(x^s) + m_c^s = \psi^s(x^s). \tag{A.17}$$

For any U that rationalizes $\mathcal{D}$, the compensating variation $\mu_c \geq m_c^s$. Indeed, if $m < m_c^s$, then $m \neq \mu_c$ for any utility function rationalizing $\mathcal{D}$ because

$$\begin{aligned}\max_{x \in \mathbb{R}_+^L} U(x, -\psi^{t_2}(x) - m) \geq U(x^s, -\psi^{t_2}(x^s) - m) > U(x^s, -\psi^{t_2}(x^s) - m_c^s) \\ = U(x^s, -\psi^s(x^s)) \geq U(x^{t_1}, -\psi^{t_1}(x^{t_1})) = V(\psi^{t_1}).\end{aligned}$$

Thus $\inf(\mu_c) \geq m_c^s$ for all $s \in S$. In fact, by adapting the argument we provided for the case of linear prices in the earlier part of this section, we could show that

$$\inf(\mu_c) = \max\{m_c^s : m_c^s \text{ satisfies (A.17) for some } s \in S\}. \tag{A.18}$$

Since the right side of this equation can be computed from the data, we have found a practical way of calculating $\inf(\mu_c)$.

Notice that if ψ^{t_2} is revealed preferred to ψ^{t_1}, in the sense that there is $s' \in S$ such that $m_c^{s'} \geq 0$, then $\inf(\mu_c) \geq 0$; in other words, a lump sum *tax* of $\inf(\mu_c)$ will leave the agent no worse off than at t_1 and potentially better off. On the other hand, if ψ^{t_2} is *not* revealed preferred to ψ^{t_1}, that is, for every $s \in S$, we have $m_c^s < 0$, then $\inf(\mu_c) < 0$; in other words, at $\psi = \psi^{t_2}$, a lump sum *transfer* of $\inf(\mu_c)$ to the agent will guarantee that the agent no worse off than at t_1 and potentially better off.

APPENDIX A.6. PROOF OF THEOREM 3

Given a deterministic data set of the form $\mathcal{D} = \{(p^t, x^t)\}_{t=1}^T$, we can construct its *iso-expenditure version* $\breve{\mathcal{D}} = \{(p^t, \breve{x}^t)\}_{t=1}^T$, where $\breve{x}^t = x^t / p^t \cdot x^t$ (so $p^t \cdot \breve{x}^t = 1$ for all t). Suppose that $\breve{x}^t$ does not lie on the intersection of budget planes, that is, there is i^t such that $\breve{x}^t \in \text{int}(B^{i^t,t})$. We make two observations. First, Proposition 1 tells us that $\mathcal{D}$ satisfies GAPP if and only if $\breve{\mathcal{D}}$ satisfies GARP. Second, if $\mathcal{D}$ satisfies GAPP then so does $\mathcal{D}' = \{(p^t, y^t)\}_{t \in T}$ if y^t has the property that its re-scaled version $\breve{y}^t$ satisfies $\breve{y}^t \in \text{int}(B^{i^t,t})$; this is because the revealed preference relations (over the bundles $\breve{y}^t$) are determined only by where $\breve{y}^t$ lies on the budget set relative to its intersection with another budget. It follows from these observations that we may classify deterministic data sets that obey GAPP according to the patch occupied by the scaled bundle $\breve{x}^t$ at each B^t. Formally, each $\mathcal{D}$ that obeys GAPP is associated with an iso-expenditure $\breve{\mathcal{D}}$ that obeys GARP, which is in turn associated with a vector $a = \left(a^{1,1}, \ldots, a^{I_T,T}\right) \in \mathcal{A}$ (as defined in Section 5.2 of the main paper).

Given a repeated cross-sectional data set $\mathscr{D}$, we can construct $\mathcal{A}$ and the matrix A. (Recall that A denotes the matrix whose columns consist of every $a \in \mathcal{A}$, arranged in an arbitrary order.) Suppose that this data set can be rationalized by some distribution μ. Let ν_a denote the mass of consumers of type a in the population, that is

$$\nu_a = \mu\left(\left\{\omega \in \Omega : \frac{\chi^t(\omega)}{p^t \cdot \chi^t(\omega)} \in B^{i_t,t} \text{ if } a^{i_t,t} = 1, \text{for all } t \in T\right\}\right).$$

At a given observation t, let $\mathcal{A}^{i_t,t} = \{a : a^{i_t,t} = 1\}$; this is the subset of GARP-consistent types that have their re-scaled demands in the patch $B^{i_t,t}$ at observation t. The proportion of the population whose types belong to $\mathcal{A}^{i_t,t}$ is

$$\mu\left(\left\{\omega \in \Omega : \frac{\chi^t(\omega)}{p^t \cdot \chi^t(\omega)} \in B^{i_t,t}\right\}\right) = \sum_{a \in \mathcal{A}^{i_t,t}} \nu_a = \sum_{a \in \mathcal{A}} \nu_a\, a^{i_t,t}.$$

Since $\mathscr{D}$ is rationalized by μ,

$$\mathring{\pi}^t(Y) = \mu(\{\omega \in \Omega : \chi^t(\omega) \in Y\}) \text{ for any measurable } Y \subset \mathbb{R}^L_+. \tag{A.19}$$

Setting $Y = \{x \in \mathbb{R}^L_+ : x/(p^t \cdot x) \in B^{i_t,t}\}$, we obtain

$$\pi^{i_t,t} = \sum_{a \in \mathcal{A}} \nu_a\, a^{i_t,t} \tag{A.20}$$

where $\pi^{i_t,t}$ is defined by equation (17) (in the main paper). In other words, the observed probability of choices that land on $B^{i_t,t}$ after scaling must equal to the mass of GARP-consistent types implied by μ. This condition must hold for all patches $B^{i_t,t}$, so (A.20) can be more succinctly written as $A\nu = \pi$, where ν is the column vector $(\nu_a)_{a \in \mathcal{A}}$. (Recall that π is the vector of observed patch probabilities.) So we have established that if $\mathscr{D}$ can be RAUM-rationalized then there is a distribution $\nu \in \Delta^{|\mathcal{A}|}$ that solves $A\nu = \pi$.

It remains for us to show the converse. Given $\mathring{\pi}^t$, we define $\tilde{\pi}^{i_t,t}$ to be the conditional distribution of demand at observation t when it restricted to the cone $K^{i_t,t} = \{r \cdot x : x \in B^{i_t,t},\ r > 0\}$. Thus, if Y is a measurable subset of $\mathbb{R}^L_+$, then

$$\mathring{\pi}^t(Y \cap K^{i_t,t}) = \pi^{i_t,t}\, \tilde{\pi}^{i_t,t}(Y).$$

(Recall that, by definition, $\pi^{i_t,t} = \mathring{\pi}^t(K^{i_t,t})$.) Of course, if $Y \cap K^{i_t,t} = \emptyset$ then $\tilde{\pi}^{i_t,t}(Y) = 0$.

Given a and t, there is a unique i_t' such that $a^{i_t',t} = 1$; let $K^t_a = K^{i_t',t}$ and let $\tilde{\pi}^t_a$ be the probability measure on $\mathbb{R}^L_+$ such that $\tilde{\pi}^t_a = \tilde{\pi}^{i_t',t}$. Obviously, $\tilde{\pi}^t_a(K^t_a) = 1$.

Let λ_a be the product measure on $(\mathbb{R}^L_+)^T$ given by $\lambda_a = \times_{t \in T} \tilde{\pi}^t_a$. It follows from the definition of a that

$$\times_{t \in T} K^t_a \subset \left\{x \in (\mathbb{R}^L_+)^T :\ \{(p^t, x^t)\}_{t \in T} \text{ satisfies GAPP}\right\}$$

and since $\tilde{\pi}^t_a(K^t_a) = 1$ for all t, we obtain

$$\lambda_a\left(\left\{x \in (\mathbb{R}^L_+)^T : \{(p^t, x^t))\}_{t \in T} \text{ satisfies GAPP}\right\}\right) = 1. \tag{A.21}$$

Note that x^t refers to the tth entry of x).

Define $\Omega = \mathcal{A} \times (\mathbb{R}^L_+)^T$ and the probability measure μ on Ω by $\mu(\{a\} \times Y) = \nu_a \lambda_a(Y)$ for any measurable set $Y \subseteq (\mathbb{R}^L_+)^T$, where ν_a refers to the ath entry of ν. Lastly, define $\chi : \Omega \to (\mathbb{R}^L_+)^T$ by $\chi((a,x)) = x$. Then, using (A.21), we obtain

$$\mu\left(\{\, (a,x) \in \Omega : \{(p^t, \chi^t(a,x))\}_{t\in T} \text{ satisfies GAPP } \}\right)$$
$$= \sum_{a\in\mathcal{A}} \nu_a \lambda_a \left(\left\{ x \in (\mathbb{R}^L_+)^T : \{(p^t, \chi^t(a,x))\}_{t\in T} \text{ satisfies GAPP} \right\}\right) = \sum_{a\in A} \nu_a = 1.$$

It remains for us to show that (A.19) holds. Let Y be a measurable set in $\mathbb{R}^L_+$. For any $K^{i_t,t}$,

$$\begin{aligned}
\mu(\{(a,x) \in \Omega : \chi^t(a,x) \in Y \cap K^{i_t,t}\}) &= \sum_{a\in\mathcal{A}} \nu_a \lambda_a(\{x \in (\mathbb{R}^L_+)^T : \chi^t(a,x) \in Y \cap K^{i_t,t}\}) \\
&= \sum_{a\in\mathcal{A}} \nu_a \lambda_a(\{x \in (\mathbb{R}^L_+)^T : x^t \in Y \cap K^{i_t,t}\}) \\
&= \sum_{a\in\mathcal{A}} \nu_a \tilde{\pi}^t_a(\{x^t \in \mathbb{R}^L_+ : x^t \in Y \cap K^{i_t,t}\})
\end{aligned}$$

Recall that $\mathcal{A}^{i_t,t} = \{a \in \mathcal{A} : a^{i_t,t} = 1\}$, so $\tilde{\pi}^t_a(\{x^t \in \mathbb{R}^L_+ : x^t \in Y \cap K^{i_t,t}\}) = 0$ for any $a \notin \mathcal{A}^{i_t,t}$. Thus

$$\begin{aligned}
\sum_{a\in\mathcal{A}} \nu_a \tilde{\pi}^t_a(\{x^t \in \mathbb{R}^L_+ : x^t \in Y \cap K^{i_t,t}\}) &= \sum_{a\in\mathcal{A}^{i_t,t}} \nu_a \tilde{\pi}^t_a(\{x^t \in \mathbb{R}^L_+ : x^t \in Y \cap K^{i_t,t}\}) \\
&= \frac{\mathring{\pi}^t(Y \cap K^{i_t,t})}{\pi^{i_t,t}} \sum_{a\in\mathcal{A}^{i_t,t}} \nu_a \\
&= \mathring{\pi}^t(Y \cap K^{i_t,t}),
\end{aligned}$$

where the last equation follows from $A\nu = \pi$. Thus we have shown that, for all $K^{i_t,t}$,

$$\mu(\{(a,x) \in \Omega : \chi^t(a,x) \in Y \cap K^{i_t,t}\}) = \mathring{\pi}^t(Y \cap K^{i_t,t}).$$

This in turn guarantees that (A.19) holds. ■

Appendix A.7. Omitted Details from Section 6

In this section, we formally develop our bootstrap procedure from Section 6.2. We begin by describing Weyl-Minkowski duality[5]which is used for the equivalent (dual) restatement (26) of our test (24). As we mentioned earlier, we will also appeal to this duality in the proof of the asymptotic validity of our testing procedure.

Theorem A.7.1. *(Weyl-Minkowski Theorem for Cones) A subset $\mathcal{C}$ of $\mathbb{R}^I$ is a finitely generated cone*

$$\mathcal{C} = \{\nu_1 a_1 + ... + \nu_{|\mathcal{A}|} a_{|\mathcal{A}|} : \nu_h \geq 0\} \text{ for some } A = [a_1, ..., a_H] \in \mathbb{R}^{I \times |\mathcal{A}|} \tag{A.22}$$

[5]See, for example, Theorem 1.3 in Ziegler (1995).

if, and only if, it is a finite intersection of closed half spaces

$$\mathcal{C} = \{t \in \mathbb{R}^I | Bt \leq 0\} \text{ for some } B \in \mathbb{R}^{m\times I}. \tag{A.23}$$

The expressions in (A.22) and (A.23) are called a $\mathcal{V}$-representation (as in "vertices") and a $\mathcal{H}$-representation (as in "half spaces") of $\mathcal{C}$, respectively. In what follows, we use an $\mathcal{H}$-representation of cone(A) corresponding to a $m \times I$ matrix B as implied by Theorem A.7.1.

We are now in a position to show that the bootstrap procedure defined in Section 6.2 is asymptotically valid. Note first that $\Theta = [\underline{\theta}, \bar{\theta}]$, where

$$\bar{\theta} = \max_{\nu\in\Delta^{|\mathcal{A}|-1}} \rho \cdot \nu = \max_{1\leq j\leq|\mathcal{A}|} \rho_j \tag{A.24}$$

$$\underline{\theta} = \min_{\nu\in\Delta^{|\mathcal{A}|-1}} \rho \cdot \nu = \min_{1\leq j\leq|\mathcal{A}|} \rho_j, \tag{A.25}$$

where ρ_j denotes the jth component of ρ. We normalize (ρ, θ) such that $\Theta = [\underline{\theta}, \underline{\theta} + 1]$. Next, define

$$\mathcal{H} := \{1, 2, ..., |\mathcal{A}|\} \tag{A.26}$$

$$\overline{\mathcal{H}} := \{j \in \mathcal{H} \mid \rho_j = \bar{\theta}\} \tag{A.27}$$

$$\underline{\mathcal{H}} := \{j \in \mathcal{H} \mid \rho_j = \underline{\theta}\} \tag{A.28}$$

$$\mathcal{H}_0 := \mathcal{H} \setminus (\overline{\mathcal{H}} \cup \underline{\mathcal{H}}). \tag{A.29}$$

Recall that τ_N is a tuning parameter chosen such that $\tau_N \downarrow 0$ and $\sqrt{N}\tau_N \uparrow \infty$. For $\theta \in \Theta_I$, we now formally define the τ_N-tightened version of $\mathcal{S}$ as

$$\mathcal{S}_{\tau_N}(\theta) := \{A\nu \mid \rho\nu = \theta, \nu \in \mathcal{V}_{\tau_N}(\theta)\},$$

where

$$\mathcal{V}_{\tau_N}(\theta) := \left\{ \nu \in \Delta^{|\mathcal{A}|-1} \;\middle|\; \begin{array}{l} \nu_j \geq \frac{(\bar{\theta}-\theta)\tau_N}{|\underline{\mathcal{H}}\cup\mathcal{H}_0|}, j \in \underline{\mathcal{H}},\ \nu_{j'} \geq \frac{(\theta-\underline{\theta})\tau_N}{|\overline{\mathcal{H}}\cup\mathcal{H}_0|},\ j' \in \overline{\mathcal{H}}, \\ \\ \nu_{j''} \geq \left[1 - \frac{(\bar{\theta}-\theta)|\underline{\mathcal{H}}|}{|\underline{\mathcal{H}}\cup\mathcal{H}_0|} - \frac{(\theta-\underline{\theta})|\overline{\mathcal{H}}|}{|\overline{\mathcal{H}}\cup\mathcal{H}_0|}\right] \frac{\tau_N}{|\mathcal{H}_0|},\ j'' \in \mathcal{H}_0 \end{array} \right\}.$$

In applications where ρ is binary, the above notation simplifies. Specifically, in our empirical application on deriving the welfare bounds, $\rho = \mathbb{1}_{t\succeq^*_p t'}$ and $\theta = \mathcal{N}_{t\succeq^*_p t'}$. Here, $\bar{\theta} = 1$, $\underline{\theta} = 0$, and $\bar{\theta} - \underline{\theta} = 1$ holds without any normalization. Also, $\overline{\mathcal{H}}$ ($\underline{\mathcal{H}}$) is just the set of indices for the types that (do not) prefer price p^t compared to $p^{t'}$, while $\mathcal{H}_0$ is empty. We then have:

$$\mathcal{S}_{\tau_N}(\mathcal{N}_{t\succeq^*_p t'}) = \left\{ A\nu \;\middle|\; \mathbb{1}'_{t\succeq^*_p t'}\nu = \mathcal{N}_{t\succeq^*_p t'}, \nu \in \mathcal{V}_{\tau_N}(\mathcal{N}_{t\succeq^*_p t'}) \right\},$$

where

$$\mathcal{V}_{\tau_N}(\mathcal{N}_{t \succeq_p^* t'}) = \left\{ \nu \in \Delta^{|\mathcal{A}|-1} \;\middle|\; \nu_j \geq \frac{(1-\mathcal{N}_{t \succeq_p^* t'})\tau_N}{|\underline{\mathcal{H}}|}, j \in \underline{\mathcal{H}},\ \nu_{j'} \geq \frac{\mathcal{N}_{t \succeq_p^* t'}\tau_N}{|\overline{\mathcal{H}}|},\ j' \in \overline{\mathcal{H}} \right\}.$$

We now state the mild data assumptions.

Assumption 1. For all $t = 1, ..., T$, $\frac{N_t}{N} \to \kappa_t$ as $N \to \infty$, where $\kappa_t > 0, 1 \leq t \leq T$.

Assumption 2. The econometrician observes T independent cross-sections of i.i.d. samples $\left\{ x^t_{n(t)} \right\}_{n(t)=1}^{N_t}, t = 1, ..., T$ of consumers' choices corresponding to the known price vectors $\{p_t\}_{t=1}^T$.

Next, let $\mathbf{d}^{i,t}_{n(t)} := \mathbf{1}\{x^t_{n(t)} \in B^{i,t}\}$, $\mathbf{d}^t_{n(t)} = [\mathbf{d}^{1,t}_{n(t)}, ..., \mathbf{d}^{I_t,t}_{n(t)}]$, and $\mathbf{d}^t_n = [\mathbf{d}^{1,t}_n, ..., \mathbf{d}^{I_t,t}_n]$. Let $\mathbf{d}_t$ denote the choice vector of a consumer facing price p^t (we can, for example, let $\mathbf{d}_t = \mathbf{d}^t_1$). Define $\mathbf{d} = [\mathbf{d}'_1, ..., \mathbf{d}'_T]'$: note, $\mathrm{E}[\mathbf{d}] = \pi$ holds by definition. Among the rows of B some of them correspond to constraints that hold trivially by definition, whereas some are for non-trivial constraints. Let $\mathcal{K}^R$ be the index set for the latter. Finally, let

$$\begin{aligned} g &= B\mathbf{d} \\ &= [g_1, ..., g_m]'. \end{aligned}$$

With these definitions, consider the following requirement:

Condition 1. For each $k \in \mathcal{K}^R$, $\mathrm{var}(g_k) > 0$ and $\mathrm{E}[|g_k/\sqrt{\mathrm{var}(g_k)}|^{2+c_1}] < c_2$ hold, where c_1 and c_2 are positive constants.

This guarantees the Lyapunov condition for the triangular array CLT used in establishing asymptotic uniform validity. This type of condition has been used widely in the literature of moment inequalities; see Andrews and Soares (2010).

Proof of Theorem 4. Define $\mathcal{C} = \mathrm{cone}(A)$ and

$$\mathcal{T}(\theta) = \{\pi = A\nu : \rho'\nu = \theta, \nu \in \mathbb{R}^{|\mathcal{A}|}\},$$

an affine subspace of $\mathbb{R}^I$. It is convenient to rewrite $\mathcal{T}(\theta)$ as $\mathcal{T}(\theta) = \{t \in \mathbb{R}^I : \tilde{B}t = d(\theta)\}$ where $\tilde{B} \in \tilde{m} \times \mathbb{R}^I$, $d(\cdot) \in \tilde{m} \times 1$, and $\tilde{m}$ all depend on (ρ, A). We let $\tilde{b}_j$ denote the j-th row of $\tilde{B}$. Then

$$\mathcal{S}(\theta) = \mathcal{C} \cap \Delta^{|\mathcal{A}|-1} \cap \mathcal{T}(\theta).$$

By Theorem A.7.1, $\mathcal{C} = \{\pi : B\pi \leq 0\}$, therefore

$$\mathcal{S}(\theta) = \{t \in \mathbb{R}^{|\mathcal{A}|} : Bt \leq 0, \tilde{B}t = d(\theta), \mathbf{1}'_H t = 1\}. \quad \text{(A.30)}$$

Let

$$\psi(\theta) = [\psi_1(\theta), ..., \psi_H(\theta)]' \quad \theta \in \Theta$$

with

$$\psi_j(\theta) = \begin{cases} \frac{(\bar{\theta}-\theta)}{|\underline{\mathcal{H}}\cup\mathcal{H}_0|} & \text{if } j \in \underline{\mathcal{H}}, \\ \frac{(\theta-\bar{\theta})}{|\overline{\mathcal{H}}\cup\mathcal{H}_0|} & \text{if } j \in \overline{\mathcal{H}}, \\ \left[1 - \frac{(\bar{\theta}-\theta)|\underline{\mathcal{H}}|}{|\underline{\mathcal{H}}\cup\mathcal{H}_0|} - \frac{(\theta-\underline{\theta})|\overline{\mathcal{H}}|}{|\overline{\mathcal{H}}\cup\mathcal{H}_0|}\right]\frac{1}{|\mathcal{H}_0|} & \text{if } j \in \mathcal{H}_0, \end{cases}$$

where terms are defined in (A.26)-(A.29). Then

$$\mathcal{S}_{\tau_N}(\theta) = \{\pi = A\nu : \nu \geq \tau_N \psi(\theta), \nu \in \Delta^{|\mathcal{A}|-1}, \rho'\nu = \theta\}.$$

Finally, let

$$\mathcal{C}_{\tau_N} = \{\pi = A\nu : \nu \geq \tau_N \psi(\theta)\}.$$

Then

$$\mathcal{S}_{\tau_N}(\theta) = \mathcal{C}_{\tau_N} \cap \Delta^{|\mathcal{A}|-1} \cap \mathcal{T}(\theta).$$

Proceeding as in the proof of Lemma 4.1 in KS, we can express the set $\mathcal{C}_{\tau_N}$ as

$$\mathcal{C}_{\tau_N} = \{t : Bt \leq -\tau_N \phi(\theta)\}$$

where

$$\phi(\theta) = -BA\psi(\theta).$$

As in Lemma 4.1 in KS, let the first $\bar{m}$ rows of B represent inequality constraints and the rest equalities, and also let Φ_{kh} the (k, h)-element of the matrix $-BA$. We have

$$\phi_k = \sum_{h=1}^{|\mathcal{A}|} \Phi_{kh}\psi_h(\theta)$$

where, for each $k \leq \bar{m}$, $\{\Phi_{kh}\}_{h=1}^{|\mathcal{A}|}$ are all nonnegative, with at least some of them being strictly positive, and $\Phi_{kh} = 0$ for all h if $\bar{m} < k \leq m$. Since $\psi_h(\theta) > 0, 1 \leq h \leq |\mathcal{A}|$ for every $\theta \in \Theta$ by definition, we have $\phi_j(\theta) \geq C, 1 \leq j \leq \bar{m}$ for some positive constant C, and $\phi_j(\theta) = 0, \bar{m} < j \leq m$ for every $\theta \in \Theta$. Putting these together, we have

$$\mathcal{S}_{\tau_N}(\theta) \quad = \quad \{t \in \mathbb{R}^{|\mathcal{A}|} : Bt \leq -\tau_N \phi(\theta), \tilde{B}t = d(\theta), \mathbf{1}_H' t = 1\}$$

where $\mathbf{1}_H$ denotes the $|\mathcal{A}|$-vector of ones. Define the $\mathbb{R}^I$-valued random vector

$$\pi^*_{\tau_N} := \frac{1}{\sqrt{N}}\zeta + \hat{\eta}_{\tau_N},\ \zeta \sim \mathrm{N}(0, \hat{S})$$

where $\hat{S}$ is a consistent estimator for the asymptotic covariance matrix of $\sqrt{N}(\hat{\pi} - \pi)$. Then (conditional on the data) the distribution of

$$\delta^*(\theta) := N \min_{\eta \in \mathcal{S}_{\tau_N}(\theta)} [\pi^*_{\tau_N} - \eta]'\Omega[\pi^*_{\tau_N} - \eta]$$

corresponds to that of the bootstrap test statistics. Let

$$B_* := \begin{bmatrix} B \\ \tilde{B} \\ \mathbf{1}'_H \end{bmatrix}$$

Define $\ell = \text{rank}(B_*)$ for the augmented matrix B_* instead of B in KS, and let the $\ell \times m$-matrix K be such that KB_* is a matrix whose rows consist of a basis of the row space row (B_*). Also let M be an $(I-\ell) \times I$ matrix whose rows form an orthonormal basis of $\ker B_* = \ker(KB_*)$, and define $P = \binom{KB_*}{M}$. Finally, let $\hat{g} = B_* \hat{\pi}$.

Define

$$\begin{aligned} T(x,y) &:= \begin{pmatrix} x \\ y \end{pmatrix}' P^{-1\prime} \Omega P^{-1} \begin{pmatrix} x \\ y \end{pmatrix}, \quad x \in \mathbb{R}^{\ell}, y \in \mathbb{R}^{I-\ell} \\ t(x) &:= \min_{y \in \mathbb{R}^{I-\ell}} T(x,y) \\ s(g) &:= \min_{\gamma = [\gamma^{\leq\prime}, \gamma^{=\prime}]', \gamma^{\leq} \leq 0, \gamma' \in \text{col}(B)} t(K[g-\gamma]) \end{aligned}$$

with

$$\gamma^{=} = \begin{bmatrix} \mathbf{0}_{m-\bar{m}} \\ d(\theta) \\ 1 \end{bmatrix}$$

where $\mathbf{0}_{m-\bar{m}}$ denotes the $(m-\bar{m})$-vector of zeros. It is easy to see that $t : \mathbb{R}^{\ell} \to \mathbb{R}_+$ is a positive definite quadratic form. By (A.30), we can write $\delta_N(\theta) = Ns(\hat{g}) = s(\sqrt{N}\hat{g})$. Likewise, for the bootstrapped version of δ we have

$$\begin{aligned} \delta^*(\theta) &= N \min_{\eta \in \mathcal{S}_{\tau_N}(\theta)} [\pi^*_{\tau_N} - \eta]' \Omega [\pi^*_{\tau_N} - \eta] \\ &= s(\varphi_N(\hat{\xi}) + \zeta), \end{aligned}$$

where $\hat{\xi} = B_* \hat{\pi} / \tau_N$. Note the function $\varphi_N(\xi) = [\varphi^1_N(\xi), ..., \varphi^m_N(\xi)]$ for $\xi = (\xi_1, ..., \xi_m)' \in \text{col}(B_*)$. Moreover, its k-th element φ^k_N for $k \leq \bar{m}$ satisfies

$$\varphi^k_N(\xi) = 0$$

if $|\xi^k| \leq \delta$ and $\xi^j \leq \delta, 1 \leq j \leq m$, $\delta > 0$, for large enough N and $\varphi^k_N(\xi) = 0$ for $k > \bar{m}$. This follows (we use some notation in the proof of Theorem 4.2 in KS, which the reader is referred to) by first noting that it suffices to show that for small enough $\delta > 0$, every element x^* that fulfills equation (9.2) in KS with its RHS intersected with $\cap_{j=1}^{\bar{m}} \tilde{S}_j(\delta), \tilde{S}_j(\delta) = \{x : |\tilde{b}'_j x - d_j(\theta)| \leq \tau\delta\}$ satisfies

$$x^* | \mathcal{S}(\theta) \in \cap_{j=1}^{q} H^{\tau}_j \cap L \cap \mathcal{T}(\theta).$$

If not, then there exists $(\tilde{a}, \tilde{x}) \in F \cap \mathcal{T}(\theta) \times \cap_{j=1}^{q} H_j \cap L \cap \mathcal{T}(\theta)$ such that

$$(\tilde{a} - \tilde{x})'(\tilde{x}|\mathcal{S}_\tau(\theta) - \tilde{x}) = 0,$$

where $\tilde{x}|\mathcal{S}_\tau(\theta)$ denotes the orthogonal projection of $\tilde{x}$ on $\mathcal{S}_\tau(\theta)$. This, in turn, implies that there exists a triplet $(a_0, a_1, a_2) \in \mathcal{A} \times \mathcal{A} \times \mathcal{A}$ such that $(a_1 - a_0)'(a_2 - a_0) < 0$. But as shown in the proof of Theorem 4.2 in KS, this cannot happen. The conclusion then follows by Theorem 1 of Andrews and Soares (2010). ■

References

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

ANDREWS, D. W., AND G. SOARES (2010): "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica*, 78, 119–157.

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

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

DEATON, A., AND J. MUELLBAUER (1980): "An Almost Ideal Demand System," *American Economic Review*, 70(3), 312–326.

FOSTEL, A., H. E. SCARF, AND M. J. TODD (2004): "Two New Proofs of Afriat's Theorem," *Economic Theory*, 24(1), 211–219.

GORMAN, W. M. (1961): "On a class of preference fields," *Metroeconomica*, 13(2), 53–56.

HALEVY, Y., D. PERSITZ, AND L. ZRILL (2018): "Parametric recoverability of preferences," *Journal of Political Economy*, 126(4), 1558–1593.

KITAMURA, Y., AND J. STOYE (2018): "Nonparametric Analysis of Random Utility Models," *Econometrica*, 86(6), 1883–1909.

MUELLBAUER, J. (1976): "Community Preferences and the Representative Consumer," *Econometrica*, pp. 979–999.

NISHIMURA, H., E. A. OK, AND J. K.-H. QUAH (2017): "A Comprehensive Approach to Revealed Preference Theory," *American Economic Review*, 107(4), 1239–1263.

POLISSON, M., J. K.-H. QUAH, AND L. RENOU (2020): "Revealed Preferences over Risk and Uncertainty," *American Economic Review*, 110(6), 1782–1820.

QUAH, J. K.-H. (2014): "A Test for Weakly Separable Preferences," *Oxford University Discussion Paper*.

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

VARIAN, H. R. (1990): "Goodness-of-Fit in Optimizing Models," *Journal of Econometrics*, 46(1), 125–140.

ZIEGLER, G. M. (1995): *Lectures on polytopes*, Graduate texts in mathematics. Springer, New York.