Testable Implications of Models of Intertemporal Choice: Exponential Discounting and Its Generalizations

By Federico Echenique, Taisuke Imai, and Kota Saito *

We present revealed-preference characterizations of the most common models of intertemporal choice: the model of exponentially discounted concave utility, and some of its generalizations. Our characterizations take consumption data *as primitives, and provide nonparametric revealed-preference tests. We apply our tests to data from two recent experiments and find that our axiomatization delivers new insights and perspectives on datasets that had been analyzed by traditional parametric methods.*

Exponentially discounted utility is the standard model of intertemporal choice in economics. It is a ubiquitous model, used in all areas of economics. Our paper is a revealed preference investigation of exponential discounting: We give a necessary and sufficient "revealed-preference axiom" that a dataset must satisfy in order to be consistent with exponential discounting. The revealed-preference axiom sheds light on the behavioral assumptions underlying the standard model of discounting. It also yields a nonparametric test of the theory, applicable in different empirical investigations of exponential discounting.

Consider an agent who chooses among intertemporal consumptions of a single good. One general theory is that the agent has a utility function $U(x_0, \dots, x_T)$ for the consumption of x_t on each date t. The Generalized Axiom of Revealed Preference (GARP) tells us whether the agent's choices are consistent with the maximization of some general utility function U. The empirical content of general utility maximization is well understood, but utility maximization is too broad, and GARP is too weak, to capture exponential discounting. The *exponentially discounted utility* (EDU) model assumes a specific form of U, namely

$$U(x_0, \dots, x_T) = \sum_{t=0}^{T} \delta^t u(x_t).$$

In this paper, we focus on concave EDU, in which u is a concave function. A concavity of u is widely used to capture a motive for consumption smoothing over time. The empirical

* Echenique: Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125, `fede@hss.caltech.edu`. Imai: Department of Economics, LMU Munich, Munich 80539, Germany, `taisuke.imai@econ.lmu.de`. Saito: Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125, `saito@caltech.edu`. We thank Kim Border and Chris Chambers for inspiration and advice. This paper subsumes the paper "Testable Implication of Exponential Discounting" (2014) circulated as Caltech Social Science Working Paper 1381. The authors wish to thank Jim Andreoni and Charlie Sprenger for their many detailed comments on a previous draft, and discussions that really benefited our paper. We are also grateful for the feedback provided by seminar audiences in the many different places where we have presented the paper. Echenique thanks the National Science Foundation for its support through the grants SES 1558757 and CNS-1518941. Imai is grateful for financial supports by the Nakajima Foundation and by the Deutsche Forschungsgemeinschaft through CRC TRR 190. Saito thanks the National Science Foundation for its support through the grant SES 1558757.

content of concave EDU maximization is different from that of general utility maximization, and not well understood in the literature.

The first and most important question addressed in our paper is: *What is the version of GARP that allows us to decide whether data are consistent with concave EDU?* The revealed-preference axiom that characterizes concave EDU is obviously going to be stronger than GARP. Despite the ubiquity of EDU in economics, the literature on revealed preference has not provided an answer. Our main result is that a certain revealed-preference axiom, termed the "Strong Axiom of Revealed Exponentially Discounted Utility" (SAR-EDU), describes the choice data that are consistent with concave EDU preferences.

SAR-EDU is a version of the "downward-sloping demand" property. It says that, with certain qualifications, prices and quantities must be inversely related. At face value, downward-sloping demand says that consumption may be higher in one period than in another as long as the price of consumption in that period is cheaper. In EDU, the utility from later consumption is discounted. So it is possible that consumption is high, even if it is expensive, as long as it happens early in time: Early consumption is more valuable than later consumption. So SAR-EDU qualifies downward-sloping demand. If consumption is higher in some periods than in others, and these periods occur later in time, then consumption must be cheaper in the later periods with high consumption.

In the paper, we study the empirical content of more general models of time discounting as well, including the quasi-hyperbolic discounting model (QHD; Phelps and Pollak, 1968; Laibson, 1997), $U(x_0, \dots, x_T) = u(x_0) + \beta \sum_{t=1}^{T} \delta^t u(x_t)$, and time-separable utility (TSU), $U(x_0, \dots, x_T) = \sum_{t=0}^{T} u_t(x_t)$, where u and u_t are concave. In the following, we do not explicitly use the concave modifier when there is no risk of confusion. For example, we say EDU to mean concave EDU.

The contribution of our paper is to characterize the empirical content of EDU and its generalizations. We provide revealed-preference axioms (axioms like GARP but stronger) characterizing EDU, QHD, and TSU. Our axioms shed new insights into the behavioral assumptions behind each of these models, and also constitute nonparametric tests. There are, of course, other axiomatizations of these models but they start from different primitives. The well-known axiomatization of EDU by Koopmans (1960), for example, starts from complete preferences over infinite consumption streams.

As additional contributions, we provide a revealed-preference characterization for the general time discounting (GTD) model: $U(x_0, \dots, x_T) = \sum_{t=0}^{T} D(t)u(x_t)$, where u is concave. GTD is more general than EDU and QHD, while it remains a special case of TSU. We also provide a characterization of the monotone time discounting (MTD) model, in which the discounting function $D(t)$ is decreasing because of impatience. We believe that these characterizations are useful to understand the behavioral meaning of impatience. MTD includes the models of diminishing impatience and its variations proposed by Halevy (2008) and Chakraborty, Halevy and Saito (forthcoming).

To illustrate the usefulness of our results for empirical work, we carry out an application to data from two recent experiments conducted by Andreoni and Sprenger (2012) (hereafter AS) and Carvalho, Meier and Wang (2016) (hereafter CMW). AS propose the Convex Time Budget (CTB) experimental design, in which subjects are asked to choose from an intertemporal budget set. CMW adopt the CTB design, and study the effect of financial resources on intertemporal decision-making.[1]

The applications of our methods to AS's and CMW's data are, we believe, fruitful. We uncover features of individual subjects' behavior that are masked by traditional parametric econometric techniques. Despite some clear differences between the AS and CMW designs, our methods yield similar results. First, the numbers of EDU-rational agents are rather small. Second, there is very little added scope for QHD. In the case of the AS experiment, all subjects rationalized as QHD are also rationalized as EDU. In the case of CMW, a very small number of subjects are QHD, but not EDU, rationalizable. Finally, the number of TSU-rational subjects is about half; the rest of the subjects are not rationalized even by the TSU model.

It should be said that our methods rest on nonparametric revealed-preference tests. As such, the tests are independent of functional form assumptions. The tests are also simple, and tightly connected to economic theory. The methodology used currently by experimentalists rests instead on parametrically estimating a given utility function. Our setup fits the experimental design of AS and CMW, and other CTB experiments, very well, but our results are also applicable more broadly, including to non-experimental field data.

Related literature. — There are different behavioral axiomatizations of EDU in the literature, starting with Koopmans (1960), and followed by Fishburn and Rubinstein (1982), Fishburn and Edwards (1997), and Bleichrodt, Rohde and Wakker (2008).

All of them take preferences or utility functions as primitive. The idea is that the relevant behavior consists of all pairwise comparisons of consumption streams. From an empirical perspective, this assumes an infinite dataset of pairwise comparisons. The difference with our work is that we start from a finite dataset of choices from "economic" budgets instead of pairwise comparisons. One advantage of infinite datasets is that one can talk about the model being identified. There is very little hope to obtain identification with a finite dataset.

In the continuous-time setup, Weibull (1985) gives a general characterization of EDU, also taking preferences as primitives. A more recent paper by Kopylov (2010) also provides a simple axiomatization of EDU in a continuous-time setup.

The QHD model was first proposed by Phelps and Pollak (1968), who did not provide an axiomatization. Several recent studies present a behavioral characterization of QHD, but all take preferences and infinite time horizons as their primitives and therefore differ from our results. See Hayashi (2003), Montiel Olea and Strzalecki (2014), and Galperti and Strulovici (2017) for axiomatizations.

The recent work of Dziewulski (2018) gives a characterization of EDU and QHD for finitely many pairwise comparisons of one-time consumptions in a setup similar to Fishburn and Rubinstein's (1982), but with finite data.

[1]Several recent experimental studies use the CTB design, both in the laboratory and in the field setting, including Andreoni, Kuhn and Sprenger (2015), Augenblick, Niederle and Sprenger (2015), Balakrishnan, Haushofer and Jakiela (forthcoming), Barcellos and Carvalho (2014), Brocas, Carrillo and Tarrasó (2018), Carvalho, Prina and Sydnor (2016), Giné et al. (2018), Janssens, Kramer and Swart (2017), Kuhn, Kuhn and Villeval (2017), Liu, Meng and Wang (2014), Lührmann, Serra-Garcia and Winter (2018), Sawada and Kuroishi (2015), and Sun and Potters (2016). Our methods are largely applicable to data from these experiments.

Time-separable utility is the most general model we axiomatize. In the application of our test to AS's and CMW's data, however, we found that a significant number of subjects are not TSU rational. This would suggest the importance of a non-time-separable model. The result by Varian (1983) can be interpreted within our context as providing a test of time-separable utility, although he does not deal with intertemporal choice. His characterization is in terms of the existence of a solution to a system of linear Afriat inequalities. Our characterization is a combinatorial condition. Gilboa (1989) has provided an elegant axiomatization of a non-time-separable utility model. In the paper, by using Anscombe and Aumann's (1963) framework and studying preferences over finite sequences of lotteries, Gilboa (1989) axiomatizes a utility function that can capture a preference for (or an aversion to) variation of utility levels across periods. The paper by Quah (2014) studies (general, non-additive) separability from a revealed-preference perspective. His approach is notable in that he does not need to assume the convexity of preferences.

A few papers focus on data from consumption surveys and Afriat inequalities. Browning (1989) provides a revealed-preference axiom for EDU with $\delta = 1$ and a single observation. Crawford (2010) investigates intertemporal consumption and discusses a particular violation of TSU, namely habit formation. Crawford (2010) presents Afriat inequalities for the model of habit formation and uses Spanish consumption data to carry out the test (see also Crawford and Polisson, 2014). Adams et al. (2014) work with the Spanish dataset and test EDU within a model of collective decision making at the household level. Aguiar and Kashaev (2018) provide a stochastic revealed-preference approach that is applicable to consumer survey data with measurement error.

It is important to emphasize that the papers on survey data allow for the existence of many goods in each period, but they do not allow for more than one (intertemporal) purchase for each agent. This assumption makes sense because in consumption surveys one typically has a single observation per household. We have instead assumed that there is only one good (money) in each period, but we allow for more than one intertemporal purchase per agent. Allowing for multiple purchases is crucial in order to apply our tests to experimental data. In experiments, a subject is usually required to make many decisions (one choice is chosen randomly to determine the payment to the subject).

I. Intertemporal Choice and Discounted Utility

A. *Notational Conventions*

For vectors $x, y \in \mathbf{R}^n$, $x \leq y$ means that $x_i \leq y_i$ for all $i = 1, \ldots, n$, $x < y$ means that $x \leq y$ and $x \neq y$, and $x \ll y$ means that $x_i < y_i$ for all $i = 1, \ldots, n$. The set of all $x \in \mathbf{R}^n$ with $0 \leq x$ is denoted by $\mathbf{R}^n_+$ and the set of all $x \in \mathbf{R}^n$ with $0 \ll x$ is denoted by $\mathbf{R}^n_{++}$.

Let T be a strictly positive integer; T will be the (finite) duration of time, or *time horizon*. We abuse notation and use T to denote the set $\{0, 1, \ldots, T\}$. A sequence $(x_0, \ldots, x_T) = (x_t)_{t \in T} \in \mathbf{R}^T_+$ will be called a *consumption stream*. There is a single good in each period; the good can be thought of as money. Note that the cardinality of the set T is $T + 1$, but this never leads to confusion.

REMARK: We can assume more generally that time takes the values $0, \tau_1, \ldots, \tau_T$, where $\tau_i < \tau_{i+1}$ for all $i < T-1$. Our results hold without changes. The only requirement on the set of time periods is that it contains 0. We use a general set of time periods in our application to experimental data (see Section III.A).

B. The Model

The objects of choice in our model are consumption streams. We assume that an agent has a budget $I > 0$, faces prices $p \in \mathbf{R}^T_{++}$, and chooses an affordable consumption stream $(x_t)_{t \in T} \in \mathbf{R}^T_+$. Prices can be thought of as interest rates.

A *model* is a class of utility functions $U : \mathbf{R}^T_+ \to \mathbf{R}$. Classical revealed preference theory focuses on the class M of locally non-satiated utility functions. An *M-rational* agent behaves as if she solves the problem:

$$\max_{x \in B(p,I)} U(x)$$

when faced with prices $p \in \mathbf{R}^T_{++}$ and budget $I > 0$. The set $B(p, I) = \{y \in \mathbf{R}^T_+ : p \cdot y \leq I\}$ is the *budget set* defined by p and I.

The focus in our paper is on more restrictive models. A first model of interest is the class GTD of *general time discounting* utility functions. This is the class of utility functions $U : \mathbf{R}^T_+ \to \mathbf{R}$ for which there exist functions $D : T \to \mathbf{R}_+$ and $u : \mathbf{R}_+ \to \mathbf{R}$ such that u is monotone increasing and concave, and

$$U((x_t)_{t \in T}) = \sum_{t \in T} D(t) u(x_t).$$

As mentioned in the introduction, we restrict attention to concave utility. Our results will be silent about the non-concave case. In consequence, we focus on agents who seek to smooth out their consumption over time.[2]

A second model of interest is *exponentially discounted utility*, the class of utility functions $EDU \subset GTD$ for which $D(t) = \delta^t$ for some $\delta \in (0, 1]$. The EDU model is the standard workhorse model of intertemporal choice and ubiquitous in economic theory.

A third model is *quasi-hyperbolic discounted utility*, obtained as $QHD \subset GTD$ by setting $D(0) = 1$ and $D(t) = \beta\delta^t$ for $t \geq 1$, where $\beta > 0$ and $\delta \in (0, 1]$.

C. The Data

We have said that a model postulates *as-if* behavior by some agent. To explain what we mean, we have to state what can be observed.

DEFINITION 1: *A* dataset *is a finite collection of pairs* $(x, p) \in \mathbf{R}^T_+ \times \mathbf{R}^T_{++}$.

[2]Strictly speaking, the quasiconcavity of the overall utility function (or the convexity of preferences over consumption streams) captures the notion of consumption smoothing, but in our case, quasiconcavity is equivalent to the concavity of u.

A dataset is our notion of observable behavior. The interpretation of a dataset $(x^k, p^k)_{k=1}^K$ is that it describes K observations of a consumption stream $x^k = (x_t^k)_{t \in T}$ at some given vector of prices $p^k = (p_t^k)_{t \in T}$, and budget $p^k \cdot x^k = \sum_{t \in T} p_t^k x_t^k$. We sometimes use K to denote the set $\{1, \dots, K\}$.

Let us clarify the meaning of a dataset by considering two examples. If we have field consumption data, collected through a consumption survey, then $K = 1$. There is one dataset for each agent or household. This is the setup of Browning (1989), for example. On the other hand, if, in an experiment, one subject is asked to make a choice from 45 different budget sets, as in Andreoni and Sprenger (2012), then K is 45. It is important to note that our framework allows, but does not require, that $K > 1$. Even if $K = 1$, our axioms may be violated, and the models are testable.

Our next definition formalizes the concept of as-if choices. Given a model $M' \subseteq M$, an agent is consistent with M', or chooses *as if* M', if some element of M' can be used to generate her choices.

DEFINITION 2: *Given a model $M' \subseteq M$, a dataset $(x^k, p^k)_{k=1}^K$ is M'-*rational *if there is $U \in M'$ such that, for all k,*

$$y \in B(p^k, p^k \cdot x^k) \Longrightarrow U(y) \leq U(x^k).$$

D. Results

The characterization of M-rational data is well-known since Afriat (1967): a dataset is M-rational if and only if it satisfies the Generalized Axiom of Revealed Preference (GARP). The Weak Axiom of Revealed Preference (WARP) is necessary, but not sufficient, for M-rationality. See Varian (1983), or Chambers and Echenique (2014), for definitions and an exposition of the basic theory. The starting point for our analysis is a characterization of GTD-rational dataset.

We need a few definitions first. Given a dataset $(x^k, p^k)_{k=1}^K$, and an observation k, we say that the pair $(x_t^k, x_{t'}^k)$ has the *downward-sloping demand property* if $x_t^k > x_{t'}^k$ implies that $p_t^k \leq p_{t'}^k$. This notion is intuitive enough: larger quantities are associated with lower prices. For reasons that shall become clear in Section II, we need to generalize the notion of downward-sloping demand. In particular, we shall generalize the property to collections, or sequences, of pairs that may not be drawn from the same observation k.

DEFINITION 3: *A sequence of pairs $(x_{t_i}^{k_i}, x_{t'_i}^{k'_i})_{i=1}^n$ is* balanced *if each k appears as k_i (on the left of the pair) the same number of times it appears as k'_i (on the right).*[3]

The meaning of a balanced sequence of pairs is simply that the sequence is obtained from rearranging n observations. We introduced the idea of downward-sloping demand for a single observation, and we now extend it to a collection of n observations, arranged to form a balanced sequence of pairs:

[3]That is, $\#\{i : k_i = k\} = \#\{i : k'_i = k\}$.

DEFINITION 4: *A sequence of pairs* $(x^{k_i}_{t_i}, x^{k'_i}_{t'_i})_{i=1}^n$ *has the* downward-sloping demand property *if*

$$x^{k_i}_{t_i} > x^{k'_i}_{t'_i} \text{ for all } i \text{ implies that } \prod_{i=1}^n \frac{p^{k_i}_{t_i}}{p^{k'_i}_{t'_i}} \leq 1.$$

For now, let us just remark that the downward-sloping demand property of a balanced sequence of pairs is a possible generalization of a single pair $(x^k_t, x^k_{t'})$ having this property. In Section II we shall explain the ideas behind the definition.

We can now state the first result, due to Echenique and Saito (2015), which serves as the starting point of our analysis. The result characterizes GTD-rational choices by means of an axiom:

Strong Axiom of Revealed General Time Discounted Utility (SAR-GTD): *For any balanced sequence of pairs* $(x^{k_i}_{t_i}, x^{k'_i}_{t'_i})_{i=1}^n$*, if each* t *appears as* t_i *(on the left of the pair) the same number of times it appears as* t'_i *(on the right), then the sequence has the downward-sloping demand property.*

THEOREM 0: *A dataset is GTD rational if and only if it satisfies SAR-GTD.*

Theorem 0 was obtained by Echenique and Saito (2015) as a characterization of subjective expected utility, in a model of choice under uncertainty. It is straightforward to interpret their result in the context of intertemporal choice.[4] Our main interest in the present paper is in EDU and QHD; two models of intertemporal choice that are far more important in economics than GTD.

Strong Axiom of Revealed Exponentially Discounted Utility (SAR-EDU): *For any balanced sequence of pairs* $(x^{k_i}_{t_i}, x^{k'_i}_{t'_i})_{i=1}^n$*, if* $\sum_{i=1}^n t_i \geq \sum_{i=1}^n t'_i$*, then the sequence has the downward-sloping demand property.*

THEOREM 1: *A dataset is EDU rational if and only if it satisfies SAR-EDU.*

Observe the relation between SAR-EDU and SAR-GTD. Both axioms require the downward-sloping demand property to hold in different circumstances. SAR-GTD imposes the property on sequences where each time period appears as t'_i the same number of times it appears as t_i. For those sequences we will obviously have $\sum_{i=1}^n t_i = \sum_{i=1}^n t'_i$. SAR-EDU, a more restrictive axiom, requires the downward-sloping property to hold when $\sum_{i=1}^n t_i \geq \sum_{i=1}^n t'_i$. Again, we develop an intuition for these axioms in Section II. The proof of Theorem 1 is in Appendix A. It follows ideas introduced in Echenique and Saito (2015).

Next, we turn to QHD. QHD is often proposed as a relaxation of EDU to accommodate situations where EDU is rejected empirically (Ericson and Laibson, 2019). It is therefore

[4]States become time periods. Subjective beliefs turn into a discount function D.

important to understand its testable implications. We show that these are captured by the following axiom.

Strong Axiom of Revealed Quasi-Hyperbolic Discounted Utility (SAR-QHD): *For any balanced sequence of pairs* $(x^{k_i}_{t_i}, x^{k'_i}_{t'_i})_{i=1}^n$, *if*

(i) $\sum_{i=1}^n t_i \geq \sum_{i=1}^n t'_i$ *and*

(ii) $\#\{i : t_i > 0\} = \#\{i : t'_i > 0\}$,

then the sequence has the downward-sloping demand property.

THEOREM 2: *A dataset is QHD rational if and only if it satisfies SAR-QHD.*

The proof of Theorem 2 is presented in online appendix A.2.

The interpretation of a dataset is more complicated when we consider tests of QHD. In the case of QHD, we assume that each x^k is a consumption stream that the agent *commits* to at $t = 0$. The reason is that a QHD agent may be dynamically inconsistent, and revise their planned consumption.

The commitment assumption fits perfectly the application in Section III to data from CTB experiments, but it will be violated by field data taken from consumption surveys. It is important to emphasize that the assumption of commitment is not necessary to test the EDU model, which is dynamically consistent.

E. Discussion

It is easy to propose examples of data that satisfy WARP or GARP, but fail SAR-EDU. In a sense, WARP and GARP are about the levels of expenditure, while SAR-EDU is about the slope of the budget sets. It is interesting that WARP and GARP can often be interpreted as capturing an inverse relation between prices and quantities. SAR-EDU is stronger than WARP or GARP, and it captures a more structured relation between prices and quantities.

It is not obvious from the syntax of SAR-EDU that one can verify whether a particular dataset satisfies SAR-EDU in finitely many steps. We can show that not only is SAR-EDU decidable in finitely many steps, but there is in fact an efficient algorithm that decides whether a dataset satisfies SAR-EDU. Another way to test SAR-EDU is based on the linearized Afriat inequalities (see Lemma 1 in Appendix A). In fact, this is how we proceed in Section III; see in particular the discussion at the end of Section III.A.

LIMITATIONS. — We should emphasize that our test does not extend easily to non-concave utilities, or to environments with many goods in each period. As will become clear in Section II, the source of our result is a calculation based on first-order conditions. These are not sufficient for optimality in the absence of concavity, and our approach cannot get off the ground. See Quah (2014) for tests that do not require concavity, and that should be applicable in the intertemporal setting.

As for many goods, the procedure of linearizing Afriat inequalities does not work with many goods, and our approach cannot be applied. For experimental datasets, the one-good assumption is not a limitation, but the application of our methods to survey data requires some aggregation.

II. Intuition Behind Theorem 1

The point of this section is to "derive" SAR-EDU from the assumption that a dataset is EDU rational. We shall introduce the axiom by deriving the implications of EDU in specific instances and hence develop the basic intuition behind our results. A secondary benefit of this discussion is that the specific instances will be very relevant to our empirical results in Section III.C.

Here we assume, for ease of exposition, that u is differentiable, but our results do not depend on differentiability.

The first-order condition for maximization of EDU is, for each $k \in K$ and $t \in T$,

$$\delta^t u'(x_t^k) = \lambda^k p_t^k. \tag{1}$$

Here λ^k is the Lagrange multiplier for the problem in the kth observation.[5]

This means, letting $\text{mrs}(x, x') = u'(x)/u'(x')$ denote the marginal rate of substitution (MRS) between x and x', that:

$$\frac{\delta^t}{\delta^{t'}} \text{mrs}(x_t^k, x_{t'}^k) = \frac{p_t^k}{p_{t'}^k}. \tag{2}$$

The first-order conditions in the form of (2) involve two unobservables: the discount factor δ and marginal utilities $u'(x_t^k)$. Quantities x_t^k and prices p_t^k are observable. Our approach proceeds by finding that certain implications of the model for the observables x^k and p^k must hold, regardless of the values of the unobservables. The implications are that quantities x^k and prices p^k be in some sense inversely related.

We derive the axiom by considering increasingly general cases. First, we consider the case of no discounting and one observation ($\delta = 1$ and $K = 1$). Then, we study the case of no discounting ($\delta = 1$ and $K \geq 1$). Finally, in Section II.C we discuss the general case (δ is unknown and $K \geq 1$) and present the revealed-preference axiom for EDU.

A. *No Discounting and One Observation: $\delta = 1$ and $K = 1$*

Suppose that $\delta = 1$ and $K = 1$. That is, we seek to impose EDU rationality in the special case when δ is known, equals 1, and our dataset has a single observation. Under these assumptions (omitting the k superindex, as $K = 1$) the first-order condition (1) becomes

[5] Our informal derivation of the axiom focuses on interior solutions. Our formal result does not depend on interiority, and allows for observations with corner solutions.

$u'(x_t) = \lambda p_t$ for each $t \in T$. For each pair $t, t' \in T$, (2) takes the form:

$$\text{mrs}(x_t, x_{t'}) = \frac{p_t}{p_{t'}}.$$

By concavity of u, for each pair $t, t' \in T$, we know that $\text{mrs}(x_t^k, x_{t'}^k) \leq 1$ when $x_t > x_{t'}$. Therefore,

$$x_t > x_{t'} \Longrightarrow \frac{p_t}{p_{t'}} \leq 1. \tag{3}$$

Thus we obtain a simple implication of EDU rationality: (3) means that *demand must slope down*. For our agent to consume more in period t than in period t', consumption in period t must be cheaper than in t'. This "downward-sloping demand axiom" coincides with the axiom obtained by Browning (1989) for the $\delta = K = 1$ case, and it is the property we referred to above before defining balanced collections of pairs.

Notice that property (3) is a special case of Definition 4 of downward-sloping demand property. The definition is more complicated than (3), and redundant for now, but will prove useful in the sequel. In fact, that every sequence of pairs has the downward-sloping demand property is not only a necessary condition but also a sufficient condition for EDU rationality in the case of $\delta = 1$ and $K = 1$.

B. *No Discounting:* $\delta = 1$

We now take one step towards our general result. Continue to assume that $\delta = 1$, but now allow that $K \geq 1$. The agent does not discount future utilities, but the dataset may contain multiple observations. The first-order condition (1) becomes $u'(x_t^k) = \lambda^k p_t^k$ for each $t \in T$ and each $k \in K$. As we discuss later in Section III.C, the case of $\delta = 1$ is relevant empirically.

If we try to proceed as in the previous section, we might consider two consumption values for observation k: x_t^k, and $x_{t'}^k$. For any two such values, we can consider the first-order condition $\text{mrs}(x_t^k, x_{t'}^k) = p_t^k / p_{t'}^k$ and conclude that, if $x_t^k > x_{t'}^k$, then $p_t^k / p_{t'}^k \leq 1$. This downward-sloping demand implication is the same as (3), and holds *within* each observation k, when we compare quantities across periods.

Downward-sloping demand within observations is one implication of the model, but it is not the only one. There are additional implications *across* observations. Consider consumption values for two different observations k and k': $x_{t_1}^k, x_{t_2}^{k'}, x_{t_3}^{k'}$, and $x_{t_4}^k$. We could consider the two marginal rates of substitution, $\text{mrs}(x_{t_1}^k, x_{t_4}^k)$ and $\text{mrs}(x_{t_3}^{k'}, x_{t_2}^{k'})$ and proceed as above to obtain implications within each observation k and k', but we can obtain further implications. We can rearrange MRSs to obtain

$$\frac{p_{t_1}^k}{p_{t_2}^{k'}} \frac{p_{t_3}^{k'}}{p_{t_4}^k} = \frac{p_{t_1}^k}{p_{t_4}^k} \frac{p_{t_3}^{k'}}{p_{t_2}^{k'}} = \text{mrs}(x_{t_1}^k, x_{t_4}^k) \cdot \text{mrs}(x_{t_3}^{k'}, x_{t_2}^{k'}) = \text{mrs}(x_{t_1}^k, x_{t_2}^{k'}) \cdot \text{mrs}(x_{t_3}^{k'}, x_{t_4}^k).$$

Moreover, if $x_{t_1}^k > x_{t_2}^{k'}$ and $x_{t_3}^{k'} > x_{t_4}^k$, then $\text{mrs}(x_{t_1}^k, x_{t_2}^{k'}) \leq 1$ and $\text{mrs}(x_{t_3}^{k'}, x_{t_4}^k) \leq 1$. Thus, we obtain an implication *across* observations:

$$x^k_{t_1} > x^{k'}_{t_2} \text{ and } x^{k'}_{t_3} > x^k_{t_4} \Longrightarrow \frac{p^k_{t_1}}{p^{k'}_{t_2}} \frac{p^{k'}_{t_3}}{p^k_{t_4}} \leq 1.$$

This also deserves to be called "downward-sloping demand." The larger consumption in periods t_1 and t_3, compared to t_2 and t_4, must be explained by cheaper prices in these periods.

Note the role of balanced collections of pairs: if each k appears as k_i (on the left of the pair) the same number of times it appears as k'_i (on the right), then we are able to derive an implication of EDU rationality. In the example above, we had $n = 2$ and the two observations were x^k and $x^{k'}$. We rearranged the marginal rates of substitutions from each to obtain additional implications of the model.

When $K = 1$ the model only has "within-observation" implications, and is characterized by all sequences having the downward-sloping demand property. Now, with $K \geq 1$, there are additional "across-observations" implications. Such implications derive from rearranging a collection of n observations. The relevant condition, or axiom, is that *any balanced sequence has the downward-sloping demand property.* As a corollary of the main theorem, we can show the following result.

PROPOSITION 1: *A dataset is EDU rational with $\delta = 1$ if and only if any balanced sequence has the downward-sloping demand property.*

We omit the proof of Proposition 1.

C. General K and δ

We now turn to the main situation of interest, when K can be arbitrary and δ is unknown. We first considered $K = \delta = 1$, and saw that it was enough to consider all possible "within-observation" marginal rates of substitutions: EDU rationality is characterized by downward-sloping demand. When $K \geq 1$ we saw that we needed to impose downward-sloping demand for balanced sequences, so as to capture the "across-observations" implications of EDU. When δ is unknown we need to further restrict the sequences that are required to satisfy downward-sloping demand.

When δ is unknown, larger consumption in one period *may not* be justified by lower prices. An agent can consume more in period t than in period t', even when the price of consumption is higher in period t, simply because t is sooner than t'. Consumption in t' is less valuable than in t by the effect of discounting. The relevant notion of "downward-sloping demand" is that if consumption is larger in *later* periods, then it must be explained by cheaper prices. Thus $\sum_{i=1}^n t_i \geq \sum_{i=1}^n t'_i$ as a condition in SAR-EDU.

In SAR-EDU, $\sum_{i=1}^n t_i \geq \sum_{i=1}^n t'_i$ means that the consumption quantities $x^{k_i}_{t_i}$ occur later in time than the quantities $x^{k'_i}_{t'_i}$. If we assume that the $x^{k_i}_{t_i}$ quantities are always larger than the $x^{k'_i}_{t'_i}$ quantities, then the explanation for such *larger* consumption quantities in *later* periods must be cheaper prices.

As in Sections II.A and II.B, a key idea behind Theorem 1 is to control the effects of the unknowns u and δ, by focusing on particular configurations of the data. For example, consider two observations x^{k_1} and x^{k_2} and choose four points in time, t_1, t_2, t_3, and t_4. By rearranging marginal rates of substitution we obtain that:

$$\begin{aligned}\operatorname{mrs}(x^{k_1}_{t_1}, x^{k_1}_{t_4}) \cdot \operatorname{mrs}(x^{k_2}_{t_3}, x^{k_2}_{t_2}) &= \frac{u'(x^{k_1}_{t_1})}{u'(x^{k_2}_{t_2})} \cdot \frac{u'(x^{k_2}_{t_3})}{u'(x^{k_1}_{t_4})} \\ &= \left(\frac{\delta^{t_2}}{\delta^{t_1}} \frac{p^{k_1}_{t_1}}{p^{k_2}_{t_2}}\right) \cdot \left(\frac{\delta^{t_4}}{\delta^{t_3}} \frac{p^{k_2}_{t_3}}{p^{k_1}_{t_4}}\right) \\ &= \delta^{(t_2+t_4)-(t_1+t_3)} \frac{p^{k_1}_{t_1}}{p^{k_2}_{t_2}} \frac{p^{k_2}_{t_3}}{p^{k_1}_{t_4}}.\end{aligned}$$

Notice that the pairs $(x^{k_1}_{t_1}, x^{k_2}_{t_2})$ and $(x^{k_2}_{t_3}, x^{k_1}_{t_4})$ constitute a balanced sequence of pairs because they arise from a rearrangement of two marginal rates of substitutions, one taken from observation k_1 and another from k_2. Now, if $x^{k_1}_{t_1} > x^{k_2}_{t_2}$ and $x^{k_2}_{t_3} > x^{k_1}_{t_4}$ then the concavity of u implies that the product $\delta^{(t_2+t_4)-(t_1+t_3)}(p^{k_1}_{t_1}/p^{k_2}_{t_2})(p^{k_2}_{t_3}/p^{k_1}_{t_4})$ cannot exceed 1.

Suppose now that the four points in time were chosen so that $t_1 + t_3 \geq t_2 + t_4$. Then the discount factor unambiguously increases the value on the left hand side: $\delta^{(t_2+t_4)-(t_1+t_3)} \geq 1$ for any $\delta \in (0, 1]$. Thus $(p^{k_1}_{t_1}/p^{k_2}_{t_2})(p^{k_2}_{t_3}/p^{k_1}_{t_4})$ cannot exceed 1. The punchline is that $(p^{k_1}_{t_1}/p^{k_2}_{t_2})(p^{k_2}_{t_3}/p^{k_1}_{t_4}) \leq 1$ follows from knowledge (or imposition) of the concavity of u and $\delta \in (0, 1]$. In this fashion we again obtain an implication of EDU for prices, an observable entity.

The argument just made extends to arbitrary balanced sequences, and essentially gives the proof of the necessity of the SAR-EDU in Theorem 1. The argument simply amounts to verifying a rather basic consequence of EDU: the consequence of EDU for those situations in which unobservables either do not matter or have a known effect (the effect either resulting from u' being decreasing or from $\delta \in (0, 1]$).

What is surprising is that such a basic consequence of the theory is sufficient as well as necessary. The proof of Theorem 1 is in Appendix A. Necessity is, as we have remarked, simple, and follows along the lines described above. The proof of sufficiency is more complicated and follows ideas introduced in Echenique and Saito (2015). We start from first-order conditions, as in the discussion leading up to SAR-EDU. These can be formulated as "Afriat inequalities" (Afriat, 1967), as in many studies of revealed preference. The problem here is that the Afriat inequalities are non-linear, and must be linearized. A key result is then an approximation result, which is complicated because the unknown quantities in the Afriat inequalities take values in a non-compact set.

D. Additional Models

The exposition so far has emphasized EDU, the canonical model of intertemporal choice. We now turn to other important models that can be analyzed through our techniques. Some of these models will turn out to be quite important in our empirical applications.

The first model is time-separable utility. It is the most general class of utility functions we consider, and a natural benchmark to understand the empirical failures of EDU. We will want to know when a dataset fails EDU simply because it is not TSU-rational, and when it fails other aspects of EDU.

The *time-separable utility* (TSU) model is the class of utility functions U for which there exists concave and strictly increasing functions $u_t : \mathbf{R}_+ \to \mathbf{R}$, for $t \in T$ such that $U(x) = \sum_{t \in T} u_t(x_t)$.

Strong Axiom of Revealed Time-Separable Utility (SAR-TSU): *For any balanced sequence of pairs* $(x_{t_i}^{k_i}, x_{t'_i}^{k'_i})_{i=1}^n$*, if* $t_i = t'_i$ *for all* i*, then the sequence has the downward-sloping demand property.*

SAR-TSU imposes the downward-sloping demand property on fewer sequences than those constrained by SAR-EDU or SAR-QHD. Note that, in contrast with EDU and QHD, TSU imposes no *within-observation* constraints. All the constraints must be *across* observations (and within periods). This means, for example, that all datasets with $K = 1$ are TSU rational. The across-observations constraints are also present in SAR-EDU and SAR-QHD to reflect that EDU and QHD are time-separable models. But SAR-EDU and SAR-QHD have additional within-observation constraints.

It is easy to observe that TSU model can be seen as the state-dependent utility (SDU) model of choice under uncertainty if we reinterpret the set of periods as the set of states. Echenique and Saito (2015) characterize SDU model by Strong Axiom of Revealed State-Dependent Utility. This axiom is equivalent to SAR-TSU under the reinterpretation of the set of periods as the set of states.

Finally, we turn to two special cases of models that we have already considered. One special case is *monotone time discounting* (MTD): the class of utility functions $MTD \subset GTD$ for which $D(t)$ is a monotone decreasing sequence. MTD includes the models of diminishing impatience and strong diminishing impatience (Chakraborty, Halevy and Saito, forthcoming; Halevy, 2008).

Strong Axiom of Revealed Monotone Time Discounted Utility (SAR-MTD): *For any balanced sequence of pairs* $(x_{t_i}^{k_i}, x_{t'_i}^{k'_i})_{i=1}^n$*, if there is a permutation* π *of* $\{1, 2, \ldots, n\}$ *such that* $t_i \geq t'_{\pi(i)}$*, then the sequence has the downward-sloping demand property.*

The other special case is *present-biased QHD*: the class of utility functions $PQHD \subset QHD$ in which $\beta \leq 1$.

Strong Axiom of Revealed Quasi-Hyperbolic Present-Biased Utility (SAR-PQHD): *For any balanced sequence of pairs* $(x_{t_i}^{k_i}, x_{t'_i}^{k'_i})_{i=1}^n$*, if*

(i) $\sum_{i=1}^n t_i \geq \sum_{i=1}^n t'_i$ *and*

(ii) $\#\{i : t_i > 0\} \geq \#\{i : t'_i > 0\}$,

then the sequence has the downward-sloping demand property.

THEOREM 3: *For $M' \in \{TSU, MTD, PQHD\}$, a dataset is M'-rational if and only if it satisfies SAR-M'.*

E. Discussion

The axiomatizations in Theorems 0 to 3 serve three different purposes. First, they describe the behaviors that are consistent with EDU and its generalizations; as we have seen these behaviors involve versions of a "qualified" downward-sloping demand property. Second, the precise form of the qualifications involved reflects how each theory imposes weaker or stronger properties. Thus GTD requires the property to hold for a balanced sequence of pairs where each t appears the same number of times on the left and of the right of each pair, while EDU requires if for sequences that satisfy a weaker property. These restrictions make sense given the functional for in each representation. For GTD all time periods have the same "standing," while in EDU the amount of time that has elapsed since the first time period matters. Third, the axioms function as nonparametric tests. As we shall see, for practical purposes it is often more convenient to analyze the underlying Afriat inequalities. But the axioms still describe useful simple patterns of violations that help categorize how the data violates EDU: see the results in Section III.C.

III. Empirical Illustration

We use our theoretical framework to analyze data from two experiments: Andreoni and Sprenger (2012, AS) and Carvalho, Meier and Wang (2016, CMW). Section III.A presents a quick summary of the data from these experiments.

The two experiments differ in the number of subjects and the number of choices made by a subject. The number of subjects is 97 for AS, and over 1,000 for CMW. The number of questions asked is also quite different, with AS asking subjects to choose in 45 different situations, and CMW asking for 12 choices. Despite such differences, CMW follow AS's basic design, and our methods are directly applicable to data from either experiment.

A. Description of the Data

AS introduce an experimental method called the Convex Time Budget (CTB). The CTB design fits our framework very well. In AS's experiment, subjects were asked to allocate 100 experimental tokens between "sooner" (time τ) and "later" (time $\tau + d$) accounts. Tokens allocated to each account had a value of a_τ and $a_{\tau+d}$, converting experimental currency unit into real monetary value for final payments. The gross interest rate over d days is given by $a_{\tau+d}/a_\tau$. There were three possible sooner dates $\tau \in \{0, 7, 35\}$, three possible delays $d \in \{35, 70, 98\}$ (both in days), and five different pairs of conversion rates $(a_\tau, a_{\tau+d})$ for each (τ, d) pair. Each subject completed 45 decisions. See Figure B.1 in the online appendix for an illustration.

Each subject's decision in a trial is characterized by a tuple $(\tau, d, a_\tau, a_{\tau+d}, c_\tau)$: the first four elements $(\tau, d, a_\tau, a_{\tau+d})$ characterize the budget set she faces in this trial, and c_τ is the number of tokens she decides to allocate to the sooner payment. In the experiment, subjects make a two-period choice. They choose $(x_\tau, x_{\tau+d})$ subject to $p_\tau x_\tau + x_{\tau+d} = I$. We need to formulate the problem as choosing $(x_0, \ldots, x_T)$ subject to $\sum_{t \in T} p_t x_t = I$. We set prices to be $p_\tau = a_{\tau+d}/a_\tau$ and $p_{\tau+d} = 1$ (a normalization), and we define consumptions (monetary amounts) $x_\tau = c_\tau a_\tau$ and $x_{\tau+d} = (100 - c_\tau)a_{\tau+d}$. We shall implicitly set the prices of periods that are not offered to be very high so that agents choose zero consumption in those periods. We present a more detailed explanation in online appendix B.

Two features of the CTB design make their experiment ideal for our exercise. First and most importantly, the experimental setup is precisely the situation our model tries to capture: subjects choose an intertemporal consumption from a budget set. Secondly, the CTB design has subjects *committing* to a payoff stream. Recall that to test for QHD and more general models (although not for EDU) we need to assume that agents commit to a consumption stream. In the CTB design, the commitment assumption is satisfied.

CMW administered incentivized intertemporal choice tasks on an internet panel with respondents aged 18 and over living in the United States. Subjects in CMW's experiment were asked to allocate \$500 into two payments with pre-specified dates, the second of which included interest. The sooner payment date (τ) was either now or in four weeks. The delay length (d) was either four weeks or eight weeks. The four interest rates used in the survey were 0%, 0.5%, 1%, and 3%. Each subject made 12 decisions in total.

Before reporting the results of our empirical analysis, we briefly describe our empirical methods. From the experimental datasets, we set up a linear programming problem so that finding a solution to the problem is equivalent to finding a rationalization of each model such as EDU, QHD, and TSU. We describe the method in detail in online appendix B. Note that typical CTB design does not allow us to test GARP because all budget sets are nested.

B. Results

We test whether each individual subject passes our axioms. The test is applied for all subjects in both the AS and CMW experiments. We shall label a subject as "M'-rational" if her choices pass the revealed-preference test for model M' and "M' non-rational" otherwise. The models can be ordered by the tightness of the associated axioms. Essentially, we have that

$$\begin{tikzcd}[column sep=small, row sep=small] & QHD \arrow[dr, phantom, "\subset" sloped] & \\ EDU \subset PQHD \arrow[ur, phantom, "\subset" sloped] \arrow[dr, phantom, "\subset" sloped] & & GTD \subset TSU. \\ & MTD \arrow[ur, phantom, "\subset" sloped] & \end{tikzcd}$$

For this reason, when we find that a subject is EDU rational, she is also M'-rational for all other models $M' \in \{PQHD, QHD, MTD, GTD, TSU\}$.[6]

We sometimes label a subject as "strictly M'-rational" for the most restrictive model M'

[6]QHD is not a subset of MTD due to the presence of future-biased QHD ($\beta > 1$).

Table 1—Pass Rates

Data	# subjects	# choices	EDU	PQHD	QHD	MTD	GTD	TSU
AS	97	45	0.299	0.299	0.299	0.392	0.423	0.516
CMW	1,060	12	0.210	0.216	0.222	0.278	0.299	0.433

such that the agent is M'-rational. For example, a subject is strictly QHD rational if her dataset passes the QHD test but not the EDU test.

Table 1 reports the pass rates for each model. Pass rates are the percentage of subjects in each experiment that pass the test, for each of the models: EDU, QHD, TSU, and so on.

Four aspects of Table 1 stand out. First, the numbers of EDU-rational agents are rather small: 30% and 21% in AS and CMW, respectively. Second, there is very little added scope for QHD. In the case of the AS experiment, *all* subjects rationalized as QHD are also rationalized as EDU; and there are very few subjects in the CMW experiment that are QHD, but not EDU, rational.[7][8] The number of TSU-rational subjects is 52% in AS and 43% in CMW. These numbers may be viewed as small as well.

The third aspect is that some violations of QHD are captured by MTD and GTD. As we mentioned above, MTD includes the models of diminishing impatience and strong diminishing impatience. Our methods can be applied to these models: see online appendix D.2.[9]

The fourth aspect is that the results for the AS and CMW experiments are quite similar, despite some notable differences in implementation and population sampled. The sample sizes are very different in these experiments: 97 for AS and over 1,000 for CMW. The number of questions asked is also quite different, with AS asking subjects to choose from 45 different budgets, while CMW asking only for 12 choices.[10] We found that the power of the tests applied on each data are similar (and, is in fact, very high; see online appendix C), so it makes sense to compare the pass rates in these two experiments.

C. Analysis of Violations of EDU, QHD, and TSU Rationalities

We use the theoretical results in the paper to study the violations of EDU, QHD, and TSU reported in Table 1. Our theoretical results help uncover the particular patterns in the data that underly subjects' violations of the different models.

We have discussed the empirical content of EDU, and of the special case of EDU with $\delta = 1$,

[7] Note the difference in pass rates for PQHD and QHD in CMW data. It means that there are six subjects who are rationalized by *future-biased* QHD.

[8] It is possible to analyze the QHD model where more than one period is regarded as "the present," as in $\sum_{t=0}^{\bar{\tau}} \delta^{\tau} u(x_t) + \beta \sum_{t=\bar{\tau}+1}^{T} \delta^t u(x_t)$. By varying $\bar{\tau}$, we obtain slightly more QHD-rational subjects, but the qualitative conclusions reported do not change. See online appendix D.1 for details. We thank an anonymous referee for suggesting the exercise.

[9] Arguably, diminishing impatience captures the essence of the behaviors that PQHD seeks to explain. In that sense, it is worth noting that the pass rates for MTD with diminishing impatience in the AS data are 37.1%, up from PQHD's 29.9% (out of 97 subjects). The corresponding pass rates in CMW data are 26.6% for MTD with diminishing impatience and 21.6% for PQHD (out of 1,060 subjects).

[10] One might, for example, conjecture that fatigue in an experiment with many choices could affect pass rates, but the comparison between AS and CMW gives no indication that fatigue matters.

the case where there is no discounting. It is obvious that there are datasets that are only rationalizable when we allow for $\delta < 1$; so EDU is a strictly weaker, more permissive, theory of intertemporal choice than EDU with $\delta = 1$. There are, however, conditions under which a dataset is EDU rational if and only if it is EDU rational with $\delta = 1$.

When we introduced SAR-EDU, we mentioned how relatively larger consumption could occur even if prices are relatively high, due to the role of discounting. Our notion of *strict impatience* is meant to capture this phenomenon.

DEFINITION 5: *A dataset* $(x^k, p^k)_{k=1}^K$ *is* strictly impatient *if for all balanced sequences* $(x_{t_i}^{k_i}, x_{t_i'}^{k_i'})_{i=1}^n$ *such that* $x_{t_i}^{k_i} > x_{t_i'}^{k_i'}$ *for all* i *and* $\sum_{i=1}^n t_i > \sum_{i=1}^n t_i'$,

$$\prod_{i=1}^{n} \frac{p_{t_i}^{k_i}}{p_{t_i'}^{k_i'}} < 1. \tag{4}$$

To interpret strict impatience, consider a pair $(x_t^k, x_{t'}^{k'})$ exhibiting a violation of (4). Thus, $x_t^k > x_{t'}^{k'}$, t is *later* than t', and $p_t^k / p_{t'}^{k'} \geq 1$. This means that the later consumption at date t is weakly more expensive than the sooner consumption at date t', but the agent still consumes more at the later date than at the sooner date. Such an agent cannot be strictly impatient.

The notion of strict impatience is important because many experimental subjects are not strictly impatient. To such subjects, the following result applies:

PROPOSITION 2: *Suppose that the dataset* $(x^k, p^k)_{k=1}^K$ *is not strictly impatient. Then the following statements are equivalent:*

(i) The dataset is EDU rational.

(ii) The dataset is EDU rational with $\delta = 1$.

(iii) All balanced sequences in the dataset have the downward-sloping demand property.

We find that *almost all* EDU non-rational subjects display a *particular kind* of violation of EDU. They carry out a two-pronged violation of EDU. First, their choices imply that they are *not* strictly impatient. Second, their choices fail to satisfy the downward-sloping demand property. Therefore, by Proposition 2, such subjects cannot be EDU rational. The finding holds true for both the AS and CMW data, and it is present in almost all subjects that violate EDU.

It is important to emphasize that each pattern in isolation does not imply a violation of EDU. Many subjects exhibit a violation of downward-sloping demand that is consistent with EDU. At the same time, they make choices that mean that they cannot be discounting. The *conjunction of both patterns* implies a violation of EDU.

VIOLATION OF EDU RATIONALITY. — As mentioned, we find that many subjects are not discounting—are not strictly impatient. In particular, we concentrate on the following patterns of choices:

Table 2—Number of Subjects Who Display Patterns (P1)–(P4) by Rationality

			Patterns of choices	
Data	Rationality	Total #	(P1) or (P2)	[(P1) or (P2)] and [(P3) or (P4)]
AS	EDU	29	5 (17%)	0 (0%)
	Strict QHD	0	N/A	N/A
	Strict TSU	21	13 (62%)	13 (62%)
	Non TSU	47	45 (96%)	45 (96%)
CMW	EDU	223	74 (33%)	0 (0%)
	Strict QHD	12	7 (58%)	7 (58%)
	Strict TSU	224	152 (68%)	152 (68%)
	Non TSU	601	570 (95%)	570 (95%)

Note: For each panel, the second column shows the category of strict rationality and the total number of subjects for each category. The third column shows the number of subjects who display the behaviors (P1) or (P2), and hence are not strictly impatient. The fourth column shows the number of subjects who display the behaviors "(P3) or (P4)", as well as "(P1) or (P2)".

(P1) A pair $(x^k_{t_1}, x^k_{t_2})$ with $x^k_{t_1} > x^k_{t_2}$, $t_1 > t_2$, and $p^k_{t_1}/p^k_{t_2} \geq 1$.

(P2) Pairs $((x^{k_1}_{t_1}, x^{k_2}_{t_3}), (x^{k_2}_{t_4}, x^{k_1}_{t_2}))$ with $x^{k_1}_{t_1} > x^{k_2}_{t_3}$, $x^{k_2}_{t_4} > x^{k_1}_{t_2}$, $t_1 + t_4 > t_3 + t_2$, and $(p^{k_1}_{t_1}/p^{k_2}_{t_3}) \cdot (p^{k_2}_{t_4}/p^{k_1}_{t_2}) \geq 1$.

The behaviors in (P1) and (P2) are special cases of the condition in Proposition 2, and they have a simple economic meaning. Behavior (P1) means that an agent consumes more in the more expensive period, as $x^k_{t_1} > x^k_{t_2}$ and $p^k_{t_1}/p^k_{t_2} \geq 1$. The more expensive period is later, as $t_1 > t_2$. Therefore, the agent cannot dislike later consumption, and must have $\delta = 1$, if she is to be EDU rational.

Behavior (P2) has a similar meaning, but is slightly more involved. Think of t_1 as being after t_2, and t_3 as being after t_4. Then $t_1 - t_2 > t_3 - t_4$ means that the time elapsed from the sooner to the later period is larger in observation k_1 than in observation k_2. Moreover, $p^{k_1}_{t_1}/p^{k_1}_{t_2} \geq p^{k_2}_{t_3}/p^{k_2}_{t_4}$, so the price of postponing consumption is higher in observation k_1 than in k_2. But then, if an agent chooses to consume relatively more later in observation k_1 than in k_2 ($x^{k_1}_{t_1} > x^{k_2}_{t_3}$ and $x^{k_2}_{t_4} > x^{k_1}_{t_2}$), she cannot dislike postponing consumption if she is EDU rational. Again, she must have $\delta = 1$ if she is to be EDU rational.

By Proposition 2, to test EDU rationality when a subject is not strictly impatient, all we need to check is the downward-sloping demand property. Again we concentrate on very simple violations of downward-sloping demand:

(P3) A pair $(x^k_{t_1}, x^k_{t_2})$ with $x^k_{t_1} > x^k_{t_2}$ and $p^k_{t_1}/p^k_{t_2} > 1$.

(P4) Pairs $((x^{k_1}_{t_1}, x^{k_2}_{t_3}), (x^{k_2}_{t_4}, x^{k_1}_{t_2}))$ with $x^{k_1}_{t_1} > x^{k_2}_{t_3}$, $x^{k_2}_{t_4} > x^{k_1}_{t_2}$ and $(p^{k_1}_{t_1}/p^{k_2}_{t_3}) \cdot (p^{k_2}_{t_4}/p^{k_1}_{t_2}) > 1$.

Table 2 reports the numbers of subjects that display the behaviors (P1), (P2), (P3), or (P4), for the AS and CMW experiments, and classified by the most stringent theory passed by each subject.

The results in Table 2 have interesting implications. Firstly, the fraction of subjects displaying (P1) or (P2) is not small, even for EDU-rational agents (17% in AS and 33% in CMW). By Proposition 2, such subjects are EDU rational with $\delta = 1$.

Secondly, among EDU non-rational agents, the fraction of subjects who display the behavior (P1) or (P2) is very large. In AS's data, it is 62% of strict TSU rational subjects, and 96% of TSU non-rational subjects. In CMW's data, it is 58% of strict QHD-rational subjects, 68% of strict TSU-rational subjects, and 95% of TSU non-rational subjects. It is interesting that the percentages are similar in AS and CMW. It is also interesting that the percentage increases as the "level of rationality" is lower.

Thirdly, and surprisingly, *all of the EDU non-rational subjects who display (P1) or (P2) also display the behaviors (P3) or (P4).* The numbers in the two columns for strict TSU and TSU non-rational subjects are the same because it is actually the same subjects that display (P1) or (P2), and (P3) or (P4). This means that the majority of the violations of EDU found in AS and CMW have a straightforward explanation in the data: they make some choices that are incompatible with discounting, and then they violate simple downward-sloping demand.

There is additional structure to the violations uncovered by Table 2. Many of the instances of (P3) and (P4) are not a violation of SAR-EDU. That is, in (P3), $t_1 \geq t_2$ does not hold and in (P4), $t_1 + t_4 \geq t_3 + t_2$ does not hold. This means that the subjects who display the behavior [(P1) or (P2)] and [(P3) or (P4)] are not EDU rational because they make choices that are not strictly impatient, which force $\delta = 1$, and in the second place they make choices that are incompatible with downward-sloping demand property. This suggests that, for most subjects, the lack of compliance with EDU boils down to the downward-sloping demand property.

Violation of QHD rationality and corner choices. — One consequence of Theorems 1 and 2 is that, under certain circumstances, EDU and PQHD are *observationally equivalent.* These circumstances are very relevant for the discussion of experiments in this section. Our next result, Proposition 3, shows that if an agent does not consume at the soonest date (i.e., $x_0^k = 0$ for all $k \in K$), then EDU and PQHD are observationally equivalent.

PROPOSITION 3: *Suppose that a dataset $(x^k, p^k)_{k=1}^K$ satisfies that $x_0^k = 0$ for all $k \in K$. Then $(x^k, p^k)_{k=1}^K$ is PQHD rational if and only if it is EDU rational.*

No subjects are strictly QHD-rational in the AS experiment. This is related to agents' peculiar pattern of choices. Proposition 3 shows that if an agent does not consume at the soonest date (i.e., $x_0^k = 0$ for all $k \in K$), then EDU and PQHD are observationally equivalent. In AS's experiment, more than 82.8% of the subjects who satisfy SAR-EDU (i.e., 25% of the total subjects) do not consume at the soonest date. So Proposition 3 means that QHD has no scope beyond EDU for such subjects.

Moreover, the same subjects in AS satisfy the condition in Proposition 4 (discussed in online appendix A.4.3); they did not consume a positive amount on the sooner date whenever the price for the sooner consumption is higher than the price for the later consumption. Therefore, Proposition 4 implies that those subjects are EDU rational.

The conclusion should be qualified by the findings using CMW's data. CMW's subjects chose corner allocations much less often than AS's: fewer than 10% of subjects never chose interior allocations, and more than 50% of them chose interior allocations in all 12 questions. And the number of strictly QHD rational agents in CMW is also very small, so the lack of scope for QHD in AS's data may not be driven by agents' tendency to choose corner allocations.

Violation of TSU rationality. — We have seen that many violations of EDU correspond to a simple pattern in the data (configurations (P1)-(P4) discussed above). There is also a simple pattern behind the violations of TSU.

Consider a pair of observations $k, k' \in K$ and time periods $s, t \in T$ such that $x_t^k > x_t^{k'}$, $x_s^{k'} > x_s^k$, and $p_t^k/p_s^k > p_t^{k'}/p_s^{k'}$. This is a 2×2 violation of the TSU axiom, with the two pairs $((x_t^k, x_t^{k'}), (x_s^{k'}, x_s^k))$. All the 601 TSU non-rational subjects in CMW display such a simple violation of the TSU axiom. In AS, 36 subjects (out of 47; 76.6%) exhibit the behavior.

IV. Concluding Remarks

We present revealed-preference characterizations, or tests, of the most common models of intertemporal choice: EDU, QHD, and TSU. We apply our tests to data from experiments by Andreoni and Sprenger (2012) and Carvalho, Meier and Wang (2016), and find that our axiomatization delivers new insights and perspectives on datasets that had been analyzed by parametric methods. Two experiments are different in important ways (such as the number of choices, the number of subjects, and the frequency of interior choices); still, the main findings for each experiment are surprisingly similar. The pass rates for EDU, QHD, and TSU are relatively low.

We believe that our results are useful for understanding what experimental subjects do, and also to design future experiments. For example, Proposition 2 shows that EDU rationality is equivalent to a much simpler property, the downward-sloping demand property, under natural conditions. This result is useful to understand the behaviors of experimental subjects, as we have shown in Section III.C. Proposition 3 is helpful in designing experiments to distinguish EDU subjects from QHD subjects. The proposition tells us that we need to set prices for the soonest consumption (i.e., consumption at period 0) cheap enough so that a subject will choose positive consumption at the soonest date. Proposition 4 (online appendix A.4.3), which shows that corner choices lead to EDU rationality, is important for CTB experiments. To test EDU rationality in a meaningful way, one should have enough variety of prices so that a subject may not choose corner allocations. In a similar way, Proposition 5 (online appendix A.4.4), which shows the equivalence between EDU and TSU rationality under certain conditions, can also be important in designing experiments to distinguish EDU from TSU subjects.

We consider a few additional issues in the online appendix. The important issue of the power of the tests is discussed in online appendix C. The robustness of revealed-preference tests to small perturbations in the data is considered in online appendix F.

REFERENCES

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

Aguiar, Victor H, and Nail Kashaev. 2018. "Stochastic Revealed Preferences with Measurement Error." Unpublished manuscript.

Andreoni, James, and Charles Sprenger. 2012. "Estimating Time Preferences from Convex Budgets." *American Economic Review*, 102(7): 3333–3356.

Andreoni, James, Michael A. Kuhn, and Charles Sprenger. 2015. "Measuring Time Preferences: A Comparison of Experimental Methods." *Journal of Economic Behavior & Organization*, 116: 451–464.

Anscombe, Francis J., and Robert J. Aumann. 1963. "A Definition of Subjective Probability." *Annals of Mathematical Statistics*, 34(1): 199–205.

Augenblick, Ned, Muriel Niederle, and Charles Sprenger. 2015. "Working Over Time: Dynamic Inconsistency in Real Effort Tasks." *Quarterly Journal of Economics*, 130(3): 1067–1115.

Balakrishnan, Uttara, Johannes Haushofer, and Pamela Jakiela. forthcoming. "How Soon Is Now? Evidence of Present Bias from Convex Time Budget Experiments." *Experimental Economics*.

Barcellos, Silvia Helena, and Leandro Carvalho. 2014. "Information about Self-Control and Intertemporal Choices." Unpublished manuscript.

Bleichrodt, Han, Kirsten I. M. Rohde, and Peter P. Wakker. 2008. "Koopmans' Constant Discounting for Intertemporal Choice: A Simplification and a Generalization." *Journal of Mathematical Psychology*, 52(6): 341–347.

Border, Kim C. 2013. "Alternative Linear Inequalities." mimeo, Caltech.

Brocas, Isabelle, Juan D. Carrillo, and Jorge Tarrasó. 2018. "How Long is a Minute?" *Games and Economic Behavior*, 111: 305–322.

Browning, Martin. 1989. "A Nonparametric Test of the Life-Cycle Rational Expectations Hypothesis." *International Economic Review*, 30(4): 979–992.

Carvalho, Leandro, Silvia Prina, and Justin Sydnor. 2016. "The Effects of Savings on Risk Attitudes and Intertemporal Choices." *Journal of Development Economics*, 120: 41–52.

Carvalho, Leandro, Stephan Meier, and Stephanie W. Wang. 2016. “Poverty and Economic Decision Making: Evidence from Changes in Financial Resources at Payday.” *American Economic Review*, 106(2): 260–284.

Chakraborty, Anujit, Yoram Halevy, and Kota Saito. forthcoming. “The Relation between Behavior under Risk and over Time.” *American Economic Review: Insights.*

Chambers, Christopher P., and Federico Echenique. 2014. “On the Consistency of Data with Bargaining Theories.” *Theoretical Economics*, 9(1): 137–162.

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

Crawford, Ian, and Matthew Polisson. 2014. “Testing for Intertemporal Nonseparability.” *Journal of Mathematical Economics*, 52: 46–49.

Dziewulski, Pawel. 2018. “Revealed Time-Preference.” *Games and Economic Behavior*, 118: 67–77.

Echenique, Federico, and Kota Saito. 2015. “Savage in the Market.” *Econometrica*, 83(4): 1467–1495.

Ericson, Keith Marzilli, and David Laibson. 2019. “Intertemporal Choice.” In *Handbook of Behavioral Economics - Foundations and Applications*. Vol. 2, , ed. B. Douglas Bernheim, Stefano DellaVigna and David Laibson, 1–67. Amsterdam:Elsevier/North-Holland.

Fishburn, Peter, and Ward Edwards. 1997. “Discount-Neutral Utility Models for Denumerable Time Streams.” *Theory and Decision*, 43(2): 139–166.

Fishburn, Peter C., and Ariel Rubinstein. 1982. “Time Preference.” *International Economic Review*, 23(3): 677–694.

Galperti, Simone, and Bruno H. Strulovici. 2017. “A Theory of Intergenerational Altruism.” *Econometrica*, 85(4): 1175–1218.

Gilboa, Itzhak. 1989. “Expectation and Variation in Multi-Period Decisions.” *Econometrica*, 57(5): 1153–1169.

Giné, Xavier, Jessica Goldberg, Dan Silverman, and Dean Yang. 2018. “Revising Commitments: Time Preference and Time-Inconsistency in the Field.” *Economic Journal*, 128(608): 159–188.

Halevy, Yoram. 2008. “Strotz Meets Allais: Diminishing Impatience and the Certainty Effect.” *American Economic Review*, 98(3): 1145–1162.

Hayashi, Takashi. 2003. “Quasi-Stationary Cardinal Utility and Present Bias.” *Journal of Economic Theory*, 112(2): 343–352.

Janssens, Wendy, Berber Kramer, and Lisette Swart. 2017. “Be Patient When Measuring Hyperbolic Discounting: Stationarity, Time Consistency and Time Invariance in a Field Experiment.” *Journal of Development Economics*, 126: 77–90.

Koopmans, Tjalling C. 1960. “Stationary Ordinal Utility and Impatience.” *Econometrica*, 28(2): 287–309.

Kopylov, Igor. 2010. “Simple Axioms for Countably Additive Subjective Probability.” *Journal of Mathematical Economics*, 46(5): 867–876.

Kuhn, Michael, Peter Kuhn, and Marie Claire Villeval. 2017. “Decision-Environment Effects on Intertemporal Financial Choices: How Relevant are Resource-Depletion Models?” *Journal of Economic Behavior & Organization*, 137: 72–89.

Laibson, David. 1997. “Golden Eggs and Hyperbolic Discounting.” *Quarterly Journal of Economics*, 112(2): 443–478.

Liu, Elaine M., Juanjuan Meng, and Joseph Tao-Yi Wang. 2014. “Confucianism and Preferences: Evidence from Lab Experiments in Taiwan and China.” *Journal of Economic Behavior & Organization*, 104: 106–122.

Lührmann, Melanie, Marta Serra-Garcia, and Joachim Winter. 2018. “The Impact of Financial Education on Adolescents’ Intertemporal Choices.” *American Economic Journal: Economic Policy*, 10(3): 309–332.

Montiel Olea, José Luis, and Tomasz Strzalecki. 2014. “Axiomatization and Measurement of Quasi-Hyperbolic Discounting.” *Quarterly Journal of Economics*, 129(3): 1449–1499.

Phelps, E. S., and Robert A. Pollak. 1968. “On Second-Best National Saving and Game-Equilibrium Growth.” *Review of Economic Studies*, 35(2): 185–199.

Quah, John K.-H. 2014. “A Test for Weakly Separable Preferences.” Unpublished manuscript.

Rockafellar, R. Tyrell. 1997. *Convex Analysis.* Princeton University Press.

Sawada, Yasuyuki, and Yusuke Kuroishi. 2015. “How Does a Natural Disaster Affect People’s Preferencec? An Investigation of a Large Scale Flood in the Philippines using Convex Time Budget Experiments.” Unpublished manuscript.

Stoer, Josef, and Christoph Witzgall. 1970. *Convexity and Optimization in Finite Dimensions.* Springer-Verlag Berlin.

Sun, Chen, and Jan Potters. 2016. “Magnitude Effect in Intertemporal Allocation Tasks.” Unpublished manuscript.

Varian, Hal R. 1983. “Non-Parametric Tests of Consumer Behaviour.” *Review of Economic Studies*, 50(1): 99–110.

Weibull, Jörgen W. 1985. “Discounted-Value Representations of Temporal Preferences.” *Mathematics of Operations Research*, 10(2): 244–250.

Appendix A: Proof of Theorem 1

We present the proof of the equivalence between EDU rationality and SAR-EDU.

The proof is based on using the first-order conditions for maximizing a utility with the EDU over a budget set. Our first lemma ensures that we can without loss of generality restrict attention to the first-order conditions. The proof of the lemma is the same as that of Lemma 7 in Echenique and Saito (2015), with the change of $\{\mu_s\}_{s\in S}$ to $\{\delta^t\}_{t\in T}$ (in Echenique and Saito (2015), μ_s is the subjective probability that state s realizes).

We use the following notation in the proofs: $\mathcal{X} = \{x_t^k : k \in K, t \in T\}$.

LEMMA 1: *Let $(x^k, p^k)_{k=1}^K$ be a dataset. The following statements are equivalent:*

(a) $(x^k, p^k)_{k=1}^K$ is EDU rational.

(b) There are strictly positive numbers v_t^k, λ^k, and $\delta \in (0,1]$, for $t = 1, \ldots, T$ and $k = 1, \ldots, K$, such that

$$\delta^t v_t^k = \lambda^k p_t^k, \quad x_t^k > x_{t'}^{k'} \Longrightarrow v_t^k \le v_{t'}^{k'}.$$

A1. Necessity

LEMMA 2: *If a dataset $(x^k, p^k)_{k=1}^K$ is EDU rational, then it satisfies SAR-EDU.*

PROOF:

Let $(x^k, p^k)_{k=1}^K$ be EDU rational, and let $\delta \in (0,1]$ and $u : \mathbf{R}_+ \to \mathbf{R}$ be as in the definition of EDU rationality. By Lemma 1, there exists a strictly positive solution v_t^k, λ^k, δ to the system in statement (b) of Lemma 1 with $v_t^k \in \partial u(x_t^k)$ when $x_t^k > 0$, and $v_t^k \ge \underline{w} \in \partial u(x_t^k)$ when $x_t^k = 0$.

Let $(x_{t_i}^{k_i}, x_{t'_i}^{k'_i})_{i=1}^n$ be a balanced sequence satisfying the conditions in SAR-EDU. Then $\sum_{i=1}^n t_i \ge \sum_{i=1}^n t'_i$ and $x_{t_i}^{k_i} > x_{t'_i}^{k'_i}$ for all i. Suppose that $x_{t'_i}^{k'_i} > 0$. Then, $v_{t_i}^{k_i} \in \partial u(x_{t_i}^{k_i})$ and $v_{t'_i}^{k'_i} \in \partial u(x_{t'_i}^{k'_i})$. By the concavity of u, it follows that $\lambda^{k_i} \delta^{t'_i} p_{t_i}^{k_i} \le \lambda^{k'_i} \delta^{t_i} p_{t'_i}^{k'_i}$ (see Theorem 24.8 of Rockafellar, 1997). Similarly, if $x_{t'_i}^{k'_i} = 0$, then $v_{t_i}^{k_i} \in \partial u(x_{t_i}^{k_i})$ and $v_{t'_i}^{k'_i} \ge \underline{w} \in \partial u(x_{t'_i}^{k'_i})$. Hence $\lambda^{k_i} \delta^{t'_i} p_{t_i}^{k_i} \le \lambda^{k'_i} \delta^{t_i} p_{t'_i}^{k'_i}$. Therefore,

$$1 \ge \prod_{i=1}^n \frac{\lambda^{k_i} \delta^{t'_i} p_{t_i}^{k_i}}{\lambda^{k'_i} \delta^{t_i} p_{t'_i}^{k'_i}} = \frac{1}{\delta^{(\sum t_i - \sum t'_i)}} \prod_{i=1}^n \frac{p_{t_i}^{k_i}}{p_{t'_i}^{k'_i}} \ge \prod_{i=1}^n \frac{p_{t_i}^{k_i}}{p_{t'_i}^{k'_i}},$$

as the sequence is balanced and satisfies the condition in SAR-EDU, i.e., $\sum_{i=1}^n t_i \ge \sum_{i=1}^n t'_i$ and the numbers λ^k appear the same number of times in the denominator as in the numerator of this product.

A2. Theorem of the Alternative

To prove sufficiency, we shall use the following lemma, which is a version of the Theorem of the Alternative. This is Theorem 1.6.1 in Stoer and Witzgall (1970). We shall use it here in the cases where $\mathbf{F}$ is either the real or the rational numbers.

LEMMA 3: *Let A be an $m \times n$ matrix, B be an $l \times n$ matrix, and E be an $r \times n$ matrix. Suppose that the entries of the matrices A, B, and E belong to the commutative ordered field $\mathbf{F}$. Exactly one of the following alternatives is true.*

1) *There is $u \in \mathbf{F}^n$ such that $A \cdot u = 0$, $B \cdot u \geq 0$, $E \cdot u \gg 0$.*
2) *There is $\theta \in \mathbf{F}^r$, $\eta \in \mathbf{F}^l$, and $\pi \in \mathbf{F}^m$ such that $\theta \cdot A + \eta \cdot B + \pi \cdot E = 0$; $\pi > 0$ and $\eta \geq 0$.*

We also use the following lemma, which follows from Lemma 3 (see Border (2013) or Chambers and Echenique (2014)):

LEMMA 4: *Let A be an $m \times n$ matrix, B be an $l \times n$ matrix, and E be an $r \times n$ matrix. Suppose that the entries of the matrices A, B, and E are rational numbers. Exactly one of the following alternatives is true.*

1) *There is $u \in \mathbf{R}^n$ such that $A \cdot u = 0$, $B \cdot u \geq 0$, and $E \cdot u \gg 0$.*
2) *There is $\theta \in \mathbf{Q}^r$, $\eta \in \mathbf{Q}^l$, and $\pi \in \mathbf{Q}^m$ such that $\theta \cdot A + \eta \cdot B + \pi \cdot E = 0$; $\pi > 0$ and $\eta \geq 0$.*

A3. Sufficiency

We proceed to prove the sufficiency direction. An outline of the argument is as follows. We know from Lemma 1 that it suffices to find a solution to the Afriat inequalities (actually first-order conditions), written as statement (b) in the lemma. We set up the problem to find a solution to a system of linear inequalities obtained from using logarithms to linearize the Afriat inequalities in Lemma 1.

Lemma 5 establishes that SAR-EDU is sufficient for SEU rationality when the logarithms of the prices are rational numbers. The role of rational logarithms comes from our use of a version of the theorem of the alternative (Lemma 4).

The next step in the proof (Lemma 6) establishes that we can approximate any dataset satisfying SAR-EDU with a dataset for which the logarithms of prices are rational, and for which SAR-EDU is satisfied. This step is crucial, and somewhat delicate.[11]

Finally, Lemma 7 establishes the result by using another version of the theorem of the alternative, stated as Lemma 3 above.

The statement of the lemmas follow. The rest of the paper is devoted to the proof of these lemmas.

[11]One might have tried to obtain a solution to the Afriat inequalities for "perturbed" systems (with prices that are rational after taking logs), and then considered the limit. This does not work because the solutions to our systems of inequalities are in a non-compact space. It is not clear how to establish that the limits exist and are well-behaved. Lemma 6 avoids the problem.

LEMMA 5: *Let data* $(x^k, p^k)_{k=1}^k$ *satisfy SAR-EDU. Suppose that* $\log(p_t^k) \in \mathbf{Q}$ *for all* k *and* t. *Then there are numbers* v_t^k, λ^k, δ, *for* $t \in T$ *and* $k = 1, \ldots, K$ *satisfying (b) in Lemma 1.*

LEMMA 6: *Let data* $(x^k, p^k)_{k=1}^k$ *satisfy SAR-EDU. Then for all positive numbers* $\bar{\varepsilon}$, *there exists* $q_t^k \in [p_t^k - \bar{\varepsilon}, p_t^k]$ *for all* $t \in T$ *and* $k \in K$ *such that* $\log q_t^k \in \mathbf{Q}$ *and the dataset* $(x^k, q^k)_{k=1}^k$ *satisfy SAR-EDU.*

LEMMA 7: *Let data* $(x^k, p^k)_{k=1}^k$ *satisfy SAR-EDU. Then there are numbers* v_t^k, λ^k, δ, *for* $t \in T$ *and* $k = 1, \ldots, K$ *satisfying (b) in Lemma 1.*

The proofs of Lemma 6 and 7 are similar to the proofs of Lemmas 12 and 13 in Echenique and Saito (2015). The proofs are in the online appendix.

A4. Proof of Lemma 5

We linearize the equation in statement (b) of Lemma 1. The result is:

$$\log v_t^k + t \log \delta - \log \lambda^k - \log p_t^k = 0, \tag{A1}$$

$$x_t^k > x_{t'}^{k'} \Longrightarrow v_{t'}^{k'} \geq v_t^k, \tag{A2}$$

$$\log \delta \leq 0. \tag{A3}$$

In the system comprised by (A1), (A2), and (A3), the unknowns are the real numbers $\log v_t^k$, $\log \delta$, $k \in K$, and $t \in T$.

First, we are going to write the system of inequalities (A1) and (A2) in a matrix form. We shall define a matrix A such that there are positive numbers v_t^k, λ^k, δ, the logs of which satisfy equation (A1) if and only if there is a solution $w \in \mathbf{R}^{K\times(T+1)+1+K+1}$ to the system of equations

$$A \cdot w = 0,$$

and for which the last component of w is strictly positive.

Let A be a matrix with $K \times (T+1) + 1 + K + 1$ columns, defined as follows: We have one row for every pair (k, t); one column for every pair (k, t); one column for each k; and two additional columns. Organize the columns so that we first have the $K \times (T+1)$ columns for the pairs (k, t), then one of the single columns mentioned in last place, which we shall refer to as the δ-column, then K columns (one for each k), and finally one last column. In the row corresponding to (k, t) the matrix has zeroes everywhere with the following exceptions: it has a 1 in the column for (k, t), t in the δ column, -1 in the column for k, and $-\log p_t^k$ in the very last column.

Thus, matrix A looks as follows:

$$\begin{array}{c} \\ \vdots \\ (k,t) \\ \vdots \end{array}
\begin{array}{c}
\begin{array}{ccccccccccccc} (1,0) & \cdots & (k,t) & \cdots & (K,T) & \delta & 1 & \cdots & k & \cdots & K & p \end{array} \\
\left[\begin{array}{ccccc|c|ccccc|c}
\vdots & & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots \\
0 & \cdots & 1 & \cdots & 0 & t & 0 & \cdots & -1 & \cdots & 0 & -\log p_t^k \\
\vdots & & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots
\end{array}\right]
\end{array}.$$

Consider the system $A \cdot w = 0$. If there are numbers solving equation (A1), then these define a solution $w \in \mathbf{R}^{K\times(T+1)+1+K+1}$ for which the last component is 1. If, on the other hand, there is a solution $w \in \mathbf{R}^{K\times(T+1)+1+K+1}$ to the system $A \cdot w = 0$ in which the last component is strictly positive, then by dividing through by the last component of w we obtain numbers that solve equation (A1).

In second place, we write the system of inequalities (A2) and (A3) in matrix forms. Let B be a matrix with $K \times (T+1) + 1 + K + 1$ columns. Define B as follows: One row for every pair (k,t) and (k',t') with $x_t^k > x_{t'}^{k'}$; in the row corresponding to (k,t) and (k',t') we have zeroes everywhere with the exception of a -1 in the column for (k,t) and a 1 in the column for (k',t'). These rows captures the inequality (A2). Finally, in the last row, we have zeroes everywhere with the exception of a -1 at $K \times (T+1) + 1$th column. We shall refer to this last row as the δ-row, which capturing the inequality (A3).

In the third place, we have a matrix E that captures the requirement that the last component of a solution be strictly positive. The matrix E has a single row and $K\times(T+1)+1+K+1$ columns. It has zeroes everywhere except for 1 in the last column.

To sum up, there is a solution to the system (A1), (A2), and (A3) if and only if there is a vector $w \in \mathbf{R}^{K\times(T+1)+1+K+1}$ that solves the system of equations and linear inequalities:

$$(S1): \ A \cdot w = 0, \ B \cdot w \geq 0, \ E \cdot w \gg 0.$$

The entries of A, B, and E are integer numbers, with the exception of the last column of A. Under the hypothesis of the lemma we are proving, the last column consists of rational numbers.

By Lemma 4, then, there is such a solution w to $S1$ if and only if there is no rational vector (θ, η, π) that solves the system of equations and linear inequalities

$$(S2): \ \theta \cdot A + \eta \cdot B + \pi \cdot E = 0, \ \eta \geq 0, \ \pi > 0.$$

In the following, we shall prove that the non-existence of a solution w implies that the data must violate SAR-EDU. Suppose then that there is no solution w and let (θ, η, π) be a rational vector as above, solving system $S2$.

By multiplying (θ, η, π) by any positive integer we obtain new vectors that solve $S2$, so we can take (θ, η, π) to be integer vectors.

Henceforth, we use the following notational convention: For a matrix D with $K\times(T+1)+1+K+1$ columns, write D_1 for the submatrix of D corresponding to the first $K \times (T+1)$ columns, let D_2 be the submatrix corresponding to the following one column (i.e., δ-column), D_3 correspond to the next K columns, and D_4 to the last column. Thus, $D = [D_1|D_2|D_3|D_4]$.

CLAIM 1: *(i)* $\theta\cdot A_1+\eta\cdot B_1 = 0$*; (ii)* $\theta\cdot A_2+\eta\cdot B_2 = 0$*; (iii)* $\theta\cdot A_3 = 0$*; and (iv)* $\theta\cdot A_4+\pi\cdot E_4 = 0$.

Proof. Since $\theta \cdot A + \eta \cdot B + \pi \cdot E = 0$, then $\theta \cdot A_i + \eta \cdot B_i + \pi \cdot E_i = 0$ for all $i = 1, \ldots, 4$. Moreover, since B_3, B_4, E_1, E_2, and E_3 are zero matrices, we obtain the claim. □

For convenience, we transform the matrices A and B using θ and η. We transform the matrices A and B as follows. Let us define a matrix A^* from A by letting A^* have $K \times (T+$

$1) + 1 + K + 1$ columns that consists of the rows as follows: for each row in $r \in A$ (i) have θ_r copies of the rth row when $\theta_r > 0$; (ii) omit row r when $\theta_r = 0$; and (iii) have θ_r copies of the rth row multiplied by -1 when $\theta_r < 0$.

We refer to rows that are copies of some r in A with $\theta_r > 0$ as *original* rows. We refer to rows that are copies of some r in A with $\theta_r < 0$ as *converted* rows.

Similarly, we define the matrix B^* from B by including the same columns as B and η_r copies of each row (and thus omitting row r when $\eta_r = 0$; recall that $\eta_r \geq 0$ for all r).

CLAIM 2: *For any (k,t), all the entries in the column for (k,t) in A_1^* are of the same sign.*

Proof. By definition of A, the column for (k,t) will have zero in all its entries with the exception of the row for (k,t). In A^*, for each (k,t), there are three mutually exclusive possibilities: the row for (k,t) in A can (i) not appear in A^*, (ii) it can appear as original, or (iii) it can appear as converted. This shows the claim. $\square$

CLAIM 3: *There exists a sequence of pairs $(x_{t_i}^{k_i}, x_{t_i'}^{k_i'})_{i=1}^{n^*}$ that satisfies a condition in SAR-EDU: $x_{t_i}^{k_i} > x_{t_i'}^{k_i'}$ for all $i = 1, \ldots, n^*$.*

Proof. We define such a sequence by induction. Let $B^1 = B^*$. Given B^i, define B^{i+1} as follows.

Denote by $>^i$ the binary relation on $\mathcal{X}$ defined by $z >^i z'$ if $z > z'$ and there is at least one pair (k,t) and (k',t') for which (i) $x_t^k > x_{t'}^{k'}$, (ii) $z = x_t^k$ and $z' = x_{t'}^{k'}$, and (iii) the row corresponding $x_t^k > x_{t'}^{k'}$ in B has strictly positive weight in η.

The binary relation $>^i$ cannot exhibit cycles because $>^i \subseteq >$. There is therefore at least one sequence $z_1^i, \ldots z_{L_i}^i$ in $\mathcal{X}$ such that $z_j^i >^i z_{j+1}^i$ for all $j = 1, \ldots, L_i - 1$ and with the property that there is no $z \in \mathcal{X}$ with $z >^i z_1^i$ or $z_{L_i}^i >^i z$.

Observe that B^i has at least one row corresponding to $z_j^i >^i z_{j+1}^i$ for all $j = 1, \ldots, L_i - 1$. Let the matrix B^{i+1} be defined as the matrix obtained from B^i by omitting one copy of the row corresponding to $z_j^i > z_{j+1}^i$, for all $j = 1, \ldots L_i - 1$.

The matrix B^{i+1} has strictly fewer rows than B^i. There is therefore n^* for which B^{n^*+1} either has no more rows, or $B_1^{n^*+1}$ has only zeroes in all its entries (its rows are copies of the δ-row which has only zeroes in its first $K \times (T+1)$ columns).

Define a sequence of pairs $(x_{t_i}^{k_i}, x_{t_i'}^{k_i'})_{i=1}^{n^*}$ by letting $x_{t_i}^{k_i} = z_1^i$ and $x_{t_i'}^{k_i'} = z_{L_i}^i$. Note that, as a result, $x_{t_i}^{k_i} > x_{t_i'}^{k_i'}$ for all i. Therefore the sequence of pairs $(x_{t_i}^{k_i}, x_{t_i'}^{k_i'})_{i=1}^{n^*}$ satisfies one of the conditions in SAR-EDU. $\square$

We shall use the sequence of pairs $(x_{t_i}^{k_i}, x_{t_i'}^{k_i'})_{i=1}^{n^*}$ as our candidate violation of SAR-EDU.

Consider a sequence of matrices A^i, $i = 1, \ldots, n^*$ defined as follows. Let $A^1 = A^*$, $B^1 = B^*$, and $C^1 = \begin{bmatrix} A^1 \\ B^1 \end{bmatrix}$. Observe that the rows of C^1 add to the null vector by Claim 1.

We shall proceed by induction. Suppose that A^i has been defined, and that the rows of $C^i = \begin{bmatrix} A^i \\ B^i \end{bmatrix}$ add to the null vector.

Recall the definition of the sequence

$$x_{t_i}^{k_i} = z_1^i > \ldots > z_{L_i}^i = x_{t_i'}^{k_i'}.$$

There is no $z \in \mathcal{X}$ with $z >^i z_1^i$ or $z_{L_i}^i >^i z$, so in order for the rows of C^i to add to zero there must be a -1 in A_1^i in the column corresponding to (k_i', t_i') and a 1 in A_1^i in the column corresponding to (k_i, t_i). Let r_i be a row in A^i corresponding to (k_i, t_i), and r_i' be a row corresponding to (k_i', t_i'). The existence of a -1 in A_1^i in the column corresponding to (k_i', t_i'), and a 1 in A_1^i in the column corresponding to (k_i, t_i), ensures that r_i and r_i' exist. Note that the row r_i' is a converted row while r_i is original. Let A^{i+1} be defined from A^i by deleting the two rows, r_i and r_i'.

CLAIM 4: *The sum of r_i, r_i', and the rows of B^i which are deleted when forming B^{i+1} (corresponding to the pairs $z_j^i > z_{j+1}^i$, $j = 1, \ldots, L_i - 1$) add to the null vector.*

Proof. Recall that $z_j^i >^i z_{j+1}^i$ for all $j = 1, \ldots, L_i - 1$. Thus, when we add the rows corresponding to $z_j^i >^i z_{j+1}^i$ and $z_{j+1}^i >^i z_{j+2}^i$, then the entries in the column for (k, t) with $x_t^k = z_{j+1}^i$ cancel out and the sum is zero in that entry. Thus, when we add the rows of B^i that are not in B^{i+1} we obtain a vector that is 0 everywhere except the columns corresponding to z_1^i and $z_{L_i}^i$. This vector cancels out with $r_i + r_i'$, by definition of r_i and r_i'. $\square$

CLAIM 5: *The matrix A^* can be partitioned into pairs (r_i, r_i'), in which the rows r_i' are converted and the rows r_i are original.*

Proof. For each i, A^{i+1} differs from A^i in that the rows r_i and r_i' are removed from A^i to form A^{i+1}. We shall prove that A^* is composed of the $2n^*$ rows r_i, r_i'.

First note that since the rows of C^i add up to the null vector, and A^{i+1} and B^{i+1} are obtained from A^i and B^i by removing a collection of rows that add up to zero, then the rows of C^{i+1} must add up to zero as well.

By way of contradiction, suppose that there exist rows left after removing r_{n^*} and r_{n^*}'. Then, by the argument above, the rows of the matrix C^{n^*+1} must add to the null vector. If there are rows left, then the matrix C^{n^*+1} is well defined.

By definition of the sequence B^i, however, B^{n^*+1} has all its entries equal to zero, or has no rows. Therefore, the rows remaining in $A_1^{n^*+1}$ must add up to zero. By Claim 2, the entries of a column (k, t) of A^* are always of the same sign. Moreover, each row of A^* has a non-zero element in the first $K \times (T + 1)$ columns. Therefore, no subset of the columns of A_1^* can sum to the null vector. $\square$

CLAIM 6: *(i) For any k and t, if $(k_i, t_i) = (k, t)$ for some i, then the row r_i corresponding to (k, t) appears as original in A^*. Similarly, if $(k_i', t_i') = (k', t')$ for some i, then the row corresponding to (k, t) appears converted in A^*. (ii) If the row corresponding to (k, t) appears as original in A^*, then there is some i with $(k_i, t_i) = (k, t)$. Similarly, if the row corresponding to (k, t) appears converted in A^*, then there is i with $(k_i', t_i') = (k, t)$.*

Proof. (i) is true by definition of $(x^{k_i}_{t_i}, x^{k'_i}_{t'_i})$. (ii) is immediate from Claim 5 because if the row corresponding to (k,t) appears as original in A^* then it equals r_i for some i, and then $x^k_t = x^{k_i}_{t_i}$. Similarly when the row appears converted. □

CLAIM 7: *The sequence* $(x^{k_i}_{t_i}, x^{k'_i}_{t'_i})_{i=1}^{n^*}$ *satisfies conditions in SAR-EDU: (a)* $\sum_{i=1}^{n^*} t_i \geq \sum_{i=1}^{n^*} t'_i$ *and (b) the number of times* k *appears as* k_i *equals the number of times it appears as* k'_i*.*

Proof. We first establish condition (a). Note that A^*_2 is a vector, and in row r the entry of A^*_2 is as follows. There must be a raw (k,t) in A of which the raw r is a copy. Therefore, the component at the row r of A^*_2 is t if r is original and $-t$ if r is converted. Now, by the construction of the sequence when r appears as original there is some i for which $t = t_i$, when r appears as converted there is some i for which $t = t'_i$. Thus, for each r there is i such that $(A^*_4)_r$ is either t_i or $-t'_i$. By Claim 1 (ii), $\theta \cdot A_2 + \eta \cdot B_2 = 0$. Recall that $\theta \cdot A_2$ equals the sum of the rows of A^*_2. Moreover, B_2 is a vector that has zeroes everywhere except a -1 in the δ row (i.e., $K \times (T+1)+1$th row). Therefore, the sum of the rows of A^*_2 equals $\eta_{K\times(T+1)+1}$, where $\eta_{K\times(T+1)+1}$ is the $K \times (T+1)+1$th element of η. Since $\eta \geq 0$, therefore, $\sum_{i=1}^{n^*} t_i \geq \sum_{i=1}^{n^*} t'_i$, and condition (a) is satisfied.

Now we turn to condition (b). By Claim 1 (iii), the rows of A^*_3 add up to zero. Therefore, the number of times that k appears in an original row equals the number of times that it appears in a converted row. By Claim 6, then, the number of times k appears as k_i equals the number of times it appears as k'_i. Therefore, condition (b) is satisfied. □

Finally, in the following, we show that $\prod_{i=1}^{n^*} p^{k_i}_{t_i}/p^{k'_i}_{t'_i} > 1$, which finishes the proof of Lemma 5 as the sequence $(x^{k_i}_{t_i}, x^{k'_i}_{t'_i})_{i=1}^{n^*}$ would then exhibit a violation of SAR-EDU.

CLAIM 8: $\prod_{i=1}^{n^*} p^{k_i}_{t_i}/p^{k'_i}_{t'_i} > 1$.

Proof. By Claim 1 (iv) and the fact that the submatrix E_4 equals the scalar 1, we obtain

$$0 = \theta \cdot A_4 + \pi E_4 = \left(\sum_{i=1}^{n^*}(r_i + r'_i)\right)_4 + \pi,$$

where $(\sum_{i=1}^{n^*}(r_i + r'_i))_4$ is the (scalar) sum of the entries of A^*_4. Recall that $-\log p^{k_i}_{t_i}$ is the last entry of row r_i and that $\log p^{k'_i}_{t'_i}$ is the last entry of row r'_i, as r'_i is converted and r_i original. Therefore the sum of the rows of A^*_4 are $\sum_{i=1}^{n^*} \log(p^{k'_i}_{t'_i}/p^{k_i}_{t_i})$. Then,

$$\sum_{i=1}^{n^*} \log(p^{k'_i}_{t'_i}/p^{k_i}_{t_i}) = -\pi < 0.$$

Thus $\prod_{i=1}^{n^*} p_{t_i}^{k_i} / p_{t'_i}^{k'_i} > 1$. $\square$