Robust Inference on Discount Factors

Charles Gauthier*

This version: September 2021

Abstract

This paper develops a revealed preference methodology to set identify the discount factor in the exponential discounting model. My approach makes no parametric assumption on the utility function, allows for unrestricted heterogeneity, and accounts for measurement error. Using longitudinal data from checkout scanners, I bound household-specific discount factors and assess their sensitivity to measurement error. I find that accounting for unobserved heterogeneity is important as observable characteristics fail to capture differences in discounting.

JEL Classification: D11, D12.

Keywords: exponential discounting, discount factor, revealed preference

1 Introduction

The exponential discounting model is a predominant tool for analyzing dynamic choice in applied work. Its attractiveness rests in that time preferences are summarized by a single parameter—the discount factor. This allows one to tractably analyze a decision maker's intertemporal choices, which is crucial in a vast range of applications. Accordingly, many studies have tried to recover its key time parameter. However, a common feature in this literature is the specification of the consumer's preferences.[1] This constitutes a potentially important limitation as erroneously specifying preferences may lead to spurious estimates of the discount factor.

*Department of Economics, University of Western Ontario, Social Science Centre, London, Ontario, N6A 5C2 (email: cgauth26@uwo.ca). I wish to thank Victor Aguiar for his helpful comments and suggestions that greatly improved the quality of the paper. I also wish to thank Roy Allen, Audra Bowlus, Maria Goltsman, Nail Kashaev, David Rivers, and anonymous referees for useful comments. The author declares that he has no relevant or material financial interests that relate to the research described in this paper.

[1]For an overview of this large literature, see Frederick et al. (2002).

At its core, the exponential discounting model assumes that the utility function is additively time-separable and stationary. Under these assumptions, the transitivity of preferences can be characterized by the well-known Generalized Axiom of Revealed Preference (GARP). In particular, Afriat (1967) showed that for any finite data set $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ of discounted prices and demands, GARP is necessary and sufficient for the existence of a well-behaved utility function that rationalizes the data. The distinctive feature of exponential discounting, though, is the prediction that consumers will be time consistent. Namely, it requires consumers to commit to their initial plan as time unfolds.[2]

I show that the exponential discounting model, which is normally stated as a dynamic maximization problem with an intertemporal budget constraint, may be expressed as a repeated static utility maximization problem without any budget constraint. Specifically, a consumer is an exponential discounter if and only if there exists a locally nonsatiated instantaneous utility function $u : \mathbb{R}^L_+ \to \mathbb{R}$ and a discount factor $\delta \in (0, 1]$ such that

$$\boldsymbol{c}_t \in \operatorname*{arg\,max}_{\boldsymbol{c}\in\mathbb{R}^L_+} u(\boldsymbol{c}) + \delta^{-t}(y^d_t - \boldsymbol{\rho}'_t\boldsymbol{c}) \quad \forall t \in \mathcal{T},$$

where $y^d_t > 0$ denotes discounted income in period t. Letting $s^d := y^d_t - \boldsymbol{\rho}'_t\boldsymbol{c}$ denote discounted savings and $U_t(\boldsymbol{c}, s^d) := u(\boldsymbol{c}) + \delta^{-t}s^d$, the objective function may be seen as an additively separable time-dependent augmented utility function $U_t : \mathbb{R}^L_+ \times \mathbb{R} \to \mathbb{R}$. The dynamics of the model is captured through the incorporation of savings into the consumer's consideration. Indeed, the amount a consumer is willing to consume in any time period is regulated by his desire to save for future consumption.

My methodology exploits the theory of revealed preference popularized by Afriat (1967) and Varian (1982). This approach obtains sharp conditions that any demand data must satisfy in order to be consistent with utility maximization, and reciprocally, any behavior stemming from utility maximization must satisfy them.[3] In the exponential discounting model, for a given set of observations $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$, these conditions yield a set of linear inequalities that are known up to the discount factor. Since revealed preference conditions are exact, the main

[2] In experimental settings, a preference reversal occurs when the consumer chooses a sooner-smaller reward over a later-larger one and then switches to the later-larger reward when an equal delay is added to both outcomes. This behavior violates time consistency if the consumer deviates from his plan and chooses the sooner-smaller reward in the future (Halevy, 2015).

[3] Although it is possible to impose additional constraints on the utility function, the revealed preference framework does not require it.

requirement maintained in this study is that consumers have perfect foresight. In addition, I impose the marginal utility of discounted expenditure to be constant across time as it is necessary for exponential discounting to have implications beyond GARP (Browning, 1989).

A data set either satisfies or violates the revealed preference inequalities that characterize exponential discounting. This makes the direct implementation of these inequalities of limited applicability as they fail to handle innocuous deviations that may arise in the data. As such, I propose a statistical test that allows for measurement error in variables as in Varian (1985). While this statistical test can be inverted to recover nonparametric bounds on the discount factor, it has the undesirable property to be extremely conservative, thus hindering one's ability to make informative inference. I address this caveat by meaningfully disciplining measurement error in terms of percentage of wasted income.

In my empirical application, I apply my methodology to the checkout scanner panel data set on food expenditures from Echenique, Lee, and Shum (2011). I find that many consumers behave consistently with exponential discounting when measurement error in prices is taken into account. Moreover, I show that bounds on the discount factor get tighter as the extent of measurement error decreases. Finally, I find that observable characteristics such as income, education and age fail to capture heterogeneity in discounting.

The remainder of the paper is organized as follows. Section 2 reviews the related literature. Section 3 formally defines the exponential discounting model, obtains its time-dependent augmented utility representation, and derives its testable implications. Section 5 introduces the statistical test and provides a confidence set for the discount factor. Section 6 contains the empirical application and Section 7 concludes. The main proofs and supplemental material can be found in the Appendices.

2 Related Literature

This paper builds on the exponential discounting characterization of Browning (1989) in order to derive a novel representation of the model in terms of a time-dependent augmented utility function. The use of an augmented utility function has also been used by Deb et al. (2018) in a different framework. They consider the concept of revealed price preference and obtain a consistency condition called the Generalized Axiom of Price Preference (GAPP). The augmented utility function I derive is distinct from theirs as it has the peculiarity of being time-dependent; a notable implication is that exponential discounting can be thought of as a static model with reference-dependent preferences.[4]

This new representation lends itself to a partial efficiency analysis similar to that of Afriat (1973) which allows me to measure the severity of a violation from exponential discounting in the data. In this respect, my result relates to existing partial efficiency results such as those for static utility maximization (Halevy, Persitz and Zrill, 2018; De Clippel and Rozen, 2018), homothetic rationalizability (Heufer and Hjertstrand, 2017), and expected utility maximization (Echenique, Imai and Saito, 2018). I complement these papers by bringing partial efficiency to a dynamic setting. Notably, my extension allows one to use the statistical test of Cherchye et al. (2020) to exponential discounting.

My endeavor is complementary to that of Adams et al. (2014) who extend the analysis of the exponential discounting model for preference heterogeneity and renegotiations within the household. It also relates to models of habit formation such as the one proposed in Crawford (2010) and Demuynck and Verriest (2013) who examine the fit of richer life-cycle models. More generally, my approach is similar to that of Blow, Browning and Crawford (2017) who develop a test for the quasi-hyperbolic model. My work differs from theirs in that I focus on improving the applicability of the standard version of exponential discounting.

My methodology is close to that of Brown and Calsamiglia (2007) who provide conditions for quasilinear utility rationalization, and to Echenique, Imai and Saito (2020) who provide an axiomatic characterization of exponential discounting for experimental data.[5] Instead, my test is aimed to be applied to survey or scanner data where choices are made over multidimensional consumption bundles.

3 Exponential Discounting

In this section, I introduce the notation used throughout the paper, formally define the exponential discounting model, and show how to get nonparametric bounds on the discount factor.

[4] In this light, my representation relates to the literature on reference-dependent utility functions popularized by the seminal work of Kahneman and Tversky (1979).

[5] Their test applies to a single good, a case that more naturally occurs in experiments.

3.1 Notation

The typical scenario under consideration is that of purchases made by a consumer over a certain time window. Let $\mathcal{L} \in \{1, ..., L\}$ denote the number of observed commodities and $\mathcal{T} = \{0, ..., T\}$ the periods for which data on consumers are observable. For any good $l \in \mathcal{L}$ and time period $t \in \mathcal{T}$, denote discounted price by $\rho_{l,t} = p_{l,t}/\prod_{i=0}^{t}(1+r_i)$, where $p_{l,t}$ is the spot price and r_i is the interest rate, and denote consumption by $c_{l,t}$.[6] An observation is therefore a pair $(\boldsymbol{\rho}_t, \boldsymbol{c}_t) \in \mathbb{R}^L_{++} \times \mathbb{R}^L_{+}$, and accordingly, a data set is written as $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$.

3.2 Exponential Discounting Rationalizability

The objective function faced by an exponential discounting (ED) consumer at time $\tau \in \mathcal{T}$ is given by

$$U_\tau(\boldsymbol{c}_\tau, ..., \boldsymbol{c}_{T-\tau}) = u(\boldsymbol{c}_\tau) + \sum_{j=1}^{T-\tau} \delta^j u(\boldsymbol{c}_{\tau+j}),$$

where $u(\cdot)$ is the instantaneous utility function and $\delta \in (0, 1]$ is the discount factor. Moreover, consumption satisfies the linear budget constraint

$$\boldsymbol{\rho}_t'\boldsymbol{c}_t + s_t^d = y_t^d + a_t^d \quad \forall t \in \{\tau, \dots, T\},$$

where s_t^d denotes discounted savings, $y_t^d > 0$ denotes discounted income and a_t is the discounted value of assets held at period t.[7] The assets evolve according to the law of motion: $a_t = (1+r_t)s_{t-1}$. A data set is consistent with exponential discounting if it can be thought of as stemming from the model.

Definition 1. A data set $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ is ED-rationalizable if there exists a locally nonsatiated, continuous, monotonic, and concave instantaneous utility function $u(\cdot)$, an income stream $(y_t^d)_{t\in\mathcal{T}} \in \mathbb{R}^{|\mathcal{T}|}_{++}$, an initial asset level $a_0 \geq 0$, and a discount factor $\delta \in (0, 1]$ such that the consumption stream $(\boldsymbol{c}_t)_{t\in\mathcal{T}}$ solves

$$\max_{(\boldsymbol{c}_t)_{t\in\mathcal{T}}\in\mathbb{R}_+^{L\times|\mathcal{T}|}} u(\boldsymbol{c}_0) + \sum_{t=1}^{T} \delta^t u(\boldsymbol{c}_t) \text{ s.t. } \boldsymbol{\rho}_0'\boldsymbol{c}_0 + \sum_{t=1}^{T} \boldsymbol{\rho}_t'\boldsymbol{c}_t = y_0 + \sum_{t=1}^{T} y_t^d + a_0.$$

The empirical implications of exponential discounting is captured by the following result due to Browning (1989).

[6]The interest rate in the first period is set to zero, that is, $r_0 = 0$.
[7]That is, $s_t^d = s_t/\prod_{i=0}^{t}(1+r_i)$, $y_t^d = y_t/\prod_{i=0}^{t}(1+r_i)$ and $a_t^d = a_t/\prod_{i=0}^{t}(1+r_i)$.

Proposition 1. *The following statements are equivalent:*

(i) The data set $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ *is ED-rationalizable.*

(ii) There exist numbers u_t*,* $t = 0, \ldots, T$*, and a discount factor* $\delta \in (0, 1]$ *such that*

$$u_s \leq u_t + \delta^{-t}\boldsymbol{\rho}_t'(\boldsymbol{c}_s - \boldsymbol{c}_t) \quad \forall s, t \in \mathcal{T}.$$

(iii) There exists a discount factor $\delta \in (0, 1]$ *such that for any subset of indices* $\tau = \{t_i\}_{i=1}^m$ *with* $t_i \in \mathcal{T}$ *and* $m \geq 2$*,*

$$0 \leq \delta^{-t_1}\boldsymbol{\rho}_{t_1}'(\boldsymbol{c}_{t_2} - \boldsymbol{c}_{t_1}) + \ldots + \delta^{-t_m}\boldsymbol{\rho}_{t_m}'(\boldsymbol{c}_{t_1} - \boldsymbol{c}_{t_m}). \tag{CM}$$

Proposition 1 gives two alternative tests for the exponential discounting model. Conditional on $\delta \in (0, 1]$, condition (ii) is a set of linear inequalities and can be solved using linear programming. Turning to (iii), note that pairs of indices $s, s+h \in \mathcal{T}$, where $h \geq 1$, provide bounds on the discount factor. Accordingly, I define the greatest lower bound and the least upper bound on the discount factor as

$$glb := \max_{s,s+h\in\mathcal{T}} \left\{ \left(\frac{\boldsymbol{\rho}_{s+h}'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h})}{\boldsymbol{\rho}_s'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h})} \right)^{1/h} \right\} \text{ such that } \boldsymbol{\rho}_s'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h}) < 0$$

and

$$lub := \min_{s,s+h\in\mathcal{T}} \left\{ \left(\frac{\boldsymbol{\rho}_{s+h}'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h})}{\boldsymbol{\rho}_s'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h})} \right)^{1/h} \right\} \text{ such that } \boldsymbol{\rho}_s'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h}) > 0,$$

whenever such bounds exist, and $glb = 0$, $lub = 1$, otherwise.

To gain some intuition on these bounds, note that when $\boldsymbol{\rho}_s'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h}) < 0$, the consumer does *not* reveal a preference for the earlier bundle over the later one.[8] This can only happen if he is somewhat patient and therefore yields a lower bound on the discount factor. In the case where $\boldsymbol{\rho}_s'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h}) > 0$, the consumer reveals a preference for the earlier bundle over the later one. In turn, this can only happen if he is somewhat impatient and therefore yields an upper bound on the discount factor.

Furthermore, note that the size of $\boldsymbol{\rho}_s'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h}) < 0$ gives an indication of how enticing $\boldsymbol{c}_{s+h}$ is compared to $\boldsymbol{c}_s$ at time s. Likewise, $\boldsymbol{\rho}_{s+h}'(\boldsymbol{c}_s - \boldsymbol{c}_{s+h}) < 0$

[8] For any $t \in \mathcal{T}$, a bundle $\boldsymbol{c}_t$ is said to be revealed preferred to a bundle $\boldsymbol{c}$ if $\boldsymbol{\rho}_t'(\boldsymbol{c} - \boldsymbol{c}_t) \leq 0$. See Appendix A for a detailed review of revealed preference concepts.

gives an indication of how enticing $\boldsymbol{c}_{s+h}$ is compared to $\boldsymbol{c}_s$ at time $s+h$. The more enticing $\boldsymbol{c}_{s+h}$ becomes at time $s+h$ relative to time s, the more patient the consumer gets. Intuitively, when $\boldsymbol{c}_{s+h}$ becomes an increasingly better option at time $s+h$ relative to time s, the consumer's willingness to leave $\boldsymbol{c}_{s+h}$ for later strengthens. In other words, the lower bound takes on larger positive values. A similar interpretation holds for upper bounds.

With these bounds in hand, condition (*iii*) allows me to derive necessary conditions that yield additional intuition on the ED model and will prove useful for computational purposes.

Corollary 1. *The data set $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ is ED-rationalizable only if GARP holds and*

$$glb \leq lub \;;\; glb \leq 1 \;;\; lub > 0. \tag{CD}$$

Corollary 1 states that the exponential discounting model has an additional testable implication compared to static utility maximization. As for the latter, GARP captures within-period consistency. That is, it ensures that the bundle chosen at time t is the best among all feasible bundles in that period. In contrast, condition CD represents the dynamic of the model and guarantees that the intertemporal choices of the consumer are pairwise time consistent. However, these conditions are not sufficient for ED-rationalizability, as the following example displays.

Example 1. *Consider a bivariate demand ($\mathcal{L} = \{1, 2\}$) with three time periods ($\mathcal{T} = \{0, 1, 2\}$). The consumer has a data set where $(\boldsymbol{\rho}_0, \boldsymbol{c}_0) = ([1, 1]', [4, 3]')$, $(\boldsymbol{\rho}_1, \boldsymbol{c}_1) = ([2, 5]', [1, 2]')$ and $(\boldsymbol{\rho}_2, \boldsymbol{c}_2) = ([4, 2]', [3, 6]')$. It is easy to verify that GARP holds. Indeed, $\mathbf{c}_0 R^D \mathbf{c}_1$, $\mathbf{c}_2 R^D \mathbf{c}_0$ and $\mathbf{c}_2 R^D \boldsymbol{c}_1$ so no cycle exists. To see that CD is also satisfied, note that*

$$\frac{\boldsymbol{\rho}_1'(\boldsymbol{c}_0 - \boldsymbol{c}_1)}{\boldsymbol{\rho}_0'(\boldsymbol{c}_0 - \boldsymbol{c}_1)} = \frac{11}{4}, \left(\frac{\boldsymbol{\rho}_2'(\boldsymbol{c}_0 - \boldsymbol{c}_2)}{\boldsymbol{\rho}_0'(\boldsymbol{c}_0 - \boldsymbol{c}_2)}\right)^{1/2} = \frac{-2}{-2} = 1, \textit{ and } \frac{\boldsymbol{\rho}_2'(\boldsymbol{c}_1 - \boldsymbol{c}_2)}{\boldsymbol{\rho}_1'(\boldsymbol{c}_1 - \boldsymbol{c}_2)} = \frac{-16}{-24} = \frac{2}{3}.$$

Clearly, $glb = 1$ and $lub = 11/4$ so the conditions of CD are met. However, the data set does not satisfy CM since when δ is equal to one[9]*,*

$$f(\delta) := \delta^{-2}\boldsymbol{\rho}_2'(\boldsymbol{c}_1 - \boldsymbol{c}_2) + \delta^{-1}\boldsymbol{\rho}_1'(\boldsymbol{c}_0 - \boldsymbol{c}_1) + \delta^{-0}\boldsymbol{\rho}_0'(\boldsymbol{c}_2 - \boldsymbol{c}_0) = -16 + 11 + 2 < 0.$$

In practice, Corollary 1 only involves testing GARP and checking the set

[9]It is sufficient to check $\delta = 1$ as the first-order condition of $f(\delta)$ is strictly positive for all $\delta \in (0, 1]$.

of inequalities in CD. As highlighted by Varian (1982), one can use an efficient algorithm called Warshall's algorithm to get the transitive closure of the direct revealed preference relation. Importantly, these conditions can be parallelized, thus greatly reducing the computational burden when exponential discounting has to be tested repeatedly.

4 Exponential Discounting under Partial Efficiency

This section shows that the exponential discounting model has a time-dependent augmented utility function representation that can be used to account for inconsistent choices in the observed data.

4.1 Time-dependent Augmented Utility Function

The main problem with the results in the previous section is that, when a data set is not exactly ED-rationalizable, it becomes impossible to recover bounds on the discount factor. This is highly prohibitive as the observed data are often inconsistent with the model. For example, in the presence of measurement error the observed data could be inconsistent with the model even if the true data are ED-rationalizable.

To remedy this problem, I provide a novel characterization of the exponential discounting model that will allow me to generalize the results introduced in the previous section.

Theorem 1. *The following statements are equivalent:*

(i) The data set $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ is ED-rationalizable.

(ii) There exists a locally nonsatiated, continuous, monotonic and concave instantaneous utility function $u(\cdot)$ and a discount factor $\delta \in (0, 1]$ such that for all $t \in \mathcal{T}$ and $\boldsymbol{c} \in \mathbb{R}^L_+$

$$u(\boldsymbol{c}_t) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}_t \geq u(\boldsymbol{c}) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}.$$

A quick comparison of the last condition in Theorem 1 with the standard formulation of exponential discounting highlights two major differences. First, there is no budget constraint in the latter. Second, the consumer's problem is much simpler as it only requires solving for optimal consumption bundles rather than the whole consumption stream. To interpret condition (ii), it is useful to rewrite it as

$$\boldsymbol{c}_t \in \arg\max_{\boldsymbol{c}\in\mathbb{R}^L_+} u(\boldsymbol{c}) + \delta^{-t}(y_t^d - \boldsymbol{\rho}_t'\boldsymbol{c}) \quad \forall t \in \mathcal{T}.$$

This formulation emphasizes that exponential discounting can be seen as a repeated static utility maximization problem. Letting $s^d := y_t^d - \boldsymbol{\rho}_t'\boldsymbol{c}$ denote savings and $U_t(\boldsymbol{c}, s^d) := u(\boldsymbol{c}) + \delta^{-t}s^d$, the objective function can be interpreted as a time-dependent augmented utility function $U_t : \mathbb{R}^L_+ \times \mathbb{R} \to \mathbb{R}$. It indicates that, in any given time period, the consumer values both current consumption and savings. This compromise between current consumption and savings captures the idea that increasing consumption today leaves a lesser amount of wealth for future periods, thus diminishing future consumption. In the absence of a budget constraint, the mechanism by which an interior solution is achieved therefore relies on the trade-off between the two.

4.2 Exponential Discounting under Partial Efficiency

In the revealed preference literature, it is standard to deal with deviations from a given model by slightly relaxing its constraints. Following this approach, I shall adopt the novel representation of Theorem 1 for exponential discounting rationalizability under partial efficiency.

Definition 2. Let $e \in (0, 1]$. The e-ED model rationalizes the data $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ if there exists a locally nonsatiated, continuous, monotonic and concave utility function $u(\cdot)$ and a discount factor $\delta \in (0, 1]$ such that for all $t \in \mathcal{T}$ and $\boldsymbol{c} \in \mathbb{R}^L_+$

$$u(\boldsymbol{c}_t) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}_t \geq u(\boldsymbol{c}) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}/e.$$

This definition accounts for digressions from exponential discounting by considering an efficiency level e that rationalizes every choice of a consumer at once.[10] In particular, note that any consumption behavior may be rationalized by the e-ED model for an e arbitrarily close to zero.[11] To see the economic intuition behind e, note that for a given time period t, the expression in the

[10]This choice follows the same suggestion as Afriat (1973) for static utility maximization. Alternatively, one could have an efficiency index for each choice as in Varian (1990), and then consider some aggregator function (Dziewulski, 2018) to determine the overall level of inefficiency. Interestingly, Dziewulski (2018) provides a formal link between efficiency levels and the notion of just-noticeable difference.

[11]That is, e may capture many sources of violation occurring simultaneously, as well as consumption behavior outside of the exponential discounting framework.

definition may be written as

$$\delta^t(\tilde{u}(\boldsymbol{c}_t) - \tilde{u}(\boldsymbol{c})) \geq e\boldsymbol{\rho}_t'\boldsymbol{c}_t - \boldsymbol{\rho}_t'\boldsymbol{c},$$

where the utilities have been scaled by a factor e. That is, the efficiency level ensures that the discounted benefit from consuming $\boldsymbol{c}_t$ rather than $\boldsymbol{c}$ is greater than the additional cost incurred from purchasing $\boldsymbol{c}_t$ instead of $\boldsymbol{c}$. The difference between the actual cost of acquiring $\boldsymbol{c}_t$ and what it should have been for it to be worthwhile therefore gives a measure of wasted income. Namely, for some $e \in (0, 1]$ and period $t \in \mathcal{T}$, the consumer wastes an amount equal to $\boldsymbol{\rho}_t'\boldsymbol{c}_t - e\boldsymbol{\rho}_t'\boldsymbol{c}_t$ or $(1-e)\%$ of his income by making an inefficient choice.[12] The following result extends Proposition 1 to a partial efficiency setting.

Proposition 2. *For a given $e \in (0, 1]$, the following statements are equivalent:*

(i) There exists a locally nonsatiated, continuous, monotonic and concave utility function $u(\cdot)$ and a discount factor $\delta \in (0, 1]$ e-ED rationalizing the data $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$.

(ii) There exist numbers u_t, $t = 0, \dots, T$, and a discount factor $\delta \in (0, 1]$ such that

$$u_s \leq u_t + \delta^{-t}\boldsymbol{\rho}_t'(\boldsymbol{c}_s/e - \boldsymbol{c}_t) \quad \forall s, t \in \mathcal{T}.$$

(iii) There exists a discount factor $\delta \in (0, 1]$ such that for any subset of indices $\tau = \{t_i\}_{i=1}^m$ with $t_i \in \mathcal{T}$ and $m \geq 2$,

$$0 \leq \delta^{-t_1}\boldsymbol{\rho}_{t_1}'(\boldsymbol{c}_{t_2}/e - \boldsymbol{c}_{t_1}) + \ldots + \delta^{-t_m}\boldsymbol{\rho}_{t_m}'(\boldsymbol{c}_{t_1}/e - \boldsymbol{c}_{t_m}). \qquad (\mathrm{CM}(e))$$

Proposition 2 gives a way to gauge the severity of departure from exponential discounting by finding an efficiency index $e \in (0, 1]$ e-ED rationalizing the data.[13] Conditional on $(e, \delta) \in (0, 1]^2$, the existence of a solution can be checked by solving the set of inequalities in Proposition 2 (ii) using linear programming. A data set that needs a small efficiency level to be e-ED rationalizable is farther away from exponential discounting than one with a large efficiency level. In particular, if $e = 1$ then the data set is ED-rationalizable.

[12] Since I consider a common rationalizing efficiency level for all time periods, the consumer wastes up to $(1-e)\%$ of his lifetime income.

[13] Appendix B discusses efficiency indices of interest such as the largest $e \in (0, 1]$ e-ED rationalizing the data.

Remark. By imposing $e = \delta = 1$ in Proposition 2, I recover the conditions for quasilinear utility maximization from Brown and Calsamiglia (2007). This observation makes clear that quasilinear utility maximization can be viewed as a special instance of exponential discounting. To test quasilinear utility maximization under partial efficiency, it suffices to find a solution to the inequalities in Proposition 2 conditional on $\delta = 1$.

5 Inference on the Discount Factor

In this section, I introduce measurement error in prices, present the statistical test of Varian (1985) when applied to the exponential discounting model, and propose a constrained statistical test based on e-ED rationalizability. I then show how the test can be inverted to construct a confidence set for the discount factor.

5.1 Statistical Test

Suppose prices are mismeasured such that observed prices $\boldsymbol{\rho}_t$ differ from true prices $\boldsymbol{\rho}_t^*$.[14] Specifically, suppose that

$$\boldsymbol{\rho}_t = \boldsymbol{\rho}_t^* / (1 + \boldsymbol{\epsilon}_t),$$

where $\boldsymbol{\epsilon}_t$ is assumed to be a random vector whose components follow independent normal distributions with mean zero and unknown variance $\sigma^2 > 0$.[15] That is, $\epsilon_{l,t} \sim N(0, \sigma^2)$ for all $l \in \mathcal{L}$ and all $t \in \mathcal{T}$. It is useful to note that, under this assumption, the test statistic

$$T(\sigma^2) := \sum_{t=0}^{T} \sum_{l=1}^{L} (\rho_{l,t}^* / \rho_{l,t} - 1)^2 / \sigma^2 \tag{1}$$

follows a chi-square distribution. Since true prices are unobservable, the idea consists of obtaining a lower bound on $T(\sigma^2)$ by considering the following quadratic programming problem:

$$S(\sigma^2, \delta^*) := \min_{\boldsymbol{\pi} \in \mathbb{R}_{++}^{L \times |\mathcal{T}|}} \sum_{t=0}^{T} \sum_{l=1}^{L} (\pi_{l,t} / \rho_{l,t} - 1)^2 / \sigma^2, \tag{2}$$

[14] In my empirical application, I use scanner data on food expenditures in which measurement error in prices prevails.

[15] Alternatively, one can assume an additive error: $\boldsymbol{\rho}_t = \boldsymbol{\rho}_t^* + \boldsymbol{\epsilon}_t$. I consider proportional measurement error as the scale of discounted prices changes significantly across time.

subject to

$$u_s - u_t \leq \delta^{*-t}\boldsymbol{\pi}_t'(\boldsymbol{c}_s - \boldsymbol{c}_t) \quad \forall s,t \in \mathcal{T},$$

where $(u_t)_{t\in\mathcal{T}}$ are real numbers and δ^* is the true discount factor. If the data set is consistent with exponential discounting under true prices, then one can always pick $\boldsymbol{\pi}_t = \boldsymbol{\rho}_t^*$ for all $t \in \mathcal{T}$. Accordingly, $S(\sigma^2, \delta^*) \leq T(\sigma^2)$ such that

$$\mathbb{P}\left[S(\sigma^2, \delta^*) \leq \chi^2_\alpha\right] \geq \mathbb{P}\left[T(\sigma^2) \leq \chi^2_\alpha\right] = 1 - \alpha,$$

where χ^2_α is the critical value of the chi-square distribution for some prespecified confidence level $\alpha \in (0,1)$.

5.2 Constrained Statistical Test

The main disadvantage of the previous test is that prices that solve the problem (2) may be much closer to observed prices than actual true prices. That is, the test is extremely conservative. To alleviate this limitation, I propose to further restrict the set of allowable true prices to those satisfying e-ED, therefore yielding the following constraints:

$$u_s - u_t \leq \delta^{*-t}\boldsymbol{\pi}_t'(\boldsymbol{c}_s - \boldsymbol{c}_t) \leq \delta^{*-t}\boldsymbol{\rho}_t'(\boldsymbol{c}_s/e - \boldsymbol{c}_t) \quad \forall s,t \in \mathcal{T}, \tag{3}$$

where $e \in (0,1]$, $(u_t)_{t\in\mathcal{T}}$ are real numbers, and $(\boldsymbol{\pi}_t)_{t\in\mathcal{T}}$ are candidate true prices.

The first inequality in (3) requires the solution $(\boldsymbol{\pi}_t)_{t\in\mathcal{T}}$ to be consistent with exponential discounting. The second inequality further ensures that the solution is not an overly distorted version of observed prices as measured by e-ED. The latter can be viewed as requiring that measurement error be small in the sense that the econometrician is not led to believe that the consumer wastes more than $1-e\%$ of his income. The resulting constrained optimization problem is given by

$$S^C(\sigma^2, \delta^*, e) := \min_{\boldsymbol{\pi}\in\mathbb{R}^{L\times|\mathcal{T}|}_{++}} \sum_{t=0}^{T}\sum_{l=1}^{L}(\pi_{l,t}/\rho_{l,t} - 1)^2/\sigma^2, \tag{4}$$

subject to

$$u_s - u_t \leq \delta^{*-t}\boldsymbol{\pi}_t'(\boldsymbol{c}_s - \boldsymbol{c}_t) \leq \delta^{*-t}\boldsymbol{\rho}_t'(\boldsymbol{c}_s/e - \boldsymbol{c}_t) \quad \forall s,t \in \mathcal{T},$$

where $(u_t)_{t\in\mathcal{T}}$ and $(\boldsymbol{\pi}_t)_{t\in\mathcal{T}}$ may take different values than those that solve the optimization problem (2).

It is useful to note that $S^C(\sigma^2, \delta^*, e) \geq S(\sigma^2, \delta^*)$ such that $S^C(\sigma^2, \delta^*, e)$ is a less conservative test statistic. Nevertheless, the restriction that the solution of the constrained optimization problem satisfies e-ED becomes immaterial when $e \in (0, 1]$ approaches zero. Thus, the optimization problem (2) is a special case of the constrained optimization problem (4).

5.3 Inference

Let $\sigma^2 > 0$ and $e \in (0, 1]$ be fixed numbers. Since the true discount factor is unknown, the key to get a feasible test statistic is to set $\delta \in (0, 1]$ and solve

$$V(\delta, e) := \min_{\boldsymbol{\pi} \in \mathbb{R}_{++}^{L \times |\mathcal{T}|}} \sum_{t=0}^{T} \sum_{l=1}^{L} (\pi_{l,t}/\rho_{l,t} - 1)^2, \tag{5}$$

subject to

$$u_s - u_t \leq \delta^{-t} \boldsymbol{\pi}_t'(\boldsymbol{c}_s - \boldsymbol{c}_t) \leq \delta^{-t} \boldsymbol{\rho}_t'(\boldsymbol{c}_s/e - \boldsymbol{c}_t) \quad \forall s, t \in \mathcal{T}.$$

Defining $S^C(\sigma^2, \delta, e) := V(\delta, e)/\sigma^2$, the constrained confidence set is obtained by inverting the constrained test statistic:

$$CS^C := \left\{\delta \in (0, 1] : S^C(\sigma^2, \delta, e) \leq \chi^2_\alpha\right\}.$$

Likewise, the confidence set associated with the unconstrained test statistic is defined by

$$CS := \left\{\delta \in (0, 1] : S(\sigma^2, \delta) \leq \chi^2_\alpha\right\}.$$

Noting that $\lim_{e \to 0} S^C(\sigma^2, \delta, e) = S(\sigma^2, \delta)$, we have

$$\mathbb{P}\left[\lim_{e \to 0} \delta^* \in CS^C\right] = \mathbb{P}\left[\delta^* \in CS\right] = \mathbb{P}\left[S(\sigma^2, \delta^*) \leq \chi^2_\alpha\right] \geq 1 - \alpha.$$

In words, the constrained test statistic converges to the unconstrained test statistic when $e \in (0, 1]$ approaches zero. In that case, the probability that the constrained confidence set covers the true discount factor becomes at least $1 - \alpha$. Conditional on (σ^2, e), the constrained confidence set can be recovered by solving (5) for each $\delta \in (0, 1]$. The choice of these variables should be chosen from prior knowledge of the data set.

6 Empirical Application

6.1 Data

In my empirical analysis, I implement the methodology developed in the previous sections using the Stanford Basket Dataset, which is a panel data set containing expenditures of 494 households between June 1991 and June 1993.[16] Specifically, I use the transformed data set of Echenique, Lee and Shum (2011). As such, goods for which prices are observed in every week are retained and aggregated by brand for periods of four weeks. This yields a total of 375 distinct goods belonging to one of the following 14 categories: bacon, barbecue sauce, butter, cereal, coffee, crackers, eggs, ice cream, nuts, analgesics, pizza, snacks, sugar and yogurt.

Since none of these categories contain goods purchased for special events (e.g., turkey for Thanksgiving) or products whose quality may change with seasons (e.g., fruits), I expect preferences to be roughly stable over the time window considered. Additionally, due to the focus on food items, I do not expect consumers' purchases to vary considerably in response to changes in income. Finally, aggregation to monthly expenditure should mitigate stockpiling associated with sales.

The data set is prone to measurement error since it contains shelf prices instead of transaction prices. Thus, observed prices differ from actual prices paid whenever a consumer uses discounts such as coupons. As the data do not contain information on interest rates, I include interest rates on personal loans at commercial banks from the Federal Reserve Bank of St. Louis.[17] I report demographic information about households in the data set in Table 1. For a comprehensive description of the scanner data set, I refer the reader to Echenique, Lee and Shum (2011).

Table 1: Demographic Variables

[16] I treat households as unitary entities even though they may have many members. As such, I refer to a household as a consumer or an individual.

[17] Since the data on interest rates are quarterly, I use a linear interpolation to obtain monthly observations.

Variable	Number of Households
Family size:	
Midsize (3,4 members)	187
Large size (> 4 members)	65
Income:	
Mid annual income ($\in$ [\$20k, \$45k])	200
High annual income (> \$45k)	141
Age:[a]	
Middle-aged	201
Old-aged	157
Education:[b]	
High school	197
College	255
Total households	480

[a] Middle-aged households are those in which the average of the spouses' ages is between 30 and 65; in old-aged households, this average exceeds 65.
[b] If both spouses are present in a household, the average education of both spouses is reported.

6.2 Specification

In what follows, I restrict the range of the monthly discount factor to $[0.75, 1.0]$ and use a step size of 0.01. This support restriction is essentially without loss of generality as the resulting support of the annualized discount factor becomes approximately $[0.02, 1.0]$.[18] For ease of comparison, I report the confidence sets for the annualized discount factor.

I also restrict the sample to consumers that do not appear to waste more than 15% of their expenditures at the observed data. That is, consumers that are e-ED rationalizable for $e \geq 0.85$. This choice is motivated by the fact that the main source of measurement error in the data is from coupons and the broader empirical evidence suggesting that almost no consumer saves more than 15% of their expenditures from sales such as price promotions (Griffith et al., 2009). For the same reason, I set the standard deviation to a conservative $\sigma = 0.15$.

Lastly, I set the significance level to $\alpha = 0.05$. Thus, a consumer is said to be consistent with the exponential discounting model if there exists a monthly discount factor $\delta \in [0.75, 1.0]$ and an efficiency level $e \in [0.85, 1.0]$ such that $S^C(\sigma^2, \delta, e) \leq \chi^2_{0.05}$, where $S^C(\sigma^2, \delta, e)$ is obtained by solving the feasible constrained optimization problem (5).[19]

[18] To obtain annualized rates, I raise the monthly discount factor to the power 13. The reason being that data are aggregated to 4-week periods, hence yielding 13 time periods in a year.

[19] Let R be the number of goods that are never purchased by a consumer. Since changing the price of a good never purchased has no effect on the constraint in (3), the number of degrees

6.3 Inference with Measurement Error and Individual Heterogeneity

In this subsection, I implement my methodology in the data according to the previous specification. I find that 144 out of the 494 consumers have data sets consistent with exponential discounting. As such, the following analysis focuses on those consumers exclusively.

In Figure 1, I show how the average constrained confidence set evolves when the restrictions on measurement error decrease from $e = 0.85$ to $e = 0$.[20] The average constrained confidence set is obtained by averaging over consumers constrained confidence sets. In particular, the average constrained confidence set at $e = 0$ corresponds to the average unconstrained confidence set obtained by applying Varian's (1985) method.

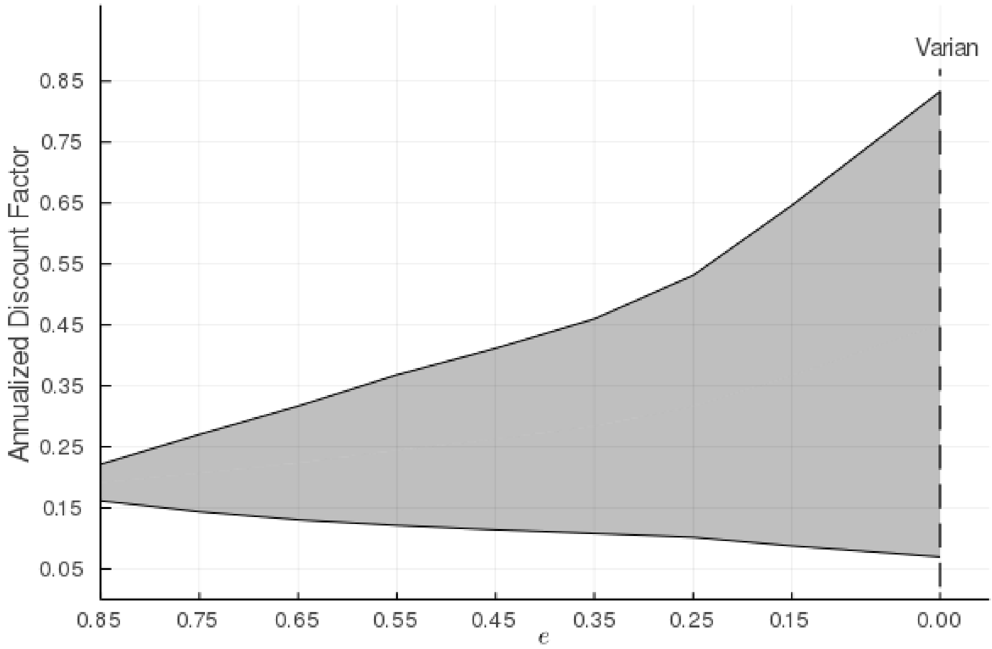

Figure 1: Average Constrained Confidence Set by Efficiency Level.

Figure 1 shows that reasonable restrictions on measurement error can significantly decrease the size of the confidence set. For Varian's (1985) method to be preferable, one would have to believe that measurement error causes the consumer to waste up to 100% of his expenditure at the observed data. By choosing $e > 0$, one recognizes that the observed data carry information that can be used for identification.

of freedom is equal to $|\mathcal{T}|(L - R)$.

[20] The step size is 0.1 from $e = 0.85$ to $e = 0.15$.

Next, Figure 2 displays how the average constrained confidence set changes by demographic. The efficiency level e is set to 0.85 such that measurement error is allowed to cause the consumer to waste up to 15% of his expenditure at the observed data. A summary of the demographic variables is given in Table 1.

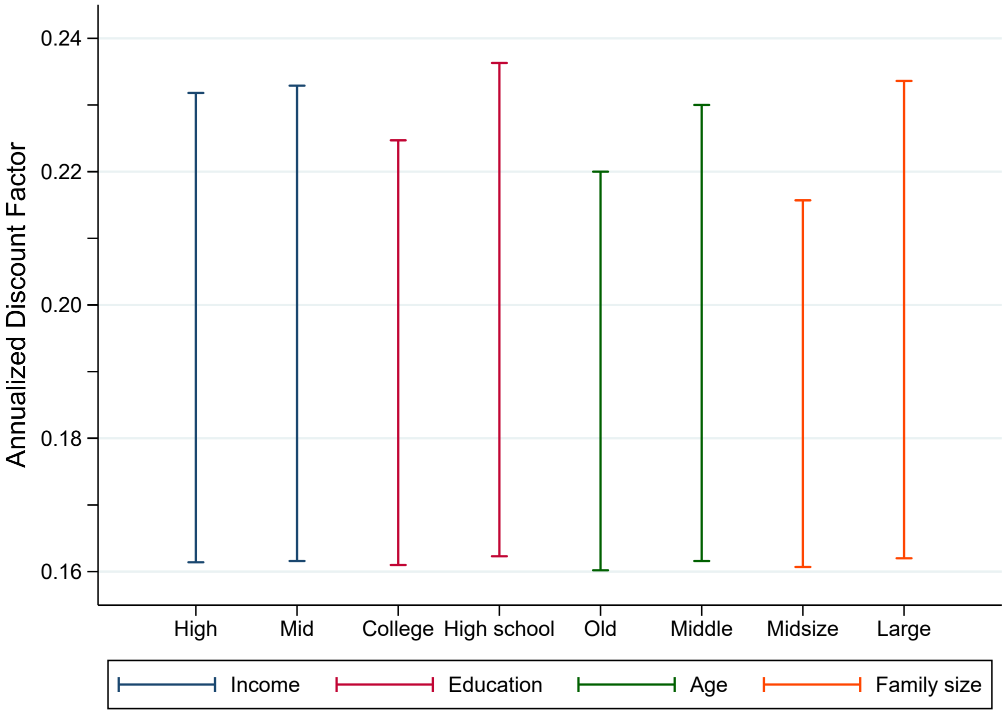

Figure 2: Average Constrained Confidence Set by Demographic.

Overall, Figure 2 shows that the relationship between the discount factor and demographics is relatively weak in the sample. Figure 2 suggests that households with high school education have slightly larger discount factors compared to those with college education. Furthermore, it suggests that middle-aged households and large size households have slightly larger discount factors compared to old-aged households and midsize households, respectively. However, income does not appear to have any significant impact on the discount factor.

Although Figure 2 suggests that discount factors are mostly homogeneous, there may be heterogeneity that is not captured by observable characteristics. Accordingly, I compare the constrained confidence set at various quantiles of the sample in Figure 3, where consumers were ordered by the midpoint of their constrained confidence set. Contrary to the previous analysis, Figure 3 reveals a fair degree of heterogeneity once individual unobserved heterogeneity is fully acknowledged.

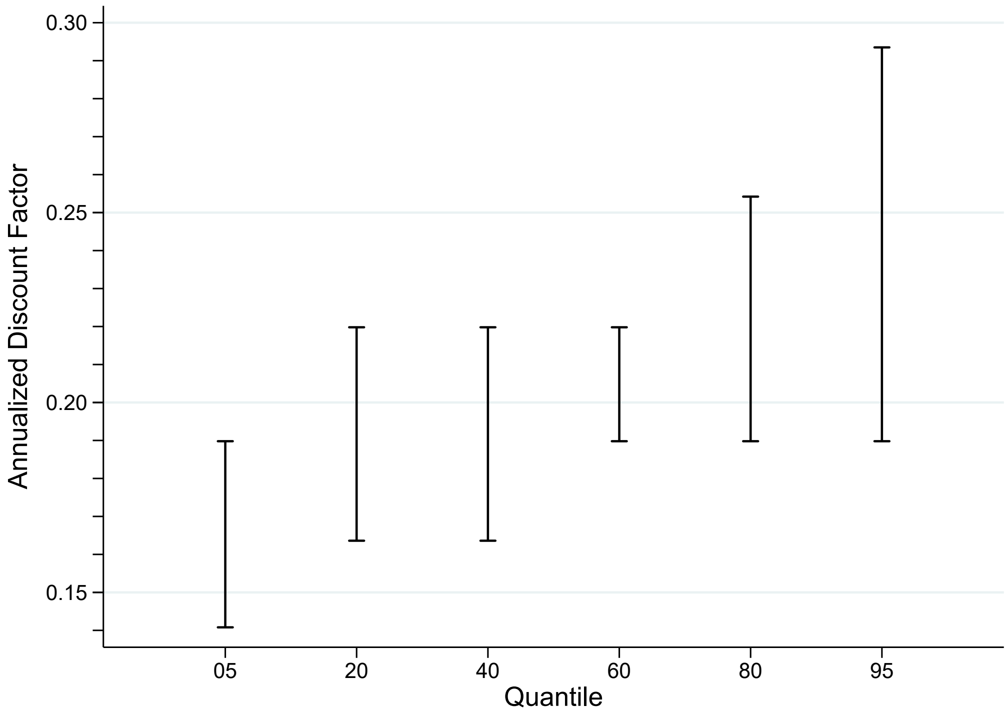

Figure 3: Constrained Confidence Set by Quantile.

6.4 Discussion

The annualized discount factors displayed in my application are well below the values usually assumed in the literature. However, the average constrained confidence set is compatible with other studies using analogous data.[21] For example, Ackerberg (2003) estimates a weekly discount factor of 0.98 with scanner data on yogurt, therefore giving an annualized discount factor of 0.35.[22]

My empirical results show that the discount factor is only weakly correlated with observable characteristics. At the same time, I find sizable heterogeneity in discount factors once unobserved heterogeneity is accounted for. This finding suggests that inference based on observable characteristics alone may downplay the amount of heterogeneity in the data. The presence of heterogeneity implies that welfare analysis and counterfactuals that make homogeneity assumptions may poorly apply to a meaningful fraction of consumers.

[21] My results are not directly comparable to those obtained using survey data such as in Blow, Browning and Crawford (2017) as the type of data differs.

[22] See Yao et al. (2012) for further estimates from field data. More generally, see Frederick et al. (2002) for a comprehensive review of the literature.

7 Conclusion

My results allow one to set identify individual discount factors while avoiding the misspecification of preferences. Inference can be made whether or not a data set contains exact information about the variance in measurement error. Once the discount factor is elicited, one could use the revealed preference inequalities to bound a consumer's response to changes in prices. That is, one could undertake a counterfactual analysis in a similar fashion as Blundell, Browning and Crawford (2003, 2008). My methodology could therefore be used to do robust welfare analysis, estimate market power in empirical industrial organization, or determine optimal pricing schemes in marketing.

In this paper, I recognize that deviations from exponential discounting may naturally arise due to imperfect measurement.[23] An interesting extension would be to consider a random utility version of exponential discounting to account for changes in preferences. Towards this goal, Appendix C provides an axiom for exponential discounting in the paradigm of revealed price preference of Deb et al. (2018) that appears compatible with the statistical framework laid out by Kitamura and Stoye (2018). I leave to future work the analysis of a random utility model of exponential discounting.

Appendix A Elementary Revealed Preference Theory

This section presents the revealed preference terminology and reviews an extension of the static utility maximization model permitting for violations from optimal behavior.[24]

For $e \in (0,1]$, a consumption bundle $\boldsymbol{c}_t$ is said to be *directly revealed preferred* to a bundle $\boldsymbol{c}_s$ if and only if $\boldsymbol{\rho}_t'(\boldsymbol{c}_s/e - \boldsymbol{c}_t) \leq 0$, where e is designed to remove revealed preference information generating cyclic preferences. Let $R^D(e)$ denote the direct revealed preference relation and let $R(e)$ denote its transitive closure.[25] When the inequality is strict, $\boldsymbol{c}_t$ is said to be *directly revealed strictly preferred* to $\boldsymbol{c}_s$ and is denoted $P^D(e)$. In the case where there is a sequence $\boldsymbol{c}_t R^D(e) \boldsymbol{c}_{t_1}, \boldsymbol{c}_{t_1} R^D(e) \boldsymbol{c}_{t_2}, \ldots, \boldsymbol{c}_{t_m} R^D(e) \boldsymbol{c}_s$ of directly revealed preferences, where

[23] A general framework to tackle measurement error in utility maximization models is provided by Aguiar and Kashaev (2020).

[24] As noted by Blow, Browning and Crawford (2017), the fact that discounting prices does not change relative prices implies that static rationalizability is the same with either spot prices $(\boldsymbol{p}_t)_{t\in\mathcal{T}}$ or discounted prices $(\boldsymbol{\rho}_t)_{t\in\mathcal{T}}$. I choose to define static utility maximization with discounted prices for notational consistency.

[25] The transitive closure $R(e)$ of a relation $R^D(e)$ is the smallest relation containing $R^D(e)$ satisfying transitivity.

$t, t_1, \ldots, t_m, s \in \mathcal{T}$, $\boldsymbol{c}_t$ is said to be *revealed preferred* to $\boldsymbol{c}_s$. Naturally, if any of those preference relations is strict, then $\boldsymbol{c}_t$ is said to be *revealed strictly preferred* to $\boldsymbol{c}_s$. The preceding notation allows me to succinctly define two important concepts.

Definition 3. Let $e \in (0, 1]$. A locally nonsatiated utility function $u(\cdot)$ e-rationalizes the data $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ if for every observed bundle $\boldsymbol{c}_t \in \mathbb{R}^L_+$, $\boldsymbol{c}_t R^D(e)\boldsymbol{c}$ implies $u(\boldsymbol{c}_t) \geq u(\boldsymbol{c})$ and $\boldsymbol{c}_t P^D(e)\boldsymbol{c}$ implies $u(\boldsymbol{c}_t) > u(\boldsymbol{c})$.

Definition 4. Let $e \in (0, 1]$. A data set $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ satisfies the Generalized Axiom of Revealed Preference (GARP(e)) if for all $s, t \in \mathcal{T}$, $\boldsymbol{c}_t R(e)\boldsymbol{c}_s$ implies not $\boldsymbol{c}_s P^D(e)\boldsymbol{c}_t$.

The generalized axiom gives an intuitive necessary condition for rationalizability by requiring the consumer to have transitive preferences. In particular, note that GARP(e) is a natural generalization of GARP(1) and simply eliminates cycles by reducing the number of revealed preferences. The following result from Halevy, Persitz and Zrill (2018) and Heufer and Hjertstrand (2017) extends the influential theorem of Afriat (1967) to consumers violating the model of atemporal utility maximization.[26,27]

e-Afriat's Theorem. *For a given $e \in (0, 1]$, the following statements are equivalent:*

(1) There exists a locally nonsatiated utility function e-rationalizing the data $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$.

(2) The data $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ satisfy GARP(e).

(3) There exist numbers u_t, $\lambda_t > 0$, $t = 0, \ldots, T$, such that

$$u_s \leq u_t + \lambda_t \boldsymbol{\rho}_t'(\boldsymbol{c}_s/e - \boldsymbol{c}_t) \quad \forall s, t \in \mathcal{T}.$$

(4) There exists a locally nonsatiated, continuous, monotonic and concave utility function e-rationalizing the data $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$.

For practical purposes, the second condition is the most convenient. As highlighted by Varian (1982), one can use an efficient algorithm called Warshall's

[26] For the case where $e = 1$, an accessible proof is given by Fostel, Scarf, and Todd (2004). An alternative and insightful proof is offered by Geanakoplos (2013).

[27] This theorem, as well as the proposition to follow, could all be written using the efficiency measure of Varian (1990). For my purposes, it is sufficient to consider a common index for all observations.

algorithm to get the transitive closure $R(e)$ of the direct revealed preference relation $R^D(e)$. For a given $e \in (0,1]$, one can then directly check for a contradiction of GARP(e) in the data. Alternatively, conditional on $e \in (0,1]$, one can use linear programming to solve the system of inequalities given in the third condition. The goal is then to verify the existence of a pair $(u_t, \lambda_t)_{t\in\mathcal{T}}$ satisfying it.

From a theoretical standpoint, however, the substance of e-Afriat's Theorem lies in the last condition. It implies that, if the consumer's choices can be thought of as generated by a locally nonsatiated utility function, then it can further be assumed to be continuous, monotonic and concave.[28,29]

Appendix B Efficiency Indices

This section presents the exponential efficiency index, proposes an efficiency index for time-consistency, and investigates which assumption of the exponential discounting model is the most problematic in the data.

B.1 Exponential Efficiency Index

For the model of static utility maximization under partial efficiency, it is common to consider the largest efficiency level rationalizing the data. This index is known as the CCEI and was suggested by Afriat (1973).[30] In a similar fashion, one can consider the largest efficiency level rationalizing the data for the exponential discounting model. Formally, I define the exponential efficiency index as

$$\text{EEI} := \sup\{e \in [0,1] : \{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}} \text{ is } e\text{-ED rationalizable}\}.$$

From the previous analysis, it is clear that the EEI can be interpreted as the smallest proportion of wasted income arising from the selection of a suboptimal consumption stream. Moreover, note that the exponential efficiency index is well-defined as the inequalities in Proposition 2 (ii) will be trivially satisfied for an e arbitrarily close to zero.

[28] A mapping $f : \mathbb{R}^L \to \mathbb{R}^L$ is said to be concave if and only if $f(\boldsymbol{c}_s) \leq f(\boldsymbol{c}_t) + \nabla f(\boldsymbol{c}_t)'(\boldsymbol{c}_s - \boldsymbol{c}_t)$ for all $s, t \in \mathcal{T}$.

[29] Concavity can only be assumed without loss of generality in finite data. In the case of infinite data, it has to be substituted by the weaker assumption of quasiconcavity. I refer the reader to Reny (2015) for a more detailed discussion.

[30] The critical cost efficiency index is defined as CCEI $:= \sup\{e \in [0,1] : \{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ satisfies GARP(e)$\}$.

Although the EEI provides a measure of distance between a data set and the exponential discounting model, it does not differentiate between deviations arising from within-period consistency and time consistency. To disentangle their respective contributions to the EEI, an efficiency measure that controls for violations of static utility maximization is needed. For convenience, denote such an index the time consistency efficiency index (TCEI). In what follows, I derive the TCEI based on the 2-step rationalization procedure of Heufer and Hjertstrand (2017) for homothetic rationalizability.

The first stage consists in finding the largest efficiency level rationalizing the data with respect to static utility maximization and yields

$$u_s \leq u_t + \lambda_t \boldsymbol{\rho}_t'(\boldsymbol{c}_s/\text{CCEI} - \boldsymbol{c}_t) \quad \forall s,t \in \mathcal{T}.$$

Imposing the additional restriction of the exponential discounting model to the CCEI-Afriat inequalities by setting $\lambda_t = \delta^{-t}$ yields

$$u_s \leq u_t + \delta^{-t} \boldsymbol{\rho}_t'(\boldsymbol{c}_s/\text{CCEI} - \boldsymbol{c}_t) \quad \forall s,t \in \mathcal{T}.$$

The TCEI then corresponds to the largest efficiency level rationalizing the previous system of inequalities with respect to the e-ED model:

$$u_s \leq u_t + \delta^{-t} \boldsymbol{\rho}_t' \left(\frac{\boldsymbol{c}_s}{\text{CCEI} \cdot \text{TCEI}} - \boldsymbol{c}_t \right) \quad \forall s,t \in \mathcal{T}.$$

That is, the TCEI gives the additional adjustment required to the CCEI-adjusted data set to satisfy the e-ED model. Since the largest efficiency level solving the e-ED model is the EEI, it follows that EEI $=$ CCEI $\cdot$ TCEI. One can therefore recover the TCEI by first obtaining the CCEI and the EEI. Moreover, taking the natural logarithm of the previous expression yields the following relationship.

Identity 1. Let EEI < 1, then $\frac{\log(\text{CCEI})}{\log(\text{EEI})} + \frac{\log(\text{TCEI})}{\log(\text{EEI})} = 1$.

This identity allows one to obtain the respective contribution of static utility maximization and time consistency to the exponential efficiency index.

Definition 5. The contribution of the CCEI to the EEI and of the TCEI to the EEI are respectively given by

$$\mathcal{C}_g := \frac{\log(\text{CCEI})}{\log(\text{EEI})} \quad \text{and} \quad \mathcal{C}_t := \frac{\log(\text{TCEI})}{\log(\text{EEI})}.$$

In particular, note that the contribution of each index is always between zero and one, strictly increases as its efficiency index decreases, and that the combined contribution of each index must always sum up to one. Finally, note that the sum of $\mathcal{C}_g$ and $\mathcal{C}_t$ gives the contribution of the EEI to itself.

B.2 Empirical Application

In what follows, I am interested in the efficiency indices for static utility maximization, time consistency and exponential discounting, as well as their respective contributions to the EEI. Let μ_i denote the mean and σ_i the standard deviation, where $i \in \{e, \mathcal{C}\}$ refers to the object over which the operation is applied. Any object $i \in \{e, \mathcal{C}\}$ underlined or overlined represents its smallest and largest value across consumers, respectively.

Using the CCEI for static utility maximization, the TCEI for time consistency and the EEI for exponential discounting, Table 2 presents summary statistics on the efficiency indices and contributions of each index. These results are obtained with a grid search over $\delta \in (0, 1]$ with a step size of 0.01 and a binary search algorithm for the efficiency indices that guarantees them to be within 2^{-10} of their true values.

Table 2: Rationalizability Results

Efficiency index	$\underline{e}$	$\overline{e}$	μ_e	σ_e	$\underline{\mathcal{C}}$	$\overline{\mathcal{C}}$	$\mu_{\mathcal{C}}$	$\sigma_{\mathcal{C}}$
CCEI	0.6865	1.0000	0.9551	0.0502	0.0000	1.0000	0.2057	0.2017
TCEI	0.4758	1.0000	0.8365	0.0802	0.0000	1.0000	0.7943	0.2017
EEI	0.3878	0.9561	0.7984	0.0820	1.0000	1.0000	1.0000	0.0000

Notes: The sample size is $N = 494$. $\underline{e}$ denotes the lowest efficiency index, $\overline{e}$ the largest efficiency index, μ_e the average efficiency index, and σ_e the standard deviation of the efficiency index. $\underline{\mathcal{C}}$ denotes the lowest contribution of the efficiency index to the EEI, $\overline{\mathcal{C}}$ the largest contribution of the efficiency index to the EEI, $\mu_{\mathcal{C}}$ the average contribution of the efficiency index to the EEI, and $\sigma_{\mathcal{C}}$ the standard deviation of the efficiency index's contribution.

Overall, the results in Table 2 indicate that time consistency is a more stringent assumption than GARP, with an average efficiency level for the TCEI below that of the CCEI by approximately 0.10. The significance of this difference is better grasped by looking at the average contribution of each index to the EEI. Markedly, on average, GARP is responsible for about 20% of a violation from exponential discounting, while 80% of it can be attributed to time consistency.

Appendix C Extension: Theory of Revealed Price Preference

The key insight of Theorem 1 was to provide a time-dependent augmented utility representation of the exponential discounting model. The simplified formulation then made possible to analyze the empirical implications of the exponential discounting model under partial efficiency. While I focused on the implications of the model in the standard framework of revealed preference, I now derive the implications of the exponential discounting model under partial efficiency according to the notion of revealed price preference developed by Deb et al. (2018).[31]

First, define a time-dependent augmented utility function as a function $U_t : \mathbb{R}^L_+ \times \mathbb{R}_- \to \mathbb{R}$ that is assumed strictly increasing in its second argument. The interpretation is that the consumer dislikes current expenditure as it removes income that could be used for later consumption. The consumer's problem can be summarized as picking $\boldsymbol{c}_t$ such that for all $t \in \mathcal{T}$ and $\boldsymbol{c} \in \mathbb{R}^L_+$

$$U_t(\boldsymbol{c}_t, -\boldsymbol{c}_t'\delta^{-t}\boldsymbol{\rho}_t) \geq U_t(\boldsymbol{c}, -\boldsymbol{c}'\delta^{-t}\boldsymbol{\rho}_t/e).$$ [32]

Intuitively, the utility cost of current expenditure becomes more salient as the consumer gets more patient ($\delta \to 1$). Conveniently, the terminology of Section A can be applied to the notion of revealed price preference. The difference is that instead of comparing consumption bundles for a given set of prices, we now compare prices for a given consumption bundle. Namely, prices $\delta^{-t}\boldsymbol{\rho}_t/e$ are said to be *directly revealed preferred* to prices $\delta^{-s}\boldsymbol{\rho}_s$ if and only if $\boldsymbol{c}_s'(\delta^{-t}\boldsymbol{\rho}_t/e - \delta^{-s}\boldsymbol{\rho}_s) \leq 0$, and likewise for other revealed price preference relations. The rationale is that at prices $\delta^{-t}\boldsymbol{\rho}_t/e$ the consumer can buy the bundle $\boldsymbol{c}_s$ while having some money left to spend in the future. Exponential discounting rationalizability and the corresponding Genereralized Axiom of Price Preference (δ-GAPP(e)) can then be defined as follows.

Definition 6. Let $e \in (0, 1]$. The e-ED model rationalizes the data $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ if there exists a time-dependent augmented utility function $U_t(\cdot, \cdot)$ and a discount factor $\delta \in (0, 1]$ such that for all $t \in \mathcal{T}$ and $\boldsymbol{c} \in \mathbb{R}^L_+$

$$U_t(\boldsymbol{c}_t, -\boldsymbol{c}_t'\delta^{-t}\boldsymbol{\rho}_t) \geq U_t(\boldsymbol{c}, -\boldsymbol{c}'\delta^{-t}\boldsymbol{\rho}_t/e).$$

Definition 7. Let $e \in (0, 1]$. A data set $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ satisfies δ-GAPP(e) if for

[31] This section introduces the notions of rationalizability and revealed price preference for the exponential discounting model under partial efficiency. The definitions as found in Deb et al. (2018) correspond to the case where $e = \delta = 1$.

[32] This definition slightly generalizes the one presented in Section 3 as the second argument is not presumed additively separable.

all $s,t \in \mathcal{T}$, there exists a discount factor $\delta \in (0,1]$ such that $\delta^{-t}\boldsymbol{\rho}_t R(e)\delta^{-s}\boldsymbol{\rho}_s$ implies not $\delta^{-s}\boldsymbol{\rho}_s P^D(e)\delta^{-t}\boldsymbol{\rho}_t$.

The next theorem derives the empirical implications of the exponential discounting model in the revealed price preference framework.

Theorem 2. *Let $e \in (0,1]$. For a given data set $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$, the following are equivalent:*

(i) The data are e-ED rationalized by a time-dependent augmented utility function.

(ii) There exists $\delta \in (0,1]$ such that the data satisfy δ-GAPP(e).

(iii) The data are e-ED rationalized by a time-dependent augmented utility function that is continuous, strictly increasing, concave, and such that a maximum exists for all $\boldsymbol{\rho} \in \mathbb{R}^L_{++}$.

Theorem 2 shows that δ-GAPP(e) is both necessary and sufficient for e-ED rationalizability in the revealed price preference paradigm. Thus, conditional on a specific value of the discount factor, the methodology of Kitamura and Stoye (2018) and Deb et al. (2018) should apply to δ-GAPP(e) and, as such, exponential discounting. In fact, the axiom tested in Deb et al. (2018) is the "quasilinear" version of the one presented in this section, 1-GAPP(1). Further note that when $U_t(\boldsymbol{c}_t, -\boldsymbol{c}_t'\delta^{-t}\boldsymbol{\rho}_t) := u(\boldsymbol{c}_t) - \boldsymbol{c}_t'\delta^{-t}\boldsymbol{\rho}_t$ with $u(\cdot)$ a locally nonsatiated function, the notion of e-ED rationalizability coincides with that of Definition 4. Therefore, δ-GAPP(e) can be seen as a generalization of the exponential discounting model to a nonseparable time-dependent augmented utility function.

Appendix D Preliminaries

Let $e, \delta \in (0,1]$. Let each entry (s,t) of a square matrix A be given by $a_{s,t} := \boldsymbol{c}_t'(\delta^{-s}\boldsymbol{\rho}_s/e - \delta^{-t}\boldsymbol{\rho}_t)$ for all $s,t \in \mathcal{T}$.

Definition 8. A square matrix A of dimension $T+1$ is cyclically consistent if for every chain $\{t_1, t_2, \ldots, t_m\} \subset \{0, 1, \ldots, T\}$, $a_{t_1,t_2} \leq 0, a_{t_2,t_3} \leq 0, \ldots, a_{t_{m-1},t_m} \leq 0$, $a_{t_m,t_1} \leq 0$ implies that all terms are zero.

Lemma 1. *A data set $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ satisfies δ-GAPP(e) if and only if the square matrix A is cyclically consistent.*

Proof

Suppose A is cyclically consistent and $\delta^{-t_1}\boldsymbol{\rho}_{t_1}R(e)\delta^{-t_m}\boldsymbol{\rho}_{t_m}$. This means that there is a sequence of revealed price preferences such that $\delta^{-t_1}\boldsymbol{\rho}_{t_1}R^D(e)\delta^{-t_2}\boldsymbol{\rho}_{t_2}$, $\delta^{-t_2}\boldsymbol{\rho}_{t_2}R^D(e)\delta^{-t_3}\boldsymbol{\rho}_{t_3}, \ldots, \delta^{-t_{m-1}}\boldsymbol{\rho}_{t_{m-1}}R^D(e)\delta^{-t_m}\boldsymbol{\rho}_{t_m}$. By definition, these imply $a_{t_1,t_2} \leq 0$, $a_{t_2,t_3} \leq 0$, ..., $a_{t_{m-1},t_m} \leq 0$. Note that if $\delta^{-t_m}\boldsymbol{\rho}_{t_m}P^D(e)\delta^{-t_1}\boldsymbol{\rho}_{t_1}$, then $a_{t_m,t_1} < 0$. Cyclical consistency then requires that $a_{t_1,t_2} = a_{t_2,t_3} = \cdots = a_{t_m,t_1} = 0$. However, this contradicts the assumption that $a_{t_m,t_1} < 0$. As such, we can't have $\delta^{-t_m}\boldsymbol{\rho}_{t_m}P^D(e)\delta^{-t_1}\boldsymbol{\rho}_{t_1}$, i.e. δ-GAPP(e) holds.

Suppose now that δ-GAPP(e) is satisfied. Construct the matrix A of revealed price preferences and note that $a_{t,t} \geq 0$ for all $t \in \mathcal{T}$. Consider any chain $\{t_1, t_2, \ldots, t_m\} \subset \{0, 1, \ldots, T\}$ such that $a_{t_1,t_2} \leq 0$, $a_{t_2,t_3} \leq 0$, ..., $a_{t_{m-1},t_m} \leq 0$, $a_{t_m,t_1} \leq 0$. For any element $a_{s,t}$ pertaining to that chain, we have $\delta^{-s}\boldsymbol{\rho}_s R^D(e)\delta^{-t}\boldsymbol{\rho}_t$. Moreover, by going along the chain we also obtain $\delta^{-t}\boldsymbol{\rho}_t R(e)\delta^{-s}\boldsymbol{\rho}_s$. Since δ-GAPP(e) requires to not have $\delta^{-s}\boldsymbol{\rho}_s P^D(e)\delta^{-t}\boldsymbol{\rho}_t$, it must be that $a_{s,t} = 0$.

Lemma 2. *CM(e) implies GARP(e).*

Proof

I proceed by contraposition. Fix $e \in (0, 1]$ and suppose GARP(e) is violated. Then, for some indices $t_1, t_m \in \mathcal{T}$, $\boldsymbol{c}_{t_1}R(e)\boldsymbol{c}_{t_m}$ and $\boldsymbol{c}_{t_m}P^D(e)\boldsymbol{c}_{t_1}$. Thus, there is a sequence of revealed preferences such that $\boldsymbol{c}_{t_1}R^D(e)\boldsymbol{c}_{t_2}$, $\boldsymbol{c}_{t_2}R^D(e)\boldsymbol{c}_{t_3}$, ..., $\boldsymbol{c}_{t_{m-1}}R^D(e)\boldsymbol{c}_{t_m}$, where $t_1, t_2, \ldots, t_m \in \mathcal{T}$. By definition, the above implies $\boldsymbol{\rho}'_{t_1}(\boldsymbol{c}_{t_2}/e - \boldsymbol{c}_{t_1}) \leq 0$, $\boldsymbol{\rho}'_{t_2}(\boldsymbol{c}_{t_3}/e - \boldsymbol{c}_{t_2}) \leq 0$, ..., $\boldsymbol{\rho}'_{t_{m-1}}(\boldsymbol{c}_{t_m}/e - \boldsymbol{c}_{t_{m-1}}) \leq 0$ and $\boldsymbol{\rho}'_{t_m}(\boldsymbol{c}_{t_1}/e - \boldsymbol{c}_{t_m}) < 0$. Given $\delta \in (0, 1]$, we also have $\delta^{-t_i}\boldsymbol{\rho}'_{t_i}(\boldsymbol{c}_{t_{i+1}}/e - \boldsymbol{c}_{t_i}) \leq 0$ for all $i \in \{1, \ldots, m-1\}$ and $\delta^{-t_m}\boldsymbol{\rho}'_{t_m}(\boldsymbol{c}_{t_1}/e - \boldsymbol{c}_{t_m}) < 0$. Summing up the resulting inequalities yields

$$0 > \delta^{-t_1}\boldsymbol{\rho}'_{t_1}(\boldsymbol{c}_{t_2}/e - \boldsymbol{c}_{t_1}) + \delta^{-t_2}\boldsymbol{\rho}'_{t_2}(\boldsymbol{c}_{t_3}/e - \boldsymbol{c}_{t_2}) + \cdots + \delta^{-t_m}\boldsymbol{\rho}'_{t_m}(\boldsymbol{c}_{t_1}/e - \boldsymbol{c}_{t_m})$$

which violates CM(e).

Appendix E Proofs

Proof of Theorem 1

(i) $\implies$ *(ii)*

From the first-order condition, we have

$$\nabla u(\boldsymbol{c}_t) \leq \delta^{-t}\boldsymbol{\rho}_t \quad \forall t \in \mathcal{T}$$

where $\nabla u(\boldsymbol{c}_t)$ is some supergradient of $u(\cdot)$ at $\boldsymbol{c}_t$. By continuity and concavity of the instantaneous utility function, we know that for all $t \in \mathcal{T}$ and $\boldsymbol{c} \in \mathbb{R}^L_+$

$$u(\boldsymbol{c}) \leq u(\boldsymbol{c}_t) + \nabla u(\boldsymbol{c}_t)'(\boldsymbol{c} - \boldsymbol{c}_t)$$

Let N be a set of indices such that $\nabla u(\boldsymbol{c}_t)_j = \delta^{-t}\boldsymbol{\rho}_{t,j}$ for all $j \in N$. It follows that $\nabla u(\boldsymbol{c}_t)_j \leq \delta^{-t}\boldsymbol{\rho}_{t,j}$ for all $j \notin N$. Thus, $\boldsymbol{c}_{t,j} = 0$ is a corner solution for all $j \notin N$. We therefore have

$$\begin{aligned} u(\boldsymbol{c}) - u(\boldsymbol{c}_t) \leq \nabla u(\boldsymbol{c}_t)'(\boldsymbol{c} - \boldsymbol{c}_t) &= \sum_{j \in N} \nabla u(\boldsymbol{c}_t)_j(\boldsymbol{c}_j - \boldsymbol{c}_{t,j}) + \sum_{j \notin N} \nabla u(\boldsymbol{c}_t)_j(\boldsymbol{c}_j - \boldsymbol{c}_{t,j}) \\ &= \sum_{j \in N} \delta^{-t}\boldsymbol{\rho}_{t,j}(\boldsymbol{c}_j - \boldsymbol{c}_{t,j}) + \sum_{j \notin N} \nabla u(\boldsymbol{c}_t)_j(\boldsymbol{c}_j - \boldsymbol{c}_{t,j}) \\ &\leq \sum_{j \in N} \delta^{-t}\boldsymbol{\rho}_{t,j}(\boldsymbol{c}_j - \boldsymbol{c}_{t,j}) + \sum_{j \notin N} \delta^{-t}\boldsymbol{\rho}_{t,j}(\boldsymbol{c}_j - \boldsymbol{c}_{t,j}) \end{aligned}$$

where the last inequality holds since $\boldsymbol{c}_{t,j} = 0$ and $\boldsymbol{c}_j \geq 0$ for all $j \notin N$. As a result, for all $t \in \mathcal{T}$ and $\boldsymbol{c} \in \mathbb{R}^L_+$

$$u(\boldsymbol{c}) \leq u(\boldsymbol{c}_t) + \delta^{-t}\boldsymbol{\rho}_t'(\boldsymbol{c} - \boldsymbol{c}_t)$$

Rearranging gives that for all $t \in \mathcal{T}$ and $\boldsymbol{c} \in \mathbb{R}^L_+$

$$u(\boldsymbol{c}_t) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}_t \geq u(\boldsymbol{c}) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}$$

where, by assumption, the instantaneous utility function is locally nonsatiated, continuous, monotonic, and concave and $\delta \in (0, 1]$.

$(ii) \implies (i)$
The instantaneous utility function is locally nonsatiated, continuous, monotonic, and concave and the discount factor satisfies $\delta \in (0, 1]$. For all $t \in \mathcal{T}$ and $\boldsymbol{c} \in \mathbb{R}^L_+$, we also have

$$u(\boldsymbol{c}_t) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}_t \geq u(\boldsymbol{c}) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}$$

Rearranging gives that for all $t \in \mathcal{T}$ and $\boldsymbol{c} \in \mathbb{R}^L_+$

$$u(\boldsymbol{c}) \leq u(\boldsymbol{c}_t) + \delta^{-t}\boldsymbol{\rho}_t'(\boldsymbol{c} - \boldsymbol{c}_t)$$

This inequality corresponds to the definition of concavity and, therefore, it follows that $\delta^{-t}\boldsymbol{\rho}_t$ is a supergradient of $u(\cdot)$ at $\boldsymbol{c}_t$ for all $t \in \mathcal{T}$.

Proof of Proposition 2

(i) $\implies$ *(ii)*

Since the data set $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$ is e-ED rationalizable, it is the case that for all $t \in \mathcal{T}$ and $\boldsymbol{c} \in \mathbb{R}^L_+$

$$u(\boldsymbol{c}_t) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}_t \geq u(\boldsymbol{c}) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}/e$$

for some $\delta \in (0, 1]$. By the same argument as in the proof of Theorem 1 (ii), we can obtain

$$u(\boldsymbol{c}_s) \leq u(\boldsymbol{c}_\mathbf{t}) + \delta^{-t}\boldsymbol{\rho}_t'(\boldsymbol{c}_s/e - \boldsymbol{c}_t) \quad \forall s, t \in \mathcal{T}$$

where one may define $u_t := u(\boldsymbol{c}_t)$ for all $t \in \mathcal{T}$.

(ii) $\implies$ *(iii)*

Starting from the e-ED Afriat inequalities, we have

$$u_s \leq u_t + \delta^{-t}\boldsymbol{\rho}_t'(\boldsymbol{c}_s/e - \boldsymbol{c}_t) \quad \forall s, t \in \mathcal{T}$$

Considering any sequence of indices $\tau \in \mathcal{I}$ and summing up the inequalities for the resulting cycle yields

$$0 \leq \delta^{-t_1}\boldsymbol{\rho}_{t_1}'(\boldsymbol{c}_{t_2}/e - \boldsymbol{c}_{t_1}) + \ldots + \delta^{-t_m}\boldsymbol{\rho}_{t_m}'(\boldsymbol{c}_{t_1}/e - \boldsymbol{c}_{t_m})$$

which corresponds to CM(e).

(iii) $\implies$ *(i)*

For some $e \in (0, 1]$, define

$$u(\boldsymbol{c}) = \inf_{\tau\in\mathcal{I}}\Big\{\delta^{-\tau(m)}\boldsymbol{\rho}_{\tau(m)}'\big(\boldsymbol{c}/e - \boldsymbol{c}_{\tau(m)}\big) + \sum_{i=1}^{m-1}\delta^{-\tau(i)}\boldsymbol{\rho}_{\tau(i)}'\big(\boldsymbol{c}_{\tau(i+1)}/e - \boldsymbol{c}_{\tau(i)}\big)\Big\}$$

This utility function is locally nonsatiated, continuous, monotonic and concave as it is the pointwise minimum of a collection of affine functions. Moreover, the infimum defining $u(\boldsymbol{c})$ has no cycle of indices. To see this, let $s \in \mathcal{T}$ and note that by CM(e) we have

$$0 \leq \delta^{-\tau(1)}\boldsymbol{\rho}_{\tau(1)}'(\boldsymbol{c}_{\tau(2)}/e - \boldsymbol{c}_{\tau(1)}) + \ldots + \delta^{-\tau(m)}\boldsymbol{\rho}_{\tau(m)}'(\boldsymbol{c}_s/e - \boldsymbol{c}_{\tau(m)}) + \delta^{-s}\boldsymbol{\rho}_s'(\boldsymbol{c}_{\tau(1)}/e - \boldsymbol{c}_s)$$

for all $\tau \in \mathcal{T}$. Consider $\boldsymbol{c} \in \mathbb{R}^L_+$ such that $\boldsymbol{c} \neq \boldsymbol{c}_t$ and let $\tau_t \in \mathcal{I}$ be a minimizing sequence for $\boldsymbol{c}_t$. It follows that

$$\begin{aligned}
u(\boldsymbol{c}) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}/e &\leq \delta^{-t}\boldsymbol{\rho}_t'\big(\boldsymbol{c}/e - \boldsymbol{c}_t\big) + \delta^{-\tau_t(m_t)}\boldsymbol{\rho}_{\tau_t(m_t)}'\big(\boldsymbol{c}_t/e - \boldsymbol{c}_{\tau_t(m_t)}\big) \\
&\quad + \sum_{i=1}^{m_t-1} \delta^{-\tau_t(i)}\boldsymbol{\rho}_{\tau_t(i)}'\big(\boldsymbol{c}_{\tau_t(i+1)}/e - \boldsymbol{c}_{\tau_t(i)}\big) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}/e \\
&= \delta^{-\tau_t(m_t)}\boldsymbol{\rho}_{\tau_t(m_t)}'\big(\boldsymbol{c}_t/e - \boldsymbol{c}_{\tau_t(m_t)}\big) \\
&\quad + \sum_{i=1}^{m_t-1} \delta^{-\tau_t(i)}\boldsymbol{\rho}_{\tau_t(i)}'\big(\boldsymbol{c}_{\tau_t(i+1)}/e - \boldsymbol{c}_{\tau_t(i)}\big) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}_t \\
&= u(\boldsymbol{c}_t) - \delta^{-t}\boldsymbol{\rho}_t'\boldsymbol{c}_t
\end{aligned}$$

where the first inequality holds since $u(\boldsymbol{c})$ uses the sequence achieving the infimum for $\boldsymbol{c}$, the first equality is a mere simplification, and the last equality is a consequence of τ_t being a minimizing sequence for $\boldsymbol{c}_t$. I thus have shown the existence of a locally nonsatiated, continuous, monotonic and concave utility function and a discount factor $\delta \in (0, 1]$ e-ED rationalizing the data.

Proof of Theorem 2

$(i) \implies (ii)$

Note that if $\delta^{-t}\boldsymbol{\rho}_t R(e)\delta^{-s}\boldsymbol{\rho}_s$, then $\boldsymbol{c}_s'(\delta^{-t}\boldsymbol{\rho}_t/e - \delta^{-s}\boldsymbol{\rho}_s) \leq 0$. As such, we have that $U_t(\boldsymbol{c}_t, -\boldsymbol{c}_t'\delta^{-t}\boldsymbol{\rho}_t) \geq U_t(\boldsymbol{c}_s, -\boldsymbol{c}_s'\delta^{-t}\boldsymbol{\rho}_t/e) \geq U_s(\boldsymbol{c}_s, -\boldsymbol{c}_s'\delta^{-s}\boldsymbol{\rho}_s)$, where the first inequality follows by definition of rationalizability and the second from the revealed price preference relation. The same argument can be made for the strict relation $P(e)$. Suppose now that δ-GAPP(e) were violated. Then, there would be $t_1, t_m \in \mathcal{T}$ such that $\delta^{-t_1}\boldsymbol{\rho}_{t_1} R(e)\delta^{-t_m}\boldsymbol{\rho}_{t_m}$ and $\delta^{-t_m}\boldsymbol{\rho}_{t_m} P^D(e)\delta^{-t_1}\boldsymbol{\rho}_{t_1}$. This implies that there also exist $t_2, t_3, \ldots, t_{m-1} \in \mathcal{T}$ such that $\delta^{-t_1}\boldsymbol{\rho}_{t_1} R^D(e)\delta^{-t_2}\boldsymbol{\rho}_{t_2}$, $\delta^{-t_2}\boldsymbol{\rho}_{t_2} R^D(e)\delta^{-t_3}\boldsymbol{\rho}_{t_3}$, $\ldots$, $\delta^{-t_{m-1}}\boldsymbol{\rho}_{t_{m-1}} R^D(e)\delta^{-t_m}\boldsymbol{\rho}_{t_m}$. Thus, we would have $U_{t_1}(\boldsymbol{c}_{t_1}, -\boldsymbol{c}_{t_1}'\delta^{-t_1}\boldsymbol{\rho}_{t_1}) \geq U_{t_2}(\boldsymbol{c}_{t_2}, -\boldsymbol{c}_{t_2}'\delta^{-t_2}\boldsymbol{\rho}_{t_2}) \geq \cdots \geq U_{t_m}(\boldsymbol{c}_{t_m}, -\boldsymbol{c}_{t_m}'\delta^{-t_m}\boldsymbol{\rho}_{t_m}) > U_{t_1}(\boldsymbol{c}_{t_1}, -\boldsymbol{c}_{t_1}'\delta^{-t_1}\boldsymbol{\rho}_{t_1})$, an obvious contradiction.

$(ii) \implies (iii)$

In any given time period $t \in \mathcal{T}$, denote the consumer's discounted income by $y > 0$ and let $s_t^d := y - \boldsymbol{c}_t'\delta^{-t}\boldsymbol{\rho}_t$ denote discounted savings. Extending the data set for savings, one obtains $\{(\boldsymbol{\rho}_t, 1), (\boldsymbol{c}_t, s_t^d)\}_{t\in\mathcal{T}}$ with the value of money priced to 1. Since variables are discounted by interest rates, one can also think of the price for money as evolving according to interest rates but where consumption prices and income are in nominal terms. For all $s, t \in \mathcal{T}$, note that we have

$$(\delta^{-t}\boldsymbol{\rho}_t/e, 1)'\big((\boldsymbol{c}_s, y - \boldsymbol{c}_s'\delta^{-s}\boldsymbol{\rho}_s) - (\boldsymbol{c}_t, y - \boldsymbol{c}_t'\delta^{-t}\boldsymbol{\rho}_t/e)\big) \leq 0 \text{ iff } \boldsymbol{c}_s'(\delta^{-t}\boldsymbol{\rho}_t/e - \delta^{-s}\boldsymbol{\rho}_s) \leq 0,$$

where the same equivalence applies with strict inequalities. Now, define $a_{s,t} := \boldsymbol{c}_t'(\delta^{-s}\boldsymbol{\rho}_s/e - \delta^{-t}\boldsymbol{\rho}_t)$ and let the matrix A be defined by $A_{s,t} := a_{s,t}\ \forall s, t \in \mathcal{T}$. Likewise, define $\tilde{a}_{s,t} := (\delta^{-t}\boldsymbol{\rho}_t/e, 1)'\big((\boldsymbol{c}_s, y - \boldsymbol{c}_s'\delta^{-s}\boldsymbol{\rho}_s) - (\boldsymbol{c}_t, y - \boldsymbol{c}_t'\delta^{-t}\boldsymbol{\rho}_t/e)\big)$ and let the matrix $\tilde{A}$ be defined by $\tilde{A}_{s,t} := \tilde{a}_{s,t}\ \forall s, t \in \mathcal{T}$. By Lemma 1, we know that δ-GAPP(e) holds if and only if A is cyclically consistent. Since the expression after "iff" corresponds to the notion of revealed price preference, $\tilde{A}$ must also satisfy cyclical consistency if and only if δ-GAPP(e) holds. An application of Fostel, Scarf and Todd (2004) (Sections 2 and 3) on the matrix $\tilde{A}$ then guarantees the existence of Afriat inequalities given by

$$u_s - u_t \leq \lambda_t(\delta^{-t}\boldsymbol{\rho}_t/e, 1)'\big((\boldsymbol{c}_s, y - \boldsymbol{c}_s'\delta^{-s}\boldsymbol{\rho}_s) - (\boldsymbol{c}_t, y - \boldsymbol{c}_t'\delta^{-t}\boldsymbol{\rho}_t/e)\big) \quad \forall s, t \in \mathcal{T}.$$

Summing up the above inequalities, it is clear that an analogue of CM(e) is satisfied. We can thus construct a well-behaved time-dependent utility function on $\{(\boldsymbol{\rho}_t, 1), (\boldsymbol{c}_t, s_t^d)\}_{t\in\mathcal{T}}$ following similar steps as those in the proof of Theorem 1 and Proposition 2. To this end, for any $\boldsymbol{c} \in \mathbb{R}^L$ and $t \in \mathcal{T}$, define

$$\begin{aligned}
&\tilde{U}_t(\boldsymbol{c}, y - \boldsymbol{c}'\delta^{-t}\boldsymbol{\rho}_t) = \\
&\inf_{\tau\in\mathcal{I}}\Big\{\lambda_{\tau(m)}\left[\delta^{-\tau(m)}\boldsymbol{\rho}_{\tau(m)}'/e\big(\boldsymbol{c} - \boldsymbol{c}_{\tau(m)}\big) + \big(y - \boldsymbol{c}'\delta^{-t}\boldsymbol{\rho}_t\big) - \big(y - \boldsymbol{c}_{\tau(m)}'\delta^{\tau(m)}\boldsymbol{\rho}_{\tau(m)}/e\big)\right] \\
&+ \sum_{i=1}^{m-1}\Big(\lambda_{\tau(i)}\Big[\delta^{-\tau(i)}\boldsymbol{\rho}_{\tau(i)}'/e\big(\boldsymbol{c}_{\tau(i+1)} - \boldsymbol{c}_{\tau(i)}\big) \\
&+ \big(y - \boldsymbol{c}_{\tau(i+1)}'\delta^{-\tau(i+1)}\boldsymbol{\rho}_{\tau(i+1)}\big) - \big(y - \boldsymbol{c}_{\tau(i)}'\delta^{-\tau(i)}\boldsymbol{\rho}_{\tau(i)}/e\big)\Big)\Big]\Big\}.
\end{aligned}$$

To see why $\tilde{U}_t : \mathbb{R}^L \times \mathbb{R} \to \mathbb{R}$ is time-dependent, simply note that for two distinct time periods $s, t \in \mathcal{T}$, $\tilde{U}_s(\boldsymbol{c}_s, y - \boldsymbol{c}_s'\delta^{-s}\boldsymbol{\rho}_s) \neq \tilde{U}_t(\boldsymbol{c}_t, y - \boldsymbol{c}_t'\delta^{-t}\boldsymbol{\rho}_t)$ even if $\boldsymbol{\rho}_s = \boldsymbol{\rho}_t$ and $\boldsymbol{c}_s = \boldsymbol{c}_t$. Moreover, note that $\tilde{U}_t(\cdot,\cdot)$ defines a continuous, strictly increasing in $(\boldsymbol{c}, y - \boldsymbol{c}'\delta^{-t}\boldsymbol{\rho}_t)$, and concave time-dependent utility function. Consider $\boldsymbol{c} \in \mathbb{R}^L$ such that $\boldsymbol{c} \neq \boldsymbol{c}_t$ and let $\tau_t \in \mathcal{I}$ be a minimizing sequence for $\boldsymbol{c}_t$. It follows that

$$
\begin{aligned}
&\tilde{U}_t(\boldsymbol{c}, y - \boldsymbol{c}'\delta^{-t}\boldsymbol{\rho}_t/e) \leq \lambda_t \left[\delta^{-t}\boldsymbol{\rho}_t'/e\big(\boldsymbol{c} - \boldsymbol{c}_t\big) + \big(y - \boldsymbol{c}'\delta^{-t}\boldsymbol{\rho}_t/e\big) - \big(y - \boldsymbol{c}_t'\delta^{-t}\boldsymbol{\rho}_t/e\big)\right] \\
&+ \lambda_{\tau_t(m_t)} \Big[\delta^{-\tau_t(m_t)}\boldsymbol{\rho}'_{\tau_t(m_t)}/e\big(\boldsymbol{c}_t - \boldsymbol{c}_{\tau_t(m_t)}\big) + \big(y - \boldsymbol{c}_t'\delta^{-t}\boldsymbol{\rho}_t\big) - \big(y - \boldsymbol{c}'_{\tau(m_t)}\delta^{-\tau(m_t)}\boldsymbol{\rho}_{\tau(m_t)}/e\big)\Big] \\
&+ \sum_{i=1}^{m_t-1} \lambda_{\tau_t(i)} \Big[\delta^{-\tau_t(i)}\boldsymbol{\rho}'_{\tau_t(i)}/e\big(\boldsymbol{c}_{\tau_t(i+1)} - \boldsymbol{c}_{\tau_t(i)}\big) + \big(y - \boldsymbol{c}'_{\tau_t(i+1)}\delta^{-\tau_t(i+1)}\boldsymbol{\rho}_{\tau_t(i+1)}\big) \\
&- \big(y - \boldsymbol{c}'_{\tau_t(i)}\delta^{-\tau_t(i)}\boldsymbol{\rho}_{\tau_t(i)}/e\big)\Big] \\
&= \lambda_{\tau_t(m_t)} \Big[\delta^{-\tau_t(m_t)}\boldsymbol{\rho}'_{\tau_t(m_t)}/e\big(\boldsymbol{c}_t - \boldsymbol{c}_{\tau_t(m_t)}\big) + \big(y - \boldsymbol{c}_t'\delta^{-t}\boldsymbol{\rho}_t\big) - \big(y - \boldsymbol{c}'_{\tau(m_t)}\delta^{-\tau(m_t)}\boldsymbol{\rho}_{\tau(m_t)}/e\big)\Big] \\
&+ \sum_{i=1}^{m_t-1} \lambda_{\tau_t(i)} \Big[\delta^{-\tau_t(i)}\boldsymbol{\rho}'_{\tau_t(i)}/e\big(\boldsymbol{c}_{\tau_t(i+1)} - \boldsymbol{c}_{\tau_t(i)}\big) + \big(y - \boldsymbol{c}'_{\tau_t(i+1)}\delta^{-\tau_t(i+1)}\boldsymbol{\rho}_{\tau_t(i+1)}\big) \\
&- \big(y - \boldsymbol{c}'_{\tau_t(i)}\delta^{-\tau_t(i)}\boldsymbol{\rho}_{\tau_t(i)}/e\big)\Big] \\
&= \tilde{U}_t(\boldsymbol{c}_t, y - \boldsymbol{c}_t'\delta^{-t}\boldsymbol{\rho}_t),
\end{aligned}
$$

where the first inequality holds since $\tilde{U}_t(\boldsymbol{c}, y - \boldsymbol{c}'\delta^{-t}\boldsymbol{\rho}_t/e)$ uses the sequence achieving the infimum for $\boldsymbol{c}$, the first equality is a mere simplification, and the last equality is a consequence of τ_t being a minimizing sequence for $\boldsymbol{c}_t$. Finally, defining $U_t : \mathbb{R}^L_+ \times \mathbb{R}_- \to \mathbb{R}$ by $U_t(\boldsymbol{c}, -\boldsymbol{c}'\delta^{-t}\boldsymbol{\rho}_t) := \tilde{U}_t(\boldsymbol{c}, y - \boldsymbol{c}'\delta^{-t}\boldsymbol{\rho}_t)$ yields a time-dependent augmented utility function rationalizing the data $\{(\boldsymbol{\rho}_t, \boldsymbol{c}_t)\}_{t\in\mathcal{T}}$, as desired. By an identical argument as in Theorem 1 of Deb et al. (2018), it is possible to modify the aforementioned utility function in order to further guarantee the existence of a solution to any set of prices $\boldsymbol{\rho} \in \mathbb{R}^L_{++}$. Since $U_t(\cdot,\cdot)$ shares the same properties as $\tilde{U}_t(\cdot,\cdot)$, we get that $(iii) \implies (i)$.

REFERENCES

Ackerberg, Daniel A. 2003. “Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination.” *International Economic Review*, 44(3): 1007–1040.

Adams, Abi, Laurens Cherchye, Bram De Rock, and Ewout Verriest. 2014. “Consume Now or Later? Time Inconsistency, Collective Choice, and Revealed Preference.” *American Economic Review*, 104(12): 4147–4183.

Afriat, Sydney N. 1967. “The Construction of Utility Functions from Expenditure Data.” *International Economic Review*, 8(1): 67–77.

Afriat, Sydney N. 1972. “Efficiency Estimation of Production Functions.” *International Economic Review*, 13(3): 568–598.

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

Aguiar, Victor H., and Nail Kashaev. 2020. “Stochastic Revealed Preferences with Measurement Error.” *Review of Economic Studies*, forthcoming.

Allen, Roy, and John Rehbeck. 2020. “Satisficing, Aggregation, and Quasilinear Utility.” http://dx.doi.org/10.2139/ssrn.3180302.

Blow, Laura, Martin Browning, and Ian Crawford. 2017. “Nonparametric Analysis of Time-inconsistent Preferences.” Economics Series Working Papers 835, University of Oxford, Department of Economics.

Blundell, Richard, Martin Browning, and Ian Crawford. 2003. “Nonparametric Engel Curves and Revealed Preference.” *Econometrica*, 71(1): 205–240.

Blundell, Richard, Martin Browning, and Ian Crawford. 2008. “Best Nonparametric Bounds on Demand Responses.” *Econometrica*, 76(6): 1227–1262.

Brown, Donald J., and Caterina Calsamiglia. 2007. “The Nonparametric Approach to Applied Welfare Analysis.” *Economic Theory*, 31(1): 183–188.

Browning, Martin. 1989. “A nonparametric Test of the Life-cycle Rational Expectations Hypothesis.” *International Economic Review*, 30(4): 979–992.

Board of Governors of the Federal Reserve System (US). 1972-2018. “Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan.” https://fred.stlouisfed.org/series/TERMCBPER24NS. Accessed August 1, 2018.

Cherchye, Laurens, Thomas Demuynck, Bram de Rock, and Joshua Lanier. 2020. “Are consumers rational? Shifting the burden of proof.” Ecares working paper, 2020-19.

Crawford, Ian. 2010. "Habits Revealed." *Review of Economic Studies*, 77(4): 1382–1402.

de Clippel, Geoffroy, and Kareen Rozen. 2018. "Consumer Theory with Misperceived Tastes." Working Papers 2018-10, Brown University, Department of Economics.

Demuynck, Thomas, and Ewout Verriest. 2013. "I'll Never Forget my First Cigarette: A Revealed Preference Analysis of the 'Habits as Durables' Model.", *International Economic Review*, 54(2): 717–738.

Dziewulski, Pawel. 2020. "Just-noticeable difference as a behavioural foundation of the critical cost-efficiency index" *Journal of Economic Theory*, 188: 105071.

Deb, Rahul, Yuichi Kitamura, John K.-H. Quah, and Jörg Stoye. 2018. "Revealed Price Preference: Theory and Empirical Analysis." https://arxiv.org/abs/1801.02702.

Echenique, Federico, Sangmok Lee, and Matthew Shum. 2011. "The Money Pump as a Measure of Revealed Preference Violations." *Journal of Political Economy*, 119(6): 1201–1223.

Echenique, Federico, Taisuke Imai, and Kota Saito. 2020. "Testable Implications of Models of Intertemporal Choice: Exponential Discounting and Its Generalizations." *American Economic Review: Microeconomics*, forthcoming.

Echenique, Federico, Taisuke Imai, and Kota Saito. 2018. "Approximate Expected Utility Rationalization." Technical report, CRC TRR 190 Rationality and Competition.

Fostel, Ana, Herbert Scarf, and Michael Todd. 2004. "Two New Proofs of Afriat's Theorem." *Economic Theory*, 24(1): 211–219.

Frederick, Shane, George Loewenstein, and Ted O'Donoghue. 2002. "Time Discounting and Time Preference: A Critical Review." *Journal of Economic Literature*, 40(2): 351–401.

Geanakoplos, John D. 2013. "Afriat from MaxMin." *Economic Theory*, 54(3): 443–448.

Griffith, Rachel, Ephraim Leibtag, Andrew Leicester, and Aviv Nevo. 2009. "Consumer shopping behavior: how much do consumers save?" *Journal of Economic Perspectives*, 23(2): 99—120.

Halevy, Yoram. 2015. "Time Consistency: Stationarity and Time Invariance." *Econometrica*, 83(1): 335–352.

Halevy, Yoram, Dotan Persitz, and Lanny Zrill. 2018. "Parametric

Recoverability of Preferences." *Journal of Political Economy*, 126(4): 1558–1593.

Heufer, Jan, and Per Hjertstrand. 2017. "Homothetic Efficiency: Theory and Applications." *Journal of Business & Economic Statistics*, 0(0): 1–13. doi:10.1080/07350015.2017.1319372.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect theory: An Analysis of Decisions under Risk." *Econometrica*, 47(2): 263–292.

Kitamura, Yuichi, and Jörg Stoye. 2018 "Nonparametric Analysis of Random Utility Models." *Econometrica*, 86(6): 1883–1909.

Reny, Philip J. 2015. "A Characterization of Rationalizable Consumer Behavior." *Econometrica*, 83(1): 175–192.

Rockafellar, R. Tyrell. 1970. *Convex analysis.* Princeton university press.

Varian, Hal R. 1982. "The Nonparametric Approach to Demand Analysis." *Econometrica*, 50(4): 945–973.

Varian, Hal R. 1985. "Non-parametric Analysis of Optimizing Behavior with Measurement Error." *Journal of Econometrics*, 30(1-2): 445–458.

Varian, Hal R. 1990. "Goodness-of-fit in Optimizing Models." *Journal of Econometrics*, 46(1-2): 125–140.

Yao, Song, Carl F. Mela, Jeongwen Chiang, and Yuxin Chen. 2012. "Determining Consumers' Discount Rates with Field Studies." *Journal of Marketing Research*, 49(6): 822-841.