

Price Heterogeneity as a source of Heterogenous Demand*

John K.-H. Quah[†] Gerelt Tserenjigmid[‡]

February 7, 2025

Abstract

We explore heterogenous prices as a source of heterogenous or stochastic demand. Heterogenous prices could arise either because there is actual price variation among consumers or because consumers (mis)perceive prices differently. Our main result says the following: if heterogenous prices have a distribution among consumers that is (in a sense) stable across observations, then a model where consumers have a common utility function but face heterogenous prices has precisely the same implications as a heterogenous preference/random utility model (with no price heterogeneity).

Keywords. random utility, stochastic demand, augmented utility, misperceived prices, reference prices, equivalence scales, Afriat's Theorem.

1 Introduction

Why is demand heterogenous? Two consumers with the same budget and facing the same prices would often choose to spend their money differently. The same consumer, when

* We thank Kwok Ping Tsang for helpful discussions. We also thank the seminar participants at U of Connecticut, CUHK-HKU-HKUST, KAIST, UC Santa Cruz, U of Sussex, Purdue U, U of Pennsylvania, and U of Manitoba.

[†] Department of Economics, National University of Singapore. E-mail: ecsqkhj@nus.edu.sg.

[‡]Department of Economics, University of California, Santa Cruz. Email: gtserenj@ucsc.edu.

presented with the same budgetary constraints, may also choose different purchases at different times. The most common and obvious way of modelling heterogenous behavior across consumers (or stochastic behavior in the case of a single consumer) is to assume that demand is governed by a distribution of preferences (rather than a single preference). When the distribution of preferences is *stable* across multiple observations, this leads to a random utility model of stochastic choice (see, for example, McFadden and Richter (1990), Kitamura and Stoye (2018), and Deb et al. (2023)). But preference heterogeneity is not the only source of variation in choice behavior. More recently, the literature on stochastic choice has highlighted the presence of other sources of heterogenous choice. For example, choice could be stochastic due to random consideration/attention sets (Manzini and Mariotti (2014), Brady and Rehbeck (2016), Cattaneo et al. (2020)) or due to deliberate randomization (Fudenberg et al. (2015), Cerreia-Vioglio et al. (2019)). In this paper, we wish to highlight another reason why choices can be heterogenous: the presence of heterogenous prices across consumers. Price heterogeneity could occur for at least two reasons.

Firstly, even when the economist uses only a single set of prices (for different goods) for all the consumers being studied, consumers *in fact* face different prices due to **price dispersion** (Stigler (1961)), which arises because they live in different locations and/or shop at different stores, and it is not always worthwhile for them to search for the cheapest price.¹ Related to this, empirical demand analysis often involves the aggregation of highly granular demand information into broad categories (such as ‘food’ or ‘clothing’); the price index constructed for a given category of goods is not tailored to each consumer and so may not accurately represent the prices encountered by a given consumer for goods in that category.

Secondly, even when consumers face the same prices, they may still (mis)perceive prices differently, leading to different demand behavior. **Price misperception** could occur because of *inattention to prices* (Chetty et al. (2009), Gabaix (2014), Matějka (2015)) or because consumers over- or under-estimate their usage of a prepaid good or service (DellaVigna and Malmendier (2006)). A related phenomenon is that the demand for a product could be affected by its *reference price* (see Koszegi and Rabin (2006), Heidhues and Köszegi (2008), Bordalo et al. (2013)), with the same price being perceived as high or low depending on the reference

¹ For empirical evidence on price dispersion, see Sorensen (2000), Brown and Goolsbee (2002), Baye et al. (2004), Van Nieuwerburgh and Weill (2010), and Kaplan et al. (2019).

price. The variation in reference prices across consumers then leads to heterogenous demand behavior.²

In our formal analysis, we assume that the analyst has information on the expenditure allocations (across K goods) of N consumers at different price indices. Formally, a data set with T observations can be represented by

$$\mathcal{E} = \{(\mathbf{e}^{i,t})_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T},$$

where $\mathbf{e}^{i,t} = (e_k^{i,t})_{k \in K}$ is the expenditure allocation of consumer i and $\bar{\mathbf{p}}^t = (\bar{p}_k^t)_{k \in K}$ are the prices of the K goods at observation t .³ If there is no price heterogeneity then consumer i 's consumption at observation t is the bundle $\mathbf{x}^{i,t}$, where the k th entry, $x_k^{i,t} = e_k^{i,t}/\bar{p}_k^t$. We investigate the implications of price heterogeneity in two different models of demand.

Constrained utility-maximization. Afriat's Theorem (Afriat (1967), Varian (1982)) allows us to check whether a consumer's demand behavior is consistent with the maximization of a utility function subject to a budget constraint, i.e., whether, for a consumer i , there is a utility function u^i such that $u^i(\mathbf{x}^{i,t}) \geq u^i(\mathbf{x})$ for all \mathbf{x} satisfying $\bar{\mathbf{p}}^t \cdot \mathbf{x} \leq \sum_{k \in K} e_k^{i,t}$. The theorem states that this holds if and only if the implied consumption bundles of consumer i , $\{\mathbf{x}^{i,t}\}_{t \in T}$, along with the observed prices $\{\bar{\mathbf{p}}^t\}_{t \in T}$, satisfy a property called the generalized axiom of revealed preference (or GARP, for short). If this holds for each agent i , then we say that \mathcal{E} can be *rationalized by heterogenous preferences*

The problem we pose is different. Instead of allowing consumers to have different utility functions, we require them to have the *same* utility function but allow each consumer to face idiosyncratic (whether real or perceived) prices $\mathbf{p}^{i,t}$. We say that the data set \mathcal{E} can be *rationalized by heterogenous prices* if there is a common utility function u such that consumer i 's expenditure allocations at different observations maximize u , given the idiosyncratic prices $\{\mathbf{p}^{i,t}\}_{t \in T}$ encountered by consumer i ; formally, we require $u(\hat{\mathbf{x}}^{i,t}) \geq u(\mathbf{x})$, where $\hat{x}_k^{i,t} = e_k^{i,t}/p_k^{i,t}$ and \mathbf{x} satisfies $\mathbf{p}^{i,t} \cdot \mathbf{x} \leq \sum_{k \in K} e_k^{i,t}$.

The interpretation of the bundles $\{\hat{\mathbf{x}}^{i,t}\}_{t \in T}$ depends on the interpretation we give to $\{\mathbf{p}\}^{i,t}$.⁴

² There is a large literature in marketing on reference-price purchasing. See Mazumdar et al. (2005) for a survey.

³Note that T (similarly N and K) denotes both a set and the number of elements in the set.

⁴In section 4, we also study the implications of price heterogeneity when preferences are heterogenous.

- If we interpret $\{\mathbf{p}^{i,t}\}_{t \in T}$ as the true prices faced by consumer i , then $\hat{\mathbf{x}}^{i,t}$ is the true bundle purchased by this consumer at observation t ; this bundle is not observed by the researcher, who only observes i 's expenditure allocation $\mathbf{e}^{i,t}$ and the price index $\bar{\mathbf{p}}^t$. Under this interpretation, if \mathcal{E} can be rationalized by heterogenous prices, then all consumers share the same utility function and each of them is rational, in the sense that there are idiosyncratic prices at which each consumer's observed expenditure allocations are optimal.
- If we interpret $\{\mathbf{p}^{i,t}\}_{t \in T}$ as consumer i 's misperceived prices, then $\hat{\mathbf{x}}^{i,t}$ is the bundle that this consumer *thinks* she is consuming when making her expenditure allocation decision across goods. We are assuming that the expenditure allocation $\mathbf{e}^{i,t}$ is the one consumer i implements so that the consumer's actual consumption at observation t is $\mathbf{x}^{i,t}$. Under this interpretation, if \mathcal{E} can be rationalized by heterogenous prices, then all consumers share the same utility function and each consumer's expenditure allocations are also rational, once her misperception of prices is taken into account.

Our notion of rationalization, in itself, is too permissive to lead to meaningful implications on data; what is needed are sensible restrictions on how heterogenous prices are distributed across consumers. In Section 3, we require the heterogenous prices to be *correct on average*, in the sense that, at each observation t and for each good k , we have $(\sum_{i \in N} p_k^{i,t})/N = \bar{p}_k^t$. In this case, we show that *any* data set \mathcal{E} can be rationalized by heterogenous prices. In particular, price heterogeneity, in this sense, is more powerful than preference heterogeneity in its ability to explain observed behavior, since each consumer's expenditure allocation (along with the prevailing price indices) must satisfy GARP in the latter model.

In Section 4, we impose a stronger restriction on price heterogeneity. We assume that price heterogeneity has a *stable structure* across observations. This assumption requires that the distribution of prices for good k (across consumers) at some observation t is the same as the distribution of prices for good k at observation t' , after normalizing by the average price of k at each observation. When this condition is imposed on the price distributions, we find that stable price heterogeneity leads to the *same* restrictions on \mathcal{E} as preference heterogeneity. In other words, *a data set \mathcal{E} can be rationalized by heterogenous preferences*

if and only if it can be rationalized by heterogenous prices with a stable structure.⁵

Augmented utility-maximization. In Section 5, we study the impact of price heterogeneity in a second model of demand. Instead of a consumer who maximizes utility subject to a budget constraint, we posit a consumer who maximizes a utility function where expenditure enters directly. Formally, the utility of a bundle \mathbf{x} acquired at cost e is $V(\mathbf{x}, -e)$, where V is strictly increasing in the last argument (in other words, higher expenditure gives strictly lower utility). Following Deb et al. (2023), we refer to V as an (*expenditure-*)*augmented utility function*. As a special case, V can have the familiar quasilinear form, where $V(x, -e) = U(x) - e$ for some function U .

Setting aside price misperception for now, at prices \mathbf{p} the consumer chooses the bundle \mathbf{x} to maximize $V(\mathbf{x}, -\mathbf{p} \cdot \mathbf{x})$. Notice that she is deterred from buying an arbitrarily large bundle because of the disutility that higher expenditure incurs. An augmented utility model is particularly suited for the study of demand under price misperception because it allows for any bundle to be purchased, even a bundle chosen under misperceived prices; on the other hand, in a model with constrained optimization, such a bundle will not be implementable if it violates the budget constraint.⁶ Given a data set \mathcal{E} , we could ask whether \mathcal{E} is rationalizable by heterogeneous augmented utility functions, assuming that (at each observation) prices are homogeneous across consumers. Deb et al. (2023) provides a characterization of those data sets \mathcal{E} which can be rationalized in this sense.

We are interested in investigating the conditions under which \mathcal{E} can be rationalized by a model where all consumers have the same augmented utility function but associates to each bundle \mathbf{x} an expenditure level that is not simply $\bar{\mathbf{p}}^t \cdot \mathbf{x}$. We allow for heterogenous prices $\{\mathbf{p}^{i,t}\}_{t \in T}$, with consumer i (mentally) associating with the bundle \mathbf{x} a *behavioral expenditure level* $\phi(\bar{\mathbf{p}}^t \cdot \mathbf{x}, \mathbf{p}^{i,t} \cdot \mathbf{x})$. Note that this value depends on the true cost of the bundle, $\bar{\mathbf{p}}^t \cdot \mathbf{x}$, as well as on $\mathbf{p}^{i,t} \cdot \mathbf{x}$. Different functional restrictions on ϕ allow the behavioral expenditure

⁵ Our result also establishes another equivalence statement, which may be interesting in its own right. We show that if \mathcal{E} can be rationalized by heterogenous preferences, then all the preferences can be drawn from a family of utility functions related to each other via equivalence scales. Equivalence scales are widely used in empirical models of demand to capture (among other things) the effects of household composition (see, for example, Lewbel (1989), Blundell and Lewbel (1991), Deaton (1997), and Lewbel and Pendakur (2017)) and in models of market demand they can be used to endow a parametric structure to preference heterogeneity (see, for example, Grandmont (1992) and Quah (1997)).

⁶ Our solution to this problem is to assume that the consumer always implements her *expenditure* allocation decision. For another solution to this problem, see Gabaix (2014).

level to capture different notions of price (and, hence, expenditure) misperception, including price inattention and reference prices; in the former case, $\mathbf{p}^{i,t}$ are consumer i 's idiosyncratic perceived prices and in the latter, $\mathbf{p}^{i,t}$ are consumer i 's idiosyncratic reference prices. The data set \mathcal{E} is said to be rationalized by an augmented utility function V and heterogenous prices if $\mathbf{x}^{i,t}$ (consumer i 's actual consumption bundle given the true prices $\bar{\mathbf{p}}^t$) maximizes $V(\mathbf{x}, -\phi(\mathbf{p} \cdot \mathbf{x}, \mathbf{p}^{i,t} \cdot \mathbf{x}))$ at each observation t . We show that, *provided ϕ satisfies some mild conditions, \mathcal{E} can be rationalized by an augmented utility function and heterogenous prices with a stable structure if and only if it can be rationalized by heterogenous augmented utility functions*. Loosely speaking, heterogeneity in price misperception has the same observable implications as heterogeneity in preferences when consumers maximize augmented utility.

2 Model and basic definitions

There are N consumers who purchase bundles of K goods. We denote by $e_k^{i,t}$ the amount that consumer i spends on good k at observation t . We assume that $e_k^{i,t} \geq 0$ and $\sum_{k \in K} e_k^{i,t} > 0$, i.e., consumer i 's total expenditure is always strictly positive though expenditure on a given good may equal zero. These expenditure allocations are observed by an analyst. Consumer i at observation t bases his expenditure allocation decision on the prices $\mathbf{p}^{i,t} = (p_k^{i,t})_{k \in K}$. Notice that these prices can be idiosyncratic; we assume that they are *not* observed by the analyst.⁷ However, the analyst does observe \bar{p}_k^t , which is an aggregate (or index) of the prices of good k at observation t . Assuming that the analyst makes $T < \infty$ observations, we may denote the data set collected in the form

$$\mathcal{E} = \{(\mathbf{e}^{i,t})_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}, \quad (1)$$

where $\mathbf{e}^{i,t} = (e_k^{i,t})_{k \in K}$ and $\bar{\mathbf{p}}^t = (\bar{p}_k^t)_{k \in K}$. (Notice that we have abused notation by using T to denote both the largest number of observations and the set $\{1, 2, \dots, T\}$, and similarly with K and N .) We assume that $p_k^{i,t} > 0$ and $\bar{p}_k^t > 0$, i.e., for every good and at every observation, the prices faced by the consumers, and the price index are *strictly* positive.

There are two possible interpretations of the idiosyncratic prices:

(1) Observational Errors The analyst's price index $\bar{\mathbf{p}}^t$ represents a simplification of a

⁷ For example, in the panel study of income dynamics (PSID) survey, which is one of the most applied data sets in economics, households report their expenditures but not the individual prices that they face.

more complex environment; $\mathbf{p}^{i,t}$ are the true prices faced by consumer i and they could be different from $\bar{\mathbf{p}}^t$ because of geographical variation in prices, variation in prices within a particular observation period t , or unobserved costs which are incurred by the consumer when purchasing a particular good. Price heterogeneity is also plausible when each product represents a broad category (such as ‘food’ or ‘clothing’). In these situations, the data available to the observer takes the form of expenditure on a product category k by each consumer and a price index for goods in that category. Simply deflating each consumer’s expenditure by this index can be misleading if consumers are systematically different in the types of goods they purchase within category k .

(2) Price Misperception Consumer i makes his expenditure allocation decision based on an idiosyncratic perception of prices at observation t . The prices $\mathbf{p}^{i,t}$ could be incorrect or misperceived, in the sense that they are different from the true prices $\bar{\mathbf{p}}^t$, which (unlike in the first interpretation) are perfectly observed by the analyst. Price misperception could occur because consumers are inattentive to prices (Chetty et al. (2009), Gabaix (2014), Matějka (2015)), for example, failing to incorporate sales taxes or shipping charges. Sometimes the phenomenon could also be exacerbated by firms engaging in price shrouding (Gabaix and Laibson (2006), Brown et al. (2010)). Phenomena, where people misperceive their usage of a prepaid good/service such as a gym membership, could also be considered as a form of price misperception (see DellaVigna and Malmendier (2006)).⁸

We would like to find conditions under which the data set \mathcal{E} has the following properties: (i) at each observation, the heterogeneous prices across consumers are consistent with the observed price indices in some sense and (ii) there is a single utility function that could explain the observed expenditure patterns. To explain this more precisely, some definitions are in order.

For given expenditure and price vectors $\mathbf{e} \in \mathbb{R}_+^K$ and $\mathbf{p} \in \mathbb{R}_{++}^K$, we denote the implied consumption bundle by $\mathbf{x}(\mathbf{e}, \mathbf{p})$; formally, $\mathbf{x}(\mathbf{e}, \mathbf{p}) \equiv (e_k/p_k)_{k \in K}$. The bundles which cost as

⁸ Price misperception is a form of bounded rationality, but it is formally different from models (such as Masatlioglu et al. (2012) and Manzini and Mariotti (2014)) where agents only consider some of the options open to them. The ‘consideration set’ in these models is always a subset of the budget set, whereas misperceived prices lead to a misperceived budget that is typically *not* a subset of the true budget set.

much or less than $\mathbf{x}(\mathbf{e}, \mathbf{p})$ is given by the *budget set*

$$\mathcal{B}(\mathbf{p}, m) = \{\mathbf{x} \in \mathbb{R}_+^K \mid \mathbf{p} \cdot \mathbf{x} \leq m\},$$

where $m = \sum_{k \in K} e_k$ is the total expenditure.

Utility Maximization. Let $\mathcal{P} = \{\mathbf{p}^{i,t}\}_{(i,t) \in N \times T}$ be a collection of heterogenous prices. We say that \mathcal{P} **rationalizes** $\mathcal{E} = \{(\mathbf{e}^{i,t})_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}$ if there is a utility function $u : \mathbb{R}_+^K \rightarrow \mathbb{R}$ such that for each $(i, t) \in N \times T$,

$$u(\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t})) \geq u(\mathbf{x}) \text{ for any } \mathbf{x} \in \mathcal{B}(\mathbf{p}^{i,t}, m^{i,t}), \quad (2)$$

$$\text{where } m^{i,t} = \sum_{k \in K} e_k^{i,t}.$$

In other words, there is a utility function (common to all consumers) such that each consumer's demand at observation t (implied by the observed expenditure shares and the idiosyncratic prices) gives greater utility than any bundle that costs as much or less. The data set \mathcal{E} is **rationalizable with heterogenous prices** if there are heterogenous prices \mathcal{P} that rationalize \mathcal{E} , typically subject to some consistency condition on \mathcal{P} which we shall discuss later.

Notice that in this notion of rationalization, there is *heterogeneity in prices across consumers* but no heterogeneity in preferences. The precise interpretation of $\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t})$ depends on our interpretation of idiosyncratic prices.

In the first (Errors) interpretation of idiosyncratic prices, these prices are true prices and so $\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t})$ is the actual bundle purchased by the consumer, while (2) guarantees that each consumer is truly optimizing at the utility function u .

In the second (Misperceptions) interpretation, the bundle $\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t})$ is what consumer i expects to be able to buy at his perceived prices and (2) guarantees that, with utility function u , consumer i thinks that his expenditure allocation decision is optimal, given his perception of prices. Since our model assumes that the allocated expenditures are *observed*, we are assuming that the consumer implements the expenditure allocation decision, so that (with $\bar{\mathbf{p}}^t$ being the true prices) the actual level of good k consumed by i at observation t is $e_k^{i,t} / \bar{p}_k^t$. Notice that the expenditure allocation decision by the consumer will typically not be optimal given the true prices, and the consumer may even become aware of the true prices

as he implements his expenditure allocation decision, but we are assuming that he does not re-optimize: he retains the expenditure allocation decision and simply adjusts by buying more or less of the good.⁹

Price Consistency. To render this model interesting, there must be some way of disciplining the sort of price heterogeneity permitted by the model. We do this via two notions of price consistency.

We call a function $W : \mathbb{R}_{++}^N \rightarrow \mathbb{R}_{++}$ a **price aggregator** if it is strictly increasing, continuous, and $W(p, \dots, p) = p$ for any $p > 0$. For example, the arithmetic average across all consumers

$$W((p^i)_{i \in N}) = \frac{\sum_{i \in N} p^i}{N} \quad (3)$$

is a price aggregator. We assume that, at each observation t and for each good k , there is an aggregator function W_k^t that is known to the analyst. In our first result (Proposition 1), we assume that the aggregator functions $\mathcal{W} = \{W_k^t\}_{k \in K, t \in T}$ satisfy one of the following regularity conditions:

- i) $\lim_{p^i \rightarrow +\infty} W_k^t(p^i, (p^j)_{j \neq i}) = +\infty$ for any $(p^j)_{j \neq i} \in \mathbb{R}_{++}^{N-1}$ and $(t, k) \in T \times K$, or
- ii) $\lim_{p^i \rightarrow 0} W_k^t(p^i, (p^j)_{j \neq i}) = 0$ for any $(p^j)_{j \neq i} \in \mathbb{R}_{++}^{N-1}$ and $(t, k) \in T \times K$.

Given the price aggregators $\mathcal{W} = \{W_k^t\}_{(k,t) \in K \times T}$, the idiosyncratic prices $\mathcal{P} = \{\mathbf{p}^{i,t}\}_{(i,t) \in N \times T}$ are **consistent in expectation** with the price indices $\{\bar{\mathbf{p}}^t\}_{t \in T}$ if, for any $(k, t) \in K \times T$,

$$\bar{p}_k^t = W_k^t((p_k^{i,t})_{i \in N}). \quad (4)$$

In the Errors interpretation of idiosyncratic prices, \mathcal{W} could be the actual formulae used in the calculation of the observed price indices, so that (4) requires that the hypothesized \mathcal{P} agrees on average with the aggregate price data. In the Misperceptions interpretation, (4) says that the consumers are on average correct in their price perceptions (according to some suitable notion of average imposed by the modeler via \mathcal{W}).¹⁰

⁹ For another way of reconciling price misperception with a binding budget constraint see Gabaix (2014). In Section 5, we present a different model of consumer demand where there is no binding budget constraint and so even a choice under misperceived prices can be implemented exactly.

¹⁰ Misperception restrictions of this type are not unique to our model; indeed, our notion is akin to the ‘rational expectations’ assumption made, for example, in Koszegi and Rabin (2006, 2007).

The second notion of price consistency we consider is to require \mathcal{P} to have a **stable distribution** with respect to the price indices $\{\bar{\mathbf{p}}^t\}_{t \in T}$. This means that there is $\lambda_{i,k}$ such that for any $t \in T$,

$$p_k^{i,t} = \lambda_{i,k} \bar{p}_k^t. \quad (5)$$

For each good, the distribution of heterogenous prices is stable across observations; indeed, if consumer i assigns a price to k that is 10% above \bar{p}_k^t at *some* observation t , then he will do the same at *every* observation. It is clear that a stable price distribution will also be consistent in expectation under the following assumptions: (a) W_k^t does not vary with t and thus can be denoted by W_k ; and (b) W_k is homogeneous of degree 1, with the normalization that $W_k((\lambda_{i,k})_{i \in N}) = 1$ for all k . Obviously, these conditions hold if the price aggregator is the arithmetic average (3) for every good and every observation.

Preference heterogeneity. Of course, the standard way of imposing structure on a data set is to require that all agents face the same prices but allow them to have different preferences. The data set $\mathcal{E} = \{(\mathbf{e}^{i,t}, \bar{\mathbf{p}}^t)\}_{(i,t) \in N \times T}$ is **rationalizable with heterogenous preferences** if there are strictly increasing and continuous utility functions $u^i : \mathbb{R}_+^K \rightarrow \mathbb{R}$ such that for each $(i, t) \in N \times T$,

$$u^i(\mathbf{x}(\mathbf{e}^{i,t}, \bar{\mathbf{p}}^t)) \geq u^i(\mathbf{x}) \text{ for any } \mathbf{x} \in \mathcal{B}(\bar{\mathbf{p}}^t, m^{i,t}). \quad (6)$$

Cross-Sectional Data Environment. As we explore the implications of price heterogeneity in this paper, we shall often compare our results with what is known in the random utility model (RUM).¹¹

The starting point in these models is a data set with repeated cross-sections of the demand distribution at different prices. To keep our discussion focused on the essentials, we shall assume a discrete cross-sectional data set. With minor modifications, our results carry through to the case where the demand distributions are not necessarily discrete; this is considered in detail in the Appendix. Given the discreteness assumption, the data set, consisting of T observations, can be written as

$$\mathcal{D} = \{E^t, \bar{\mathbf{p}}^t\}_{t \in T}. \quad (7)$$

¹¹ Random utility models are studied in Luce (1959), Block and Marshak (1960), McFadden (1973), Falmagne (1978), McFadden and Richter (1990), Kitamura and Stoye (2018), among others.

At each observation, t , the analyst observes the price index $\bar{\mathbf{p}}^t$ as well as N expenditure allocation decisions over K goods, which form the set E^t . It is usual in this model to assume that, for any \mathbf{e} and \mathbf{e}' in E^t , $\sum_{k=1}^K e_k = \sum_{k=1}^K e'_k = m^t$. In other words, the total expenditure is the same for different elements in E^t and their implied demand bundles $\mathbf{x}(\mathbf{e}, \bar{\mathbf{p}}^t)$ and $\mathbf{x}(\mathbf{e}', \bar{\mathbf{p}}^t)$ are both bundles in the budget set $\mathcal{B}(\bar{\mathbf{p}}^t, m^t)$.

We could think of E^t *either* as the allocation decisions of a population with N consumers *or* as N allocation decisions made by the same consumer whose demand behavior is stochastic. We shall refer to a data set of the form \mathcal{D} as a **cross-sectional data set**.

A sorting function at t , σ^t , is a one-to-one map from the set $\{1, 2, \dots, N\}$ to E^t ; in other words, σ^t attaches an index to each expenditure allocation in E^t . Note that $\sigma^t(i)$ is an element of E^t that corresponds to $\mathbf{e}^{i,t}$. We say that a cross-sectional data set \mathcal{D} is **RUM-rationalizable** (or admits a RUM-rationalization) if there are sorting functions $\{\sigma^t\}_{t \in T}$ such that

$$\mathcal{E}(\{\sigma^t\}_{t \in T}) = \{(\sigma^t(i))_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}, \quad (8)$$

(which has the same form as (1)) is rationalizable with heterogenous preferences. In other words, \mathcal{D} is RUM-rationalizable if there is a way of sorting the expenditure allocations in E^t so that the entire data set could be understood as arising from N utility-maximizing consumers, all of whom face prices $\bar{\mathbf{p}}^t$ at each observation t , but have possibly heterogenous preferences. We can compare this notion of rationalization with that under the **random price model** (RPM). The cross-sectional data set \mathcal{D} is **RPM-rationalizable** if there are sorting functions $\{\sigma^t\}_{t \in T}$ such that $\mathcal{E}(\{\sigma^t\}_{t \in T})$ (as defined by (8)) is rationalizable by heterogenous prices, subject to some consistency condition on the heterogenous prices.

3 Price Heterogeneity with Consistent Expectations

In this section, we show that price heterogeneity, even when it is required to be consistent in expectation, is powerful enough to explain almost any data set. Our first result states this claim formally.

PROPOSITION 1. *Let $\mathcal{W} = \{W_k^t\}_{k \in K, t \in T}$ be a collection of aggregator functions and $\mathcal{E} = \{(\mathbf{e}^{i,t})_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}$ be a data set, with $e_k^{i,t} > 0$ for all $(i, t) \in N \times T$ and $k = 1, 2$. Then there*

are heterogenous prices $\mathcal{P} = \{\mathbf{p}^{i,t}\}_{(i,t) \in N \times T}$, which are consistent in expectation w.r.t \mathcal{W} and satisfy $p_k^{i,t} = \bar{p}_k^t$ for all $k > 2$, that rationalize \mathcal{E} .

In empirical studies of demand, heterogeneity in preferences is often employed as a way of capturing heterogeneity in observed behavior. This proposition highlights the possibility of an alternative way to explain demand heterogeneity. It states that if there are two goods for which demand is strictly positive at all observations and for all agents, then price heterogeneity in *just those two goods* can rationalize any pattern of expenditure, even if consistency in expectation is imposed. In the Errors interpretation of heterogenous prices, the heterogeneity in demand arises from heterogeneity in prices that the analyst could not observe because of imperfections in the data collected. In the Misperceptions interpretation of heterogenous prices, the heterogeneity in demand arises from mistakes in consumers' decision-making. Either way, heterogenous prices could explain heterogenous demand behavior. Of course, whether such an explanation is a plausible one would depend on the specific case being considered – we are not suggesting that it is always and everywhere a better explanation, only that it is a potentially powerful tool for explaining demand heterogeneity.

In the context of a cross-sectional data set $\mathcal{D} = \{E^t, \bar{\mathbf{p}}^t\}_{t \in T}$, it is well-known that RUM imposes observable restrictions on \mathcal{D} (see the example in Section 4). On the other hand, Proposition 1 guarantees that *any* data set \mathcal{D} with the property that $e_k^t > 0$ for $k = 1, 2$ and all $\mathbf{e}^t \in E^t$ is RPM-rationalizable by heterogenous prices with consistent price expectations.¹² This is good news if the objective is to have a fully flexible model that could capture any set of observed demand behavior on \mathcal{D} . Of course, any explanation of the data with such a method will have implications on the unobserved heterogeneity in prices, and it would then be a matter of (context-dependent) judgment whether the unobserved heterogeneity required to explain the data is, in fact, reasonable. Another approach is to impose stronger restrictions on the types of price heterogeneity permitted, which will temper the explanatory power of the heterogenous-prices model; this approach is explored in Section 4.

¹² Indeed, with *any* set of sorting functions, $\{\sigma^t\}_{t \in T}$, Proposition 1 guarantees that $\mathcal{E}(\{\sigma^t\}_{t \in T}) = \{(\sigma^t(i))_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}$ is rationalizable with heterogenous prices.

3.1 Proof of Proposition 1 and related results

The proof of Proposition 1, as well as other results in this paper, relies on Afriat's Theorem (Afriat (1967), Varian (1982)). The theorem considers a data set with T observations, where observation t consists of a bundle $\mathbf{x}^t \in \mathbb{R}_+^K$ purchased when the prices are $\mathbf{p}^t \in \mathbb{R}_{++}^K$. Thus the data set may be written as $\mathcal{O} = \{(\mathbf{x}^t, \mathbf{p}^t)\}_{t \in T}$. A strictly increasing and continuous utility function $u : \mathbb{R}_+^K \rightarrow \mathbb{R}$ is said to rationalize \mathcal{O} if, for each $t \in T$,

$$u(\mathbf{x}^t) \geq u(\mathbf{x}) \text{ for all } \mathbf{x} \in \mathcal{B}(\mathbf{p}^t, \mathbf{p}^t \cdot \mathbf{x}^t).$$

Afriat's Theorem characterizes rationalizable data sets through a property called the **generalized axiom of revealed preference**, or **GARP** for short.¹³

GARP is basically a type of no-cycling condition on $\mathcal{X} = \{\mathbf{x}^t\}_{t \in T}$, the set of observed demand bundles. Let $\mathbf{x}^t, \mathbf{x}^s$ be two elements in \mathcal{X} . We say \mathbf{x}^t is **directly revealed preferred to** \mathbf{x}^s , denoted by $\mathbf{x}^t \succsim_D \mathbf{x}^s$, if $\mathbf{p}^t \cdot \mathbf{x}^t \geq \mathbf{p}^t \cdot \mathbf{x}^s$. Similarly, \mathbf{x}^t is **directly revealed strictly preferred to** \mathbf{x}^s , denoted by $\mathbf{x}^t \succ_D \mathbf{x}^s$, if $\mathbf{p}^t \cdot \mathbf{x}^t > \mathbf{p}^t \cdot \mathbf{x}^s$. We say \mathbf{x}^t is **revealed preferred to** \mathbf{x}^s , denoted by $\mathbf{x}^t \succsim_R \mathbf{x}^s$, if there is a sequence $\{\mathbf{x}^{t_l}\}_{l=1}^L$ such that $\mathbf{x}^{t_1} = \mathbf{x}^t$, $\mathbf{x}^{t_L} = \mathbf{x}^s$, and $\mathbf{x}^{t_l} \succsim_D \mathbf{x}^{t_{l+1}}$ for each $l < L$. GARP requires that for any $s, t \in T$, $\mathbf{x}^t \succsim_R \mathbf{x}^s$ implies $\mathbf{x}^s \not\succ_D \mathbf{x}^t$.

In essence, the proof of Proposition 1 consists of constructing heterogenous prices for goods 1 and 2 that are consistent in expectation and have the property that the collection of observations

$$\mathcal{O}^* = \{(\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}), \mathbf{p}^{i,t})\}_{(i,t) \in N \times T} \quad (9)$$

obeys GARP. This is a notional data set with $N \times T$ observations, where each observation consists of a bundle and a price vector, and is precisely of the type where one may apply Afriat's Theorem. The theorem guarantees that there is a utility function that rationalizes this \mathcal{O}^* (equivalently, rationalizes \mathcal{E} with heterogenous prices) if \mathcal{O}^* satisfies GARP. In our proof, we choose the idiosyncratic prices $p_k^{i,t}$, for $k = 1, 2$, carefully so that the direct revealed preference relation \succsim_D on $\mathcal{X}^* = \{\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t})\}_{(i,t) \in N \times T}$ has the following property:

$$\text{if } \mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}) \succsim_D \mathbf{x}(\mathbf{e}^{j,s}, \mathbf{p}^{j,s}), \text{ then either } i > j \text{ or } i = j \text{ and } t \geq s. \quad (10)$$

¹³ To be precise, \mathcal{O} satisfies GARP if it can be rationalized by a locally nonsatiated preference. Conversely, when \mathcal{O} obeys GARP, then it is rationalizable by a strictly increasing, continuous, and concave utility function.

In other words, if $\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t})$ is directly revealed preferred to $\mathbf{x}(\mathbf{e}^{j,s}, \mathbf{p}^{j,s})$, then (i, t) dominates (j, s) in the lexicographic order on $N \times T$.¹⁴ This in turn guarantees that if $\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}) \succeq_R \mathbf{x}(\mathbf{e}^{j,s}, \mathbf{p}^{j,s})$, then either $i > j$ or $i = j$ and $t \geq s$. Thus, it is impossible for $\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}) \succeq_R \mathbf{x}(\mathbf{e}^{j,s}, \mathbf{p}^{j,s})$ and for $\mathbf{x}(\mathbf{e}^{j,s}, \mathbf{p}^{j,s}) >_D \mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t})$, which means that GARP holds. The details of the proof can be found in the Appendix.

Relationship to Varian (1988). This paper shows that GARP becomes a vacuous condition in certain circumstances when prices or demand are only partially observed. To be specific, suppose that at each observation t , the prices of all K goods, \mathbf{p}^t , are observed, but only the consumer's demand for goods $2, 3, \dots, K$, which we denote by $\mathbf{x}_{-1}^t = (x_2^t, \dots, x_K^t)$, are observed. In this case, there would always exist $\hat{x}_1^t \in \mathbb{R}_+$ (a hypothetical demand for good 1 at observation t) such that $\{(\mathbf{x}^t, \mathbf{p}^t)\}_{t \in T}$ obeys GARP, where $\mathbf{x}^t = (\hat{x}_1^t, \mathbf{x}_{-1}^t)$. In other words, when the demand of one good is missing from the data, rationalizability can never be rejected. The proof provided by Varian involves picking a sequence x_1^t (for $t = 1, 2, \dots$) in such a way that it is impossible for \mathbf{x}^t to be revealed preferred to $\mathbf{x}^{t'}$ if $t < t'$.

While not explicitly considered in Varian (1988), it is clear that Varian's result has a variation of the following form (which is also clear from our proof of Proposition 1): *suppose each observation t consists of the prices and demand for goods $2, 3, \dots, K$ (denoted by \mathbf{p}_{-1}^t and \mathbf{x}_{-1}^t respectively) as well as the expenditure on good 1, e_1^t , where $e_1^t > 0$ for all t ; then there is \hat{p}_1^t and \hat{x}_1^t such that $\hat{p}_1^t \hat{x}_1^t = e_1^t$ and, with $\mathbf{x}^t = (\hat{x}_1^t, \mathbf{x}_{-1}^t)$ and $\mathbf{p}^t = (\hat{p}_1^t, \mathbf{p}_{-1}^t)$, the data set $\{(\mathbf{x}^t, \mathbf{p}^t)\}_{t \in T}$ obeys GARP (and is thus rationalizable).*

Compared to this result, the setup of Proposition 1 is more complicated because $\mathcal{E} = \{(\mathbf{e}^{i,t})_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}$ collects observations from multiple agents and the average price of each good (across all agents) is observed. Thus, while for a *given* agent i we are free to 'break up' the expenditure on good 1, $e_1^{i,t}$, in any way between the price and demand for good 1, a higher idiosyncratic price of good 1 for one agent has to be balanced by a lower price of good 1 for another agent. With this added consistency requirement, the rationalizability of \mathcal{E} can only be guaranteed if there is price heterogeneity in at least two goods. In the Appendix, we provide an example of a data set \mathcal{E} that cannot be rationalized with idiosyncratic prices in only one good, even if agents are permitted to have distinct preferences.

¹⁴ The lexicographic order ranks (i, t) higher than (j, s) if either $i > j$ or $i = j$ and $t > s$.

Relationship to Brown and Matzkin (1996). Brown and Matzkin (1996) show that the Walrasian model of an exchange economy is refutable when the economy's total endowment, the equilibrium prices, and the equilibrium income distribution are observable. Their result could be re-formulated (as we do below) as a result on the dis-aggregation of demand.

The data set consists of T observations with each observation t consisting of the following: *aggregate* expenditure (across all consumers) on each good k , $E_k^t > 0$; the aggregate expenditure of each consumer, $m^{i,t} > 0$; and the prevailing price vector $\bar{\mathbf{p}}^t \gg 0$. Formally,

$$\mathcal{E}^{**} = \{(E_k^t)_{k \in K}, (m^{i,t})_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}.$$

(For consistency, we require $\sum_{k \in K} E_k^t = \sum_{i \in N} m^{i,t}$ at each t ; in other words, total expenditure on all goods equals total expenditure of all consumers, at each t .) We refer to $(e_k^{i,t})_{(i,k,t) \in N \times K \times T} \geq 0$ as a disaggregation of $(E_k^t)_{(k,t) \in K \times T}$ if $\sum_{i \in N} e_k^{i,t} = E_k^t$ for each (t, k) and $\sum_{k \in K} e_k^{i,t} = m^{i,t}$ for each (i, t) . We say that \mathcal{E}^{**} admits a **disaggregation with heterogenous preferences** if there is a disaggregation $(e_k^{i,t})_{(i,k,t) \in N \times K \times T}$ such that $\mathcal{E} = \{(\mathbf{e}^{i,t})_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}$ is rationalizable with heterogenous preferences; similarly, \mathcal{E}^{**} admits a **disaggregation with heterogenous prices** if there is a disaggregation such that \mathcal{E} is rationalizable with heterogenous prices.

One may be tempted to think that *every* \mathcal{E}^{**} admits a disaggregation with heterogenous preferences, but Brown and Matzkin (1996) show that there exists \mathcal{E}^{**} that do not admit such a disaggregation; they also set out a procedure that could test if such a disaggregation exists. This result is sensitive to the assumption of uniform pricing across consumers and price heterogeneity on any two goods will guarantee that ‘anything goes.’

COROLLARY 1. *Let $\mathcal{W} = \{W_k^t\}_{k \in K, t \in T}$ be a collection of aggregator functions. For every data set $\mathcal{E}^{**} = \{(E_k^t)_{k \in K}, (m^{i,t})_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}$ there is a disaggregation $(e_k^{i,t})_{(i,k,t) \in N \times K \times T}$ with the following property: for any $k', k'' \in K$, there are heterogenous prices $\mathcal{P} = \{\mathbf{p}^{i,t}\}_{(i,t) \in N \times T}$ that are consistent in expectation w.r.t \mathcal{W} , satisfy $p_k^{i,t} = \bar{p}_k^t$ for all $k \neq k', k''$, and rationalize $\mathcal{E} = \{(\mathbf{e}^{i,t})_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}$.*

This proposition follows immediately from Proposition 1, by simply choosing the disaggregation $e_k^{i,t} = [m^{i,t} / (\sum_{i \in N} m^{i,t})] E_k^t > 0$. In the next section, we show that stronger restrictions on price heterogeneity will diminish the explanatory power of price heterogeneity and allow us to restore the Brown-Matzkin result.

4 Stable Price Heterogeneity and Preference Heterogeneity

In the previous section, we analyzed price heterogeneity with consistent expectations and found that it gives rise to a very flexible model that imposes no restrictions on observed demand. In this section, we explore the implications of stable price heterogeneity. In Section 4.1 we explain why stable price heterogeneity does not enlarge the observable implications of a model with heterogenous preferences and homogenous prices. Section 4.2 contains the main result of the paper: we show that stable price heterogeneity has precisely the same explanatory power as heterogenous preferences.

4.1 Preference heterogeneity when prices are heterogenous

In the result below, we ask what happens when there is stable price heterogeneity (i.e., $p_k^{i,t} = \lambda_{i,k} \bar{p}_k^t$) in a market where consumers have heterogenous preferences.

PROPOSITION 2. *For any stable price distribution $\mathcal{P} = \{\mathbf{p}^{i,t}\}_{(i,t) \in N \times T}$, the data set $\mathcal{E} = \{(\mathbf{e}^{i,t})_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}$ is rationalizable with heterogenous preferences if and only if, for each $i \in N$, $\{(\mathbf{e}^{i,t}, \mathbf{p}^{i,t})\}_{t \in T}$ is rationalizable; i.e., there is a strictly increasing and continuous utility function $u^i : \mathbb{R}_+^K \rightarrow \mathbb{R}$ such that for each $t \in T$,*

$$u^i(\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t})) \geq u^i(\mathbf{x}) \text{ for any } \mathbf{x} \in \mathcal{B}(\mathbf{p}^{i,t}, m^{i,t}).$$

The proof of this proposition uses Afriat's Theorem and can be found in the Appendix. To interpret this result, we first consider the case where prices vary across consumers (*in fact*) with $\bar{\mathbf{p}}^t$ being the average price vector at observation t . Then Proposition 2 tells us that, provided each agent's demand is rationalizable and price heterogeneity is stable, the average price $\bar{\mathbf{p}}^t$ is a good substitute for the true price $\mathbf{p}^{i,t}$, in the sense that it does not destroy the rationalizability of the data. In the second interpretation, where $\mathbf{p}^{i,t}$ is agent i 's perceived price, this proposition tells us that stable price misperception is undetectable: the agent's behavior is indistinguishable from someone who perceives prices correctly and is maximizing utility.

Cross-Sectional Data Environment. Proposition 2 has the following immediate corollary for cross-sectional data sets: *for any stable price distribution $\mathcal{P} = \{\mathbf{p}^{i,t}\}_{(i,t) \in N \times T}$, the cross*

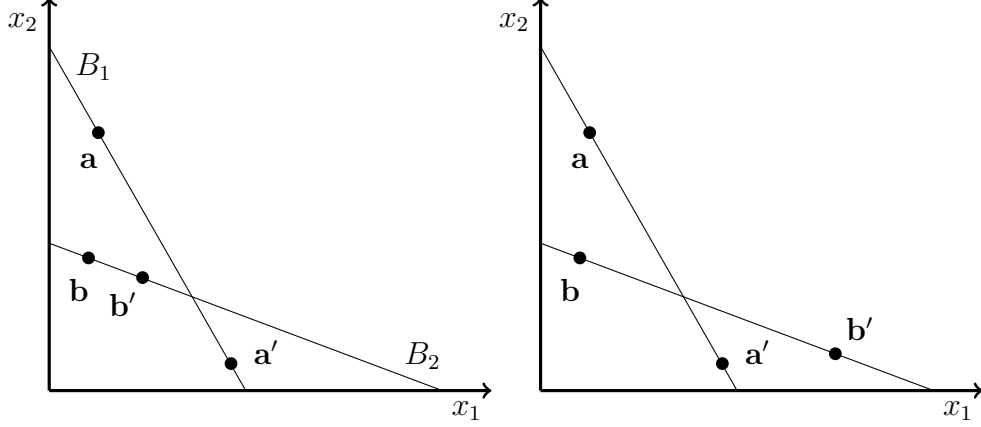


Figure 1: Observable restrictions in the Random Utility Model

The cross-sectional data set depicted on the right, with $E^1 = \{\mathbf{a}, \mathbf{a}'\}$ and $E^2 = \{\mathbf{b}, \mathbf{b}'\}$ is RUM-rationalizable since one could pair up \mathbf{a} with \mathbf{b} and \mathbf{a}' with \mathbf{b}' and each of the agent-level data sets obeys GARP. On the other hand, the data set on the left is not RUM-rationalizable because it is impossible to sort the elements of E^1 and E^2 in such a way that both agent-level data sets obey GARP.

sectional data set $\mathcal{D} = \{E^t, \bar{\mathbf{p}}^t\}_{t \in T}$ is RUM-rationalizable if and only if there are sorting functions $\{\sigma^t\}_{t \in T}$ and strictly increasing and continuous utility functions $u^i : \mathbb{R}_+^K \rightarrow \mathbb{R}$ such that for each $i \in N$ and $t \in T$,

$$u^i(\mathbf{x}(\sigma^t(i), \mathbf{p}^{i,t})) \geq u^i(\mathbf{x}) \text{ for any } \mathbf{x} \in \mathcal{B}(\mathbf{p}^{i,t}, m^t). \quad (11)$$

In other words, RUM-rationalizability is not destroyed by the insertion of stable price heterogeneity: if \mathcal{D} is a cross-sectional data set collected from a population of consumers who maximize (heterogenous) increasing and continuous utility functions under prices with stable heterogeneity, then \mathcal{D} will be RUM-rationalizable. It follows that the observable implications of the random utility model (such as that depicted in Figure 1) remain precisely the same even if there is stable price heterogeneity among agents in the population.

Relationship to Brown and Matzkin (1996) (continued). In Corollary 1, we show that *every* dataset $\mathcal{E}^{**} = \{(E_k^t)_{k \in K}, (m^{i,t})_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}$ admits a decomposition with heterogeneous prices which are consistent in expectation. This ‘anything goes’ result no longer holds if we confine ourselves to stable price heterogeneity. Indeed, it follows immediately from Proposition 2 that, for any stable price distribution $\mathcal{P} = \{\mathbf{p}^{i,t}\}_{(i,t) \in N \times T}$, the dataset \mathcal{E}^{**} admits a disaggregation $\mathcal{E} = \{(\mathbf{e}^{i,t})_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}$ that is rationalizable with heterogeneous preferences if and only if, for each $i \in N$, $\{(\mathbf{e}^{i,t}, \mathbf{p}^{i,t})\}_{t \in T}$ is rationalizable (by a strictly increasing

and continuous utility function). In particular, the observable restrictions on \mathcal{E}^{**} imposed by heterogenous preferences and homogenous prices are *exactly the same* as the observable restrictions imposed by heterogenous prices and stable price heterogeneity. Thus the Brown-Matzkin result is robust to the presence of stable price heterogeneity.

4.2 Stable Price Heterogeneity and Equivalence Scale Utilities

Our objective in this section is to establish a threefold equivalence. Our first equivalence is to show that if $\mathcal{E} = \{(\mathbf{e}^{i,t})_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}$ is rationalizable with heterogenous preferences, then it suffices to choose these preferences from a semiparametric family of utility functions where (in a specific sense) any two utility functions in the family are related by a linear transformation. The second equivalence is that \mathcal{E} is rationalizable with heterogenous preferences if and only if it is rationalizable with heterogenous prices with a stable distribution.

Let R be a nonempty subset of K . For a utility function $U : \mathbb{R}_+^K \rightarrow \mathbb{R}$ and $\beta \in \mathbb{R}_{++}$, let

$$U_R^\beta(\mathbf{x}) \equiv U(\beta \mathbf{x}_R, \mathbf{x}_{-R}) \quad \text{for any } \mathbf{x} \in \mathbb{R}_+^K.$$

We refer to U_R^β as a *R-scale transformation* of U . A collection \mathcal{U}_R of utility functions form an **R-scale family** if there is a strictly increasing and concave utility function U such that $\mathcal{U}_R = \{U_R^\beta : \beta \in \mathbb{R}_{++}\}$. Scale transformations first appeared in the econometric literature on demand estimation, where it is used to capture (among other things) the effects of household composition on demand (see, for example, Barten (1964), Prais and Houthakker (1971), and Muellbauer (1980)). They have also been used in the study of demand aggregation, where they help to obtain properties on *market* demand which are not necessarily present at the agent-level demand correspondence (see Mas-Colell and Neufeind (1977), Dierker et al. (1984), Grandmont (1987, 1992), and Quah (1997)).

A basic reason for the use of scale transformations is that there is a convenient relationship between the demand generated by U and that generated by U_R^β . Let $\mathbf{x}(\mathbf{p}, m)$ and $\mathbf{x}^\beta(\mathbf{p}, m)$ be their demand sets at price \mathbf{p} and income/expenditure m , i.e., $\mathbf{x}(\mathbf{p}, m) = \text{argmax}\{U(\mathbf{x}) | \mathbf{x} \in \mathcal{B}(\mathbf{p}, m)\}$ and $\mathbf{x}^\beta(\mathbf{p}, m) = \text{argmax}\{U_R^\beta(\mathbf{x}) | \mathbf{x} \in \mathcal{B}(\mathbf{p}, m)\}$. Then it is straightforward to check that \mathbf{x} and \mathbf{x}^β are related in the following way:

$$\mathbf{x}^\beta(\mathbf{p}, m) = \left(\frac{\mathbf{x}_R(\mathbf{p}, \beta m)}{\beta}, \mathbf{x}_{-R}(\mathbf{p}, \beta m) \right).$$

Note that in the case where $R = K$, we have $\mathbf{x}^\beta(\mathbf{p}, m) = \mathbf{x}(\mathbf{p}, \beta m)/\beta$.¹⁵

We have explained in Section 2 what it means for a data set $\mathcal{E} = \{(\mathbf{e}^{i,t})_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}$ to be rationalizable with heterogenous preferences. We say that \mathcal{E} is rationalizable with heterogenous preferences **drawn from an R -scale family** if the preferences needed to explain \mathcal{E} can all be induced by utility functions drawn from an R -scale family; more formally, there are real numbers $\beta_1, \dots, \beta_N \in \mathbb{R}_{++}$ and a strictly increasing and concave utility function U such that for each $(i, t) \in N \times T$,

$$U_R^{\beta_i}(\mathbf{x}(\mathbf{e}^{i,t}, \bar{\mathbf{p}}^t)) \geq U_R^{\beta_i}(\mathbf{x}) \text{ for any } \mathbf{x} \in \mathcal{B}(\bar{\mathbf{p}}^t, m^{i,t}). \quad (12)$$

The next result states the threefold equivalence between three distinct notions of rationalizability.

THEOREM 1. *Let $\mathcal{E} = \{(\mathbf{e}^{i,t})_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}$ be a data set and suppose that for some $R \subseteq K$, we have $\mathbf{e}_R^{i,t} > 0$ for each $(i, t) \in N \times T$. Then the following are equivalent.*

- (i) \mathcal{E} is rationalizable with heterogenous preferences.
- (ii) \mathcal{E} is rationalizable with heterogenous preferences drawn from an R -scale family.
- (iii) Let W be any linear homogeneous aggregator function. Then there are real numbers $\lambda_1, \dots, \lambda_N > 0$ such that the stable price distribution $\mathcal{P} = \{\mathbf{p}^{i,t}\}_{(i,t) \in N \times T}$ with $\mathbf{p}^{i,t} = (\lambda_i \bar{\mathbf{p}}_R^t, \bar{\mathbf{p}}_{-R}^t)$ is consistent in expectation w.r.t. W and rationalizes \mathcal{E} .

Note: Since our maintained assumption is that \mathcal{E} satisfies $\sum_{k \in K} e_k^{i,t} > 0$ for all (i, t) , we could always choose $R = K$, but the theorem allows for R to be strictly smaller. If (say) good 1 satisfies $e_1^{i,t} > 0$ for all (i, t) , then we could choose $R = \{1\}$. Statement (ii) then refers to rationalizability with preference heterogeneity obtained by scaling good 1, while statement (iii) refers to rationalizability with heterogeneity in just the price of good 1; each of these statements is equivalent to rationalizability with heterogenous preferences.

In the literature on demand aggregation, it is fairly common to assume that all agents in a market (or some segment of the market) have preferences that belong to an R -scale family

¹⁵ Moreover, when $R = K$, the Independence of a Base Utility (IB) property in Lewbel (1989) (also called *Equivalence Scale Exactness*) is satisfied. The IB property is commonly assumed in empirical demand work to justify the use of the cost of living indices.

(see, for example, Grandmont (1987, 1992) and Quah (1997)). The equivalence of statements (i) and (ii) in this proposition tells us that this modeling practice is (in essence) without loss of generality since any data set \mathcal{E} that is rationalizable with heterogenous preferences can also be rationalized by preferences drawn from an R -scale family. Of course, in that literature, further conditions are imposed on the distribution of the scales (that is, the distribution of β in our notation) and it is these further distributional assumptions that drive the strong conclusions on market demand obtained in that literature.¹⁶

Recall that Proposition 1 tells us that *any* data set \mathcal{E} can be explained by heterogenous prices, even when this heterogeneity is required to satisfy a consistency condition. The equivalence of (i) and (iii) in Theorem 1 tells us that the imposition of stable heterogeneity on prices disciplines the implications of heterogenous prices; in fact, the explanatory power of such a model (in the sense of the data sets that it could explain) coincides precisely with that of a model with heterogenous preferences. In the case where there is a good – say, good 1 – where the expenditure is strictly positive for all agents and at all observations, then (stable) price heterogeneity for that good alone is sufficient to capture the effects of random preferences. Lastly, note that the combination of Proposition 2 and Theorem 1 means that even when preference heterogeneity and stable price heterogeneity are *both* present in a given environment, the data it generates is equivalent to data from an environment where *only* preference heterogeneity or price heterogeneity is present.

The next result is an obvious corollary of Theorem 1 obtained by transposing the theorem to the setting of a cross-sectional data set. The punchline is that *RUM-rationalizability is equivalent to RPM-rationalizability with stable heterogenous prices*. In the Appendix, we provide a more general formulation of this equivalence in which the cross-sectional data set is allowed to be non-discrete (as in McFadden and Richter (1990) or Kitamura and Stoye (2018)).

COROLLARY 2. *Let $\mathcal{D} = \{E^t, \bar{\mathbf{p}}^t\}_{t \in T}$ be a cross-sectional data set and suppose that for some $R \subseteq K$, we have $\mathbf{e}_R^{i,t} > 0$ for each $(i, t) \in N \times T$. Then, the following are equivalent.*

- (i) \mathcal{D} is RUM-rationalizable.

¹⁶ For example, sufficient dispersion in the distribution of β leads to the approximate linearity of the market demand function with respect to total expenditure/income (see Grandmont (1992) and Quah (1997)).

- (ii) \mathcal{D} is RUM-rationalizable with heterogenous preferences drawn from an R -scale family.
- (iii) Let W be any linear homogeneous aggregator function. Then \mathcal{D} is RPM-rationalizable with a price distribution \mathcal{P} that is stable and consistent in expectation w.r.t. W , with heterogenous prices only for the goods in R .

Proof sketch of Theorem 1. The detailed proof is found in the Appendix, but the following quick sketch may be helpful.

Obviously, (ii) implies (i). To show that (iii) implies (ii), let $R \subseteq K$ with $e_R^{i,t} > 0$ for each $(i, t) \in N \times T$. Then there are real numbers $\lambda_1, \dots, \lambda_N > 0$ and an increasing and continuous utility function U such that, with $\mathbf{p}^{i,t} = (\lambda_i \bar{\mathbf{p}}_R^t, \bar{\mathbf{p}}_{-R}^t)$,

$$U(\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t})) \geq U(\mathbf{x}) \text{ for any } \mathbf{x} \in \mathcal{B}(\mathbf{p}^{i,t}, m^{i,t}) \quad (13)$$

for all (i, t) . We show that this implies that

$$U_R^{\lambda_i}(\mathbf{x}(\mathbf{e}^{i,t}, \bar{\mathbf{p}}^t)) \geq U_R^{\lambda_i}(\tilde{\mathbf{x}}) \text{ for any } \tilde{\mathbf{x}} \in \mathcal{B}(\bar{\mathbf{p}}^t, m^{i,t})$$

and thus \mathcal{E} is rationalizable with heterogenous preferences drawn from an R -scale family. It remains to show that (i) implies (iii). Suppose \mathcal{E} is rationalizable with heterogenous preferences. We need to find real numbers $\lambda_1, \dots, \lambda_N > 0$ such that

$$\mathcal{O}^* = \{(\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}), \mathbf{p}^{i,t})\}_{(i,t) \in N \times T} \text{ where } \mathbf{p}^{i,t} = (\lambda_i \bar{\mathbf{p}}_R^t, \bar{\mathbf{p}}_{-R}^t)$$

obeys GARP. Obviously, $\mathcal{O}^* = \bigcup_{i \in N} \mathcal{O}^i$, where $\mathcal{O}^i = \{(\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}), \mathbf{p}^{i,t})\}_{t \in T}$. Note that \mathcal{O}^i obeys GARP for any $\lambda_i > 0$; this follows from Proposition 2 and the fact that \mathcal{E} is rationalizable with heterogenous preferences. We then show that it is possible to choose λ_{i+1} sufficiently smaller than λ_i so that, for any two consumers i and j with $j > i$, their budget sets are sufficiently distinct so that no bundle chosen by i is revealed preferred to a bundle chosen by j . We could then conclude that \mathcal{O}^* obeys GARP. \square

5 Expenditure-Augmented Utility and Stable Price Heterogeneity

Let $\mathcal{O} = \{(\mathbf{x}^t, \mathbf{p}^t)\}_{t \in T}$ be the data set collected from a single consumer, where \mathbf{x}^t is the bundle purchased at the price \mathbf{p}^t . In Section 3.1 we recounted Afriat's Theorem, which provides necessary and sufficient conditions for \mathcal{O} to be rationalizable, in the sense that there is a

utility function such that the chosen bundle \mathbf{x}^t is weakly preferred to alternative bundles in the budget set $\mathcal{B}(\mathbf{p}^t, \mathbf{p}^t \cdot \mathbf{x}^t)$. Afriat's Theorem addresses the problem of rationalizability in the context of the constrained-optimization model of consumer demand, but there are other ways in which consumer demand can be modeled.

In partial equilibrium settings, where the set of goods being studied is just a subset of all the goods consumed by a consumer, it is common to model demand with a utility function of the quasilinear form; the consumer's choice is then the bundle $\mathbf{x} \in \mathbb{R}_+^K$ that maximizes $\bar{U}(\mathbf{x}) - \mathbf{p} \cdot \mathbf{x}$, when \mathbf{p} is the price vector for the K goods. More generally, we can model the consumer's choice as arising from the unconstrained maximization of a utility function where expenditure enters as an argument of the utility function. Formally, the utility of acquiring the bundle \mathbf{x} at cost $e \geq 0$ is $V(\mathbf{x}, -e)$ where $V : \mathbb{R}_+^K \times \mathbb{R}_- \rightarrow \mathbb{R}$ is required to be strictly increasing in the last argument (so that increasing expenditure strictly reduces utility). Quasilinear utility-maximization would then correspond to the special case where $V(\mathbf{x}, -e) = \bar{U}(\mathbf{x}) - e$. We refer to V as an **expenditure-augmented utility function** or simply an augmented utility function.

Consumer demand arising from the maximization of such a utility function is studied in Deb et al. (2023). This study provides necessary and sufficient conditions under which a data set \mathcal{O} is augmented utility-rationalizable, in the sense that there is an augmented utility function V for which

$$V(\mathbf{x}^t, -\mathbf{p}^t \cdot \mathbf{x}^t) \geq V(\mathbf{x}, -\mathbf{p}^t \cdot \mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathbb{R}_+^K.^{17}$$

In this section, we study the implications of price heterogeneity in the augmented utility model (AUM). We focus on two notions of price heterogeneity, both of which are inspired by the behavioral economics literature. The first notion is price heterogeneity arising from **inattention to prices** among consumers, leading them to optimize according to idiosyncratic prices that differ from the observed price. While we have already introduced this idea in the previous sections, one could argue that the augmented utility model is a particularly convenient setting in which to study this phenomenon: since this model is not a constrained-optimization model, we can assume that a consumer who has been inattentive

¹⁷The characterization of those data sets \mathcal{O} that are quasilinear-rationalizable, in the sense that there is \bar{U} such that $\bar{U}(\mathbf{x}^t) - \mathbf{p}^t \cdot \mathbf{x}^t \geq \bar{U}(\mathbf{x}) - \mathbf{p}^t \cdot \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}_+^K$, can be found in Brown and Calsamiglia (2007).

to prices and who has chosen a bundle based on a wrong impression of prices will simply *purchase the bundle he intended*; in contrast, in a constrained-optimization model, such an assumption cannot be made because a consumer who chooses a bundle based on the wrong prices could violate the budget constraint and thus a reconciliation procedure must be added to the model.

For example, imagine a consumer i who on a shopping trip to a grocery store chooses a bundle that maximizes an augmented utility function. However, consumer i is not fully attentive to prices and so chooses the bundle \mathbf{x} that maximizes $V(\mathbf{x}, -\mathbf{p}^i \cdot \mathbf{x})$, where \mathbf{p}^i is the consumer's impression of prices, which may differ from the true prices \mathbf{p} . We assume that if $\tilde{\mathbf{x}}$ is the outcome of this maximization problem, then $\tilde{\mathbf{x}}$ is the bundle bought. The consumer may realize at some point that he has made a mistake and total expenditure is different from what he expected, but since groceries form just a subset of the goods purchased by this consumer, it is plausible to assume that he could adjust for this mistake when purchasing other goods in the future.

The second notion of price heterogeneity we have in mind arises from **reference price dependence**, of the type investigated in Koszegi and Rabin (2006) and Heidhues and Kőszegi (2008). In this case, the consumer observes prices perfectly but has some expectation (or a reference) of what prices should be, and his demand is affected by the extent to which the actual prices he encounters differ from his reference. Heterogenous reference prices are then a source of heterogeneity in demand.

5.1 AU-rationalizability with heterogenous prices

We now explain how we can formally incorporate both price inattention and reference price dependence into the augmented utility model. We assume that a consumer i has in mind a price vector \mathbf{p}^i , which may be different from the true price vector \mathbf{p} . We interpret \mathbf{p}^i as either what the consumer thinks are the prevailing prices (which could be wrong because he is inattentive to prices) or as the consumer's reference prices. Let $e = \mathbf{p} \cdot \mathbf{x}$ and $e^i = \mathbf{p}^i \cdot \mathbf{x}$. We assume that the consumer behaves as though he attributes to the bundle \mathbf{x} an expenditure level that depends on e and e^i ; formally, the expenditure associated with the bundle \mathbf{x} is $\phi(\mathbf{p} \cdot \mathbf{x}, \mathbf{p}^i \cdot \mathbf{x})$, where $\phi : \mathbb{R}_{++}^2 \rightarrow \mathbb{R}_+$ has as its arguments e and e^i . We refer to ϕ as the

behavioral expenditure function. The consumer i chooses \mathbf{x} to maximize

$$V(\mathbf{x}, -\phi(\mathbf{p} \cdot \mathbf{x}, \mathbf{p}^i \cdot \mathbf{x})),$$

where V is the augmented utility function.

For example, if we are modeling consumers who are inattentive to prices, we let $\phi(e, e^i) = e^i$, so the consumer chooses \mathbf{x} to maximize $V(\mathbf{x}, -\mathbf{p}^i \cdot \mathbf{x})$. If we are modeling price reference dependence, we could let $\phi(e, e^i) = ef(e/e^i)$, where $f(1) = 1$. The effective expenditure attributed a bundle \mathbf{x} is then $\mathbf{p} \cdot \mathbf{x} f((\mathbf{p} \cdot \mathbf{x})/(\mathbf{p}^i \cdot \mathbf{x}))$. If f is increasing, then the consumer penalizes a bundle \mathbf{x} when its actual cost $\mathbf{p} \cdot \mathbf{x}$ exceeds the benchmark cost $\mathbf{p}^i \cdot \mathbf{x}$ in the sense that the consumer attributes to it an expenditure $\mathbf{p} \cdot \mathbf{x} f((\mathbf{p} \cdot \mathbf{x})/(\mathbf{p}^i \cdot \mathbf{x}))$ that is greater than the actual expenditure $\mathbf{p} \cdot \mathbf{x}$; similarly, a bundle that costs less than its benchmark is (psychologically) mapped to a cost that is even lower than the actual cost.

For a given ϕ , we say that a collection of heterogenous prices $\mathcal{P} = \{\mathbf{p}^{i,t}\}_{(i,t) \in N \times T}$ **AU-rationalizes** $\mathcal{E} = \{(\mathbf{e}^{i,t})_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}$ if there is an augmented utility function V such that for each $(i, t) \in N \times T$,

$$V(\mathbf{x}^{i,t}, -\phi(\bar{\mathbf{p}}^t \cdot \mathbf{x}^{i,t}, \mathbf{p}^{i,t} \cdot \mathbf{x}^{i,t})) \geq V(\mathbf{x}, -\phi(\bar{\mathbf{p}}^t \cdot \mathbf{x}, \mathbf{p}^{i,t} \cdot \mathbf{x})) \text{ for all } \mathbf{x} \in \mathbb{R}_+^K,$$

where $\mathbf{x}^{i,t} = \mathbf{x}(\mathbf{e}^{i,t}, \bar{\mathbf{p}}^t)$. The main result of this section establishes the equivalence between this notion of rationalizability and **AU-rationalizability with heterogenous preferences**, which means the following: there are augmented utility functions V^1, V^2, \dots, V^N such that for each $(i, t) \in N \times T$,

$$V^i(\mathbf{x}^{i,t}, -\bar{\mathbf{p}}^t \cdot \mathbf{x}^{i,t}) \geq V^i(\mathbf{x}, -\bar{\mathbf{p}}^t \cdot \mathbf{x}) \text{ for all } \mathbf{x} \in \mathbb{R}_+^K.$$

We say that ϕ is *regular* if

- (a) $\phi(e, \alpha e)$ is strictly increasing in e for every $\alpha > 0$;
- (b) *either* ϕ is increasing in its second argument and, for any compact set I with $0 \notin I$ such that $a, b, x \in I$, there is $m > 0$ such that $\phi(a, x) > \phi(b, y)$ for any $y < m$ *or* ϕ is decreasing in its second argument and for any compact set I with $0 \notin I$ such that $a, b, y \in I$, there is $m > 0$ such that $\phi(a, x) > \phi(b, y)$ for any $x < m$.

Condition (b) requires that if x and y are sufficiently far apart, then $\phi(a, x) - \phi(b, y)$ is positive. Recall the two examples of ϕ we have discussed above. When $\phi(e, e') = e'$, ϕ is obviously regular. If $\phi(e, e') = e f(e/e')$ with $f(1) = 1$, then it clearly satisfies property (a). If f is increasing, continuous, and $\lim_{r \rightarrow \infty} f(r) = \infty$, then ϕ is decreasing in its second argument and satisfies property (b); if f is decreasing, continuous, and $\lim_{r \rightarrow \infty} f(r) = 0$, then ϕ is increasing in its second argument and satisfies property (b).

The following result gives the precise sense in which preference and price heterogeneity have equivalent observable restrictions, in the context of the augmented utility model.

THEOREM 2. *For any data set $\mathcal{E} = \{(\mathbf{e}^{i,t})_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}$ the following are equivalent.*

- (i) \mathcal{E} is AU-rationalizable with heterogenous preferences.
- (ii) Suppose ϕ is a regular behavioral expenditure functions and W is a linear homogeneous aggregator function satisfying $\lim_{p^i \rightarrow \infty} W(p^i, (p^j)_{j \neq i}) = \infty$ for any $(p^j)_{j \neq i} \in \mathbb{R}_{++}^{N-1}$. Then there is a collection of prices $\mathcal{P} = \{\mathbf{p}^{i,t}\}_{(i,t) \in N \times T}$ that has a stable distribution (i.e., $\mathbf{p}^{i,t} = \lambda_i \bar{\mathbf{p}}^t$ for some $\lambda_i > 0$), is consistent in expectation w.r.t. W , and AU-rationalizes \mathcal{E} (for the given ϕ).

There is an analogous version of Theorem 2 in the setting of a cross-sectional data set. We say that a cross sectional data set \mathcal{D} is **random AU-rationalizable** (RAU-rationalizable) if there are sorting functions $\{\sigma^t\}_{t \in T}$ such that $\mathcal{E}(\{\sigma^t\}_{t \in T}) = \{(\sigma^t(i))_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}$ is AU-rationalizable with heterogenous preferences.¹⁸ We can compare this notion of AU-rationalization with that under the **random price model** (RPM). The cross sectional data set \mathcal{D} is **AU-RPM-rationalizable** by a collection of heterogenous prices $\mathcal{P} = \{\mathbf{p}^{i,t}\}_{(i,t) \in N \times T}$ if there are sorting functions $\{\sigma^t\}_{t \in T}$ such that $\mathcal{E}(\{\sigma^t\}_{t \in T}) = \{(\sigma^t(i))_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}$ is AU-rationalized by \mathcal{P} . The following result is an immediate consequence of Theorem 2.

COROLLARY 3. ¹⁹ *For any cross-sectional data $\mathcal{D} = \{E^t, \bar{\mathbf{p}}^t\}_{t \in T}$, the following are equivalent.*

- (i) \mathcal{D} is RAU-rationalizable.

¹⁸ There is a procedure to test for the RAU-rationalizability of a cross-sectional dataset \mathcal{D} , which is similar to the test of RUM-rationalizability presented in Kitamura and Stoye (2018)) (see Deb et al. (2023)).

¹⁹ It is possible to formulate a version of this theorem that allows for \mathcal{D} to have non-discrete distributions. Such a result would be analogous to the equivalence of RUM- and RPM-rationalizability when \mathcal{D} has non-discrete distributions (see Theorem 3 in the Appendix), and it will have a similar proof. We will not present such a result since it does not convey a message not already conveyed by Corollary 3.

(ii) Suppose ϕ is a regular behavioral expenditure function and W is linear homogeneous aggregator function satisfying $\lim_{p^i \rightarrow \infty} W(p^i, (p^j)_{j \neq i}) = \infty$ for any $(p^j)_{j \neq i} \in \mathbb{R}_{++}^{N-1}$. Then \mathcal{D} is AU-RPM rationalizable (for the given ϕ) by some $\mathcal{P} = \{\mathbf{p}^{i,t}\}_{(i,t) \in N \times T}$, where \mathcal{P} has a stable distribution and is consistent in expectation w.r.t. W .

To prove Theorem 2, we use Theorem 2 of Deb et al. (2023), which characterizes a more general version of the AUM that allows for non-linear pricing. Consider a collection of functions (or price systems) $\{f^t\}_{t \in T}$ on \mathbb{R}_+^K , where $f^t(\mathbf{x})$ is interpreted as the cost of purchasing \mathbf{x} at observation t . For example, prices are linear and $\bar{\mathbf{p}}^t$ are the prevailing prices, then $f^t(\mathbf{x}) = \bar{\mathbf{p}}^t \cdot \mathbf{x}$. We say a collection of consumption bundles and price systems $\{(\mathbf{x}^t, f^t)\}_{t \in T}$ is AU-rationalizable if there is an augmented utility function V such that for each $t \in T$,

$$V(\mathbf{x}^t, -f^t(\mathbf{x}^t)) \geq V(\mathbf{x}, -f^t(\mathbf{x})) \text{ for all } \mathbf{x} \in \mathbb{R}_+^K.$$

Abusing notation somewhat, we say that a collection of consumption bundles and price vectors $\{(\mathbf{x}^t, \bar{\mathbf{p}}^t)\}_{t \in T}$ is AU-rationalizable if it is AU-rationalizable with f^t given by $f^t(\mathbf{x}) = \bar{\mathbf{p}}^t \cdot \mathbf{x}^t$.

Theorem 2 of Deb et al. (2023) shows that AU-rationalization is characterized by a type of no-cycling condition, which we call *generalized axiom of price preference* (GAPP). We say f^t is **directly revealed (strictly) preferred to** f^s , denoted by $f^t \succsim_P f^s$, if $f^t(\mathbf{x}^s) \leq (<) f^s(\mathbf{x}^s)$. We say f^t is **revealed preferred to** f^s , denoted by $f^t \succsim_P^* f^s$, if there is a sequence $\{f^{t_l}\}_{l=1}^L$ such that $f^{t_1} = f^t$, $f^{t_L} = f^s$, and $f^{t_l} \succsim_P f^{t_{l+1}}$ for each $l < L$. We say that $\{(\mathbf{x}^t, f^t)\}_{t \in T}$ satisfies GAPP if, for any $s, t \in T$, $f^t \succsim_P^* f^s$ implies $f^s \not\succsim_P f^t$; GAPP is necessary and sufficient condition for $\{(\mathbf{x}^t, f^t)\}_{t \in T}$ to be AU-rationalizable.

Proof sketch of Theorem 2. The detailed proof can be found in the Appendix, but the following quick sketch may be helpful. Fix $\lambda_1, \dots, \lambda_N$ and let $f^{i,t}(\mathbf{x}) \equiv \phi(\bar{\mathbf{p}}^t \cdot \mathbf{x}, \lambda_i \bar{\mathbf{p}}^t \cdot \mathbf{x})$. Then by Theorem 2 of Deb et al. (2023), \mathcal{P} AU-rationalizes \mathcal{E} iff $\{(\mathbf{x}^{i,t}, f^{i,t})\}_{(i,t) \in N \times T}$ satisfies GAPP. On the other hand, \mathcal{E} is AU-rationalizable with heterogenous preferences iff $\{(\mathbf{x}^{i,t}, \bar{\mathbf{p}}^t)\}_{t \in T}$ satisfies GAPP for each i . It is now easy to see that (ii) implies (i). Indeed, if $\{(\mathbf{x}^{i,t}, f^{i,t})\}_{(i,t) \in N \times T}$ satisfies GAPP, then $\{(\mathbf{x}^{i,t}, f^{i,t})\}_{t \in T}$ satisfies GAPP for each i . Furthermore, ϕ is regular and thus satisfies property (a), which implies that $f^t(\mathbf{x}^s) \leq (<) f^s(\mathbf{x}^s)$ if and only if $\bar{\mathbf{p}}^t \cdot \mathbf{x}^s \leq (<) \bar{\mathbf{p}}^s \cdot \mathbf{x}^s$. We conclude that $\{(\mathbf{x}^{i,t}, f^{i,t})\}_{t \in T}$ satisfies GAPP if and only if $\{(\mathbf{x}^{i,t}, \bar{\mathbf{p}}^t)\}_{t \in T}$ satisfies GAPP.

To prove that (i) implies (ii), we construct $\lambda_1, \dots, \lambda_N$ such that λ_i/λ_{i+1} is sufficiently large (or small) so that $f^{j,s}$ cannot be revealed preferred to $f^{i,t}$ when $i < j$; in other words, there will be no revealed preference cycle involving observations from different agents. We show how this can be done in the Appendix, relying on the regularity of ϕ (and property (b) in particular). Therefore, checking whether $\{(\mathbf{x}^{i,t}, f^{i,t})\}_{(i,t) \in N \times T}$ satisfies GAPP is equivalent to checking whether $\{(\mathbf{x}^{i,t}, f^{i,t})\}_{t \in T}$ satisfies GAPP for each i . By property (a) of the regularity of ϕ , $\{(\mathbf{x}^{i,t}, f^{i,t})\}_{t \in T}$ satisfies GAPP if and only if $\{(\mathbf{x}^t, \bar{\mathbf{p}}^t)\}_{t \in T}$ satisfies GAPP. Lastly, the latter property holds since (i) holds. \square

Appendix

Proof of Proposition 1. By Afriat's theorem, we shall construct heterogenous prices $\mathcal{P} = \{\mathbf{p}^{i,t}\}_{(i,t) \in N \times T}$, which are consistent in expectation and satisfy $p_k^{i,t} = \bar{p}_k^t$ for all $k > 2$, such that a data set $\mathcal{O}^* = \{(\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}), \mathbf{p}^{i,t})\}_{(i,t) \in N \times T}$ satisfies GARP. For any $(i, t) \in N \times T$ and $k \in K$ with $k \geq 3$, let $p_k^{i,t} \equiv \bar{p}_k^t$. Hence, $W_k^t(p_k^{1,t}, \dots, p_k^{N,t}) = W_k^t(\bar{p}_k^t, \dots, \bar{p}_k^t) = \bar{p}_k^t$ when $k \geq 3$.

Let $>^*$ be a lexicographic order on $N \times T$ such that $(j, s) >^* (i, t)$ if $j > i$ or $i = j$ and $s > t$. In other words, $(i, t+1) >^* (i, t)$ and $(i+1, 1) >^* (i, T)$. We now construct $\{p_1^{i,t}, p_2^{i,t}\}_{(i,t) \in N \times T}$ such that the bundle $\mathbf{x}(\mathbf{e}^{j,s}, \mathbf{p}^{j,s})$ for (j, s) is not affordable for (i, t) whenever $(j, s) >^* (i, t)$. That is, $\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}) \not\preceq_D \mathbf{x}(\mathbf{e}^{j,s}, \mathbf{p}^{j,s})$ (i.e., $\mathbf{p}^{i,t} \cdot \mathbf{x}(\mathbf{e}^{j,s}, \mathbf{p}^{j,s}) > \mathbf{p}^{i,t} \cdot \mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}) = m^{i,t}$) when $(j, s) >^* (i, t)$.

We first construct $p_1^{i,t}$ with $i < N$ as follows. Take any $p_1 > 0$ and let $p_1^{1,1} \equiv p_1$. Moreover, take any small enough $\epsilon > 0$ and we can construct the rest of $p_1^{i,t}$ such that

$$\epsilon p_1^{i,t} \geq p_1^{j,s} \text{ whenever } (j, s) >^* (i, t).$$

In words, the price of good 1 is lower in later periods and is also lower for agents with higher indexes; i.e., $p_1^{i,t}$ is decreasing exponentially in i as well as t . In particular, set $p_1^{i,t} = p_1 \epsilon^{(i-1)T+t-1}$. Therefore, $\mathbf{x}^{j,s}(\mathbf{e}^{j,s}, \mathbf{p}^{j,s})$ is not affordable for (i, t) when $(j, s) >^* (i, t)$ and $i, j < N$. More formally, when $i, j < N$, when ϵ is small enough we have

$$\mathbf{p}^{i,t} \cdot \mathbf{x}(\mathbf{e}^{j,s}, \mathbf{p}^{j,s}) = \sum_{k \in K} p_k^{i,t} \frac{e_k^{j,s}}{p_k^{j,s}} > \frac{p_1^{i,t}}{p_1^{j,s}} e_1^{j,s} \geq \frac{e_1^{j,s}}{\epsilon} > m^{i,t} \text{ since } \frac{p_1^{i,t}}{p_1^{j,s}} \geq \frac{1}{\epsilon}.$$

Second, let $p_1^{N,t}$ be the solution to $W_1^t(p_1^{1,t}, \dots, p_1^{n-1,t}, p_1^{N,t}) = \bar{p}_1^t$. Third, take any $p_2 > 0$ and let $p_2^{N,t} \equiv \epsilon^t p_2$. Similar to the previous argument, the bundle $\mathbf{x}(\mathbf{e}^{N,t+1}, \mathbf{p}^{N,t+1})$ for $(N, t+1)$ is not affordable for (N, t) (i.e., $\mathbf{x}(\mathbf{e}^{N,t}, \mathbf{p}^{N,t}) \not\preceq_D \mathbf{x}(\mathbf{e}^{N,t+1}, \mathbf{p}^{N,t+1})$) since the price of good 2 for agent N is exponentially decreasing in t . Fourth, for any (i, t) with $i < N$, let $p_2^{i,t} \equiv \bar{p}_2^t$ where \bar{p}_2^t solves $W_2^t(\bar{p}_2^t, \dots, \bar{p}_2^t, \epsilon^t p_2) = \bar{p}_2^t$.

Case 1. Suppose for any $(p^j)_{j \neq i} \in \mathbb{R}_{++}^{N-1}$ and $(t, k) \in T \times K$, $\lim_{p^i \rightarrow +\infty} W_k^t(p^i, (p^j)_{j \neq i}) = +\infty$.

In this case, the equation $W_1^t(p_1^{1,t}, \dots, p_1^{n-1,t}, p_1^{N,t}) = \bar{p}_1^t$ has a solution when p_1 is small enough so that $\bar{p}_1^t \geq p_1^{i,t} = p_1 \epsilon^{(i-1)T+t-1}$ for each $i < N$. Moreover, $W_2^t(\bar{p}_2^t, \dots, \bar{p}_2^t, \epsilon^t p_2) = \bar{p}_2^t$ has a solution when p_2 is small enough.

To prove GARP, it is sufficient to show that $\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}) \not\preceq_D \mathbf{x}(\mathbf{e}^{j,s}, \mathbf{p}^{j,s})$ whenever $(j, s) >^* (i, t)$. By the previous arguments, it is enough to consider the case in which $j = N$ and $i < N$. Since $p_2^{i,t} = \bar{p}_2^t \geq \bar{p}_2^t$, we have $\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}) \not\preceq_D \mathbf{x}(\mathbf{e}^{N,s}, \mathbf{p}^{N,s})$; i.e.,

$$\mathbf{p}^{i,t} \cdot \mathbf{x}(\mathbf{e}^{N,s}, \mathbf{p}^{N,s}) = \sum_{k \in K} p_k^{i,t} \frac{e_k^{N,s}}{p_k^{N,s}} > \frac{p_2^{i,t}}{p_2^{N,s}} e_2^{N,s} \geq \frac{\bar{p}_2^t}{\epsilon^s p_2} e_2^{N,s} > m^{i,t} \text{ when } \epsilon \text{ and } p_2 \text{ are small enough.}$$

Case 2. Suppose for any $(p^j)_{j \neq i} \in \mathbb{R}_{++}^{N-1}$ and $(t, k) \in T \times K$, $\lim_{p^i \rightarrow +0} W_k^t(p^i, (p^j)_{j \neq i}) = 0$.

In this case, the equation $W_1^t(p_1^{1,t}, \dots, p_1^{n-1,t}, p_1^{N,t}) = \bar{p}_1^t$ has a solution when p_1 is large enough so that $\frac{\bar{p}_1^t}{\epsilon} \leq p_1^{i,t} = p_1 \epsilon^{(i-1)T+t-1}$ for each $i < N$. Moreover, $W_2^t(\bar{p}_2^t, \dots, \bar{p}_2^t, \epsilon^t p_2) = \bar{p}_2^t$ has a solution when p_2 is large enough.

To prove GARP, it is sufficient to show that $\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}) \not\preceq_D \mathbf{x}(\mathbf{e}^{j,s}, \mathbf{p}^{j,s})$ whenever $(j, s) >^* (i, t)$. By the previous arguments, it is enough to consider the case in which $j = N$ and $i < N$. Since $p_1^{i,t} \geq \frac{\bar{p}_1^t}{\epsilon}$ and $p_1^{N,s} \leq \bar{p}_1^s$, we have $\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}) \not\preceq_D \mathbf{x}(\mathbf{e}^{N,s}, \mathbf{p}^{N,s})$; i.e.,

$$\mathbf{p}^{i,t} \cdot \mathbf{x}(\mathbf{e}^{N,s}, \mathbf{p}^{N,s}) = \sum_{k \in K} p_k^{i,t} \frac{e_k^{N,s}}{p_k^{N,s}} > \frac{p_1^{i,t}}{p_1^{N,s}} e_1^{N,s} \geq \frac{\bar{p}_1^t}{\epsilon \bar{p}_1^s} e_1^{N,s} > m^{i,t}$$

when ϵ is small enough. \square

Below, we provide an example of a data set that cannot be rationalized by price heterogeneity in one good. This shows that, in Proposition 1, it is essential that there be at least two goods with price heterogeneity. In fact, the example establishes a stronger conclusion: the data set it considers cannot be rationalized by price heterogeneity in a single good alone, even if we allow the two agents to have distinct preferences.

Example. We assume that there are two agents, four goods, and four observations. The two agents have identical expenditure allocations, and the observations are as follows.

$$\begin{aligned} \mathbf{e}^{i,1} &= (10, 1, 1, 1) & \mathbf{e}^{i,2} &= (1, 10, 1, 1) & \mathbf{e}^{i,3} &= (1, 1, 10, 1) & \mathbf{e}^{i,4} &= (1, 1, 1, 10) \\ \bar{\mathbf{p}}^1 &= (2, 1, 1, 1) & \bar{\mathbf{p}}^2 &= (1, 2, 1, 1) & \bar{\mathbf{p}}^3 &= (1, 2, 1, 1) & \bar{\mathbf{p}}^{i,2} &= (1, 1, 1, 2) \end{aligned}$$

Proposition 1 guarantees that this data can be rationalized by heterogenous prices with consistent expectations, so long as there is price heterogeneity in at least two goods. However, assuming that prices are aggregated by taking the arithmetic average, we claim that this data *cannot* be rationalized with price heterogeneity in good 1 alone, even if we allow the two agents to have different preferences. Of course, given the symmetry, it is clear that this data set cannot be rationalized by price heterogeneity in any other good alone.

Let $x_k^{i,t} = e_k^{i,t}/p_k^{i,t}$. Note that for any $i \in N$ and distinct $t, t' \in \{2, 3, 4\}$, $\mathbf{x}^{i,t}$ is directly revealed strictly preferred to $\mathbf{x}^{i,t'}$ if and only if

$$p_1^{i,t} \times \frac{1}{p_1^{i,t'}} + 1 \times 1 + 1 \times 5 + 2 \times 1 < 13$$

or $p_1^{i,t}/p_1^{i,t'} < 5$. Therefore, $\mathbf{x}^{i,t}$ is directly revealed strictly preferred to $\mathbf{x}^{i,t'}$ and vice versa if and only if

$$\frac{1}{5} < \frac{p_1^{i,t'}}{p_1^{i,t}} < 5. \quad (14)$$

Since $p_1^{1,t} + p_1^{2,t} = 1$ for any $t = 2, 3, 4$, there is an agent i and distinct $t, t' \in \{2, 3, 4\}$ such that $1 \geq p^{i,t}, p^{i,t'} \geq 1/2$. Therefore,

$$\frac{1}{2} < \frac{p_1^{i,t}}{p_1^{i,t'}} < 2$$

which means (given (14)) that agent i has violated GARP. \square

Proof of Proposition 2. Take any stable price distribution $\mathcal{P} = \{\mathbf{p}^{i,t}\}_{(i,t) \in N \times T}$ where $p_k^{i,t} = \lambda_{i,k} \bar{p}_k^t$ for some $\{\lambda_{i,k}\}_{(i,k) \in N \times K}$. By Afriat's Theorem, $\mathcal{E} = \{(\mathbf{e}^{i,t})_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}$ is rationalizable with heterogenous preferences if and only if $\mathcal{O}^i = \{(\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}), \mathbf{p}^{i,t})\}_{t \in T}$ satisfies GARP for each $i \in N$. Similarly, $\{(\mathbf{e}^{i,t}, \mathbf{p}^{i,t})\}_{(i,t) \in N \times T}$ is rationalizable with heterogenous preferences if and only if $\bar{\mathcal{O}}^i = \{(\mathbf{x}(\mathbf{e}^{i,t}, \bar{\mathbf{p}}^t), \bar{\mathbf{p}}^t)\}_{t \in T}$ satisfies GARP for each $i \in N$. To establish our result, we need only show that for any $i \in N$, \mathcal{O}^i satisfies GARP if and only if $\bar{\mathcal{O}}^i$ satisfies GARP. Indeed, note that the revealed preference relations on \mathcal{O}^i and $\bar{\mathcal{O}}^i$ are the 'same' in the following sense: for any $s, t \in T$, $\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t})$ is directly revealed preferred to $\mathbf{x}(\mathbf{e}^{i,s}, \mathbf{p}^{i,s})$

if and only if $\mathbf{x}(\mathbf{e}^{i,t}, \bar{\mathbf{p}}^t)$ is directly revealed preferred to $\mathbf{x}(\mathbf{e}^{i,s}, \bar{\mathbf{p}}^s)$. This is simply because, by the definition of revealed preference relations, $\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}) \succeq_D \mathbf{x}(\mathbf{e}^{i,s}, \mathbf{p}^{i,s})$ is equivalent to

$$\mathbf{p}^{i,t} \cdot \mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}) = m^{i,t} \geq \mathbf{p}^{i,t} \cdot \mathbf{x}(\mathbf{e}^{i,s}, \mathbf{p}^{i,s}) = \sum_{k \in K} p_k^{i,s} \frac{e_k^{i,t}}{p_k^{i,t}} = \sum_{k \in K} \lambda_{i,k} \bar{p}_k^s \frac{e_k^{i,t}}{\lambda_{i,k} \bar{p}_k^t} = \sum_{k \in K} \bar{p}_k^s \frac{e_k^{i,t}}{\bar{p}_k^t}$$

while $\mathbf{x}(\mathbf{e}^{i,t}, \bar{\mathbf{p}}^t) \succeq_D \mathbf{x}(\mathbf{e}^{i,s}, \bar{\mathbf{p}}^s)$ is also equivalent to

$$\bar{\mathbf{p}}^t \cdot \mathbf{x}(\mathbf{e}^{i,t}, \bar{\mathbf{p}}^t) = m^{i,t} \geq \bar{\mathbf{p}}^t \cdot \mathbf{x}(\mathbf{e}^{i,t}, \bar{\mathbf{p}}^s) = \mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}) = \sum_{k \in K} \bar{p}_k^s \frac{e_k^{i,t}}{\bar{p}_k^t}.$$

Hence, for each $i \in N$, \mathcal{O}^i satisfies GARP if and only if $\bar{\mathcal{O}}^i$ satisfies GARP. \square

Proof of Theorem 1. Clearly, (ii) implies (i). To show that (iii) implies (ii), take any $R \subseteq K$ with $e_R^{i,t} > 0$ for each $(i, t) \in N \times T$. Suppose there are real numbers $\lambda_1, \dots, \lambda_N > 0$ such that stable price distribution $\mathcal{P} = \{\mathbf{p}^{i,t}\}_{(i,t) \in N \times T}$ with $\mathbf{p}^{i,t} = (\lambda_i \bar{\mathbf{p}}_R^t, \bar{\mathbf{p}}_{-R}^t)$ rationalizes \mathcal{E} .²⁰ Hence, there is a utility function $U : \mathbb{R}_+^K \rightarrow \mathbb{R}$ such that for each $(i, t) \in N \times T$,

$$U(\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t})) \geq U(\mathbf{x}) \text{ for any } \mathbf{x} \in \mathcal{B}(\mathbf{p}^{i,t}, m^{i,t}). \quad (15)$$

Since $\mathbf{x}_R(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}) = \mathbf{x}_R(\mathbf{e}^{i,t}, \bar{\mathbf{p}}^t)/\lambda_i$ and $\mathbf{x}_{-R}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}) = \mathbf{x}_{-R}(\mathbf{e}^{i,t}, \bar{\mathbf{p}}^t)$, we have

$$U(\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t})) = U\left(\frac{\mathbf{x}_R(\mathbf{e}^{i,t}, \bar{\mathbf{p}}^t)}{\lambda_i}, \mathbf{x}_{-R}(\mathbf{e}^{i,t}, \bar{\mathbf{p}}^t)\right) = U_R^{\frac{1}{\lambda_i}}(\mathbf{x}(\mathbf{e}^{i,t}, \bar{\mathbf{p}}^t)).$$

Let $\tilde{\mathbf{x}} = (\lambda_i \mathbf{x}_R, \mathbf{x}_{-R})$. Then $U(\mathbf{x}) = U_R^{\frac{1}{\lambda_i}}(\lambda_i \mathbf{x}_R, \mathbf{x}_{-R}) = U_R^{\frac{1}{\lambda_i}}(\tilde{\mathbf{x}})$. Moreover,

$$\mathbf{x} \in \mathcal{B}(\mathbf{p}^{i,t}, m^{i,t}) \text{ if and only if } \tilde{\mathbf{x}} \in \mathcal{B}(\bar{\mathbf{p}}^t, m^{i,t})$$

since $\mathbf{p}^{i,t} \cdot \mathbf{x} = \lambda_i \bar{\mathbf{p}}_R^t \cdot \mathbf{x}_R + \bar{\mathbf{p}}_{-R}^t \cdot \mathbf{x}_{-R} = \bar{\mathbf{p}}^t \cdot \tilde{\mathbf{x}}$. Finally, the above three observations show that (15) is equivalent to

$$U_R^{\frac{1}{\lambda_i}}(\mathbf{x}(\mathbf{e}^{i,t}, \bar{\mathbf{p}}^t)) \geq U_R^{\frac{1}{\lambda_i}}(\tilde{\mathbf{x}}) \text{ for any } \tilde{\mathbf{x}} \in \mathcal{B}(\bar{\mathbf{p}}^t, m^{i,t}).$$

Therefore, \mathcal{E} is rationalizable with heterogenous preferences drawn from an R -scale family.

It remains to show that (i) implies (iii). Suppose \mathcal{E} is rationalizable with heterogenous preferences. We shall find $\lambda_1, \dots, \lambda_N > 0$ such that stable price distribution $\mathcal{P} = \{\mathbf{p}^{i,t}\}_{(i,t) \in N \times T}$ with $\mathbf{p}^{i,t} = (\lambda_i \bar{\mathbf{p}}_R^t, \bar{\mathbf{p}}_{-R}^t)$ is consistent in expectation and that rationalizes \mathcal{E} .

²⁰ We will not use the fact that prices are consistent in expectation.

Let $\lambda_i = \beta \epsilon^i$ where $\epsilon > 0$ and $\beta = 1/W((\epsilon^i)_{i \in N})$. By homogeneous of degree 1 of W , we have $W((\lambda_i \bar{p}_k^t)_{i \in N}) = W((\beta \epsilon^i \bar{p}_k^t)_{i \in N}) = \bar{p}_k^t \beta W((\epsilon^i)_{i \in N}) = \bar{p}_k^t$ for any $k \in R$ and $t \in T$. By the property of W , we also have $W((p_k^{i,t})_{i \in N}) = W((\bar{p}_k^t)_{i \in N}) = \bar{p}_k^t$ for any $k \notin R$ and $t \in T$. Therefore, $\mathcal{P} = \{\mathbf{p}^{i,t}\}_{(i,t) \in N \times T}$ is consistent in expectation.

We now shall show that $\mathcal{P} = \{\mathbf{p}^{i,t}\}_{(i,t) \in N \times T}$ rationalizes \mathcal{E} when ϵ is small enough. By Afriat's theorem, it is enough to show that a data set $\mathcal{O}^* = \{(\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}), \mathbf{p}^{i,t})\}_{(i,t) \in N \times T}$ obeys GARP when ϵ is small enough. We write $\mathcal{O}^* = \bigcup_{i \in N} \mathcal{O}^i$, where $\mathcal{O}^i = \{(\mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}), \mathbf{p}^{i,t})\}_{t \in T}$.

Let $\succsim_{\overline{D}}$ be the revealed preference relation on \mathcal{O}^* . We claim that when ϵ is sufficiently small, for any $i, j \in N$ with $i > j$ and $s, t \in T$, $\mathbf{x}(\mathbf{e}^{j,s}, \mathbf{p}^{j,s}) \not\succsim_{\overline{D}} \mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t})$. Since

$$\mathbf{p}^{j,s} \cdot \mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t}) = \frac{\lambda_j}{\lambda_i} \sum_{k \in R} \bar{p}_k^s \frac{e_k^{i,t}}{\bar{p}_k^t} + \sum_{k \notin R} \bar{p}_k^s \frac{e_k^{i,t}}{\bar{p}_k^t} \geq \frac{1}{\epsilon^{i-j}} \sum_{k \in R} \bar{p}_k^s \frac{e_k^{i,t}}{\bar{p}_k^t}$$

and $\mathbf{e}_R^{i,t} > 0$, $\mathbf{p}^{j,s} \cdot \mathbf{x}(\mathbf{e}^{i,t}, \mathbf{p}^{i,t})$ will exceed $\mathbf{p}^{j,s} \cdot \mathbf{x}(\mathbf{e}^{j,s}, \mathbf{p}^{j,s}) = \sum_{k \in K} e_k^{j,s}$ when ϵ is small enough and $i > j$. Therefore, there is no revealed preference cycle involving observations from \mathcal{O}^i and \mathcal{O}^j when $i \neq j$. Lastly, since \mathcal{E} is rationalizable with heterogenous preferences, Proposition 2 guarantees that \mathcal{O}^i satisfies GARP for each $i \in N$. Therefore, \mathcal{O}^* satisfies GARP. \square

Proof of Theorem 2. We have already shown that (ii) implies (i) in the proof sketch provided in Section 5.1. It remains for us to show that (i) implies (ii). First, observe that, for any $\lambda_i > 0$, $\{(\mathbf{x}^{i,t}, f^{i,t})\}_{t \in T}$ satisfies GAPP. Indeed, if (i) holds, then $\{(\mathbf{x}^{i,t}, \bar{\mathbf{p}}^t)\}_{t \in T}$ satisfies GAPP and then the regularity of ϕ (specifically, property (a)) guarantees that $\{(\mathbf{x}^{i,t}, f^{i,t})\}_{t \in T}$ also satisfies GAPP (see the argument in the proof sketch). Furthermore, we claim that we can construct $\lambda_1, \dots, \lambda_N$ such that there are no revealed preference cycles involving $f^{i,t}$ and $f^{j,s}$ with $i \neq j$; we establish this claim below using property (b) in the definition of the regularity of ϕ . We can thus conclude that $\{(\mathbf{x}^{i,t}, f^{i,t})\}_{(i,t) \in N \times T}$ satisfies GAPP and, by Theorem 2 of Deb et al. (2023), it is AU-rationalizable, as required by (ii). We break down our proof into two cases, corresponding to the two conditions required for the regularity of ϕ under property (b).

Case 1. ϕ is increasing in its second argument and, for any compact set I with $0 \notin I$ such that $a, b, x \in I$, there is $m > 0$ such that $\phi(a, x) > \phi(b, y)$ for any $y < m$.

It is straightforward to show that this combination of assumptions on ϕ guarantees that for any $a, b \in I$ such that I is a compact set with $0 \notin I$, and $x \in S$ where S is uniformly

bounded away from zero, there is $m > 0$ such that $\phi(a, x) > \phi(b, y)$ for any $y < m$. It is this property on ϕ specifically that we need in our proof.

We now explain how we can construct positive numbers $\lambda_{N-1}, \dots, \lambda_2, \lambda_1$ such that, for any $j > i$ and $s, t \in T$, $f^{j,s} \not\preceq_P f^{i,t}$ for any $\lambda_N \geq 1$. First, note that $\{\bar{\mathbf{p}}^s \cdot \mathbf{x}^{N-1,t}\}_{s,t \in T}$ is a compact set not containing zero, and the set $\{\lambda_N \bar{\mathbf{p}}^s \cdot \mathbf{x}^{N-1,t}\}_{s,t \in T, \lambda_N \geq 1}$ is uniformly bounded away from zero. Therefore, there is $0 < \lambda_{N-1} < 1$ such that

$$f^{N,s}(\mathbf{x}^{N-1,t}) = \phi(\bar{\mathbf{p}}^s \cdot \mathbf{x}^{N-1,t}, \lambda_N \bar{\mathbf{p}}^s \cdot \mathbf{x}^{N-1,t}) > f^{N-1,t}(\mathbf{x}^{N-1,t}) = \phi(\bar{\mathbf{p}}^t \cdot \mathbf{x}^{N-1,t}, \lambda_{N-1} \bar{\mathbf{p}}^t \cdot \mathbf{x}^{N-1,t})$$

when λ_{N-1} is small enough. Note that λ_{N-1} can be chosen independently of the value of λ_N . Assuming that we have chosen $\lambda_{N-1}, \dots, \lambda_{i+1}$, we now explain how λ_i can be chosen. Observe that $\{\bar{\mathbf{p}}^s \cdot \mathbf{x}^{i,t}\}_{s,t \in T}$ is a compact set not containing zero and the set $\{\lambda_N \bar{\mathbf{p}}^s \cdot \mathbf{x}^{i,t}\}_{s,t \in T, \lambda_N \geq 1} \cup \{\lambda_j \bar{\mathbf{p}}^s \cdot \mathbf{x}^{i,t}\}_{s,t \in T, N > j > i}$ is uniformly bounded away from zero. Therefore, there is $\lambda_i \in (0, 1)$ such that for all $j > i$,

$$f^{j,s}(\mathbf{x}^{i,t}) = \phi(\bar{\mathbf{p}}^s \cdot \mathbf{x}^{i,t}, \lambda_j \bar{\mathbf{p}}^s \cdot \mathbf{x}^{i,t}) > f^{i,t}(\mathbf{x}^{i,t}) = \phi(\bar{\mathbf{p}}^t \cdot \mathbf{x}^{i,t}, \lambda_i \bar{\mathbf{p}}^t \cdot \mathbf{x}^{i,t})$$

when λ_i is sufficiently small. Note that λ_i can be chosen independently of λ_N .

We have shown that there is $\lambda_{N-1}, \dots, \lambda_2, \lambda_1$ such for any $\lambda_N \geq 1$ and $i, j \in N$ with $i < j$ and $s, t \in T$, $f^{j,s} \not\preceq_P f^{i,t}$. Thus there can no revealed preference cycles involving $f^{i,t}$ and $f^{j,s}$ with $i \neq j$. We also know that $W(\lambda_1, \lambda_2, \dots, 1) \leq 1$, so for λ_N sufficiently large, $W(\lambda_1, \lambda_2, \dots, \lambda_N) = 1$, as required.

Case 2. ϕ is decreasing in its second argument, and for any compact set I with $0 \notin I$ such that $a, b, y \in I$, there is $m > 0$ such that $\phi(a, x) > \phi(b, y)$ for any $x < m$.

The proof is analogous to the proof for Case 1. Note that the combination of assumptions on ϕ in Case 2 guarantees that for any $a, b \in I$ such that I is a compact set with $0 \notin I$ and $y \in S$ where S is uniformly bounded away from zero, there is $m > 0$ such that $\phi(a, x) > \phi(b, y)$ for any $x < m$. It is this property on ϕ specifically that we need in our proof.

We now explain how we can construct positive numbers $\lambda_{N-1}, \dots, \lambda_2, \lambda_1$ such that, for any $j > i$ and $s, t \in T$, $f^{i,s} \not\preceq_P f^{j,t}$ for any $\lambda_N \geq 1$. First, note that $\{\bar{\mathbf{p}}^s \cdot \mathbf{x}^{N,t}\}_{s,t \in T}$ is a compact set bounded away from zero and the set $\{\lambda_N \bar{\mathbf{p}}^t \cdot \mathbf{x}^{N,t}\}_{t \in T, \lambda_N \geq 1}$ is uniformly bounded away from zero. Therefore, there is $0 < \lambda_{N-1} < 1$ such that

$$f^{N,t}(\mathbf{x}^{N,t}) = \phi(\bar{\mathbf{p}}^t \cdot \mathbf{x}^{N,t}, \lambda_N \bar{\mathbf{p}}^t \cdot \mathbf{x}^{N,t}) < f^{N-1,s}(\mathbf{x}^{N,t}) = \phi(\bar{\mathbf{p}}^s \cdot \mathbf{x}^{N,t}, \lambda_{N-1} \bar{\mathbf{p}}^s \cdot \mathbf{x}^{N,t})$$

when λ_{N-1} is small enough. Note that λ_{N-1} can be chosen independently of the value of λ_N . So this guarantees that $f^{N-1,s} \not\preceq_P f^{N,t}$ for any $\lambda_N \geq 1$. Assuming that we have chosen $\lambda_{N-1}, \dots, \lambda_{i+1}$, we now explain how λ_i can be chosen. Observe that $\{\bar{\mathbf{p}}^s \cdot \mathbf{x}^{j,t}\}_{s,t \in T, j > i}$ is a compact set not containing zero and the set $\{\lambda_N \bar{\mathbf{p}}^t \cdot \mathbf{x}^{N,t}\}_{t \in T, \lambda_N \geq 1} \cup \{\lambda_j \bar{\mathbf{p}}^t \cdot \mathbf{x}^{j,t}\}_{t \in T, N > j > i}$ is uniformly bounded away from zero. Therefore, there is $\lambda_i \in (0, 1)$ such that for all $j > i$,

$$f^{j,t}(\mathbf{x}^{j,t}) = \phi(\bar{\mathbf{p}}^t \cdot \mathbf{x}^{j,t}, \lambda_j \bar{\mathbf{p}}^t \cdot \mathbf{x}^{j,t}) < f^{i,s}(\mathbf{x}^{j,t}) = \phi(\bar{\mathbf{p}}^s \cdot \mathbf{x}^{j,t}, \lambda_i \bar{\mathbf{p}}^s \cdot \mathbf{x}^{j,t})$$

when λ_i is sufficiently small. Note that λ_i can be chosen independently of λ_N .

We have shown that there is $\lambda_{N-1}, \dots, \lambda_2, \lambda_1$ such for any $\lambda_N \geq 1$ and $i, j \in N$ with $i < j$ and $s, t \in T$, $f^{i,s} \not\preceq_P f^{j,t}$. We know that $W(\lambda_1, \lambda_2, \dots, 1) \leq 1$, so for λ_N sufficiently large, $W(\lambda_1, \lambda_2, \dots, \lambda_N) = 1$, as required. \square

Random Utility and Random Price Models. So far, we have considered discrete cross-sectional data sets. In this section, we show an equivalence between price and preference heterogeneity in a cross-sectional data environment which has continuous probability distributions on expenditure bundles. Let $\mathcal{E}^t = \{\mathbf{e} \in \mathbb{R}_+^K : \sum_{k \in K} e_k = m^t\}$ where m^t is the income level. Then, our data set takes the form

$$\mathcal{D} = \{(\mu_e^t, \bar{\mathbf{p}}^t)\}_{t \in T},$$

where μ_e^t is a probability distribution over \mathcal{E}^t . Let \mathcal{U} be the set of all strictly increasing, strictly concave, and continuous utility functions on \mathbb{R}_+^K . We then say that the dataset \mathcal{D} is **RUM-rationalizable** if there is a probability distribution ρ on \mathcal{U} such that for any $t \in T$ and any measurable subset $E \subseteq \mathbb{R}_+^K$,

$$\mu_e^t(E) = \rho(\{u \in \mathcal{U} : \arg \max_{\mathbf{e} \in \mathcal{E}^t} u(\mathbf{x}(\mathbf{e}, \bar{\mathbf{p}}^t)) \in E\}).$$

We say that the dataset \mathcal{D} is **RPM-rationalizable** if there is a probability distribution η on \mathbb{R}_{++} satisfying $\int \lambda d\eta(\lambda) = 1$ and a utility function $u \in \mathcal{U}$ such that, for any $t \in T$ and any measurable subset $E \subseteq \mathbb{R}_+^K$,

$$\mu_e^t(E) = \eta(\{\lambda \in \mathbb{R}_{++} : \arg \max_{\mathbf{e} \in \mathcal{E}^t} u(\mathbf{x}(\mathbf{e}, \mathbf{p}^{\lambda,t})) \in E\}),$$

where $\mathbf{p}^{\lambda,t} \equiv \lambda \bar{\mathbf{p}}^t$. (Note that our definition of RPM-rationalizability here incorporates a stable price distribution and consistency in expectation w.r.t. the arithmetic mean.)

The next result states that \mathcal{D} is RUM-rationalizable if it is RPM-rationalizable; furthermore, whenever \mathcal{D} is RUM-rationalizable, then there is another dataset $\hat{\mathcal{D}}$, arbitrarily close to \mathcal{D} by the Levy-Prokhorov metric,²¹ such that $\hat{\mathcal{D}}$ is RPM-rationalizable. In this sense, RUM- and RPM-rationalizability are approximately equivalent.

THEOREM 3. *Suppose $\mathcal{D} = \{(\mu_e^t, \bar{\mathbf{p}}^t)\}_{t \in T}$ where μ_e^t are probability distributions which are absolutely continuous with respect to the Lebesgue measure on \mathcal{E}^t .²² If \mathcal{D} is RPM-rationalizable, then it is RUM-rationalizable. Conversely, if \mathcal{D} is RUM-rationalizable, then for any $\epsilon > 0$, there are discrete probability distributions $\hat{\mu}_e^t \in \Delta(\mathcal{E}^t)$ such that $\max_{t \in T} d_{LP}(\mu_e^t, \hat{\mu}_e^t) < \epsilon$ and $\hat{\mathcal{D}} = \{(\hat{\mu}_e^t, \bar{\mathbf{p}}^t)\}_{t \in T}$ is RPM-rationalizable.*

Proof of Theorem 3. We first show that RPM implies RUM. Suppose \mathcal{D} is RPM-rationalizable; i.e., there is a distribution η on $\Delta(\mathbb{R}_{++})$ and $u \in \mathcal{U}$ such that, for any $t \in T$ and any measurable subset $E \subseteq \mathbb{R}_+^K$, $\mu_e^t(E) = \eta(\{\lambda \in \mathbb{R}_{++} : \arg \max_{\mathbf{e} \in \mathcal{E}^t} u(\mathbf{x}(\mathbf{e}, \mathbf{p}^{\lambda, t})) \in E\})$. For each $\lambda > 0$ and $t \in T$, let $\mathbf{e}^{\lambda, t} \equiv \arg \max_{\mathbf{e} \in \mathcal{E}^t} u(\mathbf{x}(\mathbf{e}, \mathbf{p}^{\lambda, t}))$. Then the data set $\mathcal{O}_\lambda = \{(\mathbf{x}(\mathbf{e}^{\lambda, t}, \mathbf{p}^{\lambda, t}), \mathbf{p}^{\lambda, t})\}_{t \in T}$ is rationalized by utility function u . Then the data set $\bar{\mathcal{O}}_\lambda = \{(\mathbf{x}(\mathbf{e}^{\lambda, t}, \bar{\mathbf{p}}^t), \bar{\mathbf{p}}^t)\}_{t \in T}$ is also rationalizable since \mathcal{O}_λ obeys GARP if and only if $\bar{\mathcal{O}}_\lambda$ obeys GARP (see the the proof of Proposition 2). Hence, there is a utility function u_λ such that $u_\lambda(\mathbf{x}(\mathbf{e}^{\lambda, t}, \bar{\mathbf{p}}^t)) \geq u_\lambda(\mathbf{x})$ for every $\mathbf{x} \in \mathcal{B}(\bar{\mathbf{p}}^t, m^t)$. Let ρ be a probability distribution over \mathcal{U} such that $\rho(\mathcal{U}) = \eta(\{\lambda \in \mathbb{R}_{++} : u_\lambda \in \mathcal{U}\})$ for any measurable subset $\mathcal{U} \in \mathcal{U}$. Therefore, we have $\mu_e^t(E) = \rho(\{u_\lambda \in \mathcal{U} : \arg \max_{\mathbf{e} \in \mathcal{E}^t} u_\lambda(\mathbf{x}(\mathbf{e}, \bar{\mathbf{p}}^t)) \in E\})$ for any $t \in T$ and any measurable subset $E \subseteq \mathbb{R}_+^K$. In other words, \mathcal{D} is RUM-rationalizable.

It remains to show that RUM-rationalizability implies approximate RPM-rationalizability. Our proof uses Theorem 1 of Kitamura and Stoye (2018). Suppose \mathcal{D} is RUM-rationalizable; i.e., there is $\rho \in \Delta(\mathcal{U})$ such that for any $t \in T$ and any measurable subset $E \subseteq \mathbb{R}_+^K$, $\mu_e^t(E) = \rho(\{u \in \mathcal{U} : \arg \max_{\mathbf{e} \in \mathcal{E}^t} u(\mathbf{x}(\mathbf{e}, \bar{\mathbf{p}}^t)) \in E\})$.

Let B^t be the budget plane at observation t , i.e., $B^t = \{\mathbf{x} \in \mathbb{R}_+^K : \mathbf{x} \cdot \bar{\mathbf{p}}^t = m^t\}$. We can find $Y_{1,t}, Y_{2,t}, \dots, Y_{L_t,t}$ such that $B^t = \bigcup_{l=1}^{L_t} Y_{l,t}$ where $Y_{l,t}$ has the following property: for any $t' \neq t$, either $\bar{\mathbf{p}}^{t'} \mathbf{x} - m^{t'} \geq 0$ for all $x \in Y_{l,t}$ or $\bar{\mathbf{p}}^{t'} \mathbf{x} - m^{t'} \leq 0$ for all $x \in Y_{l,t}$. Following

²¹ For any probability distributions μ and μ' , the Levy-Prokhorov metric d_{LP} is given by $d_{LP}(\mu, \mu') = \inf\{\epsilon > 0 : \mu'(E^\epsilon) + \epsilon > \mu(E) \text{ and } \mu(E^\epsilon) + \epsilon > \mu'(E) \text{ for any } E\}$ where $E^\epsilon = \bigcup_{x \in E} \{y \in \mathbb{R}^K : \|x - y\| < \epsilon\}$.

²² The absolute continuity assumption on μ_e^t is not strictly necessary but it simplifies the exposition because it guarantees that the intersection of patches has probability measure zero (see Kitamura and Stoye (2018)).

Kitamura and Stoye, we refer to $Y_{l,t}$ as a *patch*. It is clear that two patches on B^t intersect only at their boundaries.

Let μ^t be the probability measure on B^t induced by the measure μ_e^t on \mathcal{E}^t and let $\pi_{l,t} = \mu^t(Y_{l,t})$ for each (l, t) , i.e., $\pi_{l,t}$ is the probability that a consumption bundle lies in $Y_{l,t}$. Note that $\sum_{l=1}^{L_t} \pi_{l,t} = 1$ since (by the absolute continuity assumption) $\mu^t(Y_{l,t} \cap Y_{l',t}) = 0$ for any $l \neq l'$. Given any $\epsilon > 0$, for each $t \in T$, there is a discrete probability distribution with finite support, $\hat{\mu}_e^t$, on \mathcal{E}^t such that $d_{LP}(\hat{\mu}_e^t, \mu_e^t) < \epsilon$, with $\hat{\mu}_e^t$ inducing a discrete probability distribution $\hat{\mu}^t$ on B^t such that $\hat{\mu}^t(Y_{l,t}) = \pi_{l,t}$ and $\hat{\mu}^t(Y_{l,t} \cap Y_{l',t}) = 0$ for any $l \neq l'$.

By Theorem 1 in Kitamura and Stoye (2018), we know that the RUM-rationalizability of \mathcal{D} only depends on the values $\mu^t(Y_{l,t})$ at all patches $Y_{l,t}$. Since $\hat{\mu}^t$ satisfies $\hat{\mu}^t(Y_{l,t}) = \mu^t(Y_{l,t})$ and $\hat{\mu}^t(Y_{l,t} \cap Y_{l',t}) = \mu^t(Y_{l,t} \cap Y_{l',t}) = 0$ whenever $l \neq l'$, we conclude that $\hat{\mathcal{D}} = \{(\hat{\mu}_e^t, \bar{\mathbf{p}}^t)\}_{t \in T}$ is also RUM-rationalizable. Since $\hat{\mu}^1, \dots, \hat{\mu}^T$ are discrete probability distributions, there is a discrete probability distribution $\hat{\rho}$ on \mathcal{U} such that, for any $t \in T$ and any measurable subset $E \subseteq \mathbb{R}_+^K$, $\hat{\mu}_e^t(E) = \hat{\rho}(\{u \in \mathcal{U} : \arg \max_{\mathbf{e} \in \mathcal{E}^t} u(\mathbf{x}(\mathbf{e}, \bar{\mathbf{p}}^t)) \in E\})$. Let $\text{supp}(\hat{\rho}) = \mathcal{U} = \{u^1, \dots, u^N\}$.

For each $u^i \in \mathcal{U}$ and $t \in T$, let $\mathbf{e}^{i,t} = \arg \max_{\mathbf{e} \in \mathcal{E}^t} u^i(\mathbf{x}(\mathbf{e}, \bar{\mathbf{p}}^t))$. Then

$$\mathcal{O}^i = \{(\mathbf{x}(\mathbf{e}^{i,t}, \bar{\mathbf{p}}^t), \bar{\mathbf{p}}^t)\}_{t \in T}$$

is rationalized by u^i . In other words, the data set $\mathcal{E} = \{(\mathbf{e}^{i,t})_{i \in N}, \bar{\mathbf{p}}^t\}_{t \in T}$ is rationalizable with heterogenous preferences. By Theorem 1, there are real numbers $\lambda_1, \dots, \lambda_N > 0$ such that the stable price distribution $\mathcal{P} = \{\mathbf{p}^{i,t}\}_{(i,t) \in N \times T}$ with $\mathbf{p}^{i,t} = \lambda_i \bar{\mathbf{p}}^t$ rationalizes \mathcal{E} and satisfies the consistency condition $\sum_{i \in N} \lambda_i \hat{\rho}^i = 1$ (where $\hat{\rho}^i$ denotes the probability of u^i under $\hat{\rho}$). It follows that $\hat{\mathcal{D}} = \{(\hat{\mu}_e^t, \bar{\mathbf{p}}^t)\}_{t \in T}$ is RPM-rationalizable, where the probability of λ_i under η is $\hat{\rho}^i$. \square

References

- AFRIAT, S. N. (1967): “The construction of utility functions from expenditure data,” *International Economic Review*, 8, 67–77.
- BARTEN, A. P. (1964): “Consumer demand functions under conditions of almost additive preferences,” *Econometrica*, 1–38.

- BAYE, M. R., J. MORGAN, AND P. SCHOLTEN (2004): “Price dispersion in the small and in the large: Evidence from an internet price comparison site,” *The Journal of Industrial Economics*, 52, 463–496.
- BLOCK, D. AND J. MARSHAK (1960): “Random Orderings and Stochastic Theories of Responses,” *Contributions to Probability and Statistics*, ed. by I Olkin et al. Stanford, 2381–2391.
- BLUNDELL, R. AND A. LEWBEL (1991): “The information content of equivalence scales,” *Journal of Econometrics*, 50, 49–68.
- BORDALO, P., N. GENNAIOLI, AND A. SHLEIFER (2013): “Salience and consumer choice,” *Journal of Political Economy*, 121, 803–843.
- BRADY, R. L. AND J. REHBECK (2016): “Menu-Dependent Stochastic Feasibility,” *Econometrica*, 84, 1203–1223.
- BROWN, D. J. AND C. CALSAMIGLIA (2007): “The nonparametric approach to applied welfare analysis,” *Economic Theory*, 31, 183–188.
- BROWN, D. J. AND R. L. MATZKIN (1996): “Testable restrictions on the equilibrium manifold,” *Econometrica*, 1249–1262.
- BROWN, J., T. HOSSAIN, AND J. MORGAN (2010): “Shrouded attributes and information suppression: Evidence from the field,” *The Quarterly Journal of Economics*, 125, 859–876.
- BROWN, J. R. AND A. GOOLSBEE (2002): “Does the Internet make markets more competitive? Evidence from the life insurance industry,” *Journal of Political Economy*, 110, 481–507.
- CATTANEO, M. D., X. MA, Y. MASATLIOGLU, AND E. SULEYMANOV (2020): “A random attention model,” *Journal of Political Economy*, 128, 2796–2836.
- CERREIA-VIOGLIO, S., D. DILLENBERGER, P. ORTOLEVA, AND G. RIELLA (2019): “Deliberately stochastic,” *American Economic Review*, 109, 2425–45.
- CHETTY, R., A. LOONEY, AND K. KROFT (2009): “Salience and taxation: Theory and evidence,” *American Economic Review*, 99, 1145–77.
- DEATON, A. (1997): *The analysis of household surveys: a microeconomic approach to development policy*, The World Bank.
- DEB, R., Y. KITAMURA, J. K. QUAH, AND J. STOYE (2023): “Revealed price preference: theory and empirical analysis,” *The Review of Economic Studies*, 90, 707–743.

- DELLAVIGNA, S. AND U. MALMENDIER (2006): “Paying not to go to the gym,” *American Economic Review*, 96, 694–719.
- DIERKER, E., H. DIERKER, AND W. TROCKEL (1984): “Price-dispersed preferences and C1 mean demand,” *Journal of Mathematical Economics*, 13, 11–42.
- FALMAGNE, J.-C. (1978): “A representation theorem for finite random scale systems,” *Journal of Mathematical Psychology*, 18, 52–72.
- FUDENBERG, D., R. IJIMA, AND T. STRZALECKI (2015): “Stochastic choice and revealed perturbed utility,” *Econometrica*, 83, 2371–2409.
- GABAIX, X. (2014): “A sparsity-based model of bounded rationality,” *The Quarterly Journal of Economics*, 129, 1661–1710.
- GABAIX, X. AND D. LAIBSON (2006): “Shrouded attributes, consumer myopia, and information suppression in competitive markets,” *The Quarterly Journal of Economics*, 121, 505–540.
- GRANDMONT, J.-M. (1987): “Distributions of preferences and the” Law of Demand”,” *Econometrica*, 155–161.
- (1992): “Transformations of the commodity space, behavioral heterogeneity, and the aggregation problem,” *Journal of Economic Theory*, 57, 1–35.
- HEIDHUES, P. AND B. KŐSZEGI (2008): “Competition and price variation when consumers are loss averse,” *American Economic Review*, 98, 1245–68.
- KAPLAN, G., G. MENZIO, L. RUDANKO, AND N. TRACHTER (2019): “Relative price dispersion: Evidence and theory,” *American Economic Journal: Microeconomics*, 11, 68–124.
- KITAMURA, Y. AND J. STOYE (2018): “Nonparametric analysis of random utility models,” *Econometrica*, 86, 1883–1909.
- KOSZEGI, B. AND M. RABIN (2006): “A model of reference-dependent preferences,” *The Quarterly Journal of Economics*, 121.
- (2007): “Reference-dependent risk attitudes,” *American Economic Review*, 97, 1047–1073.
- LEWBEL, A. (1989): “Household equivalence scales and welfare comparisons,” *Journal of Public Economics*, 39, 377–391.
- LEWBEL, A. AND K. PENDAKUR (2017): “Unobserved preference heterogeneity in demand

- using generalized random coefficients,” *Journal of Political Economy*, 125, 1100–1148.
- LUCE, R. D. (1959): *Individual choice behavior: A theoretical analysis*, John Wiley and sons.
- MANZINI, P. AND M. MARIOTTI (2014): “Stochastic choice and consideration sets,” *Econometrica*, 82, 1153–1176.
- MAS-COLELL, A. AND W. NEUEFEIND (1977): “Some generic properties of aggregate excess demand and an application,” *Econometrica*, 591–599.
- MASATLIOGLU, Y., D. NAKAJIMA, AND E. Y. OZBAY (2012): “Revealed attention,” *The American Economic Review*, 102, 2183–2205.
- MATĚJKA, F. (2015): “Rigid pricing and rationally inattentive consumer,” *Journal of Economic Theory*, 158, 656–678.
- MAZUMDAR, T., S. P. RAJ, AND I. SINHA (2005): “Reference price research: Review and propositions,” *Journal of Marketing*, 69, 84–102.
- McFADDEN, D. (1973): “Conditional logit analysis of qualitative choice behavior,” *Frontiers in Econometrics*, 105–142.
- McFADDEN, D. AND M. K. RICHTER (1990): “Stochastic rationality and revealed stochastic preference,” *Preferences, Uncertainty, and Optimality, Essays in Honor of Leo Hurwicz*, Westview Press: Boulder, CO, 161–186.
- MUELLBAUER, J. (1980): “The estimation of the Prais-Houthakker model of equivalence scales,” *Econometrica*, 153–176.
- PRAIS, S. J. AND H. S. HOUTHAKKER (1971): *The analysis of family budgets*, vol. 4, CUP Archive.
- QUAH, J. K.-H. (1997): “The law of demand when income is price dependent,” *Econometrica*, 1421–1442.
- SORENSEN, A. T. (2000): “Equilibrium price dispersion in retail markets for prescription drugs,” *Journal of Political Economy*, 108, 833–850.
- STIGLER, G. J. (1961): “The economics of information,” *Journal of Political Economy*, 69, 213–225.
- VAN NIEUWERBURGH, S. AND P.-O. WEILL (2010): “Why has house price dispersion gone up?” *The Review of Economic Studies*, 77, 1567–1606.
- VARIAN, H. R. (1982): “The nonparametric approach to demand analysis,” *Econometrica*,

945–973.

——— (1988): “Revealed preference with a subset of goods,” *Journal of Economic Theory*, 46, 179–185.