REVEALED PRICE PREFERENCE: THEORY AND EMPIRICAL ANALYSIS

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

DECEMBER 18, 2021

ABSTRACT. To determine the welfare implications of price changes in demand data, we introduce a revealed preference relation over prices. We show that the absence of cycles in this relation characterizes a consumer who trades off the utility of consumption against the disutility of expenditure. Our model can be applied whenever a consumer's demand over a strict subset of all available goods is being analyzed; it can also be extended to settings with discrete goods and nonlinear prices. To illustrate its use, we apply our model to a single-agent data set and to a data set with repeated cross-sections. We develop a novel test of linear hypotheses on partially identified parameters to estimate the proportion of the population who are revealed better off due to a price change in the latter application. This new technique can be used for nonparametric counterfactual analysis more broadly.

1. INTRODUCTION

A central issue in economic analysis is the determination of the welfare effect of price changes. As an example, suppose we observe a consumer's purchases of two goods, gasoline and food, from two separate trips to a grocery store with an on-site gasoline retailer. In the first instance t, the prices are $p^t = (2, 2)$ of gasoline and food respectively and she buys a bundle $x^t = (6, 3)$. In her second trip t', the prices are $p^{t'} = (3, 1)$ and she purchases $x^{t'} = (5, 4)$. The most basic welfare question one can ask here is whether the consumer is better off at the prices prevailing at t or at t' (keeping fixed the prices of all other goods she consumes)? In this paper, we introduce a theoretical framework based

[◊]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`.

We are very grateful for comments provided by the editor, the referees and numerous seminar, conference participants. We also thank Samuel Norris and Krishna Pendakur for generously sharing their data with us and John Hoddinott for patiently explaining Progresa. Deb thanks the SSHRC for their continued financial support and the Cowles Foundation for their hospitality. Stoye acknowledges support from NSF grants SES-1260980 and SES-1824375 as well as from Cornell University's Economics Computer Cluster Organization, which was partially funded through NSF Grant SES-0922005. Part of this work was done while Stoye was at Bonn University and the Hausdorff Center for Mathematics. We thank Michael Sullivan and Matthew Thirkettle for excellent research assistance.

on revealed preference, along with a nonparametric econometric technique, that would allow us to answer questions of this type.

A typical approach to this problem is to model the consumer as having a quasilinear utility $\widetilde{U}(x) - p \cdot x$ since, in particular, this allows for simple "sufficient statistics" analysis of welfare gains or losses using a Harberger formula (see Chetty, 2009 and most recently Kleven, 2021 for an overview of this approach). The second term $(-p \cdot x)$ in the quasilinear utility captures the fact that the goods being analyzed (food and gasoline in our example) do *not* constitute the universe of the consumer's consumption, so that expenditure lowers utility because it reduces the consumption of an outside (numeraire) good.

The point of departure of our analysis is the simple observation that, even without modeling the consumer's preference as quasilinear (or taking any other functional form) we can still conclude that she is better off at t compared to t'. This is because $p^{t'} \cdot x^{t'} = 19$ whereas $p^t \cdot x^{t'} = 18$. If the prices at t' were p^t instead of $p^{t'}$, the consumer would be better off since purchasing the same bundle $x^{t'}$ would cost less, leaving her with more money to buy other goods (apart from gasoline and food).[1] More generally, the consumer has a *preference over prices* that an analyst could at least partially discern from the data: if at observations t and t', we find that $p^t \cdot x^{t'} \leq (<) p^{t'} \cdot x^{t'}$, then

the consumer has revealed that she (strictly) prefers the price p^t to the price $p^{t'}$.

Welfare comparisons made in this way are consistent only if the revealed preference relation over prices has no cycles, a property we call the *generalized axiom of price preference* (GAPP). A natural question then arises: what does GAPP mean for consumer behavior?

Augmented Utility Functions: To answer this question, we assume that the analyst collects a data set $\mathcal{D} = \{p^t, x^t\}_{t=1}^T$ from a consumer; each observation t consists of the prices $p^t \in \mathbb{R}^L_{++}$ of L goods (representing some but not all the goods she consumes) and the consumer's demand $x^t \in \mathbb{R}^L_+$ at those prices. We show that GAPP (on $\mathcal{D}$) is both necessary and sufficient for the existence of a strictly increasing function $U : \mathbb{R}^L_+ \times \mathbb{R}_- \to \mathbb{R}$ that *rationalizes* $\mathcal{D}$ in the following sense:

$$x^t \in \operatorname{argmax}_{x \in \mathbb{R}^L_+} U(x, -p^t \cdot x) \text{ for all } t = 1, 2, ..., T.$$

The function U should be interpreted as an *expenditure-augmented utility function*, where $U(x, -e)$ is the consumer's utility when she purchases x after spending e. Note that the consumer's optimal expenditure on the observed goods is dependent on prices: she could in principle spend more than what she actually spent (as she optimizes over $x \in \mathbb{R}^L_+$) but the trade-off is the disutility of greater expenditure. Observe that quasilinear utility $U(x, -p \cdot x) = \widetilde{U}(x) - p \cdot x$ is a special case of an augmented utility function.

[1] Another way of seeing this is the following. Suppose t' is a supermarket where the prices are $p^{t'}$ and we observe the bundle $x^{t'}$ being bought by a consumer. If at supermarket t, the prices are p^t, then we know that the consumer would prefer this supermarket, since the same purchases at t' would cost less at t.

Below are features of the augmented utility model that make it widely applicable.

(1) Being more general, it does not have the overly strong implications of the quasilinear model on consumer demand (see Section 2.3). In particular, it is broad enough to accommodate several prominent behavioral economics models such as reference dependence, mental budgeting and inattention to prices. We briefly describe the first of these here, Section 2.4 has a more detailed discussion. Kőszegi and Rabin (2006) and Heidhues and Kőszegi (2008) argue that consumption decisions can depend not just on the actual prices but also on the prices the consumer *expected* to pay. Specifically, the disutility from spending is greater if the expected price was lower than the sticker price and vice versa. A simple way they propose of capturing this phenomenon is the following function

$$U(x, -px) = \widetilde{U}(x) - p \cdot x - F(p \cdot x - \tilde{p} \cdot x).$$

The first two terms capture standard quasilinear preferences whereas the third term captures a general form of reference dependence.[2] In Koszegi and Rabin's terminology, the consumer gets "gain-loss utility" by comparing the expenditure $p \cdot x$ she incurs on a bundle x against the expenditure $\tilde{p} \cdot x$ she expected to incur, where $\tilde{p}$ are her reference prices.[3] While the introduction of the third term means that this utility is no longer quasilinear, it remains an augmented utility function (when $\tilde{p}$ is kept fixed).

(2) Since a consumer's utility at prices p is given by $\max_{x \in \mathbb{R}^L_+} U(x, -p \cdot x)$, this obviously leads to a ranking or preference on prices. Going further, it is possible to develop notions analogous to compensating and equivalent variations which gives us a *quantitative* sense of how much one set of prices is preferred to another (see Section 3.3).

(3) Readers familiar with Afriat's Theorem (Afriat, 1967) will no doubt have already noticed that we are working in a similar framework. That theorem characterizes a data set $\mathcal{D} = \{p^t, x^t\}_{t=1}^T$ that could be rationalized in the following sense: there is $\widetilde{U} : \mathbb{R}^L_+ \to \mathbb{R}$ such that $\widetilde{U}(x^t) \geq \widetilde{U}(x)$ for all $x \in \mathbb{R}^L_+$ that satisfy $p^t \cdot x \leq p^t \cdot x^t$. The notion of rationalization in our model is distinct from that in Afriat's Theorem (even the utility functions have different domains) and there are data sets that could be rationalized in one sense but not the other. We explain these differences in Sections 2.2 and 3.1.

(4) In Section 4, we show that our framework generalizes to accommodate discrete choice, characteristics models, and nonlinear prices.

Random Augmented Utility Model (RAUM): In the second part of the paper, we develop the random version of the augmented utility model, in order to study the demand distribution of a population of consumers drawn from repeated cross-sectional data. We first devise a test to check if the data are consistent with the RAUM. We then develop a

[2]For a related model of reference prices leading to a similar functional form, see Sakovics (2011).

[3]A common choice for F is $F(p \cdot x - \tilde{p} \cdot x) = \max\{\bar{k}(p \cdot x - \tilde{p} \cdot x), 0\} + \min\{\underline{k}(p \cdot x - \tilde{p} \cdot x), 0\}$ with $\bar{k} > \underline{k} > 0$ or that the consumer feels losses relative to the reference point more severely than commensurate gains.

procedure to *estimate* the *proportion* of consumers who are made better or worse off by a given change in prices; welfare analysis of this kind under general preference heterogeneity is a challenging empirical issue, and has attracted considerable recent research (see, for example, Hausman and Newey (2016) and its references).

Unlike the case of data collected from a single individual, it is worth noting that, in this case, both model testing and welfare analysis are statistical since we need to account for sampling error inherent in repeated cross sectional data. Our RAUM test uses existing (though recently developed) econometric methods. On the other hand, to carry out the welfare analysis, we have to develop new theoretical econometric results. This is a standalone contribution that has applications beyond this paper.

We argue that testing the RAUM on actual repeated cross-sectional data (such as household survey data) is easier than testing the random utility model (RUM), that is, the random version of the standard model where consumers maximize utility subject to a budget constraint. A test for the RUM is described in McFadden and Richter (1991) (henceforth MR), but two hurdles must be overcome before their test can be implemented. First, MR do not account for finite sample issues as they assume that population distributions of demand are observed; recently, Kitamura and Stoye (2018) (henceforth KS) developed a testing procedure which incorporates sampling error. Second, the test suggested by MR requires the observation of large samples of consumers who face not only the same prices but also make *identical* total expenditures (on the observed goods). This feature is not true of any real observational data where a consumer's demand (and thus total expenditure) on the observed goods is typically price dependent. Thus to implement their test, KS first estimate demand distributions at a fixed (median) level of total expenditure, which requires the use of an instrumental variable technique (with all its attendant assumptions) to adjust for the dependence of the total expenditure on prices

In contrast, the RAUM can be tested directly on household survey data, even when the demand distribution at a given price vector implies *heterogenous levels of total expenditure across consumers*.[4] This allows us to estimate the demand distribution by simply using sample frequencies and we can avoid the above-mentioned additional layer of demand estimation needed for testing the RUM.

The reason for this remarkable simplification is somewhat ironic: we show that a data set is consistent with the RAUM if, and only if, a converted version of the data set which has identical total expenditures among consumers at each price passes the RUM test devised by MR. In other words, we apply the test suggested by MR, but not for the model they have in mind. This trick also means that we can use, and in a more straightforward way, the econometric techniques in KS.

[4]We allow two demands x, x' in the support of the demand distribution at prices p^t to satisfy $p^t \cdot x \neq p^t \cdot x'$.

Further, we can evaluate the welfare impact of an observed prices change. If we observe the true distribution of demand at each price, theoretical bounds can be derived for the population proportion who are revealed better or worse off after a price change. Of course, for finite samples, these bounds instead have to be estimated. To do so, we develop new econometric techniques to estimate confidence intervals on these population proportions; our methodology builds on the econometric theory in KS but is novel.

We emphasize that these new econometric techniques can be more generally applied to linear hypothesis testing of parameter vectors that are partially identified, even in models that are unrelated to demand theory (see, for example, Lazzati, Quah, and Shirai (2018)). They provide a new method for estimation and inference in nonparametric counterfactual analysis and, since the evaluation of counterfactuals is an important goal of empirical research, they are potentially very useful to practitioners.

Empirical Applications: We use separate data sets to demonstrate how welfare analysis can be done using both the deterministic and random versions of our model. First, we use the deterministic model to analyze panel data from the Mexican conditional cash transfer program Progresa. We show that price changes by sellers in response to the cash transfers benefit the untreated households. Second, we conduct welfare analysis with the RAUM using repeated cross-sectional data on household expenditures from Canada and the U.K. This demonstrates how to operationalize our novel econometric methodology to conduct inference for counterfactuals.

2. The Deterministic Model

We consider an econometrician who is studying a consumer's demand for L goods. We assume an idealized environment suitable for partial equilibrium analysis, where the consumer's demand for these goods at different prices are observed, while the consumer's wealth and the prices of all other goods are held fixed.[5]

Specifically, the econometrician collects a data set with a finite number of observations; each observation t can be represented as (p^t, x^t), where $p^t \in \mathbb{R}^L_{++}$ are the prices of the L goods and $x^t \in \mathbb{R}^L_+$ is the bundle of those goods purchased by the consumer.[6] We denote the data set by $\mathcal{D} := \{(p^t, x^t)\}_{t=1}^T$. (We shall slightly abuse notation and use T to refer both to the (finite) number of observations and to the set $\{1, \ldots, T\}$; similarly, L could denote both the number, and the set, of commodities.)

We begin with a basic question: given $\mathcal{D}$, can the econometrician sign the welfare impact of a price change from p^t to $p^{t'}$? Perhaps the most intuitive welfare comparison that can be made is as follows: if at prices $p^{t'}$, the econometrician finds that $p^{t'} \cdot x^t < p^t \cdot x^t$

[5]Under fairly standard (but strong) assumptions, changes to the external environment can be precisely justified by deflating the prices of the L goods (see Section 3.5).

[6]We postpone the discussion of discrete consumption spaces and nonlinear pricing to Section 4.

then he may conclude that the agent is better off at the price vector $p^{t'}$ compared to p^t. This is because, at the price $p^{t'}$ the consumer can, if she wishes, buy the bundle bought at p^t and she would still have money left over to buy other things, so she must be strictly better off at $p^{t'}$. This ranking is eminently sensible, but can it lead to inconsistencies?

Example 1. Consider a two observation data set

$$p^t = (2,1)\,,\; x^t = (4,0) \;\text{ and }\; p^{t'} = (1,2)\,,\; x^{t'} = (0,1).$$

Since $p^{t'} \cdot x^t < p^t \cdot x^t$, it seems that the consumer is better off at prices $p^{t'}$ than at p^t; however, it is also true that $p^{t'} \cdot x^{t'} > p^t \cdot x^{t'}$, which gives the opposite conclusion.

This example shows that a consistent welfare comparison at different prices requires the imposition of some restriction on the data set. To be precise, define the binary relations $\succeq_p$ and $\succ_p$ on $\mathcal{P} := \{p^t\}_{t\in T}$, the set of price vectors observed in $\mathcal{D}$, in the following manner:

$$p^{t'} \succeq_p (\succ_p) p^t \text{ if } p^{t'} \cdot x^t \leq (<) p^t \cdot x^t.$$

We say that price $p^{t'}$ is *directly (strictly) revealed preferred* to p^t if $p^{t'} \succeq_p (\succ_p) p^t$, that is, whenever the bundle x^t is (strictly) cheaper at prices $p^{t'}$ than at prices p^t. We denote the transitive closure of $\succeq_p$ by $\succeq_p^*$, that is, for $p^{t'}$ and p^t in $\mathcal{P}$, we have $p^{t'} \succeq_p^* p^t$ if there are t_1, t_2,...,t_N in T such that $p^{t'} \succeq_p p^{t_1}$, $p^{t_1} \succeq_p p^{t_2}$,..., $p^{t_{N-1}} \succeq_p p^{t_N}$, and $p^{t_N} \succeq_p p^t$; in this case we say that $p^{t'}$ is *revealed preferred* to p^t. If anywhere along this sequence, $\succeq_p$ can be replaced with $\succ_p$ then we say that $p^{t'}$ is *revealed strictly preferred* to p^t and denote that relation by $p^{t'} \succ_p^* p^t$.[7] The following restriction, which excludes circularity in the assessment of the consumer's wellbeing at different prices, is a bare minimum condition to impose on $\mathcal{D}$.

Definition 2.1. The data set $\mathcal{D} = \{(p^t, x^t)\}_{t=1}^T$ satisfies the *Generalized Axiom of Price Preference* or *GAPP* if there are no observations $t, t' \in T$ such that $p^{t'} \succeq_p^* p^t$ and $p^t \succ_p^* p^{t'}$.

This in turn leads naturally to the following question: if a consumer's observed demand behavior obeys GAPP, what could we say about her decision making procedure?

2.1. *The Expenditure-Augmented Utility Model*

An *expenditure-augmented utility function* or simply, an *augmented utility function*, is a function $U : \mathbb{R}_+^L \times \mathbb{R}_- \to \mathbb{R}$, where $U(x, -e)$ is the consumer's utility when she spends e to purchase bundle x. We require that $U(x, -e)$ is strictly increasing in the last argument (in other words, is strictly decreasing in expenditure), which captures the tradeoff the consumer faces between consuming x and consuming other goods (outside the set L).

[7]Notice that it makes sense to write $\hat{p} \succeq_p p^t$ even if $\hat{p}$ is not in $\mathcal{P}$, since the demand at $\hat{p}$ is not needed in the definition revealed preference. Similarly, it is possible to define $\hat{p} \succ_p p^t$ and the transitive extensions $\hat{p} \succeq_p^* p^t$ and $\hat{p} \succ_p^* p^t$. This observation is useful later on, in Sections 3.3 and 5.3. .

At a given price p, the consumer chooses a bundle x to maximize $U(x, -p \cdot x)$. We denote the *indirect utility at price* p by

$$V(p) := \sup_{x \in \mathbb{R}^L_+} U(x, -p \cdot x). \tag{1}$$

If the consumer's augmented utility maximization problem has a solution at every price vector $p \in \mathbb{R}^L_{++}$, then V is also defined at those prices and this induces a reflexive, transitive, and complete preference over prices in $\mathbb{R}^L_{++}$.

A data set $\mathcal{D} = \{(p^t, x^t)\}_{t=1}^T$ is *rationalized by an augmented utility function* if there exists such a function $U : \mathbb{R}^L_+ \times \mathbb{R}_- \to \mathbb{R}$ with

$$x^t \in \operatorname{argmax}_{x \in \mathbb{R}^L_+} U(x, -p^t \cdot x) \qquad \text{for all } t \in T. \tag{2}$$

It is straightforward to see that GAPP is necessary for a data set to be rationalized by an augmented utility function. First, notice that if $p^{t'} \succeq_p p^t$, then $p^{t'} \cdot x^t \leq p^t \cdot x^t$, and so

$$V(p^{t'}) \geq U(x^t, -p^{t'} \cdot x^t) \geq U(x^t, -p^t \cdot x^t) = V(p^t).$$

Furthermore, $U(x^t, -p^{t'} \cdot x^t) > U(x^t, -p^t \cdot x^t)$ if $p^{t'} \succ_p p^t$, and in that case $V(p^{t'}) > V(p^t)$. Suppose GAPP is not satisfied and there are two observations $t, t' \in T$ such that $p^{t'} \succeq_p^* p^t$ and $p^t \succ_p^* p^{t'}$. Then there exist $t_1, t_2, \ldots, t_N \in T$ which yield the contradiction

$$V(p^{t'}) \geq V(p^{t_1}) \geq \cdots \geq V(p^{t_N}) \geq V(p^t) > V(p^{t'}).$$

Our main theoretical result, stated next, also establishes the sufficiency of GAPP for rationalization. Moreover, whenever $\mathcal{D}$ can be rationalized, it can be rationalized by an augmented utility function U with a list of properties that make it convenient for analysis.

Theorem 1. *Given a data set $\mathcal{D} = \{(p^t, x^t)\}_{t=1}^T$, the following are equivalent:*

(1) $\mathcal{D}$ is rationalized by an augmented utility function.

(2) $\mathcal{D}$ satisfies GAPP.

(3) $\mathcal{D}$ is rationalized by an augmented utility function U that is strictly increasing, continuous, and concave. Moreover, U is such that $\max_{x \in \mathbb{R}^L_+} U(x, -p \cdot x)$ has a solution for all $p \in \mathbb{R}^L_{++}$.

2.2. *Afriat's Theorem and Proof of Theorem 1*

Before we prove Theorem 1, it is worth providing a short description of the standard revealed preference theory of consumer demand and its central result, Afriat's Theorem. This will be useful not just because we will invoke the result several times but also since it will serve as an important point of contrast for our axiom and results.

The standard theory due to Afriat (1967) is built formally on the same primitives as our model: a finite data set of prices and corresponding consumption bundles. Unlike our model however, it is assumed that the observed goods correspond to the *universe* of the consumer's consumption. Formally, a data set $\mathcal{D}$ is said to be *rationalized by a utility function* if there exists a locally nonsatiated[8] utility function $\widetilde{U} : \mathbb{R}^L_+ \to \mathbb{R}$ such that

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

In words, this criterion asks whether there is a utility function defined over the L observed goods such that the consumer is utility maximizing at every observation t over the fixed budget $p^t \cdot x^t$ corresponding to the observed expenditure.

Of course, data sets (outside of laboratory data) almost never contain the universe of consumed goods and the consumer's *true* budget set is not observed, especially when one takes into account the possibility of borrowing and saving. Given this, when checking if a data set $\mathcal{D}$ can be rationalized in the sense of (3), we are effectively testing whether the consumer is maximizing a sub-utility function $\widetilde{U} : \mathbb{R}^L_+ \to \mathbb{R}$ defined specifically on those L goods (or equivalently, has weakly separable preferences).

It should be clear that rationalization in the sense of (3) is distinct from rationalization by an augmented utility function. The augmented utility model specifically takes into account the impact of the prices of these L goods on the consumption of other goods: it is *necessarily* a partial equilibrium model, and designed for partial equilibrium welfare analysis of the type carried out in empirical industrial organization or public economics. An example is the study of the welfare impact of a sales tax levied on a subset of goods.

Revealed preference in Afriat's setting is captured by two binary relations, $\succeq_x$ and $\succ_x$ which are defined as follows on the set $\mathcal{X} := \{x^t\}_{t \in T}$ of chosen bundles observed in $\mathcal{D}$:

$$x^{t'} \succeq_x (\succ_x)\, x^t \text{ if } p^{t'} \cdot x^{t'} \geq (>)\, p^{t'} \cdot x^t.$$

We say that the bundle $x^{t'}$ is *directly revealed (strictly) preferred* to the bundle x^t if $x^{t'} \succeq_x (\succ_x)\, x^t$, that is, whenever x^t is (strictly) cheaper at prices $p^{t'}$ than $x^{t'}$. This terminology is intuitive: if the agent is maximizing some locally nonsatiated utility function $\widetilde{U} : \mathbb{R}^L_+ \to \mathbb{R}$, then $x^{t'} \succeq_x x^t$ ($x^{t'} \succ_x x^t$) must imply that $\widetilde{U}(x^{t'}) \geq (>)\, \widetilde{U}(x^t)$.

We denote the transitive closure of $\succeq_x$ by $\succeq^*_x$, that is, for $x^{t'}, x^t \in \mathcal{X}$, we have $x^{t'} \succeq^*_x x^t$ if there are t_1, t_2,...,t_N in T such that $x^{t'} \succeq_x x^{t_1}$, $x^{t_1} \succeq_x x^{t_2}, \ldots, x^{t_{N-1}} \succeq_x x^{t_N} \succeq_x x^t$, and $x^{t_N} \succeq_x x^t$; in this case, we say that $x^{t'}$ is *revealed preferred* to x^t. If any $\succeq_x$ in this sequence can be replaced with $\succ_x$, we say that $x^{t'}$ is *revealed strictly preferred* to x^t and denote that relation by $x^{t'} \succ^*_x x^t$. Clearly, if $\mathcal{D}$ is rationalizable by some locally nonsatiated utility function $\widetilde{U}$, then $x^{t'} \succeq^*_x (\succ^*_x)\, x^t$ implies that $\widetilde{U}(x^{t'}) \geq (>)\, \widetilde{U}(x^t)$. Thus, a necessary condition for $\mathcal{D}$'s rationalizability is that the revealed preference relation has no cycles.

Definition 2.2. A data set $\mathcal{D} = \{(p^t, x^t)\}_{t=1}^T$ satisfies the *Generalized Axiom of Revealed Preference* or *GARP* if there are no observations $t, t' \in T$ such that $x^{t'} \succeq^*_x x^t$ and $x^t \succ^*_x x^{t'}$.

[8] This means that at any bundle x and open neighborhood of x, there is a bundle x' in the neighborhood with strictly higher utility.

The main claim of Afriat's Theorem is that this condition is also *sufficient* (the formal statement can be found in the online Appendix A.1.1). Having described Afriat's Theorem, we are now in a position to prove Theorem 1.

PROOF OF THEOREM 1. We will show that (2) $\implies$ (3). We have already argued that (1) $\implies$ (2) and (3) $\implies$ (1) by definition.

Choose a number $M > \max_t p^t \cdot x^t$ and define the augmented data set $\widetilde{\mathcal{D}} = \{(p^t, 1), (x^t, M - p^t \cdot x^t)\}_{t=1}^T$. This data set augments $\mathcal{D}$ since we have introduced an $L+1^{\text{th}}$ good, which we have priced at 1 across all observations, with the demand for this good equal to $M - p^t \cdot x^t$.

The crucial observation to make here is that

$$(p^t, 1)(x^t, M - p^t \cdot x^t) \geq (p^t, 1)(x^{t'}, M - p^{t'} \cdot x^{t'}) \text{ if and only if } p^{t'} \cdot x^{t'} \geq p^t \cdot x^{t'},$$

which means that

$$(x^t, M - p^t \cdot x^t) \succeq_x (x^{t'}, M - p^{t'} \cdot x^{t'}) \text{ if and only if } p^t \succeq_p p^{t'}.$$

Similarly,

$$(p^t, 1)(x^t, M - p^t \cdot x^t) > (p^t, 1)(x^{t'}, M - p^{t'} \cdot x^{t'}) \text{ if and only if } p^{t'} \cdot x^{t'} > p^t \cdot x^{t'},$$

and so

$$(x^t, M - p^t \cdot x^t) \succ_x (x^{t'}, M - p^{t'} \cdot x^{t'}) \text{ if and only if } p^t \succ_p p^{t'}.$$

Consequently, $\mathcal{D}$ satisfies GAPP if and only if $\widetilde{\mathcal{D}}$ satisfies GARP. Applying Afriat's Theorem when $\widetilde{\mathcal{D}}$ satisfies GARP, there is $\widetilde{U} : \mathbb{R}^{L+1} \to \mathbb{R}$ (notice that $\widetilde{U}$ is defined on $\mathbb{R}^{L+1}$ and not just $\mathbb{R}_+^{L+1}$; see Remark 3 in Appendix A.1.1) such that

$$(x^t, M - p^t \cdot x^t) \in \operatorname{argmax}_{\{(x,m) \in \mathbb{R}_+^L \times \mathbb{R} \,:\, p^t \cdot x + m \leq M\}} \widetilde{U}(x, m) \qquad \text{for all } t \in T. \tag{4}$$

The function $\widetilde{U}$ can be chosen to be strictly increasing, continuous, and concave, and the lower envelope of a finite set of affine functions. Clearly, the augmented utility function $\overline{U} : \mathbb{R}_+^L \times \mathbb{R}_- \to \mathbb{R}$ defined by $\overline{U}(x, -e) := \widetilde{U}(x, M - e)$ is strictly increasing in $(x, -e)$, continuous, concave and rationalizes $\mathcal{D}$.

Define $\widehat{U} : \mathbb{R}_+^L \times \mathbb{R}_- \to \mathbb{R}$ by

$$\widehat{U}(x, -e) := \overline{U}(x, -e) - h(\max\{0, e - M\}), \tag{5}$$

where $h : \mathbb{R}_+ \to \mathbb{R}$ is a differentiable function satisfying $h(0) = 0$, $h'(k) > 0$, $h''(k) \geq 0$ for $k \in \mathbb{R}_+$, and $\lim_{k \to \infty} h'(k) = \infty$. (For example, $h(k) = k^3$.) Like $\overline{U}$, the function $\widehat{U}$ is strictly increasing in $(x, -e)$, continuous and concave and x^t solves $\max_{x \in \mathbb{R}_+^L} \widehat{U}(x, -p^t x)$ (because $\widehat{U}(x, -e) \leq \overline{U}(x, -e)$ for all $(x, -e)$, and $\widehat{U}(x^t, -p^t \cdot x^t) = \overline{U}(x^t, -p^t \cdot x^t)$). Furthermore, for every $p \in \mathbb{R}_{++}^L$, $\operatorname{argmax}_{x \in X} \widehat{U}(x, -p \cdot x)$ is nonempty.[9] ■

[9]Choose a sequence $x^n \in \mathbb{R}_+^L$ such that $\widehat{U}(x^n, -p \cdot x^n)$ tends to $\sup_{x \in \mathbb{R}_+^L} \widehat{U}(x, -p \cdot x)$ (which we allow to be infinity). It is impossible for $p \cdot x^n \to \infty$ because the piecewise linearity of $U(x, -e)$ in x and the assumption

Note that GARP imposes testable restrictions distinct from GAPP (we postpone a detailed discussion to Section 3.1). This can be seen from Figure 1 which plots not just the observed consumption bundles from Example 1 but also the corresponding budget sets (derived from the observed prices and expenditures). As we argued, GAPP does not hold

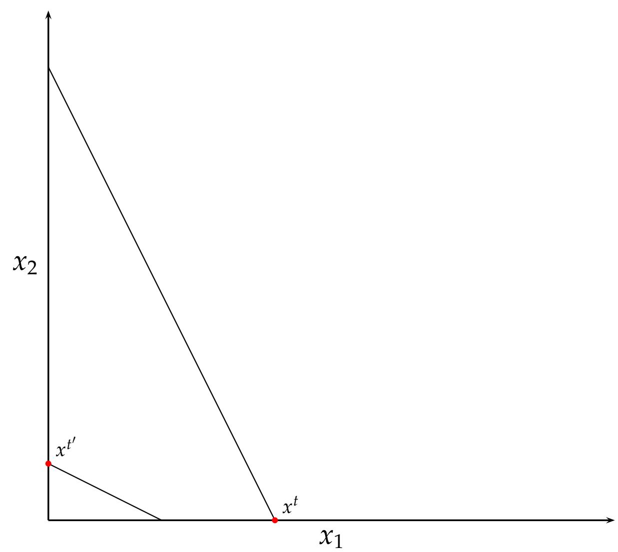


FIGURE 1. GARP consistent choices that don't allow for consistent welfare predictions.

but, since the budget sets do not even cross, it is clear that GARP does.

From this point onwards, when we refer to 'rationalization' without additional qualifiers, we shall mean rationalization by an augmented utility function, that is, in the sense given by (2) rather than in the sense given by (3).

2.3. *'Standard' consumer theory and the augmented utility function*

Perhaps the clearest motivation for our model is to think of it as a generalization of the quasilinear utility model, in which the consumer derives utility $\widetilde{U}(x)$ from the bundle x and maximizes utility net of expenditure, that is, she chooses x to maximize

$$U(x, -e) := \widetilde{U}(x) - e, \tag{6}$$

where $e = p \cdot x$. The familiar textbook way of justifying this objective function is to think of the consumer as having a utility function $\overline{U}$ defined over $L+1$ goods, with the last 'outside' good entering additively and linearly into the utility function, so that $\overline{U}(x, z) = \widetilde{U}(x) + z$. If the consumer has total wealth W, the utility of buying bundle $x \in \mathbb{R}^L_+$ is then

$$\overline{U}(x, W - p \cdot x) = \widetilde{U}(x) - p \cdot x + W.$$

that $\lim_{k\to\infty} h'(k) \to \infty$ implies that $\widehat{U}(x^n, -p \cdot x^n) \to -\infty$. So the sequence $p \cdot x^n$ is bounded, which in turn means that there is a subsequence of x^n that converges to $x^\star \in \mathbb{R}^L_+$. By the continuity of $\widehat{U}$, we obtain $\widehat{U}(x^\star, -p \cdot x^\star) = \sup_{x \in \mathbb{R}^L_+} \widehat{U}(x, -p \cdot x)$.

Ignoring boundary issues, the consumer is effectively maximizing (6).

Despite widespread use in partial equilibrium analysis, the complete absence of income effects makes the quasilinear model unsuitable for certain empirical applications. For this reason, it is common to relax the linearity of $\overline{U}$ while retaining the assumption that outside consumption is captured by a *single* outside good; this is true, for example, in the literature on modeling the demand for differentiated goods.[10] Then the utility of purchasing bundle $x \in \mathbb{R}^L_+$ is $\overline{U}(x, W - p \cdot x)$ and provided W is fixed, the consumer effectively maximizes an augmented utility function: simply let $U(x, -e) = \overline{U}(x, W - e)$.

Obviously, a consumer's outside consumption opportunities would in reality involve more than one good, and the prices of those outside goods could change as well. Within the familiar constrained-optimal model of consumer theory, there are known conditions that justify the representation of those consumption opportunities by a representative good (with its corresponding price index). This is explained in detail in Section 3.5.

Finally, it is worth mentioning that the augmented utility function captures, as a special case, quasilinear utility maximization subject to certain constraints. One example is consumption with a subsistence constraint, which we describe in the empirical application in Section 7.1. Loosely speaking, we can capture constraints on $(x, -e)$ with an augmented utility $U(x, -e)$ that assigns very low values at $(x, -e)$ that violate the constraint.

2.4. *Behavioral preferences captured by the augmented-utility model*

The central feature of the augmented utility model is that consumers experience disutility from expenditure. As we explained in the previous subsection, this disutility could be interpreted in a purely opportunity cost sense – more expenditure on the consumed goods imply less money available for other goods. In this understanding, the augmented utility function is a reduced form of a broader 'true' utility function defined on all goods.

However, it is also reasonable to think that the consumer has – *directly* – a preference over bundles of the observed L goods and their associated expenditure, which she has developed as a way of guiding purchasing decisions. Thus it is the basic object of analysis and not the reduced form of something more fundamental. This understanding of choice behavior is exploited in the behavioral economics literature and the following quote from Prelec and Loewenstein (1998) is effectively a description of augmented utility:

> *each time a consumer engages in an episode of consumption, we assume she asks herself: "How much is this pleasure costing me?" The answer to this question is the imputed cost of consumption. This imputed cost is "real" in the sense that it actually detracts from consumption pleasure.*

[10] For example, in Berry, Levinsohn, and Pakes (1995) and in Nevo (2000), $\overline{U}$ is additively separable between the L goods and the outside good; in the former, the utility of consuming y units of the outside good is $\alpha \ln y$, for some $\alpha > 0$, whereas in the latter it is αy (in other words, $\overline{U}$ is quasilinear). In Bhattacharya (2015), $\overline{U}$ is allowed to be a general function defined on $L + 1$ goods.

In this understanding, the disutility of expenditure is still related to opportunity cost, but the relationship is more flexible than what is permitted in a classical framework.

In the Introduction, we described how reference-dependent preferences could be captured by an augmented utility function. In the remainder of this section, we describe how our model relates to two other prominent themes in the behavioral literature.

Inattention to Prices and Expenditure

Chetty, Looney, and Kroft (2009) (and the literature that followed) observe that consumers often misperceive prices: specifically, they show that shoppers at grocery stores often do not internalize the price effect of taxes. Gabaix (2019) summarizes this literature and argues that many behavioral biases take the form of inattention.

Our model captures a version of price inattention discussed in Bordalo, Gennaioli, and Shleifer (2013) and Gabaix (2014). Here, a consumer perceives the expenditure of a bundle x at prices p as $f(x, p \cdot x)$, where f is increasing in the true expenditure $p \cdot x$ and can potentially depend on x. With this misperception, a consumer with a quasilinear preference chooses $x \in \mathbb{R}^L_+$ to maximize

$$U(x, -p \cdot x) = \widetilde{U}(x) - f(x, p \cdot x) \tag{7}$$

A special case is where a consumer has a default price p^d and misperceives the actual price p to be $ap + (1-a)p^d$ where $a \in [0,1]$ is the 'attention parameter.' The perceived expenditure is then $f(x, p \cdot x) = ap \cdot x + (1-a)p^d \cdot x$. More generally, the model accommodates $f(x, p \cdot x) = a(x)p \cdot x + (1 - a(x))p^d \cdot x$, where the attention parameter $a(x) \in [0,1]$ varies across bundles.[11] Among other things, this allows a consumer to be more attentive if she is purchasing large bundles compared to small ones (so that $a(x)$ tends to 1 when x is large). Another possibility is that the consumer pays attention to marginal increases in expenditure only when certain thresholds are crossed; this would correspond to the case where f is a step function that depends only on the expenditure $e = p \cdot x$. Clearly, inattention as modeled by (7) is an instance where the agent has an augmented utility function, even though it will typically not be quasilinear (in actual expenditure).[12]

Notice that using an augmented utility function (such as (7)) to capture price inattention is particularly apt because, as noted by Gabaix (2014), the numeraire serves as "the shock absorber that adjusts to the budget constraint." The alternative is to model the consumer as having *both* price misperception over a given set of goods *and* a budget on those goods that must be satisfied, which inevitably leads to the added complication of modelling

[11]This formulation of perceived expenditure is more general than Gabaix (2014) in that it allows the attention parameter a to depend on x but is less general in that the parameter does not vary across goods.

[12]We should add that formulae such as (7) would typically have observable implications that are stronger than GAPP. In other words, price inattention models do *not* have precisely the same empirical content as the augmented utility model.

how the agent adjusts her intended demand when she realizes it violates (because prices are misperceived) the budget constraint at the true prices.

Budgeting and Mental Budgeting

As explained in Section 2.3, it is common in partial equilibrium analysis to introduce a numeraire good and assume that the agent has a (standard) utility function and budget set defined on $L+1$ goods, with price and income information used to determine the level of the numeraire consumed. Obviously, this approach requires income information which is not always in the data[13] and even when it is available, it is strictly speaking not the right value to use as the global budget if the consumer can save and borrow (as acknowledged, for example, in Hausman and Newey (2016)). More generally, determining 'the real budget' is not always straightforward, even in a classical setting.

Regularities highlighted by behavioral economists add a further wrinkle to the concept of a budget. It has been widely observed that households do not always treat money as fungible and instead create separate accounts for various categories of goods (Thaler, 1999). This is not only true for consumption decisions (see, for instance, Hastings and Shapiro (2013, 2018)) but also for savings decisions, which is why consumers often save more when they have access to commitment savings options (important theoretical and empirical contributions are Amador, Werning, and Angeletos (2006) and Feldman (2010), Dupas and Robinson (2013) respectively).

Now consider a researcher modeling the demand for a subset of L goods. If mental accounting effects are important, she must allow for the fact that she cannot observe how the agent categorizes goods, nor does she know the true mental budget that determines expenditure (on the L observed goods and their perceived alternatives). In this situation, augmented utility is a natural way to model the demand for those L goods: it is consistent with constrained utility maximization incorporating an outside good (see Section 2.3) but *does not require* the researcher to take a stand on the unobserved mental budget.[14]

3. Properties of the augmented utility model

In this section, we explore various aspects of the augmented utility model, beginning with a discussion of the relationship between GAPP and GARP.

[13] Several widely used data sets, such as supermarket scanner panel data, that contain rich purchase information, do not have accurate income measures. Here, income information is typically the self-reported category (income ranges) when households apply for loyalty cards (so even this information becomes dated).

[14] Here, the assumption is that the data spans a period over which the mental budget for the observed and unobserved goods is stable. Varying mental budgets would manifest as GAPP violations (see Example 3).

3.1. *Comparing GAPP and GARP*

Recall that Example 1 in Section 2 is an example of a data set that obeys GARP but fails GAPP. We now present an example of a data set that satisfies GAPP but fails GARP.

Example 2. Consider the data set consisting of the following two choices:

$$p^t = (2,1)\,,\; x^t = (2,1) \;\text{ and }\; p^{t'} = (1,4)\,,\; x^{t'} = (0,2).$$

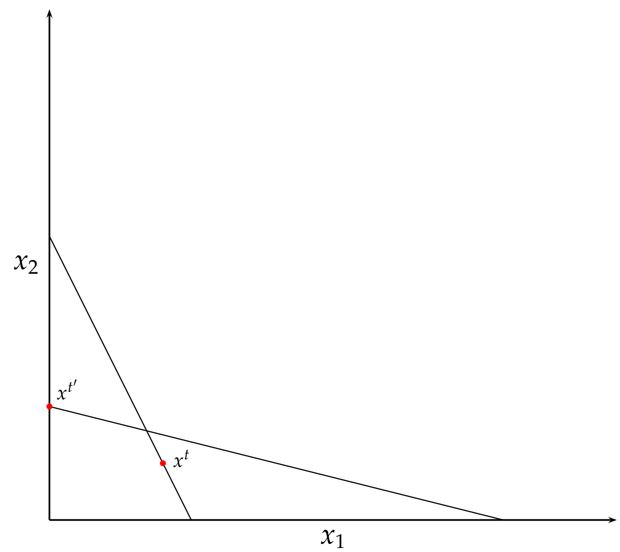


FIGURE 2. Choices that satisfy GAPP but not GARP

These choices, as shown in Figure 2, violate GARP as $p^t \cdot x^t = 5 > 2 = p^t \cdot x^{t'}$ ($x^t \succ_x x^{t'}$) and $p^{t'} \cdot x^{t'} = 8 > 6 = p^{t'} \cdot x^t$ ($x^{t'} \succ_x x^t$). However, these choices satisfy GAPP as $p^{t'} \cdot x^{t'} = 8 > 2 = p^t \cdot x^{t'}$ ($p^t \succ_p p^{t'}$) but $p^t \cdot x^t = 5 \not\geq 6 = p^{t'} \cdot x^t$ ($p^{t'} \not\succeq_p p^t$).

GAPP and GARP are distinct conditions but they coincide in data sets where $p^t \cdot x^t = 1$ for all $t \in T$. This is because $x^t \succeq_x (\succ_x) x^{t'}$ if and only if $p^t \succeq_p (\succ_p) p^{t'}$ as both require $1 \geq (>)\, p^t \cdot x^{t'}$. Given a data set $\mathcal{D} = \{(p^t, x^t)\}_{t=1}^T$, we define the *iso-expenditure version of* $\mathcal{D}$ as a data set $\breve{\mathcal{D}} := \left\{(p^t, \breve{x}^t)\right\}_{t=1}^T$, such that $\breve{x}^t = x^t/(p^t \cdot x^t)$ and note that $p^t \cdot \breve{x}^t = 1$ for all $t \in T$. Observe that the revealed price preference relations $\succeq_p$, $\succ_p$ remain unchanged when consumption bundles are scaled. Thus a data set obeys GAPP if and only if its iso-expenditure version obeys GAPP, which in this case is equivalent to GARP.[15]

[15] There is a similar 'GARP-version' of Proposition 1 and that result has been exploited before in the literature (see, for example, Sakai (1977)). Suppose $\mathcal{D} = \{(p^t, x^t)\}_{t=1}^T$ obeys GARP. Then GARP holds even if each observed price vector p^t is arbitrarily scaled. In particular, $\mathcal{D}$ obeys GARP if and only if $\hat{\mathcal{D}} = \{(\hat{p}^t, x^t)\}_{t\in T}$, where $\hat{p}^t = p^t/(p^t \cdot x^t)$, obeys GARP (equivalently, GAPP) since $\hat{p}^t \cdot x^t = 1$ for all $t \in T$. The latter perspective is useful because it highlights the possibility of applying Afriat's Theorem on $\hat{\mathcal{D}}$, *in the space of prices* (in other words, with the roles of prices and bundles reversed). This immediately gives us a different, 'dual' rationalization of $\mathcal{D}$ in terms of indirect utility, that is, there is a continuous, strictly decreasing, and convex function $\widetilde{V} : \mathbb{R}^L_{++} \to \mathbb{R}$ such that $\hat{p}^t \in \arg\min_{\{p \in \mathbb{R}^L_{++} : p \cdot x^t \geq 1\}} \widetilde{V}(p)$.

Proposition 1. *Let $\mathcal{D} = \{(p^t, x^t)\}_{t=1}^T$ be a data set and let $\breve{\mathcal{D}} = \{(p^t, \breve{x}^t)\}_{t\in T}$, where $\breve{x}^t = x^t/(p^t \cdot x^t)$. Then the revealed preference relations $\succeq_p^*$ and $\succ_p^*$ on $\mathcal{P} = \{p^t\}_{t=1}^T$ and the revealed preference relations $\succeq_x^*$ and $\succ_x^*$ on $\breve{\mathcal{X}} = \{\breve{x}^t\}_{t=1}^T$ are related in the following manner:*

(1) $p^t \succeq_p^* p^{t'}$ if and only if $\breve{x}^t \succeq_x^* \breve{x}^{t'}$.

(2) $p^t \succ_p^* p^{t'}$ if and only if $\breve{x}^t \succ_x^* \breve{x}^{t'}$.

As a consequence, $\mathcal{D}$ obeys GAPP if and only if its iso-expenditure version, $\breve{\mathcal{D}}$, obeys GARP.

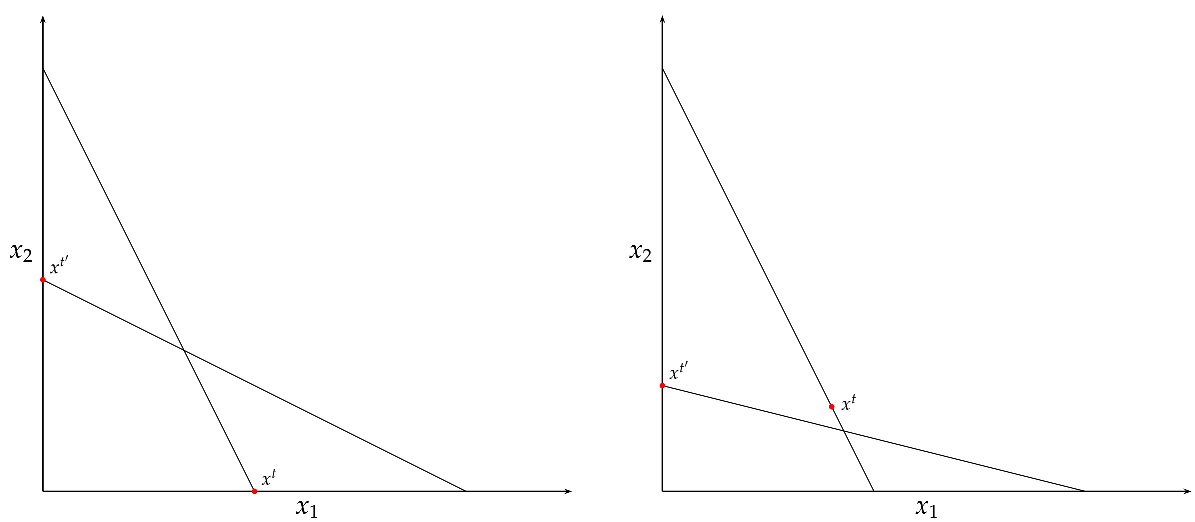
Proof. Notice that

$$p^t \cdot \frac{x^t}{p^t \cdot x^t} \geq (>) p^t \cdot \frac{x^{t'}}{p^{t'} \cdot x^{t'}} \iff p^{t'} \cdot x^{t'} \geq (>) p^t \cdot x^{t'}.$$

The left side of the equivalence says that $\breve{x}^t \succeq_x \breve{x}^{t'}$ while the right side says that $p^t \succeq_p p^{t'}$. This implies (1) since $\succeq_p^*$ and $\succeq_x^*$ are the transitive closures of $\succeq_p$ and $\succeq_x$ respectively.

Similarly, it follows (from the strict inequality version of the above equivalence) that $\breve{x}^t \succ_x \breve{x}^{t'}$ if and only if $p^t \succ_p p^{t'}$, which leads to (2). The claims (1) and (2) together guarantee that there is a sequence of observations in $\mathcal{D}$ that lead to a GAPP violation if and only if the analogous sequence in $\breve{\mathcal{D}}$ lead to a GARP violation. ■

As an illustration, compare the data sets in Figure 1 and Figure 2 to the iso-expenditure data sets in Figure 3a and Figure 3b. It can be clearly observed that the iso-expenditure data in Figure 3a contains a GARP violation (which implies it does not satisfy GAPP) whereas the data in Figure 3b does not violate GARP (and, hence, satisfies GAPP).



(a) Example 1 (b) Example 2

FIGURE 3. Expenditure-Normalized Choices

Proposition 1 implies that the augmented utility model can be tested in two ways: we can either test GAPP directly or test GARP on its iso-expenditure version. If we are simply interested in testing GAPP on a single-agent data set $\mathcal{D}$, normalization brings no advantage: the test is computationally straightforward in either case. However, as we shall see in Section 5, iso-expenditure scaling plays an important role in the test we develop (on repeated cross-sectional demand data) for the random utility version of our model.

GARP and GAPP are distinct but not mutually exclusive properties. For instance, data collected from a consumer who maximizes a quasilinear augmented utility satisfies both properties.[16] However, even when both properties are satisfied, demand predictions at an out-of-sample price will differ based on the property employed. Further discussion on the relationship between GAPP and GARP is found in online Appendix A.1.

Proposition 1 and the fact that scaling consumption bundles does not affect the revealed price preference relation makes it natural to wonder about the relationship between the augmented-utility model and the standard model (as in (3)) with homothetic preferences. A data set that is rationalized in the latter sense (Varian (1983) provides a characterization) has the feature that it must satisfy GARP for *any* arbitrary scaling of consumption bundles and thus will satisfy GAPP. By contrast, a data set that satisfies GAPP must only satisfy GARP for the particular scaling that equalizes expenditure across observations. Thus, GAPP is a less stringent property; that it is *strictly* less stringent is clear from Example 2, which satisfies GAPP but violates GARP and therefore cannot be rationalized in Afriat's sense for any locally nonsatiated preference, let alone a homothetic preference.[17]

3.2. *Preference over Prices*

We know from Theorem 1 that if $\mathcal{D}$ obeys GAPP then it can be rationalized by an augmented utility function with an indirect utility V that is defined at all price vectors in $\mathbb{R}^L_{++}$. It is easy to check that any V as defined by (1) has the following properties:

(a) V is *nonincreasing*; that is, if $p' \geq p$ (element by element) then $V(p') \leq V(p)$, and
(b) V is *quasiconvex*; that is, $V(\beta p + (1-\beta)p') \leq \max\{V(p), V(p')\}$ for all $\beta \in [0,1]$.

Any rationalizable data set $\mathcal{D}$ is potentially rationalizable by many augmented utility functions, with each one leading to a different indirect utility. We denote this set of indirect utilities by $\mathbf{V}(\mathcal{D})$. We previously noted that if $p^t \succeq_p^* (\succ_p^*)\, p^{t'}$ then $V(p^t) \geq (>)\, V(p^{t'})$ for any $V \in \mathbf{V}(\mathcal{D})$; in other words, the conclusion that the consumer prefers the prices p^t to $p^{t'}$ is *nonparametric* in that it is independent of the precise augmented utility used to rationalize $\mathcal{D}$. The next result (proved in Appendix A.2) says that, absent further information on the augmented utility, this is *all* the information on the consumer's preference

[16] When U has the form (6), x^t maximizes $U(x, -p^t \cdot x)$ only if x^t maximizes $\widetilde{U}(x)$ in $\{x \in \mathbb{R}^L_+ : p^t \cdot x \leq p^t \cdot x^t\}$. Thus $\mathcal{D}$ must also obey GARP. A broader class of augmented utility functions that satisfy both GAPP and GARP is given in Section A.1.2 of the online appendix.

[17] Example A.5 in the online appendix contrasts demand predictions using the augmented utility model and the constrained-optimization model (both with and without imposing homotheticity on the preference).

over prices in $\mathcal{P}$ that we can glean from the data. Thus, the revealed price preference relation contains the most detailed information for welfare comparisons.

Proposition 2. *Suppose $\mathcal{D} = \{(p^t, x^t)\}_{t=1}^T$ is rationalizable by an augmented utility function. Then for any p^t, $p^{t'}$ in $\mathcal{P}$:*

(1) $p^t \succeq_p^* p^{t'}$ if and only if $V(p^t) \geq V(p^{t'})$ for all $V \in \mathbf{V}(\mathcal{D})$.

(2) $p^t \succ_p^* p^{t'}$ if and only if $V(p^t) > V(p^{t'})$ for all $V \in \mathbf{V}(\mathcal{D})$.

3.3. *Compensation for a price change*

In the standard consumer model, the compensating and equivalent variations are used to quantify the welfare impact of a price change. Analogous concepts exist for the augmented utility model and bounds for them can be recovered from the data.[18]

Let U be the consumer's augmented utility function. Suppose, the price changes from p^{t_1} to p^{t_2} leading to a change of x^{t_1} to x^{t_2} in consumption. Then, there exists μ_c such that

$$\max_{x \in \mathbb{R}^L_+} U(x, -p^{t_2} \cdot x - \mu_c) = V(p^{t_1}). \tag{8}$$

Note that μ_c is unique since U is strictly increasing in the last argument. We can think of μ_c as the lump sum transferred from the consumer after the price change that makes her just indifferent between the situation before and after the change.

Suppose we interpret U as arising from an overall utility function $\widetilde{U}(x, z)$ (that depends on the observed goods x and the level z of an outside good), given the consumer's wealth of M, so that $U(x, -e) = \widetilde{U}(x, M - e)$. Since μ_c solves (8), it will also satisfy

$$\max_{\{x \in \mathbb{R}^L_+ :\, p^{t_2} \cdot x \leq M - \mu_c\}} \widetilde{U}(x, (M - \mu_c) - p^{t_2} \cdot x) \;=\; \widetilde{U}(x^{t_1}, M - p^{t_1} \cdot x^{t_1}).$$

In other words, μ_c is the reduction in total wealth that will leave the consumer's overall utility at p^{t_2} the same as it was at p^{t_1}. Thus, with this interpretation of the augmented utility function, μ_c coincides with what is called the *compensating variation* in standard consumer theory and we shall use this term to also refer to μ_c (defined by (8)).

Pushing the analogy further, we can use the compensating variation in our model in the same way it is typically used. For example, a price change from p^{t_1} to p^{t_2} may benefit some but hurt others. The Kaldor criterion deems this change an overall improvement if the sum of the compensating variations across consumers is positive as it guarantees that those who benefit could, in principle, compensate the losers and still be better off.

[18]Calculations of the compensating and equivalent variations in either the standard model or in ours assume that the preference recovered from consumption data is indeed the consumer's true preference and that is the position we take in this subsection. If behavioral considerations of the type discussed in Section 2.4 are at play, then one could conceptually distinguish between an agent's true preference and that which is recovered from the data under some type of bounded rationality (such as price inattention). The latter preference will still be useful for a positive analysis of demand (for example, in estimating behavior at out-of-sample prices) but compensation which is calculated based on such a preference could be problematic.

In a similar way, we can define the *equivalent variation* as the value μ_e that solves

$$\max_{x\in\mathbb{R}^L_+} U(x, -p^{t_1}\cdot x + \mu_e) = V(p^{t_2}). \tag{9}$$

If $U(x, -e) = \widetilde{U}(x, M - e)$, μ_e coincides with the usual equivalent variation as it solves

$$\max_{\{x\in\mathbb{R}^L_+ :\, p^{t_2}\cdot x\leq M+\mu_e\}} \widetilde{U}(x, (M + \mu_e) - p^{t_1}\cdot x) = \widetilde{U}(x^{t_2}, M - p^{t_2}\cdot x^{t_2}).$$

If $\mathcal{D}$ satisfes GAPP and contains observation (p^{t_1}, x^{t_1}), what can we say about the compensating variation of a price change from p^{t_1} to p^{t_2} (where the latter may not be a price observed in $\mathcal{D}$)? Even though there is typically a range of these values since there is more than one augmented utility that rationalizes $\mathcal{D}$, it is possible to obtain a *tight* lower bound for the set of possible compensating variation values. This is given by

$$\inf\{\mu_c : \mu_c \text{ solves (8) for some augmented utility function } U \text{ that rationalizes } \mathcal{D}\}.$$

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

We now describe how to compute this bound; we omit the analogous exercise for the equivalent variation. Let $S \subset T$ be the set of observations such that $s \in S$ if $p^s \succeq^*_p p^{t_1}$. This set is nonempty since it contains p^{t_1} itself. For each $s \in S$, there is m^s_c such that

$$p^{t_2}\cdot x^s + m^s_c = p^s x^s. \tag{10}$$

We claim that for any U that rationalizes $\mathcal{D}$, the compensating variation $\mu_c \geq m^s_c$. This is because if $m < m^s_c$, then $m \neq \mu_c$ for any utility function rationalizing $\mathcal{D}$. Indeed,

$$\begin{aligned}\max_{x\in\mathbb{R}^L_+} U(x, -p^{t_2}\cdot x - m) \geq U(x^s, -p^{t_2}\cdot x^s - m) > U(x^s, -p^{t_2}\cdot x^s - m^s_c)\\ = U(x^s, -p^s\cdot x^s) \geq U(x^{t_1}, -p^{t_1}\cdot x^{t_1}) = V(p^{t_1}).\end{aligned}$$

Thus $\inf(\mu_c) \geq m^s_c$ for all $s \in S$. In fact, it is possible to obtain a stronger conclusion:

$$\inf(\mu_c) = \max\{m^s_c : m^s_c \text{ satisfies (10) for some } s \in S\}. \tag{11}$$

Note that the right side of this equation can be easily computed from the data.

Notice that if p^{t_2} is revealed preferred to p^{t_1} (equivalently, that there is $s' \in S$ such that $m^{s'}_c \geq 0$),[19] then $\inf(\mu_c) \geq 0$; in other words, at $p = p^{t_2}$, 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 p^{t_2} is *not* revealed preferred to p^{t_1}, that is, for every $s \in S$, we have $m^s_c < 0$, then $\inf(\mu_c) < 0$; in other words, at $p = p^{t_2}$, a lump sum *transfer* of $\inf(\mu_c)$ to the agent will leave the agent no worse off than at t_1 and potentially better off.

Appendix A.5 has a fuller discussion and includes a proof of (11).

[19] Recall that $p^{t_2} \succeq^*_p p^{t_1}$ makes sense even if p^{t_2} is not observed in the data set; see footnote 7.

3.4. *Measuring departures from rationality*

GARP is frequently violated in empirical applications. The *extent* of such violations for a data set $\mathcal{D}$ is typically measured by the *critical cost efficiency index* (Afriat, 1973). This is the largest $e \in (0,1]$ for which there is a locally-nonsatiated utility function $\widetilde{U}$ such that $\widetilde{U}(x^t) \geq \widetilde{U}(x)$ for all x in the 'shrunken' budget set $B_e^t = \{x \in \mathbb{R}_+^L : \ p^t \cdot x \leq ep^t \cdot x^t\}$. Calculating this index is straightforward via a modified version of GARP (see Afriat, 1973). Rationality is imperfect if $e < 1$ since the consumer behaves as though she ignores bundles x' that satisfy $ep^t \cdot x^t < p^t \cdot x' \leq p^t \cdot x^t$ and, there could be some observation $\hat{t}$ and bundle x' in this range for which $\widetilde{U}(x') > \widetilde{U}(x^{\hat{t}})$.

We can similarly measure the extent to which a data set $\mathcal{D}$ fails to be rationalized by an augmented utility function. We say that $\mathcal{D}$ is ϑ-rationalized by an augmented utility function if there is an augmented utility $U : \mathbb{R}_+^L \times \mathbb{R}_- \to \mathbb{R}$ such that, at each t,

$$U(x^t, -p^t \cdot x^t) \geq U(x, -\vartheta^{-1} p^t \cdot x) \text{ for all } x \in \mathbb{R}_+^L.$$

Note that if $\mathcal{D}$ can be ϑ-rationalized then it can be ϑ'-rationalized for any $\vartheta' < \vartheta$, since U is strictly decreasing in expenditure. A consumer who is ϑ-rational (for $\vartheta < 1$) has limited rationality in the sense that there could be a bundle x' and an observation $\hat{t}$ such that

$$U(x^{\hat{t}}, -p^{\hat{t}} \cdot x') > U(x^{\hat{t}}, -p^{\hat{t}} \cdot x^{\hat{t}}) \geq U(x', -\vartheta^{-1} p^{\hat{t}} \cdot x').$$

In words, the consumer fails to recognize that x' is superior to $x^{\hat{t}}$ at $t = \hat{t}$ because she has inflated (by ϑ^{-1}) the cost of x'. Any data set can be ϑ-rationalized for some $\vartheta \in (0,1]$ and the supremum ϑ^* over these values provides a natural measure of rationality which we shall refer to as the *rationality index*. The next result (proved in Appendix A.4.3) establishes a connection between this rationality index and Afriat's efficiency index.

Proposition 3. *Let $\mathcal{D} = \{(p^t, x^t)\}_{t=1}^T$ be a data set and let $\check{\mathcal{D}} = \{(p^t, \check{x}^t)\}_{t \in T}$, where $\check{x}^t = x^t/(p^t \cdot x^t)$, be its iso-expenditure version. Then ϑ^* is the rationality index for $\mathcal{D}$ if and only if it is the critical cost efficiency index for $\check{\mathcal{D}}$.*

A consequence of this proposition is that the rationality index (like the efficiency index) is easy to compute. Appendix A.4.2 discusses its computation in more general settings.

3.5. *Deflating prices*

For a data set $\mathcal{D} = \{(p^t, x^t)\}_{t=1}^T$ spanning a long period, nominal expenditure may not be an accurate opportunity cost measure due to price changes of both the observed and unobserved goods. This can be accounted for by deflating the prices of the L goods with a general price index. In other words, we can check if $\widetilde{\mathcal{D}} = \{(p^t/k^t, x^t)\}_{t=1}^T$ obeys GAPP, where $k^t \in \mathbb{R}_{++}$ is an index of the general price level. If it does, it implies that there is an augmented utility function U that rationalizes the data after deflation; in other words,

$$x^t \in \underset{x \in \mathbb{R}^L_+}{\operatorname{argmax}}\, U\left(x, -(p^t \cdot x)/k^t\right) \text{ for all } t \in T.$$

This simple adjustment can be precisely justified when the augmented utility is the reduced form of a larger constrained optimization problem. Indeed, suppose that the consumer is maximizing an overall utility $\widetilde{U}(x, y)$ subject to $p \cdot x + q \cdot y \leq M$, where x is the observed bundle and y is a bundle of other goods with prices q. Keeping q and M fixed, let $U(x, -e)$ be the greatest overall utility the consumer can achieve by choosing y optimally, subject to expenditure $M - e$ and conditional on consuming x, that is,

$$U(x, -e) = \max\{\widetilde{U}(x, y) : y \geq 0 \text{ and } q \cdot y \leq M - e\}. \tag{12}$$

At the prices p^t for the observed goods and q for the outside goods, the consumer chooses a bundle (x^t, y^t) to maximize $\widetilde{U}$ subject to $p^t \cdot x + q \cdot y \leq M$. Then $\mathcal{D} = \{(p^t, x^t)\}_{t=1}^T$ will obey GAPP, since x^t maximizes $U(x, -p^t \cdot x)$, with U as defined by (12).

Now suppose that the prices of the other goods and global budget are changing and consider the simplest case where they move proportionately: at t, they are $k^t q$ and $k^t M$ for some scalar $k^t > 0$. In other words, the consumer's nominal wealth is keeping pace with price inflation. Then, at t, the consumer maximizes $\widetilde{U}(x, y)$ subject to (x, y) obeying $p^t \cdot x + k^t q \cdot y \leq k^t M$. Dividing this inequality by k^t, we see that the consumer's choice is identical to the case where the price of the observed goods is p^t / k^t, with constant external prices q and total wealth M respectively. Therefore, $\widetilde{\mathcal{D}}$ obeys GAPP.

In Appendix A.3, we derive a price index (so GAPP holds after deflating p^t) when the relative prices of the outside goods change but this requires stronger assumptions on $\widetilde{U}$.

4. General consumption spaces and nonlinear pricing

We now assume that the consumption space is a set $X \subseteq \mathbb{R}^L_+$ and we define a *price system* as a map $\psi : X \to \mathbb{R}_+$, where $\psi(x)$ is the cost of purchasing $x \in X$. A special case of a price system is $\psi(x) = p \cdot x$ but the more general formulation with ψ allows for quantity discounts, bundle pricing and other pricing features that can be important in certain contexts (such as our empirical application in Section 7.1). A data set $\mathcal{D} = \{(\psi^t, x^t)\}_{t=1}^T$ consists of price systems and the observed consumption bundles; $\mathcal{D}$ is rationalized by an augmented utility function $U : X \times \mathbb{R}_- \to \mathbb{R}$ if

$$x^t \in \underset{x \in X}{\operatorname{argmax}}\, U(x, -\psi^t(x)) \text{ for all } t \in T. \tag{13}$$

The notion of revealed preference over prices can be extended to a revealed preference over price systems. We say that $\psi^{t'}$ is *directly revealed preferred* (*directly revealed strictly* *preferred*) to ψ^t if $\psi^{t'}(x^t) \leq (<) \psi^t(x^t)$; we denote this by $\psi^{t'} \succeq_p (\succ_p) \psi^t$. We denote the transitive closure of $\succeq_p$ by $\succeq_p^*$, that is, $\psi^{t'} \succeq_p^* \psi^t$ if there are $t_1, t_2, \ldots, t_N$ in T such that $\psi^{t'} \succeq_p \psi^{t_1}$, $\psi^{t_1} \succeq_p \psi^{t_2}, \ldots, \psi^{t_{N-1}} \succeq_p \psi^{t_N}$, and $\psi^{t_N} \succeq_p \psi^t$; in this case we say that $\psi^{t'}$ is *revealed preferred* to ψ^t. If anywhere along this sequence, it is possible to replace $\succeq_p$ with $\succ_p$ then we denote that relation by $\psi^{t'} \succ_p^* \psi^t$ 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, then it obeys the following generalization of GAPP to price systems:

there do not exist observations $t, t' \in T$ such that $\psi^{t'} \succeq_p^* \psi^t$ and $\psi^t \succ_p^* \psi^{t'}$.

The next result shows that the converse is also true. The proof (in fact, of a more general result that allows for errors) is in Appendix A.4.1.

Theorem 2. *A data set $\mathcal{D} = \{(\psi^t, x^t)\}_{t=1}^T$ can be rationalized by an augmented utility function if and only if satisfies GAPP.*

Furthermore, suppose that $\mathcal{D}$ satisfies GAPP, X is closed and that, for all $t \in T$, the price systems have the following properties: (i) ψ^t is a continuous function; (ii) for any number M, $\{x \in X : \psi^t(x) \leq M\}$ is a compact set; and (iii) ψ^t is strictly increasing in x_K for some $K \subseteq L$.[20] *Then, for any closed set $Y \subseteq \mathbb{R}_+^L$ containing X, there is a continuous augmented utility function $U : Y \times \mathbb{R}_- \to \mathbb{R}$ that rationalizes $\mathcal{D}$, with $U(x, -e)$ strictly increasing in x_K.*

REMARKS: (1) Note that condition (ii) is a weak assumption requiring that there be no arbitrarily large bundles with a bounded price. (2) By definition, an augmented utility function is strictly decreasing in expenditure, but in certain cases it may be natural to require U to be strictly increasing in x_K for some set K (which can be empty). The theorem says that this is possible, so long as the price systems are also strictly increasing in x_K. (3) Note the theorem establishes the additional properties for domains larger than X. We show in Section 4.1 that this is natural in certain applications.

The literature on mental accounting has emphasized the possibility of actors in the economy manipulating the mental budgets of agents. The following example shows how a nonlinear GAPP test can be used to detect such phenomena.

Example 3. A store initially prices two goods at $p^t = (1, 2)$ and a shopper purchases $x^t = (10, 20)$ from the store. The store introduces a scheme where regular customers receive a voucher of 12 dollars for in-store purchases; prices are changed to $p^{t'} = (2, 3/2)$ and the shopper buys $x^{t'} = (20, 20)$.[21] What is the impact of the voucher?

Since the value of the voucher is small in terms of total income, the shopper could spread this reward widely across all purchases (including purchases from other stores)

[20]This means that if $x', x \in X$, $x' \neq x$, $x_\ell = x'_\ell$ for all $\ell \notin K$, and $x'_\ell \geq x_\ell$ for all $\ell \in K$, then $\psi^t(x') > \psi^t(x)$.

[21]If good 1 is cheap to procure, this scheme is advantageous to the store, since in the first instance, the shopper spends 50 dollars while in the second, she spends 58 (net of the voucher).

and this should result in no (or at least a very small) impact on demand for the store's products. On the other hand, she may have a mental budget for purchases at that store, and the voucher represents an appreciable increase in *that* mental budget by 12 dollars.

A revealed preference analysis supports the latter hypothesis. If we ignore the voucher, the data are not compatible with the maximization of an augmented utility function since $p^t \cdot x^t = 50 = p^{t'} \cdot x^t$ and $p^{t'} \cdot x^{t'} = 70 > 60 = p^t \cdot x^{t'}$, which violates GAPP. On the other hand, at observation t', we could model the shopper as mentally discounting 12 dollars from her expenditure at the shop. In formal terms, the price system at t' is a function $\psi^{t'}(x) = \max\{p^{t'} \cdot x - 12, 0\}$, so $\psi^{t'}(x^{t'}) = 58$. In this case, we have $\psi^{t'} \succ_p^* \psi^t$ (where $\psi^t(x) = p^t \cdot x$), but now $\psi^t \not\succeq_p^* \psi^{t'}$ since $\psi^t(x^{t'}) = 60 > \psi^{t'}(x^{t'}) = 58$. So the data satisfies GAPP, but with a nonlinear pricing system based on the shopper's mental accounting.[22]

4.1. *Discrete consumption spaces*

Below are three instances where Theorem 2 could be applied.

(1) Suppose that the consumption space $X = \mathbb{N}^K \times \mathbb{R}_+^{L-K}$ (where $\mathbb{N}$ is the set of natural numbers) consists of L goods of which the first K can only be consumed in discrete quantities. Theorem 2 is applicable whether or not prices are linear. In the latter case, the price system is $\psi^t(x) = p^t \cdot x$, which is strictly increasing in x. Theorem 2 guarantees that, if GAPP holds, then there is a continuous augmented utility function that is strictly increasing in x and rationalizes $\mathcal{D}$.

(2) Another natural environment is one where the consumer purchases a subset of objects from a set with L items. Then each subset can be represented as an element of $X = \{0,1\}^L$. For $x \in X$, the ℓ^{th} entry x_ℓ equals 1 if and only if the ℓ^{th} object is chosen. If only certain subsets are permissible, then X would be a strict subset of $\{0,1\}^L$. The price system ψ specifies the cost of different bundles in X. Let e_ℓ denote the vector with 1 in the ℓ^{th} entry and zero everywhere else. Then $\psi(e_\ell)$ is the price of purchasing good ℓ alone. The price system is nonlinear if $\psi(x) \neq \sum_{\ell=1}^{L} x_\ell \psi(e_\ell)$ for some $x \in X$.

(3) Empirical demand models of differentiated goods typically model goods in terms of their characteristics (see Nevo (2000)). Suppose that there are L characteristics and I goods. Let $Y_\ell \subseteq \mathbb{R}_+$ be the set of values that characteristic ℓ can take. Then, the characteristics space is $Y = \times_{\ell=1}^{L} Y_\ell$.[23] Each good i has characteristics $x^i \in Y$. Assuming (as is common) that a consumer purchases only one good, the consumption space is $X = \{x^i\}_{i=1}^I$ and a price system $\psi : X \to \mathbb{R}_{++}$ is just a list of prices for the different goods.

[22] Notice (in connection with our discussion of mental accounting in Section 2.4) that the total mental budget of the shopper remains unknown, though the researcher observes an event that has altered that budget.

[23] If characteristic 1 naturally takes on continuous values (such as calories) then we let $Y_\ell = \mathbb{R}_+$. Characteristic 2 could be the brand. Suppose there are two brands, then $Y_2 = \{1,2\}$, and so on.

Here, it is natural to model the consumer with an augmented utility function defined on characteristics and expenditures $Y \times \mathbb{R}_-$, even though the products available are only those in X. Furthermore, among the characteristics, there could be those where higher values are unambiguously better, in which case it is natural to require that utility is strictly increasing in those characteristics. Theorem 2 allows for these considerations. If $\mathcal{D}$ obeys GAPP then it can be rationalized by a continuous augmented utility function $U : Y \times \mathbb{R}_- \to \mathbb{R}$. Additionally, for a set of characteristics $K \subseteq L$, one could guarantee that $U(y, -e)$ is strictly increasing in y_K so long as $\psi^t(x)$ is strictly increasing in x_K, for all t.

In models of differentiated goods, it is also common to allow for the introduction of new goods and for changes to a product's characteristics.[24] Changes to a product's characteristics could potentially lead to a change in the product's utility which, unless taken into account by the test, could lead to a spurious rejection of augmented utility-maximization. When changes to product characteristics are observable, they can be formally captured by allowing the set of alternatives to depend on t; in Section A.4.4, we explain how it is possible to modify the GAPP test in Theorem 2 to account for changes of this type.

4.2. *Characteristics models with continuous consumption spaces*

Now consider the characteristics space $Y = \mathbb{R}^L_+$, with each product i represented by a vector of characteristics $x^i \in Y$. Goods can be bought in bundles, so the consumption space is the convex cone X generated by $\{x^i\}_{i=1}^I$.[25] We assume that the vectors $\{x^i\}_{i=1}^I$ are linearly independent; this guarantees that for each $\hat{x} \in X$, there is a *unique* bundle of goods, $\hat{\alpha} = (\hat{\alpha}_i)_{i=1}^I \in \mathbb{R}^I_+$ such that $\sum_{i=1}^I \hat{\alpha}_i x^i = \hat{x}$. We denote $\hat{\alpha}$ by $\alpha(\hat{x})$. Let $p^t \in \mathbb{R}^I_{++}$ be the prices of the I goods at observation t and so a bundle $x \in X$ costs $\psi^t(x) = p^t \cdot \alpha(x)$.

The researcher observes prices p^t and bundles $\alpha^t \in \mathbb{R}^I_+$. We assume $\{x^i\}_{i=1}^I$ is known so the consumption in characteristics space, $x^t = \sum_{i=1}^I \alpha_i^t x^i$, and the price system ψ^t can be imputed. Theorem 2 guarantees that if $\mathcal{D} = \{(\psi^t, x^t)\}_{t \in T}$ satisfies GAPP then it can be rationalized by a continuous augmented utility $U : \mathbb{R}^L_+ \times \mathbb{R}_- \to \mathbb{R}$. If $\psi^t(x)$ is strictly increasing in $x \in X$ for each t, we can also ensure that $U(y, -e)$ is increasing in y.

5. The Random Augmented Utility Model

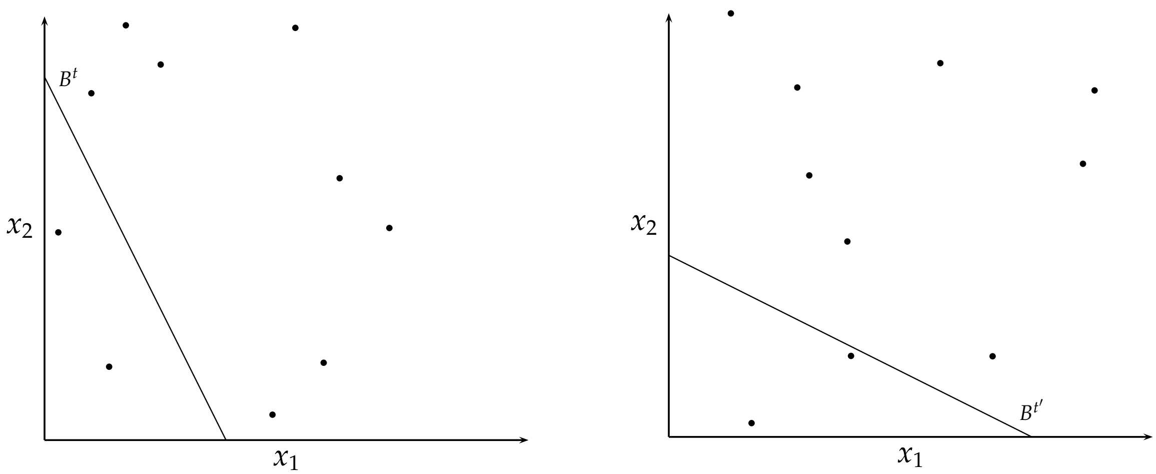
In this section, we develop the random version of the expenditure-augmented utility model, beginning with an example in which computations can be done in closed form.

[24] These changes could be substantive (for example, a change to a breakfast cereal formula) or it could be a change in advertising expenditure that serves as a proxy for a change in a product's public profile.

[25] For a GARP-based test of a model of this type, see Blow, Browning, and Crawford (2008).

5.1. *An Illustrative Example*

Example 4. Suppose we have repeated cross-sectional data consisting of the demand of a population of ten consumers at two price vectors. This is illustrated in Figure 4, where the collection of points in the left and right panels indicate the demand bundles at $p^t = (2,1)$ and $p^{t'} = (1,2)$ respectively. The lines in Figure 4 merely indicate relative prices.



(a) Choices under prices p^t (b) Choices under prices $p^{t'}$

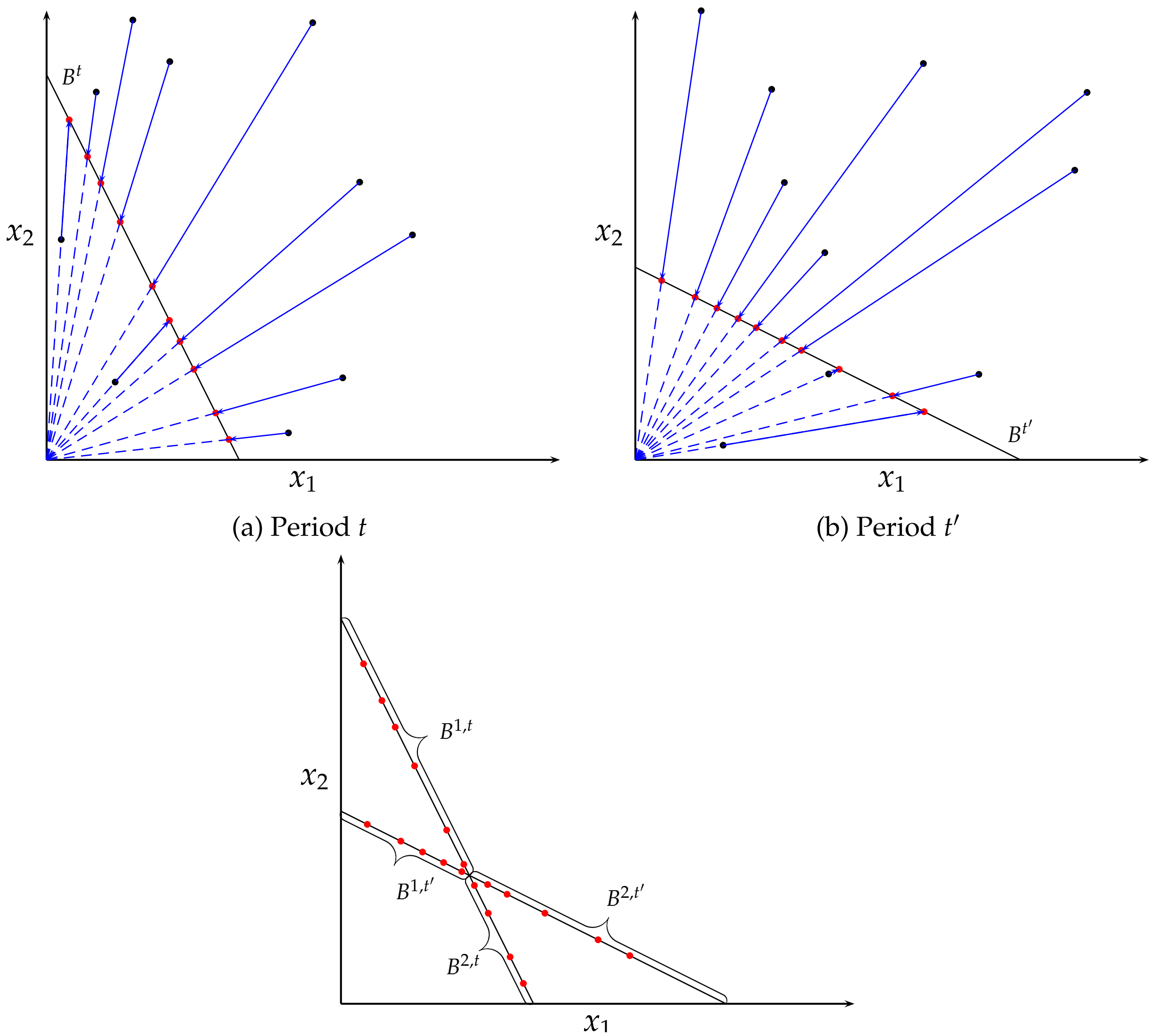
FIGURE 4. The Data Set

As this is a repeated cross-section, we cannot match consumption bundles across the two panels by consumer identity. The question is whether this data set can be *rationalized*, by which we mean the following:

Can the choices at t and t' be matched to form ten distinct pairs such that each pair is rationalized by an augmented utility function (or, equivalently, satisfies GAPP)?

The interpretation is that each pair of choices corresponds to the demand of a single consumer and so rationalization requires the existence of a (time-invariant) distribution of consumer types or, equivalently, a *random augmented utility*. For more conventional models of utility maximization, this question was analyzed by McFadden and Richter (1991, for discrete choice) and McFadden (2005, for linear budgets). While these settings appear quite different from ours, the tight link established in Proposition 3 allows us to build on these results.

Specifically, a pair $\mathcal{D} = \{(p^t, x), (p^{t'}, x')\}$ created by choosing bundle x from observation t and x' from observation t' obeys GAPP if and only if its *iso-expenditure* analog, $\breve{\mathcal{D}} = \{(p^t, \breve{x}), (p^{t'}, \breve{x}')\}$ as defined in Proposition 3, obeys GARP. This is visualized in Figures 5a and 5b: replacing data points by iso-expenditure analogs is equivalent to projecting them along origin rays onto the budget corresponding to unit expenditure. Figure 5c superimposes the scaled bundles from both cross-sections. Rationalizability by a random augmented utility is therefore equivalent to asking whether these rescaled observations can be rationalized by a standard random utility model (which in this simple case means sorting them into pairs of observations that each obey GARP).



(a) Period t

(b) Period t'

(c) Patches generated by budget intersections

FIGURE 5. Observed and Rescaled Choices

In this example, $\check{\mathcal{D}}$ satisfies GARP if, and only if, it is *not* the case that $\check{x} \in B^{2,t}$ and $\check{x}' \in B^{1,t'}$, where $(B^{2,t}, B^{1,t'})$ are indicated in Figure 5c. Instead $\check{\mathcal{D}}$ must be one of three individually rationalizable choice types: either $(\check{x}, \check{x}') \in B^{1,t} \times B^{2,t'}$ (no revealed preference), or $(\check{x}, \check{x}') \in B^{1,t} \times B^{1,t'}$ ($\check{x}$ revealed preferred to $\check{x}'$), or $(\check{x}, \check{x}') \in B^{2,t} \times B^{2,t'}$ ($\check{x}'$ revealed preferred to $\check{x}$).

Denote the population share of these types by (ν_1, ν_2, ν_3). They must generate the observed proportion of choices on the segments $(B^{1,t}, B^{2,t}, B^{1,t'}, B^{2,t'})$. Figure 6a visualizes how observed choice probabilities relate to (ν_1, ν_2, ν_3); Figure 6b gives the corresponding sample proportions

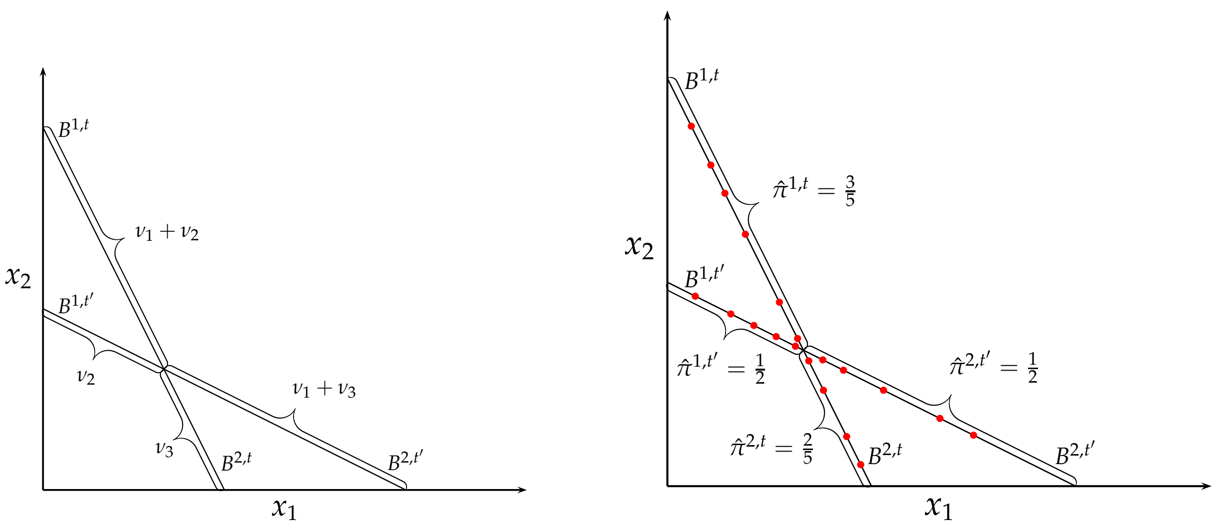
$$\hat{\pi} = \left(\hat{\pi}^{1,t}, \hat{\pi}^{1,t'}, \hat{\pi}^{2,t}, \hat{\pi}^{2,t'}\right)' = \left(\frac{3}{5}, \frac{2}{5}, \frac{1}{2}, \frac{1}{2}\right)'. \tag{14}$$

The empirical choice frequencies are rationalizable by a random augmented utility if

$$\nu_1 + \nu_2 = \hat{\pi}^{1,t},\ \nu_1 + \nu_3 = \hat{\pi}^{2,t'},\ \nu_2 = \hat{\pi}^{1,t'},\ \nu_3 = \hat{\pi}^{2,t} \tag{15}$$

can be solved for nonnegative (ν_1, ν_2, ν_3).

This is indeed the case in the example, where the equations are uniquely solved by $(\nu_1, \nu_2, \nu_3)' = \left(\frac{1}{10}, \frac{1}{2}, \frac{2}{5}\right)'$.[26] To further confirm this, we could pair up the bundles on the two budget lines in order of the consumption of x_2. Then it is easily seen that each pair satisfies GARP (and hence the corresponding un-scaled pairs satisfy GAPP).



(a) Resulting Distribution of Choices (b) The Scaled Empirical Choice Probabilities $\hat{\pi}$

FIGURE 6. Choice Distribution and Empirical Frequency

5.2. *Rationalization by Random Augmented Utility*

Consider now a *repeated cross-sectional data set*, $\mathscr{D} := \{(p^t, \mathring{\pi}^t)\}_{t=1}^T$, where each observation consists of a price p^t and a probability measure $\mathring{\pi}^t$ on $\mathbb{R}^L_+$ representing the demand distribution in the population at that price. We now provide a general definition of rationalizability for such a data set.

Definition 5.1. The repeated cross-sectional data set $\mathscr{D} = \{(p^t, \mathring{\pi}^t)\}_{t=1}^T$ is *rationalized* by the *random augmented utility model* (RAUM) if there exists a probability space $(\Omega, \mathcal{F}, \mu)$

[26] In general, the solution is not unique. In this simple example, it is straightforward to check that the data can be rationalized by a random augmented utility if, and only if, $\hat{\pi}^{1,t} + \hat{\pi}^{2,t'} \geq 1$.

and a random variable $\chi : \Omega \to (\mathbb{R}^L_+)^T$ such that, almost surely, $\{(p^t, \chi^t(\omega))\}_{t\in T}$ can be rationalized by an augmented utility function (equivalently, obeys GAPP) and

$$\mathring{\pi}^t(Y) = \mu(\{\omega \in \Omega : \chi^t(\omega) \in Y\}) \text{ for any measurable } Y \subseteq \mathbb{R}^L_+. \tag{16}$$

In this definition, one could interpret Ω as the population of consumers and $\chi^t(\omega)$ as the demand of consumer type ω at price p^t. All consumer types in the support of μ must be consistent with the augmented utility model and, for all t, the observed distribution of demand $\mathring{\pi}^t$ must coincide with that induced by the distribution μ over consumer types. Alternatively, individuals' augmented utility functions might change over time but in such a way that the population distribution is stationary.

In Example 4, the data set contains two cross-sectional distributions, both of which are discrete with 10 mass points. A RAUM-rationalization involves matching observations in t with those in t', so that each pair obeys GAPP. In the general case with T cross-sections, the function χ solves a T-fold matching problem, where each group $\{\chi^t(\omega)\}_{t\in T}$ (along with the associated prices) satisfies GAPP and agrees with the observations (that is, (16) is satisfied).[27]

Theorem 3 below characterizes the rationalizability of $\mathscr{D} = \{(p^t, \mathring{\pi}^t)\}_{t=1}^T$. It is proved in the online appendix, but we now explain it heuristically. Let $B^t := \{x \in \mathbb{R}^L_+ \,:\, p^t \cdot x = 1\}$ be the budget plane at prices p^t and expenditure 1 and let $\breve{\pi}^t$ be the distribution that obtains after projecting $\mathring{\pi}^t$ onto B^t.[28] We refer to $\breve{\mathscr{D}} = \{(p^t, \breve{\pi}^t)\}_{t=1}^T$ as the iso-expenditure analog of $\mathscr{D}$. We say that $\breve{\mathscr{D}}$ is rationalized by the random utility model (RUM) if there is a probability space $(\Omega, \mathcal{F}, \mu)$ and a random variable $\chi : \Omega \to (\mathbb{R}^L_+)^T$ such that, almost surely, $\{(p^t, \chi^t(\omega))\}_{t\in T}$ obeys GARP and $\breve{\pi}^t(Y) = \mu(\{\omega \in \Omega : \chi^t(\omega) \in Y\})$ for any measurable $Y \subseteq \mathbb{R}^L_+$. Crucially, reasoning along the lines of Proposition 3 establishes that $\mathscr{D}$ can be RAUM-rationalized if, and only if, $\breve{\mathscr{D}} = \{(p^t, \breve{\pi}^t)\}_{t=1}^T$ can be RUM-rationalized. Finally, to check the latter, we simply adopt the procedure laid out in McFadden (2005) and KS for testing RUM on iso-expenditure data sets, which we now explain.

For ease of exposition, we impose the following assumption.

Assumption 1. For all $t, t' \in T$ with $B^t \neq B^{t'}$,

$$\mathring{\pi}^t \left(\left\{ x \in \mathbb{R}^L_+ : \frac{x}{p^t \cdot x} \in B^t \text{ and } \frac{x}{p^{t'} \cdot x} \in B^{t'} \right\} \right) = 0.$$

This assumption excludes (with probability 1) choices on the intersection of budget planes. It is not required for any of our results but simplifies the exposition because it

[27]It is straightforward to check that, with two observations, finding a rationalization is equivalent to finding a zero-cost solution to the transportation problem (see Galichon and Henry (2011)) where the cost of a pair of bundles is 0 if it obeys GAPP and 1 otherwise.

[28]Formally, given a measurable set C in B^t, $\breve{\pi}^t(C) = \mathring{\pi}^T(\overline{C})$, where $\overline{C}$ is the cone generated by C.

forces revealed preferences to be strict.[29] It is always satisfied if $\mathring{\pi}^t$ is absolutely continuous with respect to Lebesgue measure and is unlikely to be violated in any application with a continuous consumption space and linear prices.

Next, for any budget B^t, let $\{B^{1,t}, \ldots, B^{I_t,t}\}$ denote the collection of subsets such that each subset has as its boundaries the intersection of B^t with other budget sets and/or the boundary planes of the positive orthant. These are the higher-dimensional and multi-period analogs to the line segments in Figure 5c. Formally, for all $t \in T$ and $i_t \neq i'_t$, each set in $\{B^{1,t}, \ldots, B^{I_t,t}\}$ is closed and convex and satisfies the following conditions:

(i) $\cup_{1 \leq i_t \leq I_t} B^{i_t,t} = B^t$,
(ii) $\text{int}(B^{i_t,t}) \cap B^{t'} = \phi$ for all $t' \neq t$ that satisfy $B^t \neq B^{t'}$ (where $\text{int}(B^{i_t,t})$ denotes the relative interior of $B^{i_t,t}$),
(iii) $B^{i_t,t} \cap B^{i'_t,t} \neq \phi$ implies that $B^{i_t,t} \cap B^{i'_t,t} \subset B^{t'}$ for some $t' \neq t$ that satisfies $B^t \neq B^{t'}$.

We will henceforth call these sets *patches*. For each patch $B^{i_t,t}$, let

$$\pi^{i_t,t} := \mathring{\pi}^t \left(\left\{ x \in \mathbb{R}^L_+ : \frac{x}{p^t \cdot x} \in B^{i_t,t} \right\} \right) \tag{17}$$

be the probability that a period-t choice after scaling lies on patch $B^{i_t,t}$. Denote by π^t the vector $(\pi^{i_t,t})_{i_t=1}^{I_t}$ and by π the column vector $(\pi^1, \pi^2, \ldots, \pi^T)'$ of *observed patch probabilities*. (Assumption 1 here causes simplification because it guarantees that $\sum_{i_t=1}^{I_t} \pi^{i_t,t} = 1$.) These patch probabilities are relevant because, as we shall explain, rationalizability depends only on the patch probabilities and not on the distribution within each patch.

To any deterministic, iso-expenditure data set $\breve{\mathcal{D}} = \{(p^t, \breve{x}^t)\}$ we can associate a vector $a = \left(a^{1,1}, \ldots, a^{I_T,T}\right)$, where $a^{i_t,t} = 1$ if $\breve{x}^t \in B^{i_t,t}$ and 0 otherwise. The first crucial observation is that any two deterministic iso-expenditure data sets that are represented by the same vector a would either *both* obey or fail GARP (because their revealed preference relations have the same structure).[30] Thus, even though there are infinitely many deterministic iso-expenditure data sets $\breve{\mathcal{D}}$ that obey GARP, they belong to a *finite* set of equivalence classes, each of which is represented by a vector a. We gather these vectors into the set $\mathcal{A}$. The second crucial observation is that $\breve{\mathscr{D}} = \{(p^t, \breve{\pi}^t)\}_{t=1}^T$ can be RUM-rationalized if, and only if, π is in the the convex hull of vectors in $\mathcal{A}$ (so π is generated by a distribution over GARP-consistent types).

[29]If we allow for mass at budget intersections, then we would have to include them in our definition of patches. This is notationally cumbersome but once included our arguments (and Theorem 3) remain correct.
[30]This observation is made in Section 3.1 of McFadden (2005).

To state this a bit more formally, collect all distinct GARP-consistent vectors $a \in \mathcal{A}$ into the columns of a matrix A. In Example 4, this matrix equals

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \tag{18}$$

with column j representing type j ($j = 1, 2, 3$).[31] Then the cross-sectional data set $\mathscr{D}$ can be RAUM-rationalized if, and only if, its iso-expenditure analog $\breve{\mathscr{D}}$ can be RUM-rationalized and the latter holds if, and only if, there is $\nu_a \geq 0$ (corresponding to weight of type $a \in \mathcal{A}$), such that $\sum_{a \in \mathcal{A}} \nu_a = 1$ and

$$\pi^{i_t,t} = \sum_{a \in \mathcal{A}} \nu_a \, a^{i_t,t} \tag{19}$$

for all $B^{i_t,t}$. More succinctly, $A\nu = \pi$, where ν is the column vector $(\nu_a)_{a \in \mathcal{A}}$.[32] The next theorem summarizes our discussion.

Theorem 3. *Let $\mathscr{D} = \{(p^t, \mathring{\pi}^t)\}_{t=1}^{T}$ be a repeated cross-sectional data set obeying Assumption 1. Then $\mathscr{D}$ can be RAUM-rationalized if and only if there exists a $\nu \in \Delta^{|\mathcal{A}|-1}$ such that $A\nu = \pi$.*

It is worth reiterating that even though the RAUM-test given by this theorem requires the (straightforward) calculation of π from an iso-expenditure analog of $\mathscr{D}$, the data set $\mathscr{D}$ need not be an iso-expenditure data set. In contrast, the RUM-test set out in McFadden (2005) requires $\mathscr{D}$ *itself* to be an iso-expenditure data set. Of course, actual data generated by heterogeneous consumers will typically not be iso-expenditure, and indeed the Family Expenditure Survey data used by KS in their implementation of the RUM-test are not. For this reason, their empirical analysis starts by *estimating* an iso-expenditure data set (corresponding to the median expenditure), before implementing the RUM-test on this estimated data set. Obviously, this requires additional econometric work and, therefore, introduces both more assumptions and (albeit asymptotically negligible) noise.

5.3. *Welfare Comparisons*

Since the test for rationalizability involves finding a distribution ν over different types, it is possible to use this distribution for welfare analysis. To be specific, suppose that a government is contemplating a change in sales tax that could lead to prices changing from its current value $p^{t'}$ to $\hat{p}$. Relevant to the government's re-election prospects is the

[31] Each column has four entries because there are four patches in total. For example, the first column represents type 1 data sets $\breve{\mathcal{D}} = \{(p^t, \breve{x}), (p^{t'}, \breve{x}')\}$, where $(\breve{x}, \breve{x}') \in B^{1,t} \times B^{2,t'}$ (see Figure 5c).
[32] In Example 4, A and π are given by (18) and (14) respectively.

proportion of consumers who will be better off as a result of this price change.[33] Our methods allow us to obtain information on this proportion.

So consider a data set $\mathscr{D}$ that contains among its observations the prevailing prices $p^{t'}$ and the demand distribution $\mathring{\pi}^{t'}$. To determine the welfare effect of a price change from $p^{t'}$ to $\hat{p}$, let $\mathbb{1}_{\hat{p}\succeq_p^* p^{t'}}$ denote the row vector with its length equal to the number of rational types ($|\mathcal{A}|$), such that the j^{th} element is 1 if $\hat{p} \succeq_p^* p^{t'}$ for the rational type corresponding to column j of A and 0 otherwise.[34] In words, $\mathbb{1}_{\hat{p}\succeq_p^* p^{t'}}$ enumerates the set of rational types for which $\hat{p}$ is revealed preferred to $p^{t'}$. If $\mathscr{D}$ is rationalizable, Theorem 3 guarantees that

$$\underline{\mathcal{N}}_{\hat{p}\succeq_p^* p^{t'}} := \min_\nu \ \mathbb{1}_{\hat{p}\succeq_p^* p^{t'}}\, \nu, \quad \text{subject to } \ A\nu = \pi, \tag{20}$$

is the lower bound on the proportion of consumers who are revealed better off at prices $\hat{p}$ compared to $p^{t'}$, while the upper bound is

$$\overline{\mathcal{N}}_{\hat{p}\succeq_p^* p^{t'}} := \max_\nu \ \mathbb{1}_{\hat{p}\succeq_p^* p^{t'}}\, \nu, \quad \text{subject to } \ A\nu = \pi. \tag{21}$$

Since (20) and (21) are both linear programs (which have solutions if, and only if, $\mathscr{D}$ is rationalizable), they are easy to implement in practice. Suppose that the solutions are $\underline{\nu}$ and $\overline{\nu}$ respectively; then for any $\beta \in [0,1]$, $\beta\underline{\nu} + (1-\beta)\overline{\nu}$ is also a solution to $A\nu = \pi$ and, in this case, the proportion of consumers who are revealed better off at $\hat{p}$ compared to $p^{t'}$ is exactly $\beta\,\underline{\mathcal{N}}_{\hat{p}\succeq_p^* p^{t'}} + (1-\beta)\,\overline{\mathcal{N}}_{\hat{p}\succeq_p^* p^{t'}}$. In other words, the proportion of consumers who are revealed better off can take any value in the interval $[\underline{\mathcal{N}}_{\hat{p}\succeq_p^* p^{t'}}, \overline{\mathcal{N}}_{\hat{p}\succeq_p^* p^{t'}}]$.

Proposition 2 tells us that the revealed preference relations are tight, in the sense that if, for some consumer, $\hat{p}$ is not revealed preferred to $p^{t'}$ then there exists an augmented utility function which rationalizes her consumption choices and for which she strictly prefers $p^{t'}$ to $\hat{p}$. Given this, we know that, amongst all rationalizations of $\mathscr{D}$, $\underline{\mathcal{N}}_{\hat{p}\succeq_p^* p^{t'}}$ is also the infimum on the proportion of consumers who are better off at $\hat{p}$ compared to $p^{t'}$.

The following proposition summarizes these observations.

Proposition 4. *Let $\mathscr{D} = \{(p^t, \mathring{\pi}^t)\}_{t=1}^T$ be a repeated cross-sectional data set that satisfies Assumption 1 and is rationalized by the RAUM. Then, for every $\eta \in [\underline{\mathcal{N}}_{\hat{p}\succeq_p^* p^{t'}}, \overline{\mathcal{N}}_{\hat{p}\succeq_p^* p^{t'}}]$, there is a rationalization of $\mathscr{D}$ for which η is the proportion of consumers who are revealed better off at $\hat{p}$ compared to $p^{t'}$. Furthermore, $\underline{\mathcal{N}}_{\hat{p}\succeq_p^* p^{t'}}$ is the infimum of the proportion of consumers who are better off at $\hat{p}$ compared to $p^{t'}$, among all the rationalizations of $\mathscr{D}$.*

It is helpful to apply Proposition 4 to Example 4. There, the solution to $A\nu = \pi$ is unique. Of the three types discussed in Section 5.1, only the second one reveals $p^t \succeq_p^* p^{t'}$, so the proportion of consumers revealed better off at p^t compared to $p^{t'}$ equals ν_2.

[33] We would like to thank an anonymous referee for suggesting this motivation.

[34] Even though $\hat{p}$ is not among the observed prices, one could still define $\hat{p} \succeq_p^* p^{t'}$; see footnote 7.

Formally, we have $\mathbb{1}_{p^t \succeq_p^* p^{t'}} = (0,1,0)$, $\mathbb{1}_{p^t \succeq_p^* p^{t'}} \cdot \nu = 1/2$, and $\underline{\mathcal{N}}_{p^t \succeq_p^* p^{t'}} = \overline{\mathcal{N}}_{p^t \succeq_p^* p^{t'}} = 1/2$. By similar reasoning, we have $\underline{\mathcal{N}}_{p^{t'} \succeq_p^* p^t} = \overline{\mathcal{N}}_{p^{t'} \succeq_p^* p^t} = 2/5$. The point identification of these quantities is due to the uniqueness of ν, which is specific to that simple example.

6. Statistical Test of RAUM and Inference for Counterfactuals

This section outlines our econometric methodologies. First, Section 6.1 provides a statistical test of the RAUM (presented in Section 5). Second, and more importantly, Section 6.2 develops a new method for obtaining asymptotically uniformly valid confidence intervals for counterfactual objects. This result applies to a general class of random utility models, including the RAUM. It can be used for statistical analyses of welfare comparisons and we use it for that purpose in our empirical study in Section 7.2.

6.1. *Testing the Random Augmented Utility Model*

Recall from Theorem 3 that, given a set of prices and corresponding demand distributions $\mathscr{D} = \{(p^t, \mathring{\pi}^t)\}_{t=1}^T$ and an implied vector π of choice probabilities on rescaled and discretized budgets, a test of the random augmented utility model is a test of

$$H_0 : \exists \nu \in \Delta^{|\mathcal{A}|-1} \text{ such that } A\nu = \pi \iff \min_{\nu \in \mathbb{R}_+^{|\mathcal{A}|}} [\pi - A\nu]' \Omega [\pi - A\nu] = 0, \tag{22}$$

where Ω is a positive definite matrix. The equivalence was noted and exploited in KS.[35]

In practice, we estimate π by its sample analog $\hat{\pi} = (\hat{\pi}^1, \ldots, \hat{\pi}^T)$ obtained by rescaling the empirical distribution of choices $\{x_{n_t}^t\}_{n_t=1}^{N_t}$ where N_t is the number of observed choices in the data in period t. This gives rise to test statistic

$$J_N := N \min_{\nu \in \mathbb{R}_+^{|\mathcal{A}|}} [\hat{\pi} - A\nu]' \Omega [\hat{\pi} - A\nu], \tag{23}$$

where $N = \sum_{t=1}^T N_t$ denotes the total number of observations. We use the modified bootstrap procedure of KS to compute critical values for this test.

6.2. *Inference for Counterfactuals in a General Class of Random Utility Models*

A counterfactual quantity in a random utility model can be generally regarded as a function of the underlying distribution ν of individual preferences. This section focuses on the case where this mapping is linear, so that we are concerned with statistical inference for $\theta = \rho \cdot \nu$, where $\rho \in \mathbb{R}^{|\mathcal{A}|}$ is a known vector which varies with the counterfactual of interest. Our analysis of welfare comparisons in Section 5.3 falls into this framework,

[35]The strategy to configure H_0 as a quadratic program also appears in De Paula, Richards-Shubik, and Tamer (2018), albeit for a different program and in a different context.

by letting θ be the proportion of consumers who are revealed better off at prices $\hat{p}$ compared to $p^{t'}$, with $\rho = \mathbb{1}_{\hat{p} \succeq^*_p p^{t'}}$. It is worth emphasizing that the methodology developed in this section has broad applicability: it can be used to study other random utility models (such as the model in Kitamura and Stoye (2019)) and to investigate other objects of interest in random utility models; for example, Lazzati, Quah, and Shirai (2018) applies our technique to estimate the proportion of non-strategic players in a game.

Note that θ is partially identified as follows:

$$\theta \in \Theta_I \quad \text{where} \quad \Theta_I := \{\rho \cdot \nu | \nu \geq 0, A\nu = \pi\}.$$

Our confidence interval inverts a test of

$$\pi \in \mathcal{S}(\theta) \quad \text{where} \quad \mathcal{S}(\theta) := \left\{ A\nu \,|\, \rho \cdot \nu = \theta,\ \nu \in \Delta^{|\mathcal{A}|-1} \right\} \tag{24}$$

or equivalently,

$$\min_{\nu \in \Delta^{|\mathcal{A}|-1}, \theta = \rho \cdot \nu} [\pi - A\nu]'\Omega[\pi - A\nu] = 0.$$

The test statistic is a scaled sample analog

$$J_N(\theta) = N \min_{\nu \in \Delta^{|\mathcal{A}|-1}, \theta = \rho \cdot \nu} [\hat{\pi} - A\nu]'\Omega[\hat{\pi} - A\nu] = N \min_{\eta \in \mathcal{S}(\theta)} [\hat{\pi} - \eta]'\Omega[\hat{\pi} - \eta], \tag{25}$$

where the second equality follows from (24). The naive bootstrap fails to deliver valid critical values for (25) as its asymptotic distribution changes discontinuously, depending on the location of π relative to the polytope $\mathcal{S}(\theta)$. A simple application of the modified bootstrap algorithm in KS does not work, as their method relies on, among other things, the polytope $\{A\nu : \nu \geq 0\}$ being a cone. This is not necessarily the case for counterfactual analysis, and we need to deal with $\mathcal{S}(\theta)$ without relying on conical properties.

That said, as in KS, we do gain an insight from Weyl-Minkowski duality. In Appendix A.7, we show that there exist nonstochastic matrices B, $\tilde{B}$ and a nonstochastic vector-valued function $d(\theta)$ such that $\pi \in \mathcal{S}(\theta)$ if, and only if,

$$B\pi \leq 0,\ \tilde{B}\pi = d(\theta) \text{ and } \mathbf{1} \cdot \pi = 1, \tag{26}$$

where $\mathbf{1}$ is the I-vector of ones where $I = \sum_{t=1}^{T} I_t$ is the total number of patches. Thus, in principle this is a linear (in)equality testing problem. There is a rich literature on such problems. However, we cannot directly invoke that literature because we cannot compute $(B, \tilde{B})$ in practice for a problem with a relevant scale.

While we therefore need to work with representation (24), representation (26) is useful. It illustrates that the inference problem is non-standard; in particular, the limiting distribution of the test statistic depends on how close to binding each of the constraints encoded in $(B, \tilde{B}, d(\theta))$ is. From analogy to the moment inequalities literature, it also pretty much implies that the constraints' slackness cannot be pre-estimated with sufficient accuracy; the reason being that it enters the test's asymptotic representation scaled by $\sqrt{N}$. However, we also know that certain existing procedures which shrink the estimated slack of all inequalities to zero before computing the distribution of J_N will work. Our proposal is inspired by these but must implement the idea with the computationally feasible representation (24) instead of (26), which is only theoretically available. This means that we cannot calculate the empirical slack, which is explicit in (the empirical version of) representation (26) but not in (24), which is why a new method is called for.

Intuitively, we contract (or "tighten") the polytope $\mathcal{S}(\theta)$ toward a point in its relative interior, thereby effectively (but non-obviously) reducing the empirical slack in any inequality constraint. This forces all the constraints with small slacks to be binding after "tightening". Note that, unlike in KS, we face substantial added complications because (i) we need to deal with a non-conical $\mathcal{S}(\theta)$, and (ii) the appropriate way to tighten the polytope $\mathcal{S}(\theta)$ varies with the value of θ through the dependence of $\mathcal{S}(\theta)$ on θ. This leads to a *restriction-dependent tightening* approach which we now describe in broad strokes.

Choose a sequence τ_N such that $\tau_N \downarrow 0$ and $\sqrt{N}\tau_N \uparrow \infty$ (we make a specific proposal in the appendix) and define

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

where $\mathcal{V}_{\tau_N}(\theta)$ is obtained by appropriately constricting $\Delta^{|\mathcal{A}|-1}$; in particular, some components of ν are forced to be boundedly above 0. Note that $\mathcal{S}_{\tau_N}(\theta)$ depends on θ through the equation $\rho \cdot \nu = \theta$ but also because, as the notation suggests, the construction of $\mathcal{V}_{\tau_N}(\theta)$ will change with θ, a key feature of our algorithm. The definition of $\mathcal{V}_{\tau_N}(\theta)$ for general ρ is rather involved and thus deferred to Appendix A.7, but it considerably simplifies for binary ρ as in our application.

The set $\mathcal{S}_{\tau_N}(\theta)$ replaces $\mathcal{S}(\theta)$ in the bootstrap population. The precise algorithm proceeds as follows. For each $\theta \in \Theta$:

(i) Compute the τ_N-*tightened restricted* estimator of the empirical choice distribution

$$\hat{\eta}_{\tau_N} := \underset{\eta \in \mathcal{S}_{\tau_N}(\theta)}{\operatorname{argmin}} N[\hat{\pi} - \eta]'\Omega[\hat{\pi} - \eta].$$

(ii) Define the τ_N-*tightened recentered* bootstrap estimators

$$\hat{\pi}^{*(r)}_{\tau_N} := \hat{\pi}^{*(r)} - \hat{\pi} + \hat{\eta}_{\tau_N}, \quad r = 1, ..., R,$$

where $\hat{\pi}^{*(r)}$ is a bootstrap analog of $\hat{\pi}$ and R is the number of bootstrap samples. For instance, in our application, $\hat{\pi}^{*(r)}$ is generated by the simple nonparametric bootstrap of choice frequencies.

(iii) For each $r = 1, ..., R$, compute

$$J^{*(r)}_{N,\tau_N}(\theta) = \min_{\eta \in \mathcal{S}_{\tau_N}(\theta)} N[\hat{\pi}^{*(r)}_{\tau_N} - \eta]'\Omega[\hat{\pi}^{*(r)}_{\tau_N} - \eta].$$

(iv) Use the empirical distribution of $J^{*(r)}_{N,\tau_N}(\theta)$ to obtain the critical value for $J_N(\theta)$.

A confidence interval for θ collects values of θ that are not rejected.

Theorem 4 below (proved in Appendix A.7) establishes asymptotic validity of the above procedure. Let

$$\mathcal{F} := \{(\theta, \pi) \mid \theta \in \Theta, \pi \in \mathcal{S}(\theta) \cup \mathcal{P}\}$$

where $\mathcal{P}$ denote the set of all π that satisfy Condition 1 in Appendix A.7.

Theorem 4. *Choose τ_N so that $\tau_N \downarrow 0$ and $\sqrt{N}\tau_N \uparrow \infty$. Also, let Ω be diagonal. Then under Assumptions 2 and 3 stated in Appendix A.7,*

$$\liminf_{N\to\infty} \inf_{(\theta,\pi)\in\mathcal{F}} \Pr\{J_N(\theta) \le \hat{c}_{1-\alpha}\} = 1 - \alpha,$$

where $0 \le \alpha \le \frac{1}{2}$ and $\hat{c}_{1-\alpha}$ is the $1-\alpha$ quantile of J^*_{N,τ_N}.

Note that we have assumed the prices are exogenous throughout our analysis in this section. If the exogeneity condition holds *conditional* on some observable covariates, it is straightforward to incorporate it by replacing the sample analogue $\hat{\pi}$ of π with an appropriate conditional choice probability estimator. Or, in situations where the control function approach is applicable, we can deal with endogeneity along the line of analysis in KS (see their Theorem 5.2). On the other hand, a satisfactory treatment of, for example, price endogeneity caused by unobserved characteristics calls for further extension of our approach. If an appropriate instrumental variable is available, then it might be possible to generalize the "$\mathcal{V}$-representations" of identified sets to accommodate it, though we leave such analysis to future research.

7. Empirical Applications

We now present two separate applications meant to show how both the deterministic and random versions of our model can be tested and employed for welfare analysis.

7.1. *Augmented utility model: testing and welfare analysis on Progresa data*

We apply the deterministic model to the Progresa-Oportunidades data set, a workhorse of the treatment evaluation literature. Progresa was a conditional cash transfer program aimed at poor communities in Mexico. The program was remarkable in that it was rolled out in random order so the causal effect of the cash transfers could be studied. For brevity, we do not describe the program in detail; information on the program is widely available including in the paper we discuss next.

Our application builds on recent work of Attanasio and Pastorino (2020) (henceforth AP) who analyze whether the program led to changes in the market prices for basic staples: rice, kidney beans, and sugar. This is an important question because the welfare effect of these transfers would clearly depend in part on their impact on prices. While the previous literature had documented that *average* prices were not affected by the program (Hoddinott, Skoufias, and Washburn, 2000), AP argue that sellers charge nonlinear prices and that these nonlinear price schedules had changed.

Because treatment was randomized across villages but means-tested at the household level, some households faced a changing price schedule but no shock to their own income. In our study, we focus our attention on these households because we can be more confident that their augmented utility functions are unchanged across the observation periods. Our objectives are, firstly, to test the augmented utility model and, secondly, to evaluate the welfare impact of price changes using that model. This data set is well suited for analysis using our deterministic model because its panel structure means that we can study each household separately. Following AP, we consider nonlinear prices, which allows us to implement the results in Section 4.

The theoretical part of AP derives the optimal (nonlinear) pricing schedule under the assumption that there is a heterogenous population of households, each of which maximizes a quasilinear utility function subject to a subsistence constraint. This constraint requires a household to consume a minimum number of calories which can be obtained from either the observed bundle x or the numeraire; given x, $\underline{z}(x)$ denotes the minimum amount of the numeraire good needed to meet the calorie threshold. Thus the household can only choose bundles x that satisfy $\psi(x) + \underline{z}(x) \leq M$, where ψ is the price system and M is household wealth. It is worth noting that the augmented utility framework is sufficiently flexible to accommodate this behavior. Indeed, the household could be thought of as maximizing an augmented utility function of the modified-quasilinear form

$$U(x, -e) = \widetilde{U}(x) - \mathbf{K}(e + \underline{z}(x) - M)\, e,$$

where $\mathbf{K}(w) = 1$ if $w \leq 0$ and $\mathbf{K}(w)$ is a very large positive number if $w > 0$. In this way, any $(x, -e)$ (a bundle and its associated expenditure) that leads to a violation of the subsistence constraint incurs a very large disutility and so will never be chosen.

We work with AP's data and refer to them for a detailed explanation. Compared to their analysis, we restrict ourselves to the narrower definition of village ("locality") because the larger units of analysis ("municipality") may not be contained in either the treatment or the control group. Also, because we are interested in intertemporal within-village price variation, we estimate separate price schedules for the same village in different waves as opposed to one price schedule (estimated across waves) per village. This necessitates

	10/98	**03/99**	**11/99**	**11/00**	**2003**
10/98		.035	.024	.006	0
03/99	.913		.198	.052	.015
11/99	.936	.686		.105	.034
11/00	.981	.914	.847		.240
2003	1	.980	.927	.520	

TABLE 1. Fraction of GAPP rationalizable consumers revealed preferring the row wave to the column wave.

being slightly more permissive about data needs, and we estimate prices for all village-good-wave triples that have 20 or more (as opposed to 75 or more) observations. We follow AP in rejecting data for villages where prices strictly increase with quantity sold and where there is insufficient variation in quantities purchased.

We estimate the price schedule for good i in village v at wave t by applying Ordinary Least Squares to

$$\log(\psi_{vti}(q_{vtih})) = b_{vti0} + b_{vti1}\log(x_{vtih}) + \varepsilon_{vtih}. \tag{27}$$

Here h indexes households and $\psi_{vti}(q_{vtih}) = \mathbb{E}[p_{vti}(x_{vtih})|x_{vtih}]\, x_{vtih}$, where $p_{vti}(x_{vtih})$ is the unit price corresponding to quantity x_{vtih}, ε is measurement error, and the expected value is taken over the empirical distribution of reported unit prices corresponding to the same quantity purchased of good i in village-wave (v,t). This is exactly Equation (15) in AP except for being estimated at a less aggregated level.

We test GAPP on untreated households in treated villages (for which we estimate price schedules) with observations in more than one wave and who purchased at least one of the three goods. In our final sample, this leaves us with 2488 households in 177 villages.[36]

We emphasize that GAPP is not vacuously satisfied on these data. Recall that GAPP cannot be violated when two price systems ψ, ψ' are ranked, in the sense that $\psi(x) \geq \psi'(x)$ for all $x \in \mathbb{R}^L_+$. Of the 20556 possible combinations of pairs of waves encountered by households in the data, about 4% have this feature, and only 20 out of 2488 households exclusively face such price pairs and therefore satisfy GAPP vacuously. Nonetheless, 83% of households pass the GAPP test. Most violations were small in the sense of the rationality index ϑ (defined in Section 3.4) being close to 1: fewer than 1% of households were below .9, and fewer than 4% were below .95.

We carried out some illustrative welfare analysis, the results of which are displayed in Tables 1 and 2. Table 1 displays the fractions of GAPP-compliant households that reveal prefer a given wave to another wave. Specifically, each cell in the table corresponds to the

[36]For 554 of these households we have two observations, for 840 households we have three, for 934 households we have four, and for 160 households we have five. There are so few with five observations because many households were enrolled into the program in the final wave and thus removed from our sample.

	03/99	11/99	11/00
75th percentile	5.36	7.26	11.65
Median	3.27	4.66	6.98
25th percentile	1.58	2.44	4.12

TABLE 2. Lower bound of the compensating variation, with 10/98 as the base

fraction of GAPP-rationalizable consumers who reveal prefer (directly or indirectly) the price system in the row wave to the price system in the corresponding column wave.[37] Notice that the data indicates a strong tendency to prefer price systems in later waves. For example, 91.3% of households reveal prefer prices in 03/99 to those in 10/98; the same is true even more strongly when 10/98 is compared against later waves.

Table 2 provides scale for this welfare improvement. We calculate, for each household, the lower bound on the compensating variation, with the price system faced by the household at 10/98 as the base.[38] These values are then ranked. Since more than 90% of households reveal prefer (price systems at) subsequent waves to 10/98, the lower bound of the compensating variation must be positive for more than 90% of households. For example, between 03/99 and 10/98, the median compensating variation is 3.27; thus, based on its observed behavior, one could remove 3.27 from this household in 03/99 and still leave it as well off in 03/99 as in 10/98. Note that the values in this table are not small, given that the household median expenditure in 10/98 on the items considered is 27.48.

These results are consistent with AP's finding that the change in the income distribution induced by Progresa caused a change in sellers' intensity of price discrimination. As a result, poorer households faced higher average prices and wealthier households faced lower ones; since Progresa was means-tested, untreated households fall into the latter category. Thus, the general equilibrium effects of the program could be the reason for the welfare improvements observed in untreated households.

7.2. *RAUM: Testing and welfare analysis on household expenditure data*

We test the RAUM and conduct welfare analyses on two repeated cross-sectional data sets: the U.K. Family Expenditure Survey (FES) and the Canadian Surveys of Household Spending (SHS). Our aim is to show that the data supports the model and to demonstrate that the estimated welfare bounds are informatively tight.

We first analyze the FES which is widely used in the nonparametric demand estimation literature (for instance, by Blundell, Browning, and Crawford (2008), KS, Hoderlein and Stoye (2014) and Adams (2020)). In the FES, about 7000 households are interviewed each

[37]Note that the (indirect) revealed preference relation $\succeq_p^*$ uses demand information at *all* waves in each binary comparison; see the definition of $\succeq_p^*$ in Section 4.

[38]The formula for the lower bound when prices are nonlinear is in Section A.5.

year and they report their consumption expenditures in different commodity groups. Following Blundell, Browning, and Crawford (2008), we derive the real consumption level for each commodity group by deflating it with a price index for that group (which is taken from the annual Retail Prices index). Again following them, we restrict attention to households with cars and children, leaving us with roughly 25% of the original data. We implement tests for 3, 4, and 5 composite goods. Blundell, Browning, and Crawford (2008) analyze the coarsest partition of 3 goods—food, services, and nondurables—and we use their replication files. As in KS, we introduce more commodities by first separating out clothing and then alcoholic beverages from the nondurables.

The data is the sample analog of $\mathscr{D} = \{(p^t, \mathring{\pi}^t)\}_{t=1}^T$ (see Section 5.2). We reiterate the point that, even though this data set is *not* iso-expenditure, we can directly test the RAUM on this data; this contrasts with testing the RUM on this data, which cannot be done directly and must involve a further procedure to estimate an iso-expenditure data set.

We implement the test in blocks of 6 years, i.e., we set $T = 6$. We avoid covering a longer period partly due to the computational demands of calculating A (the matrix of GARP-consistent types; see (18)),[39] but also because a time-invariant distribution of augmented utility functions is only plausible over shorter time horizons, for example because of long term first-order changes to the U.K. income distribution (Jenkins, 2016).

Table 3 displays our results: columns correspond to different blocks of 6 years and rows contain the values of the test statistic and the corresponding p-values. The test statistic J_N is defined by (23), with the identity matrix serving as Ω. Note that for the years 90-95, the test statistic is zero; this means that the sample distribution $\hat{\pi}$ satisfies the rationality condition in Theorem 3 exactly. That is, there is a distribution ν on GARP-consistent types such that $\hat{\pi} = A\nu$. Apart from this case, the sample distribution does not exactly satisfy the rationality condition and so the test statistic is strictly positive; nonetheless, the p-values make it very clear that, overall, our model is not rejected by the FES data.

We also estimated the bounds $[\underline{\mathcal{N}}_{p^t \succeq_p^* p^{t'}}, \overline{\mathcal{N}}_{p^t \succeq_p^* p^{t'}}]$ (as defined by (20) and (21)) on the proportion of households that are revealed better off at prices p^t than at prices $p^{t'}$. For brevity, Table 4 presents a few representative estimates using data from 1975-1980. The second column are the bounds obtained by calculating $\mathbb{1}_{p^t \succeq_p^* p^{t'}} \nu$ from the (not necessarily unique) values of ν that minimize the test statistic (23). In two cases this estimate is unique. Applying the procedure for calculating confidence intervals in Section 6.2, we

[39] That said, new techniques developed in Smeulders, Cherchye, and de Rock (2021) have significantly reduced the computational demands of the problem.

		Year Blocks									
		75-80	76-81	77-82	78-83	79-84	80-85	81-86	82-87	83-88	84-89
3 Goods	Test Statistic (J_N)	0.337	0.917	0.899	0.522	0.018	0.082	0.088	0.095	0.481	0.556
	p-value	0.04	0.34	0.55	0.59	0.99	0.67	0.81	0.91	0.61	0.48
4 Goods	Test Statistic (J_N)	0.4	0.698	0.651	0.236	0.056	0.036	0.037	0.043	0.043	0.232
	p-value	0.25	0.58	0.63	0.91	0.96	0.99	0.96	0.95	0.99	0.68
5 Goods	Test Statistic (J_N)	0.4	0.687	0.705	0.329	0.003	0.082	0.088	0.104	0.103	0.144
	p-value	0.3	0.66	0.68	0.88	0.999	0.96	0.79	0.85	0.9	0.83

		Year Blocks									
		85-90	86-91	87-92	88-93	89-94	90-95	91-96	92-97	93-98	94-99
3 Goods	Test Statistic (J_N)	0.027	1.42	2.94	1.51	1.72	0	0.313	0.7	0.676	0.26
	p-value	0.69	0.3	0.18	0.24	0.21	1	0.59	0.48	0.6	0.83
4 Goods	Test Statistic (J_N)	0.227	0.025	0.157	0.154	0.004	1.01	0.802	0.872	0.904	0.604
	p-value	0.48	0.96	0.8	0.73	0.97	0.21	0.31	0.57	0.65	0.74
5 Goods	Test Statistic (J_N)	0.031	0.019	0.018	0.019	0.023	0.734	0.612	0.643	0.634	0.488
	p-value	0.85	0.98	0.97	0.91	0.83	0.22	0.4	0.72	0.78	0.79

TABLE 3. Test Statistics, p-values for 6 budget sequences of the FES. Bootstrap size is $R = 1000$.

Comparison	Estimated Bounds	Confidence Interval
$p^{1976} \succ_p^* p^{1977}$	$[.150, .155]$	$[.13, .183]$
$p^{1977} \succ_p^* p^{1976}$	$\{.803\}$	$[.784, .831]$
$p^{1979} \succ_p^* p^{1980}$	$[.517, .530]$	$[.487, .56]$
$p^{1980} \succ_p^* p^{1979}$	$\{.463\}$	$[.436, .497]$

TABLE 4. Estimated bounds and confidence intervals for the proportion of consumers who reveal prefer one price to another one in the FES data. Data used are for 1975-1980. Bootstrap size is $R = 1000$.

obtain the intervals displayed (which necessarily contain the estimated bounds). We note that the width of these intervals is less than .1 throughout, so they are quite informative.[40]

For our second empirical application using Canadian data, we use the replication kit of Norris and Pendakur (2013, 2015). Like the FES, the SHS is a publicly available, annual data set of household expenditures in different categories. We study annual expenditures in 5 categories that constitute a large share of the overall expenditure on nondurables:

[40]Note that, the true values of the proportion of the population satisfying $p^t \succ_p^* p^{t'}$ and $p^{t'} \succ_p^* p^t$ typically add up to strictly less than 1 because there is no revealed price preference relation between p^t and $p^{t'}$ for some share of the population. See, for example, type 1 consumers in Example 4.

Province		Year Blocks							
		97-02	98-03	99-04	00-05	01-06	02-07	03-08	04-09
Alberta	Test Statistic (J_N)	.07	0	0	0	0	0	.003	4.65
	p-value	.94	1	1	1	1	1	.98	.04
British Columbia	Test Statistic (J_N)	.89	.56	.48	.07	.05	6.23	8.87	8.71
	p-value	.47	.47	.98	.96	.97	.05	.02	.01
Manitoba	Test Statistic (J_N)	0	0	0	0	0	0	.01	.01
	p-value	1	1	1	1	1	1	1	1
New Brunswick	Test Statistic (J_N)	.08	.05	0	0	0	.60	.58	.57
	p-value	.94	.94	1	1	1	.58	.79	.68
Newfoundland	Test Statistic (J_N)	.10	.32	.29	.29	.38	3.08	2.30	2.08
	p-value	.85	.90	.91	.87	.81	.21	.35	.27
Nova Scotia	Test Statistic (J_N)	.05	.03	0	0	0	0	.93	1.02
	p-value	.97	.98	1	1	1	1	.69	.58
Ontario	Test Statistic (J_N)	.064	.040	.035	0	0	0	0	0
	p-value	.98	.95	.91	1	1	1	1	1
Quebec	Test Statistic (J_N)	.11	0	0	0	0	.51	.54	.49
	p-value	.88	1	1	1	1	.67	.67	.65
Saskatchewan	Test Statistic (J_N)	0	0	0	0	0	.02	.02	0
	p-value	1	1	1	1	1	1	1	1

TABLE 5. Test Statistics and p-values for sequences of 6 budgets of the SHS. Bootstrap size is $R = 1000$.

Comparison	Estimated Bounds	Confidence Interval
$p^{1998} \succ_p^* p^{2001}$	$\{.099\}$	$[.073, .125]$
$p^{2001} \succ_p^* p^{1998}$	$\{.901\}$	$[.875, .927]$
$p^{1999} \succ_p^* p^{2002}$	$[.299, .341]$	$[.272, .385]$
$p^{2002} \succ_p^* p^{1999}$	$[.624, .701]$	$[.594, .728]$

TABLE 6. Estimated bounds and confidence intervals for the proportion of consumers who reveal prefer one price to another one in the SHS data. Data used are for 1997-2002 in British Columbia. Bootstrap size is $R = 1000$.

food purchased (at home and in restaurants), clothing and footwear, health and personal care, recreation, and alcohol and tobacco. As Table 4 shows, the SHS data allows us to analyze the data separately for the nine most populous provinces. The number of households in each province-year range from 291 (Manitoba, 1997) to 2515 (Ontario, 1997). We use province-year prices indices (as constructed by Norris and Pendakur (2015)) and deflate them using province-year CPI data from Statistics Canada to get real price indices.

Table 6 displays the test statistics and associated p-value for each province and every 6 year block. Compared to the FES data, there are two notable differences. The first is that many more test statistics are exactly zero; that is, the observed choice frequencies are rationalized by the random augmented utility model. The second is that, for a small proportion of year blocks, there are statistically significant positive test statistics (in particular, the last three columns for British Columbia). Nonetheless, the p-values taken together do not reject the model if multiple testing is taken into account; for example, step-down procedures would terminate at the first step (that is, Bonferroni adjustment). Finally, we can also estimate the proportion of the population with a revealed preference for one year's prices over another. We provide an illustration in Table 6; notice that the confidence intervals are informative, with a width no greater than 0.15.

8. Conclusion

We propose a revealed price preference relation that generates a nonparametric ranking of price vectors; a consistency (no-cycles) condition in this relation characterizes an augmented utility model in which consumers get utility from consumption and disutility from expenditure. This model is a natural generalization of quasilinearity and, furthermore, captures some prominent behavioral models of consumption. The model is also flexible enough to accommodate nonlinear prices, discrete choice and other consumption environments. We develop the theoretical basis for welfare analysis in our model.

We generalize our model to a random utility context which is suitable for welfare analysis using repeated cross-sectional (as opposed to single-agent) data and show how to statistically test this random augmented utility model. A strength of this model is that it can be directly taken to household expenditure data in contrast to the standard random utility model which requires an additional round of estimation to account for the dependence of expenditure on prices. We develop novel econometric theory to determine the proportion of consumers who are made better or worse off by a price change. This theory—which derives bounds on linear transforms of partially identified vectors—is a standalone contribution which has broader applications beyond those in this paper.

Finally, we operationalize both the deterministic and random versions of our model in separate applications to single-agent and repeated cross-sectional data. We confirm that our model is supported by data and can be used for meaningful welfare analysis.

References

ADAMS, A. (2020): "Mutually Consistent Revealed Preference Bounds," *American Economic Journal: Microeconomics*, 12(1), 42–74.

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

AFRIAT, S. N. (1973): "On a System of Inequalities in Demand Analysis: An Extension of the Classical Method," *International Economic Review*, 14(2), 460–472.

AMADOR, M., I. WERNING, AND G.-M. ANGELETOS (2006): "Commitment vs. Flexibility," *Econometrica*, 74(2), 365–396.

ATTANASIO, O., AND E. PASTORINO (2020): "Nonlinear Pricing in Village Economies," *Econometrica*, 88(1), 207–263.

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 841–890.

BHATTACHARYA, D. (2015): "Nonparametric Welfare Analysis for Discrete Choice," *Econometrica*, 83(2), 617–649.

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

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

BORDALO, P., N. GENNAIOLI, AND A. SHLEIFER (2013): "Salience and Consumer Choice," *Journal of Political Economy*, 121(5), 803–843.

CHETTY, R. (2009): "Sufficient Statistics for Welfare Analysis: A Bridge between Structural and Reduced-Form Methods," *Annual Review of Economics*, 1(1), 451–488.

CHETTY, R., A. LOONEY, AND K. KROFT (2009): "Salience and Taxation: Theory and Evidence," *American Economic Review*, 99(4), 1145–77.

DE PAULA, A., S. RICHARDS-SHUBIK, AND E. TAMER (2018): "Identifying Preferences in Networks With Bounded Degree," *Econometrica*, 86(1), 263–288.

DUPAS, P., AND J. ROBINSON (2013): "Why Don't the Poor Save More? Evidence from Health Savings Experiments," *American Economic Review*, 103(4), 1138–71.

FELDMAN, N. E. (2010): "Mental Accounting Effects of Income Tax Shifting," *Review of Economics and Statistics*, 92(1), 70–86.

GABAIX, X. (2014): "A Sparsity-Based Model of Bounded Rationality," *Quarterly Journal of Economics*, 129(4), 1661–1710.

——— (2019): "Behavioral Inattention," in *Handbook of Behavioral Economics: Applications and Foundations 1*, vol. 2, pp. 261–343.

GALICHON, A., AND M. HENRY (2011): "Set Identification in Models with Multiple Equilibria," *Review of Economic Studies*, 78(4), 1264–1298.

HASTINGS, J., AND J. M. SHAPIRO (2018): "How are SNAP Benefits Spent? Evidence from a Retail Panel," *American Economic Review*, 108(12), 3493–3540.

HASTINGS, J. S., AND J. M. SHAPIRO (2013): "Fungibility and Consumer Choice: Evidence from Commodity Price Shocks," *Quarterly Journal of Economics*, 128(4), 1449–1498.

HAUSMAN, J. A., AND W. K. NEWEY (2016): "Individual Heterogeneity and Average Welfare," *Econometrica*, 84(3), 1225–1248.

HEIDHUES, P., AND B. KŐSZEGI (2008): "Competition and Price Variation when Consumers are Loss Averse," *American Economic Review*, 98(4), 1245–68.

HODDINOTT, J., E. SKOUFIAS, AND R. WASHBURN (2000): "The Impact of PROGRESA on Consumption: A Final Report," Discussion paper.

HODERLEIN, S., AND J. STOYE (2014): "Revealed Preferences in a Heterogeneous Population," *The Review of Economics and Statistics*, 96(2), 197–213.

JENKINS, S. P. (2016): "The Income Distribution in the UK," in *Social Advantage and Disadvantage*, ed. by H. Dean, and L. Platt, pp. 135–160. Oxford University Press.

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

——— (2019): "Nonparametric Counterfactuals of Random Utility Models," *Discussion paper*, arXiv:1902.08350.

KLEVEN, H. (2021): "Sufficient Statistics Revisited," *Annual Review of Economics*, forthcoming.

KŐSZEGI, B., AND M. RABIN (2006): "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 121(4), 1133–1165.

LAZZATI, N., J. K.-H. QUAH, AND K. SHIRAI (2018): "Nonparametric Analysis of Monotone Choice," *Working Paper*.

MCFADDEN, D., AND K. RICHTER (1991): "Stochastic rationality and revealed stochastic preference," in *Preferences, Uncertainty and Rationality*, ed. by J. Chipman, D. McFadden, and K. Richter, pp. 161–186. Westview Press, Boulder.

MCFADDEN, D. L. (2005): "Revealed Stochastic Preference: A Synthesis," *Economic Theory*, 26(2), 245–264.

NEVO, A. (2000): "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," *Journal of Economics and Management Strategy*, 9(4), 513–548.

NORRIS, S., AND K. PENDAKUR (2013): "Imputing Rent in Consumption Mmeasures, with an Application to Consumption Poverty in Canada, 1997–2009," *Canadian Journal of Economics*, 46(4), 1537–1570.

——— (2015): "Consumption Inequality in Canada, 1997 to 2009," *Canadian Journal of Economics*, 48(2), 773–792.

PRELEC, D., AND G. LOEWENSTEIN (1998): "The Red and the Black: Mental Accounting of Savings and Debt," *Marketing science*, 17(1), 4–28.

SAKAI, Y. (1977): "Revealed Favorability, Indirect Utility, and Direct Utility," *Journal of Economic Theory*, 14(1), 113–129.

SAKOVICS, J. (2011): "Reference Distorted Prices," *Quantitative Marketing and Economics*, 9(4), 339–363.

SMEULDERS, N., L. CHERCHYE, AND B. DE ROCK (2021): """Nonparametric Analysis of Random Utility Models: Computational Tools for Statistical Testing," *Econometrica*, 89(1), 437–455.

THALER, R. H. (1999): "Mental Accounting Matters," *Journal of Behavioral Decision Making*, 12(3), 183–206.

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

* * * * * *

ONLINE APPENDIX

* * * * * *

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*

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

Once this is established, it is straightforward to show that

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

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).[41] 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,

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

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

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 $\{x \in \mathbb{R}_+^L : p^t \cdot x \leq p^t \cdot x^t\}$. Thus

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

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 houesholds 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.5. 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.

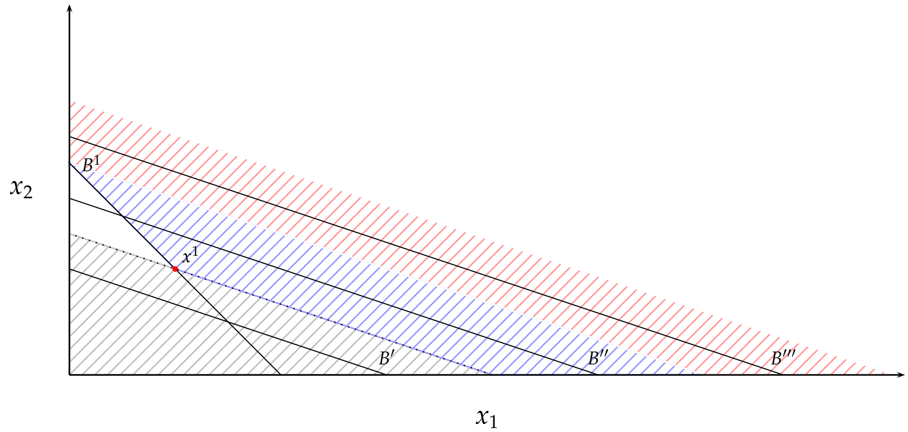
Consider first the constrained-optimization model. (a) Suppose that $w < \hat{p} \cdot x^1 = 7/4$; the line of points/bundles incurring this level of expenditure is depicted by B' in Figure A.1a. In this case, any bundle with $\hat{p} \cdot x = w$ will *not* be revealed preferred to x^1 and so x can 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.

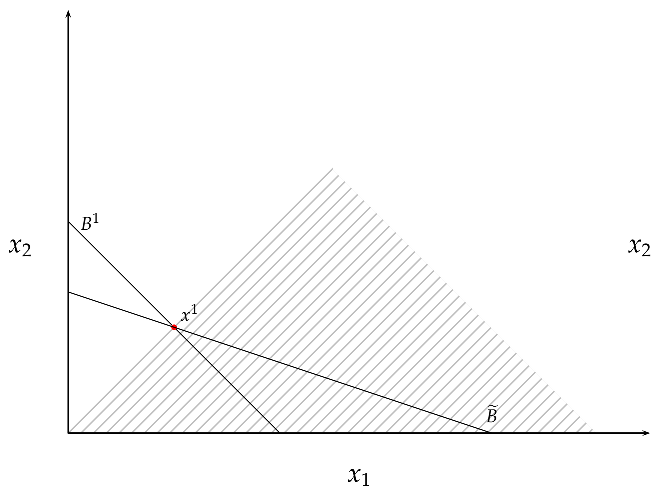
What happens to the predictions of the constrained-maximization model when the utility function is required to be homothetic? It is well known that homothetic utility functions generate demand that is linear in cones. Therefore, for any $x \in \mathbb{R}^2_+$, the data set $\{(p^1, x^1), (\hat{p}, x)\}$ 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 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.

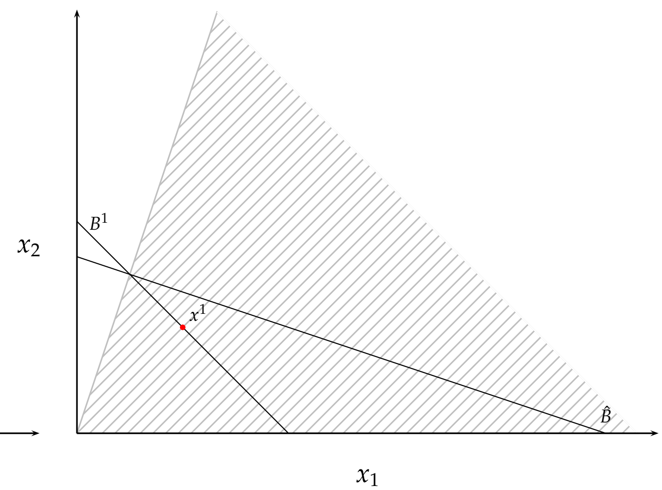
In the case of the augmented utility model, recall that if x satisfies $\hat{p} \cdot x = p^1 \cdot x^1 = 2$, then $\{(p^1, x^1), (\hat{p}, x)\}$ satisfies GAPP if and only if it satisfies GARP (see Proposition 1).



(a) Counterfactuals using GARP



(b) Counterfactuals using HARP



(c) Counterfactuals using GAPP

FIGURE A.1. Counterfactuals with different consumption models

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.6. 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:

$\psi^{t'} \succeq_{p,\overline{\vartheta}} \psi^t$ *if* $\psi^{t'}(x^t) \leq \vartheta^{t'}\psi^t(x^t)$ *and* $\psi^{t'} \succ_{p,\overline{\vartheta}} \psi^t$ *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.[42] 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 $(\vartheta, \vartheta, \dots, \vartheta)$-rationalization for some $\vartheta > 0$ (see Remark 3 following Theorem A.4.1). This guarantees that the *rationality index*, given by

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

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 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

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

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

[42] 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 $X \times X$.

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:

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

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 $\widetilde{U} : \mathbb{R}^L_+ \to \mathbb{R}$.[43] By definition, the critical cost efficiency index c^* satisfies

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

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

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

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

[43] 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 $t = 1, 2, \ldots, T$) 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.

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 $\hat{x} \geq_K x^t$.[44] 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 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.

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

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:

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

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

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 $\inf(\mu_c) \geq m^{\bar{s}}_c$, 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 $m^{\bar{s}}_c + \epsilon$.

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^s_c such that

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

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

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

Thus $\inf(\mu_c) \geq m^s_c$ 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

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

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 $m^{s'}_c \geq 0$, 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 $m^s_c < 0$, 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

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

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[45]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*

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

[45]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 2. 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 3. 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

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

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

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

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). ■

APPENDIX REFERENCES

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.

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.