Stochastic Revealed Preferences with Measurement Error: Testing for Exponential Discounting in Survey Data*

Victor H. Aguiar
vaguiar@uwo.ca

Nail Kashaev†
nkashaev@uwo.ca

This version: May 2017

Abstract Measurement error causes a loss of distributional information, preventing the researcher from applying the deterministic revealed-preference tools at the individual level. This paper proposes a new statistical revealed-preference framework that is applicable to such cases. We use our framework to establish a nonparametric test of the standard exponential-discounting, time-separable consumer model in environments with measurement error in consumption. The proposed testing procedure is asymptotically consistent, though it may be conservative. We provide Monte Carlo evidence showing that measurement error in the standard deterministic revealed-preference tests may lead to mistakenly judging an otherwise time-consistent consumer to be time-inconsistent. In contrast, our methodology does not reject the exponential-discounting model in such cases. We find support for exponential-discounting behavior in a consumption panel survey for single-individual households, while rejecting the model for the case of couples' households.

JEL classification numbers: C60, D10.
Keywords: exponential discounting, time consistency, time discount factor, revealed preferences, latent variable integration via simulation, measurement error.

*We are grateful to Elizabeth Caucutt, Lance Lochner, Jim MacGee, Salvador Navarro, Joris Pinkse and David Rivers for useful comments and encouragement.

†Department of Economics, University of Western Ontario.

1. Introduction

This paper proposes a new statistical revealed-preference framework in order to account for possible measurement error. We use our framework to formulate a nonparametric test of the classic dynamic consumer choice problem with exponential discounting and time-separable tastes, when measurement error in consumption is present. We achieve this goal by imposing a weak centering assumption on the unobserved consumption measurement error. We require that in each time period, on average, consumers recall and report accurately their total expenditures.

Measurement error in reported consumption is a well-known feature of survey data and is therefore an important issue to address when conducting field studies using the revealed-preference framework.[1] Moreover, one common and important criticism of revealed-preference studies (at least from a standard statistical approach) is the argument that they have tended to disregard measurement error in their analyses. Here we provide a solution for this key issue, and advance the revealed-preference methodology.

The leading solution to deal with measurement error in the revealed-preference framework usually consists of perturbing (minimally) any observed individual consumption streams in order to satisfy the conditions of a revealed-preference test (Adams et al. (2014)). However, this approach does not allow for standard statistical hypothesis testing. In particular, one cannot control the probability of making an error in rejecting a particular model when in fact this rejection is an artifact of noisy measurements. The closest work to our analysis is the methodology developed by Varian (1985). Varian's approach allows one to conduct standard statistical hypothesis testing under the strong assumptions of normality and additivity of consumption measurement error. In addition, Varian's methodology requires that the variance of the measurement error be known. In contrast, our methodology is fully nonparametric.

We provide Monte Carlo evidence suggesting that a consumption stream generated by exponential-discounting behavior can fail the standard deterministic revealed-preference test with high probability (close to 1) due to perturbations caused by a very moderate measurement error. In stark contrast, our methodology is able to conclude correctly, under the usual confidence levels, that the data set is consistent with the null hypothesis of interest. We highlight the possibility of the deterministic revealed-preference tests overrejecting the exponential-discounting model due to the presence of measurement error.

Our *methodological contribution* is to provide a new nonparametric statistical notion of rationalizability of a random vector of prices and consumption streams when there is measurement error in consumption. We do this on the basis of the work of Schennach (2014) on entropic latent variable integration via simulation (ELVIS). We use our main result to formulate a statistical test for the null hypothesis of exponential discounting, generalizing the work of Browning (1989) for the case with measurement error.

We wish to emphasize that our statistical methodology can be applied to any revealed-preference

[1] See Varian (1985), Tsur (1989), Gross (1995), Blundell et al. (2003), Beatty and Crawford (2011), Adams et al. (2014), Boccardi (2016).

model. However, we focus on the intertemporal consumption model with exponential discounting due to its importance for applied work. 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 that the exponential-discounting consumer is time-consistent. 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. Nevertheless, this model is under increasing scrutiny in the light of experimental evidence that tends to find that the behavior of experimental subjects is time-inconsistent. Unfortunately, there is less evidence from the field, and in particular, from survey data. Several authors such as Browning (1989), DellaVigna and Malmendier (2006), and Blow et al. (2013), have provided suggestive evidence against the validity of the exponential-discounting model in the field. Our methodology addresses, in a nonparametric fashion, the presence of measurement error in survey-gathered data, in order to examine the robustness of these findings.

Our *empirical contribution* is to apply our methodology to a consumer panel data set gathered from single-individual and couples households in Spain.[2] In contrast to Blow et al. (2013), we find support for exponential-discounting behavior for single-individual households. At the same time, we reject the null hypothesis of exponential discounting for the case of couples. Hence, we provide evidence against the existence of a household consumer with exponential-discounting behavior while accounting for measurement error. We also examine a generalized version of the exponential-discounting model in the form of the collective household consumer unit proposed by Adams et al. (2014). In this more general model, each individual within the couple is an exponential discounter, and the individuals jointly decide how much to consume privately and publicly by maximizing a weighted sum of individual discounted utilities. Using a deterministic test, Adams et al. (2014) conclude that this model is accepted. Our testing procedure leads to the opposite conclusion, we reject the collective model with exponential-discounting behavior for the couples case.

Outline

The paper proceeds as follows. Section 2 presents a brief literature review about previous studies to test the exponential-discounting model in survey data. Section 3 presents the exponential-discounting model and the deterministic revealed-preference test that we use as a benchmark. Section 4 contains the main contribution: we present our new statistical nonparametric test for random exponential discounting in the presence of measurement error. Here, we introduce a centering condition on the additive measurement error on consumption streams. We formulate a new statistical notion of rationalizability on the basis of the ELVIS methodology (Schennach (2014)). Section 5 establishes a testing and inference framework for exponential-discounting behavior. Section 6 implements our empirical test of exponential discounting in a well-known consumer panel-survey

[2]This data set was used in Beatty and Crawford (2011), Blow et al. (2013), Adams et al. (2014).

data set used in Adams et al. (2014). Here we report the results for the testing procedure for the single-individual households, and for the couples' households. Section 7 performs two Monte Carlo experiments to assess the performance of our testing procedure in a finite sample. Section 8 presents several important extensions for the proposed methodology. The first one corresponds to the collective household model presented in Adams et al. (2014); the second deals with the case of uncertainty about future income level; and the third corresponds to the sophisticated quasi-hyperbolic model. Finally, we conclude in Section 9. All proofs are collected in the appendix.

2. Literature Review

Several studies that use laboratory setups have provided evidence against exponential discounting. Andreoni and Sprenger (2012) reject the null hypothesis of exponential discounting in favor of hyperbolic discounting (present bias). Montiel Olea and Strzalecki (2014) set-identify the distribution of hyperbolic discounting and find evidence against exponential discounting in an experimental pilot study. More recently, Echenique et al. (2014) develop a revealed-preference methodology (for one good) with several consumption streams (nonparametric and deterministic); they use Andreoni and Sprenger (2012) data set to conclude that the majority of subjects are not consistent with exponential-discounting. Nonetheless, it is important to explore to what extent this finding has external validity. To address this issue, researchers have turned to survey data in the form of household consumption panels. Most of the existing work from the field has found evidence against exponential discounting. 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 hypothesis testing. (Additional problems with the existing evidence are strong parametric assumption on preferences and 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 revealed preference by Browning (1989). In particular, it avoids making parametric assumptions about the functional form of instantaneous utility. However, this work does not take into consideration the fact that consumption quantities can be mismeasured. Blow et al. (2013) apply Browning (1989) to a survey data (similar to what we use) to find that less than 1 percent of the households pass the revealed-preference test for exponential discounting. However, the low success rate of the deterministic test for exponential discounting may be due to measurement error. In the present work, we tackle this issue by extending their revealed-preference results to account for measurement error in consumption. From the revealed preferences perspective, our approach is a generalization of Browning (1989) and Blow et al. (2013). It differs from the work in Echenique et al. (2014). The latter investigates the exponential-discounting model when the researcher observes several consumption streams for the same consumer with one good or commodity. In this different environment, Echenique et al. (2014) develop in depth the revealed-preference theory of the exponential-discounting model as well as the quasi-hyperbolic discounting model and other time-separable models of consumer theory. However, this approach is deterministic in nature, and does not allow for multiple commodities. For that reason, it is better suited for experimental data settings. The work of Quah et al. (2015) provides a test of exponential time discounting for a case without measurement error and without time-separable preferences.

The exponential-discounting model, as well as a variant of this model that allows for collective household decision making, has also been tested in Adams et al. (2014). Adams et al. (2014) tackle the issue of measurement error. They find the additive perturbation with minimal norm that renders the individual consumption streams compatible with the revealed-preference 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 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 (given the subjectivity of the threshold). In addition, it is unclear what the probability of rejecting the null hypothesis is when the null hypothesis is in fact true. This is undesirable for a statistical test. In contrast, our approach allows one to perform traditional statistical hypothesis testing without compromising the generality of the revealed-preference conditions under a weak centering assumption on the measurement error. This is important because in our methodology the critical value used to conduct the test arises from the asymptotic behavior of our test statistic. Moreover, we establish empirically that, when measurement error is present, we must reject the collective household exponential-discounting model with intra-household preference heterogeneity that is proposed by Adams et al. (2014). This finding stands in contrast to Adams et al. (2014). They suggest that the collective model rationalizes the data set we use in our empirical study by applying their deterministic methodology.

The work of DellaVigna and Malmendier (2006) studies preferences over contracts, and time-discounting behavior in the field. This study uses a parametric utility specification and the assumption that time discount factors are common across individuals. They allow for measurement error with parametric restrictions. Under these assumptions they find evidence against the exponential-discounting model. More importantly, the work by DellaVigna and Malmendier (2006) differs from ours in the following respect. Their work is a quasi-experimental study that uses external sources of variation to differentiate between the exponential and hyperbolic cases. In contrast, our work deals with survey data sets in which there is no other information beyond interest rates, prices, and consumption streams. In this environment, we establish that we can differentiate exponential-discounting behavior from sophisticated hyperbolic-discounting behavior. However, we do not reject the null hypothesis of exponential discounting for the case of single households. We note that consumers may be hyperbolic discounters in one-shot decisions (e.g., signing a gym contract), but may be time-consistent in making frequent decisions (e.g., buying food, recreation, or transportation), as in our application.

Our work 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.[3] The lack of consensus can be attributed in some degree to the lack of identification of 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. In this situation, only the support of the distribution can be set-identified. However, the methodology presented in this work allows us to test for exponential-discounting behavior even in this setting (i.e., without identifying the discount factor distribution).[4]

3. The Exponential-Discounting Consumer: Imposing Shape Constraints through the Revealed-Preference Approach

The main objective of this section is to provide a brief summary of the deterministic exponential-discounting consumer model and its revealed-preference characterization. We start from a deterministic baseline, following Browning (1989) and Rockafellar (1970). All quantities used here are assumed to be measured precisely.

3.1. The Deterministic Benchmark

We assume that an individual consumer has preferences over a stream of dated consumption bundles $(c_t)_{t\in\mathcal{T}}$, where $\mathcal{T} = \{\tau, \cdots, T\}$, $\tau, T \in \mathbb{N}_+$, $\tau \leq T$, and $c_t \in \mathbb{R}^L_+$. (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

$$p_t' c_t - y_t + s_t - a_t = 0,$$

where $d \in (0, 1]$ is the discount factor, $p_t \in \mathbb{R}^L_{++}$ is the price vector, $y_t \in \mathbb{R}_{++}$ 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.

[3] We refer the reader to the survey by Frederick et al. (2002), for its extensive references, while focusing here only on the immediate antecedents to our work.

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

Moreover, the assets evolve according to the following law of motion:

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

where $r_{t+1} \in \mathbb{R}$ is the interest rate that is accessible for the consumer. The initial ($t = 0$) holdings of assets 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 elements of interest, or 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. If we identify the pair (d, u), then we identify the value function V_t for all t. The exponential-discounting consumer is time-consistent. That is, she will solve the same problem at any point of the time window.

Under additional assumptions such as the differentiability of the instantaneous utility function, and the strict positivity of the consumption vector, the well-known solution to this problem can be characterized by the Euler equation:

$$\frac{\partial u(c_t)}{\partial c_{t,l}} = d(1 + r_{t+1})\frac{p_{t,l}}{p_{t+1,l}}\frac{\partial u(c_{t+1})}{\partial c_{t+1,l}}.$$

In general, the optimization problem produces the following condition:

$$\nabla u(c_t) \leq \lambda \frac{1}{d^t}\rho_t,$$

where $\lambda \in \mathbb{R}_{++}$ is a Lagrange multiplier, $\rho_t = \frac{p_t}{\prod_{i=1}^{t}(1+r_i)}$ is a discounted price, and $\nabla u(c_t)$ is a subgradient of u at the point c_t. (Under differentiability, $\nabla u(c_t)$ is a gradient.)

Lemma 1. *(Browning (1989)) A deterministic array $(\rho_t, c_t)_{t\in\mathcal{T}}$ can be generated by an exponential-discounting consumer if and only if there exist $\lambda > 0$, $d \in (0, 1]$, and concave, locally non-satiated and continuous $u : \mathbb{R}_+^L \to \mathbb{R}$ such that $\nabla u(c_t) \leq \lambda \frac{1}{d^t}\rho_t$ for all $t \in \mathcal{T}$.*

3.2. The Elimination of a Latent Infinite-Dimensional Parameter

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.[5] Since our objective is to test, not to estimate, the exponential-discounting model, we follow a different path. The theorists of revealed preference have devised ingenious ways to eliminate the latent infinite-dimensional parameters (e.g., the utility functions) by exploiting their shape restrictions. We follow Browning (1989) and Rockafellar (1970) to formulate a result that eliminates u for the

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

case of exponential discounting.

The cost of doing this is that the Euler equation has to be replaced by a set of inequalities. However, these inequalities do not depend on the infinite-dimensional parameter u anymore. Moreover, they require only the concavity of u. As a result, the inequalities are exact and do not involve any form of approximation, which is an improvement upon the nonparametric methods (e.g., sieves, kernel estimators) or the parametric approach used in many applied papers.

First, we recall the definition of the concavity of u.

Definition 1. (Concavity) A utility function u is said to be concave if and only if $u(c_s) - u(c_t) \leq \nabla u(c_t)'(c_s - c_t)$, for all $s, t \in \mathcal{T}$.

The nonparametric test for exponential-discounting consumption without measurement error is captured by the following result.

Theorem 1. *(Browning (1989)) Given $(\rho_t, c_t)_{t\in\mathcal{T}}$, the following are equivalent:*

(i) The deterministic array $(\rho_t, c_t)_{t\in\mathcal{T}}$ can be generated by the exponential-discounting model.

(ii) There exist a constant $d \in (0, 1]$ and a positive vector $(v_t)_{t\in\mathcal{T}}$ such that:

$$v_t - v_s \geq d^{-t}\rho_t'(c_t - c_s) \quad \forall t, s \in \mathcal{T}.$$

Observe that the formulation in Browning (1989) has transformed a unique equality that depends on d and the infinite-dimensional u (i.e., the Euler equation) to a set of inequality conditions that depend only on d and on a constant finite-dimensional vector $(v_t)_{t\in\mathcal{T}}$. Nonetheless, this set of conditions is satisfied if and only if we can find a pair (d, u) that satisfies the conditions in Lemma 1. Given d, this is a parametric linear programming problem that will tell us whether a consumption stream is consistent with exponential discounting (i.e., equivalently testing the cyclical monotonicity condition in Rockafellar (1970)).

This methodology is traditionally applied at the individual level in survey data, assuming that the data contains no measurement error. The results are usually disappointing with high rates of rejections for the exponential discounting modelwith high rates of rejections for the exponential discounting model (see Blow et al. (2013), and the more optimistic Adams et al. (2014)). The popular exponential-discounting model is found to be inconsistent with the data. We argue that this may be an overly pessimistic conclusion, and possibly an artifact of measurement error. For that reason, we will extend the revealed-preference framework to a noisy or stochastic environment in the next section.

4. The Exponential-Discounting Model with Heterogeneity in Tastes and Measurement Error

In this section, we introduce the Exponential Discounting Random Utility model (henceforth, EDRUM) with mismeasured consumption. We provide a new statistical notion of rationalizability (i.e., consistency of the observed data with the EDRUM), and provide a result similar to Theorem 1 in the presence of measurement error.

4.1. Exponential Discounting with Heterogeneous Discount Factors and Heterogeneous Utility

We are interested in testing a model of dynamic random utility with exponential discounting such that each individual is an independent identically distributed (i.i.d.) draw from some stochastic consumption rule. By using Lemma 1, we can directly define the EDRUM as follows. From here on, we use bold face font to denote random objects and regular font for deterministic ones. Let $\boldsymbol{\rho}_t$ and $\mathbf{c}_t^*$ denote random vectors of discounted prices and true consumption at time t, respectively (supported on $P_t \subseteq \mathbb{R}^L_{++}$, and $C_t^* \subseteq \mathbb{R}^L_+$).

Definition 2. (EDRUM) A random array $(\boldsymbol{\rho}_t, \mathbf{c}_t^*)_{t\in\mathcal{T}}$ is consistent with the EDRUM if there exists a triple $(\mathbf{u}, \mathbf{d}_\delta, \boldsymbol{\lambda})$, where *(i)* $\mathbf{u}$ is a random, concave, locally nonsatiated, and continuous utility function; *(ii)* $\mathbf{d}_\delta \in D_\delta \subseteq (0, 1]$ is a random discount factor; *(iii)* $\boldsymbol{\lambda} \in \mathbb{R}_{++}$ is a random variable interpreted as a Lagrange multiplier capturing consumption patterns; and such that $\nabla \boldsymbol{u}(\mathbf{c}_t^*) \leq \boldsymbol{\lambda}\mathbf{d}_\delta^{-t}\boldsymbol{\rho}_t$ with probability 1.

This definition means that for a given realization of *(i)* the utility function, *(ii)* the discount factor, and *(iii)* the Lagrange multipliers, the realized discounted prices and the realized true consumption should fulfill the inequality $\nabla u(c_t^*) \leq \lambda d^{-t}\rho_t$. This is a special case of the random utility model in which the preferences $(\mathbf{u}, \mathbf{d}_\delta)$ and the distribution of wealth captured by $\boldsymbol{\lambda}$ are drawn at some initial time for each consumer, and then are kept fixed over time. In practice, we can normalize $\boldsymbol{\lambda}$ to a constant equal to 1 since it is time-invariant.

Remark 1. The definition of the EDRUM allows for nondifferentiable utility functions. So, the subgradient $\boldsymbol{u}(\mathbf{c}_t^*)$ may be set-valued. In this case one should read the condition $\nabla u(c_t^*) \leq \lambda d^{-t}\rho_t$ as "there exists $\xi \in \nabla u(c_t^*)$ such that $\xi \leq \lambda d^{-t}\rho_t$."

Given the definition of the EDRUM, 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 consistent with the EDRUM.

(ii) There exist a random variable $\mathbf{d}_\delta \in D_\delta$ and a random vector $\mathbf{v} = (\mathbf{v}_t)_{t\in\mathcal{T}} \in \mathbb{R}^{|\mathcal{T}|}_+$ such that

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

Lemma 2 allows us to statistically test the rationalizability of $(\boldsymbol{\rho}_t, \mathbf{c}_t^*)_{t\in\mathcal{T}}$. However, any test based on this notion of rationalizability cannot differentiate between quasi-consistent cases and exact exponential discounting (an issue first identified by Galichon and Henry (2013)). The reason is that there can be time-inconsistent random choice rules that are arbitrarily close to the EDRUM. The issue arises because the set of the exponential-discounting behaviors is not closed. That is why we need to extend the notion of consistency of the data set with the EDRUM.

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 exponential-discounting behavior.

Definition 3. (Approximate EDRUM) We say that $(\boldsymbol{\rho}_t, \mathbf{c}_t^*)_{t\in\mathcal{T}}$ is **approximately consistent** with the EDRUM if there exists a sequence of random variables $(\mathbf{d}_\delta^m, (\mathbf{v}^m)')' \in D_\delta \times \mathbb{R}_+^{|\mathcal{T}|}$, $m = 1, 2, \ldots$, such that

$$\mathbb{P}\left(\mathbb{1}\left(\mathbf{v}_t^m - \mathbf{v}_s^m \geq (\mathbf{d}_\delta^m)^{-t}\boldsymbol{\rho}_t'[\mathbf{c}_t^* - \mathbf{c}_s^*]\right)\right) \to_{m\to+\infty} 1,$$

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

4.2. Introducing Measurement Error

Theorem 1 and Lemma 2 allow us to get testable implications of the EDRUM. These implications depend solely on the distribution of $\mathbf{d}_\delta$ and $\mathbf{v} = (\mathbf{v}_t)_{t\in\mathcal{T}}$. The usual approach to testing the EDRUM would amount to solving a linear programming problem corresponding to Theorem 1 at the level of the individual consumers, for a fixed discount factor. However, this common practice does not work any more in the presence of **measurement error**. When **true consumption** is measured erroneously, we observe not $\mathbf{c}_t^*$ but rather a perturbed version of it.

Remark 2. Measurement error in consumption in surveys arises because consumers are asked to report past consumption patterns in a periodic nature. Consumers may fail to correctly report the consumption level due to bounded recall, social desirability, and mistakes in communication. Thus, self-reported consumption is likely mismeasured.

Define the **measurement error** $\mathbf{w} = (\mathbf{w}_t)_{t\in\mathcal{T}} \in W$, where $\mathbf{w}_t \in W \subseteq \mathbb{R}^L$, $t \in \mathcal{T}$, as the difference between reported consumption, $\mathbf{c} = (\mathbf{c}_t)_{t\in\mathcal{T}}$, and true consumption, $(\mathbf{c}_t^*)_{t\in\mathcal{T}}$. That is,

$$\mathbf{w}_t = \mathbf{c}_t - \mathbf{c}_t^*,\ t \in \mathcal{T}. \tag{1}$$

It is important to note that we *define* the measurement error. We do not make any assumptions about how the difference between $\mathbf{c}$ and $\mathbf{c}^*$ arises (i.e., we allow for measurement error in consumption to be multiplicative or additive).[6]

By Lemma 2 we can immediately conclude that the observed $\mathbf{x} = (\boldsymbol{\rho}_t, \mathbf{c}_t)_{t\in\mathcal{T}}$ can be generated by the EDRUM if and only if there exist $(\boldsymbol{d}_\delta, \mathbf{v}, \mathbf{w})$ such that

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

However, we know that without restrictions on the distribution of the measurement error, revealed-preference tests have no power. That is, there always exists a $\mathbf{w}$ such that the observed $\mathbf{x}$ is consistent with the exponential-discounting model. We therefore impose reasonably weak centering restrictions on the measurement error, namely, the mean budget neutrality. In this case, we show that the restrictions on measurement error interact with the revealed-preference restrictions in order to provide a test for the EDRUM with asymptotic power equal to 1.

Assumption 1. *(Mean Budget Neutrality)* $\mathbb{E}\left[\mathbf{d}_\delta^{-t}\boldsymbol{\rho}_t'\mathbf{w}_t\right] = 0$, *for all* $t \in \mathcal{T}$.

Assumption 1 implies that the measurement error does not alter the mean discounted value of total expenditure, $\mathbb{E}\left[\mathbf{d}_\delta^{-t}\boldsymbol{\rho}_t'\mathbf{c}_t\right] = \mathbb{E}\left[\mathbf{d}_\delta^{-t}\boldsymbol{\rho}_t'\mathbf{c}_t^*\right]$. In other words, mean budget neutrality captures the idea that consumers, on average, may remember the total expenditure level better than the actual details.

Our assumption about measurement error does not imply the classical measurement error assumptions that may fail in the consumer environment. In fact, it is compatible with nonclassical measurement error, such as $\mathbb{E}\left[\mathbf{w}_t|\mathbf{c}_t^*\right] \leq 0$ (Carroll et al. (2014)). Assumption 1 is implied by the following conditions: *(i)* $\mathbb{E}[\,\boldsymbol{\rho}_t'\mathbf{w}_t|\mathbf{c}_t^*\,] = 0$, and *(ii)* $\mathbf{d}_\delta \perp\!\!\!\perp \mathbf{w}_t|\mathbf{c}_t^*$, where condition *(i)* means that the consumers correctly recall, on average, their total expenditure on all goods conditional on true consumption, and condition *(ii)* means that measurement error is independent of the random discount factor conditional on true consumption.

Example 2. (Additive Measurement Error) Consider a case where $\mathbf{c}_{t,l} = \mathbf{c}_{t,l}^* + \boldsymbol{\epsilon}_{t,l}$ for $l = 1, \ldots, L$; and where $\boldsymbol{\epsilon}_{t,l} \sim TN_{[-a,a]}(0, \sigma)$ 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.. Assume, moreover, that $\boldsymbol{\epsilon}_{t,L} = -\frac{1}{\rho_{t,L}}\sum_{l=1}^{L-1}\boldsymbol{\rho}_{t,l}\boldsymbol{\epsilon}_{t,l}$ a.s.. In words, the consumers report correctly the total expenditure with probability 1 while making some mistakes about the budget shares for each good. Measurement error is defined as $\boldsymbol{w}_{t,l} = \boldsymbol{\epsilon}_{t,l}$. Note that $\mathbf{w}_{t,l}$ is independent from $\mathbf{d}_\delta$ conditional on $\mathbf{c}_t^*$, and by construction $\sum_{l=1}^{L}\boldsymbol{\rho}_{t,l}\boldsymbol{\epsilon}_{t,l} = 0$ a.s.. This means that Assumption 1 holds.

Example 3. (Price and Expenditures with Measurement Error) Our definition of measurement error is general enough to cover special cases of measurement error in prices and expenditures. Say that both reported good-level expenditures $\boldsymbol{\rho}_{t,l}\mathbf{c}_{t,l}$, and discounted prices $\boldsymbol{\rho}_{t,l}$ are measured with error. In particular, the error enters multiplicatively, such that $\boldsymbol{\rho}_{t,l}\mathbf{c}_{t,l} = \boldsymbol{\rho}_{t,l}^*\mathbf{c}_{t,l}^*\boldsymbol{\epsilon}_{1,t,l}$, and $\boldsymbol{\rho}_{t,l} = \boldsymbol{\rho}_{t,l}^*\boldsymbol{\epsilon}_{2,l}$ (with $\boldsymbol{\epsilon_i}$ for $i = 1, 2$ denoting a positive random vectors). In practice, observed consumption is computed as the ratio of these mismeasured expenditures and prices: $\mathbf{c}_{t,l} = \boldsymbol{\rho}_{t.l}^*\mathbf{c}_{t,l}^*\boldsymbol{\epsilon}_{1,t,l}/\boldsymbol{\rho}_{t,l}^*\boldsymbol{\epsilon}_{2,l}$. The result is that

[6] Formally this makes the support W depend on the support of the observed and true consumption. For simplicity we omit this dependence from the notation.

observed consumption is also mismeasured and is of the form $\mathbf{c}_{t,l} = \mathbf{c}^*_{t,l}\boldsymbol{\epsilon}_{3,t,l}$, where $\boldsymbol{\epsilon}_{3,t,l} = \boldsymbol{\epsilon}_{1,t,l}/\boldsymbol{\epsilon}_{2,l}$. Here we assume that the price measurement error has to be time-invariant. Note that by defining measurement error as $\hat{\mathbf{w}}_{t,l} = \mathbf{c}^*_{t,l}\boldsymbol{\epsilon}_{3,t,l} - \mathbf{c}^*_{t,l}/\boldsymbol{\epsilon}_{2,l}$ instead of $\mathbf{w}_{t,l} = \mathbf{c}^*_{t,l}\boldsymbol{\epsilon}_{3,t,l} - \mathbf{c}^*_{t,l}$, we can rationalize the data set $(\boldsymbol{\rho}_t, \mathbf{c}_t)_{t\in\mathcal{T}}$, because the equality $\boldsymbol{\rho}_t'(\mathbf{c}_t - \mathbf{c}_s + \hat{\mathbf{w}}_s - \hat{\mathbf{w}}_t) = \sum_{l=1}^{L} \boldsymbol{\rho}^*_{t,l}\boldsymbol{\epsilon}_{2,l}(\mathbf{c}^*_{t,l}/\boldsymbol{\epsilon}_{2,l} - \mathbf{c}^*_{s,l}/\boldsymbol{\epsilon}_{2,l})$ simplifies to $\boldsymbol{\rho}_t'(\mathbf{c}_t - \mathbf{c}_s + \hat{\mathbf{w}}_s - \hat{\mathbf{w}}_t) = (\boldsymbol{\rho}^*_t)'(\mathbf{c}^*_t - \mathbf{c}^*_s)$. Our methodology works without any change for this case since $\mathbf{w}_t$ and $\hat{\mathbf{w}}_t$ are latent random variables from the econometrician's point of view, and their only role is to perturb the data set to make it consistent with the revealed-preference restrictions. This means that the case presented in this example with price measurement error is indistinguishable from a case with a perfectly measured price vector and error in measuring consumption when measurement error in prices is time-invariant. Assumption 1 can be satisfied under standard conditional independence assumptions between $\boldsymbol{d}_\delta$, $\boldsymbol{\epsilon}_{1,t,l}$, and $\boldsymbol{\epsilon}_{2,l}$, in addition to a restriction that conditional mean of all multiplicative disturbance terms is equal to 1.

4.3. Characterization of the EDRUM via Moment Conditions

Now, we recast the empirical content of the EDRUM in a form amenable to statistical testing. In particular, we write a set of moment conditions that will summarize its empirical content. Recall that the $\mathbf{x} \in X$ denotes observed quantities. Let $\mathbf{e} = (\mathbf{d}_\delta, \mathbf{v}', \mathbf{w}')' \in E$ denote the vector of latent random variables. 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 bold face font letters denote random objects). Define the following moment functions

$$\begin{aligned} g_{I,t,s}(\mathbf{x},\mathbf{e}) &= \mathbb{1}\left(\mathbf{v}_t - \mathbf{v}_s - \mathbf{d}_\delta^{-t}\boldsymbol{\rho}_t'[\mathbf{c}_t - \mathbf{w}_t - \mathbf{c}_s + \mathbf{w}_s] \geq 0\right) - 1, \quad t \neq s \in \mathcal{T}, \\ g_{M,t}(\mathbf{x},\mathbf{e}) &= \mathbf{d}_\delta^{-t}\boldsymbol{\rho}_t'\mathbf{w}_t, \quad t \in \mathcal{T}, \\ g_I(\mathbf{x},\mathbf{e}) &= (g_{I,t,s}(\mathbf{x},\mathbf{e}))_{t\neq s\in\mathcal{T}}, \\ g_M(\mathbf{x},\mathbf{e}) &= (g_{M,t}(\mathbf{x},\mathbf{e}))_{t\in\mathcal{T}}, \\ g(\mathbf{x},\mathbf{e}) &= (g_I(\mathbf{x},\mathbf{e})', g_M(\mathbf{x},\mathbf{e})')'. \end{aligned}$$

We end up having $k = |\mathcal{T}|^2 - |\mathcal{T}|$ and $q = |\mathcal{T}|$ moment functions corresponding to inequality conditions and the mean budget neutrality conditions, respectively. 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 $\pi \in \mathcal{P}_X$.

Lemma 3. *The following are equivalent:*

(i) A random vector $\mathbf{x} = (\boldsymbol{\rho}_t, \mathbf{c}_t)_{t\in\mathcal{T}}$ is approximately consistent with the EDRUM 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 $\mathbf{x}$.

Lemma 3 establishes the equivalence between the EDRUM with the mean budget-neutral measurement error and a system of moment conditions. However, it does not put any restrictions on the support of the distribution of the discount factor since $\mathcal{P}_{E|X}$ is not closed. Any distribution of the discount factor with a support contained in $(0, 1]$ can be achieved as the limit of a sequence of distributions with support equal $(0, 1]$. To allow for more informative, but still flexible, support of $\mathbf{d}_\delta$, we use Assumption 2.

Assumption 2. *(Compact Support) The random discount factor has support contained in $D_\delta = [\delta_1, \delta_2]$, where $\delta = (\delta_1, \delta_2)'$, and $0 < \epsilon \leq \delta_1 \leq \delta_2 \leq 1$ for some $\epsilon > 0$.*

The parametrization in Assumption 2 is not restrictive, and can easily be replaced by any parametric restriction on the support of the discount factor. For instance, instead of a compact interval, one can assume that support is discrete or a finite union of disjoint compact intervals.

Under Assumption 2, $\mathcal{P}_{E|X}$ has to be modified. We now define $\mathcal{P}_{E|X}(\delta) \subseteq \mathcal{P}_{E|X}$ as a set of all probability measures μ such that $\mathbb{E}_{\mu\times\pi_0}\left[\mathbb{1}\left(\delta_1 \leq \mathbf{d}_\delta \leq \delta_2\right)\right] = 1$, where π_0 is the observed distribution of $\mathbf{x}$. To emphasize the dependence of the model on the support parameter δ, we reformulate the moment functions as:

$$g(\mathbf{x}, \mathbf{e}; \delta) = \mathbb{1}\left(\delta_1 \leq \mathbf{d}_\delta \leq \delta_2\right) g(\mathbf{x}, \mathbf{e}).$$

Theorem 2. *The following are equivalent:*

(i) A random vector $\mathbf{x} = (\boldsymbol{\rho}_t, \mathbf{c}^*_t)_{t\in\mathcal{T}}$ is approximately consistent with the EDRUM such that assumptions 1 and 2 hold for some δ.

(ii) There exists a parameter value δ such that

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

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

Theorem 2 tells us that the observed consumption pattern in time, captured by the random vector $(\boldsymbol{\rho}_t, \mathbf{c}_t)_{t\in\mathcal{T}}$, can be generated by the EDRUM under the restrictions on measurement error and the discount factor distributions if and only if there exists a distribution of latent variables conditional on observables that satisfies the revealed-preference inequalities with probability 1.

Our statistical notion of rationalizability makes clear that when one is dealing with measurement error, there is no revealed-preference test that can decide whether a finite sample is consistent with exponential discounting. We can decide only that the data set is asymptotically consistent with the model as $n \to \infty$. Moreover, even asymptotically, there is no way to differentiate between the approximate EDRUM consumer and the true EDRUM consumer. Nonetheless, we can do traditional hypothesis testing and we can decide at a fixed confidence level whether we reject the null hypothesis of (approximate) exponential discounting under assumptions 1 and 2 for a given sample. Conceptually, our notion of rationalizability corresponds to the extended notion of 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 survey data that suffers from measurement error of the type described only in terms of observables.

5. 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 (MEM) conditions using Schennach (2014). Next, we provide a semi-analytic solution to the MEM conditions. Finally, we propose several procedures to test for the EDRUM and to construct confidence intervals for the discount support parameter δ.

5.1. MEM Conditions and the Semi-analytic Solution

Following Schennach (2014), we define the MEM as follows:

Definition 4. (Maximum Entropy Moment, MEM) The MEM s of the moment $g(x, e; \delta)$, for fixed δ and x is

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

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

In words, the MEM h is the marginal moment of the moment g, at which the latent variable has been integrated out. The MEM depends only on the observable random variables.[7] The importance of the MEM is captured in our main result.

Theorem 3. *The following are equivalent:*

(i) A random vector $\mathbf{x} = (\boldsymbol{\rho}_t, \mathbf{c}^*_t)_{t \in \mathcal{T}}$ is approximately consistent with the EDRUM such that assumptions 1 and 2 hold for some δ.

(ii) There is a parameter value δ such that

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

[7] The choice of the $\eta(e|x; \delta)$ is inessential as it only affects the nuisance parameter value; and leaves testing for the existence of δ unaffected (Schennach (2014)).

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

Proof. The result is a direct application of Theorem 2 and Theorem 2.1 in Schennach (2014). ■

We emphasize that Theorem 3 provides both *necessary and sufficient* conditions for the observed data to be (approximately) generated by a random exponential-discounting consumer. This represents an important gain in power with respect to any of the averaging-based tests of exponential discounting that are usually used in the presence of measurement error.

We must stress that we could impose high-level technical assumptions to ensure that the sequence of random latent variables that approximates random exponential-discounting behavior converges to a proper random variable. Thus, this limiting random variable would ensure that the infimum in Theorem 3 is attained, and that the notion of approximate rationalizability collapses to exact rationalizability. However, this obscures the fact that any assumption in that direction has no testable implications.

The remarkable advantage of applying the results of Schennach (2014) to the dynamic revealed-preference approach is that it marginalizes out the latent random variables. More importantly, we have a robust statistical framework to test for exponential discounting in the presence of measurement error. In particular, we have not made any strong distributional assumptions about $\mathbf{d}_\delta$ or $\mathbf{u}$ (the heterogeneous tastes). The only assumptions are these of exponential discounting and concavity of the utility function, and a centering assumption about the measurement error.

In short, the proposed methodology allows us to test exponential discounting in a robust manner without parametric assumptions about preferences or strong distributional assumptions about the measurement error.[8]

Remark 3. Theorem 3 does not imply that the discount-factor support parameter δ is point identified. In fact, it will always be set identified. Indeed, if δ is such that the moment conditions are equal to zero, then the moment condition will also be satisfied by $\tilde{\delta} = (\delta_1 - \xi, \delta_2 + \xi)'$ for small enough $\xi > 0$.

5.2. Semi-analytic Solution for the MEM for the EDRUM with Measurement Error

One can directly employ the MEM in Theorem 3 to test the EDRUM, and to construct confidence intervals for δ. However, doing so is potentially problematic. One possible concern is that the number of MEM, $k + q = |\mathcal{T}|^2$, 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), 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 correspond to the revealed-preference inequality constraints is equal to $+\infty$,

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

and every component of γ_0 that correspond to the mean budget-neutrality constraint is finite and unique. Thus, for testing purposes (under the null hypothesis of the EDRUM), we can restrict the first k components of γ to be equal to $+\infty$, and then minimize a convex objective function over a parameter space of lower dimensionality.

Assumption 3. *(Nondegeneracy) The conditional distribution of $\mathbf{w}_t|\mathbf{x}$ is not degenerate almost surely in $\mathbf{x}$ for all $t \in \mathcal{T}$.*

Assumption 3 rules out cases in which there is no measurement error. Our methodology still works for cases without measurement error, but in those case, it is preferable to use the equivalent deterministic revealed-preference benchmark.

Definition 5. (User-specified MEM distribution) For every δ, almost surely in $\mathbf{x}$, the user-specified distribution $\eta(\cdot|\mathbf{x};\delta)$ satisfies the following:

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

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

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

$$\int_X \int_{E|X} \exp(\gamma_M' g_M(x, e; \delta)) d\eta(e|x; \delta) d\pi_0 < \infty.$$

The first condition on Definition 5 requires that the support of η allows the inequalities to be satisfied. The second conditions is parallel to assumptions 3. This is a definition and not an assumption, as we are free to choose $\eta(\cdot|x;\delta)$.

Theorem 4. *The following are equivalent:*

(i) A random vector $\mathbf{x} = (\boldsymbol{\rho}_t, \mathbf{c}_t^*)_{t\in\mathcal{T}}$ is approximately consistent with the EDRUM such that assumptions 1-3 hold for some δ.

(ii) There is a parameter value δ such that for any sequence $\{\gamma_{I,m}\}_{m=1}^{\infty}$ that componentwise diverges to $+\infty$,

$$\lim_{m\to+\infty} \min_{\gamma_M\in\mathbb{R}^q} \left\| \mathbb{E}_{\pi_0}\left[h(\mathbf{x}; \delta, (\gamma_{I,m}', \gamma_M')') \right] \right\| = 0, \tag{2}$$

where the user-specified measure $\eta(\cdot|\cdot;\delta)$ in the definition of h satisfies conditions in Definition 5. The sequence of minimizers of (2), $\{\gamma_{M,m}\}$, converges to some finite $\gamma_{0,M}$ that does not depend on $\{\gamma_{I,m}\}$.

(iii) There is a parameter value δ such that

$$\min_{\gamma_M\in\mathbb{R}^q} \left\| \mathbb{E}_{\pi_0}\left[\tilde{h}_M(\mathbf{x}; \delta, \gamma_M) \right] \right\| = 0, \tag{3}$$

where

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

and the new measure $d\tilde{\eta}(\cdot|x) = \mathbb{1}\left(g_I(x,\cdot)=0\right)d\eta(\cdot|x)$ and the user-specified measure $\eta(\cdot|\cdot;\delta)$ satisfy the conditions in Definition 5. Moreover, the minimizer of (3) is finite, and is equal to $\gamma_{0,M}$.

Theorem 4 substantially simplifies the conclusion of Theorem 3. First, we need to minimize the norm of the MEM over a much smaller parameter space ($\mathbb{R}^q = \mathbb{R}^{|\mathcal{T}|}$ instead of $\mathbb{R}^{k+q} = \mathbb{R}^{|\mathcal{T}|^2}$). Second, if the data is consistent with the EDRUM, then the minimizer, $\gamma_{0,M}$, has to be finite and unique. Finally, for the implementation purposes, it suffices to replace γ_I by a very large number.[9]

5.3. Testing and Inference

Theorem 4 provides moment conditions that are necessary and sufficient for the data $\{\mathbf{x}_i\}_{i=1}^n = \{(\boldsymbol{\rho}_i', \mathbf{c}_i')'\}_{i=1}^n$ to be approximately consistent with the exponential-discounting model. Let $\theta = (\delta', \gamma') \in \Theta = [\epsilon, 1]^2 \times \mathbb{R}^q$. Recall that δ is the discount-factor support parameter, and γ is a nuisance parameter. For any sequence $\{\gamma_{I,m}\}$ that componentwise diverges to $+\infty$ as $m \to +\infty$, define

$$\hat{h}(\theta) = \frac{1}{n}\sum_{i=1}^n h(\mathbf{x}_i, \delta, (\gamma_{I,m}', \gamma')),$$
$$\hat{\Omega}(\theta) = \frac{1}{n}\sum_{i=1}^n h(\mathbf{x}_i,\theta)h(\mathbf{x}_i,\theta)' - \hat{h}(\theta)\hat{h}(\theta)'.$$

Let $\Omega^-(\theta)$ be the generalized inverse of the matrix Ω. The testing procedure we propose is due to Schennach (2014), and is based on

$$T_n(\delta_0) = n \inf_{\gamma\in\mathbb{R}^{d_q}} \hat{h}(\delta_0,\gamma)'\hat{\Omega}^-(\delta_0,\gamma)\hat{h}(\delta_0,\gamma).$$

The confidence sets for δ can be obtained by inverting $T_n(\delta_0)$. That is, the $(1-\alpha)$-confidence set for δ is $\{\delta_0 \in \Delta : T_n(\delta_0) \le \chi^2_{q,1-\alpha}\}$, where $\chi^2_{q,1-\alpha}$ denotes the $(1-\alpha)$ quantile of χ^2 distribution with q degrees of freedom (χ^2_q). One can also reject the exponential-discounting model if the confidence set is empty.

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

The assumption that the observed data has bounded support can be weakened. It is only needed to guarantee that the variance of the moment conditions is bounded for all δ and every finite γ.

[9] Alternatively, we can generate the new measure $d\tilde{\eta}(\cdot|x) = \mathbb{1}\left(g_I(x,\cdot)=0\right)d\eta(\cdot|x)$ by sampling from $\eta(\cdot|x)$ and then accepting a draw only if the revealed-preference inequalities captured by $\mathbb{1}\left(g_I(x,\cdot)=0\right)$ are satisfied. The last part amounts to solve a linear program as in Adams et al. (2014) treatment of measurement error.

Theorem 5. *Suppose assumptions 1–4 hold. Then under the null hypothesis that the data is approximately consistent with the EDRUM with* $\delta = \delta_0$

$$\lim_{n\to\infty} \mathbb{P}\left(T_n(\delta_0) > \chi^2_{q,1-\alpha}\right) \leq \alpha, \tag{4}$$

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

If, moreover, the minimal eigenvalue of $\mathbb{V}[\tilde{h}_M(\mathbf{x},\delta,\gamma)]$ *is uniformly bounded away from zero and the maximal eigenvalue of* $\mathbb{V}[\tilde{h}_M(\mathbf{x},\delta,\gamma)]$ *is uniformly bounded from above, then under the alternative hypothesis that the data is not approximately consistent with the EDRUM with* $\delta = \delta_0$

$$\lim_{n\to\infty} \mathbb{P}\left(T_n(\delta_0) > \chi^2_{q,1-\alpha}\right) = 1.$$

Theorem 5 allows us to construct confidence sets for the discount-factor support parameter δ and to test the EDRUM. If we test a particular value of δ, then the procedure may be conservative due to a potential rank deficiency of the variance matrix. If the variance matrix is of full rank, then the testing procedure is consistent. Even in the case of rank deficiency, one can get smaller confidence sets by using the singularity-robust, nonlinear Anderson-Rubin test of Andrews and Guggenberger (2015).

6. Empirical Application: Testing Exponential Discounting in Survey Data

In this section, we use our methodology to test the exponential-discounting hypothesis in a consumer panel data set gathered from single-individual and couples households in Spain. 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* 1985-1997). The data set consists of 185 individuals and 2004 couples expenditures and prices for 17 commodities (categories of goods) recorded over four consecutive quarters. 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 data set also contains information on the nominal interest rate on consumer loans faced by the household in any particular quarter.[10]

Some notable studies in the deterministic revealed-preference context used this Spanish household consumer survey (Beatty and Crawford (2011), Blow et al. (2013), Adams et al. (2014)). The shared conclusion drawn from these previous deterministic studies is that the exponential-discounting hypothesis is rejected for a substantial fraction of single-individual and couples households. That is, a sizable fraction of households behave in a manner inconsistent with the predictions of the model. In contrast, we find that the exponential-discounting model with measurement error cannot

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

be rejected at the 95 percent confidence level for single-individual households. This fact indicates that deterministic tests may not be very informative about the behavior in population due to measurement error. Small violations of the deterministic revealed-preference inequalities will lead to big rejection rates. Introducing the measurement error into the analysis takes these small violations into account.

In what follows, we also establish that our test rejects a special version of the collective household consumption problem presented in Adams et al. (2014) (see Section 8.1). The rejection of this particular version of the collective household consumption problem stands in contrast with the conclusion of Adams et al. (2014). In the later work, a deterministic framework without measurement error finds the collective household consumption problem to be accepted for virtually all individuals in the considered sample.

To the best of our knowledge, we are the first to perform traditional statistical hypothesis testing of the exponential-discounting consumer model in the presence of measurement error in a completely nonparametric fashion. Our empirical findings suggest that the revealed-preference practitioners should take into account measurement error in order to be able to conduct traditional hypothesis testing. Furthermore, we wish to emphasize that our methodology does not require identification of the distribution of the discount factor. Nevertheless, we still can construct confidence sets for its support.

We make one additional assumption to simplify the problem in computational terms.

Assumption 5. *(Fixed Upper Bound for the Random Discount Factor) The upper bound of the support of the random discount factor $max(supp(D_\delta)) = 1$ or, equivalently, $\delta_2 = 1$.*

There is some evidence in experimental setups that supports the restriction that the discount factor distribution upper bound is close to unity (e.g., Montiel Olea and Strzalecki (2014)); thus, we choose $\delta_2 = 1$. This is a weak assumption that helps to improve the visualization of the present method and simplifies the optimization procedure. More important, our results are still valid even if this assumption does not hold. In case this assumption fails and the null hypothesis is not rejected, we can only conclude that there is a distribution of the discount factor with support on or inside $[\delta_1, 1]$.

6.1. The Results

Single-Individual Households

We apply the deterministic methodology of Browning (1989) to single-individual households. Our conclusion is that 81.1 percent of the single-individual households behave inconsistently with exponential discounting. Next, we revisit this conclusion using our methodology which addresses measurement error while allowing a heterogeneous discount factor. We find that at the 95 percent confidence level we cannot reject exponential discounting. First, for every fixed δ_1 on the grid $\{0.1, 0.15, \ldots, 1\}$, we compute $T_n(\delta_1)$ and compare it with $\chi^2_{4,0.95} = 9.5$. The results are presented in figure 1. The smallest value of the lower bound of the support of the discount factor to pass the test is 0.1 ($T_n(0.1) = 3.05$); for $\delta_1 \geq 0.55$ the test is rejected. That means that the support of $\mathbf{d}_\delta$ cannot be a subset of $[0.55, 1]$ with at least 95 percent confidence level.

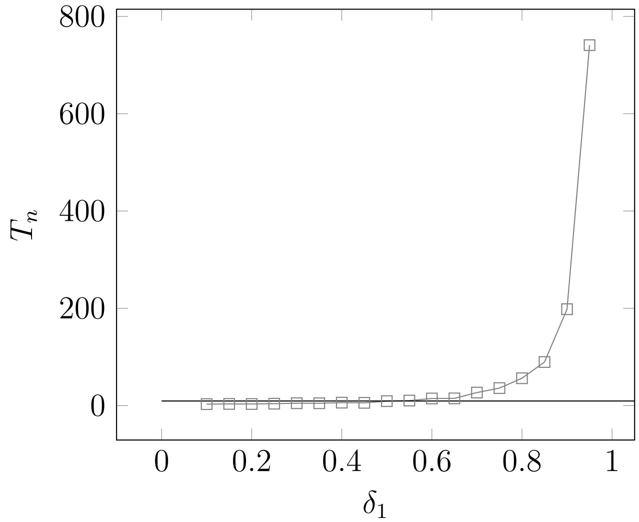

Figure 1 – Values of T_n. Horizontal line corresponds to 95 percent quantile of χ^2_4.

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 to be 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, our method does not reject the exponential-discounting model for single-individuals households. But we do reject the model for couples' households. In the case of couples' households, the test statistic is above the conservative critical value based on the χ^2_4 distribution.

The test statistic takes its minimum value in the grid search at $\delta_1 = 0.1$.[11] The value of the test statistic is $T_n(0.1) = 9.94$. This is above the conservative critical value $\chi^2_{4,0.95} = 9.5$.[12] Hence, we reject the null hypothesis of exponential discounting for all values of δ_1 at the 95 percent confidence level.

This is the first nonparametric evidence robust to measurement error that rejects the existence of a population of household consumers that are consistent with exponential-discounting. We must emphasize that the EDRUM is rejected with a very weak assumption about measurement error in the couples case, while it is not rejected in the single-individual household case. This should convince practitioners about the importance of modeling intrahousehold decision making when

[11]We searched for $\delta_1 \in [0.1, 1]$, with a grid definition of 0.3. Since we reject for $\underline{\hat{\delta}} = 0.1$, we do not need a finer grid, since it will be rejected for the rest of the points, as the numerical exercise confirms.

[12]The test statistic for the rest of the grid jumps to 13.34 for $\delta_1 = 0.15$, and then is increasing in the values of the grid $\{0.20.40.70.9\}$ until it reaches 1084.279.

dealing with intertemporal choice. While the exponential-discounting model seems to be a good model for single individuals, it fails for couples.

In what follows, we establish that a suitable application of our methodology, also rejects the collective household consumption problem presented in Adams et al. (2014). In the model in Adams et al. (2014), each member of a couples' household is consistent with the EDRUM under two assumptions: *(i)* full efficiency in the provision of household public goods, and *(ii)* common support for the distribution of preferences, and random discount factors (as well as relative household power, as captured by random Pareto weights to be defined in Section 8.1).

7. 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). In the second experiment, we provide evidence for the power of our testing procedure with respect to a fixed alternative.

7.1. 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 (Blow et al., 2013, Browning, 1989). However, our test will not reject it.

We choose our simulation configuration setup to match those of the single-individual household characteristics in our application. The Monte Carlo exercise will deal with a moderate size data set of $n = 200$ individuals to show that it works in a data set of the same size in our application. The sample size is $n \in \{200, 1500\}$, 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).[13] These are the prices faced by the single-individual households in our application. We consider consumers with the constant 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. The consumers are exponential

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

discounters with heterogeneous random discount factor $\mathbf{d}_\delta \sim U[0.8, 1]$.

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

$$c^*_{t,l} = \left(\frac{\lambda}{d^t_\delta}\rho_{t,l}\right)^{-1/\sigma},$$

for all $l = 1, \cdots, L$ and $t \in \mathcal{T}$. We normalize the Lagrange multiplier λ to be equal to one. Measurement error is drawn from the following distribution $\boldsymbol{\epsilon}_{t,l} \sim U[0.98, 1.02]$ which implies that $E[\boldsymbol{\epsilon}_{t,l}] = 1$. Then observed consumption is equal to true consumption times the multiplicative perturbation $c_{t,l} = c^*_{t,l}\epsilon_{t,l}$, and we define measurement error as $w_{t,l} = c_{t,l} - c^*_{t,l}$. Note that the implied random measurement error $\mathbf{w}_{t,l}$ has $E[\mathbf{w}_{t,l}] = 0$ by construction. The random vector $\boldsymbol{\epsilon}$ captures incorrect consumption reporting or recording, and can be as high as 1.02 times the true consumption. This means that relative measurement error is around 2 percent. Also, we observe that assumption 1 is satisfied by the proposed measurement error:

$$\mathbb{E}\left[\mathbf{d}_\delta^{-t}\boldsymbol{\rho}_t'\mathbf{w}_t\right] = 0 \quad \forall t \in \mathcal{T}.$$

This is true because $\mathbb{E}\left[\boldsymbol{\rho}_t'\mathbf{w}_t | \boldsymbol{\rho}_t, \mathbf{d}_\delta\right] = 0$, given that $\mathbb{E}\left[\mathbf{w}_{t,l} | \boldsymbol{\rho}_t, \mathbf{d}_\delta\right] = 0$ a.s.. This produces a data set of $(\boldsymbol{\rho}_{t,i}, \mathbf{c}_{t,i})_{i=1, t\in\mathcal{T}}^{i=n}$. We replicate the experiment $m = 1000$ times. The deterministic test in Browning (1989) rejects the exponential-discounting model in 61.5 (62.3) percent of the cases on average across the samples for $n = 200$ ($n = 1500$), while our methodology accepts the null hypothesis that all single households are consistent with random exponential-discounting (as seen in Section 7.2).

7.2. Power Analysis

We choose our simulation configuration setup to match Section 7.1 (with a sample size of $n = 1500$). However, the consumers are sophisticated quasi-hyperbolic discounters with heterogeneous random discount factor $\mathbf{d}_\delta \sim U[0.8, 1]$. The quasi-hyperbolic discounting behavior is controlled by the present bias parameter $\beta \in \{0.5, 0.6, \ldots, 1\}$, which is the same for all consumers. When $\beta = 1$, the consumers are consistent with the EDRUM and our testing procedure should accept the null hypothesis asymptotically at least with probability $1 - \alpha$. For $\beta \neq 1$ the consumers are not consistent with the EDRUM. Thus, for $\beta \neq 1$ we should reject the EDRUM asymptotically with probability 1.

Following Blow et al. (2013) the consumption rule for each consumer and realized d_δ is given by:

$$c^*_{t,l} = \left(\frac{\lambda}{d^t}\rho_{t,l}\prod_{i=1}^{t}\frac{1}{[1-(1-\beta)\mu_i]}\right)^{-1/\sigma}, \quad l = 1, \cdots, L;\ t \in \mathcal{T},$$

where λ is the Lagrange multiplier normalized to be $\lambda = 1$, and $\mu_t \in [0, 1]$ for all $t \in \mathcal{T}$ captures the individual (realized) wealth effects for each realization of income level at time t. Note that

$\mu_t = \sum_{l=1}^{L} p_{t,l} \frac{\partial c_{t,l}^*}{\partial a_t}$, where a_t represents the assets at time t. Since the CES utility function implies that there is at least one normal good, it follows that $\mu_t \in [0,1]$. Therefore, we generate the data set by letting $\boldsymbol{\mu}_t$ be uniformly distributed on $[0,1]$. The randomness of $\boldsymbol{\mu}_t$ captures here the differences in wealth levels across time and across consumers. The data generating process for measurement error coincides with the one presented in Section 7.1. We conduct the experiment $m = 1000$ times for each value of β.

The results are presented in figure 2. For $\beta = 1$, as expected, the rejection rate is close to 5

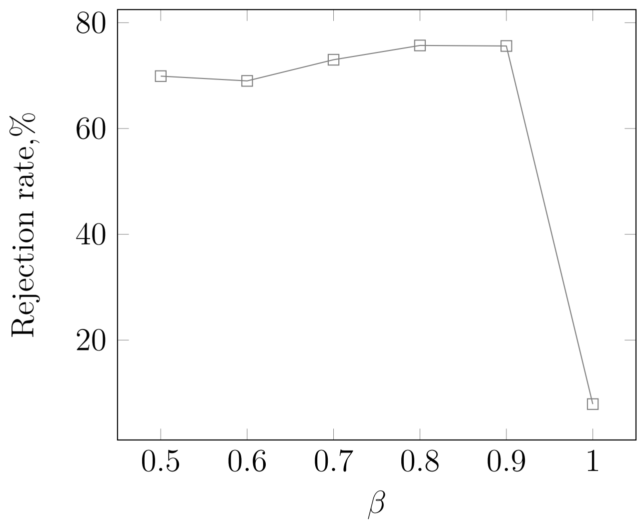

Figure 2 – Rejection rates for different levels of the present bias parameter β.

percent (7.9 percent). For $\beta \neq 1$ the rejection rates are greater or equal than 69 percent.

8. Extensions

In this section we study variants and generalizations of the exponential-discounting model within our framework. This illustrates the wide applicability of our methodology.

8.1. Testing for the Collective Exponential-Discounting Model in Our Framework

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). In this sense, 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. Adams et al. (2014) find that the couples' household data set we are studying is rationalized by this dynamic collective consumption model. They establish this conclusion using the deterministic approach without measurement error. In the present context, it is natural to ask if this conclusion is robust to the presence of measurement error. We formulate a test for the collective model using our methodology. In contrast to Adams et al. (2014), we reject the null hypothesis of consistency of the data set with the dynamic collective model.

This finding is important for at least the following reasons. First, the collective exponential-discounting model presented in Adams et al. (2014) could be considered as a potential alternative to the household exponential-discounting model. Second, even if the deterministic approach rationalizes the data set, the presence of measurement error may change these results.

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 that 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 $p_{t,A} + p_{t,B} = p_{t,H}$.

The existence of Lindahl prices allows us to think of members of the household as autonomous (but interlinked) exponential discounters. Under full efficiency in the collective household problem 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 $\mathbf{d}_{\delta,A}$ and $\mathbf{d}_{\delta,B}$ 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 random exponential-discounting case, namely, we assume that the preferences and Pareto weights remain stable for each household after being drawn from the joint distribution of $(\mathbf{u}_A, \mathbf{d}_{\delta,A}, \boldsymbol{\omega}_A)$ and $(\mathbf{u}_B, \mathbf{d}_{\delta,B}, \boldsymbol{\omega}_B)$ at the first time period. With these preliminaries in hand, we can establish and prove a stochastic analogue to the result in Adams et al. (2014).

Theorem 7. *If a random vector* $(\boldsymbol{\rho}_t, \mathbf{c}_t^*)_{t\in\mathcal{T}}$ *is generated by a collective household with random exponential discountings under full efficiency, then there exist random variables* $\mathbf{d}_{\delta,A}, \mathbf{d}_{\delta,B}$ *which are both supported on or inside* $[\delta_1, \delta_2] \subset (0, 1]$*, and strictly positive random vectors* $(\mathbf{v}_{t,A})_{t\in\mathcal{T}}$*,* $(\mathbf{v}_{t,B})_{t\in\mathcal{T}}$ *that satisfy*

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

Theorem 7 does not provide sufficient conditions for collective rationalizability. It must be clear that 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. Without loss of generality we assume that household member A is asked to report the total household consumption expenditure level (which is in fact what the econometrician observes). In that case, we replace the mean budget neutrality condition (Assumption 1) by the analogous collective mean budget neutrality condition.

Assumption 6. *(Collective Mean Budget Neutrality)* $\mathbb{E}\left[\mathbf{d}_{A,\delta}^{-t}\boldsymbol{\rho}'_t\mathbf{w}_t\right] = 0$*, for all* $t \in \mathcal{T}$.

Under these conditions we find that the minimal value attained by the test statistic for the collective exponential-discounting model is 1950.54, which exceeds even the 99$^{\text{th}}$ quantile of the χ^2_4. Thus, we conclude that the couples' household data set is not consistent with the collective exponential-discounting model under the assumptions of full efficiency, common support for preferences, and the collective mean budget constraint. From this set of assumptions, we believe that full efficiency in the household provision of goods must be relaxed in order to properly assess the collective model. Once again, our results, underscore the importance of taking measurement error into account when assessing the consistency of a mismeasured data set with a model.

8.2. Exponential Discounting with Uncertainty

Here we explore a test proposed by Browning (1989) in the case of uncertainty in income. The income that the consumer receives at time $t \in \mathcal{T}$, y_t, is assumed to be an i.i.d. draw from a random variable $\mathbf{y}_t$. The consumer chooses the consumption level before the uncertainty is resolved. In this case, the revealed-preference theory has to be modified. In particular, we need a martingale difference condition.

Theorem 8. *(Browning (1989)) An array $(\rho_t, c_t)_{t\in\mathcal{T}}$ can be generated by an exponential-discounting model if and only if there exists a constant $\delta \in (0, 1]$, a strictly positive vector $(v_t)_{t\in\mathcal{T}}$, and a stochastic process $\{\boldsymbol{\psi}_t\}_{t\in\mathcal{T}}$, $\boldsymbol{\psi}_t \geq 0$ for all $t \in \mathcal{T}$, such that*

$$\mathbb{E}\left[\boldsymbol{\psi}_t - \boldsymbol{\psi}_{t-1} | \boldsymbol{\psi}_{t-2}, \cdots, \psi_0\right] = 0 \text{ a.s.},$$

$$v_t - v_s \geq \frac{\boldsymbol{\psi}_t}{\delta^t} \rho_t' [c_t - c_s] \text{ a.s.}.$$

It is evident from Theorem 8, that exponential discounting with uncertainty in income has no testable implications. This is true even in the case in which there is no measurement error. In fact, we can always come up with an idiosyncratic income shock as captured by the random vector $(\boldsymbol{\psi}_t)_{t\in\mathcal{T}}$ to rationalize any realized data set $(\rho_t, c_t)_{t\in\mathcal{T}}$.

Keeping this result in mind is essential for understanding our empirical findings and our methodology. Even when the model for exponential discounting with uncertainty is used in the parametric literature, it does not impose any meaningful nonparametric restrictions on the data set we are considering here. Evidently, under parametric assumptions about income shocks we may have some testable implications, but these implications are beyond the scope of this work.

Similarly, an analogous generalization of the EDRUM for the case of uncertainty in future income has no empirical content.

8.3. The Case of Quasi-Hyperbolic Discounting

Here we explore the extension of the revealed-preference theory for the quasi-hyperbolic discounting framework provided by Blow et al. (2013) for the deterministic case. In particular, we consider the case in which the consumer is not an exponential discounter but behaves like a quasi-hyperbolic discounter instead. There are two well-known cases of this in the literature: the naive quasi-hyperbolic discounter and the sophisticated discounter.

Blow et al. (2013) establish that the testable implications of the naive quasi-hyperbolic discounter are exactly the same as those for the exponential discounter; for that reason it is not possible to distinguish between the two models of consumer behavior in the decision environment of interest here.

In contrast, the sophisticated quasi-hyperbolic consumer model has deterministic testable implications that are distinct from those of the exponential-discounting model.

Theorem 9. *(Blow et al. (2013)) An array $(\rho_t, c_t)_{t\in\mathcal{T}}$ can be generated by the sophisticated quasi-hyperbolic discounting model if and only if there exist a constant $d \in (0, 1]$, a positive vector $(v_t)_{t\in\mathcal{T}}$, and a vector $(\psi_t)_{t\in\mathcal{T}}$, $\psi_t \geq 0$ for all $t \in \mathcal{T}$, such that*

$$
\begin{aligned}
&1 \leq \psi_t \leq \psi_{t-1}, \quad \forall t \in \{1, \cdots, |\mathcal{T}| - 1\},\\
&v_t - v_s \geq \frac{\psi_t}{d^t}\rho_t'[c_t - c_s], \quad \forall t, s \in \mathcal{T}.
\end{aligned}
$$

We can apply the ELVIS methodology (Schennach (2014)) for an analogous reformulation of the sophisticated quasi-hyperbolic case. However, our empirical results point in a different direction. For the case of single-individual households the exponential-discounting model cannot be rejected. For the case of couples' households it seems unlikely that present bias is the reason for the rejection of the EDRUM. Instead, we leave as an open question the nature of the interaction between time inconsistency and collective decision making for households in the lines first proposed by Adams et al. (2014) in the case of measurement error.

8.4. A Population Average-Based Test for Exponential-Discounting and the Collective Model

Here, we show that a special version of the model in Adams et al. (2014) has the same average implications as the EDRUM. In the special version, these two conditions must hold: *(i)* full efficiency in the provision of household public goods, and *(ii)* common distribution of preferences, random discount factors, and relative Pareto weights.

A direct implication of Lemma 2 is the following.

Corollary 1. *If a random vector $(\boldsymbol{\rho}_t, \mathbf{c}_t^*)_{t\in\mathcal{T}}$ is consistent with the EDRUM, then there exist a random variable $\mathbf{d}_\delta$ supported on $[\delta_1, \delta_2] \subset (0, 1]$, and a strictly positive random vector $(\mathbf{v}_t)_{t\in\mathcal{T}}$ that satisfies*

$$
\mathbb{E}[\mathbf{v}_t] - \mathbb{E}[\mathbf{v}_s] \geq \mathbb{E}\left[\mathbf{d}_\delta^{-t}\boldsymbol{\rho}_t'[\mathbf{c}_t^* - \mathbf{c}_s^*]\right] \geq 0
$$

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

Next we formalize the assumption we need to establish the equivalence of the EDRUM to the collective household model.

Assumption 7. *(Commonly Distributed Preferences and Pareto Weights) $(\mathbf{u}_A, \mathbf{d}_{\delta,A}, \boldsymbol{\omega}_A)$ and $(\mathbf{u}_B, \mathbf{d}_{\delta,B}, \boldsymbol{\omega}_B)$ have the same distribution.*

Assumption 7 does not imply that $\mathbf{d}_{\delta,A}$ and $\mathbf{d}_{\delta,B}$ are almost surely equivalent. However, this assumption does imply that all quantities that depend only on the distribution (e.g., expectations) are the same. We believe that this assumption is not overly restrictive. In fact, it allows for the deterministic heterogeneity that the collective household model in Adams et al. (2014) requires. Intuitively, each household member is assumed to be a realization of a random exponential discounter given some Lindahl prices (discounted by the random interest rate).

Theorem 10. *If a random vector $(\boldsymbol{\rho}_t, \mathbf{c}^*_t)_{t\in\mathcal{T}}$ is generated by a collective household with random exponential discountings, full efficiency, and which satisfies Assumption 7, then there exist a random variable $\mathbf{d}_\delta$ supported on $[\delta_1, \delta_2] \subset (0, 1]$, and a strictly positive random vector $(\mathbf{v}_t)_{t\in\mathcal{T}}$ that satisfy*

$$\mathbb{E}[\,\mathbf{v}_t\,] - \mathbb{E}[\,\mathbf{v}_s\,] \geq \mathbb{E}\left[\,\mathbf{d}_\delta^{-t}\boldsymbol{\rho}_t'[\mathbf{c}^*_t - \mathbf{c}^*_s]\,\right] \geq 0,$$

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

Surprisingly, Theorem 10 implies that our average-based empirical test of exponential discounting is also a test of a collective household model with exponential discounting (assuming full efficiency, common distribution of the random discount factors, the utility functions, and the Pareto weights). The average-based moment inequalities can be implemented using the ELVIS methodology by means of a (random) slack variable that converts the inequalities to equalities. This result can be seen as a simple and unified test of both models. Nonetheless, the results in this section also warn us about the loss of power incurred by using only averages – instead of all the distributional information, as our main test does.[14]

9. Conclusion

We extend the revealed-preference approach to a stochastic environment in order to test for exponential discounting. We provide Monte Carlo evidence suggesting that the deterministic revealed-preference approach may fail if the observed data has a measurement error. The methodology presented here is able to outperform the deterministic approach in this particular environment.

We apply our methodology to a widely used consumption panel household survey, and we find robust evidence against the exponential-discounting model in the case of couples' households. More surprisingly, we cannot reject the hypothesis that the data set was generated by random exponential discounting for the case of single-individual households. This finding goes against the conclusions of the treatments of the same data sets that use a deterministic revealed-preference approach (Adams et al. (2014), Beatty and Crawford (2011), Blow et al. (2013)).

Our findings provide a guide for future research. Acceptance of the exponential-discounting model means that in some situations, such as in our application, consumers may behave in a time-consistent manner. This may happen when the decision is made frequently giving space to developing expertise. More important, the decisions that we study, in contrast to low-stakes decisions in laboratory situations, involve moderate-stakes. The consumers may be more likely to exert self-control and behave in a time-consistent manner when making decisions about purchasing food and services. In our application the consumers decide over their budget shares among different good-categories such as food, clothing, and transportation. Arguably, this type of decision is both

[14] In a very preliminary version of this paper, we implemented the average-based test with exactly the same results that we obtained using the distributional test. This version is available upon request.

repetitive and important enough for the consumers to be time-consistent. In contrast, for the case of couples' households we reject the unitary exponential-discounting model, and the collective exponential-discounting model with full efficiency. Our results when compared with the single household evidence suggest that time-inconsistencies in the consumption behavior in the couples' case arise due to preference aggregation. We believe that the collective model may still be a suitable candidate for explaining this behavior, but the assumption of full efficiency must be relaxed. We left as a future avenue of research the study of a collective household model where consumption is no longer on the Pareto frontier but rather departs from it stochastically. Inefficiencies on the provision of household public goods may arise due to strategic behavior, miscommunication or mistakes.

References

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

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

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

Andrews, D. W. and Guggenberger, P. (2015). Identification-and singularity-robust inference for moment condition models. *Cowles Foundation Discussion Paper No. 1978.*

Beatty, T. K. and 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., and Crawford, I. (2013). Never Mind the Hyperbolics: Nonparametric Analysis of Time-Inconsistent Preferences. *Unpublished manuscript.*

Blundell, R. W., Browning, M., and 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.*

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

Campbell, J. Y. and 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., Crossley, T. F., and Sabelhaus, J. (2014). Introduction to" improving the measurement of consumer expenditures". In *Improving the Measurement of Consumer Expenditures*, pages 1–20. University of Chicago Press.

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

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

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

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

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

Dunn, K. B. and 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., and Saito, K. (2014). Testable Implications of Quasi-Hyperbolic and Exponential Time Discounting.

Ekeland, I., Galichon, A., and 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., and Srisuma, S. (2016). Nonparametric euler equation identification and estimation. *Working Paper*.

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

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

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

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

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. and Singleton, K. J. (1982). Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica: Journal of the Econometric Society*, pages 1269–1286.

Mastrobuoni, G. and Rivers, D. A. (2016). Criminal discount factors and deterrence.

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

Quah, J., Nishimura, H., and Ok, E. A. (2015). A Comprehensive Approach to Revealed Preference Theory. Technical report, University of Oxford, Department of Economics.

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

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

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

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

10. Appendix

10.1. Proof of Lemma 2

Proof. First we establish that (ii) implies (i). Given $\mathbf{d}_\delta$ and $\mathbf{v}_t$ define the random utility $\mathbf{u}(\cdot)$ as

$$\mathbf{u}(c) = \min_{t\in\mathcal{T}}\{\mathbf{v}_t + \mathbf{d}_\delta^{-t}\boldsymbol{\rho}_t'(c - \mathbf{c}_t^*)\} \text{ a.s..}$$

Then the triple $(\mathbf{u}, \mathbf{d}_\delta, \boldsymbol{\lambda})$ such that $\boldsymbol{\lambda} = 1$ a.s. satisfies the definition of EDRUM. Indeed, by construction $\mathbf{u}(\cdot)$ is continuous and concave with probability 1. Moreover, since

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

We have that $\mathbf{u}(\mathbf{c}_\mathbf{t}^*) = \mathbf{v}_t$ a.s.. Also, we have that $\mathbf{u}(c^*) \leq \mathbf{v}_t$ a.s. for c^* affordable under the flow constraints. Hence, $\mathbf{d}_\delta^{-t}\boldsymbol{\rho}_t \in \nabla\mathbf{u}(\mathbf{c}_t^*)$ a.s. when c_t is strictly positive, and in general there exist a $\xi \in \mathbf{u}(\mathbf{c}_t^*)$ such that $\xi \leq d_\delta^{-t}\boldsymbol{\rho}_t$ a.s..

To prove that (i) implies (ii), first note that concavity of $\mathbf{u}(\cdot)$ and the definition of an EDRUM implies that with probability 1 for any s, t and some $\xi \in \nabla \mathbf{u}(\mathbf{c}_t)$

$$\mathbf{u}(\mathbf{c}_s^*) - \mathbf{u}(\mathbf{c}_t^*) \leq \xi'(\mathbf{c}_s^* - \mathbf{c}_t^*),$$
$$\xi \leq \boldsymbol{\lambda} \mathbf{d}_\delta^{-t} \boldsymbol{\rho}_t.$$

Let N be the set of indices such that $d_\delta^{-t} \boldsymbol{\rho}_{ti} = \xi_i$ for every $i \in N$. So $d_\delta^{-t} \boldsymbol{\rho}_{ti} > \xi_i$ for every $i \notin N$. As a result, $\mathbf{c}_{ti}^*$ has to be a corner solution for every $i \notin N$. That is, $\mathbf{c}_{ti}^* = 0$. Thus,

$$\mathbf{u}(\mathbf{c}_s^*) - \mathbf{u}(\mathbf{c}_t^*) \leq \xi'(\mathbf{c}_s^* - \mathbf{c}_t^*) = \sum_{i \in N} \xi_i (\mathbf{c}_{si}^* - \mathbf{c}_{ti}^*) + \sum_{i \notin N} \xi_i \mathbf{c}_{si}^* =$$
$$= \sum_{i \in N} \boldsymbol{\lambda} d_\delta^{-t} \boldsymbol{\rho}_{ti} (\mathbf{c}_{si}^* - \mathbf{c}_{ti}^*) + \sum_{i \notin N} \xi_i \mathbf{c}_{si}^* \leq \sum_{i \in N} \boldsymbol{\lambda} d_\delta^{-t} \boldsymbol{\rho}_{ti} (\mathbf{c}_{si}^* - \mathbf{c}_{ti}^*) + \sum_{i \notin N} \boldsymbol{\lambda} d_\delta^{-t} \boldsymbol{\rho}_{ti} \mathbf{c}_{si}^*,$$

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

$$\mathbf{u}(\mathbf{c}_t^*) - \mathbf{u}(\mathbf{c}_s^*) \geq \boldsymbol{\lambda} \mathbf{d}_\delta^{-t} \boldsymbol{\rho}_t' [\mathbf{c}_t^* - \mathbf{c}_s^*] \quad \forall s, t \in \mathcal{T}.$$

We let $\mathbf{v}_t = \mathbf{u}(\mathbf{c}_t^*)/\boldsymbol{\lambda}$ a.s.. The random variable $\mathbf{v}_t$ is well-defined since $\boldsymbol{\lambda} \in \mathbb{R}_{++}$ (by local non-satiation), and we obtain the conditions in (ii). ■

10.2. Proof of Lemma 3

Proof. The result follows directly from observing that:

$$\inf_{\mu \in \mathcal{P}_{E|X}} \|\mathbb{E}_{\mu \times \pi_0} [g(\mathbf{x}, \mathbf{e})]\| = 0 \iff \exists\, \mu_m \in \mathcal{P}_{E|X} \lim_{m \to +\infty} \mathbb{E}_{\mu_m \times \pi_0} [g(\mathbf{x}, \mathbf{e})] = 0 \iff$$
$$\exists\, \mathbf{e}^m : \forall\, s, t \in \mathcal{T}, \lim_{m \to +\infty} \mathbb{P} \left(\mathbb{1} \left(\mathbf{v}_t^m - \mathbf{v}_s^m - (\mathbf{d}_\delta^m)^{-t} \boldsymbol{\rho}_t' [\mathbf{c}_t^* - \mathbf{c}_s^*] \geq 0 \right) \right) = 1 \;\&\; \lim_{m \to +\infty} \mathbb{E} \left[(\mathbf{d}_\delta^m)^{-t} \boldsymbol{\rho}_t' \mathbf{w}_t^m \right] = 0.$$

■

10.3. Proof of Theorem 4

First we need some preliminaries, and without loss of generality the following assumption.

Assumption 8. *For any $\gamma_t \in \mathbb{R}$, $t \in \mathcal{T}$*

$$\mathbb{E} \left[\exp \left(\sum_{t=1}^{q} \gamma_t \mathbf{p}_t' \mathbf{w}_t \right) \right] < \infty.$$

This assumption is done without loss of generality as $\mathbf{w}_t$ has bounded support for all $t \in \mathcal{T}$, given the boundedness of both true consumption and observed consumption. Both consumptions are bounded because they have to be non-negative and they have to satisfy the flow budget-constraints with strictly (discounted) positive prices.

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

Proof. Take δ as in the statement of the Theorem and some $x \in X$. We will abuse notation by dropping δ in definitions of the objects below.

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

$$\sup_{x,e} \left|\exp(\gamma_{I,m}' g_I(x,e)) - \mathbb{1}\left(g_I(x,e) = 0 \right)\right| \leq \exp(-\min_{i=1,\ldots,k} \gamma_{I,m,i}) \to_{m\to+\infty} 0,$$

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

$$\begin{aligned}&\left\|\int f(e)\exp(\gamma_{I,m}' g_I(x,e))d\eta(e|x) - \int f(e)\mathbb{1}\left(g_I(x,e) = 0 \right) d\eta(e|x)\right\| \leq \\ &\leq \exp(-\min_{i=1,\ldots,k} \gamma_{I,m,i}) \int \|f(e)\|\, d\eta(e|x) \to_{m\to+\infty} 0.\end{aligned}$$

Hence, the sequence of measures $\exp(\gamma_{I,m}' g_I(x,\cdot))d\eta(\cdot|x)$ converges to the measure $\mathbb{1}\left(g_I(x,\cdot) = 0 \right) d\eta(\cdot|x)$ in total variation. The later measure is well defined and nontrivial since we assume that $\tilde{E}|X = \{e : \mathbb{1}\left(g_I(x,e) = 0 \right)\}$ has a positive measure under $\eta(\cdot|x)$. Let $d\tilde{\eta}(\cdot|x)$ denote $\mathbb{1}\left(g_I(x,\cdot) = 0 \right) d\eta(\cdot|x)$.

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

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

Assumption 5.(iii) together with Step 1 imply that for any compact set $\Gamma \in \mathbb{R}^q$, uniformly in $\gamma_M \in \Gamma$

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

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

$$\inf_{\gamma_M\in\mathbb{R}^q} \left\|\mathbb{E}_{\pi_0}\left[\tilde{h}_M(\mathbf{x};\delta,\gamma) \right]\right\|. \tag{5}$$

Step 3. Note that (5) is equivalent to the optimization problem in Theorem 3. Hence, infimum in (5) is equal to 0 if and only if the data is approximately consistent with the EDRUM.

We assumed that every component of g_M takes both positive and negative values on some non-zero measure subsets of $\tilde{E}|X$. Hence, following the proof of Theorem 2.1 and Lemma A.1 in Schennach (2014), we can conclude that if infimum in (5) is equal to 0, then it is achieved at some finite and unique $\gamma_{0,M}$. Otherwise, $\|\gamma_M\|$ diverges to infinity. ■

10.4. Proof of Theorem 5

Proof. First, note that by the proof of Theorem 4, both under the null and the alternative hypothesis,

$$\mathbb{E}_{\pi_0}\left[h(\mathbf{x};\delta,(\gamma'_{I,m},\gamma'_M)')\right] = \left(0_{1\times k}, \mathbb{E}_{\pi_0}\left[\tilde{h}_M(\mathbf{x};\delta,\gamma_M)'\right]\right)' + o(1),$$
$$\Omega_M(\theta) = \tilde{\Omega}_M(\theta) + o(1),$$

where $\tilde{\Omega}_M(\theta)$ is a variance matrix of $\tilde{h}_M$. Hence,

$$T_n(\delta_0) = n \inf_{\gamma\in\mathbb{R}^q} \hat{\tilde{h}}_M(\delta_0,\gamma)'\hat{\tilde{\Omega}}^-_M(\delta_0,\gamma)\hat{\tilde{h}}_M(\delta_0,\gamma) + o_p(1),$$

where

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

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

Theorem 11. *(Theorem F.1, Schennach (2014)) Let data be i.i.d.. If (i) the parameter space for δ, Δ, is compact; (ii) the set*

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

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

$$\lim_{n\to\infty} \mathbb{P}\left(T_n(\delta_0) > \chi^2_{q,\alpha}\right) \le \alpha.$$

Compactness of the parameter space and i.i.d. sample are assumed. Conditions (ii) and (iii) follow from Assumption 4, compactness of $\Delta = [\epsilon, 1]^2$ and continuity of $\tilde{h}$ in δ and γ. Indeed, for every $\delta\in\Delta$ and any finite γ there exist finite positive constants C_1, C_2 and C_3 that only depend on the support of $\boldsymbol{\rho}_t$ and $\mathbf{w}_t$ such that

$$\mathbb{E}\left[\left\|\tilde{h}(\mathbf{x},\delta,\gamma)\right\|\right] \le \frac{C_1}{\epsilon^q}\frac{\exp(\max_j|\gamma_j|\,C_2)}{\exp(-\min_j|\gamma_j|\,C_3)}.$$

Hence, for any nonempty compact set Γ_δ one can take $C = \max_{\gamma\in\Gamma_\delta}\frac{C_1}{\epsilon^q}\frac{\exp(\max_j|\gamma_j|\,C_2)}{\exp(-\min_j|\gamma_j|\,C_3)}$. Similarly, one can use C to bound $\mathbb{E}\left[\left\|\tilde{h}(\mathbf{x},\delta,\gamma)\right\|^2\right]$.

Under the alternative hypothesis, $T_n(\delta_0)/n$ either converges to a positive constant or diverges to infinity. Thus, the test is consistent. ∎

10.5. Proof of Theorem 7

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

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

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

Then we multiply the first inequality by $\mathbf{d}^t_{\delta,A}$, 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 $\mathbf{d}^t_{\delta,B}$. Then we add-up the two inequalities, to obtain:

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

■

10.6. Proof of Theorem 10

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

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

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

Observe that in the basis of the collective household problem, the distribution of (latent) $\mathbf{c}^*_{t,A}, \mathbf{c}^*_{t,B}, \mathbf{c}^*_{t,H}$ and (latent) Lindahl prices $\mathbf{p}_{t,A}, \mathbf{p}_{t,B}$ are defined by the following equations:

(i) $\boldsymbol{\omega}_A \mathbf{d}^t_{\delta,A}\frac{\partial \mathbf{u}_A(C^*_{t,A},C^*_{t,H})}{\partial C^*_{t,A}} \leq \boldsymbol{\lambda}\mathbf{p}_{t,I}$, (ii) $\boldsymbol{\omega}_B \mathbf{d}^t_{\delta,B}\frac{\partial \mathbf{u}_B(C^*_{t,B},C^*_{t,H})}{\partial C^*_{t,B}} \leq \boldsymbol{\lambda}\mathbf{p}_{t,I}$, (iii) $\boldsymbol{\omega}_A \mathbf{d}^t_{\delta,A}\frac{\partial \mathbf{u}_A(C^*_{t,A},C^*_{t,H})}{\partial C^*_{t,H}} + \boldsymbol{\omega}_B \mathbf{d}^t_{\delta,B}\frac{\partial \mathbf{u}_B(C^*_{t,B},C^*_{t,H})}{\partial C^*_{t,H}} \leq \boldsymbol{\lambda}\mathbf{p}_{t,H}$, (iv) $\mathbf{p}_{t,A} = \frac{1}{\boldsymbol{\lambda}}\mathbf{d}^t_{\delta,A}\frac{\partial \mathbf{u}_A(C^*_{t,A},C^*_{t,H})}{\partial C^*_{t,H}}$ (iv) $\mathbf{p}_{t,B} = \mathbf{p}_{t,H} - \mathbf{p}_{t,A}$. For given $\mathbf{d}^t_{\delta,A}, \mathbf{d}^t_{\delta,B}$, $\mathbf{u}_A, \mathbf{u}_B$ and $\mathbf{p}_{t,I}, \mathbf{p}_{t,H}$ and $\boldsymbol{\omega}_A, \boldsymbol{\omega}_B$, since $\boldsymbol{\lambda}$ is common it is normalized to 1 a.s. without loss of generality. Under assumption (7) we have that the conditional distribution of:

$$\mathbf{c}^*_{t,A}, \mathbf{c}^*_{t,H}, \mathbf{p}_{t,A} | \mathbf{u}_A, \mathbf{d}_{\delta,A}, \boldsymbol{\omega}_A,$$

is the same conditional distribution of

$$\mathbf{c}^*_{t,B}, \mathbf{c}^*_{t,H}, \mathbf{p}_{t,B} | \mathbf{u}_B, \mathbf{d}_{\delta,B}, \boldsymbol{\omega}_B.$$

This implies under assumption (7) that the joint distribution $\mathbf{c}^*_{t,A}, \mathbf{c}^*_{t,H}, \mathbf{p}_{t,A}, \mathbf{u}_B, \mathbf{d}_{\delta,B}, \boldsymbol{\omega}_B$ is the same as $\mathbf{c}^*_{t,B}, \mathbf{c}^*_{t,H}, \mathbf{p}_{t,B}, \mathbf{c}^*_{t,A}, \mathbf{c}^*_{t,H}, \mathbf{p}_{t,A}$. In fact, observe that given that $(\mathbf{u}_A, \mathbf{d}_{\delta,A}, \boldsymbol{\omega}_A)$ and $(\mathbf{u}_B, \mathbf{d}_{\delta,B}, \boldsymbol{\omega}_B)$ have the same CDF and given the above conclusion about the equality of the conditional CDF's of $\mathbf{c}^*_{t,A}, \mathbf{c}^*_{t,H}, \mathbf{p}_{t,A}|\mathbf{u}_A, \mathbf{d}_{\delta,A}, \boldsymbol{\omega}_A$, and $\mathbf{c}^*_{t,B}, \mathbf{c}^*_{t,H}, \mathbf{p}_{t,B}|\mathbf{u}_B, \mathbf{d}_{\delta,B}, \boldsymbol{\omega}_B$, then the previous statement follows.

We take expectations and by the previous argument we have that:

$$E[\frac{1}{\mathbf{d}^t_{\delta,A}}\boldsymbol{\rho}'_{t,I}\mathbf{c}^*_{s,B}] = E[\frac{1}{\mathbf{d}^t_{\delta,B}}\boldsymbol{\rho}'_{t,I}\mathbf{c}^*_{s,B}]$$

and $E[\frac{1}{\mathbf{d}^t_{\delta,A}}\boldsymbol{\rho}'_{t,B}\mathbf{c}^*_{s,H}] = E[\frac{1}{\mathbf{d}^t_{\delta,B}}\boldsymbol{\rho}'_{t,B}\mathbf{c}^*_{s,H}]$ for all $s, t \in \mathcal{T}$. Finally, by Theorem 6, and by adding up the two conditions, we conclude that,

$$E[\mathbf{v}_{t,A} + \mathbf{v}_{t,B}] - E[\mathbf{v}_{s,B} + \mathbf{v}_{s,B}] \geq$$

$$E[\frac{1}{\mathbf{d}^t_{\delta,A}}(\boldsymbol{\rho}'_{t,I}(\mathbf{c}^*_{t,I} - \mathbf{c}^*_{s,I}) + \boldsymbol{\rho}'_{t,H}(\mathbf{c}^*_{t,H} - \mathbf{c}^*_{s,H}))] =$$

$$E[\mathbf{d}^{-t}_{\delta,A}\boldsymbol{\rho}^{-t}_t{}'(\mathbf{c}^*_t - \mathbf{c}^*_s)].$$

We let $\mathbf{v}_t = \mathbf{v}_{t,A} + \mathbf{v}_{t,B}$ and $\mathbf{d}_\delta = \mathbf{d}_{\delta,A}$ and we have established the result.

■

10.7. Subvector Anderson-Rubin test of Andrews and Guggenberger (2015)

For every $\delta_0 \in \Delta$, let

$$\tilde{\gamma}_n = \arg\min_{\gamma \in \bar{\mathbb{R}}^{d_\gamma}} \left\| \hat{h}(\delta_0, \gamma) \right\|^2.$$

Let $\hat{\Lambda}_n(\theta)$ be a diagonal matrix with the eigenvalues of $\hat{\Omega}(\theta)$ on the diagonal in nonincreasing order, and $A^*_n(\theta)$ be an orthogonal matrix of the eigenvectors of $\hat{\Omega}(\theta)$ corresponding to $\hat{\Lambda}_n(\theta)$. Denote $\hat{r}(\theta) = rank(\hat{\Omega}(\theta))$. Let $\hat{A}_n(\theta)$ be the first $\hat{r}(\theta)$ columns of $\hat{A}^*_n(\theta)$. Define the nonredundant moment functions, sample moments and sample moment variance matrix as

$$\begin{aligned} h_{\hat{A}_n}(\mathbf{x}_i, \theta) &= \hat{A}_n(\delta_0, \tilde{\gamma}_n)' h(\mathbf{x}_i, \theta), \\ \hat{h}_{\hat{A}_n}(\theta) &= \frac{1}{n}\sum_{i=1}^n h_{\hat{A}_n}(\mathbf{x}_i, \theta), \\ \hat{\Omega}_{\hat{A}_n}(\theta) &= \frac{1}{n}\sum_{i=1}^n h_{\hat{A}_n}(\mathbf{x}_i, \theta) h_{\hat{A}_n}(\mathbf{x}_i, \theta)' - \hat{h}_{\hat{A}_n}(\theta)\hat{h}_{\hat{A}_n}(\theta)'. \end{aligned}$$

Next, define

$$\hat{\gamma}_n = \arg\min_{\gamma \in \bar{\mathbb{R}}^{d_\gamma}} \left\| A_{\hat{A}_n} \hat{h}_{\hat{A}_n}(\delta_0, \gamma) \right\|^2,$$

where $A_{\hat{A}_n}$ is a square matrix that satisfies

$$A'_{\hat{A}_n} A_{\hat{A}_n} = \hat{\Omega}^{-1}_{\hat{A}_n}(\delta_0, \tilde{\gamma}_n).$$

Define "orthogonalized" estimators of the variance of the nonredundant moment functions:

$$\begin{aligned}
\tilde{\Omega}_n(\theta) &= \left(\tilde{\Omega}_{n,1}(\theta), \tilde{\Omega}_{n,2}(\theta), \ldots, \tilde{\Omega}_{n,\hat{r}ank(\theta)}(\theta)\right), \\
\tilde{\Omega}_{n,j}(\theta) &= \frac{1}{n}\sum_{i=1}^{n} h_{\hat{A}_n}(\mathbf{x}_i, \theta) h_{\hat{A}_n,j}(\mathbf{x}_i, \theta) - \hat{\Phi}_{n,j}(\theta)\hat{\Omega}^{-1}_{\hat{A}_n}(\theta)\hat{h}_{\hat{A}_n}(\theta) - \hat{h}_{\hat{A}_n}(\theta)\hat{h}_{\hat{A}_n,j}(\theta), \\
\hat{\Phi}_{n,j}(\theta) &= \frac{1}{n}\sum_{i=1}^{n}\left[h_{\hat{A}_n}(\mathbf{x}_i, \theta) h_{\hat{A}_n,j}(\mathbf{x}_i, \theta) - \frac{1}{n}\sum_{s=1}^{n} h_{\hat{A}_n}(\mathbf{x}_s, \theta) h_{\hat{A}_n,j}(\mathbf{x}_s, \theta)\right] h'_{\hat{A}_n}(\mathbf{x}_i, \theta).
\end{aligned}$$

and the expectation of the Jacobian of the nonredundant moment functions:

$$\begin{aligned}
\tilde{J}_n(\theta) &= \tilde{\Omega}(\delta_0, \hat{\gamma}_n)^{-1/2}\tilde{G}_{\gamma n}(\theta), \\
\tilde{G}_{\gamma n}(\theta) &= \left(\tilde{G}_{\gamma_1 n}(\theta), \tilde{G}_{\gamma_2 n}(\theta), \ldots, \tilde{G}_{\gamma_{d_\gamma} n}(\theta)\right), \\
\tilde{G}_{\gamma_j n}(\theta) &= \frac{1}{n}\sum_{i=1}^{n} \partial_{\gamma_j} h_{\hat{A}_n}(\mathbf{x}_i, \theta) - \hat{F}_{n,j}(\theta)\hat{\Omega}^{-1}_{\hat{A}_n}(\theta)\hat{h}_{\hat{A}_n}(\theta), \\
\hat{F}_{n,j}(\theta) &= \frac{1}{n}\sum_{i=1}^{n}\left[\partial_{\gamma_j} h_{\hat{A}_n}(\mathbf{x}_i, \theta) - \frac{1}{n}\sum_{s=1}^{n} \partial_{\gamma_j} h_{\hat{A}_n}(\mathbf{x}_i, \theta)\right] h_{\hat{A}_n}(\mathbf{x}_i, \theta)',
\end{aligned}$$

where $\partial_{\gamma_j} h(x, \theta)$ is a vector consisting of partial derivatives of moment functions with respect to j-th component of γ.

The subvector Andersen-Rubin (SR-AR) test statistic is

$$T_{2,n}(\delta_0) = n\hat{h}_{\hat{A}_n}(\delta_0, \tilde{\gamma}_n)'\tilde{\Omega}(\delta_0, \hat{\gamma}_n)^{-1/2} M_{\tilde{J}_n(\delta_0, \hat{\gamma}_n)}\tilde{\Omega}(\delta_0, \hat{\gamma}_n)^{-1/2}\hat{h}_{\hat{A}_n}(\delta_0, \tilde{\gamma}_n), \tag{6}$$

where $M_A = I - P_A$ and P_A is a projector on the column space of A. Then the $(1-\alpha)$ confidence set based on $T_{2,n}$ is $\{\delta_0 \in \Delta : T_{1,n}(\delta_0) \leq \chi^2_{\hat{r}((\delta_0, \tilde{\gamma}_n)), 1-\alpha}\}$. Similarly to the case with $T_{1,n}$, one can conduct the model specification test by checking if the confidence set is empty.

For $\delta_1 \in \{0.1, 0.15, 0.50\}$ the value of the subvector Anderson-Rubin test statistics is greater than or equal to 3550 and clearly exceed the critical value $\chi^2_{4,0.05} = 9.49$. The values of the Anderson-Rubin test statistic for $\delta_1 \in \{0.55, 0.60, \ldots, 1\}$ are presented in Figure 3. The estimated rank of the asymptotic variance matrix for these values of δ_1 is equal to 3, so the statistic value is compared to $\chi^2_{3,0.05} = 7.82$.

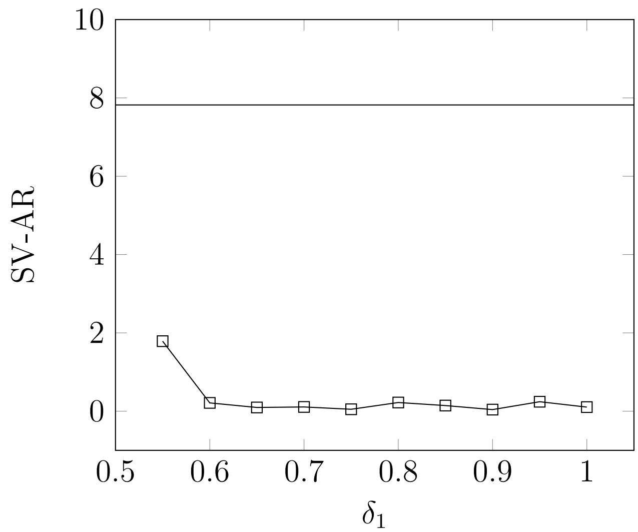

Figure 3 – Values of the subvector Anderson-Rubin test statistic. Horizontal line corresponds to 95 quantile of χ^2_3.