Online Appendix

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

Federico Echenique Taisuke Imai Kota Saito

## Contents

# A Omitted Proofs and Additional Theoretical Results

## A.1 Proof of Theorem 1 (ctd.)

### A.1.1 Proof of Lemma 1

We shall prove that (a) implies (b). Let $(x^k, p^k)_{k=1}^K$ be EDU rational. Let $\delta \in (0, 1]$ and $u : \mathbf{R}_+ \to \mathbf{R}$ be as in the definition of EDU rational data. Then (see, for example, Theorem 28.3 of Rockafellar, 1997), there are numbers $\lambda^k \geq 0$, $k = 1, \ldots, K$ such that if we let

$$v_t^k = \frac{\lambda^k p_t^k}{\delta^t},$$

then $v_t^k \in \partial u(x_t^k)$ if $x_t^k > 0$, and there is $\underline{w} \in \partial u(x_t^k)$ with $v_t^k \geq \underline{w}$ if $x_t^k = 0$. In fact, it is easy to see that $\lambda^k > 0$, and therefore $v_t^k > 0$.

By the concavity of $u$, and the consequent monotonicity of $\partial u(x_t^k)$ (see Theorem 24.8 of Rockafellar, 1997), if $x_t^k > x_{t'}^{k'} > 0$, $v_t^k \in \partial u(x_t^k)$, and $v_{t'}^{k'} \in \partial u(x_{t'}^{k'})$, then $v_t^k \leq v_{t'}^{k'}$. If $x_t^k > x_{t'}^{k'} = 0$, then $\underline{w} \in \partial u(x_{t'}^{k'})$ with $v_{t'}^{k'} \geq \underline{w}$. We thus have $v_t^k \leq \underline{w} \leq v_{t'}^{k'}$.

In the second place, we show that (b) implies (a). Suppose that the numbers $v_t^k$, $\lambda^k$, $\delta$, for $t \in T$ and $k \in K$, are as in (b).

Enumerate the elements in $\mathcal{X}$ in increasing order:

$$y_1 < y_2 < \ldots < y_n.$$

Let

$$\underline{y}_i = \min\{v_t^k : x_t^k = y_i\} \text{ and } \bar{y}_i = \max\{v_t^k : x_t^k = y_i\}.$$

Let $z_i = (y_i + y_{i+1})/2$, $i = 1, \ldots, n-1$; $z_0 = 0$, and $z_n = y_n + 1$. Let $f$ be a correspondence defined as follows:

$$f(z) = \begin{cases} [\underline{y}_i, \bar{y}_i] & \text{if } z = y_i, \\ \max\{\bar{y}_i : z < y_i\} & \text{if } y_n > z \text{ and } \forall i (z \neq y_i), \\ \underline{y}_n/2 & \text{if } y_n < z. \end{cases}$$

By assumption of the numbers $v_t^k$, we have that, when $y < y'$, $v \in f(y)$ and $v' \in f(y')$, then $v \leq v'$. Then the correspondence $f$ is monotone and there is a concave function $u$ for which $\partial u = f$ (Theorem 24.8 of Rockafellar, 1997). Given that $v_t^k > 0$ all the elements in the range of $f$ are positive, and therefore $u$ is strictly increasing.

Finally, for all $(k,t)$, $\lambda^k p_t^k/\delta^t = v_t^k \in \partial u(v_t^k)$ and therefore the first-order conditions to a maximum choice of $x$ hold at $x_t^k$. Since $u$ is concave the first-order conditions are sufficient. The dataset is therefore EDU rational. $\square$

#### A.1.2 Proof of Lemma 6

For each sequence $\sigma = (x_{t_i}^{k_i}, x_{t_i'}^{k_i'})_{i=1}^n$ that satisfies conditions in SAR-EDU, we define a vector $h_\sigma \in \mathbf{N}^{(K\times T)^2}$ as follows. To make the notation easier, we identify the pair $(x_{t_i}^{k_i}, x_{t_i'}^{k_i'})$ with $((k_i,t_i),(k_i',t_i'))$. Let $h_\sigma((k,t),(k',t'))$ be the number of times that the pair $(x_t^k, x_{t'}^{k'})$ appears in the sequence $\sigma$. One can then describe the satisfaction of SAR-EDU by means of the vectors $h_\sigma$. Define

$$H = \left\{ h_\sigma \in \mathbf{N}^{(K\times T)^2} : \sigma \text{ satisfies conditions in SAR-EDU} \right\}.$$

Observe that the set $H$ depends only on $(x^k)_{k=1}^K$ in the dataset $(x^k,p^k)_{k=1}^K$. It does not depend on prices.

For each $((k,t),(k',t')) \in (K\times T)^2$ such that $x_t^k > x_{t'}^{k'}$, define

$$\widehat{\gamma}((k,t),(k',t')) = \log\left(\frac{p_t^k}{p_{t'}^{k'}}\right),$$

and define $\widehat{\gamma}((k,t),(k',t')) = 0$ when $x_t^k \le x_{t'}^{k'}$. Then, $\widehat{\gamma}$ is a $(KT)^2$-dimensional real-valued vector. If $\sigma = (x_{t_i}^{k_i}, x_{t_i'}^{k_i'})_{i=1}^n$, then

$$\widehat{\gamma}\cdot h_\sigma = \sum_{((k,t),(k',t'))\in(K\times T)^2} \widehat{\gamma}((k,t),(k',t'))h_\sigma((k,t),(k',t')) = \log\left(\prod_{i=1}^n \frac{p_{t_i}^{k_i}}{p_{t_i'}^{k_i'}}\right).$$

Therefore, the data satisfy SAR-EDU if and only if $\widehat{\gamma}\cdot h \le 0$ for all $h \in H$.

Enumerate the elements in $\mathcal{X}$ in increasing order:

$$y_1 < y_2 < \cdots < y_N.$$

Fix an arbitrary $\underline{\xi} \in (0,1)$. We shall construct by induction a sequence $(\varepsilon_t^k(n))$ for $n = 1,\ldots,N$, where $\varepsilon_t^k(n)$ is defined for all $(k,t)$ with $x_t^k = y_n$.

By the denseness of the rational numbers, and the continuity of the exponential function, for each $(k,t)$ such that $x_t^k = y_1$, there exists a positive number $\varepsilon_t^k(1)$ such that $\log(p_t^k \varepsilon_t^k(1)) \in \mathbf{Q}$ and $\underline{\xi} < \varepsilon_t^k(1) < 1$. Let $\varepsilon(1) = \min\{\varepsilon_t^k(1) : x_t^k = y_1\}$.

In the second place, for each $(k,t)$ such that $x_t^k = y_2$, there exists a positive number $\varepsilon_t^k(2)$ such that $\log(p_t^k \varepsilon_t^k(2)) \in \mathbf{Q}$ and $\underline{\xi} < \varepsilon_t^k(2) < \varepsilon(1)$. Let $\varepsilon(2) = \min\{\varepsilon_t^k(2) : x_t^k = y_2\}$.

In the third place, and reasoning by induction, suppose that $\varepsilon(n)$ has been defined and that $\underline{\xi} < \varepsilon(n)$. For each $(k,t)$ such that $x_t^k = y_{n+1}$, let $\varepsilon_t^k(n+1) > 0$ be such that $\log(p_t^k \varepsilon_t^k(n+1)) \in \mathbf{Q}$, and $\underline{\xi} < \varepsilon_t^k(n+1) < \varepsilon(n)$. Let $\varepsilon(n+1) = \min\{\varepsilon_t^k(n+1) : x_t^k = y_n\}$.

This defines the sequence $(\varepsilon_t^k(n))$ by induction. Note that $\varepsilon_t^k(n+1)/\varepsilon(n) < 1$ for all $n$. Let $\bar{\xi} < 1$ be such that $\varepsilon_t^k(n+1)/\varepsilon(n) < \bar{\xi}$.

For each $k \in K$ and $t \in T$, let $q_t^k = p_t^k \varepsilon_t^k(n)$, where $n$ is such that $x_t^k = y_n$. We claim that the data $(x^k, q^k)_{k=1}^K$ satisfy SAR-EDU. Let $\gamma^*$ be defined from $(q^k)_{k=1}^K$ in the same manner as $\widehat{\gamma}$ was defined from $(p^k)_{k=1}^K$.

For each pair $((k,t),(k',t'))$ with $x_t^k > x_{t'}^{k'}$, if $n$ and $m$ are such that $x_t^k = y_n$ and $x_{t'}^{k'} = y_m$, then $n > m$. By the definition of $\varepsilon$,

$$\frac{\varepsilon_t^k(n)}{\varepsilon_{t'}^{k'}(m)} < \frac{\varepsilon_t^k(n)}{\varepsilon(m)} < \bar{\xi} < 1.$$

Therefore,

$$\gamma^*((k,t),(k',t')) = \log \frac{p_t^k \varepsilon_t^k(n)}{p_{t'}^{k'} \varepsilon_{t'}^{k'}(m)} < \log \frac{p_t^k}{p_{t'}^{k'}} + \log \bar{\xi} < \log \frac{p_t^k}{p_{t'}^{k'}} = \widehat{\gamma}(x_s^k, x_{t'}^{k'}).$$

Thus, for all $h \in H$, $\gamma^* \cdot h \leq \widehat{\gamma} \cdot h \leq 0$, as $h \geq 0$ and the data $(x^k, p^k)_{k=1}^K$ satisfies SAR-EDU. Thus, the data $(x^k, q^k)_{k=1}^K$ satisfies SAR-EDU. Finally, note that $\underline{\xi} < \varepsilon_t^k(n) < 1$ for all $n$ and each $k \in K, t \in T$. Thus, by choosing $\underline{\xi}$ close enough to 1, we can take the prices $(q^k)_{k=1}^K$ to be as close to $(p^k)_{k=1}^K$ as desired.

### A.1.3 Proof of Lemma 7

Consider the system comprised by (A1), (A2), and (A3) in the proof of Lemma 5. Let $A$, $B$, and $E$ be constructed from the data as in the proof of Lemma 5. The difference with respect to Lemma 5 is that now the entries of $A_4$ may not be rational. Note that the entries of $E$, $B$, and $A_i$, $i = 1, 2, 3$, are rational.

Suppose, towards a contradiction, that there is no solution to the system comprised by (A1), (A2), and (A3). Then, by the argument in the proof of Lemma 5 there is no solution to system $S1$. Lemma 3 with $\mathbf{F} = \mathbf{R}$ implies that there is a real vector $(\theta, \eta, \pi)$ such that

$$\theta \cdot A + \eta \cdot B + \pi \cdot E = 0 \text{ and } \eta \geq 0, \pi > 0.$$

Recall that $B_4 = 0$ and $E_4 = 1$, so we obtain that $\theta \cdot A_4 + \pi = 0$.

Let $(q^k)_{k=1}^K$ be vectors of prices such that the dataset $(x^k, q^k)_{k=1}^K$ satisfies SAR-EDU and $\log q_t^k \in \mathbf{Q}$ for all $k$ and $s$. (Such $(q^k)_{k=1}^K$ exists by Lemma 6.) Construct matrices $A'$, $B'$, and $E'$ from this dataset in the same way as $A$, $B$, and $E$ are constructed in the proof of Lemma 5. Note that only the prices are different in $(x^k, q^k)$ compared to $(x^k, p^k)$. Therefore, it follows that $E' = E$, $B' = B$ and $A'_i = A_i$ for $i = 1, 2, 3$. Since only prices $q^k$ are different in this dataset, only $A'_4$ may be different from $A_4$.

By Lemma 6, we can choose prices $q^k$ such that $|\theta \cdot A'_4 - \theta \cdot A_4| < \pi/2$. We have shown that $\theta \cdot A_4 = -\pi$, so the choice of prices $q^k$ guarantees that $\theta \cdot A'_4 < 0$. Let $\pi' = -\theta \cdot A'_4 > 0$.

Note that $\theta \cdot A'_i + \eta \cdot B'_i + \pi' E_i = 0$ for $i = 1, 2, 3$, as $(\theta, \eta, \pi)$ solves system $S2$ for matrices $A$, $B$, and $E$, and $A'_i = A_i$, $B'_i = B_i$, and $E_i = 0$ for $i = 1, 2, 3$. Finally, $B_4 = 0$ so

$$\theta \cdot A'_4 + \eta \cdot B'_4 + \pi' E_4 = \theta \cdot A'_4 + \pi' = 0.$$

We also have that $\eta \geq 0$ and $\pi' > 0$. Therefore $\theta$, $\eta$, and $\pi'$ constitute a solution to system $S2$ for matrices $A'$, $B'$, and $E'$.

Lemma 3 then implies that there is no solution to system $S1$ for matrices $A'$, $B'$, and $E'$. Thus, there is no solution to the system comprised by (A1), (A2), and (A3) in the proof of Lemma 5. However, this contradicts Lemma 5 because the data $(x^k, q^k)$ satisfies SAR-EDU and $\log q_t^k \in \mathbf{Q}$ for all $k \in K$ and $t \in T$.

## A.2 Proof of Theorem 2

The proofs for QHD and PQHD are similar, so we give a detailed proof for PQHD and then explain how the proof for QHD is different.

**Lemma 8.** *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 PQHD rational.*

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

$$v_t^k = \lambda^k p_t^k \text{ if } t = 0, \ \beta \delta^t v_t^k = \lambda^k p_t^k \text{ if } t > 0, \text{ and } x_t^k > x_{t'}^{k'} \Longrightarrow v_t^k \leq v_{t'}^{k'}.$$

The proof of Lemma 8 is similar to the proof of Lemma 1 and omitted.

### A.2.1 Necessity

**Lemma 9.** *If a dataset $(x^k, p^k)_{k=1}^K$ is PQHD rational, then it satisfies SAR-PQHD.*

*Proof.* Let $(x^k, p^k)_{k=1}^K$ be PQHD rational, and let $\beta \leq 1$, $\delta \in (0, 1]$, and $u : \mathbf{R}_+ \to \mathbf{R}$ be as in the definition of PQHD rational. By Lemma 8, there exists a strictly positive solution $v_t^k$, $\lambda^k$, $\beta$, $\delta$ to the system in statement (b) of Lemma 8 with $v_t^k \in \partial u(x_t^k)$ when $x_t^k > 0$, and $v_t^k \geq \underline{w} \in \partial u(x_t^k)$ when $x_t^k = 0$. Moreover, $v_t^k = \lambda^k p_t^k / D(t)$, where $D(t) = 1$ if $t = 0$ and $D(t) = \beta\delta^t$ if $t > 0$.

Let $(x_{t_i}^{k_i}, x_{t_i'}^{k_i'})_{i=1}^n$ be a sequence satisfying the four conditions in SAR-PQHD: (i) $\sum_{i=1}^n t_i \geq \sum_{i=1}^n t_i'$, (ii) $\#\{i : t_i > 0\} \geq \#\{i : t_i' > 0\}$, (iii) 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), and (iv) $x_{t_i}^{k_i} > x_{t_i'}^{k_i'}$ 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 $v_{t_i}^{k_i} \leq v_{t_i'}^{k_i'}$. Similarly, if $x_{t_i'}^{k_i'} = 0$, then $v_{t_i}^{k_i} \in \partial u(x_{t_i}^{k_i})$ and $v_{t_i'}^{k_i'} \geq \underline{w} \in \partial u(x_{t_i'}^{k_i'})$, so that $v_{t_i}^{k_i} \leq v_{t_i'}^{k_i'}$. Therefore,

$$1 \geq \prod_{i=1}^n \frac{\lambda^{k_i} D(t_i') p_{t_i}^{k_i}}{\lambda^{k_i'} D(t_i) p_{t_i'}^{k_i'}} = \prod_{i=1}^n \frac{D(t_i') p_{t_i}^{k_i}}{D(t_i) p_{t_i'}^{k_i'}} = \frac{\beta^{\#\{i:t_i'>0\} - \#\{i:t_i>0\}}}{\delta^{(\sum t_i - \sum t_i')}} \prod_{i=1}^n \frac{p_{t_i}^{k_i}}{p_{t_i'}^{k_i'}} \geq \prod_{i=1}^n \frac{p_{t_i}^{k_i}}{p_{t_i'}^{k_i'}},$$

where the first equality holds by condition (iii) of SAR-PQHD; and the numbers $\lambda^k$ appear the same number of times in the denominator as in the numerator of this product. Moreover, the last inequality holds by conditions (i) and (ii) of SAR-PQHD. $\square$

### A.2.2 Sufficiency

**Lemma 10.** *Let data $(x^k, p^k)_{k=1}^k$ satisfy SAR-PQHD. Suppose that $\log(p_t^k) \in \mathbf{Q}$ for all $k$ and $t$. Then there are numbers $v_t^k$, $\lambda^k$, $\beta$, $\delta$, for $t \in T$ and $k \in K$ satisfying (b) in Lemma 8.*

**Lemma 11.** *Let data $(x^k, p^k)_{k=1}^k$ satisfy SAR-PQHD. 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-PQHD.*

**Lemma 12.** *Let data $(x^k, p^k)_{k=1}^k$ satisfy SAR-PQHD. Then there are numbers $v_t^k$, $\lambda^k$, $\beta$, $\delta$, for $t \in T$ and $k \in K$ satisfying (b) in Lemma 8.*

Lemma 11 and 12 hold as in the proof for Theorem 1.

### A.2.3 Proof of Lemma 10

We linearize the equation in system (b) of Lemma 8. The result is:

$$\log v_t^k - \log \lambda^k - \log p_t^k = 0 \text{ if } t = 0, \tag{5}$$

$$\log v_t^k + \log \beta + t \log \delta - \log \lambda^k - \log p_t^k = 0 \text{ if } t > 0, \tag{6}$$

$$x_t^k > x_{t'}^{k'} \Longrightarrow \log v_{t'}^{k'} \geq \log v_t^k, \tag{7}$$

$$\log \beta \geq 0, \tag{8}$$

$$\log \delta \leq 0. \tag{9}$$

In the system comprised by (5), (6), (7), (8), and (9), the unknowns are the real numbers $\log \beta$, $\log \delta$, $\log \lambda^k$, and $\log v_t^k$ for all $k = 1, \ldots, K$ and $t = 1, \ldots, T$.

First, we are going to write the system of inequalities from (5) to (9) in a matrix form.

We shall define a matrix $A$ such that there are positive numbers $v_t^k$, $\lambda^k$, $\beta$, and $\delta$, the logs of which satisfy equations (5) and (6) if and only if there is a solution $w \in \mathbf{R}^{K\times(T+1)+2+K+1}$ 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 $K \times (T+1)$ rows and $K \times (T+1) + 2 + K + 1$ columns, defined as follows: We have one row for every pair $(k,t)$, one column for every pair $(k,t)$, two columns 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 two columns, which we shall refer to as the $\beta$-column and $\delta$-column, respectively, 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 $(k,t)$, it has a 1 if $t > 0$ and it has a 0 if $t = 0$ in the $\beta$-column, it has a $t$ in the $\delta$-column, it has a $-1$ in the column for $k$, and $-\log p_t^k$ in the very last column.

Thus, matrix $A$ looks as follows:

$$\begin{array}{cc}
 & \begin{array}{cccccccccccccc} (1,1) & \cdots & (k,t) & (k,t') & \cdots & (K,T) & \beta & \delta & 1 & \cdots & k & \cdots & K & p \end{array} \\
\begin{array}{c} \vdots \\ (k,t=0) \\ (k,t'>0) \\ \vdots \end{array} & \left[\begin{array}{cccccc|c|c|ccccc|c}
\vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots \\
0 & \cdots & 1 & 0 & \cdots & 0 & 0 & t & 0 & \cdots & -1 & \cdots & 0 & -\log p_t^k \\
0 & \cdots & 0 & 1 & \cdots & 0 & 1 & t' & 0 & \cdots & -1 & \cdots & 0 & -\log p_{t'}^k \\
\vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots
\end{array}\right]
\end{array}.$$

Consider the system $A \cdot w = 0$. If there are numbers solving Equations (5) and (6), then these define a solution $w \in \mathbf{R}^{K\times(T+1)+2+K+1}$ for which the last component is 1. If, on the other hand, there is a solution $w \in \mathbf{R}^{K\times(T+1)+2+K+1}$ 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 equations (5) and (6).

In the second place, we write the system of inequalities (7), (8), and (9) in a matrix form. Let $B$ be a matrix with $K \times (T+1) + 2 + K + 1$ 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 $(k',t')$. Finally, we have last two rows, where we have zero everywhere with one exception. In the first row, we have a $-1$ at $(K \times (T+1)+1)$-th column; in the second row, we have a $-1$ at $(K \times (T+1)+2)$-th column. We shall refer to the first row as the $\beta$-row, which captures (8). We also shall refer to the second row as the $\delta$-row, which captures (9). For general QHD, we do not have a $\beta$-row.

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 $K \times (T+1) + 2 + K + 1$ columns. It has zeroes everywhere except for 1 in the last column.

To sum up, there is a solution to system (5), (6), (7), (8), and (9) if and only if there is a vector $w \in \mathbf{R}^{K\times(T+1)+2+K+1}$ 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 argument now follows along the lines of the proof of Theorem 1. Suppose that there is no solution $w$ and let $(\theta, \eta, \pi)$ be an integer vector solving system $(S2): \ \theta \cdot A + \eta \cdot B + \pi \cdot E = 0, \ \eta \geq 0, \pi > 0$.

The following has the same proof as Claim 1.

**Claim 9.** *(i) $\theta \cdot A_1 + \eta \cdot B_1 = 0$; (ii) $\theta \cdot A_2 + \eta \cdot B_2 = 0$; (iii) $\theta \cdot A_3 + \eta \cdot B_3 = 0$; (iv) $\theta \cdot A_4 = 0$; and (v) $\theta \cdot A_5 + \pi \cdot E_5 = 0$.*

We transform the matrices $A$ and $B$ based on the values of $\theta$ and $\eta$, as we did in the proof of Theorem 1. Let us define a matrix $A^*$ from $A$ and $B^*$ from $B$, as we did in the proof of Theorem 1. We can prove the same claims (i.e., Claims 2, 3, 4, 5, and 6) as in the proof of Theorem 1. The proofs are the same and omitted. In particular, we can show that 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-PQHD, $x_{t_i}^{k_i} > x_{t'_i}^{k'_i}$ for all $i = 1, \ldots, n^*$. 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-PQHD.

**Claim 10.** *The sequence $(x_{t_i}^{k_i}, x_{t'_i}^{k'_i})_{i=1}^{n^*}$ satisfies conditions (i), (ii), and (iii) in SAR-PQHD.*

*Proof.* We first establish condition (i). Note that $A_3^*$ is a vector, and in row $r$ the entry of $A_3^*$ is as follows. There must be a $(k,t)$ of which $r$ is a copy. Then the component at row $r$ of $A_3^*$ is $t$ if $r$ is original and $-t$ if $r$ is converted. Now, 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 $(A_3^*)_r$ is either $t_i$ or $-t'_i$.

By Claim 9 (iii), $\theta \cdot A_3 + \eta \cdot B_3 = 0$. Recall that $\theta \cdot A_3$ equals the sum of the rows of $A_3^*$. Moreover, $B_3$ is a vector that has zeroes everywhere except a $-1$ in the $\delta$ row (i.e., $(K \times (T+1) + 2)$-th row). Therefore, the sum of the rows of $A_3^*$ equals $\eta_{K\times(T+1)+2}$, where $\eta_{K\times(T+1)+2}$ is the $(K \times (T+1) + 2)$-th element of $\eta$. Since $\eta \geq 0$, therefore, $\sum_{i:t_i>0} t_i - \sum_{i:t'_i>0} t'_i = \eta_{K\times(T+1)+2} \geq 0$, and condition (i) in the axiom is satisfied.

Next, we show condition (ii). By Claim 9 (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 $\beta$ 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 $\eta$. Since $\eta \geq 0$, therefore, $\#\{i : t_i > 0\} - \#\{i : t'_i > 0\} = \eta_{K\times(T+1)+1} \geq 0$, and condition (ii) in the axiom is satisfied. (For general QHD, $B_2$ is a zero vector in the $\beta$-row, i.e., $(K\times(T+1)+1)$-th row. Therefore, $\#\{i : t_i > 0\} - \#\{i : t'_i > 0\} = 0$, and condition (ii) in SAR-QHD is satisfied.)

Now we turn to condition (iii). By Claim 9 (iv), the rows of $A_4^*$ 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 (iii) in the axiom is satisfied. $\square$

Finally, we can show that $\prod_{i=1}^{n^*} p_{t_i}^{k_i} / p_{t'_i}^{k'_i} > 1$, which finishes the proof of Lemma 5 as the sequence $(x_{t_i}^{k_i}, x_{t'_i}^{k'_i})_{i=1}^{n^*}$ would then exhibit a violation of SAR-PQHD. The proof is the same as that of the corresponding lemma in the proof of Theorem 1.

## A.3 Proof of Theorem 3: $M' = \boldsymbol{TSU}$

The proof that SAR-TSU is equivalent to TSU rationality is similar to the proof of Theorem 1. In the following, we explain the differences.

**Lemma 13.** *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 TSU rational.*

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

$$v_t^k = \lambda^k p_t^k \textit{ and } x_t^k > x_t^{k'} \Longrightarrow v_t^k \leq v_t^{k'}.$$

The proof of Lemma 13 is similar to the proof of Lemma 1 and omitted.

To see that SAR-TSU is necessary, let $(x_{t_i}^{k_i}, x_{t_i'}^{k_i'})_{i=1}^n$ be a sequence under the conditions of the axiom. We present the proof under the assumption that $u_t$ is differentiable, but it is straightforward to use the concavity and the corresponding monotonicity of the superdifferential of $u_t$, as we did in the proof of Theorem 1. The first-order condition is $u_t'(x_t^k) = \lambda^k p_t$. Since $t_i = t_i'$ for each $i$, we obtain

$$1 \geq \prod_{i=1}^n \frac{u_{t_i}'(x_{t_i}^{k_i})}{u_{t_i}'(x_{t_i}^{k_i'})} = \prod_{i=1}^n \frac{\lambda^{k_i} p_{t_i}^{k_i}}{\lambda^{k_i'} p_{t_i}^{k_i'}} = \prod_{i=1}^n \frac{\lambda^{k_i}}{\lambda^{k_i'}} \prod_{i=1}^n \frac{p_{t_i}^{k_i}}{p_{t_i}^{k_i'}} = \prod_{i=1}^n \frac{p_{t_i}^{k_i}}{p_{t_i}^{k_i'}},$$

where the last equality holds because each $k$ appears as $k_i'$ the same number of times it appears as $k_i$.

In the following, we prove the sufficiency. The outline of the proof is the same as in the proof of Theorem 1.

**Lemma 14.** *Let data $(x^k, p^k)_{k=1}^k$ satisfy SAR-TSU. Suppose that $\log(p_t^k) \in \mathbf{Q}$ for all $k$ and $t$. Then there are numbers $v_t^k$, $\lambda^k$, $\beta$, $\delta$, for $t \in T$ and $k \in K$ satisfying (b) in Lemma 13.*

*Proof.* We linearize the equation in System (b) of Lemma 13. The result is:

$$\log v_t^k - \log \lambda^k - \log p_t^k = 0, \tag{10}$$

$$x_t^k > x_t^{k'} \Longrightarrow \log v_t^{k'} \geq \log v_t^k. \tag{11}$$

In the system comprised by (10) and (11), the unknowns are the real numbers $\lambda^k$ and $\log v_t^k$ for all $k = 1, \ldots, K$ and $t = 1, \ldots, T$.

We shall define a matrix $A$ such that there are positive numbers $v_t^k$ and $\lambda^k$, the logs of which satisfy equation (10) if and only if there is a solution $w \in \mathbf{R}^{K \times (T+1)+K+1}$ to the system of equations

$$A \cdot w = 0,$$

and for which the last component of $u$ is strictly positive.

Let $A$ be a matrix with $K \times (T+1)$ rows and $K \times (T+1) + K + 1$ columns. The matrix $A$ is similar to the matrix $A$ defined in the proof of Theorem 1, only the difference here is that we no longer have the $\delta$-column. Thus, matrix $A$ looks as follows:

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

Consider the system $A \cdot w = 0$. If there are numbers solving equation (10), then these define a solution $w \in \mathbf{R}^{K \times (T+1)+K+1}$ for which the last component is 1. If, on the other hand, there is a solution $w \in \mathbf{R}^{K \times (T+1)+K+1}$ 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 (10).

In the second place, we write the system of inequalities (11) in a matrix form. Let $B$ be a matrix with $K \times (T+1) + K + 1$ 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 $(k',t)$.

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 $K \times (T+1) + K + 1$ columns. It has zeroes everywhere except for 1 in the last column.

To sum up, there is a solution to system (10) and (11) if and only if there is a vector $w \in \mathbf{R}^{K \times (T+1)+K+1}$ 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 system $S1$ if and only if there is no vector $(\theta, \eta, \pi)$ 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-TSU. Suppose then that there is no solution $w$ and let $(\theta, \eta, \pi)$ be a rational vector as above, solving system $S2$.

By multiplying $(\theta, \eta, \pi)$ by any positive integer we obtain new vectors that solve system $S2$, so we can take $(\theta, \eta, \pi)$ to be integer vectors.

For convenience, we transform the matrices $A$ and $B$ using $\theta$ and $\eta$. We now transform the matrices $A$ and $B$ based on the values of $\theta$ and $\eta$, as we did in the proof of Theorem 1. Let us define a matrix $A^*$ from $A$ and $B^*$ from $B$, as we did in the proof of Theorem 1. We can prove the same claims (i.e., Claims 2, 3, 4, 5, and 6) as in the proof of Theorem 1. The proofs are the same and omitted. In particular, we can show that there exists a sequence of pairs $(x^{k_i}_{t_i}, x^{k'_i}_{t'_i})^{n^*}_{i=1}$ that satisfies $x^{k_i}_{t_i} > x^{k'_i}_{t'_i}$ for all $i = 1, \ldots, n^*$. Moreover, by the definition of $B$ matrix, we have $t_i = t'_i$ because in matrix $B$ we have $z >^i z'$ if there exist $t \in T$ and $k, k' \in T$ such that there exist $x^k_t = z$ and $x^{k'}_t = z'$. Moreover, as in Claim 7, we can show that in the sequence $(x^{k_i}_{t_i}, x^{k'_i}_{t'_i})^{n^*}_{i=1}$, each $k$ appears $k_i$ the same number of times it appears as $k'_i$. Finally, we can show that $\prod^{n^*}_{i=1} p^{k_i}_{t_i}/p^{k'_i}_{t'_i} > 1$, which finishes the proof of Lemma 14 as the sequence $(x^{k_i}_{t_i}, x^{k'_i}_{t'_i})^{n^*}_{i=1}$ would then exhibit a violation of SAR-TSU. The proof is the same as in the proof of Theorem 1 and omitted. □

**Lemma 15.** *Let data $(x^k, p^k)^k_{k=1}$ satisfy SAR-TSU. Then for all positive numbers $\bar{\varepsilon}$, there exists $q^k_t \in [p^k_t - \bar{\varepsilon}, p^k_t]$ for all $t \in T$ and $k \in K$ such that $\log q^k_t \in \mathbf{Q}$ and the dataset $(x^k, q^k)^k_{k=1}$ satisfy SAR-TSU.*

**Lemma 16.** *Let data $(x^k, p^k)^k_{k=1}$ satisfy SAR-TSU. Then there are numbers $v^k_t$ and $\lambda^k$ for all $t \in T$ and $k \in K$ satisfying* ((b)) *in Lemma 13.*

Lemma 15 and 16 hold as in the proof of Theorem 1.

## A.4 Observational Equivalence between Models

### A.4.1 Proof of Proposition 2

The equivalence between (ii) and (iii) are by Proposition 1. We will show the equivalence between (i) and (ii). Obviously (ii) implies (i). So we will show the converse.

To show the converse, suppose that a dataset is not EDU rational with $\delta = 1$. Then, by

Proposition 1, there must be a balanced sequence of pairs $(x_{s_i}^{m_i}, x_{s_i'}^{m_i'})_{i=1}^M$ such that

$$\prod_{i=1}^{M} \frac{p_{s_i}^{m_i}}{p_{s_i'}^{m_i'}} > 1.$$

Since the data is not strictly impatient, there exists a balanced sequence $(x_{t_i}^{k_i}, x_{t_i'}^{k_i'})_{i=1}^n$ of pairs such that $x_{t_i}^{k_i} > x_{t_i'}^{k_i'}$ for all $i$, $\sum_{i=1}^n t_i > \sum_{i=1}^n t_i'$, and

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

Now we can consider a balanced sequence of pairs by including $(x_{s_i}^{m_i}, x_{s_i'}^{m_i'})_{i=1}^M$ and enough copies of $(x_{t_i}^{k_i}, x_{t_i'}^{k_i'})_{i=1}^n$ such that the sum of $t$ on the left of the pair is greater than the sum of $t$ on the right. This is possible because $\sum_{i=1}^n t_i > \sum_{i=1}^n t_i'$. The resulting product of ratios of prices is still equal to

$$\prod_{i=1}^{n} \frac{p_{s_i}^{m_i}}{p_{s_i'}^{m_i'}} > 1.$$

Thus the dataset is not EDU rational. □

### A.4.2 Proof of Proposition 3

Of course, if the data is EDU rational then it is PQHD rational. Let us prove the converse. Choose a sequence $(x_{t_i}^{k_i}, x_{t_i'}^{k_i'})_{i=1}^n$ such that (i) $x_{t_i}^{k_i} > x_{t_i'}^{k_i'}$ for all $i \in \{1, \dots, n\}$, (ii) $\sum_{i=1}^n t_i \geq \sum_{i=1}^n t_i'$, and (iii) each $k$ appears as $k_i$ the same number of times as $k_i'$.

By (i), $x_{t_i}^{k_i} > 0$ for all $i \in \{1, \dots, n\}$. Since $x_0^k = 0$ for all $k \in K$, we obtain $t_i > 0$ for all $i \in \{1, \dots, n\}$. Therefore, $\#\{i \in \{1, \dots, n\} : t_i > 0\} = \#\{i \in \{1, \dots, n\}\} \geq \#\{i \in \{1, \dots, n\} : t_i' > 0\}$. Therefore, the sequence satisfies all of the conditions in SAR-PQHD. Since the dataset is PQHD rational, Theorem 2 shows that

$$\prod_{i=1}^{n} \frac{p_{t_i}^{k_i}}{p_{t_i'}^{k_i}} \leq 1. \tag{12}$$

Therefore, conditions (i), (ii), and (iii) imply (12), which is SAR-EDU. Therefore, by Theorem 1, the dataset must be EDU rational. □

### A.4.3 Corner Choices and EDU Rationality

Section 4.3 discusses one condition on the data for which EDU and QHD are observationally equivalent. The condition is very common in AS's dataset. Here we discuss the same condition and relate it to the pass rate for the EDU test. The condition can be stated and analyzed for datasets obtained from any CTB experimental design. We show that, in CTB experiments, if a subject does not consume a positive amount on the sooner date whenever the price for the sooner consumption is more expensive than the price for the later consumption, then the subject must be EDU rational.

We first formalize the notion of a *CTB* dataset. We say that a dataset $(x^k, p^k)_{k\in K}$ is *CTB* if for each $k \in K$, there exist $l(k), s(k) \in T$ such that $l(k) > s(k)$,

$$x_t^k = 0 \text{ for all } t \in T \setminus \{l(k), s(k)\}, \tag{13}$$

and

$$\min_{t\in T\setminus\{l(k),s(k)\}} p_t^k > p_{s(k)}^k \geq p_{l(k)}^k. \tag{14}$$

In a CTB dataset, an agent can choose positive consumptions only on the two specified dates $l(k)$ and $s(k)$. The date $l(k)$ is a *later date* and $s(k)$ is a *sooner date* (see also Section 4.1).

**Proposition 4.** *Suppose that a dataset $(x^k, p^k)_{k\in K}$ is CTB. Moreover, for each $k \in K$,*

$$p_{s(k)}^k > p_{l(k)}^k \implies x_{s(k)}^k = 0. \tag{15}$$

*Then, the dataset is EDU rational.*

*Proof.* Choose a sequence $(x_{t_i}^{k_i}, x_{t_i'}^{k_i'})_{i=1}^n$ such that (i) $x_{t_i}^{k_i} > x_{t_i'}^{k_i'}$ for all $i \in \{1, \ldots, n\}$, (ii) $\sum_{i=1}^n t_i \geq \sum_{i=1}^n t_i'$, and (iii) each $k$ appears as $k_i$ the same number of times as $k_i'$. By conditions (i) and (iii), we can arrange the sequence to obtain a new sequence $(y_{t_i}^{k_i}, y_{t_i'}^{k_i'})$ such that $y_{t_i}^{k_i} > 0$, $k_i = k_i'$ for all $i \in \{1, \ldots, n\}$, $\{y_{t_i}^{k_i}\}_{i=1}^n = \{x_{t_i}^{k_i}\}_{i=1}^n$, and $\{y_{t_i'}^{k_i'}\}_{i=1}^n = \{x_{t_i'}^{k_i'}\}_{i=1}^n$.

Define a subset $K'$ of $K$ by $K' = \{k \in K : p_{s(k)}^k > p_{l(k)}^k\}$. Choose any $i \in \{1, \ldots, n\}$. We will show that $p_{t_i}^{k_i}/p_{t_i'}^{k_i} \leq 1$.

Consider the case where $k_i \in K'$. Then, by conditions (13) and (15), $t_i = l(k_i)$ because $x_{t_i}^{k_i} > 0$. So $t_i' = s(k_i)$ or $t_i' \notin \{s(k_i), l(k_i)\}$. Therefore,

$$\frac{p_{t_i}^{k_i}}{p_{t_i'}^{k_i}} = \frac{p_{l(k_i)}^{k_i}}{p_{t_i'}^{k_i}} \leq \max\left\{\frac{p_{l(k_i)}^{k_i}}{p_{s(k_i)}^{k_i}}, \frac{p_{l(k_i)}^{k_i}}{\min_{t\in T\setminus\{l(k_i),s(k_i)\}}\{p_t^{k_i}\}}\right\} \leq 1,$$

where the first inequality holds by (14) and the second inequality holds because $k_i \in K'$.

Consider the case where $k_i \notin K'$. By (14), we have $p^{k_i}_{s(k_i)} = p^{k_i}_{l(k_i)}$. By (13), we have $t_i = l(k_i)$ or $t_i = s(k_i)$ because $x^{k_i}_{t_i} > 0$. Then,

$$\frac{p^{k_i}_{t_i}}{p^{k_i}_{t'_i}} \leq \max\left\{\frac{p^{k_i}_{l(k_i)}}{p^{k_i}_{s(k_i)}}, \frac{p^{k_i}_{s(k_i)}}{p^{k_i}_{l(k_i)}}, \frac{p^{k_i}_{s(k_i)}}{\min_{t\in T\setminus\{l(k_i),s(k_i)\}}\{p^{k_i}_t\}}, \frac{p^{k_i}_{l(k_i)}}{\min_{t\in T\setminus\{l(k_i),s(k_i)\}}\{p^{k_i}_t\}}\right\} \leq 1.$$

□

The proposition is important because condition (15) is satisfied by 82.8% (24 out of 29) of EDU rational subjects in AS dataset. Note that the condition does not require anything when $p^k_{l(k)} = p^k_{s(k)}$.

#### A.4.4 Observational equivalence between EDU and TSU

One consequence of Theorems 1 and 3 is that, under certain circumstances, EDU and TSU are *observationally equivalent.*

**Proposition 5.** *Suppose that for any $k, k' \in K$, $x^k_t > x^{k'}_{t'}$ when $t < t'$. Then $(x^k, p^k)^K_{k=1}$ is TSU rational if and only if it is EDU rational.*

*Proof.* EDU implies TSU. We will show that TSU implies EDU. Choose a sequence $(x^{k_i}_{t_i}, x^{k'_i}_{t'_i})^n_{i=1}$ such that (i) $x^{k_i}_{t_i} > x^{k'_i}_{t'_i}$ for all $i \in \{1, \dots, n\}$, (ii) $\sum^n_{i=1} t_i \geq \sum^n_{i=1} t'_i$, and (iii) each $k$ appears as $k_i$ the same number of times as $k'_i$.

Suppose that $t_j \neq t'_j$ for some $j$, then by the condition, $t_j < t'_j$. Moreover, by the condition for all $i$, $t_i \leq t'_i$.

Thus, $\sum^n_{i=1} t_i < \sum_{i=1} t'_i$. This is a contradiction. Therefore, $t_i = t'_i$ for all $i$. Since the dataset is TSU rational, Theorem 3 shows that

$$\prod^n_{i=1} \frac{p^{k_i}_{t_i}}{p^{k_i}_{t'_i}} \leq 1. \tag{16}$$

Therefore, conditions (i), (ii), and (iii) imply (16), which is SAR-EDU. Therefore, by Theorem 1, the dataset must be EDU rational. □

The condition means that, independently of $k$, the agents consume more in an earlier period than in a later period.

# B Implementation of Revealed-Preference Tests

This section presents a method to implement the revealed-preference tests for time discounting models using Matlab R2018a (MathWorks). We use Andreoni and Sprenger's (2012) experimental choice data as the model case, but our method is applicable to other empirical/experimental data sets.

**Dataset.** Subjects in the Andreoni and Sprenger's (2012) experiment completed 45 intertemporal decisions with varying starting dates $\tau$, delay lengths $d$, and gross interest rates $a_{\tau+d}/a_\tau$ and, in particular, they complete 5 decision problems for each pair of $(\tau, d)$. See Figure B.1 for an illustration of budgets. For each subject, the decision in every trial is characterized by a tuple $(\tau, d, a_\tau, a_{\tau+d}, c_\tau)$ where $c_\tau$ is the number of tokens allocated to sooner payment.

The following figure illustrates the budgets faced by the subjects in AS's experiment, fixing one time frame at $(\tau, d)$.

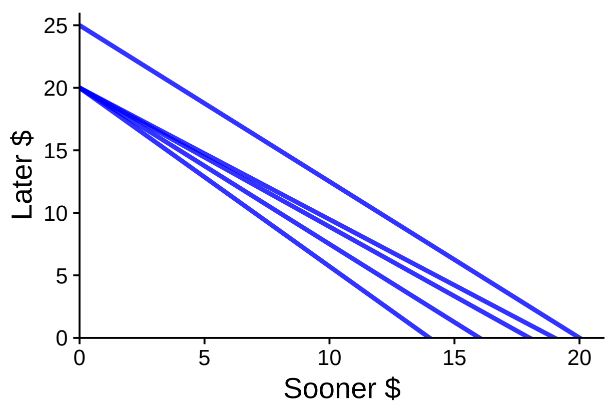


Figure B.1: An illustration of the CTB design in Andreoni and Sprenger (2012). Budget sets are represented in blue lines, fixing one time frame at $(\tau, d) = (0, 35)$.

In order to rewrite our data in price-consumption format as in the theory, we set prices $p_\tau = 1 + r = a_{\tau+d}/a_\tau$ and $p_{\tau+d} = 1$ (normalization), and define consumptions $x_\tau = c_\tau \cdot a_\tau$ and $x_{\tau+d} = (100 - c_\tau) \cdot a_{\tau+d}$. This gives us a dataset $(x^k, p^k)_{k=1}^{45}$.

As we explained in the main body of the paper, we implicitly set prices of consumption in periods that were not offered to a subject as very high in order to ensure that consumption is zero. The idea is as follows. Think of EDU for concreteness. We use first-order conditions, so that we are looking for a rationalizing $u$ and $\delta$ such that $\delta^t u'(x_t^k) = \lambda^k p_t^k$ if $x_t^k > 0$ and $\delta^t u'(x_t^k) \leq \lambda^k p_t^k$ if $x_t^k = 0$. In setting up such a system of equations we can ignore the $t$ that was not offered to the agents in trial $k$. Then whatever $u$ we construct will have a finite derivative $u'(0)$ at zero. Therefore, we can set $p_t^k$ to be high enough so that the agent finds it optimal to consume $x_t^k = 0$. By this argument is it clear that one can ignore the (zero) consumption in the periods that were not offered in trial $k$, as we think of consumption in those periods as prohibitively expensive. This is of course consistent with the fact that AS did not offer subjects any consumption in those periods; consumption in those periods is infeasible. The set of time periods we are looking at is thus $T = \{0, 7, 35, 42, 70, 77, 98, 105, 133\}$.

We are able to check whether a given dataset is consistent with TSU, QHD, PQHD, or EDU, by solving the corresponding linear programming problem. The construction of linear programming problems closely follows the argument in the proofs of Theorem 1, 2, and 3. In particular, the key to this procedure is to set up a system of linear inequalities of the form:

$$S: \begin{cases} A \cdot u = 0 \\ B \cdot u \geq 0 \\ E \cdot u > 0 \end{cases},$$

which, in the case of EDU for example, is a matrix form of the linearized system:

$$\begin{aligned} \log v(x_t^k) + t \log \delta - \log \lambda^k - \log p_t^k = 0, \\ x > x' \Longrightarrow \log v(x') \geq \log v(x), \\ \log \delta \leq 0. \end{aligned}$$

**A system of linear inequalities.** We now construct three key ingredients of the system, matrices $A$, $B$, and $E$, starting from those necessary for testing EDU. The first matrix $A$ looks as follows:

$$\begin{array}{c} \\ \vdots \\ {\scriptstyle (k,t)} \\ \vdots \end{array} \begin{array}{c} \begin{array}{ccccccccccccc} {\scriptstyle (1,0)} & {\scriptstyle \cdots} & {\scriptstyle (k,t)} & {\scriptstyle \cdots} & {\scriptstyle (45,133)} & {\scriptstyle \delta} & {\scriptstyle 1} & {\scriptstyle \cdots} & {\scriptstyle k} & {\scriptstyle \cdots} & {\scriptstyle 45} & {\scriptstyle p} \end{array} \\ \left[\begin{array}{ccccc|c|ccccc|c} \vdots & & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & t & 0 & \cdots & -1 & \cdots & 0 & -\log p_t^k \\ \vdots & & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots \end{array}\right] \end{array}.$$

Since we can ignore the $t$ that was not offered to the agents in trial $k$, the matrix has $45 \times 2 = 90$ rows and $45 \times 2 + 1 + 45 + 1 = 137$ columns. In the row corresponding to $(k, t)$ the matrix has zeroes everywhere with the following exceptions: it has a 1 in the column for $(k, t)$; it has a $t$ in the $\delta$ column; it has a $-1$ in the column for $k$; and $-\log p_t^k$ in the very last column. This finalizes the construction of $A$.

Next, we construct matrix $B$ that has 137 columns and there is 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 $(k', t')$. Finally, in the last row, we have zero everywhere with the exception of a $-1$ at 91st column. We shall refer to this last row as the $\delta$-row.

Finally, we prepare a matrix that captures the requirement that the last component of a solution be strictly positive. The matrix $E$ has a single row and 137 columns. It has zeroes everywhere except for 1 in the last column.

**Constructing matrices for other tests.** In order to test models other than EDU, we need to modify matrices $A$, $B$, and $E$ appropriately.

For the QHD test, we insert another column capturing the present/future bias parameter $\beta$, which we shall refer to the $\beta$-column. Therefore, three matrices $A$, $B$, and $E$ have $45 \times 2 + 1 + 1 + 45 + 1 = 138$ columns. In the row corresponding to $(k, t)$ of the matrix $A$, the $\beta$-column has a 1 if $t > 0$ and a 0 if $t = 0$, indicating "now" or "future".

$$
\begin{array}{cc}
 & \begin{array}{cccccccccccccc} (1,1) & \cdots & (k,t) & (k,t') & \cdots & (45,133) & \beta & \delta & 1 & \cdots & k & \cdots & K & p \end{array} \\
\begin{array}{c} \vdots \\ (k,t{=}0) \\ (k,t'{>}0) \\ \vdots \end{array} &
\left[\begin{array}{cccccc|c|c|ccccc|c}
\vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots \\
0 & \cdots & 1 & 0 & \cdots & 0 & 0 & t & 0 & \cdots & -1 & \cdots & 0 & -\log p_t^k \\
0 & \cdots & 0 & 1 & \cdots & 0 & 1 & t' & 0 & \cdots & -1 & \cdots & 0 & -\log p_{t'}^k \\
\vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots
\end{array}\right].
\end{array}
$$

The construction of matrix $B$ for testing general QHD is the same as above (although the size is now different). For the PQHD test, we add $\beta$-row which has 0 everywhere except $-1$ in the $\beta$-column to capture $\beta \leq 1$.

For the MTD and GTD tests, we have nine columns capturing discount factors $D(t)$'s.

$$
\begin{array}{cc}
 & \begin{array}{cccccccccccccc} (1,0) & \cdots & \widetilde{x}_\ell & \cdots & (45,133) & \cdots & D(t) & \cdots & 1 & \cdots & k & \cdots & 45 & p \end{array} \\
\begin{array}{c} \vdots \\ (k,t) \\ \vdots \end{array} &
\left[\begin{array}{ccccc|ccc|ccccc|c}
\vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\
0 & \cdots & 1 & \cdots & 0 & \cdots & 1 & \cdots & 0 & \cdots & -1 & \cdots & 0 & -\log p_t^k \\
\vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots
\end{array}\right].
\end{array}
$$

In the matrix $B$, we add rows

$$\begin{array}{c}
\begin{array}{ccccccccccccccc}
(1,0) & \cdots & (k,t) & \cdots & (45,133) & \cdots & D(t) & D(t{+}1) & \cdots & 1 & \cdots & k & \cdots & 45 & p
\end{array} \\
\left[\begin{array}{ccccc|cccc|ccccc|c}
\vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\
0 & \cdots & 0 & \cdots & 0 & \cdots & 1 & -1 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 0 \\
\vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots
\end{array}\right]
\end{array}$$

in testing MTD to impose the monotonicity restriction on $D(t)$'s.

The matrix $A$ for testing TSU is similar to that appears in testing EDU. The difference is that we no longer have the $\delta$-column.

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

Next, we construct $B$ 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 $(k',t)$.

**Solve the system.** Our task is to check if there is a vector $w$ that solves the following system of linear inequalities corresponding to a model $M'$

$$S_{M'} : \begin{cases} A \cdot w = 0 \\ B \cdot w \geq 0 \\ E \cdot w > 0 \end{cases} .$$

If there is a solution $w$ to this system, we say that the dataset is $M'$-rational.

We use the function `linprog` in the Optimization Toolbox of Matlab to find a solution. More precisely, we translate the systems of linear inequalities $S_{M'}$ into constraints in a linear programming problem and solve

$$LP_{M'}: \begin{cases} \min & z \cdot w \\ \text{s.t.} & A \cdot w = 0 \\ & -B \cdot w \leq 0 \\ & -E \cdot w < 0 \end{cases},$$

where $z$ is a zero vector.

It is not possible, however, to specify strict inequality constraints in `linprog`. As an alternative, we find a solution $w$ that has 1 in the last element, i.e., $w_p = 1$. In other words, we solve a normalized version of the problem,

$$LP'_{M'}: \begin{cases} \min & z \cdot w \\ \text{s.t.} & A \cdot w = 0 \\ & -B \cdot w \leq 0 \\ & w_p = 1 \end{cases},$$

where $z$ is a zero vector as above. Here, the constraint $E \cdot w > 0$ is omitted since it is automatically satisfied by our normalization $w_p = 1$.

If the given dataset is EDU rational, we can recover upper and lower bounds of the daily discount factor consistent with the observed choice data. Remember that we include the $\delta$-row in $B$. The constraint $B \cdot u \geq 0$ then implies that the 91st element of any solution $w^*$ of $LP'_{M'}$, called $u^*_\delta$, captures the daily discount factor. To be more precise, we can recover the daily discount factor $\delta$ by $\exp(w^*_\delta)$ since we normalize $w^*_p$ to be 1. Therefore, a solution (if any) of $LP'_{M'}$ in which the 91st element of $z$ is 1 and 0 elsewhere suggests an lower bound of $\delta$ and a solution (if any) of $LP'_{M'}$ in which the 91st element of $z$ is $-1$ and 0 elsewhere suggests an upper bound of $\delta$. In a similar manner, we can recover bounds of present/future bias $\beta$.

# C Power of the Tests

In this section, we discuss the power of our tests. The low power of GARP is well documented. As a result, it is common to assess the power of a test by comparing the pass rates (the fraction of choices that pass the relevant revealed preference axiom) from purely random choices.[1] Here we report the results from such an assessment using our tests and the experimental design of AS. We find no evidence of low power.

We generate 10,000 datasets in which choices are made at random and uniformly distributed on the frontier of the budget set (Method 1 of Bronars, 1987). Datasets generated in this way always fail our tests (Table C.1 shows pass rates). Next, we apply the simple bootstrap method to look at the power from an ex-post perspective, as originally introduced in Andreoni and Miller (2002). For each of the 45 budget sets, we randomly pick one choice from the set of choices observed in the entire experiment (i.e., 97 observations for each budget). We generate 10,000 such datasets and apply our revealed-preference tests. We again observe high percentages of violation.

Table C.1: Power measures.

| | AS design | | | CMW design | | |
|---|---|---|---|---|---|---|
| Sampling | EDU | QHD | TSU | EDU | QHD | TSU |
| Uniform random | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Simple Bootstrap | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

The conclusion is that our tests seem to have good power against the (admittedly crude) alternative of random choices. This is a credit to the design of AS.

We apply the same method to CMW version of CTB design (a total of 12 budgets and 1,060 observations for each budget). We again find no evidence of low power.

[1] The idea of using random choices as a benchmark is first applied to revealed preference theory by Bronars (1987). This approach is the most popular in empirical application: see, among other studies, Adams et al. (2014), Andreoni and Miller (2002), Beatty and Crawford (2011), Choi et al. (2007), Crawford (2010), Dean and Martin (2016), Fisman et al. (2007). For overview of power calculation, see discussions in Andreoni et al. (2013) and Crawford and De Rock (2014).

# D Additional Empirical Results

## D.1 Definition of the "Present" Period in QHD

A clear distinction between "the present" ($t = 0$) and "future" ($t > 0$) plays a key role in the QHD model. Experimental studies typically treat reward that is delivered within the day of the experiment as an immediate reward. We followed this approach in our empirical application (Section 4), but the question, *how soon is "now"*, is one of the topics actively debated in the literature (Balakrishnan et al., forthcoming; DellaVigna, 2018; Ericson and Laibson, 2019).

We consider a version of QHD in which the present can last more than one time period and hence the $\beta$ discount factor applies after a number of initial periods. Since our methodology forces us to choose the number of initial periods, for any given $\bar{\tau}$ we can analyze and test for the model

$$\sum_{t=0}^{\bar{\tau}} \delta^t u(x_t) + \beta \sum_{t=\bar{\tau}+1}^{T} \delta^t u(x_t). \tag{17}$$

In order to determine the boundary between the present and future time periods empirically, we run the QHD test for different candidate values of $\bar{\tau} \in T$. The set of time periods is $T = \{0, 7, 35, 42, 70, 77, 98, 105, 133\}$ (days) in AS data and $T = \{0, 4, 8, 12\}$ (weeks) in CMW data. We can now look at the "present" revealed through choice. More precisely, we find the smallest $\bar{\tau} \in T$ under which a subject is strictly QHD rational (Table D.1). Note that the QHD test with $\bar{\tau} = 0$ corresponds to the original QHD test, while the test with $\bar{\tau} = \max T$ corresponds to the EDU test.

Table D.1: Revealed "present" for strictly QHD-rational subjects.

| | $\bar{\tau}$ (days) | | | | | | | | | $\bar{\tau}$ (weeks) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AS | 0 | 7 | 35 | 42 | 70 | 77 | 98 | 105 | CMW | 0 | 4 | 8 |
| PQHD ($\beta \leq 1$) | 0 | 2 | 1 | 2 | 0 | 1 | 0 | 0 | PQHD ($\beta \leq 1$) | 6 | 4 | 5 |
| FQHD ($\beta \geq 1$) | 0 | 1 | 1 | 1 | 3 | 0 | 0 | 0 | FQHD ($\beta \geq 1$) | 6 | 7 | 61 |

In Section 4, we argue that the scope of QHD is limited since there are no subjects who are not EDU rational but QHD rational in AS data. However, when we allow subjects to treat all $t \leq 7$ as the present, three subjects become strictly QHD rational. We observe a similar pattern in CMW data.

## D.2 Diminishing Impatience

Following Halevy (2008) and Chakraborty et al. (forthcoming), we define the agent's *one-period impatience at $t$* by $D(t)/D(t+1)$, and consider the following two properties of the discount function $D : T \to \mathbf{R}_+$.

**Definition 6.** *A monotonically decreasing discount function $D : T \to \mathbf{R}_+$ exhibits* diminishing impatience (DI) *if*

$$\frac{D(0)}{D(1)} > \frac{D(t)}{D(t+1)} \tag{18}$$

*for every $t \geq 1$, and* strong diminishing impatience (SDI) *if*

$$\frac{D(t)}{D(t+1)} > \frac{D(t+1)}{D(t+2)} \tag{19}$$

*for every $t \geq 1$.*

The definitions of DI and SDI respectively capture the essence of PQHD and hyperbolic discounting. It is easy to see that the discounting functions of PQHD and hyperbolic discounting respectively satisfy DI and SDI.

We can test DI and SDI by adding additional rows capturing (log-linearized version of) condition (18) or (19) to matrix $B$ for the test of MTD.[2] Table D.2 presents the result. It is interesting to see that the pass rates for the test of DI are strictly larger than those of PQHD. This suggests that the limited scope of PQHD (beyond EDU) could result from the particular functional form of the PQHD model.

Table D.2: Pass rates.

| Data | # subjects | # choices | EDU | PQHD | SDI | DI | MTD |
|---|---|---|---|---|---|---|---|
| AS | 97 | 45 | 0.299 | 0.299 | 0.330 | 0.371 | 0.392 |
| CMW | 1,060 | 12 | 0.210 | 0.216 | 0.264 | 0.266 | 0.278 |

[2] Since Matlab cannot handle strict inequalities, we impose constraints

$$\log D(0) - \log D(1) \geq \log D(t) - \log D(t+1) + \varepsilon$$

in DI test and

$$\log D(t) - \log D(t+1) \geq \log D(t+1) - \log D(t+2) + \varepsilon$$

in SDI test, where $\varepsilon$ is a small slack term. Table D.2 presents pass rates when $\varepsilon = 10^{-12}$.

# E Ground Truth Analysis

We assess the basic performance of our revealed-preference tests using simulated choices. As in Andreoni and Sprenger (2012), we assume a decision maker has a utility function (CRRA with quasi-hyperbolic discounting) of the form:

$$U(x_0, \dots, x_T) = \frac{1}{\alpha} x_0^\alpha + \beta \sum_{t \in T \setminus \{0\}} \frac{1}{\alpha} \delta^t x_t^\alpha.$$

We simulate synthetic subjects' choice data in Andreoni and Sprenger's (2012) environment (i.e., time frames and budgets are identical to those actual subjects faced in their experiment) under all combinations of parameters $\alpha \in \{0.8, 0.82, \dots, 1\}$, $\delta \in \{0.95, 0.951, \dots, 1\}$, and $\beta \in \{0.8, 0.82, \dots, 1.2\}$, resulting the total of 11,781 such synthetic subjects. We then perform our revealed-preference tests, in particular, tests for EDU and QHD rationality, and ask following questions: (i) do our tests correctly identify EDU or QHD rational datasets, and (ii) can our tests recover "true" underlying model parameters?

A few remarks are in order. First, for some parameter specifications, it is possible that the slope of (linear) indifference curves coincide with those of budget lines. This happens 21 times when $(\alpha, \delta) = (1, 1)$.[3] If the slope of the indifference curve coincides with the budget line (i.e., every point on the budget yields the same level of utility), we randomly pick one point from the budget as the optimal choice as a tie-breaking rule. Second, in order to avoid the rounding issue in Matlab, we treat numbers less than $10^{-6}$ to be 0. In other words, if the predicted allocation is sufficiently close to a corner, we treat it as a corner choice. Third, unlike Andreoni and Sprenger's (2012) original experiment where subjects made choices from "discrete" budget sets by allocating 100 tokens, we allow simulated choices to be at any point on the continuous budget lines. We also prepare another set of simulated choices (with the same set of parameters) which mimic behavior of the Andreoni and Sprenger's (2012) experimental subjects for the purpose of comparison.

**Test results.** The results are presented in Table E.1. We first look at our baseline simulation in which choices were made from continuous budget sets. Of the 11,781 synthetic

[3] For example, consider the case when $(\alpha, \delta, \beta) = (1, 1, 0.8)$ and $(1, 1, 0.9)$. Since the utility function has the form $x_\tau + \beta x_{\tau+d}$ when $\tau = 0$, indifference curve coincides with budget line when prices are 1.11 or 1.25. Another possibility is in the time frame $(\tau, d) = (7, 70)$, where the price of 1 (tokens allocated to sooner and later payments have the same exchange rate) is offered. In this case, indifference curve coincides with budget line as long as $(\alpha, \delta) = (1, 1)$.

TABLE E.1: Test results using simulated choice data from continuous budgets (top panel) and discrete budgets (bottom panel).

| Continuous budget | Parameters | | | Total |
|---|---|---|---|---|
| | $\alpha = 1$ | $\alpha < 1$<br>$\beta = 1$ | $\alpha < 1$<br>$\beta \neq 1$ | |
| No interior choice | 1,050 | 38 | 700 | 1,788 |
| Pass EDU | 939 | 510 | 2,501 | 3,950 |
| Pass QHD | 1,071 | 510 | 10,200 | 11,781 |
| Sample size | 1,071 | 510 | 10,200 | 11,781 |

| Discrete budget | Parameters | | | Total |
|---|---|---|---|---|
| | $\alpha = 1$ | $\alpha < 1$<br>$\beta = 1$ | $\alpha < 1$<br>$\beta \neq 1$ | |
| No interior choice | 1,050 | 252 | 4,746 | 6,048 |
| Pass EDU | 939 | 510 | 6,913 | 8,362 |
| Pass QHD | 1,071 | 510 | 10,160 | 11,741 |
| Sample size | 1,071 | 510 | 10,200 | 11,781 |

subjects, 3,950 (33.5%) passed the EDU test and 11,781 (100%) passed the QHD test.

We then split the sample into three groups. The first group of subjects have the linear utility function ($\alpha = 1$). They made no interior choices (except for the knife edge case described above), and 939 of them passed the EDU test. The second group of subjects have nonlinear utility and no present/future bias ($\alpha < 1$, $\beta = 1$). They all passed the EDU test (and hence the QHD test, too), as expected. The third group of subjects have nonlinear utility and present/future bias ($\alpha < 1$, $\beta \neq 1$). We find that 2,501 of them passed the EDU test, even though their underlying preferences were strictly present/future biased.

The bottom panel of Table E.1 presents the results with simulated data when choices are assumed to be on the discrete points on the budget lines. As one can imagine, the number of synthetic subjects who make no interior choices increases and accordingly the pass rate for the EDU test increases from 33.5% to 71.0%. We also find that "perturbations" induced by discretization of budget sets is powerful enough for some of the subjects to become QHD non-rational.

# F Jittering: Perturbing Choices

We look at the robustness of the results from our revealed-preference tests. We "jitter the data by adding white noise to the observed choices, following the idea introduced by Andreoni et al. (2013).

Assume a QHD model

$$U(x_0, \ldots, x_T) = \frac{1}{\alpha} x_0^\alpha + \beta \sum_{t=1}^{T} \delta^t \frac{1}{\alpha} x_t^\alpha$$

as in AS. For each budget in the AS experiment (there are 45 of those), the model predicts demand for sooner payment, $x(p, \tau, d; \alpha, \delta, \beta)$. We then add "jitters" to these predicted demands so that we observe $\widehat{x}(p, \tau, d; \alpha, \delta, \beta, \sigma) = x(p, \tau, d; \alpha, \delta, \beta) + \varepsilon$. Jitters are assumed to be drawn from a normal distribution, but we ensure that the jittered demand $\widehat{x}(p, \tau, d)$'s are on the budget line. In other words, jitters are drawn from a truncated normal distribution.[4]

In this exercise, we take parameters from AS aggregate estimates: $\alpha = 0.897$, $\delta = 0.999$. For the present bias parameter, we take AS aggregate estimate $\beta = 1.007$ together with other "reasonable" values such as 0.974 (aggregate estimate from Augenblick et al., 2015), 0.995, 1, and 1.05. As for standard deviation of the normal distribution, we use $\sigma \in \{0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1\}$.

For each set of parameters and standard deviation of white noise $(\alpha, \delta, \beta, \sigma)$, we simulate 1,000 sets of observations $\{\widehat{x}(p_b, \tau_b, d_b; \alpha, \delta, \beta, \sigma)\}_{b=1}^{45}$. We then perform our EDU and QHD tests.

Table F.1 reports pass rates for the QHD test for each set of parameters and standard deviation. When the standard deviation is $\sigma = 0.001$, the simulated dataset always pass the QHD test. As the standard deviation increases, pass rates decrease at the speed depending on the parameter configuration.[5]

Table F.2 reports the same statistics for the EDU test. A notable feature in this simulation is that the dataset generated by non-EDU preferences (i.e., $\beta = 0.995$ and $1.007$) pass the EDU test in many occasions. As in the case of the QHD test, pass rates decrease at the speed depending on the parameter configuration.

This exercise has demonstrated that our revealed-preference tests detect irregularities

[4]Andreoni et al. (2013) note that "truncating is known to bias the frequency of corner solutions downward". An alternative approach is "censoring," which would have a bias in the opposite direction.

[5]We also confirm that predicted choices indeed pass the QHD test in the absence of jittering (4th column in the table).

Table F.1: QHD test pass rates.

| | Parameters | | | Standard deviation ($\sigma$) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| # | $\alpha$ | $\delta$ | $\beta$ | 0 | 0.001 | 0.005 | 0.010 | 0.050 | 0.100 | 0.500 | 1.000 |
| 1 | 0.897 | 0.999 | 0.974 | 1.00 | 1.00 | 0.98 | 0.84 | 0.21 | 0.01 | 0.00 | 0.00 |
| 2 | 0.897 | 0.999 | 0.995 | 1.00 | 1.00 | 1.00 | 1.00 | 0.43 | 0.17 | 0.00 | 0.00 |
| 3 | 0.897 | 0.999 | 1.000 | 1.00 | 1.00 | 1.00 | 1.00 | 0.49 | 0.17 | 0.00 | 0.00 |
| 4 | 0.897 | 0.999 | 1.007 | 1.00 | 1.00 | 1.00 | 0.98 | 0.30 | 0.08 | 0.00 | 0.00 |
| 5 | 0.897 | 0.999 | 1.050 | 1.00 | 1.00 | 1.00 | 0.91 | 0.19 | 0.05 | 0.00 | 0.00 |

Table F.2: EDU test pass rates.

| | Parameters | | | Standard deviation ($\sigma$) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| # | $\alpha$ | $\delta$ | $\beta$ | 0 | 0.001 | 0.005 | 0.010 | 0.050 | 0.100 | 0.500 | 1.000 |
| 1 | 0.897 | 0.999 | 0.974 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2 | 0.897 | 0.999 | 0.995 | 1.00 | 1.00 | 1.00 | 1.00 | 0.43 | 0.17 | 0.00 | 0.00 |
| 3 | 0.897 | 0.999 | 1.000 | 1.00 | 1.00 | 1.00 | 1.00 | 0.49 | 0.17 | 0.00 | 0.00 |
| 4 | 0.897 | 0.999 | 1.007 | 1.00 | 1.00 | 1.00 | 0.96 | 0.25 | 0.08 | 0.00 | 0.00 |
| 5 | 0.897 | 0.999 | 1.050 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

induced by white noise, but we cannot provide a definitive answer to whether the degree of irregularities necessary to violate EDU/QHD rationality is big or small (in other words, how sensitive our tests are) because we do not have a clear benchmark to compare with.

As provide standard deviation of NLS error in the aggregate estimate (corresponding to parameter set #4), which is 6.13.

Alternatively, one can use variations observed in the actual experimental data to compare with standard deviations used in this exercise. Let $x_i(p_b, \tau_b, d_b)$ denote subject $i$'s demand for sooner payment in budget $b$. Then, we calculate the root mean squared error (RMSE)

$$v_i = \sqrt{\frac{1}{45}\sum_{b=1}^{45}\big(x_i(p_b,\tau_b,d_b) - x(p_b,\tau_b,d_b;\alpha,\delta,\beta)\big)^2}$$

for each subject $i$. Table F.3 reports summary statistics for the distribution of $v_i$'s. It is clear that the variation of the observed data measured by RMSE is much higher than the standard deviation of white noise at which we achieve 50% pass rate for the QHD test. This may suggest that about 50% of the subjects are not rationalized by QHD model because of structural irregularities rather than trembling on their choices. However, we emphasize

Table F.3: Distributions of $v_i$'s.

| | Parameters | | | Percentile | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| # | $\alpha$ | $\delta$ | $\beta$ | 5-th | 10-th | 25-th | 50-th | 75-th | 90-th | 95-th |
| 1 | 0.897 | 0.999 | 0.974 | 3.00 | 3.76 | 4.68 | 5.93 | 6.33 | 7.83 | 10.50 |
| 2 | 0.897 | 0.999 | 0.995 | 2.91 | 3.66 | 4.60 | 5.93 | 6.17 | 7.94 | 10.61 |
| 3 | 0.897 | 0.999 | 1.000 | 2.93 | 3.68 | 4.63 | 5.94 | 6.15 | 7.97 | 10.64 |
| 4 | 0.897 | 0.999 | 1.007 | 2.95 | 3.71 | 4.62 | 5.91 | 6.18 | 8.02 | 10.67 |
| 5 | 0.897 | 0.999 | 1.050 | 3.10 | 3.58 | 4.48 | 5.61 | 6.13 | 8.28 | 10.92 |

again that we do not have clear guidance for the benchmark: we demonstrate the case of $v_i$'s but this may not be the right one to compare with.

# References


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

Andreoni, J., W. T. Harbaugh, and B. J. Gillen (2013): "The Power of Revealed Preference Tests: Ex-Post Evaluation of Experimental Design," Unpublished manuscript.

Andreoni, J. and J. Miller (2002): "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism," *Econometrica*, 70, 737–753.

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

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

Balakrishnan, U., J. Haushofer, and P. Jakiela (forthcoming): "How Soon Is Now? Evidence of Present Bias from Convex Time Budget Experiments," *Experimental Economics*.

Beatty, T. K. M. and I. A. Crawford (2011): "How Demanding is the Revealed Preference Approach to Demand?" *American Economic Review*, 101, 2782–2795.

Bronars, S. G. (1987): "The Power of Nonparametric Tests of Preference Maximization," *Econometrica*, 55, 693–698.

Chakraborty, A., Y. Halevy, and K. Saito (forthcoming): "The Relation between Behavior under Risk and over Time," *American Economic Review: Insights*.

Choi, S., R. Fisman, D. Gale, and S. Kariv (2007): "Consistency and Heterogeneity of Individual Behavior under Uncertainty," *American economic Review*, 97, 1921–1938.

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

Crawford, I. and B. De Rock (2014): "Empirical Revealed Preference," *Annual Review of Economics*, 6, 503–524.

Dean, M. and D. Martin (2016): "Measuring Rationality with the Minimum Cost of Revealed Preference Violations," *Review of Economics and Statistics*, 98, 524–534.

DellaVigna, S. (2018): "Structural Behavioral Economics," in *Handbook of Behavioral Economics - Foundations and Applications*, ed. by B. D. Bernheim, S. DellaVigna, and

D. Laibson, Amsterdam: Elsevier/North-Holland, vol. 1, 613–725.

Ericson, K. M. and D. Laibson (2019): "Intertemporal Choice," in *Handbook of Behavioral Economics - Foundations and Applications*, ed. by B. D. Bernheim, S. DellaVigna, and D. Laibson, Amsterdam: Elsevier/North-Holland, vol. 2, 1–67.

Fisman, R., S. Kariv, and D. Markovits (2007): "Individual Preferences for Giving," *American Economic Review*, 97, 1858–1876.

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

Rockafellar, R. T. (1997): *Convex Analysis*, Princeton University Press.