Stochastic Revealed Preferences with Measurement Error[*]

Victor H. Aguiar[†]   Nail Kashaev[†]

vaguiar@uwo.ca   nkashaev@uwo.ca

This version: June 17, 2020
First version: October 11, 2017

Abstract   A long-standing question about consumer behavior is whether individuals' observed purchase decisions satisfy the revealed preference (RP) axioms of the utility maximization theory (UMT). Researchers using survey or experimental panel data sets on prices and consumption to answer this question face the well-known problem of measurement error. We show that ignoring measurement error in the RP approach may lead to overrejection of the UMT. To solve this problem, we propose a new statistical RP framework for consumption panel data sets that allows for testing the UMT in the presence of measurement error. Our test is applicable to all consumer models that can be characterized by their first-order conditions. Our approach is nonparametric, allows for unrestricted heterogeneity in preferences, and requires only a centering condition on measurement error. We develop two applications that provide new evidence about the UMT. First, we find support in a survey data set for the dynamic and time-consistent UMT in single-individual households, in the presence of *nonclassical* measurement error in consumption. In the second application, we cannot reject the static UMT in a widely used experimental data set in which measurement error in prices is assumed to be the result of price misperception due to the experimental design. The first finding stands in contrast to the conclusions drawn from the deterministic RP test of Browning (1989). The second finding reverses the conclusions drawn from the deterministic RP test of Afriat (1967) and Varian (1982). JEL classification numbers: C60, D10.

Keywords: rationality, utility maximization, time consistency, revealed preferences, measurement error.

1. Introduction

One well-known feature of consumer panel data sets—whether they are based on surveys, experiments, or scanners—is measurement error in prices or consumption.[1] This is of significant concern because measurement error is responsible for one important limitation of the standard revealed-preference (RP) framework. Specifically, RP tests tend to overreject the utility maximization theory (UMT). To the best of our knowledge, this paper proposes the first fully nonparametric statistical RP framework for consumer panel data sets in the presence of measurement error. We show that taking measurement error into account can significantly change the conclusions about the validity of the UMT in a given context. In the two applications we develop, we cannot reject the validity of the UMT, a finding which contradicts the conclusions of the deterministic RP framework.

Measurement error is the difference between the unobserved but true value of the variable of interest and its observed but mismeasured counterpart. If the UMT is valid, the corresponding RP conditions must be satisfied by the true prices and consumption. However, there is no reason to believe that either the mismeasured consumption or mismeasured prices that we usually observe are consistent with the RP conditions. A key concern for RP practitioners is that, in the presence of measurement error, a deterministic RP test may overreject the null hypothesis that the UMT is valid. We provide Monte Carlo evidence that this concern may be relevant in practice. Measurement error in consumption may arise in survey data due to misreporting, in experimental data due to trembling-hand errors, and in scanner data due to recording errors.[2] Measurement error in prices may arise in experimental data due to misperception errors and in scanner data due to unobserved coupons.[3]

Our methodology covers as special cases the static UMT, and the classic dynamic UMT with exponential discounting.[4] When we apply our methodology to a consumer panel survey of Spanish households and allow for measurement error in consumption, we find that the dynamic UMT with exponential discounting cannot be rejected for single-individual households. This first finding contradicts the conclusions of the deterministic RP test of Browning (1989). When we apply our framework to a widely used experimental data set (Ahn et al., 2014) and allow for measurement error in prices that arises due to misperception, we find that the static UMT cannot be rejected. This second finding is the opposite of the conclusions that must be drawn when one applies the deterministic RP test of Afriat (1967) and Varian (1982) to the same data set. Taken together, these findings suggest that the negative conclusions about the validity of the UMT drawn from the deterministic RP framework may not be robust to measurement error.

---

[1]See Mathiowetz et al. (2002), Echenique et al. (2011), Carroll et al. (2014), and Gillen et al. (2017).

[2]*Trembling-hand errors* are nonsystematic mistakes incurred by a subject when trying to implement a decision because she has difficulties with the interface of an experiment.

[3]*Misperception errors* are nonsystematic mistakes incurred by a subject because she misperceives information due to the experimental design.

[4]We also cover firm-cost minimization (Varian, 1984), dynamic rationalizability with quasi-hyperbolic discounting (Blow et al., 2017), homothetic rationalizability (Varian, 1985), quasilinear rationalizability (Brown & Calsamiglia, 2007), expected utility maximization (Diewert, 2012), and static utility maximization with nonlinear budget constraints (Forges & Minelli, 2009).

The leading solution for dealing with measurement error in the RP framework usually consists of perturbing (minimally) any observed individual consumption streams in order to satisfy the conditions of an RP test (Adams et al., 2014). However, this approach does not allow for standard statistical hypothesis testing. In particular, one cannot control the probability of erroneously rejecting a particular model when such a rejection could be an artifact of noisy measurements. Other works on RP with measurement error, such as the seminal contribution of Varian (1985), allow for statistical hypothesis testing but require knowledge of the distribution of measurement error; this may be impractical because this does not align with the nonparametric nature of the RP framework. In contrast, our procedure does not suffer from these issues.

Our main result is the formulation of a statistical test for the null hypothesis that a random data set of mismeasured prices and consumption is consistent with any given model that can be characterized by first-order conditions. Based on this test we provide and implement a general methodology to make out-of-sample predictions or counterfactual analyses with minimal assumptions (e.g., sharp bounds for average or quantile demand). Our approach takes advantage of the work of Schennach (2014) on Entropic Latent Variable Integration via Simulation (ELVIS) to provide a practical implementation of our test.

Our RP methodology is fully nonparametric and admits unrestricted heterogeneity in preferences. In addition, we require only a centering condition on the unobserved measurement error. The centering condition captures the application-specific knowledge we have about measurement error. Moreover, our framework is general enough to allow for (i) *nonclassical* measurement errors in consumption in survey environments, (ii) trembling-hand errors in experimental setups, (iii) misperception of prices due to experimental designs (price measurement error), and (iv) different forms of measurement error in prices in scanner environments.

In our first application in particular, we require that consumers be accurate, on average, in recalling and reporting their total expenditures. This assumption is compatible with systematic misreporting of consumption in surveys. For our second application, to an experimental data set, we require that errors in consumption or prices be centered around zero, which is compatible with trembling-hand and misperception errors in subjects' behavior. Measurement error in experimental data sets may arise because the experimental design may fail to elicit the intended choices of consumers. In such a case, classical measurement-error assumptions that allow for nonsystematic mistakes must be taken into account to ensure the external validity of the conclusions drawn from applying any RP test to this type of data set.

Our empirical contribution is to apply our methodology to a well-known consumer panel survey of single-individual and couples' households in Spain[5] in order to test for the dynamic UMT, and to reexamine the static UMT in a widely used experimental data set (Ahn et al., 2014); in doing this reexamination, we allow (separately) for trembling-hand errors and misperceived prices.

For the first application, we note that under the exponential discounting model, the consumer's time preferences are captured by a time-invariant discount factor and a time-invariant instantaneous utility. The main feature of this model is the time-consistency of the exponential discounting

---

[5]This data set has been used in Beatty & Crawford (2011), Adams et al. (2014), and Blow et al. (2017).

consumer. In other words, if the consumer prefers consumption bundle c at time t to x at time $t + k$, then she will always prefer c at time τ to x at time $\tau + k$. The exponential discounting model remains the workhorse of a large body of applied work in economics. However, several authors, such as Browning (1989), DellaVigna & Malmendier (2006), and Blow et al. (2017), have provided suggestive evidence against the validity of this model. Our methodology addresses, in a nonparametric fashion, the presence of measurement error in survey data, in order to examine the robustness of these findings.

We find support for exponential discounting behavior for single-individual households. This contrasts with the results of applying the deterministic methodology of Browning (1989) to the same sample. At the same time, in line with the findings of Blow et al. (2017) (who also use the deterministic methodology of Browning, 1989), we reject the null hypothesis of exponential discounting for the case of couples. When compared with the single-household evidence, these results suggest that time inconsistencies in consumer behavior in the couples' case arise due to preference aggregation.

In our second application, we apply our methodology to test for the validity of the classical static UMT. Since the work of Afriat (1967) and Varian (1982), researchers have used the deterministic RP framework to examine the validity of the UMT in experimental data sets. The experimental design proposed by Ahn et al. (2014) is particularly useful for this task since it provides a controlled environment with substantial price variation, variation that guarantees that the UMT has empirical bite.

When we apply the deterministic RP test to this experimental data set, we conclude that the UMT is rejected for most subjects. Nonetheless, the external validity of the conclusions drawn from the deterministic RP tests applied to these experimental data sets may be limited. One key reason for this is that the elicitation of consumer behavior may have been subject to measurement error. Gillen et al. (2017) argue that experimental elicitations of choices are subject to random variation in participants' perception and focus. Moreover, RP practitioners since Afriat (1967) have recognized that the deterministic RP test for the static UMT may be too demanding in the presence of imperfect devices for the elicitation of choices. Many researchers have studied how to allow for optimization mistakes in the RP framework and how to measure the intensity of any departure from rationality.[6] However, none of the existing approaches designed to introduce the possibility of mistakes in the RP framework has allowed for a fully nonparametric approach to doing standard statistical hypothesis testing. In our application, we allow for the possibility of nonsystematic mistakes by requiring that the measurement error in consumption or prices be mean zero.

We cannot reject the null hypothesis of the validity of the static UMT with misperception of prices. However, when we allow only for trembling-hand errors in consumption, we must strongly reject the static UMT. Our findings call into question the robustness of the deterministic RP test that is due to Afriat (1967) and Varian (1982) to measurement error in prices.

---

[6]See Afriat (1967), Varian (1990), and Echenique et al. (2011).

Outline

The paper proceeds as follows. Section 2 presents the first-order conditions approach to the deterministic RP methodology. Section 3 contains our statistical test. Section 4 presents a framework for recoverability and counterfactual analysis on the basis of our testing methodology. Section 5 provides an econometric framework for our methodology. Section 6 provides a guide specifying the centering condition in different environments. Section 7 implements our empirical test for the case of the dynamic UMT in a consumer panel-survey data set. Section 8 implements our methodology for the case of the static UMT in an experimental data set. Section 9 presents a brief discussion of related literature. Finally, we conclude in Section 10. All proofs can be found in Appendix A.

2. The Revealed-Preference Methodology and the First-Order Conditions Approach

The main objective of this section is to provide a brief summary in a united fashion of two very important deterministic consumer models and their RP characterization. In particular, we study the static UMT or rational model (R), and the dynamic UMT with exponential discounting (ED). These models are at the center of many applied and theoretical works. We show that they can be completely characterized by their first-order conditions in an RP fashion. All quantities used here are assumed to be measured precisely.

Let the consumption space be $\mathbb{R}^L_+ \setminus \{0\}$, where $L \in \mathbb{N}$ is the number of commodities.[7] Consider a consumer who is endowed with a utility function $u : \mathbb{R}^L_+ \to \mathbb{R}$ that is assumed to be concave, locally nonsatiated, and continuous. The consumer faces a sequence of decision problems indexed by $t \in \mathcal{T}$, where $\mathcal{T} = \{0, \cdots, T\}$, with a known and finite $T \in \mathbb{N}$. At each decision problem $t \in \mathcal{T}$, the consumer faces the price vector $p_t \in \mathbb{R}^L_{++}$.

Definition 1 (Static UMT, R-rationalizability)**.** A deterministic array $(p_t, c_t)_{t \in \mathcal{T}}$ is R-rationalizable (in a static sense) if there exists a concave, locally nonsatiated, and continuous function u, and some constants $y_t > 0$, $t \in \mathcal{T}$, such that the consumption bundle c_t solves:

$$\max_{c \in \mathbb{R}^L_+} u(c),$$

$$\text{s.t. } p'_t c = y_t,$$

for all $t \in \mathcal{T}$.

---

[7]We use $\mathbb{N}$ to denote the set of natural numbers. The expression $\mathbb{R}^L_+$ denotes the set of componentwise nonnegative elements of the L-dimensional Euclidean space $\mathbb{R}^L$, and $\mathbb{R}^L_+ \setminus \{0\}$ denotes the set of vectors $v \in \mathbb{R}^L_+$ that are distinct from zero ($v \neq 0$). Similarly, $\mathbb{R}^L_{++}$ denotes the set of componentwise positive elements of $\mathbb{R}^L_+$. The inner product of two vectors $v_1, v_2 \in \mathbb{R}^L$ is denoted by $v'_1 v_2$.

Next we focus on the dynamic UMT. We assume that an individual consumer has preferences over a stream of dated consumption bundles $(c_t)_{t \in \mathcal{T}}$, where $\mathcal{T} = \{0, \cdots, T\}$, $T \in \mathbb{N}$, and $c_t \in \mathbb{R}_+^L \setminus \{0\}$. (The number of goods, L, is kept the same across the time interval.) At time τ, the consumer chooses how much c_τ she will consume by maximizing

$$V_\tau(c) = u(c_\tau) + \sum_{j=1}^{T-\tau} d^j u(c_{\tau+j}),$$

subject to the linear budget or flow constraints shown here:

$$p_t' c_t - y_t + s_t - a_t = 0, \quad t = \tau, \ldots, T,$$

where $d \in (0, 1]$ is the discount factor; $p_t \in \mathbb{R}_{++}^L$ is the price vector as before; $y_t \in \mathbb{R}_{++}$ is income received by the individual at time t; s_t is the amount of savings held by the consumer at the end of time t; and a_t is the volume of assets held at the start of time t. The consumer invests all her savings. Moreover, the assets evolve according to the following law of motion:

$$a_t = (1 + r_t) s_{t-1},$$

where $r_{t+1} > -1$ is the interest rate that is accessible for the consumer. The holdings of assets in the last period ($t = T$) are set to be zero.

The intertemporal value function, $V_t : \mathbb{R}_+^{L \times (T-t+1)} \to \mathbb{R}_{++}$, represents the consumer preferences at a given time t. The components of this representation are the parameters of the model. First, $d \in (0, 1]$ is a scalar number that measures the degree of discount that the consumer gives to the future. Second, $u : \mathbb{R}_+^L \to \mathbb{R}_{++}$ is an instantaneous utility function that is assumed to be concave, locally nonsatiated, and continuous. The exponential discounting consumer is time-consistent, that is, she will solve the dynamic problem above the same way at any point of the time window.

Definition 2 (Dynamic UMT, ED-rationalizability). A deterministic array $(p_t, r_t, c_t)_{t \in \mathcal{T}}$ is ED-rationalizable if there exist a concave, locally nonsatiated, and continuous function u, a vector $(y_t)_{t \in \mathcal{T}} \in \mathbb{R}_{++}^{|\mathcal{T}|}$, and a scalar $a_0 \geq 0$ such that the consumption stream $(c_t)_{t \in \mathcal{T}}$ solves:

$$\max_{z \in \mathbb{R}_+^{L \times |\mathcal{T}|}} u(z_0) + \sum_{t=1}^{T} d^t u(z_t),$$

subject to

$$p_0' z_0 + \sum_{t=1}^{T} \frac{p_t' z_t}{\prod_{i=1}^{t}[1 + r_i]} = \sum_{t=1}^{T} \frac{y_t}{\prod_{i=1}^{t}[1 + r_i]} + a_0.$$

ED-rationalizability implicitly assumes perfect foresight (e.g., individuals know their future income) and homogeneity of consumers within a household (e.g., household members have the same discount factors). In Appendix E we show that our methodology covers two extensions of this model: the dynamic UMT with income uncertainty, and the collective model of Adams et al.

(2014).

2.1. The First-Order Conditions Approach

Now we establish that any consumer model $m \in \{R, ED\}$ can be completely characterized in terms of its first-order conditions with respect to (i) a concave, locally nonsatiated, and continuous utility function $u : \mathbb{R}_+^L \to \mathbb{R}$, (ii) the effective (or transformed) prices $\rho_t^m \in \mathbb{R}_{++}^L$, and (iii) restrictions on some constants $\lambda_t^m \in \mathbb{R}_{++}$ and $\delta_t^m \in (0, 1]$, interpreted as the marginal utility of income and the discount rate, respectively. We call this *the first-order conditions approach*. Observe that the utility function is model-independent, but the effective prices, the marginal utility of income, and the discount rate are not. We define the effective prices in Table 1.

Table 1 – Definition of ρ_t^m

m	R	ED
ρ_t^m	p_t	$p_t / \prod_{j=1}^t (1 + r_j)$

The following lemma summarizes the results in Browning (1989) for the exponential discounting case, and it is trivial for the static rationalizability case. Let $\nabla u(c_t)$ denote a supergradient of u at the point c_t.[8]

Lemma 1. *For any model* $m \in \{R, ED\}$, *a deterministic array* $(\rho_t^m, c_t)_{t \in \mathcal{T}}$ *is* m-*rationalizable if and only if there exists* $(u, (\lambda_t^m, \delta_t^m)_{t \in \mathcal{T}})$ *such that*

(i) $u : \mathbb{R}_+^L \to \mathbb{R}$ *is a concave, locally nonsatiated, and continuous utility function;*

(ii) $\delta_t^m \nabla u(c_t) \leq \lambda_t^m \rho_t^m$ *for every* $t \in \mathcal{T}$. *If* $c_{t,j} \neq 0$, *then* $\delta_t^m \nabla u(c_t)_j = \lambda_t^m \rho_{t,j}^m$, *where* $c_{t,j}$, $\nabla u(c_t)_j$, *and* $\rho_{t,j}$ *are the* j-*th components of* c_t, $\nabla u(c_t)$, *and* ρ_t, *respectively;*

(iii) $\lambda_t^R = \lambda_t > 0$ *and* $\delta_t^R = 1$ *for all* $t \in \mathcal{T}$;

(iv) $\lambda_t^{ED} = 1$ *and* $\delta_t^{ED} = d^t$, *where* $d \in (0, 1]$, *for all* $t \in \mathcal{T}$.

We want to highlight that while we focus on these two models for expositional and motivational purposes, our methodology is applicable to any model that can be characterized using the first-order conditions approach.

Remark 1. Lemma 1 allows for nondifferentiable utility functions. So, the supergradient of $u(c_t)$ may be set-valued. In this case one should read the condition $\delta_t^m \nabla u(c_t) \leq \lambda_t^m \rho_t^m$ as "there exists $\xi \in \nabla u(c_t)$ such that $\delta_t^m \xi \leq \lambda_t^m \rho_t^m$."

---

[8]The supergradient is $\nabla u(c_t) = \{\xi \in \mathbb{R}^L : u(c) - u(c_t) \leq \xi'(c - c_t), \forall c \in \mathbb{R}_+^L \setminus \{0\}\}$. Under differentiability, $\nabla u(c_t)$ is a gradient.

Remark 2. Lemma 1 specialized for ED-rationality implies that we do not need to observe consumers over all periods of their lives. The first-order conditions are the same irrespective of whether the consumer lives any finite number of time periods containing the observed time-window, or is alive only during the latter period.

2.2. The Elimination of a Latent Infinite-Dimensional Parameter

Since our objective is not to estimate but to test m-rationalizability, we will eliminate the utility function u from its characterization. We follow the theorists of RP to eliminate the latent infinite-dimensional parameters by exploiting their shape restrictions.

In particular, we follow Afriat (1967), Varian (1985), Browning (1989), and Rockafellar (1970) to formulate a result that eliminates the utility function u (an infinite dimensional parameter) from the first-order conditions. The cost of doing this is that we have to replace the first-order conditions by a set of inequalities that require only the concavity of u. As a result, the inequalities are exact and do not involve any form of approximation; this is an advantage compared to other nonparametric methods (e.g., sieves, kernel estimators) or the parametric approach used in many applied papers.

To formulate our result, we first recall the definition of the concavity of u.

Definition 3 (Concavity). A utility function u is said to be concave if and only if $u(\tilde{c}) - u(c) \leq \nabla u(c)'(\tilde{c} - c)$, for all $c, \tilde{c} \in \mathbb{R}^L_+ \setminus \{0\}$.

Remark 3. In Definition 3 we implicitly assume the existence of the supergradient of u. Since the supergradient may be set-valued, one should read the condition $u(\tilde{c}) - u(c) \leq \nabla u(c)'(\tilde{c} - c)$ as "$u(\tilde{c}) - u(c) \leq \xi'(\tilde{c} - c)$ for all $\xi \in \nabla u(c)$."

The nonparametric characterization of the m-rationalizability of observed consumption and prices without measurement error is captured by the following result.

Theorem 1. *For any $m \in \{R, ED\}$, the following are equivalent:*

(i) *The deterministic array $(\rho^m_t, c_t)_{t \in \mathcal{T}}$ is m-rationalizable.*

(ii) *There exist vectors $(\lambda^m_t)_{t \in \mathcal{T}}$, $(\delta_t)_{t \in \mathcal{T}}$, and a positive vector $(v_t)_{t \in \mathcal{T}}$ such that:*

$$v_t - v_s \geq \frac{\lambda^m_t}{\delta^m_t} \rho^{m\prime}_t (c_t - c_s),$$

with $\lambda^R_t = \lambda_t > 0$, $\delta^R_t = 1$, $\lambda^{ED}_t = 1$, and $\delta^{ED}_t = d^t$, where $d \in (0, 1]$, for all $t, s \in \mathcal{T}$.

Theorem 1 summarizes known results from the RP literature.[9] Observe that Theorem 1 has transformed the first-order conditions that depend on the infinite-dimensional u to a set of inequalities that depend only on a deterministic finite-dimensional array $(v_t, \delta^m_t, \lambda^m_t)_{t \in \mathcal{T}}$. Nonetheless,

---

[9]The proof is a consequence from the results in Afriat (1967), Varian (1985), and Browning (1989) taken together.

this set of conditions is satisfied if and only if we can find a utility function that satisfies the conditions in Lemma 1.[10] Checking the set of inequalities is a parametric problem and it tells us whether a consumption stream is m-rationalizable. This methodology is traditionally applied at the individual level in panel data sets, assuming that the data contains no measurement error. In the next section we extend the RP framework to a noisy or stochastic environment.

3. The Revealed-Preference Approach with Measurement Error

In this section, we introduce a new statistical notion of m-rationalizability (henceforth, s/m-rationalizability) with mismeasured consumption or prices, and provide a result similar to Theorem 1 in the presence of measurement error. From here on, we use boldface font to denote random objects and regular font for deterministic ones.

3.1. Statistical Rationalizability

We are interested in testing a statistical model of consumption such that each individual is an independent, identically distributed (i.i.d.) draw from some stochastic consumption rule. Note that by Lemma 1 the choice of a particular model m only affects the definition of the effective price, and the restrictions on the marginal utility of income and the discount rate. Henceforth, we fix some model such that the effective prices, and the restrictions on the marginal utility of income and the discount rate are known, and we omit the superscript m from the notation.

Using Lemma 1 as motivation, we directly define s/m-rationalizability as follows. Let $\boldsymbol{\rho}_t^* \in P_t^* \subseteq \mathbb{R}_{++}^L$ and $\mathbf{c}_t^* \in C_t^* \subseteq \mathbb{R}_+^L \setminus \{0\}$ denote random vectors of true effective prices and true consumption at time t, respectively.[11]

Definition 4 (s/m-rationalizability). A random array $(\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t \in \mathcal{T}}$ is s/m-rationalizable if there exists a tuple $(\mathbf{u}, (\boldsymbol{\lambda}_t, \boldsymbol{\delta}_t)_{t \in \mathcal{T}})$ such that

(i) $\mathbf{u}$ is a random, concave, locally nonsatiated, and continuous utility function;

(ii) $(\boldsymbol{\lambda}_t)_{t \in \mathcal{T}}$ is a positive random vector, interpreted as the marginal utility of income, supported on or inside a known set $\Lambda \subseteq \mathbb{R}_{++}^{|\mathcal{T}|}$;

---

[10]In contrast to Afriat's theorem for the static UMT, the assumption of concavity of the utility function is necessary in our framework. The reason is that concavity is testable in some cases that are different from static utility maximization (e.g., expected utility maximization, Polisson et al., 2020). Concavity guarantees that first-order conditions are necessary and sufficient in a wide variety of models beyond the static UMT. For the additional generality our result requires this additional constraint.

[11]For short, we use a.s. instead of "almost surely." We denote (i) the probability of an event A by the expression $\mathbb{P}(A)$; (ii) the indicator function by $\mathbb{1}(A) = 1$ when the statement A is true, otherwise it is zero; (iii) the mathematical expectation of any random vector $\mathbf{z}$ by the expression $\mathbb{E}[\mathbf{z}]$; (iv) the cardinality of a set $\mathcal{A}$ is given by the expression $|\mathcal{A}|$; and (v) the norm of a vector v is given by $\|v\|$.

(iii) $(\boldsymbol{\delta}_t)_{t\in\mathcal{T}}$ is a positive random vector, interpreted as time-varying discount factor, supported on or inside a known set $\Delta \subseteq (0,1]^{|\mathcal{T}|}$;

(iv) $\boldsymbol{\delta}_t\nabla\boldsymbol{u}(\boldsymbol{c}_t^*) \leq \boldsymbol{\lambda_t}\boldsymbol{\rho}_t^*$ a.s. for all $t \in \mathcal{T}$;

(v) For every $j = 1, \ldots, L$ and $t \in \mathcal{T}$, it must be the case that $\mathbb{P}\left(\boldsymbol{c}_{t,j}^* \neq 0, \boldsymbol{\delta}_t\nabla\boldsymbol{u}(\boldsymbol{c}_t^*)_j < \boldsymbol{\lambda_t}\boldsymbol{\rho}_{t,j}^*\right) = 0$, where $c_{t,j}^*$, $\rho_{t,j}^*$, and $\nabla u(c_t^*)_j$ denote the j-th components of c_t^*, ρ_t^*, and $\nabla u(c_t^*)$, respectively.

This definition means that for a given realization of (i) the utility function, (ii) the marginal utility of income, and (iii) the discount rate, the realized effective prices and the realized true consumption should fulfill the inequality $\delta_t\nabla u(c_t^*) \leq \lambda_t\rho_t^*$. This is a special case of the dynamic random utility model in which the preferences (captured by $\mathbf{u}$), the random discount factor (captured by $(\boldsymbol{\delta}_t)_{t\in\mathcal{T}}$), and the distribution of the marginal utility of income (captured by $(\boldsymbol{\lambda}_t)_{t\in\mathcal{T}}$) are drawn at some initial time for each consumer, and then are kept fixed over time.

Several consumer models can be characterized by their first-order conditions and by restrictions on the marginal utility of income, as we observed in Section 2. For instance, we define the statistical version of R-rationalizability or s/R-rationalizability by requiring that the support of the marginal utility of income be strictly positive (i.e., $\Lambda = \mathbb{R}_{++}^{|\mathcal{T}|}$), and the discount rate to be one (i.e., $\Delta = \{1\}^{|\mathcal{T}|}$). Similarly, we define s/ED-rationalizability by imposing that $\Lambda = \{1\}^{|\mathcal{T}|}$, and the support Δ be given by the restriction $\boldsymbol{\delta}_t = \mathbf{d}^t$, where $\mathbf{d}$ is a random variable supported on $(0,1]$. The effective prices in each case are to be defined according to Table 1.

Given the definition of s/m-rationalizability, we can now formulate the stochastic version of Theorem 1.

Lemma 2. *For a given random array $(\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t\in\mathcal{T}}$, the following are equivalent:*

(i) *The random array $(\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t\in\mathcal{T}}$ is s/m-rationalizable.*

(ii) *There exist positive random vector $(\mathbf{v}_t)_{t\in\mathcal{T}}$, $(\boldsymbol{\lambda}_t)_{t\in\mathcal{T}}$ supported on or inside Λ, and $(\boldsymbol{\delta}_t)_{t\in\mathcal{T}}$ supported on or inside Δ such that*

$$\mathbf{v}_t - \mathbf{v}_s \geq \frac{\boldsymbol{\lambda}_t}{\boldsymbol{\delta}_t}\boldsymbol{\rho}_t^{*\prime}(\mathbf{c}_t^* - \mathbf{c}_s^*) \quad \text{a.s.}, \quad \forall s, t \in \mathcal{T}.$$

Lemma 2 allows us to statistically test the s/m-rationalizability of $(\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t\in\mathcal{T}}$. However, as the following example demonstrates, any test based on this notion of rationalizability cannot differentiate between "almost" s/m-rationalizability and exact s/m-rationalizabilty (an issue first identified by Galichon & Henry, 2013).

Example 1 (Hyperbolic Discounting). Consider the case of a consumer who maximizes

$$V_\tau(c) = u(c_\tau) + \beta \sum_{j=1}^{T-\tau} d^j u(c_{\tau+j}),$$

where $\beta \in (0,1]$ is the present-bias parameter. It is easy to see that if $\beta \to 1$, then the consumption stream generated by this model is arbitrarily close to the ED-rationalizable behavior.

Example 1 presents a random choice rule that is not s/m-rationalizable, but is arbitrarily close to being s/m-rationalizable. In other words, there may exist sequences of random arrays that are not s/m-rationalizable that converge to random arrays that are s/m-rationalizable. That is why we need to extend the notion of the consistency of a data set that is characterized by s/m-rationalizability.

Definition 5 (Approximate s/m-rationalizability). We say that $(\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t \in \mathcal{T}}$ is **approximately consistent** with s/m-rationalizability if there exists a sequence of random variables $(\mathbf{v}_j', \boldsymbol{\lambda}_j', \boldsymbol{\delta}_j')' \in \mathbb{R}_+^{|\mathcal{T}|} \times \Lambda \times \Delta$, $j = 1, 2, \ldots$, such that

$$\mathbb{P} \left(\mathbb{1} \left(\mathbf{v}_{j,t} - \mathbf{v}_{j,s} \geq \frac{\boldsymbol{\lambda}_{j,t}}{\boldsymbol{\delta}_{j,t}} \boldsymbol{\rho}_t^{*\prime} [\mathbf{c}_t^* - \mathbf{c}_s^*] \right) = 1 \right) \to_{j \to +\infty} 1,$$

for all $s, t \in \mathcal{T}$.

3.2. Introducing Measurement Error

Theorem 1 and Lemma 2 provide testable implications of s/m-rationalizability. These implications depend solely on the distribution of $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_t)_{t \in \mathcal{T}}$, $\boldsymbol{\delta} = (\boldsymbol{\delta}_t)_{t \in \mathcal{T}}$ and $\mathbf{v} = (\mathbf{v}_t)_{t \in \mathcal{T}}$. The usual approach to testing s/m-rationalizability would amount to solving a (non)linear programming problem corresponding to Theorem 1 at the level of individual consumers. However, this common practice does not work any more in the presence of measurement error. When **true consumption** or **true prices** are measured erroneously, we observe not $\mathbf{c}_t^*$ or $\boldsymbol{\rho}_t^*$ but rather perturbed versions of them. (See Section 6 for a discussion of the reasons measurement error in consumption and prices arises in survey, experimental, and scanner data sets.)

Define the **measurement error** $\mathbf{w} = (\mathbf{w}_t)_{t \in \mathcal{T}} \in W$ as the difference between reported consumption and prices, $\mathbf{c} = (\mathbf{c}_t)_{t \in \mathcal{T}}$ and $\boldsymbol{\rho} = (\boldsymbol{\rho}_t)_{t \in \mathcal{T}}$; and true consumption and prices, $(\mathbf{c}_t^*)_{t \in \mathcal{T}}$ and $(\boldsymbol{\rho}_t^*)_{t \in \mathcal{T}}$. That is,

$$\mathbf{w}_t = \begin{pmatrix} \mathbf{w}_t^c \\ \mathbf{w}_t^p \end{pmatrix},$$

where $\mathbf{w}_t^c = \mathbf{c}_t - \mathbf{c}_t^*$ and $\mathbf{w}_t^p = \boldsymbol{\rho}_t - \boldsymbol{\rho}_t^*$ for all $t \in \mathcal{T}$.

It is important to note that we *define* the measurement error. We do not make any assumptions about how the difference between observed and true quantities arises (i.e., we allow for measurement error to be multiplicative or additive).[12] Moreover, we **do not** need to assume that measurement error is independent of other variables, independent within time periods, or independent across goods.

By Lemma 2 we can immediately conclude that the observed $\mathbf{x} = (\boldsymbol{\rho}_t, \mathbf{c}_t)_{t \in \mathcal{T}}$ can be s/m-rationalized if and only if there exist $(\boldsymbol{\lambda}_t, \boldsymbol{\delta}_t, \mathbf{v}_t, \mathbf{w}_t)_{t \in \mathcal{T}}$, with $(\boldsymbol{\lambda}_t)_{t \in \mathcal{T}}$ supported on or inside Λ, and

---

[12]Formally, this makes the support W depend on the support of both the observed and the true quantities. For simplicity we omit this dependency from the notation.

$(\boldsymbol{\delta}_t)_{t \in \mathcal{T}}$ supported on or inside Δ such that

$$\mathbf{v}_t - \mathbf{v}_s \geq \frac{\boldsymbol{\lambda}_t}{\boldsymbol{\delta}_t}(\boldsymbol{\rho}_t - \mathbf{w}_t^p)'(\mathbf{c}_t - \mathbf{c}_s + \mathbf{w}_s^c - \mathbf{w}_t^c) \quad \text{a.s.}, \forall s, t \in \mathcal{T}.$$

However, we know that without restrictions on the distribution of measurement error, RP tests have no power. That is, there always exists a measurement error $\mathbf{w}$ such that the observed $\mathbf{x}$ is consistent with s/m-rationalizability. Hence, we require access to additional validation information about measurement error. The source of measurement error is different in different applications. That is why in this section we formulate a general restriction on the measurement error distribution that can be tailored for a given empirical application.

Recall that $\mathbf{x} \in X$ denotes observed quantities. Let $\mathbf{e} = (\boldsymbol{\lambda}', \boldsymbol{\delta}', \mathbf{v}', \mathbf{w}')' \in E|X$ denote the vector of latent random variables, supported on or inside the conditional support $E|X$. We say that a mapping $g_M : X \times E|X \to \mathbb{R}^{d_M}$ is a *measurement error moment*. We only require the following condition on measurement error.

Assumption 1 (Centered Measurement Error). *(i) The random vector $\mathbf{e}$ is supported on or inside the known support $E|X$. (ii) There exists a known measurement error moment $g_M : X \times E|X \to \mathbb{R}^{d_M}$ such that*

$$\mathbb{E}\left[g_M(\mathbf{x}, \mathbf{e})\right] = 0.$$

The choice of function g_M depends on the application and the assumptions the researcher is willing to make on the basis of the available knowledge about the nature of measurement error. In Section 6 we provide examples of moment conditions in data sets that are often used in the RP literature. The objects of interest for us are measurement error in consumption, expenditure, and prices.

4. Recoverability and Counterfactuals

Varian (1982, 1984) exploits the connections between empirical content and counterfactuals. In particular, Varian (1982) seems to be the first to think of nonparametric counterfactual analysis as specification testing.[13] Following these ideas, our formulation of rationalizability allows us to answer questions about the recoverability of, and counterfactual predictions for, different objects of interest.

In Section 4.1 we show how to recover different quantities of interest (e.g., average true consumption at a given $t = \tau$) from the s/m-rationalizable data set. In Section 4.2 we demonstrate how to make out-of-sample predictions for expected consumption in a way that is analogous to Blundell et al. (2014). In the presence of measurement error, distributional information about

---

[13]Recent work building on these connections includes Blundell et al. (2003) and Allen & Rehbeck (2019) in demand analysis, and Norets & Tang (2014) in the analysis of dynamic binary choice models.

the primitives of the model of interest is inevitably lost. Hence, we cannot apply the traditional approach proposed by Varian (1982) to recover preferences and to do counterfactual analysis on an individual basis. Instead, we use this section to pose questions about the primitives of the model at the level of the population.

4.1. Recoverability

Assume that $\mathbf{x} = (\boldsymbol{\rho}_t, \mathbf{c}_t)_{t \in \mathcal{T}}$ can be s/m-rationalized and Assumption 1 holds. Suppose that there is a finite-dimensional parameter of interest $\theta_0 \in \Theta$, where Θ is a compact subset of the Euclidean space. The parameter of interest is related to the model via the user-specified moment condition

$$\mathbb{E}\left[g_R(\mathbf{x}, \mathbf{e}; \theta_0)\right] = 0 \in \mathbb{R}^{d_R}.$$

The function g_R can take different forms depending on the different questions the user wants to answer. We provide some examples here.

Example 2 (Expected True Consumption and Expected True Consumption Change)**.** If θ_0 is the expected true consumption at $t = \tau$, then $g_R(x, e; \theta_0) = c_\tau - w_\tau^c - \theta_0$. If θ_0 is the expected difference in true consumption at $t = \tau + 1$ and $t = \tau$, then $g_R(x, e; \theta_0) = c_{\tau+1} - w_{\tau+1}^c - c_\tau + w_\tau^c - \theta_0$.

The user may also be interested in testing the joint null hypothesis that (i) the consumer is s/ED-rationalizable and (ii) the random discount factor distribution has certain properties.

Example 3 (Average Random Discount Factor)**.** The user may be interested in testing whether the average value of the random discount factor is equal to a certain fixed value, in which case $g_R(x, e; \theta_0) = d - \theta_0$.

In addition, our framework allows us to have, as a special case, latent random variables with flexible support.

Example 4 (Support of the Random Discount Factor)**.** The user may be interested in whether the random time-discount factor $\mathbf{d}$ has a support on or inside $[\theta_{01}, \theta_{02}] \subseteq (0, 1]$. Then, for $\theta_0 = (\theta_{01}, \theta_{02})'$, one can define $g_R(x, e; \theta_0) = \mathbb{1}\left(\theta_{01} \le d \le \theta_{02}\right) - 1$.

4.2. Counterfactual Out-of-Sample Predictions

We consider a counterfactual situation in which the user is given an out-of-sample effective random price vector $\boldsymbol{\rho}_{T+1}^*$ (supported in $\mathbb{R}_{++}^L$), a data set $\mathbf{x} = (\boldsymbol{\rho}_t, \mathbf{c}_t)_{t \in \mathcal{T}}$ such that Assumption 1 holds, and she then asks two related questions. First, the user wants to know if there exists a counterfactual random consumption vector $\mathbf{c}_{T+1}^*$ such that the augmented random array $\{(\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t \in \mathcal{T}}, (\boldsymbol{\rho}_{T+1}^*, \mathbf{c}_{T+1}^*)\}$ is approximately s/m-rationalizable, where $\mathbf{c}_t^* = \mathbf{c}_t - \mathbf{w}_t^c$ and $\boldsymbol{\rho}_t^* = \boldsymbol{\rho}_t - \mathbf{w}_t^p$.

Second, if the answer to the first question is affirmative, then the user will be interested in constructing confidence sets for some counterfactual finite-dimensional parameter $\theta_0 \in \Theta$. The parameter θ_0 satisfies the user-specified moment condition

$$\mathbb{E}\left[g_C\left((\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t \in \mathcal{T}}, \mathbf{c}_{T+1}^*; \boldsymbol{\rho}_{T+1}^*, \theta_0 \right) \right] = 0 \in \mathbb{R}^{d_C}.$$

Both questions can be answered simultaneously with our characterization of s/m-rationalizability. Observe that the answer to the first question is negative if the random array $(\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t \in \mathcal{T}}$ is not s/m-rationalizable. In contrast, if the random array $(\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t \in \mathcal{T}}$ is s/m-rationalizable, then the counterfactual exercise is equivalent to checking that the counterfactual price/consumption distribution is simultaneously compatible with s/m-rationalizability and the user-specified moment condition. Formally, to answer both questions, we define what it means for a random array $(\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t \in \mathcal{T}}$ to be counterfactually rationalizable (C/m-rationalizability) for a given $\boldsymbol{\rho}_{T+1}^*$, θ_0, and g_C.

Definition 6 (C/m-rationalizability). For a given $\boldsymbol{\rho}_{T+1}^*$, g_C, and θ_0, a random array $(\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t \in \mathcal{T}}$ is approximately C/m-rationalizable if there exist $\mathbf{c}_{T+1}^*$ such that

(i) The augmented random array $\{(\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t \in \mathcal{T}}, (\boldsymbol{\rho}_{T+1}^*, \mathbf{c}_{T+1}^*)\}$ is approximately s/m-rationalizable;

(ii) $\mathbb{E}\left[g_C((\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t \in \mathcal{T}}, \mathbf{c}_{T+1}^*; \boldsymbol{\rho}_{T+1}^*, \theta_0) \right] = 0.$

Observe that if a random array $(\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t \in \mathcal{T}}$ is C/m-rationalizable for a given $(\boldsymbol{\rho}_{T+1}^*, \mathbf{c}_{T+1}^*)$ and θ_0, then it is also s/m-rationalizable. However, the opposite is not always true.

We can apply Lemma 2 to Definition 6 and get an extended system of the RP inequalities coupled with the counterfactual moment conditions g_C. Moreover, we can define an identified set for counterfactual parameter values Θ_0. Formally,

$$\Theta_0 = \left\{ \theta_0 \in \Theta \; : \; (\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t \in \mathcal{T}} \text{ is approximately C/m-rationalizable given } \boldsymbol{\rho}_{T+1}^*, \theta_0, \text{ and } g_C \right\}.$$

We highlight that our framework can accommodate additional support restrictions on the counterfactual objects. A classical support constraint is a target out-of-sample expenditure level (i.e. $\boldsymbol{\rho}_{T+1}^{*\prime} \mathbf{c}_{T+1}^* = 1$ a.s.) as in Varian (1982). We omit these constraints from our discussion to simplify exposition.

Example 5 (Average Varian Support Set). We consider a moment

$$g_C((\rho_t^*, c_t^*)_{t \in \mathcal{T}}, c_{T+1}^*; \rho_{T+1}^*, \theta_0) = c_{T+1}^* - \theta_0,$$

with $\theta_0 \in \Theta = \mathbb{R}_+^L \setminus \{0\}$ as a hypothesized average-demand vector. Thus, Θ_0 is the Average Varian Support Set. Given $\boldsymbol{\rho}_{T+1}^*$ this set describes the bounds on the average demand that is compatible with the s/m-rationalizability of the random array $(\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t \in \mathcal{T}}$.

Example 6 (Quantile Varian Support Set). For s/R-rationalizability, we can consider the following moment condition:

$$g_C((\rho_t^*, c_t^*)_{t \in \mathcal{T}}, c_{T+1}^*; \rho_{T+1}^*, \theta) = \mathbb{1}\left(\rho_{T+1}^{*\prime} c_{T+1}^* \leq \bar{e}_c \right) - \phi,$$

14

where $\theta = (\bar{e}_c, \phi)' \in \mathbb{R}_{++} \times [0,1]$, $\bar{e}_c$ is a fixed ϕ-quantile of the counterfactual expenditure distribution. Next we can define the ϕ-quantile Varian Support Set:

$$\left\{ c \in \mathbb{R}_+^L \setminus \{0\} \; : \; \rho'_{T+1} c = \bar{e}_c, \; (\bar{e}_c, \phi)' \in \Theta_0 \right\}.$$

This set describes the bounds of the counterfactual demand for a given $\boldsymbol{\rho}^*_{T+1}$ and ϕ-quantile of $\mathbf{u}(\mathbf{c}^*_{T+1})$ that is compatible with s/R-rationalizability.

Some counterfactual questions (e.g. the Average Varian Support Set) lead to convex identified sets Θ_0.

Proposition 1. *If the parameter space Θ is convex and g_C is such that*

$$g_C((\rho^*_t, c^*_t)_{t \in \mathcal{T}}, c^*_{T+1}; \rho^*_{T+1}, \theta) = \tilde{g}((\rho^*_t, c^*_t)_{t \in \mathcal{T}}, c^*_{T+1}; \rho^*_{T+1}) - A((\rho^*_t, c^*_t)_{t \in \mathcal{T}}; \rho^*_{T+1})' \theta$$

for some $\tilde{g}$ and A, then Θ_0 is convex.

Proposition 1 imposes two restrictions on g_C: (i) additive separability between $\mathbf{c}^*_{T+1}$ and θ; (ii) affinity of the moment condition in θ. In Section 5 we provide a framework to construct confidence sets for the counterfactual parameters by means of the test inversion. Convexity of Θ_0 can substantially simplify the computation of the confidence sets since one does not need to conduct test inversion at every point of the parameter space.

5. Econometric Framework

In Sections 3.2 and 4.2 we showed how testing, recoverability, and counterfactuals in RP models with measurement error can be framed in the form of moment conditions. In this section we recast the empirical content of the RP inequalities in a form amenable to statistical testing. To simplify the exposition we will focus on testing s/m-rationalizability in the presence of measurement error (g_M and g_I only).

5.1. Characterization of the Model via Moment Conditions

First, we write a set of moment conditions that will summarize the empirical content of s/m-rationalizability. Recall that $\mathbf{x} \in X$ denotes observed quantities and $\mathbf{e} = (\boldsymbol{\lambda}', \boldsymbol{\delta}', \mathbf{v}', \mathbf{w}')' \in E|X$ denote the vector of latent random variables. The support $E|X$ depends on the fixed supports Λ and Δ that characterizes the particular model of interest. We use $\mathcal{P}_X$, $\mathcal{P}_{E,X}$, and $\mathcal{P}_{E|X}$ to denote the set of all probability measures defined over the support of $\mathbf{x}$, $(\mathbf{e}', \mathbf{x}')'$, and $\mathbf{e}|\mathbf{x}$, respectively. (Recall

that the boldface font letters denote random objects.) Define the following moment functions:

$$g_{I,t,s}(\mathbf{x}, \mathbf{e}) = \mathbb{1}\left(\mathbf{v}_t - \mathbf{v}_s - \frac{\boldsymbol{\lambda}_t}{\boldsymbol{\delta}_t}(\boldsymbol{\rho}_t - \mathbf{w}_t^p)'[\mathbf{c}_t - \mathbf{w}_t^c - \mathbf{c}_s + \mathbf{w}_s^c] \geq 0\right) - 1, \quad t \neq s \in \mathcal{T},$$

$$g(\mathbf{x}, \mathbf{e}) = (g_I(\mathbf{x}, \mathbf{e})', g_M(\mathbf{x}, \mathbf{e})')'.$$

We have $k = |\mathcal{T}|^2 - |\mathcal{T}|$ and $q = d_M$ moment functions which correspond to inequality conditions (g_I) and the measurement error centering conditions (g_M), respectively.[14] Define $\mathbb{E}_{\mu \times \pi}[g(\mathbf{x}, \mathbf{e})] = \int_X \int_{E|X} g(x, e) d\mu d\pi$, where $\mu \in \mathcal{P}_{E|X}$ and $\pi \in \mathcal{P}_X$.

Theorem 2. *The following are equivalent:*

(i) *A random vector* $\mathbf{x} = (\boldsymbol{\rho}_t, \mathbf{c}_t)_{t \in \mathcal{T}}$ *is approximately* s/m*-rationalizable such that Assumption 1 holds.*

(ii)

$$\inf_{\mu \in \mathcal{P}_{E|X}} \|\mathbb{E}_{\mu \times \pi_0}[g(\mathbf{x}, \mathbf{e})]\| = 0,$$

where $\pi_0 \in \mathcal{P}_X$ *is the observed distribution of* $\mathbf{x}$.

Theorem 2 establishes the equivalence between (i) s/m-rationalizability with the centered measurement error condition and (ii) a system of moment conditions. In other words, the observed consumption pattern, captured by the random array $(\boldsymbol{\rho}_t, \mathbf{c}_t)_{t \in \mathcal{T}}$, can be s/m-rationalized under the restrictions on measurement error if and only if there exists a distribution of latent variables conditional on observables that satisfies the RP inequalities with probability 1, for the given supports Λ and Δ.

Our notion of s/m-rationalizability makes clear that when one is dealing with measurement error, no RP test can decide whether a finite sample is consistent with model m. We can decide only that the data set is asymptotically consistent with the model as the sample size goes to infinity. Moreover, even asymptotically, there is no way to differentiate between the notion of approximate s/m-rationalizability and the notion of exact s/m-rationalizability. Nonetheless, we can do traditional hypothesis testing and decide at a fixed significance level whether we reject the null hypothesis of (approximate) model m consistency under Assumption 1 for a given sample. Conceptually, our notion of rationalizability corresponds to the extended notion of an identified set in Schennach (2014).

Note that the test is not yet formally established. We have a set of latent random variables $\mathbf{e}$ distributed according to an unknown $\mu \in \mathcal{P}_{E|X}$. This problem can be solved nonparametrically using the Entropic Latent Variable Integration via Simulation (ELVIS) of Schennach (2014). The main advantage of the ELVIS approach is that it allows us to formulate a test that can be implemented in panel data sets suffering from measurement error of the type described only in terms of observables.

---

[14]If in addition, the user includes moments g_R or g_C, then $q = d_M + d_R$ or $q = d_M + d_C$, respectively.

5.2. ELVIS and Its Implications for Testing and Inference

We start this section by showing how the nonparametric results of Theorem 2 can be used to construct a set of (equivalent) parametric maximum-entropy moment conditions using Schennach (2014). Next, we provide a semi-analytic solution to the these conditions. Finally, we propose a procedure to test for s/m-rationalizability.

Following Schennach (2014), we define the maximum-entropy moment as follows.

Definition 7 (Maximum-Entropy Moment)**.** The maximum-entropy moment of the moment $g(x, \cdot)$, for a fixed x, is

$$h(x; \gamma) = \frac{\int_{e \in E|X} g(x, e) \exp(\gamma' g(x, e)) d\eta(e|x)}{\int_{e \in E|X} \exp(\gamma' g(x, e)) d\eta(e|x)},$$

where $\gamma \in \mathbb{R}^{k+q}$ is a nuisance parameter, and $\eta \in \mathcal{P}_{E|X}$ is an arbitrary user-input distribution function supported on $E|X$ such that $\mathbb{E}_{\pi_0} [\log \mathbb{E}_\eta [\exp(\gamma' g(\mathbf{x}, \mathbf{e}))|\mathbf{x}]]$ exists and is twice continuously differentiable in γ for all $\gamma \in \mathbb{R}^{k+q}$.

Note that

$$\left\{ d\eta^*(\cdot|x; \gamma) = \frac{\exp(\gamma' g(x, \cdot)) d\eta(\cdot|x)}{\int_{e \in E|X} \exp(\gamma' g(x, e)) d\eta(e|x)}, \ \gamma \in \mathbb{R}^{k+q} \right\}$$

is a family of exponential conditional probability measures. Thus, the maximum-entropy moment h is the marginal moment of the function g, at which the latent variable has been integrated out using one of the members from the above exponential family. The importance of the maximum-entropy moment is captured in the following result.

Theorem 3. *The following are equivalent:*

(i) *A random array $\mathbf{x} = (\boldsymbol{\rho}_t, \mathbf{c}_t)_{t \in \mathcal{T}}$ is approximately* s/m-*rationalizable such that Assumption 1 holds.*

(ii)

$$\inf_{\gamma \in \mathbb{R}^{k+q}} \|\mathbb{E}_{\pi_0} [h(\mathbf{x}; \gamma)]\| = 0,$$

where $\pi_0 \in \boldsymbol{P}_X$ is the observed distribution of $\mathbf{x}$.

The idea behind Theorem 3 is that if there exists a distribution that satisfies the moment condition, then there must be another distribution that (i) belongs to a particular finite-dimensional exponential family and (ii) satisfies the same moment condition. Since we are only interested in the existence of the former distribution, instead of searching over the set of all possible distributions we are going to only search over this "smaller" exponential family.

We emphasize that Theorem 3 provides both *necessary and sufficient* conditions for the observed data to be (approximately) s/m-rationalizable. This represents an important gain in power with respect to any of the averaging-based tests of RP models that are usually used in the presence of measurement error.

High-level technical assumptions can ensure that the sequence of random latent variables that approximates model m converges to a proper random variable. Thus, this limiting random variable would ensure (i) that the infimum in Theorem 3 is attained, and (ii) that the notion of approximate rationalizability collapses to exact rationalizability. However, this obscures the fact that any assumption made in that direction has no testable implications.

The remarkable advantage of applying the results of Schennach (2014) to the RP approach is that it marginalizes out the latent random variables. More importantly, we have a robust statistical framework with which to test our models in the presence of measurement error. In particular, we have not made any strong distributional assumptions about $\boldsymbol{\lambda}$, $\boldsymbol{\delta}$ or $\mathbf{u}$ (the heterogeneous tastes). The only restrictions are the concavity assumption on the utility function, and a centering condition on measurement error.[15]

Remark 4. Theorem 3 does not imply that the distribution of the latent variables (or their support) is point-identified. In fact, it will always be set-identified.

Remark 5. The maximum-entropy moment implicitly depends on the choice of the user-specified distribution η. However, η does not have any effects on the set of values $\mathbb{E}_{\pi_0}\left[h(\mathbf{x};\cdot)\right]$ or its sample analogue can take. That is, the choice of η will not affect the value $\inf_{\gamma\in\mathbb{R}^{k+q}}\left\|\mathbb{E}_{\pi_0}\left[h(\mathbf{x};\gamma)\right]\right\|$ takes both asymptotically and in finite samples. The choice of η affects *only* the nuisance parameter (γ) value. See Remark 2.3 in Schennach (2014) for further details.

5.3. Semi-analytic Solution for the Maximum-entropy Moment

One can directly employ the maximum-entropy moment in Theorem 3 to test model m. However, doing so is potentially problematic. One possible concern is the fact that the number of maximum-entropy moments corresponding to the (g_I) conditions, $k = |\mathcal{T}|^2 - |\mathcal{T}|$, grows quadratically with $|\mathcal{T}|$. Moreover, γ_0, the nuisance parameter value at which infimum is achieved, may be set-identified when unbounded (e.g., some of the components of γ_0 may be equal to infinity[16]), which would therefore lead to nonstandard testing procedures.

Here we show that there exists a semi-analytic solution to the optimization problem where every component of γ_0 that corresponds to the RP inequality constraints is equal to $+\infty$, and every component of γ_0 that corresponds to the measurement error centering constraint is finite and unique. Thus, for testing purposes (under the null hypothesis of model m), we can minimize an objective function over a parameter space of lower dimensionality.

Assumption 2. *(Nondegeneracy) There exist two subsets of $E|X$, E' and E'', with a positive measure, such that componentwise the measurement error moment is such that $\sup_{e\in E'} g_M(\mathbf{x},e) < 0 < \inf_{e\in E''} g_M(\mathbf{x},e)$ with positive probability.*

---

[15]At this point, we can use as an alternative the methodology presented by Ekeland et al. (2010) to deal with latent variables in our moment conditions.

[16]By "equal to infinity" we mean that the infimum is achieved along the sequence of γ that diverges to infinity along some coordinates.

Assumption 3. *(Bounded support) The random array* $\mathbf{x} = (\boldsymbol{\rho}_t, \mathbf{c}_t)_{t \in \mathcal{T}}$ *has a bounded support.*

Assumption 2 rules out cases in which there is no measurement error and allows us to a have a unique minimizer of the objective function. It can be relaxed since our methodology still works for cases without measurement error, but in those cases, it is preferable to use the equivalent deterministic RP benchmark. Assumption 3 is made to simplify the analysis and can be replaced by tail restrictions on the distribution of $\mathbf{x}$.

Note that the user-specified distribution η should obey the same restrictions as the unknown distribution of latent $\mathbf{e}$. Thus, we impose the following restrictions on η:

Definition 8 (User-specified distribution). Almost surely in $\mathbf{x}$, the user-specified distribution $\eta(\cdot|\mathbf{x})$ satisfies all of the following:

   (i) The set $\tilde{E}|X = \{e \in E|X : g_I(\mathbf{x}, e) = 0\}$ has a positive measure under $\eta(\cdot|\mathbf{x})$.

   (ii) There exist two subsets of $\tilde{E}|X$, E' and E'', with a positive measure under $\eta(\cdot|\mathbf{x})$, such that componentwise $\sup_{e \in E'} g_M(\mathbf{x}, e) < 0 < \inf_{e \in E''} g_M(\mathbf{x}, e)$.

   (iii) For every finite $\gamma_M \in \mathbb{R}^q$,

$$\int_{E|X} \|g_M(\mathbf{x}, e)\|^2 \exp(\gamma_M' g_M(\mathbf{x}, e)) d\eta(e|\mathbf{x}) < \infty.$$

The first condition in Definition 8 requires that the support of η allows the inequalities to be satisfied. The second and third conditions are regularity conditions. This is a definition and not an assumption, as we can always construct an allowable η.[17] We are ready to present our main result.

Theorem 4. *Given a user-specified measure η that satisfies the three conditions in Definition 8, the following are equivalent:*

   (i) *A random array* $\mathbf{x} = (\boldsymbol{\rho}_t, \mathbf{c}_t)_{t \in \mathcal{T}}$ *is approximately consistent with* s/m-*rationalizability such that Assumptions 1, 2, and 3 hold.*

   (ii) *For any sequence* $\{\gamma_{I,l}\}_{l=1}^{+\infty}$ *that componentwise diverges to* $+\infty$,

$$\lim_{l \to +\infty} \min_{\gamma_M \in \mathbb{R}^q} \left\| \mathbb{E}_{\pi_0} \left[h(\mathbf{x}; (\gamma_{I,l}', \gamma_M')') \right] \right\| = 0. \tag{1}$$

   The sequence of minimizers of (1), $\{\gamma_{M,l}\}$, *converges to some finite* $\gamma_{0,M}$ *that does not depend on* $\{\gamma_{I,l}\}_{l=1}^{+\infty}$.

   (iii)

$$\min_{\gamma_M \in \mathbb{R}^q} \left\| \mathbb{E}_{\pi_0} \left[\tilde{h}_M(\mathbf{x}; \gamma_M) \right] \right\| = 0, \tag{2}$$

   where

$$\tilde{h}_M(x; \gamma) = \frac{\int_{e \in E|X} g_M(x, e) \exp(\gamma' g_M(x, e)) \mathbb{1}\left(g_I(x, e) = 0\right) d\eta(e|x)}{\int_{e \in E|X} \exp(\gamma' g_M(x, e)) \mathbb{1}\left(g_I(x, e) = 0\right) d\eta(e|x)}.$$

---

[17]Schennach (2014) provides a generic construction that can be used here.

Moreover, the minimizer of (2) is finite, and is equal to $\gamma_{0,M}$.

The intuition behind Theorem 4 is that the RP inequalities presented here restrict only the conditional support of the latent variables (including the measurement error). Hence, given the support restrictions captured by the RP inequalities, only the centering condition comes in the form of moments.[18]

Theorem 4 substantially simplifies the conclusion of Theorem 3. First, we need to minimize the objective function over a much smaller parameter space ($\mathbb{R}^q$ instead of $\mathbb{R}^{k+q}$). Thus, the problem becomes computationally tractable. Second, if the data is consistent with s/m-rationalizability, then the minimizer, $\gamma_{0,M}$, has to be *finite and unique*. Finally, to compute $\tilde{h}_M(x, \gamma_M)$ in applications, one may need to use Markov Chain Monte Carlo (MCMC) methods by sampling from η. Theorem 4 implies that it suffices to sample from $d\tilde{\eta}(\cdot|x) = \mathbb{1}\left(g_I(x, \cdot) = 0 \right) d\eta(\cdot|x)$. The straightforward way to sample from $\tilde{\eta}$ is to sample from η and then reject the draw if it does not satisfy the RP inequalities captured by $\mathbb{1}\left(g_I(x, \cdot) = 0 \right)$. The last part usually amounts to solving a linear program for R-rationalizability. For the case where this rejection sampling is not efficient, in Appendix C, we describe a new hit-and-run algorithm to sample from $\tilde{\eta}(\cdot|x)$ directly.[19] This approach is particularly useful for testing s/ED-rationalizability in one of our applications.

5.4. Testing

Theorem 4 provides moment conditions that are necessary and sufficient for the data $\{\mathbf{x}_i\}_{i=1}^n = \{(\boldsymbol{\rho}_{t,i}, \mathbf{c}_{t,i})_{t \in \mathcal{T}}\}_{i=1}^n$ (where n is the sample size), to be approximately consistent with s/m-rationalizability. Now, define the following sample analogues of the maximum-entropy moment and its variance:

$$\hat{\tilde{h}}_M(\gamma) = \frac{1}{n} \sum_{i=1}^n \tilde{h}_M(\mathbf{x}_i, \gamma);$$

$$\hat{\tilde{\Omega}}(\gamma) = \frac{1}{n} \sum_{i=1}^n \tilde{h}_M(\mathbf{x}_i, \gamma)\tilde{h}_M(\mathbf{x}_i, \gamma)' - \hat{\tilde{h}}_M(\gamma)\hat{\tilde{h}}_M(\gamma)'.$$

Let Ω^- denote the generalized inverse of the matrix Ω. The testing procedure we propose is due to Schennach (2014) and is based on this test statistic:

$$\mathrm{TS}_n = n \inf_{\gamma \in \mathbb{R}^q} \hat{\tilde{h}}_M(\gamma)' \hat{\tilde{\Omega}}^-(\gamma) \hat{\tilde{h}}_M(\gamma).$$

Assumption 4. *The data $\{\mathbf{x}_i\}_{i=1}^n$ is i.i.d.*

---

[18]If Assumption 2 is violated for some component j of g_M (e.g., $\sup_{E'} g_{M,j}(x, e) \geq 0$ a.s. for all $E' \subseteq E|X$), then since $\mathbb{E}\left[g_{M,j}(\mathbf{x}, \mathbf{e})\right] = 0$ it has to be the case that $g_{M,j}(\mathbf{e}, \mathbf{x}) = 0$ a.s.. Thus, $g_{M,j}$ becomes a condition similar to g_I that can be replaced by a support restriction.

[19]The classical hit-and-run algorithm is an efficient MCMC method that generates uniform draws from a convex polytope. It initiates inside the polytope and proceeds by randomizing directions and then advancing a random distance while remaining inside the set.

Theorem 5. *Suppose Assumptions 1, 2, 3, and 4 hold. Then under the null hypothesis that the data is approximately consistent with* s/m-*rationalizability, it follows that*

$$\lim_{n \to \infty} \mathbb{P}\left(\mathrm{TS}_n > \chi^2_{q,1-\alpha}\right) \le \alpha,$$

for every $\alpha \in (0, 1)$.

If, moreover, the minimal eigenvalue of the variance matrix $\mathbb{V}[\tilde{h}_M(\mathbf{x}, \gamma)]$ is uniformly, in γ, bounded away from zero and the maximal eigenvalue of $\mathbb{V}[\tilde{h}_M(\mathbf{x}, \gamma)]$ is uniformly, in γ, bounded from above, then, under the alternative hypothesis that the data is not approximately consistent with s/m-*rationalizability, it follows that*

$$\lim_{n \to \infty} \mathbb{P}\left(\mathrm{TS}_n > \chi^2_{q,1-\alpha}\right) = 1.$$

We conclude this section by noting that we can speed up computations of the test statistic by obtaining an initial guess for the minimizer of the objective function efficiently. We make use of the fact that, although the objective function $\hat{\tilde{h}}_M(\cdot)'\hat{\Omega}^-(\cdot)\hat{\tilde{h}}_M(\cdot)$ may have several local minima, the quadratic form $\hat{\tilde{h}}_M(\cdot)'B\hat{\tilde{h}}_M(\cdot)$, where B is any conformable positive definite matrix, has a unique global minimum and has no other local minima.[20]

5.5. Confidence Sets for Parameters of Interest

The above testing procedure can be modified for construction of confidence sets for parameter θ_0 from Section 4.2. In particular, recall that one just needs to extend the set of the original moment conditions (the centering condition g_M and the RP inequalities g_I) by g_R or g_C and add (if needed) extra RP inequalities that correspond to $(\boldsymbol{\rho}^*_{T+1}, \mathbf{c}^*_{T+1})$. As in Section 5.4 we can then define $\mathrm{TS}_n(\theta)$ as the value of the test statistic computed for a fixed value of θ. Under assumptions similar to those of Theorem 5, the confidence set for θ_0 can be obtained by inverting $\mathrm{TS}_n(\theta_0)$. That is, the $(1-\alpha)$-confidence set for θ_0 is

$$\{\theta_0 \in \Theta : \mathrm{TS}_n(\theta_0) \le \chi^2_{q_{\mathrm{ext}},1-\alpha}\},$$

where $\chi^2_{q_{\mathrm{ext}},1-\alpha}$ denotes the $(1-\alpha)$ quantile of the χ^2 distribution with q_{ext} degrees of freedom $(\chi^2_{q_{\mathrm{ext}}})$. The number q_{ext} is determined by the number of the non RP moment conditions in the extended system that are not the support restrictions. Note that we do not pretest for s/m-rationalizability in order to construct the confidence set for θ_0. If the data set is not s/m-rationalizable, then the confidence set will be empty asymptotically.

---

[20]Additional details about the computational aspects of our methodology can be found in Appendix C.

6. Measurement Error in Different Data Sets

Recall that to apply our methodology we need measurement-error restrictions that come in the form of moments g_M (Assumption 1). These moments capture the knowledge the user has about measurement error in a particular data set. In this section we provide examples of such centering conditions in different data sets relevant for the RP literature.

6.1. Measurement Error in Survey Data Sets

Measurement error in surveys may arise because of errors due to the respondent, the interviewer, or the survey design itself (Carroll et al., 2014). Recall mistakes, social desirability, and recording errors, among other factors, may cause measurement error in surveys (see Meyer et al., 2015). Household surveys usually measure expenditure across different goods, but do not measure consumption or prices directly. Prices are sometimes price indexes constructed by a national statistics agency.[21] Consumption is generated by dividing reported consumption expenditures in surveys by the observed prices aggregated to a category of goods (e.g., Blundell et al., 2003, Adams et al., 2014, and Kitamura & Stoye, 2018). There is evidence that individuals, in either recall or self-reported surveys, can systematically overreport or underreport their expenditure on different categories of goods (Mathiowetz et al., 2002, Carroll et al., 2014, Pistaferri, 2015). In sum, both prices and consumption may have measurement error.

Consumption quantities are generated by dividing expenditures for a given good by its price index. This means that in the survey environment, measurement error in consumption is often *nonclassical* because it may combine measurement error of expenditures and prices in a nonlinear way (Attanasio et al., 2014). Nonetheless, we will argue that average total expenditures is well-measured in some surveys.

Abildgren et al. (2018) find that mean total expenditures are usually well-measured in the Danish interview-based household budgets survey (DHS, 2015). They use administrative registry information of the same households interviewed in the survey of interest to conclude that there are no statistically significant differences between the mean total expenditure in Danish households as reported in the registry data set and in the household survey. In addition, Kolsrud et al. (2017), using administrative registry consumption data from Sweden, also finds that mean household expenditures match those computed on the basis of household surveys. One potential caveat to this evidence is that registry data itself may be mismeasured. However, direct evidence on household expenditures from retail data shows that total household expenditure error in registry data has zero mean (Baker et al., 2018).[22]

---

[21]For instance, in data sets like the ones used in Blundell et al. (2003), Adams et al. (2014), and Kitamura & Stoye (2018), price indexes are collected through direct observation in a given market and then merged with the expenditure survey data sets.

[22]This pattern of measurement error in expenditure is also indirectly supported by evidence on measurement error in income. For instance, Angel et al. (2018), using data from Austria, show that the difference in household income

Based on this evidence for measurement error, the analyst can impose the following condition. For all $t \in \mathcal{T}$ it must be the case that

$$\mathbb{E}\left[\boldsymbol{\rho}_t'\mathbf{c}_t\right] = \mathbb{E}\left[\boldsymbol{\rho}^{*\prime}_t\mathbf{c}_t^*\right]. \tag{3}$$

Equation (3) provides a measurement error moment that allows for nonclassical measurement error in consumption in the survey environment. In particular, it requires only that total expenditure measurement error be nonsystematic. Equation (3) implies that measurement error in consumption does not alter the mean value of total expenditures. In other words, it captures the idea that consumers, on average, may remember the total expenditure level better than the actual details.

Other surveys, such as the Spanish Continuous Expenditure Survey (1985-1997), used in Beatty & Crawford (2011), Adams et al. (2014), and in one of our applications, have documented both over- and underreporting of income. Unfortunately, there is no validation data on expenditure itself (the object of our interest). However, Gradín et al. (2008) report that for the time period of interest the evolution of the mean household income and the mean household expenditure for Spain are very similar.

Spanish, Swedish, and Danish households self-report expenditures for a number of good categories in their respective household surveys. We see no reason to believe that Spanish subjects will behave differently from Danish ones, given the similar format of the surveys. For that reason, we believe it is reasonable to assume that the mean expenditure is usually well measured in the Spanish household survey.

However, in practice, the condition in Equation (3) may not be restrictive enough (i.e., the joint hypothesis of this assumption and s/m-rationality may not be rejected in general). We remind the reader that the measurement error moment is not the result of a modeling decision but should capture additional information that is usually obtained from validation exercises. Consequently, if we know too little about measurement error, then our data will not suffice to ensure a reasonable statistical power. A natural solution to this problem is to provide additional restrictions on measurement error by collecting new information about it. These additional restrictions will necessarily apply to specific settings or data sets.

A particular case of the condition captured by Equation (3), with the additional support constraint that prices are perfectly measured (i.e., $\boldsymbol{\rho}_t^* = \boldsymbol{\rho}_t$ a.s. for all $t \in \mathcal{T}$), is:

Assumption 1.1. *(Mean-Budget Neutrality for Survey Data Sets) For every $t \in \mathcal{T}$,*

$$\mathbb{E}\left[\boldsymbol{\rho}_t'\mathbf{w}_t^c\right] = 0.$$

Assumption 1.1 requires that measurement error in consumption be orthogonal to the effective prices. We maintain the assumption that prices are well measured in the rest of this subsection. We have three reasons for doing this. (i) This assumption is a common requirement in most of the

---

reported from surveys and that from registry data displays mean-reverting behavior, as high income levels are underreported and low income levels are overreported. This evidence is consistent with the literature on measurement error in earnings (Bound et al., 2001).

work dedicated to the study of consumption decisions in households using survey data. In fact, we follow several works interested in modeling consumption in assuming that total expenditures and prices are measured better (i.e., with a higher signal-to-noise ratio) than the *generated* consumption variable.[23] (ii) It is empirically plausible that prices are better measured than consumption. In fact, there is evidence (Castillo et al., 1999, Guerrero de Lizardi, 2008) that the price index in Spain (which we use in one of our applications) is usually well measured with an estimated upward bias of the total price index of 0.07 percent for the time window $1981 - 1991$ (Castillo et al., 1999), which is arguably small.[24] For context, the estimated bias for the US in the well-known Boskin report was 0.40 percent for a similar time period (Boskin et al., 1998). In contrast, there is agreement that consumption measurement in surveys is very noisy (Alan et al., 2018).[25] (iii) We show that s/m−rationalizability is robust to measurement error when it is local (Appendix D.3). Hence, we will focus on the main source of measurement error that is consumption in the case of surveys. In sum, in the self-reported expenditures survey panel data sets, we argue that the *main* source of measurement error is in the consumption variable.

The mean-budget neutrality assumption is still compatible with nonclassical measurement error in consumption, such as $\mathbf{w}_{l,t}^c \leq 0$ a.s. ($\mathbf{w}_{l,t}^c \geq 0$ a.s.) for some good category l and all decision tasks t. That is, this condition allows agents on average to underreport (overreport) consumption on some categories of goods as long as they on average equally overreport (underreport) consumption on other categories of goods. Mean-budget neutrality will fail if measurement error in expenditures is systematic (i.e., every consumer overreports expenditures or underreports expenditures simultaneously).

We want to highlight the fact that despite its simplicity, Assumption 1.1 is a generalization of commonly held parametric assumptions about measurement error in the study of consumption. Here we provide some examples.

Example 7 (Multiplicative Measurement Error)**.** When one studies consumption of a single good across time (estimation of the Euler equation), one usually assumes that consumption measurement error is multiplicative (Alan et al., 2018). Formally, $\mathbf{c}_t = \mathbf{c}_t^* \boldsymbol{\epsilon}_t$, where $\boldsymbol{\epsilon}_t$ can be assumed to be independent of, or at least orthogonal to, a set of instruments. For instance, we can assume that $\boldsymbol{\epsilon}_t$ is independent of true consumption conditional on effective prices and with mean 1. In that case, Assumption 1.1 holds. Alternatively, we can assume $\boldsymbol{\epsilon}_t$ to be independent of effective prices and true consumption, which implies Assumption 1.1.

Example 8 (Additive Measurement Error)**.** Consider a case in which $\mathbf{c}_{t,l} = \mathbf{c}_{t,l}^* + \epsilon_{t,l}$ for $l = 1, \ldots, L$; and where $\epsilon_{t,l} \sim TN_{[-a,a]}(0, \sigma)$ for $l = 1, \ldots, L-1$ (from a truncated normal with variance σ and bounds $[-a, a]$ for some positive $a > 0$) such that $\mathbf{c}_{t,l} \geq 0$ a.s.. Note that this example is similar to

---

[23]For examples of papers assuming that prices are perfectly measured and that measurement error is present only in consumption, see Cochrane (1991), Ventura (1994), Carroll (2001), Ludvigson & Paxson (2001), Alan et al. (2009), Adams et al. (2014), Carroll et al. (2014), Toda & Walsh (2015), and Alan et al. (2018).

[24]The bias is defined as the difference between the Laspeyres price index as used in Spain and the ideal Fisher price index (Diewert, 1998). This is usually known as the substitution or aggregation bias.

[25]For instance, Alan & Browning (2010) estimate that roughly 80 percent of the variation in consumption growth rate is due to measurement error (this is for the PSID in the US).

the one used in Varian (1985) with constraints to impose the nonnegativity of consumption. Note that Assumption 1.1 holds because measurement error is independent of prices and mean zero.

We conclude this section by noting that even if prices are measured correctly, there is a potential source of error coming from the fact that price indexes are the result of aggregation of commodities into categories. This aggregation implies that price indexes do not exactly reflect actual prices faced by consumers. Thus, the budget sets faced by the consumers may be different from the ones implied by price indexes.[26] This problem is common to all demand analysis using survey data sets. Nonetheless, separability of the utility function or homotheticity within a category are examples of conditions under which commodity aggregation produces consumption and price indexes that are consistent with utility maximization (Jerison & Jerison, 1994, Lewbel, 1996). If one is willing to impose such conditions, then a rejection of utility maximization, even in the presence of commodity aggregation error, means that rationality can be rejected at the disaggregated level.[27] In addition, a nonrejection of the UMT means that price indexes and consumption behave *as if* they are rational. That is, the substitution patterns observed in the data after commodity aggregation and their relation with relative changes in price-indexes can be summarized by the first-order condition approach.[28]

6.2. Measurement Error in Consumption or Prices in Experimental Data Sets

Measurement error in consumption or prices in experimental data sets arise due to difficulties in eliciting the intended choices of the experimental subjects. Indeed, the experimental elicitation of choice may be subject to random variation due to (i) the subject's misperception of some elements of the task, (ii) the level of understanding of the experimental design, and (iii) nonsystematic mistakes in implementing intended choice. In general, there is an imperfect relation between the elicited choice and the intended choice behavior that the experiment tries to measure (Gillen et al., 2017).

We consider two sources of measurement error in the experimental environment in the context of the budget allocation task due to Choi et al. (2014) and Ahn et al. (2014) (which we use in our second application). The first one is the possibility of measurement error in consumption due to trembling-hand errors and the second one is measurement error in prices due to misperception.

We capture the relation between the true intended choices and the measured choices with the following assumption that allows for trembling-hand errors.

---

[26]We thank a referee for pointing this out.

[27]Recent work by Sato (2020) provides nonparametric RP evidence in favor of weak separability and validity of several widely used price indexes for food and beverages categories using data from Japan. In addition, a survey by Shumway & Davis (2001) finds evidence in favor of the conditions in Lewbel (1996).

[28]Another possible interpretation of our analysis in survey environments is that we are testing for utility-maximizing behavior in a budgeting problem. In the budgeting problem, the consumer allocates her income into different good categories to maximize her utility at that level of aggregation. Then, the consumer maximizes utility within different categories. This nested maximization problem is similar to mental accounting (Thaler, 1985) in its modern interpretation by Montgomery et al. (2019).

Assumption 1.2. *(Trembling-Hand Errors for Experimental Data Sets) For all $t \in \mathcal{T}$, it must be the case that*

$$\mathbb{E}\left[\mathbf{c}_t\right] = \mathbb{E}\left[\mathbf{c}_t^*\right].$$

Assumption 1.2 requires that measurement error in consumption be nonsystematic or equivalently centered around zero. Formally, for all $t \in \mathcal{T}$ it must be that

$$\mathbb{E}\left[\mathbf{w}_t^c\right] = 0.$$

We use data from Kurtz-David et al. (2019) to provide direct empirical evidence supporting Assumption 1.2. In a setup with two goods Kurtz-David et al. (2019) implement a motor task by asking their subjects to click on a visual target located on a budget line. Their experimental interface is exactly the same as in Choi et al. (2014). Kurtz-David et al. (2019) record the target coordinates and the actual coordinates that subjects choose on a screen using a mouse. The difference between the target coordinates and the actual choice is the trembling-hand error. Using this data set, we verified that the average trembling-hand error is statistically not different from zero.[29] We believe this is reasonable evidence in favor of Assumption 1.2 for budget allocation experiments.

Failure to account for the possibility of subjects' misperception of the experimental task may affect the elicitation of true consumer behavior. In particular, the experimental design of Choi et al. (2014) and Ahn et al. (2014) relies on a graphical representation of the budget hyperplane to elicit consumption choices. It is therefore a visual task and an economic task at the same time. We now consider the possibility of misperception of prices that can be thought of as measurement error in prices due to experimental design.

Assumption 1.3. *(Misperception of Prices for Experimental Data Sets) For all $t \in \mathcal{T}$, it must be the case that*

$$\mathbb{E}\left[\boldsymbol{\rho}_t\right] = \mathbb{E}\left[\boldsymbol{\rho}_t^*\right].$$

Assumption 1.3 relaxes the implicit requirement in the deterministic RP framework that subjects perceive the budget constraints without any distortion. Note that the graphical experimental device used by Choi et al. (2014) and Ahn et al. (2014) may make it difficult for consumers to correctly understand the true prices. We believe that it is desirable to have an RP test of the validity of the UMT that is robust to the misperception of prices when the distortion is (i) attributable to the experimental design and (ii) nonsystematic.

We use data from Kurtz-David et al. (2019) to provide empirical evidence supporting Assumption 1.3. In contrast to the case of trembling-hand errors, there are no direct measurements on price misperception. However, Kurtz-David et al. (2019) collected data on visual misperception of coordinates that we can use to test Assumption 1.3 indirectly. In particular, we find evidence for a special case of this assumption that is compatible with the misperceived price being $\boldsymbol{\rho}_{l,t} = \boldsymbol{\rho}_{l,t}^*/\boldsymbol{\epsilon}_{l,t}$,

---

[29]Formally, we did a t-test of Assumption 1.2 across 23 subjects for 27 trials. The null hypothesis is not rejected at the 5 percent significance level for all except for one good at one trial, but the null hypothesis is not rejected in any instance at the 1 percent significance level. We did not perform a joint test because there are missing values.

for all goods $l \in \{1, \cdots, L\}$, where $\epsilon_{l,t}$ captures misperception. Assumption 1.3 holds if the misperception error has mean 1 and is independent of true prices.

In a setup with two goods, Kurtz-David et al. (2019) implement a visual task by giving their subjects a particular numerical target $z_t^* = (z_{1,t}^*, z_{2,t}^*)$. Next, the experimenters ask their subjects to locate this coordinate on a budget line. As a result, the point $z_t = (z_{1,t}^* \epsilon_{1,t}, z_{2,t}^* \epsilon_{2,t})$ is observed, where we treat ϵ_t as the misperception error. Since prices can be computed from two observed points on the same budget line, misperceived prices will be $\rho_{l,t} = \rho_{l,t}^*/\epsilon_{l,t}$ for $l \in \{1, 2\}$.[30] A sufficient condition for Assumption 1.3 is that the multiplicative error is independent of true prices and that the mean perception error across individuals is 1. We cannot check the independence condition from this validation data. Nonetheless, we cannot reject the null hypothesis that the mean (multiplicative) perception error is equal to 1 across the 25 subjects in this experiment, for each of the 27 trials (at the 5 percent significance level).[31]

In the budget allocation tasks implemented by Choi et al. (2014) and Ahn et al. (2014), the subject is forced to choose a point on the budget line. Given that in these experimental environments the total income or wealth is known, the support of measurement error $E|X$ must be such that $\boldsymbol{\rho}_t^{*\prime} \mathbf{c}_t^* = \boldsymbol{\rho}_t^\prime \mathbf{c}_t$ a.s..[32] Note that in both the trembling-hand errors and in the misperception case we have $d_M = |\mathcal{T}| \cdot L$. The number of centering conditions grows with the number of commodities and the number of decision tasks.

6.3. Measurement Error in Scanner Data Sets

Even though our main focus is on survey and experimental data sets, our methodology can be used in other types of data sets with their corresponding measurement error constraints. Here we provide a quick overview of the case of scanner data sets because of its relevance for RP practitioners.

Scanner consumer panel data sets are usually of high quality; thus, measurement error concerns may be less important. However, in some cases, like in the well-known Nielsen Homescan Scanner Data Set (NHS), there is evidence of measurement error in prices due to classical misreporting but also due to imputation done in the data collection stage (Einav et al., 2010).[33]

Einav et al. (2010) use validation data from a retailer and compare it to a subsample of the self-reported NHS (for 2004); they conclude that consumption is measured rather precisely with roughly 90 percent of all records being exactly reported. On the other hand, prices are likely to be recorded with an upward bias. In fact, prices are measured precisely in only 50 percent of records

---

[30]Formally, we assume that the perception error realization is the same when the subject observes a second point given the same budget line.

[31]We performed a t-test analogous to the case of the trembling-hand error.

[32]In our application, we want to test separately for the validity of the static UMT together with (i) trembling-hand error in consumption or (ii) misperception of prices. Hence, we will assume that prices are perfectly measured in the former case (i.e., $\mathbf{w}_t^p = 0$ a.s.), while assuming that consumption is perfectly measured in the latter case (i.e, $\mathbf{w}_t^c = 0$ a.s.).

[33]Similarly, in the older Stanford Basket Scanner Data Set (Echenique et al., 2014), there could be measurement error in prices due to unobserved coupons or discounts.

in the sample of interest. The sample mean of the logarithm of prices[34] in the NHS seems to be slightly above the same quantity for the validation data.[35] Statistically, the difference between the logarithm of prices in the total sample of interest in the NHS and in the validation data is not different from zero (at the 5 percent significance level), as reported in Einav et al. (2010).[36]

We believe that the findings of Einav et al. (2010) support the conclusion that consumption measurement error in the NHS can be treated as local perturbations (see Appendix D.3). Hence, we assume that consumption is measured precisely (i.e., $\mathbf{w}_t^c = 0$ a.s.) and impose the following centering condition for the NHS:

Assumption 1.4. *(Centered Differences in Price Measurement Error) For all $t, s \in \mathcal{T}$ and all $l = 1, 2, \ldots, L$, it must be the case that*

$$\mathbb{E}\left[\log \mathbf{p}_{l,t}^*\right] = \mathbb{E}\left[\log \mathbf{p}_{l,t}\right].$$

The above centering condition for measurement error allows for nonsystematic overreporting or underreporting of the logarithm of prices on average. Assumption 1.4 implies that the number of measurement error conditions is $d_M = |\mathcal{T}| \cdot L$.

The source of measurement error in the NHS is very different from the one in survey data. In the NHS, households report quantities and prices, while in the survey data set, households report expenditures. For this reason survey data sets are usually used for aggregate analysis (at the level of the category of goods), in cases where price indexes computed by the national statistical agencies are available. As a result, the main mismeasured object is consumption. In contrast, scanner data sets have rich disaggregated information on prices and quantities. But misreported or imputed prices lead to measurement error. For that reason, we impose our centering condition on the main source of measurement error in each of these cases.

6.4. Other Forms of Measurement Error: Moment Inequalities and Instruments

Our methodology also allows for imposing moment inequality restrictions on measurement error. Following Schennach (2014), we can handle conditions of the type

$$\mathbb{E}\left[g_M(\mathbf{x}, \mathbf{e})\right] \geq 0$$

---

[34]The mean is taken across all records in the time window of interest.

[35]It seems that Nielsen generates the price in cases in which they have access to retailer-level price data. The reason for the overreporting is that this imputation process ignores the discounts that consumers may get (Einav et al., 2010).

[36]In the subsample of records from the NHS in which the consumers did not get a sales discount, the measurement error in the logarithm of prices seems classical (i.e., centered around zero and symmetric). In contrast, in the subsample of records in which the consumers got a sales discount, the distribution of measurement error in the logarithm of prices has a fat right tail (Einav et al., 2010). As a result, the total measurement error in the logarithm of prices is not symmetric.

by introducing an additional slack positive random vector $\mathbf{s} = (\mathbf{s}_j)_{j \in \{1, \cdots, d_M\}}$ such that

$$\mathbb{E}\left[g_M(\mathbf{x}, \mathbf{e}) - \mathbf{s}\right] = 0.$$

Moment inequalities may be particularly useful for taking into account bounds on measurement-error averages (e.g., $\mathbb{E}\left[\mathbf{w}_t^p\right] \leq 0$ for all $t \in \mathcal{T}$). Imposing support constraints on measurement error (e.g., by rounding: $|\mathbf{w}_{l,t}| \leq 1/2$ a.s.) can be handled automatically by setting the support $E|X$ appropriately.

Measurement error moments can also capture exclusion/orthogonality restrictions. In other words, the analyst may have information or be willing to assume that a particular observed variable is orthogonal to measurement error in consumption, prices or expenditures. The literature of demand estimation in both the static and dynamic setups, which use moments, usually handles measurement error through exclusion restrictions.[37]

Note that Assumption 1.1 can be understood as an orthogonality restriction between prices and consumption measurement error (when prices are observed without error). Consider an instrumental variable $\mathbf{z}_t$ supported on $\mathbb{R}^L$ that is orthogonal to the consumption measurement error; this can be expressed as:

$$\mathbb{E}\left[\mathbf{z}_t' \mathbf{w}_t^c\right] = 0,$$

for all $t \in \mathcal{T}$. For Assumption 1.1, the variable $\mathbf{z}_t = \mathbf{p}_t$ a.s.. These additional restrictions will increase the power of the test simply because they contain more information about measurement error.

6.5. Asymptotic Power for s/ED-Rationalizability: Illustrative Example

In previous sections we discussed different centering conditions that can be imposed on measurement error in different data sets. Since our characterization of s/m-rationality is necessary and sufficient, we have asymptotic power of one against the alternate hypothesis of inconsistency with s/m-rationalizability. Nonetheless, we still have to show that the alternative hypothesis space is nonempty. That is, we need to make sure that the restrictions on measurement error we provide are falsifiable. Here we build an illustrative example for s/ED-rationality and the centering condition captured by Assumption 1.1 with no price measurement error. Similar examples can be built for other moments used in our applications (see Appendix D). We also provide simulation evidence in Appendix B.2.

Consider the random array $(\boldsymbol{\rho}_t, \mathbf{c}_t)_{t \in \mathcal{T}}$ such that $\mathcal{T} = \{0, 1\}$, $\boldsymbol{\rho}_0 = (1, 1)'$, $\boldsymbol{\rho}_1 = (2, 2)'$, $\mathbf{c}_0 = (1, 1)'$, and $\mathbf{c}_1 = (2, 2)'$ a.s.. This data set requires that prices and consumption be the same for all consumers in the population. Moreover, it is easy to see that in deterministic terms ED-rationality fails. We will now show that s/ED-rationality also fails with appropriate centering conditions. Assume towards a contradiction that this random array is s/ED-rationalizable. Then, there have to be

---

[37]See Lewbel & Pendakur (2009) and Alan et al. (2018).

random variables $\mathbf{d} \in (0, 1]$, $\{\mathbf{w}_t^c\}_{t=0,1}$, and $\{\mathbf{v}_t\}_{t=0,1}$ such that

$$\mathbf{v}_1 - \mathbf{v}_0 \geq \frac{\rho_1'}{\mathbf{d}}(c_1 - c_0) - \frac{\rho_1'}{\mathbf{d}}(\mathbf{w}_1 - \mathbf{w}_0) \text{ a.s.},$$

$$\mathbf{v}_0 - \mathbf{v}_1 \geq \rho_0'(c_0 - c_1) - \rho_0'(\mathbf{w}_0 - \mathbf{w}_1) \text{ a.s.}.$$

Combining these two inequalities, we get

$$\frac{\mathbf{d}}{2}(\mathbf{v}_1 - \mathbf{v}_0) \geq \mathbf{v}_1 - \mathbf{v}_0 \text{ a.s.}.$$

Thus, since $1 \geq \mathbf{d} > 0$ a.s., it follows that $\mathbf{d}(\mathbf{v}_1 - \mathbf{v}_0) \leq 0$ a.s.. However, this implies that $\mathbf{v}_0 \geq \mathbf{v}_1$ a.s.. If Assumption 1.1 holds, then it must be the case that:

$$\mathbb{E}\left[\rho_0 \mathbf{w}_0\right] = \mathbb{E}\left[\rho_1 \mathbf{w}_1\right] = 0.$$

As a result, applying the centering condition, we obtain a contradiction:

$$0 \geq \mathbb{E}\left[\mathbf{d}(\mathbf{v}_1 - \mathbf{v}_0)\right] \geq 4 + \mathbb{E}\left[\rho_1'(\mathbf{w}_1 - \mathbf{w}_0)\right] = 4 + \mathbb{E}\left[\rho_1'\mathbf{w}_1\right] - 2\mathbb{E}\left[\rho_0'\mathbf{w}_0\right] = 4.$$

This contradiction means that the random array $(\boldsymbol{\rho}_t, \mathbf{c}_t)_{t \in \mathcal{T}}$ is not s/ED-rationalizable under the centering conditions we described above.

We highlight that potential lack of statistical power of some centering conditions is not a defect of our methodology. The reason is that if the quality of the data set at hand is not good enough to credibly discern whether observed behavior is consistent with a given model, then no methodology can address this issue.

7. Empirical Application (I): Testing the Dynamic UMT with Exponential Discounting in Survey Data

In our first application, we apply our methodology to a consumer panel data set gathered from single-individual and couples' households in Spain to test for s/ED-rationalizability. This important model is under increasing scrutiny because experimental evidence tends to find that the behavior of experimental subjects is time-inconsistent.[38] Nonetheless, it is important to explore to what extent this finding has external validity.

To address this issue, some researchers have turned to survey data in the form of household consumption panels. Most of this work has found evidence against exponential discounting (Blow et al., 2017). However, the existing literature has not yet addressed the issue of measurement error in the consumption reported by households in a way that allows us to perform traditional

---

[38]See, for instance, Andreoni & Sprenger (2012), Montiel Olea & Strzalecki (2014), and Echenique et al. (2014).

hypothesis testing. (Some additional problems with the existing evidence are (i) the strong parametric assumption on preferences, and (ii) homogeneity restrictions on the discount factor and preferences.)

One solution to some of the problems in the literature can be found in the work on deterministic RP by Browning (1989). In particular, Browning's work avoids making parametric assumptions about the functional form of instantaneous utility. However, this work does not take into consideration the fact that consumption can be mismeasured. In our Monte Carlo experiment, the deterministic test in Browning (1989) rejects the correct null hypothesis of exponential discounting behavior in 79.37 percent of the cases, while our methodology correctly fails to reject the null hypothesis that all households are consistent with exponential discounting at the correct 5 percent significance level (Appendix B.1). In addition, when we applied Browning (1989) deterministic methodology to our single-individual households data set, we also obtained a very low success rate. However, this low success rate of the deterministic test for exponential discounting may be due to measurement error. In our empirical application, we found support for exponential discounting behavior in single-individual households, while at the same time, support for the negative finding in Blow et al. (2017) in the case of couples' households. This fact indicates that deterministic tests may not be very informative about the behavior in a population due to measurement error. Small violations of the deterministic RP inequalities will lead to big rejection rates. Introducing measurement error into the analysis takes these small violations into account.[39]

Our empirical application also contributes to the literature on estimating the discount factor distribution in survey data sets and in a classical consumer theory environment. This has been the topic of a large body of work which, however, has reached little or no consensus.[40] This lack of consensus can be attributed in some degree to a failure to identify the parameters of interest. Here, we show that the discount factor distribution cannot be identified solely from prices, interest rates, and consumption observations in a data set that suffers from measurement error. (For details see Section 4.) However, our methodology allows us to test for exponential discounting behavior even in this setting (i.e., without identifying the discount factor distribution).[41]

If one ignores the issues of measurement error, the Euler equation allows one to estimate the discount factor and the marginal utility either parametrically or semi-parametrically.[42] Since our objective is not to estimate but to test the exponential discounting model, we follow a different path.

In particular, we work with the data set used in Adams et al. (2014): the Spanish Continuous Family Expenditure Survey (*Encuesta Continua de Presupuestos Familiares*). The data set consists of the expenditures for 185 individuals and 2004 couples, as well as prices for 17 commodities

---

[39]In Appendix E.2, we also establish that our test fails to rejects implications of the collective household consumption problem presented in Adams et al. (2014).

[40]We refer the reader to the survey by Frederick et al. (2002) for its extensive references.

[41]In order to learn more information about the discount factor distribution, one needs additional data. One notable example is Mastrobuoni & Rivers (2016), which uses a quasi-experiment to pin down criminals' time preferences.

[42]Examples of estimators of the Euler equation and similar models include Hall (1978), Hansen & Singleton (1982), Dunn & Singleton (1986), Gallant & Tauchen (1989), Chapman (1997), Campbell & Cochrane (1999), Ai & Chen (2003), Chen & Ludvigson (2009), Darolles et al. (2011), Chen et al. (2014), and Escanciano et al. (2016).

(categories of goods) recorded over four time periods. Each household was interviewed for four consecutive quarters between 1985-1997. We construct a panel data set of consumption and prices by pooling household's quarterly data points.

The categories of goods are: all food and nonalcoholic drinks, all clothing, cleaning, nondurable articles, household services, domestic services, public transport, long-distance travel, other transport, petrol, leisure (four categories), other services (two categories), and food consumed outside the home. The data set also contains information on the nominal interest rate on consumer loans faced by the household in any particular quarter.[43]

Formally, we test for s/ED-rationalizability with (i) effective prices equal to the discounted spot prices, $\boldsymbol{\rho}_t = \boldsymbol{\rho}_t^{\text{ED}}$ (defined in Table 1), (ii) random marginal utility of income equal to the discounted value of one unit of wealth, $\boldsymbol{\lambda}_t = 1$ a.s., and (iii) $\boldsymbol{\delta}_t = \mathbf{d}^t$ where $\mathbf{d}$ is interpreted as the (time-invariant) random discount factor supported on or inside $(0, 1]$.[44] Imposing the additional Assumptions 1.1, 2, and 4, we can apply our testing methodology developed in Section 5.4. Recall that Assumption 1.1 indicates that measurement error does not alter the mean value of total expenditure, $\mathbb{E}\left[\boldsymbol{\rho}_t' \mathbf{c}_t\right] = \mathbb{E}\left[\boldsymbol{\rho}_t' \mathbf{c}_t^*\right]$.

7.1. The Results

Single-Individual Households

We apply the deterministic methodology of Browning (1989) to single-individual households. Our initial conclusion is that 81.1 percent of the single-individual households behave inconsistently with exponential discounting (even when allowing for substantially more heterogeneity than previous works).[45] Next, we revisit this conclusion using our methodology, which addresses measurement error, while allowing a heterogeneous discount factor. We find that we cannot reject exponential discounting. Formally, we find at the 5 percent significance level that the random array $\mathbf{x} = (\boldsymbol{\rho}_t, \mathbf{c}_t)_{t \in \mathcal{T}}$ is s/ED-rationalizable with a random discount factor $\mathbf{d}$ *supported on or inside* $[0.975, 1]$ ($\text{TS}_n = 6.479$, p-value = 0.166). We believe that the lower bound value of 0.975, for the quarterly discount rate, is reasonable (it corresponds to a annualized discount rate of 0.9).

Although there is no agreement in the literature about what appropriate values for the discount factor are (Frederick et al., 2002), a common benchmark in applied work is to set the discount factor according to the real interest rate in the economy.[46] In our case, the lower bound of the quarterly discount factor is 0.975. This discount factor bound corresponds to an annual real interest

---

[43]We spare the reader more details and refer them instead to Adams et al. (2014) for further information on the data set.

[44]When the discount factor is 0, it is easy to see that ED-rationality becomes equivalent to R-rationality. In practice, we pick the interval $[0.1, 1]$ as the largest possible support for discount factor. This interval contains most reasonable values in the literature (Frederick et al., 2002, Montiel Olea & Strzalecki, 2014).

[45]We search for each individual household discount factor d in the grid $\{0.1, 0.15, \cdots, 1\}$. See Crawford (2010), Adams et al. (2014), and Blow et al. (2017) for discount factor ranges close to $[0.9, 1]$.

[46]DeJong & Dave (2011) suggests setting the discount factor value to $d = 1/(1 + \bar{r})$, where $\bar{r}$ is the average (across individuals) annual real interest rate.

rate of 10.7 percent. This is roughly consistent with this benchmark for the average real interest rates observed in our sample.[47]

However, at the 5 percent significance level, we cannot reject exponential discounting, when discounting is set at 1 a.s. ($\text{TS}_n = 6.140$, p-value = 0.189). Of course, this does not mean that this is the value of the discount factor. In fact, our sample and our knowledge about measurement error are not enough to elicit the support of the distribution of discount factors. Nonetheless, we can differentiate between behavior consistent with exponential discounting and systematic departures from it (as seen in the power analysis of our method). In sum, our findings provide evidence supporting exponential discounting for singles under reasonable discount factors.

Using our methodology in data sets with more time periods or with additional information about measurement error will improve the information we can obtain about discount factors. However, we will see that identifying the support of the discount factor is not essential to provide informative bounds on average demand (see Appendix F).[48]

Couples' Households

For the couples' households, the deterministic test of Browning (1989) rejects the exponential discounting model for 88.5 percent of the observations. Although this number seems large, one should keep in mind that for single-individual households the same deterministic test rejects the model in 81.1 percent of the cases. At the same time, for $\mathbf{d} \in [0.1, 1]$, our method does not reject the exponential discounting model for single-individual households. But we do reject the model for couples' households. In the case of couples' households, the test statistic is $\text{TS}_n = 21.205$ (p-value = 0.0003).[49]

7.2. Discussion and Related Work on Testing the Exponential Discounting Model

One possible concern about our methodology is that its power is low in survey data sets (due to a small T dimension of the data) given its nonparametric nature. However, in Appendix B.2 we report, for 1000 trials with a sample size of 2000, a rejection rate greater than or equal to 71.3 percent (with a data generating process consistent with the collective model as in Adams et al., 2014).

---

[47]Using the results from Section 4.1 we also tested several candidates for the average quarterly discount factor $\theta_0 = \mathbb{E}[\mathbf{d}]$: 0.991 ($\text{TS}_n = 15.954$, p-value = 0.007), 0.992 ($\text{TS}_n = 9.367$, p-value = 0.095), and 0.995 ($\text{TS}_n = 5.124$, p-value = 0.401). The smallest θ_0 that is not rejected at the 5 percent significance level, 0.992, corresponds to the annual average discount factor of 0.968 and the average annual real interest rate of 3.3 percent.

[48]As robustness check we conducted the test for the discount factor supported on or inside $[0.1, 1]$, $[0.5, 1]$, and $[0.9, 1]$. As expected, since all three intervals contain $[0.975, 1]$, the null hypothesis is not rejected. The test statistic TS_n is equal to 6.476 with p-value = 0.166 for all three intervals.

[49]As a robustness check, we also tested the model assuming that $\mathbf{d}$ is supported on or inside $[0.5, 1]$, $[0.9, 1]$, and $\mathbf{d} = 1$ a.s.. As expected, since these intervals are contained in $[0.1, 1]$, the null hypothesis is rejected. The test statistic TS_n (p-value) is equal to 21.205 (0.0003), 21.189 (0.0003), and 21.997 (0.0003), respectively.

Our results for couples' households provide the first nonparametric evidence which is robust to measurement error and which demonstrates that not all couples' households manifests behavior consistent with exponential discounting. In Appendix E.2, we establish that a suitable extension of our methodology fails to reject the collective household consumption problem presented in Adams et al. (2014).[50] This should convince practitioners about the importance of modeling intrahousehold decision-making when dealing with intertemporal choice. The rejection of exponential discounting behavior for couples' households can be better understood given theoretical results that show that aggregating time-consistent preferences may lead to time-inconsistent behavior (Jackson & Yariv, 2015).

The deterministic methodology of Browning (1989) concludes that the fraction of households that is inconsistent with exponential discounting under the deterministic test is similar for both single individuals and couples, but our statistical test rejects in the latter case while reaching the opposite conclusion in the former case. The difference in conclusions is due to the fact that our test implicitly takes into account the severity of the violations of exponential discounting, and imposes the mean budget-neutrality assumption on the measurement error corrections.[51]

Our main empirical finding is robust to price measurement error. Adding an additional source of measurement error would make the rationalizability notion less demanding.[52]

8. Empirical Application (II): Static UMT in Experimental Data Sets with Trembling-Hand Errors and Misperception of Prices

In this section, we use our methodology to test the static UMT in the widely known experimental data set by Ahn et al. (2014). The experimental task consists of $T = 50$ independent decision trials with $n = 154$ subjects. Each decision is a portfolio problem. The subjects face three states of the world $\sigma \in \{1, 2, 3\}$. The subjects are given 100 tokens per task and they have to choose a bundle of Arrow securities, $c_t \in \mathbb{R}_+^3$, for a randomly drawn price vector $p_t \in \mathbb{R}_+^3 \setminus \{0\}$. The subjects are forced to choose a bundle that satisfies Walras' law such that for every decision task it must be that $p_t' c_t = 100$. The subjects receive a payment in tokens according to the probability of each state of the world at the end of each round. At the end of the experimental task one of the rounds was selected using a uniform distribution and the tokens payment corresponding to that round is paid in dollars.[53] The exchange rate is 0.05 dollars per token and the participation fee is 5 dollars.

This ingenious experimental device (due to Choi et al., 2014) has allowed the RP practitioners to

---

[50]Mazzocco (2007) also provides evidence in favor of the collective model using a parametric methodology.

[51]We have also tested s/ED-rationality under Assumption 1.2, which requires that consumption quantities measurement error is centered around zero as in Varian (1985). We strongly reject the null hypothesis in this case, providing an illustration of the importance of using empirically-backed centering conditions.

[52]See Appendix D.3 for results on robustness of our methodology to local perturbations in prices.

[53]The subjects are informed that the probability of state $\sigma = 2$ is 1/3, and the joint probability of the states $\sigma \in \{1, 3\}$ is 2/3.

collect a large number of observations per individual with high price variation. Beatty & Crawford (2011) highlighted the importance of rich price variation to have enough power in the experimental design to detect violations of the UMT.

The deterministic RP test for the static UMT in this data set concludes that only 12.98 percent of the experimental subjects pass the test. At first sight, this is a striking result, because the majority of subjects seems to be inconsistent with the core consumer model in economics. We reexamine the robustness of this result to measurement error in consumption due to errors in the elicitation of the intended behavior of consumers. In our application, we found support for the static UMT in this experiment in stark contrast with the findings from the deterministic RP test.

Measurement error in the experimental environment may arise due to the nature of the design. Subjects are presented with a graphical representation of a 3 dimensional budget hyperplane, and they must choose the consumption bundle by pointing to a point in this hyperplane using a computer mouse or the arrow keys in a keyboard. We must note that there is a mechanical measurement error due to the resolution of the budget hyperplane which is 0.2 tokens. More important factors, such as a lack of expertise in the decision task, could lead the consumers to make implementation mistakes when trying to choose their preferred alternative. Kurtz-David et al. (2019) provide direct evidence of trembling-hand error in a budget allocation task similar to Ahn et al. (2014), as well as indirect evidence of visual misperception, as we have previously discussed. Nonetheless, the actual reason why the experimental design fails to elicit the intended decision task is not our main concern. We take the stand that a desirable test of the static UMT has to be robust to possible nonsystematic mistakes arising from any experimental design. We consider both trembling-hand errors in consumption and nonsystematic misperception of prices. Formally, we test for s/R-rationalizability with (i) effective prices equal to the prices at each trial $\boldsymbol{\rho}_t = \boldsymbol{\rho}_t^{\mathrm{R}}$ (defined in Table 1), (ii) marginal utility of wealth $(\boldsymbol{\lambda}_t)_{t \in \mathcal{T}}$ supported on $\Lambda = \mathbb{R}_{++}^{|\mathcal{T}|}$, and (iii) random discount factor equal to 1 ($\boldsymbol{\delta}_t = 1$ a.s. for all $t \in \mathcal{T}$).

We note that there is evidence for trembling-hand errors and misperception errors in the budget allocation task we consider here (see Section 6). However, we fail to reject the null hypothesis of the static UMT when allowing only for price misperception. Evidently, it follows that we fail to reject the null hypothesis of the static UMT when allowing for both price misperception and trembling hand errors. In addition, we reject the null hypothesis of the static UMT when allowing only for trembling hand errors. Thus, we can conclude that the main source of error in this type of experimental tasks is most likely price misperception.

Misperception of Prices

We consider the possibility of measurement error in prices arising from misperception. We investigate the case where consumers behave as if they are trying to maximize a utility function subject to a misperceived vector of prices. In this regard, we take the point-of-view of Gillen et al. (2017) that points out that misperception in a low stakes experimental environments may affect the external validity of the conclusions drawn from an experimental data set. Kurtz-David et al. (2019)

provide indirect evidence of errors induced by the visual task in the budget allocation problem in the very closely related interface of Choi et al. (2014). We require that consumers' average perception of prices is unbiased, namely, for all $t \in \mathcal{T}$ it must be that

$$\mathbb{E}\left[\boldsymbol{\rho}_t\right] = \mathbb{E}\left[\boldsymbol{\rho}_t^*\right].$$

This is captured in Assumption 1.3. In order to isolate the effect of misperception on the observed violations of the static UMT, we assume that consumer behavior is measured perfectly ($\mathbf{w}_t^c = 0$ a.s.). In addition, due to the experimental design in Ahn et al. (2014), it must be that true prices are such that $\mathbf{p}_t^{*\prime}\mathbf{c}_t^* = 100$ a.s. (i.e., Walras' law holds). The value of the test statistic is $\mathrm{TS}_n = 10.211$ (p-value $> 1 - 10^{-10}$). This is below the conservative critical value $\chi_{150,0.95}^2 = 179.581$.[54] We do not reject the null hypothesis of the static UMT in the presence of misperception of prices, when the average vector of prices is equivalent to the true vector of prices.[55] More importantly, this finding puts in perspective the rejection of the static UMT in experimental data sets that use the graphical representation of budget hyperplanes (Choi et al., 2014, Ahn et al., 2014). In particular, we find evidence that prices misperception matters. When we account for this possibility, we no longer reject the null hypothesis of the static UMT.

Trembling-Hand Errors in Consumption

We say that measurement error in consumption is the result of a trembling-hand when it is centered at zero. This idea is captured in Assumption 1.2 that requires that for all $t \in \mathcal{T}$ it must be that $\mathbb{E}\left[\mathbf{w}_t^c\right] = 0$. Also, we keep the restriction that the true prices and consumption satisfy Walras' law $\mathbf{p}_t^{*\prime}\mathbf{c}_t^* = 100$ a.s., and that prices are measured perfectly $\mathbf{w}_t^p = 0$ a.s.. As we discussed before, we have obtained direct evidence for this centering condition using the replication data of Kurtz-David et al. (2019). We strongly reject the null hypothesis of the static UMT when allowing for measurement error in the elicitation of the true consumer behavior due to trembling-hand errors. The test statistic is $\mathrm{TS}_n = 214.355$ (p-value $< 5 \cdot 10^{-4}$). This is above the conservative critical value $\chi_{150,0.95}^2 = 179.581$.[56]

---

[54]We used 899 draws in the Monte Carlo computation of the maximum-entropy moment of this problem. We chose this number on the basis of the previous empirical exercise in this experimental data set.

[55] Assumption 1.3 has empirical bite. See Appendix D.2.

[56]We used 2970 draws in the Monte Carlo computation of the maximum-entropy moment of this problem. We also tried 580 and 899 with test statistics with values of 215.94 and 200.70. Which is evidence that moderate size of draws for the Monte Carlo integration step do relatively well in this setup.

9. Relation to the Literature

Afriat (1967) shows that it is sufficient to know shape constraints on the utility function (i.e. concavity) to bypass the need to know the utility function when testing for the UMT. We generalize this insight by allowing measurement error in consumption and prices. Among authors using the deterministic RP approach, the immediate antecedents to our work using the first-order approach are Browning (1989), Blow et al. (2017), and Brown & Calsamiglia (2007). Important advances have been made on testing and doing counterfactual analysis under rationalizability or random utility.[57] However, the majority of these results assume that observed quantities are measured accurately. Varian (1985) is possibly the first work to introduce the subject of measurement error into the RP approach. Varian's methodology is the closest to that of our own work; he considers precisely measured (albeit random) prices to study measurement error in consumption. Varian's work is compatible with standard statistical hypothesis testing under the strong assumptions of normality (with known variance) and additivity of consumption measurement error. In contrast, our methodology is fully nonparametric. We are able to improve upon Varian's methodology and relax its core assumptions by using a moments approach to measurement error in the RP framework.

Other papers have dealt with measurement error under different parametric assumptions about measurement error or about the heterogeneity of preferences.[58] Deb et al. (2017) consider a nonparametric model of "price preference." They propose an RP test of their model that is robust to small measurement error in prices. Boccardi (2016) considers a case of demand with error and establishes a way to account for the trade-off between the fit of the model and its predictive ability (which is a generalization of Beatty & Crawford, 2011).[59]

In practice, the RP theorists (e.g., Adams et al., 2014 and Cherchye et al., 2017) have dealt with measurement error by perturbing (minimally) the observed individual consumption in order to satisfy the conditions of an RP test. For instance, Adams et al. (2014) find the additive perturbation with a minimal norm that renders the individual consumption streams compatible with the RP restrictions. Then a subjective threshold is imposed on the maximum admissible norm of the measurement error vector. If the computed norm is above the threshold, then the model is rejected. However, their methodology has one important drawback: every data set can be made to satisfy their test or, equivalently, the test has no power. In addition, the test in Adams et al. (2014) has no size control.

Among researchers using the RP approach, Blundell et al. (2003, 2014) are the first to provide

---

[57]Relevant examples are Blundell et al. (2014), Dette et al. (2016), Lewbel & Pendakur (2017), and Kitamura & Stoye (2018).

[58]Gross (1995) assumes that random consumption is generated by consumers with similar preferences. Tsur (1989) imposes a log-normal multiplicative measurement-error structure in expenditures. Hjertstrand (2013) proposes a generalization of Varian (1985), but requires knowing the distribution of measurement error. Echenique et al. (2011) assume that measurement error in prices is a normal random variable independent across households and prices with a fixed mean and known variance.

[59]The fit of an RP model indicates whether the a data set is consistent with the model. The predictive ability of an RP model is a measure of how easy it is for a data set generated at random to be consistent with the model. See Beatty & Crawford (2011) for further details.

consumer demand bounds, under the assumption of static utility maximization in a semiparametric environment (with additive heterogeneity) in which income changes continuously. Our work differs from theirs mainly in that we allow for unrestricted heterogeneity in preferences, do not require that income be observable, and do not impose semiparametric assumptions on wealth effects to provide bounds for demand, given new prices. In addition, nonclassical measurement error is not compatible with their approach.

10. Conclusion

We propose a new stochastic and nonparametric RP approach (suitable for an environment with measurement error in consumption or prices) that is useful to test for several consumer models that can be characterized by their first-order conditions. In particular, our work can be used (but is not limited) to test for static utility maximization (Afriat, 1967), and for dynamic rationalizability with exponential discounting (Browning, 1989).

References

Abildgren, K., Kuchler, A., Rasmussen, A. S. L., & Sorensen, H. S. (2018). Consistency between household-level consumption data from registers and surveys. In *35th IARIW General Conference, Copenhagen, Danmarks Nationalbank Working Paper, no. 131*.

Adams, A., Cherchye, L., De Rock, B., & Verriest, E. (2014). Consume Now or Later? Time Inconsistency, Collective Choice, and Revealed Preference. *American Economic Review*, 104(12), 4147–4183.

Afriat, S. N. (1967). The construction of utility functions from expenditure data. *International economic review*, 8(1), 67–77.

Ahn, D., Choi, S., Gale, D., & Kariv, S. (2014). Estimating ambiguity aversion in a portfolio choice experiment. *Quantitative Economics*, 5(2), 195–223.

Ai, C. & Chen, X. (2003). Efficient estimation of models with conditional moment restrictions containing unknown functions. *Econometrica*, 71(6), 1795–1843.

Alan, S., Atalay, K., & Crossley, T. F. (2018). Euler equation estimation on micro data. *Macroeconomic Dynamics*, (pp. 1–26).

Alan, S., Attanasio, O., & Browning, M. (2009). Estimating euler equations with noisy data: two exact gmm estimators. *Journal of Applied Econometrics*, 24(2), 309–324.

Alan, S. & Browning, M. (2010). Estimating intertemporal allocation parameters using synthetic residual estimation. *The Review of Economic Studies*, 77(4), 1231–1261.

Allen, R. & Rehbeck, J. (2019). Identification with additively separable heterogeneity. *Econometrica*, 87(3), 1021–1054.

Andreoni, J. & Sprenger, C. (2012). Estimating time preferences from convex budgets. *The American Economic Review*, 102(7), 3333–3356.

Angel, S., Heuberger, R., & Lamei, N. (2018). Differences between household income from surveys and registers and how these affect the poverty headcount: evidence from the austrian silc. *Social indicators research*, 138(2), 575–603.

Attanasio, O., Hurst, E., & Pistaferri, L. (2014). The evolution of income, consumption, and leisure inequality in the united states, 1980–2010. In *Improving the measurement of consumer expenditures* (pp. 100–140). University of Chicago Press.

Baker, S. R., Kueng, L., Meyer, S., & Pagel, M. (2018). *Measurement error in imputed consumption*. Technical report, National Bureau of Economic Research.

Beatty, T. K. & Crawford, I. A. (2011). How demanding is the revealed preference approach to demand? *The American Economic Review*, 101(6), 2782–2795.

Blow, L., Browning, M., & Crawford, I. (2013). Never Mind the Hyperbolics: Nonparametric Analysis of Time-Inconsistent Preferences. *Unpublished manuscript.*

Blow, L., Browning, M., & Crawford, I. (2017). Nonparametric analysis of time-inconsistent preferences. *Mimeo.*

Blundell, R., Kristensen, D., & Matzkin, R. (2014). Bounding quantile demand functions using revealed preference inequalities. *Journal of Econometrics*, 179(2), 112–127.

Blundell, R. W., Browning, M., & Crawford, I. A. (2003). Nonparametric engel curves and revealed preference. *Econometrica*, 71(1), 205–240.

Boccardi, M. J. (2016). Predictive ability and the fit-power trade-off in theories of consumer behavior. *Mimeo.*

Boelaert, J. (2014). *revealedPrefs: Revealed Preferences and Microeconomic Rationality*. R package version 0.2.

Boskin, M. J., Dulberger, E. L., Gordon, R. J., Griliches, Z., & Jorgenson, D. W. (1998). Consumer prices, the consumer price index, and the cost of living. *Journal of economic perspectives*, 12(1), 3–26.

Bound, J., Brown, C., & Mathiowetz, N. (2001). Chapter 59 - Measurement Error in Survey Data. In J. J. Heckman & E. Leamer (Eds.), *Handbook of Econometrics*, volume 5 (pp. 3705–3843). Elsevier.

Brown, D. J. & Calsamiglia, C. (2007). The nonparametric approach to applied welfare analysis. *Economic Theory*, 31(1), 183–188.

Browning, M. (1989). A nonparametric test of the life-cycle rational expectations hypothesis. *International Economic Review*, (pp. 979–992).

Browning, M., Chiappori, P.-A., & Weiss, Y. (2010). Uncertainty and dynamics in the collective model.

Campbell, J. Y. & Cochrane, J. H. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of political Economy*, 107(2), 205–251.

Carroll, C. D. (2001). Death to the log-linearized consumption euler equation!(and very poor health to the second-order approximation). *Advances in Macroeconomics*, 1(1).

Carroll, C. D., Crossley, T. F., & Sabelhaus, J. (2014). Introduction to" improving the measurement of consumer expenditures". In *Improving the Measurement of Consumer Expenditures* (pp. 1–20). University of Chicago Press.

Castillo, J. R., Ley, E., & Izquierdo, M. (1999). *La medición de la inflación en España*. Number 17. " la Caixa: Savings and Pensions Bank of Barcelona".

Chapman, D. A. (1997). Approximating the asset pricing kernel. *The Journal of Finance*, 52(4), 1383–1410.

Chen, X., Chernozhukov, V., Lee, S., & Newey, W. K. (2014). Local identification of nonparametric and semiparametric models. *Econometrica*, 82(2), 785–809.

Chen, X. & Ludvigson, S. C. (2009). Land of addicts? an empirical investigation of habit-based asset pricing models. *Journal of Applied Econometrics*, 24(7), 1057–1093.

Cherchye, L., Demuynck, T., De Rock, B., Vermeulen, F., et al. (2017). Household consumption when the marriage is stable. *American Economic Review*, 107(6), 1507–1534.

Choi, S., Kariv, S., Müller, W., & Silverman, D. (2014). Who is (more) rational? *The American Economic Review*, 104(6), 1518–1550.

Cochrane, J. H. (1991). A simple test of consumption insurance. *Journal of political economy*, 99(5), 957–976.

Crawford, I. (2010). Habits revealed. *The Review of Economic Studies*, 77(4), 1382–1402.

Darolles, S., Fan, Y., Florens, J.-P., & Renault, E. (2011). Nonparametric instrumental regression. *Econometrica*, 79(5), 1541–1565.

Deb, R., Kitamura, Y., Quah, J. K.-H., & Stoye, J. (2017). Revealed price preference: Theory and stochastic testing. *Working paper*.

DeJong, D. N. & Dave, C. (2011). *Structural macroeconometrics*. Princeton University Press.

DellaVigna, S. & Malmendier, U. (2006). Paying not to go to the gym. *The American Economic Review*, (pp. 694–719).

Dette, H., Hoderlein, S., & Neumeyer, N. (2016). Testing multivariate economic restrictions using quantiles: the example of slutsky negative semidefiniteness. *Journal of Econometrics*, 191(1), 129–144.

Diewert, W. E. (1998). Index number issues in the consumer price index. *Journal of Economic Perspectives*, 12(1), 47–58.

Diewert, W. E. (2012). Afriat's theorem and some extensions to choice under uncertainty. *The Economic Journal*, 122(560), 305–331.

Dunn, K. B. & Singleton, K. J. (1986). Modeling the term structure of interest rates under non-separable utility and durability of goods. *Journal of Financial Economics*, 17(1), 27–55.

Echenique, F., Imai, T., & Saito, K. (2014). Testable implications of quasi-hyperbolic and exponential time discounting.

Echenique, F., Lee, S., & Shum, M. (2011). The money pump as a measure of revealed preference violations. *Journal of Political Economy*, 119(6), 1201–1223.

Einav, L., Leibtag, E., & Nevo, A. (2010). Recording discrepancies in nielsen homescan data: Are they present and do they matter? *QME*, 8(2), 207–239.

Ekeland, I., Galichon, A., & Henry, M. (2010). Optimal transportation and the falsifiability of incompletely specified economic models. *Economic Theory*, 42(2), 355–374.

Escanciano, J. C., Hoderlein, S., Lewbel, A., Linton, O., & Srisuma, S. (2016). Nonparametric euler equation identification and estimation. *Working Paper*.

Forges, F. & Minelli, E. (2009). Afriat's theorem for general budget sets. *Journal of Economic Theory*, 144(1), 135–145.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature*, 40(2), 351–401.

Galichon, A. & Henry, M. (2013). Dilation bootstrap. *Journal of Econometrics*, 177(1), 109 – 115.

Gallant, A. R. & Tauchen, G. (1989). Seminonparametric estimation of conditionally constrained heterogeneous processes: Asset pricing applications. *Econometrica: Journal of the Econometric Society*, (pp. 1091–1120).

Gauthier, C. (2018). Nonparametric identification of discount factors under partial efficiency. *Working paper.*

Gillen, B., Snowberg, E., & Yariv, L. (2017). *Experimenting with measurement error: techniques with applications to the Caltech cohort study.* Technical report, National Bureau of Economic Research.

Gradín, C., Cantó, O., & Del Río, C. (2008). Inequality, poverty and mobility: Choosing income or consumption as welfare indicators. *Investigaciones Económicas*, 32(2), 169–200.

Gross, J. (1995). Testing data for consistency with revealed preference. *The Review of Economics and Statistics*, (pp. 701–710).

Guerrero de Lizardi, C. (2008). Sesgos de medición del índice nacional de precios al consumidor, 2002-2007. *Investigación económica*, 67(266), 37–65.

Hall, R. E. (1978). Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence. *Journal of political economy*, 86(6), 971–987.

Hansen, L. P. & Singleton, K. J. (1982). Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica: Journal of the Econometric Society*, (pp. 1269–1286).

Hjertstrand, P. (2013). A simple method to account for measurement errors in revealed preference tests. *IFN Working Paper.*

Jackson, M. O. & Yariv, L. (2015). Collective dynamic choice: the necessity of time inconsistency. *American Economic Journal: Microeconomics*, 7(4), 150–178.

Jerison, D. & Jerison, M. (1994). Commodity aggregation and slutsky asymmetry. In *Models and Measurement of Welfare and Inequality* (pp. 752–764). Springer.

Kitamura, Y. & Stoye, J. (2018). Nonparametric analysis of random utility models. *Econometrica*, 86(6), 1883–1909.

Kolsrud, J., Landais, C., & Spinnewijn, J. (2017). Studying consumption patterns using registry data: lessons from swedish administrative data.

Kurtz-David, V., Persitz, D., Webb, R., & Levy, D. J. (2019). The neural computation of inconsistent choice behavior. *Nature communications*, 10(1), 1583.

Lewbel, A. (1996). Aggregation without separability: A generalized composite commodity theorem. *The American Economic Review*, 86(3), 524–543.

Lewbel, A. & Pendakur, K. (2009). Tricks with hicks: The easi demand system. *American Economic Review*, 99(3), 827–63.

Lewbel, A. & Pendakur, K. (2017). Unobserved preference heterogeneity in demand using generalized random coefficients. *Journal of Political Economy*, 125(4), 1100–1148.

Ludvigson, S. & Paxson, C. H. (2001). Approximation bias in linearized euler equations. *Review of Economics and Statistics*, 83(2), 242–256.

Mastrobuoni, G. & Rivers, D. (2016). Criminal discount factors and deterrence. *Available at SSRN 2730969.*

Mathiowetz, N., Brown, C., & Bound, J. (2002). Measurement error in surveys of the low-income population. *Studies of welfare populations: Data collection and research issues*, (pp. 157–194).

Mazzocco, M. (2007). Household Intertemporal Behaviour: A Collective Characterization and a Test of Commitment. *The Review of Economic Studies*, 74(3), 857–895.

Meyer, B. D., Mok, W. K., & Sullivan, J. X. (2015). Household surveys in crisis. *Journal of Economic Perspectives*, 29(4), 199–226.

Montgomery, A., Olivola, C. Y., & Pretnar, N. (2019). A structural model of mental accounting. *Available at SSRN 3472156.*

Montiel Olea, J. L. & Strzalecki, T. (2014). Axiomatization and Measurement of Quasi-hyperbolic Discounting. *Quarterly Journal of Economics*, 129, 1449–1499.

Norets, A. & Tang, X. (2014). Semiparametric inference in dynamic binary choice models. *Review of Economic Studies*, 81(3), 1229–1262.

Pistaferri, L. (2015). Household consumption: Research questions, measurement issues, and data collection strategies. *Journal of Economic and Social Measurement*, 40(1-4), 123–149.

Polisson, M., Quah, J. K.-H., & Renou, L. (2020). Revealed preferences over risk and uncertainty. *American Economic Review*, 110(6), 1782–1820.

Powell, M. J. (2009). The bobyqa algorithm for bound constrained optimization without derivatives. *Cambridge NA Report NA2009/06, University of Cambridge, Cambridge*, (pp. 26–46).

Rios, L. M. & Sahinidis, N. V. (2013). Derivative-free optimization: a review of algorithms and comparison of software implementations. *Journal of Global Optimization*, 56(3), 1247–1293.

Rockafellar, R. T. (1970). *Convex analysis.* Princeton university press.

Sasaki, Y. (2015). A contraction fixed point method for infinite mixture models and direct counterfactual analysis.

Sato, H. (2020). Do large-scale point-of-sale data satisfy the generalized axiom of revealed preference in aggregation using representative price indexes?: A case involving processed food and beverages. *Mimeo.*

Schennach, S. M. (2014). Entropic latent variable integration via simulation. *Econometrica*, 82(1), 345–385.

Shumway, C. R. & Davis, G. C. (2001). Does consistent aggregation really matter? *Australian Journal of Agricultural and Resource Economics*, 45(2), 161–194.

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing science*, 4(3), 199–214.

Toda, A. A. & Walsh, K. (2015). The double power law in consumption and implications for testing euler equations. *Journal of Political Economy*, 123(5), 1177–1200.

Tsur, Y. (1989). On testing for revealed preference conditions. *Economics Letters*, 31(4), 359–362.

Varian, H. R. (1982). The nonparametric approach to demand analysis. *Econometrica: Journal of the Econometric Society*, (pp. 945–973).

Varian, H. R. (1984). The nonparametric approach to production analysis. *Econometrica: Journal of the Econometric Society*, (pp. 579–597).

Varian, H. R. (1985). Non-parametric analysis of optimizing behavior with measurement error. *Journal of Econometrics*, 30(1-2), 445–458.

Varian, H. R. (1990). Goodness-of-fit in optimizing models. *Journal of Econometrics*, 46(1-2), 125–140.

Ventura, E. (1994). A note on measurement error and euler equations: An alternative to log-linear approximations. *Economics Letters*, 45(3), 305–308.

A. Appendix

A.1. Proof of Lemma 2

First we establish that (i) implies (ii). If the random array $(\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t \in \mathcal{T}}$ is s/m-rationalizable, by concavity of $\mathbf{u}(\cdot)$, with probability 1 for any s, t and some $\xi \in \nabla \mathbf{u}(\mathbf{c}_t^*)$

$$\mathbf{u}(\mathbf{c}_s^*) - \mathbf{u}(\mathbf{c}_t^*) \leq \xi'(\mathbf{c}_s^* - \mathbf{c}_t^*),$$

$$\xi \leq \frac{\boldsymbol{\lambda}_t}{\boldsymbol{\delta}_t} \boldsymbol{\rho}_t^*.$$

Let $\mathbf{N}$ be a random set of indices such that $\frac{\lambda_t}{\delta_t}\rho_{ti}^* = \xi_i$ for every $i \in \mathbf{N}$. Hence, $\frac{\lambda_t}{\delta_t}\rho_{ti}^* \geq \xi_i$ for every $i \notin \mathbf{N}$ with probability 1. As a result, $\mathbf{c}_{ti}^*$ has to be a corner solution for every $i \notin \mathbf{N}$. That is, $\mathbf{c}_{ti}^* = 0$. Thus, with probability 1,

$$
\mathbf{u}(\mathbf{c}_s^*) - \mathbf{u}(\mathbf{c}_t^*) \leq \xi'(\mathbf{c}_s^* - \mathbf{c}_t^*) = \sum_{i \in \mathbf{N}} \xi_i(\mathbf{c}_{si}^* - \mathbf{c}_{ti}^*) + \sum_{i \notin \mathbf{N}} \xi_i \mathbf{c}_{si}^* =
$$

$$
= \sum_{i \in \mathbf{N}} \frac{\lambda_t}{\delta_t}\rho_{ti}^*(\mathbf{c}_{si}^* - \mathbf{c}_{ti}^*) + \sum_{i \notin \mathbf{N}} \xi_i \mathbf{c}_{si}^* \leq \sum_{i \in \mathbf{N}} \frac{\lambda_t}{\delta_t}\rho_{ti}^*(\mathbf{c}_{si}^* - \mathbf{c}_{ti}^*) + \sum_{i \notin \mathbf{N}} \frac{\lambda_t}{\delta_t}\rho_{ti}^*\mathbf{c}_{si}^*,
$$

where the last inequality follows from $\mathbf{c}_s$ being nonnegative. As a result, with probability 1,

$$
\forall s, t \in \mathcal{T} \; : \; \mathbf{u}(\mathbf{c}_t^*) - \mathbf{u}(\mathbf{c}_s^*) \geq \frac{\lambda_t}{\delta_t}\rho_t^{*'}[\mathbf{c}_t^* - \mathbf{c}_s^*].
$$

For any $\epsilon > 0$, we let $\mathbf{v}_t = \mathbf{u}(\mathbf{c}_t^*) - \min_{s \in \mathcal{T}} \mathbf{u}(\mathbf{c}_s^*) + \epsilon$ a.s., for all $t \in \mathcal{T}$. The well-defined positive random vector $(\mathbf{v}_t)_{t \in \mathcal{T}}$ together with $(\lambda_t, \delta_t)_{t \in \mathcal{T}}$ satisfies the inequalities in (ii).

Now, we want to prove that (ii) implies (i). The result follows from Theorem 24.8 in Rockafellar (1970). For completeness of the proof we repeat the arguments of Theorem 24.8 in Rockafellar (1970). For a finite $m \in \mathbb{N}$, let $\mathbf{t} = \{t_i\}_{i=1}^m$, $t_i \in \mathcal{T}$, be a finite set of indices such that for a fixed $\hat{t} \in \mathcal{T}$, $c_{t_1}^* = c_{\hat{t}}^*$. Let $\mathcal{I}$ be the collection of all such indices (i.e., $\mathbf{t} \in \mathcal{I}$). Define

$$
\mathbf{u}(c^*) = \inf_{\mathbf{t} \in \mathcal{I}} \left\{ \frac{\lambda_{t_1}}{\delta_{t_1}}\rho_{t_1}^{*'}(\mathbf{c}_{t_2}^* - \mathbf{c}_{t_1}^*) + \cdots + \frac{\lambda_{t_m}}{\delta_{t_m}}\rho_{t_m}^{*'}(c^* - \mathbf{c}_{t_m}^*) \right\}.
$$

With probability 1, the random function $\mathbf{u}(\cdot)$ is well-defined, concave, locally nonsatiated, and continuous, since it is a pointwise minimum of a finite set of affine functions for every m. Moreover, the infimum in $\mathcal{I}$ is attained and achieved at a set of indices without repetitions (this is a consequence of (ii)). Indeed, under (ii), for any finite m, $\{t_i\}_{i=1}^m$ and $\mathbf{c}_s^*$, $s \in \mathcal{T}$, with probability 1,

$$
\frac{\lambda_{t_1}}{\delta_{t_1}}\rho_{t_1}^{*'}(\mathbf{c}_{t_2}^* - \mathbf{c}_{t_1}^*) + \cdots + \frac{\lambda_{t_m}}{\delta_{t_m}}\rho_{t_m}^{*'}(\mathbf{c}_s^* - \mathbf{c}_{t_m}^*) + \frac{\lambda_s}{\delta_s}\rho_s^{*'}(\mathbf{c}_{t_1}^* - \mathbf{c}_s^*) \geq
$$

$$
\mathbf{v}_{t_2} - \mathbf{v}_{t_1} + \mathbf{v}_{t_3} - \mathbf{v}_{t_2} + \cdots + \mathbf{v}_s - \mathbf{v}_{t_m} + \mathbf{v}_{t_1} - \mathbf{v}_s = 0.
$$

Thus,

$$
\mathbf{u}(\mathbf{c}_s^*) \geq \frac{\lambda_s}{\delta_s}\rho_s^{*'}(\mathbf{c}_s^* - \mathbf{c}_{t_1}^*) > -\infty
$$

with probability 1. (In particular, $u(c_{\hat{t}}^*) = 0$.)

It is left to show that, with probability 1, $\frac{\lambda_t}{\delta_t}\rho_t^* \in \nabla\mathbf{u}(\mathbf{c}_t^*)$ for all $t \in \mathcal{T}$. Fix some $t \in \mathcal{T}$ and $\delta > 0$. By the definition of $\mathbf{u}(\cdot)$, there exists some $\{t_i\}_{i=1}^m$ such that, with probability 1, $\mathbf{u}(\mathbf{c}_t^*) + \delta > \frac{\lambda_{t_1}}{\delta_{t_1}}\rho_{t_1}^{*'}(\mathbf{c}_{t_2}^* - \mathbf{c}_{t_1}^*) + \cdots + \frac{\lambda_{t_m}}{\delta_{t_m}}\rho_{t_m}^{*'}(\mathbf{c}_t^* - \mathbf{c}_{t_m}^*) \geq \mathbf{u}(\mathbf{c}_t^*)$. Again, by the definition of $\mathbf{u}(\cdot)$, for any c^*

$$
\frac{\lambda_{t_1}}{\delta_{t_1}}\rho_{t_1}^{*'}(\mathbf{c}_{t_2}^* - \mathbf{c}_{t_1}^*) + \cdots + \frac{\lambda_{t_m}}{\delta_{t_m}}\rho_{t_m}^{*'}(\mathbf{c}_t^* - \mathbf{c}_{t_m}^*) + \frac{\lambda_t}{\delta_t}\rho_t^{*'}(c^* - \mathbf{c}_t^*) \geq \mathbf{u}(c^*).
$$

Hence,

$$\mathbf{u}(\mathbf{c}_t^*) + \delta + \frac{\boldsymbol{\lambda}_t}{\boldsymbol{\delta}_t}\boldsymbol{\rho}_t^{*\prime}(c^* - \mathbf{c}_t^*) > \mathbf{u}(c^*).$$

Since the choice of δ, t and c^* was arbitrary, $\frac{\lambda_t}{\delta_t}\boldsymbol{\rho}_t^* \in \nabla\mathbf{u}(\mathbf{c}_t^*)$ for all $t \in \mathcal{T}$.

A.2. Proof of Proposition 1

Take any $\theta_1 \in \Theta_0$, $\theta_2 \in \Theta_0$, and $\lambda \in [0, 1]$ (if Θ_0 is empty, then the conclusion of the proposition follows trivially). Since $\theta_i \in \Theta_0$, $i = 1, 2$, by Theorems 2 and 3 there exist $\{\mu_{i,k}\}_{k=1}^{\infty}$, $i = 1, 2$, such that

$$\lim_{k\to\infty} \left\| \mathbb{E}_{\mu_{i,k}\times\pi_0} \left[g(\mathbf{x}, \tilde{\mathbf{e}}; \rho_{T+1}, \theta_i) \right] \right\| = 0$$

for both $i = 1, 2$. Consider $\theta_\lambda = \lambda\theta_1 + (1-\lambda)\theta_2$ and $\mu_{\lambda,k} = \lambda\mu_{1,k} + (1-\lambda)\mu_{2,k}$. Note that

$$\mathbb{E}_{\mu_{\lambda,k}\times\pi_0} \left[g_C(\mathbf{x}, \mathbf{c}_{T+1}; \rho_{T+1}, \theta_\lambda) \right] = \mathbb{E}_{\mu_{\lambda,k}\times\pi_0} \left[\tilde{g}(\mathbf{x}, \mathbf{c}_{T+1}; \rho_{T+1}) \right] - \mathbb{E}_{\pi_0} \left[A(\mathbf{x}; \rho_{T+1})'\theta_\lambda \right] =$$
$$\lambda\mathbb{E}_{\mu_{1,k}\times\pi_0} \left[\tilde{g}(\mathbf{x}, \mathbf{c}_{T+1}; \rho_{T+1}) \right] - \lambda\mathbb{E}_{\pi_0} \left[A(\mathbf{x}; \rho_{T+1})'\theta_1 \right] +$$
$$(1-\lambda)\mathbb{E}_{\mu_{2,k}\times\pi_0} \left[\tilde{g}(\mathbf{x}, \mathbf{c}_{T+1}; \rho_{T+1}) \right] - (1-\lambda)\mathbb{E}_{\pi_0} \left[A(\mathbf{x}; \rho_{T+1})'\theta_2 \right]$$
$$\lambda\mathbb{E}_{\mu_{1,k}\times\pi_0} \left[g_C(\mathbf{x}, \mathbf{c}_{T+1}; \rho_{T+1}, \theta_1) \right] + (1-\lambda)\mathbb{E}_{\mu_{2,k}\times\pi_0} \left[g_C(\mathbf{x}, \mathbf{c}_{T+1}; \rho_{T+1}, \theta_2) \right].$$

Hence, by the triangular inequality,

$$\lim_{k\to\infty} \left\| \mathbb{E}_{\mu_{\lambda,k}\times\pi_0} \left[g_C(\mathbf{x}, \mathbf{c}_{T+1}; \rho_{T+1}, \theta_\lambda) \right] \right\| \leq \lambda \lim_{k\to\infty} \left\| \mathbb{E}_{\mu_{1,k}\times\pi_0} \left[g_C(\mathbf{x}, \mathbf{c}_{T+1}; \rho_{T+1}, \theta_1) \right] \right\| +$$
$$(1-\lambda) \lim_{k\to\infty} \left\| \mathbb{E}_{\mu_{2,k}\times\pi_0} \left[g_C(\mathbf{x}, \mathbf{c}_{T+1}; \rho_{T+1}, \theta_2) \right] \right\| = \lambda \cdot 0 + (1-\lambda) \cdot 0 = 0.$$

Thus, since $g_{I,O}$ and g_M do not depend on θ, by the triangular inequality,

$$\lim_{k\to\infty} \left\| \mathbb{E}_{\mu_{\lambda,k}\times\pi_0} \left[g(\mathbf{x}, \tilde{\mathbf{e}}; \rho_{T+1}, \theta_\lambda) \right] \right\| \leq \lambda \lim_{k\to\infty} \left\| \mathbb{E}_{\mu_{1,k}\times\pi_0} \left[g(\mathbf{x}, \tilde{\mathbf{e}}; \rho_{T+1}, \theta_1) \right] \right\| +$$
$$(1-\lambda) \lim_{k\to\infty} \left\| \mathbb{E}_{\mu_{2,k}\times\pi_0} \left[g(\mathbf{x}, \tilde{\mathbf{e}}; \rho_{T+1}, \theta_2) \right] \right\| = 0.$$

The later means that $\theta_\lambda \in \Theta_0$. The fact that the choice of θ_1, θ_2, and λ was arbitrary implies that Θ_0 is convex.

A.3. Proof of Theorem 3

The result is a direct application of Theorem 2, and Theorem 2.1 in Schennach (2014). For completeness of the proof we present Theorem 2.1 in Schennach (2014) using our notation below.

Theorem (Theorem 2.1, Schennach, 2014). *Assume that the marginal distribution of $\mathbf{x}$ is supported on some set $X \subseteq \mathbb{R}^{d_x}$, while the distribution of $\mathbf{e}$ conditional on $\mathbf{x} = x$ is supported on or inside*

the set $E \subseteq \mathbb{R}^{d_e}$ for any $x \in X$. Let h, g and η satisfy Definition 7. Then

$$\inf_{\mu \in \mathcal{P}_{E|X}} \|\mathbb{E}_{\mu \times \pi_0}[g(\mathbf{x}, \mathbf{e})]\| = 0 \iff \inf_{\gamma \in \mathbb{R}^{k+q}} \|\mathbb{E}_{\pi_0}[h(\mathbf{x}; \gamma)]\| = 0,$$

where $\pi_0 \in \boldsymbol{P}_{\mathcal{X}}$ is the observed distribution of $\mathbf{x}$.

A.4. Proof of Theorem 4

Recall that the first $k = |\mathcal{T}|^2 - |\mathcal{T}|$ moments correspond to the inequality conditions, and the last q moments correspond to the measurement error centering conditions. Let $\gamma_I = (\gamma_j)_{j=1,\dots,k}$, $g_I = (g_j)_{j=1,\dots,k}$, $\gamma_M = (\gamma_j)_{j=k+1,\dots,k+q}$, and $g_M = (g_j)_{j=k+1,\dots,k+q}$ be sub-vectors of γ and g that correspond to inequality and the measurement error centering conditions, respectively.

Step 1. Take a sequence $\{\gamma_{I,l}\}_{l=1}^{+\infty}$ such that every component of $\gamma_{I,l}$ diverges to $+\infty$. Note that since g_I takes values in $\{-1, 0\}^k$,

$$\sup_{x,e} \left|\exp(\gamma'_{I,l}g_I(x,e)) - \mathbb{1}(g_I(x,e) = 0)\right| \le \exp(-\min_{i=1,\dots,k} \gamma_{I,l,i}) \to_{l \to +\infty} 0,$$

where $\gamma_{I,l,i}$ is the i-th component of $\gamma_{I,l}$. Hence, for any function $f \in L^1(\eta(\cdot|x))$

$$\left\|\int f(e)\exp(\gamma'_{I,l}g_I(x,e))d\eta(e|x) - \int f(e)\mathbb{1}(g_I(x,e) = 0)d\eta(e|x)\right\| \le$$
$$\le \exp(-\min_{i=1,\dots,k} \gamma_{I,l,i})\int \|f(e)\| d\eta(e|x) \to_{l \to +\infty} 0.$$

Hence, the sequence of measures $\exp(\gamma'_{I,l}g_I(x,\cdot))d\eta(\cdot|x)$ converges to the measure

$$\mathbb{1}(g_I(x,\cdot) = 0)d\eta(\cdot|x)$$

in total variation. The later measure is well-defined and nontrivial since we assume that $\tilde{E}|X = \{e : \mathbb{1}(g_I(x,e) = 0)\}$ has a positive measure under $\eta(\cdot|x)$. Let $d\tilde{\eta}(\cdot|x)$ denote $\mathbb{1}(g_I(x,\cdot) = 0)d\eta(\cdot|x)$.

Step 2. Consider the moment conditions under $d\tilde{\eta}(\cdot|x)$

$$\tilde{h}_M(x;\gamma) = \frac{\int_{e \in E|X} g_M(x,e)\exp(\gamma'g_M(x,e))d\tilde{\eta}(e|x)}{\int_{e \in E|X} \exp(\gamma'g_M(x,e))d\tilde{\eta}(e|x)}.$$

Definition 8.(iii) together with Assumption 3 and Step 1 imply that for any compact set $\Gamma \in \mathbb{R}^q$, uniformly in $\gamma_M \in \Gamma$

$$\left\|\mathbb{E}_{\pi_0}\left[h(\mathbf{x}; (\gamma'_{I,l}, \gamma'_M)')\right]\right\| = \left\|\mathbb{E}_{\pi_0}\left[\tilde{h}_M(\mathbf{x}; \gamma_M)\right]\right\| + o(1).$$

Thus, by continuity of h_M in γ_M, when l goes to infinity, we can work with the reduced optimization

problem:
$$\inf_{\gamma \in \mathbb{R}^q} \left\| \mathbb{E}_{\pi_0} \left[\tilde{h}_M(\mathbf{x}; \gamma) \right] \right\|. \tag{4}$$

Step 3. Note that (4) is equivalent to the optimization problem in Theorem 3. Hence, infimum in (4) is equal to 0 if and only if the data is approximately consistent with model m.

We assumed that every component of g_M takes both positive and negative values on some nonzero measure subsets of $\tilde{E}|X$ (Assumption 2). Hence, following the proof of Theorem 2.1 and Lemma A.1 in Schennach (2014), we can conclude that if infimum in (4) is equal to 0, then it is achieved at some finite and unique $\gamma_{0,M}$. Otherwise, $\|\gamma_M\|$ diverges to infinity.

A.5. Proof of Theorem 5

The result is a direct application of Theorem F.1 in Schennach (2014). For completeness of the proof we present the version of it that is applicable to our setting below.

Theorem (Theorem F.1, Schennach, 2014). *Let data be i.i.d.. If (i) the set*

$$\Gamma = \{\gamma \in \mathbb{R}^q : \mathbb{E} \left[\left\| \tilde{h}_M(\mathbf{x}, \gamma) \right\| \right] \leq C\}$$

is nonempty for some $C < \infty$; (ii) $\mathbb{E} \left[\left\| \tilde{h}_M(\mathbf{x}, \gamma) \right\|^2 \right] < \infty$ for all $\gamma \in \Gamma$, then

$$\lim_{n \to \infty} \mathbb{P} \left(\text{TS}_n > \chi^2_{q,\alpha} \right) \leq \alpha.$$

An i.i.d. sample is assumed. To show the validity of conditions (i) and (ii) note that since $\mathbf{x}$ has a bounded support (by Assumption 3) and $\tilde{\eta}$ satisfies conditions of Definition 8.(iii), for any finite γ there exist finite positive constant $C_1(\gamma)$ such that almost surely in $\mathbf{x}$

$$\left\| \tilde{h}_M(\mathbf{x}, \gamma) \right\|^2 \leq C_1(\gamma).$$

Hence, for any nonempty compact set Γ one can take $C = \sup_{\gamma \in \Gamma} C_1(\gamma)$. Together with Assumption 3, the later implies condition (ii). Similarly, one can use C to bound $\mathbb{E} \left[\left\| \tilde{h}_M(\mathbf{x}, \gamma) \right\| \right]$.

Under the alternative hypothesis, $\left\| \hat{\tilde{h}}_M(\gamma) \right\|$ either converges to a positive constant or diverges to infinity. Thus, since eigenvalues of $\tilde{\Omega}(\gamma)$ are bounded away from zero and are bounded from above the test is consistent.

A.6. Proof of Theorem 7

By Theorem 6 we have that the following inequalities hold a.s.:

$$\mathbf{v}_{t,A} - \mathbf{v}_{s,A} \geq \frac{1}{\mathbf{d}_A^t} [\boldsymbol{\rho}_{t,I}'(\mathbf{c}_{t,I}^* - \mathbf{c}_{t,B}^* - \mathbf{c}_{s,I}^* + \mathbf{c}_{s,B}^*) + \frac{\mathbf{p}_{t,H} - \mathbf{p}_{t,B}}{\prod_{j=1}^t (1 + \mathbf{r}_j)}'(\mathbf{c}_{t,H}^* - \mathbf{c}_{s,H}^*) \quad \forall t, s \in \mathcal{T},$$

$$\mathbf{v}_{t,B} - \mathbf{v}_{s,B} \geq \frac{1}{\mathbf{d}_B^t} [\boldsymbol{\rho}_{t,I}'(\mathbf{c}_{t,B}^* - \mathbf{c}_{s,B}^*) + \frac{\mathbf{p}_{t,B}}{\prod_{j=1}^t (1 + \mathbf{r}_j)}'(\mathbf{c}_{t,H}^* - \mathbf{c}_{s,H}^*) \quad \forall t, s \in \mathcal{T}.$$

Then we multiply the first inequality by $\mathbf{d}_A^t$, this random variable is positive a.s., so it does not alter the inequalities. We do the same for the second inequality, and multiply it by $\mathbf{d}_B^t$. Then we add-up the two inequalities, to obtain:

$$\mathbf{d}_A^t(\mathbf{v}_{t,A} - \mathbf{v}_{s,A}) + \mathbf{d}_B^t(\mathbf{v}_{t,B} - \mathbf{v}_{s,B}) \geq \boldsymbol{\rho}_{t,I}'(\mathbf{c}_t^* - \mathbf{c}_s^*) \quad \forall t, s \in \mathcal{T}.$$

B. Monte Carlo Experiments

In this section we study the behavior of our test in two Monte Carlo experiments. In the first one, we provide evidence for overrejection of the exponential discounting model by the deterministic test of Browning (1989) and correct coverage by our test. In the second experiment, we provide evidence for the power (consistency) of our test against some fixed alternatives. Finally, we conduct some robustness checks of our Markov Chain Monte Carlo (MCMC) integration.

B.1. Overrejection of Exponential Discounting for Browning's Deterministic Test

The objective of the Monte Carlo simulation exercise is to test the performance of the methodological procedure developed in this paper against the deterministic benchmark. We are going to provide evidence that a data set generated by a random exponential discounter, when contaminated with measurement error, will be erroneously rejected by deterministic methodologies at the individual level for a sizable fraction of the sample (Browning, 1989, Blow et al., 2017). However, our test will not reject it.

We choose our simulation configuration setup to match those of the household characteristics in our application. The Monte Carlo exercise will deal with a moderate size data set of $n = 2000$ individuals to show that it works in a data set of the roughly the same size as in our application. The time period is $\mathcal{T} = \{0, 1, 2, 3\}$, and we consider $L = 17$ goods. We use the same discounted prices $\{\rho_{i,t}\}_{i=1}^n$ as the ones given in Adams et al. (2014).[60] These are the prices faced by the single-individual/couples households in our application. We consider consumers with the constant

---

[60]We use the observed price matrix and sample from it uniformly with repetition at each Monte Carlo experiment.

elasticity of substitution (CES) instantaneous utility

$$u(c_t) = \sum_{l=1}^{L} \frac{c_{t,l}^{1-\sigma}}{1-\sigma},$$

where $\boldsymbol{\sigma} \sim U[1/15, 100]$ is heterogeneous across individuals.

Following Browning (1989), the true consumption rule for each consumer and each realized d is given by

$$c_{t,l}^* = \left(\frac{1}{d^t} \rho_{t,l}\right)^{-1/\sigma},$$

for all $l = 1, \cdots, L$ and $t \in \mathcal{T}$. For the discount factor we considered two different data generating processes (DGPs). For the first DGP $\mathbf{d} \sim U[0.8, 1]$ (DGP1) and for the second one $\mathbf{d} = 1$ a.s. (DGP2).

We perturb the generated consumption with $\boldsymbol{\epsilon}_{t,l} \sim$ i.i.d. $U[0.97, 1.03]$, which implies that $\mathbb{E}[\boldsymbol{\epsilon}_{t,l}] = 1$. That is, observed consumption is equal to true consumption times the multiplicative perturbation $\mathbf{c}_{t,l} = \mathbf{c}_{t,l}^* \boldsymbol{\epsilon}_{t,l}$. We define measurement error in consumption as $\mathbf{w}_{t,l}^c = \mathbf{c}_{t,l} - \mathbf{c}_{t,l}^*$ and fix $\mathbf{w}_t^p = 0$ a.s.. Note that the implied random measurement error $\mathbf{w}_{t,l}^c$ is mean-zero by construction and satisfies Assumption 1.1. The random vector perturbations $\boldsymbol{\epsilon}_{t,l}$ captures incorrect consumption reporting or recording, and can be as high as 1.03 times the true consumption. This means that relative measurement error is around 3 percent. This procedure produces a data set $(\boldsymbol{\rho}_{t,i}, \mathbf{c}_{t,i})_{i=1, t\in\mathcal{T}}^{i=n}$.

We replicate the experiment $m = 1000$ times for both DGPs. The results are presented in Table 2. For the deterministic test in Browning (1989) we use a grid search over d on $[0.1, 1]$ with a grid step 0.05. Searching over a smaller set (e.g., $[0.8, 1]$) will only weakly increase the rejection rate of the deterministic test. For DGP1 the deterministic test rejects the exponential discounting model in 53.04 percent of the cases on average across experiments. For DGP 2 the average rejection rate across experiments is 79.37 percent.

We use our methodology to test for s/ED-rationality for both DGPs assuming that the support of $\mathbf{d}$ is known. Assuming bigger support for $\mathbf{d}$ will only weakly decrease the rejection rate of our test. In other words, the design of our experiment favors the deterministic test. Nevertheless, our methodology cannot reject the correct null hypothesis that all households are consistent with s/ED-rationality at the 5 percent significance level. The rejection rates for DGP1 and DGP2 are 4.9 and 3 percent, respectively. As expected, both rejection rates are less than 5 percent.

Table 2 – Rejection Rates: ED-rationalizability. Number of replications $m = 1000$.

	rejection rate (%)	
	Deterministic test	Our methodology
DGP1	53.04	4.9
DGP2	79.37	3

B.2. Power Analysis

We choose our simulation configuration setup to match Section B.1. However, the consumer units are assumed to be couples whose behavior is described by the collective model with exponential discounting described in Adams et al. (2014). The individuals in the household are indexed by A and B. The random discount factors are $\mathbf{d}_A$ and $\mathbf{d}_B$. Individuals face different prices for good l at time period t given their bargaining power $\boldsymbol{\mu}_{t,l}$. We observe the sum of these two prices $\boldsymbol{\rho}_{t,l}$. That is, $\boldsymbol{\rho}_{t,l,A} = \boldsymbol{\mu}_{t,l}\boldsymbol{\rho}_{t,l}$, and $\boldsymbol{\rho}_{t,l,B} = (1 - \boldsymbol{\mu}_{t,l})\boldsymbol{\rho}_{t,l}$. Note that the bargaining power is good and time specific. The random price vectors $\boldsymbol{\rho}_t$ were drawn from the data set in Adams et al. (2014) as described in Section B.1.

Similar to the experimental design in Section B.1, the consumption rule for each individual and realized d_j, $j \in \{A, B\}$ is given by

$$c_{t,l,j}^* = \left(\frac{1}{d_j^t}\rho_{t,l,j}\right)^{-1/\sigma_{l,j}}, \quad l = 1, \cdots, L; \ t \in \mathcal{T},$$

where $\sigma_{l,j}$ is a realization of $\boldsymbol{\sigma}_{l,j} \sim$ i.i.d. $U[1/15, 100]$, $j \in \{A, B\}$. Then the household consumption data is the sum of individual consumption: $\mathbf{c}_t^* = \mathbf{c}_A^* + \mathbf{c}_B^*$ a.s.. The generating process for measurement error coincides with the one presented in Section B.1. As a result we generate $(\boldsymbol{\rho}_t, \mathbf{c}_t)_{t \in \mathcal{T}}$ and test whether this data is consistent with s/ED-rationality.

We consider two different DGPs for the distribution of $\mathbf{d}_j$, $j \in \{A, B\}$, and $\boldsymbol{\mu}_{t,l}$.
DGP3. $\mathbf{d}_A \sim U[0.1, 1]$, $\mathbf{d}_B \sim U[0.99, 1]$, and $\boldsymbol{\mu}_{t,l} = 1/2$ a.s.. Under this DGP household members face the same prices but may have different discount factors.
DGP 4. $\mathbf{d}_A \sim U[0.1, 1]$, $\mathbf{d}_B \sim U[0.99, 1]$, and $\boldsymbol{\mu}_{t,l} \sim$ i.i.d. $U[1/3, 2/3]$. Under this DGP household members face different prices and may have different discount factors.

We conducted the experiments with each DGP $m = 1000$ times for two sample sizes, $n = 2000$ and $n = 3000$. The supports of the discount factors were assumed unknown and contained inside $[0.1, 1]$ interval. The results are presented in Table 3.

Table 3 – Rejection Rates: Collective Model. Number of replications $m = 1000$.

	prices	discount factors	rejection rate (%) $n = 2000$	$n = 3000$
DGP3	same	different	31.6	69.5
DGP4	different	different	71.3	96.6

For DGP3 with equal bargaining power (same prices) and heterogeneous discount factors, the rejection rate is 31.6 percent for the sample size of $n = 2000$ and increases to 69.5 percent for $n = 3000$. For DGP4 with asymmetric bargaining power (different prices) and heterogeneous discount factors the rejection rate is even bigger and is equal to 71.3 and 96.6 percent for $n = 2000$ and $n = 3000$, respectively.

We highlight that DGP3 is compatible with hyperbolic discounting. It is easy to see that consumption behavior of the collective model with symmetric bargaining (i.e., same prices) satisfies the Afriat inequalities for hyperbolic discounting in Blow et al. (2013).

B.3. Robustness of MCMC integration.

Our testing procedure requires some user-specified parameters: the distribution η and the length of the MCMC chain. As mentioned in Section 5.2, the choice of η has no effect on the value of the test statistic both asymptotically and in finite samples. In other words the difference in values of the test statistics computed using two different η's can only be driven by numerical precision of the MCMC integration step and the optimization algorithm used. Thus, we focus on the performance of procedure for different MCMC chain length.

The results in Sections B.1 and B.2 were obtained using the chain length equal to cl = 10000. We decrease the chain length to cl = 5000 and for the sample size $n = 2000$ we additionally experiment with DGP2, DGP3, and DGP4. The remaining elements of the simulations remain the same as before.

Table 4 shows that halving the chain length from 1000 to 5000 changes very little the rejection rates of the three DGPs of interest. This is of course desirable as lack of robustness would suggest that the MCMC chain has not converged. This provides reassurance that our choice of chain length 10000 is appropriate.

Table 4 – Rejection Rates: ED and Collective Models. Sample size $n = 2000$.

| | rejection rate (%) | |
	cl = 10000	cl = 5000
DGP2	3	3.6
DGP3	31.6	32.8
DGP4	71.3	69.9

C. Computational Aspects

In this appendix we discuss the computational aspects of our procedure. In Appendix C.1 we provide a general pseudo-algorithm to implement our procedure. Appendix C.2 describes the MCMC procedure used for latent variable integration. Appendix C.3 provides a description of the "hit-and-run" algorithm we used in the construction of the MCMC chain. We provide the specification for η and the optimization routines used in our applications and simulations in

Appendix C.4.

C.1. Pseudo-Algorithm

This pseudo-algorithm is based on Schennach's algorithm provided in GAUSS as a supplement to Schennach (2014). The actual implementation of the algorithm has been vectorized and parallelized.

1: **Step** 0 (Setting parameters)

- Fix $\mathcal{T} = \{0, \cdots, T\}$, consumer experiments, and $\mathcal{L} = \{1, \cdots, L\}$, set of goods.

- Fix g_I and g_M.

- Fix Λ, the support of $(\boldsymbol{\lambda}_t)_{t \in \mathcal{T}}$, and Δ, the support of $(\boldsymbol{\delta}_t)_{t \in \mathcal{T}}$.

- Fix $\eta \in \mathcal{P}_{E|X}$ (See Appendix C.4 for details)

- Fix $x = (x_i)_{i=1,\ldots,n}$, where $x_i = (\rho_{i,t}, c_{i,t})_{t \in \mathcal{T}}$ is i-th observation and n is the sample size.

2: **end Step** 0.

3: **Step** 1(Integration: Evaluation of the objective function at a given $\gamma \in \mathbb{R}^{|\mathcal{T}|}$)

- Set $i = 1$.

4: **While** $i \leq n$

- Define the measure
$$\tilde{\eta}(\cdot|x_i) = \eta(\cdot|x_i)\mathbb{1}\left(g_I(x_i, \cdot) = 0 \right).$$

- Integrate latent variables using $\tilde{\eta}(\cdot|x_i)$ to obtain $\tilde{h}_M(x_i, \gamma)$ (See Appendix C.3 for implementation details).

- Set $i = i + 1$.

5: **end While.**

- Compute
$$\hat{\tilde{h}}_M(\gamma) = \frac{1}{n}\sum_{i=1}^{n}\tilde{h}_M(x_i, \gamma)$$

and
$$\hat{\tilde{\Omega}}(\gamma) = \frac{1}{n}\sum_{i=1}^{n}\tilde{h}_M(x_i, \gamma)\tilde{h}_M(x_i, \gamma)' - \hat{\tilde{h}}_M(\gamma)\hat{\tilde{h}}_M(\gamma)'.$$

- Compute $\text{ObjFun}(\gamma) = n\hat{\tilde{h}}_M(\gamma)'\hat{\tilde{\Omega}}(\gamma)^{-}\hat{\tilde{h}}_M(\gamma).$

6: **end Step** 1.

7: **Step** 3 (Optimization Step)

- Compute $\text{TS}_n = \min_\gamma \text{ObjFun}(\gamma).$

8: **end Step** 3.

C.2. Latent Variable Integration

Evaluation of the objective function requires integrating latent variables. We use MCMC methods. For completeness we provide the algorithm for MCMC integration to get $\tilde{h}_M(x_i, \gamma)$.

1: **Inputs**

- Fix cl – total MCMC chain length; nburn – number of "burned" chain elements;

- Fix η, γ, x_i, and the first element of the chain $e_{-\text{nburn}}$ that satisfies the constraints.

- Set $r = -\text{nburn} + 1$ and $\tilde{h}(x_i, \gamma)_M(\gamma) = 0$.

2: **While** $r \leq \text{nsims}$

- Draw $e_{\text{jump}} = ((v_t)_{t \in \mathcal{T}}, (\lambda_t)_{t \in \mathcal{T}}, (\delta_t)_{t \in \mathcal{T}}, w^c, w^\rho)$ proportional to $\tilde{\eta}(\cdot|x_i) = \eta(\cdot|x_i)\mathbb{1}\left(g_I(x_i, \cdot) = 0\right)$.

- Draw α from $U[0, 1]$.

- Set e_r equal to e_{jump} if $[g_M(x_i, e_{\text{jump}}) - g_M(x_i, e_{r-1})]'\gamma > \log(\alpha)$ and to e_r otherwise.

- **if** $r > 0$

- Compute $\tilde{h}_M(x_i, \gamma) = \tilde{h}_M(x_i, \gamma) + g_M(x_i, e_r))/\text{cl}$

- **end if**

- Set $r = r + 1$

5: **end While.**

To compute the chain, one always can use "rejection sampling": at every step check whether a candidate element of the chain satisfies the inequalities (support constraints). Since our constraints have a simple form, we propose to use a version of the "hit-and-run" algorithm that we describe below.

C.3. "Hit-and-run" Algorithm

Since we use the algorithm presented below in our application with survey data and for concreteness we focus on s/ED-rationalizability with consumption measurement error. Note that instead of working with measurement errors in consumption, we can equivalently work with true unobserved consumption c_t^*. Thus, the latent variables, $e = (d, (c_t^{*\prime}, v_t)_{t \in \mathcal{T}})$, have to satisfy the following set of constraints:

$$v_t - v_s \geq \frac{\rho_t'}{d^t}(c_t^* - c_s^*),$$

54

$$v_t, \ c_t^* \geq 0,$$
$$1 \geq d \geq \theta_0,$$

for all $s, t \in \mathcal{T}$.

The idea behind the "hit-and-run" algorithm is (i) to pick some initial point e^0 that satisfies the support constraints;[61] (ii) to construct a candidate point by moving along a random direction within the constrained set on a randomly chosen distance; (iii) to use a user-specified Monte-Carlo acceptance rule to assign to e^1 either the initial point e^0 or the candidate point; (iv) to apply steps (ii) and (iii) to e^1 to construct e^2; (v) to repeat until the length of the chain reaches user chosen number.

Take some arbitrary e^r that satisfies the constraints. Let ξ be a direction vector (not necessary unit vector). Thus, the candidate vector is

$$e^{r+1} = e^r + \alpha \xi,$$

where $\alpha \geq 0$ determines the scale of the perturbation $\alpha \xi$.

Sign Constraints. We start with sign constraints on consumption: $c_{t,l}^* \geq 0$ for all l and t. Let K_c be a set of indexes that correspond to $c_{t,l}^*$ in e^r. Hence, the constraints take the form

$$\alpha \xi_k \geq -e_k^r, \quad \forall k \in K_c$$

Define $K_+ = \{k \in K_c : \xi_k > 0\}$, $K_- = \{k \in K_c : \xi_k < 0\}$, and $K_0 = \{k \in K_c : \xi_k = 0\}$. Then, the sign constraints are

$$\alpha \geq -\frac{e_k^r}{\xi_k}, \quad \forall k \in K_+,$$
$$\alpha \leq -\frac{e_k^r}{\xi_k}, \quad \forall k \in K_-.$$

Note that the constraints that correspond to $k \in K_0$ are always satisfied since $e_k^r \geq 0$ (i.e. satisfies the constraints). Thus, the sign constraints can be simplified to

$$\min_{k \in K_-} -\frac{e_k^r}{\xi_k} \geq \alpha \geq \max_{k \in K_+} -\frac{e_k^r}{\xi_k}, \tag{5}$$

where $\min_{k \in \emptyset} = +\infty$, and $\max_{k \in \emptyset} = -\infty$.

Afriat Constraints. Next, we consider the Afriat inequalities. Let $e^r(v, t)$ and $e^r(c, t)$ be the components of e^r that correspond to v_t and c_t, respectively. Assume that d is fixed. Then, the Afriat inequalities are

$$e^r(v, t) - e^r(v, s) \geq \frac{\rho_t'}{d^t}(e^r(c, t) - e^r(c, s)), \forall t, s,$$

---

[61] In our application and simulations we computed initial point by minimizing the norm of g_M subject to the Afriat and sign constraints per observation.

Thus, since d is fixed (the component of ξ that corresponds to d is zero), after plugging in e^{r+1} we get

$$e^r(v,t) - e^r(v,s) + \alpha(\xi(v,t) - \xi(v,s)) \geq \rho_t'(e^r(c,t) - e^r(c,s))/d^t + \alpha\rho_t'(\xi(c,t) - \xi(c,s))/d^t, \forall t,s,$$

where $\xi(v,t)$ and $\xi(c,t)$ are the components of ξ that correspond to v_t and c_t, respectively. Hence,

$$\alpha P_t'(\xi(t) - \xi(s)) \leq -P_t'(e^r(t) - e^r(s)), \forall t, s,$$

where $P_t = (-1, \rho_t'/d^t)'$, $\xi(t) = (\xi(v,t), \xi(c,t)')'$, and $e^r(t) = (e^r(v,t), e^r(c,t)')'$. Similarly to box constraints, we can define $T_+^A = \{(t,s) \ : \ P_t'(\xi(t) - \xi(s)) > 0\}$, $T_-^A = \{(t,s) \ : \ P_t'(\xi(t) - \xi(s)) < 0\}$, and $T_0^A = \{(t,s) \ : \ P_t'(\xi(t) - \xi(s)) = 0\}$. Thus, the Afriat constraints are

$$\min_{(t,s)\in T_+^A} -\frac{P_t'(e^r(t) - e^r(s))}{P_t'(\xi(t) - \xi(s))} \geq \alpha \geq \max_{(t,s)\in T_-^A} -\frac{P_t'(e^r(t) - e^r(s))}{P_t'(\xi(t) - \xi(s))}. \tag{6}$$

In other words we characterized possible perturbations of c_t and v_t given d that are allowed under Afriat constraints. Next we want to characterize the set for d given consumption and utility numbers. Note that by assumption $d \in [\theta_0, 1]$ and that

$$d^t(v_t - v_s) \geq \rho_t(c_t - c_s).$$

Define $T_{d,+}^A = \{(t,s) : (v_t - v_s) > 0\}$, $T_{d,-}^A = \{(t,s) : (v_t - v_s) < 0\}$, and $T_{d,0}^A = \{(t,s) : (v_t - v_s) = 0\}$. Hence, the Afriat ineqaulities are equivalent to

$$\min\left\{\min_{(t,s)\in T_{d,-}^A}\left(\max\left\{\frac{\rho_t(c_t - c_s)}{(v_t - v_s)}, 0\right\}\right)^{1/t}, 1\right\} \geq d \geq \max\left\{\max_{(t,s)\in T_{d,+}^A}\left(\max\left\{\frac{\rho_t(c_t - c_s)}{(v_t - v_s)}, 0\right\}\right)^{1/t}, \theta_0\right\}. \tag{7}$$

Inequalities (5)-(7) give sharp restrictions on α that would guarantee that the next draw e^{r+1} satisfies the constraints. Below we provide algorithms how to generate (i) new consumption and utility numbers given prices and discount factor, and (ii) discount factor given prices, consumption, and utility numbers.

1: **Generating new consumption vector and utility numbers**

- Fix the discount factor and prices.

- Draw a random direction vector ξ from a uniform distribution on the $[|\mathcal{T}| + |\mathcal{T}| \cdot L]$-dimensional unit sphere.

- Compute the interval A using (5) and (6).

- Draw α uniformly from A.

- Generate new consumption vectors and utility numbers using ξ and α.

2: **Generating new discount factor**

56

- Fix prices, consumption, and utility numbers.

- Uniformly draw d from the interval that satisfies (7).

Thus, we can propose the two approaches to sample from the cone characterized by the Afriat and the sign constraints. If one decides to keep the same d for generating the chain, then one can initially draw several independent draws of d, and for every realization of d generate its own chain. The second approach can be thought of as "double-hit-and-run": first generate new consumption and utility numbers, and then generate new discount factor using these new consumption and utility numbers. In our application we use double-hit-and-run approach.

C.4. User-specified $\tilde{\eta}$

In this section we specify a particular choice of η used in our applications and simulations.

When integrating measurement error, instead of drawing measurement error (e.g., $\mathbf{w}_t^c$), we draw unobserved true variable (e.g., $\mathbf{c}_t^*$) and then constructed the measurement error by taking the difference between observed mismeasured and latent true variables (e.g., $\mathbf{w}_t^c = \mathbf{c}_t - \mathbf{c}_t^*$). Note that working with true variables allows us to easily generate measurement errors that imply correct signs for true variables (e.g. $\mathbf{c}_t^* \geq 0$). In particular, in our applications and simulations we impose sign constraints directly in the sampling stage (Step 2 in Section C.2).

For our first application (survey data) to build $\tilde{\eta}$ we used the "hit-and-run" algorithm described in Appendix C.3 to produce draws of $\mathbf{e}$. In particular, $\mathbf{c}^* = (\mathbf{c}_t^*)_{t \in \mathcal{T}}$ is such that (i) it satisfies the Afriat-like inequalities and sign constraints, (ii) the user specified distribution over $\mathbf{w}^c = (\mathbf{w}_t^*)_{t \in \mathcal{T}}$ is

$$d\tilde{\eta}(w^c|x) \propto \exp(- \|g_M(x,e)\|^2),$$

where $g_M(x,e) = (\rho_t' w_t^c)_{t \in \mathcal{T}}$. To achieve this, we use the standard Metropolis Hastings algorithm in each step of the "hit-and-run" algorithm to get the draws from the desired distribution. Note that by construction this distribution has the correct support $E|X$.

The $\tilde{\eta}$ distribution can be adapted to accommodate other moments such as those in our extensions and counterfactual analysis by using the appropriate moment conditions. If the moment conditions, which are not support constraints, include other random variables in $\mathbf{e}$, the distribution $\tilde{\eta}$ will have to be defined on them, and not only on $\mathbf{w}^c$ like in our first application.

For our second application (experimental data) we use a different strategy since (i) the panel is long ($T = 50$), (ii) the centering conditions do not depend on v_t and δ_t, and (iii) the static UMT has a simplified characterization in terms of Generalized Axiom of Revealed Preferences (GARP).[62] Hence, we can simplify our problem by considering a reduced latent random vector that consists only of true consumption or true prices. We then choose $\tilde{\eta}$ to be a uniform distribution

---

[62] One can replace the Afriat inequalities by the GARP inequalities since GARP is equivalent to R-rationalizability in this case.

over consumption or prices that satisfy GARP and that produce an expenditure level equal to 1.[63] We can do this with the support constraints that we consider in this application. The key for a good computational performance of this step is to check for GARP consistency in an efficient way for each candidate draw of prices or consumption. For this purpose, we use a recursive algorithm to check GARP using an implementation of the deep-first search algorithm with recursive tabu search (see Boelaert, 2014).

In both applications it is trivial to verify that these choices of $\tilde{\eta}$ satisfy the conditions stated in Definition 7.

C.5. Optimization

We optimize the objective function specified in the pseudo-algorithm using Bobyqa procedure as implemented in the NLopt library following Powell (2009). As an initial guess for the optimization we use the outcome of applying BlackBox Differential Evolution Algorithm to minimize our objective function. Bobyqa performs derivative-free optimization using iteratively constructed quadratic approximations of the objective. We observe that in our simulations this combination of optimizers perform the best in terms of accuracy and speed among similar NLopt alternatives.

For the second application because the number of moments is larger we use as an initial guess for the optimization the outcome of two-step GMM estimator. Since the objective function of the two-step GMM estimator has a unique minimizer and is locally convex around it, we use Bobqya here as well. Bobqya works well in convex problems as documented in Rios & Sahinidis (2013). Following Schennach (2014), we additionally verified our results using Neldermead.[64]

Another alternative to find good initial values is taking advantage of a convex problem related to our problem. As shown in Schennach (2014), the moment condition in Theorem 3 is a first-order condition of the following convex optimization problem (Lemma A.1 in Schennach, 2014):

$$\min_{\gamma \in \mathbb{R}^q} \mathbb{E}_{\pi_0} \left[\ln \mathbb{E}_{\tilde{\eta}} \left[\exp(\gamma' g_M(\mathbf{x}, \mathbf{e})) | \mathbf{x} \right] \right].$$

Moreover, the norm $\| \mathbb{E}_{\pi_0} \left[h(\mathbf{x}; \cdot) \right] \|$ has a unique global minimum, is convex in the neighborhood of the minimizer if this minimizer is finite, and has no other local minima. Hence, computationally the problem is convenient.

---

[63]We impose nonnegativity constraints on consumption and positivity constraints on prices. The requirement to produce expenditure level equal to 1 makes the support of consumption and price bounded.

[64]See Sasaki (2015) for an alternative optimization technique.

D. Analytical Power Results. Robustness to Local Perturbations

In this appendix we provide examples of DGPs that will fail to pass our test (Sections D.1 and D.2). In Section D.3 we show robustness of UMTs that we consider to local perturbations in observed quantities or prices.

D.1. s/ED-Rationalizability, Mean-budget Neutrality, Price and Consumption Measurement Error

In this section we construct the data set that can not be s/ED-rationalized by measurement error in consumption and time invariant measurement error in prices if the centering condition comes in the form

$$\mathbb{E}\left[\boldsymbol{\alpha}\boldsymbol{\rho}_t'\mathbf{c}_t\right] = \mathbb{E}\left[\boldsymbol{\alpha}\boldsymbol{\rho}_t^{*\prime}\mathbf{c}_t^*\right], \quad \forall t \in \mathcal{T},$$

where $\boldsymbol{\alpha} \in (0,1]$, represents individual specific weights. We consider the environment with 2 time periods and 2 goods. We assume that the price measurement error comes in the following form:

$$\boldsymbol{\rho}_t^* = \widetilde{\mathbf{W}}\boldsymbol{\rho}_t,$$

where

$$\widetilde{\mathbf{W}} = \begin{pmatrix} \widetilde{\mathbf{w}}_1^p & 0 \\ 0 & \widetilde{\mathbf{w}}_2^p \end{pmatrix}$$

is the matrix of time invariant multiplicative price measurement errors.

The above centering condition covers variety of measurement error. For instance, if $\boldsymbol{\alpha} = 1$ a.s., then we have the centering condition used in our application. The random weight $\boldsymbol{\alpha}$ is allowed to be correlated with all observables and measurement error.

Take $\{\mathbf{c}_t\}_{t=0,1}$ and $\{\boldsymbol{\rho}_t\}_{t=0,1}$ such that

$$\boldsymbol{\rho}_{0,1} = \boldsymbol{\rho}_{0,2} = 1 \text{ a.s.}, \quad \boldsymbol{\rho}_{1,1} = \boldsymbol{\rho}_{1,2} = 2 \text{ a.s.},$$
$$\mathbf{c}_{0,1} = \mathbf{c}_{0,2} = 1 \text{ a.s.}, \quad \mathbf{c}_{1,1} = \mathbf{c}_{1,2} = 2 \text{ a.s.}.$$

Denote $\rho_0 = (1,1)'$, $\rho_1 = 2(1,1)'$, $c_0 = (1,1)'$, and $c_1 = 2(1,1)'$. By way of contradiction suppose that there exist $\mathbf{d} \in (0,1]$, $\{\mathbf{c}_t^*, \boldsymbol{\rho}_t^*\}_{t=0,1}$, $\boldsymbol{\alpha} \in (0,1]$, nonnegative $\{\mathbf{v}_t\}_{t=0,1}$ such that the Afriat inequalities and the centering conditions are satisfied:

$$\mathbf{v}_1 - \mathbf{v}_0 \geq \frac{\boldsymbol{\rho}_1^{*\prime}}{\mathbf{d}}(\mathbf{c}_1^* - \mathbf{c}_0^*) \text{ a.s.}, \quad \mathbf{v}_0 - \mathbf{v}_1 \geq \boldsymbol{\rho}_0^{*\prime}(\mathbf{c}_0^* - \mathbf{c}_1^*) \text{ a.s.},$$
$$\mathbb{E}\left[\boldsymbol{\alpha}\boldsymbol{\rho}_0'\mathbf{c}_0\right] = \mathbb{E}\left[\boldsymbol{\alpha}\boldsymbol{\rho}_0^{*\prime}\mathbf{c}_0^*\right], \quad \mathbb{E}\left[\boldsymbol{\alpha}\boldsymbol{\rho}_1'\mathbf{c}_1\right] = \mathbb{E}\left[\boldsymbol{\alpha}\boldsymbol{\rho}_1^{*\prime}\mathbf{c}_1^*\right],$$
$$\boldsymbol{\rho}_0^* = \widetilde{\mathbf{W}}\rho_0, \quad \boldsymbol{\rho}_1^* = \widetilde{\mathbf{W}}\rho_1.$$

Note that since $\mathbf{d} > 0$ with probability 1, the inequalities can be rewritten as

$$\frac{\mathbf{d}}{2}(\mathbf{v}_1 - \mathbf{v}_0) \geq \rho_0' \widetilde{\mathbf{W}}(\mathbf{c}_1^* - \mathbf{c}_0^*) \text{ a.s.,}$$

$$\mathbf{v}_0 - \mathbf{v}_1 \geq \rho_0' \widetilde{\mathbf{W}}(\mathbf{c}_0^* - \mathbf{c}_1^*) \text{ a.s..}$$

Hence,

$$\frac{\mathbf{d}}{2}(\mathbf{v}_1 - \mathbf{v}_0) \geq \mathbf{v}_1 - \mathbf{v}_0.$$

Thus, since $\mathbf{d} \in (0, 1]$, we can deduce that $\mathbf{v}_1 - \mathbf{v}_0 \leq 0$ a.s. If we multiply the first Afriat inequality by strictly positive $\boldsymbol{\alpha}$ and take expectations from both sides, we get

$$0 \geq \mathbb{E}\left[\boldsymbol{\alpha}\frac{\mathbf{d}}{2}(\mathbf{v}_1 - \mathbf{v}_0)\right] \geq \mathbb{E}[\boldsymbol{\alpha}\rho_0'\mathbf{W}\mathbf{c}_1^*] - \mathbb{E}[\boldsymbol{\alpha}\rho_0'\mathbf{W}\mathbf{c}_0^*] =$$

$$\frac{\mathbb{E}[\boldsymbol{\alpha}\rho_1'\mathbf{W}\mathbf{c}_1^*]}{2} - \mathbb{E}[\boldsymbol{\alpha}\rho_0^{*\prime}\mathbf{c}_0^*] = \frac{\mathbb{E}[\boldsymbol{\alpha}\rho_1^{*\prime}\mathbf{c}_1^*]}{2} - \mathbb{E}[\boldsymbol{\alpha}\rho_0^{*\prime}\mathbf{c}_0^*] =$$

$$\frac{\mathbb{E}[\boldsymbol{\alpha}\rho_1'\mathbf{c}_1]}{2} - \mathbb{E}[\boldsymbol{\alpha}\rho_0'\mathbf{c}_0] = 4\mathbb{E}[\boldsymbol{\alpha}] - 2\mathbb{E}[\boldsymbol{\alpha}] = 2\mathbb{E}[\boldsymbol{\alpha}] > 0,$$

where the equalities come from from the centering conditions and the fact that $\rho_1 = 2\rho_0$, and the last inequality is implied by $\mathbb{E}[\boldsymbol{\alpha}] > 0$. The above contradiction implies that the constructed data set will never pass our test.

There are at least two implications of the example constructed in this section. First, we can further restrict $\boldsymbol{\alpha}$ by assuming that $\boldsymbol{\alpha} = 1$ a.s.. Thus, the above example demonstrates that without price measurement error the centering condition $\mathbb{E}[\rho_t'\mathbf{w}_t^c] = 0$, $t \in \mathcal{T}$ has empirical content. Second, note that the trembling-hand centering condition $(\mathbb{E}[\mathbf{w}_t^c] = 0, t \in \mathcal{T})$ implies that in our example $\mathbb{E}[\rho_t'\mathbf{w}_t^c] = 0$, $t \in \mathcal{T}$, since ρ_t has a degenerate distribution. Hence, the trembling-hand centering condition has empirical content as well.

We conclude this section by noting that our example can be used to construct an example with time invariant consumption measurement error and time varying price measurement error because the mean budget neutrality condition and the Afriat inequalities are "symmetric" in prices and consumption.

D.2. GARP and Trembling-Hand Measurement Error in Consumption or Prices

In the experimental data we use individuals are forced to pick points on the budget lines. That is, $\rho_t^* \mathbf{c}_t^* = \rho_t \mathbf{c}_t$ a.s. for all $t \in \mathcal{T}$. In this section we construct an example for the GARP with trembling-hand error in consumption. Consider 2 goods and 2 time periods environment with deterministic prices.

$$p_0 = (1, 2)', \quad p_1 = (2, 1)'.$$

The observed consumption vectors are random and satisfy

$$
\mathbf{c}_0 = \begin{cases} (\epsilon, 1 - \epsilon/2)', & \text{with probability } 1/2, \\ (\epsilon, 3/4 - \epsilon/2)', & \text{with probability } 1/2, \end{cases} \qquad \mathbf{c}_1 = \begin{cases} (1 - \epsilon/2, \epsilon)', & \text{with probability } 1/2, \\ (3/4 - \epsilon/2, \epsilon)', & \text{with probability } 1/2, \end{cases}
$$

where $0 < \epsilon < 1/8$. Hence,

$$
\mathbb{E}\left[\mathbf{c}_0\right] = (\epsilon, 7/8 - \epsilon/2)', \quad \mathbb{E}\left[\mathbf{c}_1\right] = (7/8 - \epsilon/2, \epsilon)'.
$$

Next, note that observed disposable income $\mathbf{y}_t$ is a binary random variable:

$$
\mathbf{y}_0 = \begin{cases} 2, & \text{with probability } 1/2, \\ 3/2, & \text{with probability } 1/2, \end{cases} \qquad \mathbf{y}_1 = \begin{cases} 2, & \text{with probability } 1/2, \\ 3/2 & \text{with probability } 1/2. \end{cases}
$$

First time period. Since mismeasured consumption has to belong to the true budget line ($\mathbf{p}_t'\mathbf{c}_t = \mathbf{p}_t\mathbf{c}_t^*$ a.s. for all t) and on average has to agree with the observed consumption, we can conclude that

$$
\mathbb{P}\left(\mathbf{c}_{01}^* + 2\mathbf{c}_{02}^* = 3/2 | \mathbf{y}_0 = 3/2\right) = 1, \quad \mathbb{P}\left(\mathbf{c}_{01}^* + 2\mathbf{c}_{02}^* = 2 | \mathbf{y}_0 = 2\right) = 1,
$$
$$
\mathbb{E}\left[\mathbf{c}_{01}^*\right] = \epsilon, \quad \mathbb{E}\left[\mathbf{c}_{02}^*\right] = 7/8 - \epsilon/2.
$$

Note that

$$
\mathbb{E}\left[\mathbf{c}_{01}^*\right] = \mathbb{E}\left[\mathbf{c}_{01}^* | \mathbf{y}_0 = 3/2\right] \mathbb{P}\left(\mathbf{y}_0 = 3/2\right) + \mathbb{E}\left[\mathbf{c}_{01}^* | \mathbf{y}_0 = 2\right] \mathbb{P}\left(\mathbf{y}_0 = 2\right)
$$
$$
= \frac{\mathbb{E}\left[\mathbf{c}_{01}^* | \mathbf{y}_0 = 3/2\right]}{2} + \frac{\mathbb{E}\left[\mathbf{c}_{01}^* | \mathbf{y}_0 = 2\right]}{2}.
$$

Hence, since $\mathbf{c}_t^* \geq 0$: a.s., $\mathbb{E}\left[\mathbf{c}_{01}^* | \mathbf{y}_0\right] \leq 2\epsilon$ a.s.. Thus, given that $\mathbf{p}_t'\mathbf{c}_t = \mathbf{p}_t\mathbf{c}_t^*$ a.s. for all t we get that $\mathbf{c}_0^* \in [0, 2\epsilon] \times [3/4 - \epsilon, 1]$ with positive probability. Similarly, $\mathbf{c}_1^* \in [3/4 - \epsilon, 1] \times [0, 2\epsilon]$ with positive probability. Thus, since $3/2 - 4\epsilon > 1$ ($\epsilon < 1/8$) it means that $\mathbf{c}_0^*$ is also available at $t = 1$ with positive probability. Similarly, $\mathbf{c}_1^*$ is available when $t = 0$ with positive probability. The latter violates GARP with positive probability (GARP has to be satisfied with probability 1). Thus, there is no trembling-hand measurement error that keeps consumption on the same budget and is consistent with GARP.

We conclude this section by noting that GARP conditions are symmetric in terms of price and consumption vectors. Thus, after relabeling (swapping prices with consumption) the above DGP also will not pass our test if one assumes that there is only mean-zero measurement error in prices.

D.3. Robustness to Local Perturbations

In this section we show that in many situations our approach is also robust to small measurement errors in observed quantities or prices. Suppose that we fix a model (i.e. the support restrictions on $\boldsymbol{\delta}_t$ and $\boldsymbol{\lambda}_t$, and the definition of $\boldsymbol{\rho}_t^*$). Define the measure of inequality slackness

$$\boldsymbol{\varepsilon}_{t,s}^* = \boldsymbol{\xi}_{t,s} - \boldsymbol{\rho}_t^*(\mathbf{c}_t^* - \mathbf{c}_s^*) \text{ a.s.}$$

where

$$\boldsymbol{\xi}_{t,s} = \frac{\boldsymbol{\delta}_t}{\boldsymbol{\lambda}_t}(\mathbf{v}_t - \mathbf{v}_s).$$

Suppose that there exist $\{\mathbf{v}_t, \boldsymbol{\delta}_t, \boldsymbol{\lambda}_t\}_{t \in \mathcal{T}}$ such that $\boldsymbol{\xi}_{t,s} \geq 0$. Next we perturb the true consumption and prices in order to see to what extent the RP inequalities are still valid. Note that the observed potentially mismeasured data $\{\boldsymbol{\rho}_t, \mathbf{c}_t\}_{t \in \mathcal{T}}$ satisfies the constraints with the same $\boldsymbol{\xi}_{t,s}$ if

$$\boldsymbol{\varepsilon}_{t,s}^* - \left[\mathbf{w}_t^{p\prime}(\mathbf{c}_t - \mathbf{c}_s) + \boldsymbol{\rho}_t'(\mathbf{w}_t^c - \mathbf{w}_s^c) + \mathbf{w}_t^{p\prime}(\mathbf{w}_s^c - \mathbf{w}_t^c)\right] \geq 0 \text{ a.s.}$$

for all t and s. Define

$$\boldsymbol{\alpha}^p = \max_t \|\mathbf{w}_t^p\|, \, \boldsymbol{\alpha}^c = \max_t \|\mathbf{w}_t^c\|,$$
$$\boldsymbol{\beta}^p = \max_t \|\boldsymbol{\rho}_t\|, \, \boldsymbol{\beta}^c = \max_t \|\mathbf{c}_t\|.$$

Then by the triangular inequality, the following inequality provides a sufficient restriction on the maximal perturbations of consumption and prices that will not refute the correctly specified model.

$$\boldsymbol{\varepsilon}_{t,s}^* \geq 6 \max\{\boldsymbol{\alpha}^p \boldsymbol{\beta}^c, \boldsymbol{\alpha}^c \boldsymbol{\beta}^p, \boldsymbol{\alpha}^p \boldsymbol{\alpha}^c\} \text{ a.s..}$$

In other words, if the UMT leads to an "interior solution" (i.e. $\boldsymbol{\varepsilon}_{t,s}^* > 0$ for all $t \neq s$), then small measurement errors without any centering restrictions will not affect the conclusions based on our testing procedure.

E. Extensions of s/ED-Rationalizability

In this appendix we show that our methodology can cover two important extensions of ED-rationalizability discussed in the main text: (i) ED-rationalizability with income uncertainty (Appendix E.1); (ii) the collective model of Adams et al. (2014) (Appendix E.2).

E.1. Income Uncertainty

In this section we consider a model of dynamic utility maximization with exponential discounting and income uncertainty. We start with the analysis of the deterministic model and then extend it to stochastic environments.

Definition 9 (Dynamic UMT with income uncertainty, ED-IU-rationalizability)**.** A deterministic array $(p_t, r_t, c_t)_{t \in \mathcal{T}}$ is ED-rationalizable in the presence of income uncertainty (ED-IU rationalizable) if: (i) There exists a concave, locally nonsatiated, and continuous function u. (ii) There exists a random income stream $\mathbf{y} = (\mathbf{y}_t)_{t \in \mathcal{T}}$. (iii) There exists an array of consumption and saving (policy) functions $(c_t(\cdot))_{t \in \mathcal{T}}$ and $(s_t(\cdot))_{t \in \mathcal{T}}$ such that $c_t : \mathbb{R}_+^{|\mathcal{T}|} \to \mathbb{R}_+^L \setminus \{0\}$ and $s_t : \mathbb{R}_+^{|\mathcal{T}|} \to \mathbb{R}_+$ for all $t \in \mathcal{T}$. In addition, we restrict these functions to depend only on the income history. That is, for all t, $c_t(y') = c_t(y)$ and $s_t(y') = s_t(y)$ for all y and y' such that $y'_\tau = y_\tau$ for all $\tau \leq t$. (iv) The consumption and saving policy functions maximize the expected flow of instantaneous utilities given the budget constraints and history of incomes captured by information I_t:

$$\max_{\{c_\tau(\cdot), s_\tau(\cdot)\}_{\tau=t,\dots,T}} \mathbb{E}\left[\sum_{\tau=t}^{T} \delta^{\tau-t} u(c_\tau(\mathbf{y})) \Big| I_t\right]$$

subject to

$$p'_\tau c_\tau(\mathbf{y}) + s_t(\mathbf{y}) = \mathbf{y}_\tau + (1 + r_\tau)s_{\tau-1}(\mathbf{y}) \text{ a.s.,}$$

for all $\tau = t, t+1, \dots, T$. (v) The consumption stream $(c_t)_{t \in \mathcal{T}}$ at every time period t is equal to the consumption policy function evaluated at a realization $y = (y_t)_{t \in \mathcal{T}}$ of the random income stream (i.e., $c_t = c_t(y)$). (vi) There is initial level of savings s_0.

ED-IU-rationalizability extends ED-rationalizability in an important direction: in accommodates possible uncertainty in the future income. Since the future income is unobserved, instead of a fixed vector of future consumption and savings, the agent has to come up with the whole consumption and saving functions in order to be ready for all possible realizations of the income stream.

In the case of income uncertainty we can still use the first-order conditions approach with an important modification. Instead of considering a support constraint on the space of marginal utility of wealth, we restrict its law of motion.

Lemma 3 (FOC for ED-IU-rationalizability)**.** *A deterministic array $(p_t, r_t, c_t)_{t \in \mathcal{T}}$ is ED-IU rationalizable if and only if there exists a concave, locally nonsatiated, and continuous function u, a discount factor $d \in (0, 1]$, and a positive random vector $(\boldsymbol{\lambda}_t)_{t \in \mathcal{T}}$ such that:*

(i) $\mathbb{E}[\boldsymbol{\lambda}_{t+1} | I_t] = \boldsymbol{\lambda}_t$ *a.s., where I_t the information (σ-algebra) generated by $(\boldsymbol{\lambda}_\tau)_{\tau \leq t}$.*

(ii) $d^t \nabla u(c_t) \leq \boldsymbol{\lambda}_t \rho_t$ *a.s.. If $c_{t,j} \neq 0$, then $d^t \nabla u(c_t)_j = \boldsymbol{\lambda}_t \rho_{t,j}$ a.s., where $c_{t,j}$, $\nabla u(c_t)_j$, and $\rho_{t,j}$ are the j-th components of c_t, $\nabla u(c_t)$, and ρ_t, respectively, and $\rho_t = p_t / \prod_{\tau=1}^{t}(1 + r_\tau)$.*

Proof. At every time period t the agent is maximizing the expected flow of instantaneous utilities

given the budget constraints and history of incomes captured by I_t:

$$\max_{\{c_\tau(\cdot),s_\tau(\cdot)\}_{\tau=t,\ldots,T}} \mathbb{E}\left[\sum_{\tau=t}^{T} \delta^{\tau-t} u(c_\tau(\mathbf{y})) \Big| I_t\right]$$

subject to

$$p_\tau' c_\tau(\mathbf{y}) + s_t(\mathbf{y}) = \mathbf{y}_\tau + (1+r_\tau)s_{\tau-1}(\mathbf{y}) \text{ a.s.,}$$

for all $\tau = t, t+1, \ldots, T$.

The Lagrangian function of the above problem takes the form

$$\mathbb{E}\left[\sum_{\tau=t}^{T} \delta^{\tau-t} u(c_\tau(\mathbf{y})) \Big| I_t\right] - \sum_{\tau=t}^{T} \mathbb{E}\left[\frac{\lambda_\tau(\mathbf{y})}{\delta^t \prod_{j=1}^{\tau}(1+r_j)} \left[p_\tau' c_\tau(\mathbf{y}) + s_t(\mathbf{y}) - \mathbf{y}_\tau - (1+r_\tau)s_{\tau-1}(\mathbf{y})\right] \Big| I_t\right],$$

where $\{\lambda_\tau(\cdot)\}_{\tau=t,\ldots,T}$ are lagrange multipliers. The denominator $\delta^t \prod_{j=1}^{\tau}(1+r_j)$ is needed for scaling of $\lambda_\tau(\cdot)$. If the instantenious utility function is concave, then, since the constraints are convex, the first-order condition will provide necessary and sufficient conditions for c_t and s_t to be optimal. The first-order condition with respect to c_τ is

$$\mathbb{E}\left[\left[\delta^\tau \nabla u(c_\tau(\mathbf{y})) - \lambda_\tau(\mathbf{y})\rho_\tau\right]' v_{c,\tau}(\mathbf{y}) \Big| I_t\right] = 0,$$

for all $t \in \mathcal{T}$, $\tau = t, \ldots, T$, and functions $v_{c,\tau}$, where $\rho_\tau = p_\tau / \prod_{j=1}^{\tau}(1+r_j)$. Note that, since for any $j = 1, \ldots, L$ the first order condition with respect to c_τ is satisfied with $v_c(\cdot) = (\mathbb{1}(i=j))_{i=1,\ldots,L}$, we have that the first order condition with respect to c_τ is satisfied if and only if

$$\delta^t \nabla u(c_t(\mathbf{y})) = \lambda_t(\mathbf{y})\rho_t \text{ a.s.,}$$

for all $t \in \mathcal{T}$.

Next, consider the first order condition with respect to s_τ:

$$\mathbb{E}\left[\left[\lambda_{\tau+1}(\mathbf{y}) - \lambda_\tau(\mathbf{y})\right] v_{s,\tau}(\mathbf{y}) \Big| I_t\right] = 0$$

for all $t \in \mathcal{T}$, $\tau = t, \ldots, T$, and functions $v_{s,\tau}$. Because of the law of iterated expectations the later is equivalent to

$$\mathbb{E}\left[\lambda_{t+1}(\mathbf{y}) - \lambda_t(\mathbf{y}) \Big| I_t\right] = 0$$

for all $t \in \mathcal{T}$, since $v_{s,\tau}(\cdot)$ only depends on the history up to moment τ. ∎

The first corollary of the lemma above is that without imposing any restriction on income shocks, at the population level, it is impossible to discern whether an array $(p_t, r_t, c_t)_{t \in \mathcal{T}}$ is ED-IU-rationalizable or R-rationalizable. The reason is that the only implication at the individual level of ED-IU-rationalizability is that the marginal utility of income is positive.[65] However, if we assume that the latent income shocks are i.i.d., and there is no aggregate shocks, in addition to

---

[65]This observation seems to have been noticed first by Adams et al. (2014).

assuming that preferences are i.i.d. and stable in the time window of interest, we can still have testable implications of the model. This statistical version of the ED-IU model is defined next.

Definition 10 (s/ED-IU-rationalizability). A random array $(\mathbf{p}_t^*, \mathbf{r}_t^*, \mathbf{c}_t^*)_{t \in \mathcal{T}}$ is s/ED-IU-rationalizable if there exists a tuple $(\mathbf{u}, (\boldsymbol{\lambda}_t, \mathbf{d})_{t \in \mathcal{T}})$ such that

(i) $\mathbf{u}$ is a random, concave, locally nonsatiated, and continuous utility function;

(ii) $\mathbf{d}$ is a random variable supported on $(0, 1]$ interpreted as the time discount factor;

(iii) $(\boldsymbol{\lambda}_t)_{t \in \mathcal{T}}$ is a positive random vector, interpreted as the marginal utility of income, such that for $t = 0, \cdots, |\mathcal{T}| - 1$:

$$\mathbb{E}\left[\boldsymbol{\lambda}_{t+1} \middle| (\boldsymbol{\lambda}_\tau)_{\tau \leq t}, (\boldsymbol{\rho}_\tau^*)_{\tau \leq t}, \mathbf{u}, \mathbf{d}, \right] = \boldsymbol{\lambda}_t \text{ a.s.,}$$

where $\boldsymbol{\rho}_t^* = \mathbf{p}_t^* / \prod_{\tau=1}^t (1 + \mathbf{r}_\tau)$, $t \in \mathcal{T}$;

(iv) $\mathbf{d}^t \nabla \boldsymbol{u}(\mathbf{c}_t^*) \leq \boldsymbol{\lambda}_t \boldsymbol{\rho}_t^*$ a.s. for all $t \in \mathcal{T}$;

(v) For every $j = 1, \ldots, L$ and $t \in \mathcal{T}$, it must be that $\mathbb{P}\left(\mathbf{c}_{t,j}^* \neq 0, \mathbf{d}^t \nabla \boldsymbol{u}(\mathbf{c}_t^*)_j < \boldsymbol{\lambda}_t \boldsymbol{\rho}_{t,j}^*\right) = 0$, where $c_{t,j}^*$, $\rho_{t,j}^*$, and $\nabla u(c_t^*)_j$ denote the j-th components of c_t^*, ρ_t^*, and $\nabla u(c_t^*)$, respectively.

In words, at the beginning of the time-window of interest, a consumer draws a utility function, and a discount factor that are going to remain fixed in time. Every time, given the realized prices, utility, and discount factor the consumer draws a new marginal utility of income, and chooses consumption according to her first-order conditions. The marginal utility of income is a martingale with respect to the known information, which includes the realizations of utility, discount factor, and discounted prices. With this definition at hand we can write the following lemma.

Lemma 4. *For a given random array $(\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t \in \mathcal{T}}$, the following are equivalent:*

(i) *The random array $(\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t \in \mathcal{T}}$ is s/ED-IU-rationalizable.*

(ii) *There exist positive random vectors $(\mathbf{v}_t)_{t \in \mathcal{T}}$ and $(\boldsymbol{\lambda}_t)_{t \in \mathcal{T}}$, and $\mathbf{d}$ supported on or inside $(0, 1]$ such that*

$$\mathbf{v}_t - \mathbf{v}_s \geq \frac{\boldsymbol{\lambda}_t}{\mathbf{d}^t} \boldsymbol{\rho}_t^{*\prime}(\mathbf{c}_t^* - \mathbf{c}_s^*) \quad \text{a.s.,} \quad \forall s, t \in \mathcal{T};$$

and such that for $t = 0, \cdots, |\mathcal{T}| - 1$:

$$\mathbb{E}\left[\boldsymbol{\lambda}_{t+1} \middle| (\boldsymbol{\lambda}_\tau)_{\tau \leq t}, (\boldsymbol{\rho}_\tau^*)_{\tau \leq t}, \{\mathbf{v}_\tau\}_{\tau \leq t}, \mathbf{d}\right] = \boldsymbol{\lambda}_t \text{ a.s..}$$

The proof of Lemma 4 is omitted as it follows trivially from our previous results.

Econometric Framework.– The additional restrictions implied by the income uncertainty in classical ED-rationalizability can be captured as a set of conditional moment conditions that restrict the latent distribution of the marginal utility of income. Note that our framework so far has only dealt with unconditional moments and support constraints. Fortunately, the ELVIS framework can deal with conditional moments too (Theorem 4.1 in Schennach, 2014). The main intuition behind this extension is that a finite set of conditional moments can be written as a (possibly infinite) collection of unconditional moments.

For simplicity of exposition and for practical purposes, instead of the martingale condition from Lemma 4, we will use its simplest implication:

$$\mathbb{E}\left[\boldsymbol{\lambda}_{t+1} - \boldsymbol{\lambda}_t\right] = 0$$

for all $t \in \mathcal{T} \setminus \{T\}$. Moreover, in order to be able to take expectations of marginal utility of income in the cross-section of individuals, we need to impose a normalization condition such that the marginal utility of income of all individuals is in the same units. A natural normalization (without loss of generality) in the form of a support constraint is

$$\boldsymbol{\lambda}_0 = 1 \text{ a.s.}.$$

Recall, that in the benchmark case of perfect-foresight (s/ED-rationalizability), the marginal utility of income is normalized to 1 for every time period (i.e. $\boldsymbol{\lambda}_t = 1$ a.s.). In the case of the static UMT (R-rationalizability) $\boldsymbol{\lambda}_t$ is only restricted to be positive. The s/ED-IU-rationalizability provides a framework that is less restrictive than s/ED-rationalizability but is more restrictive than R-rationalizability.

Empirical Results.– In our first application we rejected the null hypothesis of s/ED-rationalizability with perfect-foresight for the case of couples' households (at the 5 percent significance level). However, we fail to reject the implication of s/ED-IU-rationalizability captured by the above moment conditions for couples' households at the 5 percent significance level with a discount factor set at $\mathbf{d} = 1$ a.s.. We find that $\text{TS}_n = 10.18$ (p-value= 0.179) is below the 95 percent quantile of the χ_7^2 (14.07).

We have not tested all necessary and sufficient conditions for s/ED-IU-rationalizability. But the evidence we provide suggests a possible explanation of the rejection of the perfect-foresight s/ED-rationalizability. In short, it may be that couples' households face more income uncertainty than singles. Hence, not taking income uncertainty into account could be a reason why we reject the dynamic UMT in the couples' households case. Indeed, Browning et al. (2010) points out that risk sharing may be a benefit from marriage. Further exploration of this explanation is beyond the scope of this paper.

E.2. Collective Exponential Discounting Model

The important contribution of Adams et al. (2014) studies a dynamic collective consumer problem to model the behavior of couple's households. The collective model considers a case in which the household maximizes a utilitarian sum of individual utilities of each member of the couple over a vector of consumption of private and (household) public goods, given the individuals' relative power within the household (Pareto weights). Each individual member of the household is an exponential discounter but the observed consumption is a result of the collective decision making process, and may not be time-consistent. We formulate a test for the collective model using our methodology. We fail to reject the null hypothesis of consistency of the data set with the dynamic collective model assuming that the random discount factor is supported on $[0.975, 1]$ (this support is the one used in Adams et al. (2014)).

Consider a household that consists of two individuals labeled by A and B. Partition the vector of goods into publicly consumed goods indexed by H and privately consumed goods indexed by I. That is, $c_t = (c'_{t,I}, c'_{t,H})'$ and $p_t = (p'_{t,I}, p'_{t,H})'$. Let $c_{t,A}$ and $c_{t,B}$ be the consumption of the privately consumed goods of individuals A and B, respectively ($c_{t,I} = c_{t,A} + c_{t,B}$). Then the collective household problem with exponential discounting corresponds to the maximization of

$$V_\tau(c) = \omega_A u_A(c_{\tau,A}, c_{\tau,H}) + \omega_B u_B(c_{\tau,B}, c_{\tau,H}) + \sum_{j=1}^{T-\tau} [d_A^j \omega_A u_A(c_{\tau+j,A}, c_{\tau+j,H}) + d_B^j \omega_B u_B(c_{\tau+j,B}, c_{\tau+j,H})],$$

subject to this linear intratemporal budget constraint:

$$p'_{\tau,I} c_{\tau,I} + p'_{\tau,H} c_{\tau,H} + s_t - y_t - (1 + r_t)s_{t-1} = 0,$$

where $\omega_A, \omega_B > 0$ are Pareto weights that remain constant across time and represent the bargaining power of each household member. Individual utility functions, u_A and u_B, are assumed to be continuous, locally nonsatiated and concave. The individual discount factors are similarly denoted by d_A and d_B. The rest of the elements are the same as in our main model.

The quantities $c_{t,A}, c_{t,B}$ are assumed to be unobservable to the econometrician. We observe only c_t. Adams et al. (2014) propose one solution to the collective household problem above. They assume full efficiency in the sense that there are personalized Lindahl prices for the publicly consumed goods $p_{t,H}$ that perfectly decentralize the above problem. The Lindahl prices are $p_{t,A} \in \mathbb{R}_{++}^{L_H}$ for household member A and the analogous $p_{t,B}$ such that $p_{t,A} + p_{t,B} = p_{t,H}$. Adams et al. (2014) established the result which is the analog of Theorem 1. Similar to the case of the single-individual household, define $\rho_{t,h} = p_{t,h}/\prod_{j=1}^t (1 + r_j)$ for $h \in \{I, H, A, B\}$.

Theorem 6 (Adams et al., 2014). *An array $(\rho_t, c_t)_{t \in \mathcal{T}}$ can be generated by a collective household exponential discounting model with full efficiency if and only if there exist $d_A, d_B \in (0, 1]$; strictly positive vectors $(v_{t,A})_{t \in \mathcal{T}}$, $(v_{t,B})_{t \in \mathcal{T}}$; individual private consumption quantities $(c_{t,A}, c_{t,B})_{t \in \mathcal{T}}$ (with $c_{t,A} + c_{t,B} = c_{t,I}$); and personalized Lindahl prices $(p_{t,A}, p_{t,B})_{t \in \mathcal{T}}$ (with $p_{t,A} + p_{t,B} = p_{t,H}$) such that*

for all $s, t \in \mathcal{T}$:

$$v_{t,A} - v_{s,A} \geq d_A^{-t} \left[\rho'_{t,I}(c_{t,A} - c_{s,A}) + \rho'_{t,A}(c_{t,H} - c_{s,H}) \right],$$
$$v_{t,B} - v_{s,B} \geq d_B^{-t} \left[\rho'_{t,I}(c_{t,B} - c_{s,B}) + \rho'_{t,B}(c_{t,H} - c_{s,H}) \right].$$

With this result in hand, we can establish our finding in a very straightforward manner. We let $\boldsymbol{\rho}_t$ and $\mathbf{c}_t^*$ be the random vectors of deflated prices and true consumption. Finally, we define $\mathbf{d}_A$ and $\mathbf{d}_B$ as the random discount factors for household members A and B, respectively. Also, $\mathbf{u}_A, \mathbf{u}_B$ and $\boldsymbol{\omega}_A, \boldsymbol{\omega}_B$ denote the random utility functions and random Pareto weights for each household member. We keep here the assumption about the data-generating process that we maintained for the case of s/ED-rationalizability, namely, we assume that the preferences and Pareto weights remain stable for each household after being drawn from the joint distribution of $(\mathbf{u}_A, \mathbf{d}_A, \boldsymbol{\omega}_A)$ and $(\mathbf{u}_B, \mathbf{d}_B, \boldsymbol{\omega}_B)$ at the first time period. Now we can establish and prove a stochastic analogue to the result in Adams et al. (2014).

Theorem 7. *If a random array $(\boldsymbol{\rho}_t, \mathbf{c}_t^*)_{t \in \mathcal{T}}$ is generated by a collective household with random exponential discounting under full efficiency, then there exist random variables $\mathbf{d}_A, \mathbf{d}_B$ which are both supported on or inside $[\theta_0, 1]$, and strictly positive random vectors $(\mathbf{v}_{t,A})_{t \in \mathcal{T}}$, $(\mathbf{v}_{t,B})_{t \in \mathcal{T}}$ that satisfy*

$$\mathbf{d}_A^t(\mathbf{v}_{t,A} - \mathbf{v}_{s,A}) + \mathbf{d}_B^t(\mathbf{v}_{t,B} - \mathbf{v}_{s,B}) \geq \boldsymbol{\rho}'_t(\mathbf{c}_t^* - \mathbf{c}_s^*) \text{ a.s.} \quad \forall t, s \in \mathcal{T}.$$

Theorem 7 does not provide sufficient conditions for collective rationalizability. We can provide a stochastic analogue of Theorem 6, but our choice has several advantages: (i) one does not need to specify which goods are consumed privately or publicly; (ii) the inequality restrictions in Theorem 7 do not depend on the unobservable Lindahl prices and private consumption vectors, which simplifies implementation; and (iii) we can maintain Assumption 1 in a very natural form. We also assume that prices are measured precisely. Assuming that $\mathbf{d}_A$ and $\mathbf{d}_B$ are supported on or inside $[0.975, 1]$ we find that $\text{TS}_n = 0.597$ (p-value $= 0.963$), which is below the 95 percent quantile of the χ_4^2 (9.5). Thus, we fail to reject the null hypothesis that the couples' household data set is consistent with the collective exponential discounting model under the assumptions of full efficiency, common support for preferences, and the collective mean budget constraint. The test statistic value for the explored $\theta_0 = 0.975$ for the collective model is below that of the exponential discounting model for the sample of couples' households.

F. Empirical Application (I) Extended: Average Varian Support Set for Budget Shares

Here we compute bounds on counterfactual average budget shares. Since the null hypothesis of s/ED-rationalizability cannot be rejected for the case of single-individual households in our
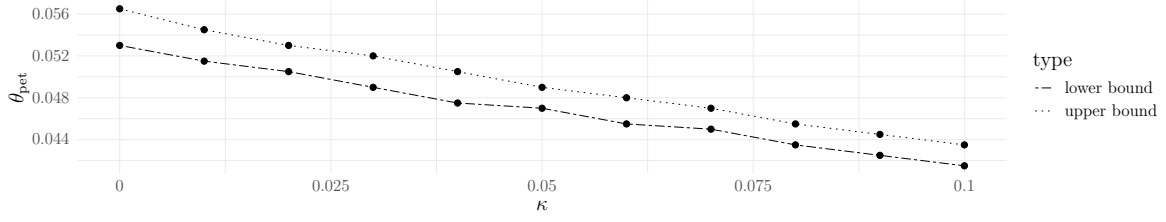
Figure 1 – Average Support Set for Budget Shares: Petrol

first application, we can compute out-of-sample forecasts of average budget shares when the price changes. We focus on one of the categories, petrol. The motivation for this exercise is twofold. First, we want to showcase that our methodology can deliver informative bounds of quantities of interest. Second, fossil fuels prices are usually highly variable. This means that understanding the effects of such price changes on the problem of budget allocation in households is empirically relevant. This question may also be important for some potential users of this methodology, like a regulator trying to impose a green-tax. We study the following counterfactual question: What would be the average budget share of petrol in time $T+1$, if the price of petrol $p_{T+1,\text{pet}}$ is $\kappa \cdot 100$ percent higher than $p_{T,\text{pet}}$ (i.e. $p_{T+1,\text{pet}} = (1+\kappa)p_{T,\text{pet}}$)?

For simplicity, we consider κ that takes values in $\{0, 0.01, \cdots, 0.10\}$ (i.e., at most a 10 percent price increase). We set the (random) interest rate faced by the single-individual households $\mathbf{r}_{T+1} = 0.6$ a.s. (roughly 1 percent increase over the average interest rate) and the support of the random discount factor to $[0.975, 1]$.[66] The counterfactual moment is

$$g_C((\boldsymbol{\rho}_t^*, \mathbf{c}_t^*)_{t \in \mathcal{T}}, \mathbf{c}_{T+1}^*; \boldsymbol{\rho}_{T+1}^*, \theta_j) = \frac{\boldsymbol{\rho}_{T+1,\text{pet}}^* \mathbf{c}_{T+1,\text{pet}}^*}{\boldsymbol{\rho}_{T+1}^{*\prime} \mathbf{c}_{T+1}^*} - \theta_{\text{pet}},$$

where $\theta_{\text{pet}} \in [0, 1]$ is the average budget share of petrol. Note that interest rate cancels out such that the budget share depends only on spot prices. For ED-rationalizability we do not need to specify the expenditure level at $T+1$, as the model endogenously predicts an expenditure level for a new price. This is because ED-rationalizability generalizes quasilinear rationalizabilty by adding discounting. However, adding discounting does not affect the model's budget-free nature (Gauthier, 2018).

The 95 percent bounds for the average counterfactual petrol share are depicted in Figure 1. Note that conditions of the Proposition 1 are satisfied. Hence, the sets of interest are connected, which means it is enough to depict the minimal and maximal shares that are not rejected by our test.[67] As expected, demand for petrol is decreasing in the price of petrol. The maximal and the minimal drops in shares associated with 10 percent price increase are 1.5 to 0.95 percentage points, respectively.[68]

---

[66]We computed our counterfactual sets with the support of the random discount factor equal to $[0.1, 1]$ and $\mathbf{d} = 1$ a.s.. The results are similar to those with $[0.975, 1]$ and available upon request.

[67]We searched for average budget share θ_{pet} in the grid $\{0.00, 0.005, \cdots, 1\}$.

[68]The empirical budget share at $t = T$ computed from our (mismeasured) data set for petrol is 6 percent.