Discrete Choice Estimation of Time Preferences

José Apesteguía
Miguel Angel Ballester

Agost 2014

Barcelona GSE Working Paper Series

Working Paper nº 787

DISCRETE CHOICE ESTIMATION OF TIME PREFERENCES*

JOSE APESTEGUIA[†] AND MIGUEL A. BALLESTER[‡]

Abstract. Discrete choice methods are often used for the estimation of time preferences. We show that these methods have pervasive problems when based on random utility models, for which cases our results establish that the probability of selecting a later option over an earlier one may be greater for higher levels of impatience. This could have profound implications, not only in the experimental estimation of time preferences, but also in a wide variety of empirical papers using such models in dynamic settings. Alternatively, we also show that discrete choice methods built on random preference models are always free of all such problems.

Keywords: Discrete Choice; Structural Estimation; Time; Discounting; Random Utility Models; Random Preference Models.
JEL classification numbers: C25; D90.

1. Introduction

The empirical assessment of time preferences is essential for a proper understanding of individual behavior and its implications in a wide range of important economic settings. The necessary estimation exercise often entails choice situations which involve an individual selecting which ever she considers the best of a finite number of options giving monetary payoffs with different degrees of delay. Thus, the literature uses a variety of micro-econometric discrete choice methods for the estimation of time preferences.[1] The purpose of this paper is to inquire into the validity of these methods.

Date: August, 2014.

*We thank Holger Gerhardt for helpful comments and Florens Odendahl and Dijana Zejcirovic for outstanding research assistance. Financial support by the Spanish Ministry of Science and Innovation (ECO2012-34202, ECO2011-25295, ECO2010-09555-E, ECO2008-04756–Grupo Consolidado C) and FEDER is gratefully acknowledged.

[†]ICREA, Universitat Pompeu Fabra and Barcelona GSE. E-mail: `jose.apesteguia@upf.edu`.

[‡]Universitat Autonoma de Barcelona and Barcelona GSE. E-mail: `miguelangel.ballester@uab.es`.

[1]See McFadden (2001) for a survey and Train (2009) for a textbook introduction to discrete choice methods in general settings, not necessarily involving time.

Random utility models are by far the most widely used discrete choice methods for the estimation of time preferences. In one strand of this literature, the probability of choice is built upon differences in the discounted utilities of the available options (RUM-DDU). This method is used in experimental papers to obtain time-discounting estimates, both in static settings, as in Chabris et al. (2008), Ida and Goto (2009), or Tanaka, Camerer and Nguyen (2010), and in conjunction with learning in a dynamic setting, as in Toubia et al. (2013). In applied settings, this method has become the key component of a large number of dynamic discrete choice models, starting with the seminal papers by Wolpin (1984) and Rust (1987). Since then, it has often been used to address a wide variety of issues such as fertility (Ahn, 1995), health (Gilleskie, 1998; Crawford and Shum, 2005), labor (Berkovec and Stern, 1991; Rust and Phelan, 1997), or political economy (Diermeier, Keane and Merlo, 2005).[2] An alternative strand of literature uses the logarithmic transformation of the discounted utilities, which ultimately implies that the probabilities of choice depend on the ratio of discounted utilities (RUM-RDU). In this line see Andersen et al. (2008) and Meier and Sprenger (2014).

In this paper, we show that both the use of RUM-DDU and that of RUM-RDU can pervasively affect the estimation exercise. To see this, consider the simplest possible scenario, in which the individual has to choose between two streams, an earlier one $\mathbf{e}$ and a later one $\mathbf{l}$, which differ only in that option $\mathbf{e}$ offers a larger monetary payoff than $\mathbf{l}$ with a shorter delay, while $\mathbf{l}$ offers a larger monetary payoff than $\mathbf{e}$ with a longer delay. Although a less patient individual should be less likely to wait for the later option $\mathbf{l}$, we show in Corollaries 2 and 3 that the RUM-DDU and the RUM-RDU probabilities of selecting option $\mathbf{l}$ may increase when the level of impatience grows. We also study the most influential parametric families of discounted utility, including the standard power function, and also the behavioral hyperbolic and $\beta - \delta$ functions.[3] All these families are built around a discount factor that measures impatience, such that larger discount factors are associated with greater levels of impatience. In Theorems 1 and 2 we show that, for a wide array of options $\mathbf{e}$ and $\mathbf{l}$, there is a discount level above which greater impatience comes with a higher probability of selecting the option that requires more patience. This obviously poses a serious practical estimation problem.

[2]The literature using dynamic discrete choice models is large; see Aguirregabiria and Mira (2010) for a survey.

[3]See Ainslie (91), Loewenstein and Prelec (1992), Laibson (1997), and O'Donoghue and Rabin (1999) for papers proposing behavioral discounted utility functions.

We then turn to the study of random preference models (RPM), where a family of discounted utilities is considered and the individual's choice behavior is modeled as a probability distribution over the family, in which the mean is interpreted as the discount factor of that individual. This probability distribution over the family, in turn determines the choice probabilities over the options. Coller and Williams (1999) and Warner and Pleeter (2001) are two examples of the use of this approach. In Theorem 3 we establish that random preference models are completely free of the above-mentioned estimation problems. That is, when the individual becomes more impatient, or equivalently, when the mean of the distribution increases, the probability of choosing the later option $\mathbf{l}$ according to the random preference model is always decreasing.

In sum, we show that there are pervasive problems in random utility models, while random preference models are completely free of the problems studied in this paper. In Apesteguia and Ballester (2014) we study the related case of the discrete choice estimation of risk aversion, showing that random utility models face problems analogous to those identified here. Specifically, the probability of taking a riskier gamble may increase with the level of risk aversion. Moreover, the findings in Apesteguia and Ballester (2014) with respect to random preference models are as positive as those reported here.

The rest of the paper is organized as follows. Section 2 introduces the notation and the main definitions used in the rest of the paper. Section 3 covers the two random utility models discussed here, and Section 4 analyzes random preference models. Section 5 concludes the paper.

2. Preliminaries

An option or stream $\mathbf{x} = (x_0, x_1, \ldots, x_T)$ describes the sum of money or income at every time t, $x_t \in \mathbb{R}_+$. The discounted utility of a stream is $U(\mathbf{x}) = \sum_t D(t)u(x_t)$, where the utility function over monetary outcomes $u : \mathbb{R}_+ \to \mathbb{R}_+$ is strictly increasing, with $u(0) = 0$, and continuous, and the discount function over time $D : \mathbb{Z}_+ \to (0, 1]$ is strictly decreasing, with $D(0) = 1$, and such that $\lim_{t\to\infty} D(t) = 0$.[4]

[4]Whether streams are finite or infinite is irrelevant for the results in this paper. The standard boundedness conditions would be required if infinite streams were used.

Notice that this is a broad family of discounted utility functions. First, it covers the standard discount function: the power function $D^{pow}_{\delta}(t) = \frac{1}{(1+\delta)^t}$.[5] Second, it also covers influential discount functions such as the hyperbolic discounting $D^{hyp}_{\delta}(t) = \frac{1}{1+\delta t}$, or the $\beta - \delta$ preference where $D^{beta}_{\delta}(0) = 1$ and $D^{beta}_{\delta}(t) = \beta D^{pow}_{\delta}(t)$ whenever $t > 0$, with $\beta \in (0, 1]$.[6]

We now describe a way to compare two utility functions in terms of impatience. To do so, we consider simple pairs of streams where there is a unique conflict between obtaining a larger monetary payoff with a shorter delay, or some other monetary payoff with a longer delay. That is, we focus on pairs, $\mathbf{e}$ and $\mathbf{l}$, for which $e_t = l_t$ except for two periods $t_e < t_l$, with $e_{t_e} > l_{t_e}$ and $e_{t_l} < l_{t_l}$. We then say that U_1 is more impatient than U_2 if, for every pair of streams $\mathbf{e}$ and $\mathbf{l}$, whenever $U_2(\mathbf{e}) \geq U_2(\mathbf{l})$ then $U_1(\mathbf{e}) \geq U_1(\mathbf{l})$. That is, whenever the less impatient utility U_2 prefers the earlier option, so does the more impatient utility U_1. This is analogous to the standard definition of more risk aversion.[7]

The main purpose of this paper is to discuss the discrete choice estimation of time preferences. In order to isolate the effect due to the discounting of time, we assume throughout the paper that the curvature of the utility function over monetary outcomes u is fixed.

3. Random Utility Models

Let $V(\mathbf{x})$ denote either the discounted utility $U(\mathbf{x})$, or the logarithmic transformation of the discounted utility $\log(U(\mathbf{x}))$. In the random utility model approach, whether based on differences in discounted utilities (RUM-DDU) or on the ratio of discounted utilities (RUM-RDU), the valuation of the stream $\mathbf{x}$ is given by the additive consideration of $V(\mathbf{x})$ and a random i.i.d. unobserved term $\epsilon(\mathbf{x})$, which follows a continuous cumulative distribution Ψ. Normalizing the variance of errors, the probability of selecting $\mathbf{l}$ over $\mathbf{e}$ is given by $f^{V}_{\Psi[\lambda]}(\mathbf{l}, \mathbf{e}) = P(\lambda V(\mathbf{l}) + \epsilon(\mathbf{l}) \geq \lambda V(\mathbf{e}) + \epsilon(\mathbf{e})) = P(\epsilon(\mathbf{e}) - \epsilon(\mathbf{l}) \leq$

---

[5]Clearly, the exponential function $D^{exp}_{\hat{\delta}}(t) = \exp^{-\hat{\delta}t}$ is equivalent to the power function by considering $\hat{\delta} = \log(1 + \delta)$. This alternative representation is therefore omitted.

[6]The alternative representation based on the exponential function is sometimes called quasi-hyperbolic. That is, $D^{qh}_{\delta}(0) = 1$ and $D^{qh}_{\delta}(t) = \beta D^{exp}_{\delta}(t)$ whenever $t > 0$, with $\beta \in (0, 1]$. For the same reason as in the previous footnote, we can omit this functional form.

[7]Benoît and Ok (2007) provide a systematic study of the notion of more impatience. See also Horowitz (1992).

$\lambda(V(\mathbf{l})-V(\mathbf{e}))) = \Psi^*(\lambda(V(\mathbf{l})-V(\mathbf{e})))$, where Ψ^* is the distribution function of the difference of errors, and hence has mean zero.[8]

When $V(\mathbf{x})$ denotes the discounted utility $U(\mathbf{x})$, the model describes the choice probability as a function of the difference between the discounted utilities of the two options. Alternatively, when $V(\mathbf{x})$ denotes the logarithmic transformation of the discounted utility $\log(U(\mathbf{x}))$, note that $\Psi^*(\lambda(\log(U(\mathbf{l}))-\log(U(\mathbf{e})))) = \Psi^*(\lambda\log(\frac{U(\mathbf{l})}{U(\mathbf{e})}))$, and hence, in this case, the choice probabilities depend on the ratio of the discounted utilities.[9] Then, a standard maximum likelihood technique is used to determine the attitude towards time that best fits the data in each case.

3.1. **Differences in Discounted Utilities.** We start the analysis of the RUM-DDU model by establishing a simple condition characterizing the case where the probability of choosing option $\mathbf{l}$ over $\mathbf{e}$ is lower for one discounted utility U_1 than for another U_2. The condition uses the function $D_{1-2} = D_1 - D_2$, that indicates how much the two utility functions differ in terms of their time-discounting behavior.

Lemma 1. *Consider any two discounted utilities U_1, U_2. Consider any distribution Ψ, and any $\lambda > 0$. Then, $f^{U_1}_{\Psi[\lambda]}(\mathbf{l},\mathbf{e}) \leq f^{U_2}_{\Psi[\lambda]}(\mathbf{l},\mathbf{e})$ if and only if $D_{1-2}(t_e)(u(l_{t_e}) - u(e_{t_e})) + D_{1-2}(t_l)(u(l_{t_l}) - u(e_{t_l})) \leq 0$.*

Proof of Lemma 1: Note that, by definition, $f^{U_1}_{\Psi[\lambda]}(\mathbf{l},\mathbf{e}) \leq f^{U_2}_{\Psi[\lambda]}(\mathbf{l},\mathbf{e})$ if and only if $\Psi^*(\lambda(U_1(\mathbf{l})-U_1(\mathbf{e}))) \leq \Psi^*(\lambda(U_2(\mathbf{l})-U_2(\mathbf{e})))$ which holds if and only if $U_1(\mathbf{l})-U_1(\mathbf{e}) \leq U_2(\mathbf{l})-U_2(\mathbf{e})$. Now, the assumptions on discounted utilities enable the latter inequality to be written as $\sum_t D_1(t)(u(l_t)-u(e_t)) \leq \sum_t D_2(t)(u(l_t)-u(e_t))$, or, equivalently, as $\sum_t D_{1-2}(t)(u(l_t)-u(e_t)) \leq 0$. Given the structure of $\mathbf{l}$ and $\mathbf{e}$, the inequality in this case can be written simply as $D_{1-2}(t_e)(u(l_{t_e}) - u(e_{t_e})) + D_{1-2}(t_l)(u(l_{t_l}) - u(e_{t_l})) \leq 0$. This proves the lemma.■

We are now in a position to consider the case where U_1 is more impatient than U_2. Our first result is positive. Corollary 1 establishes that, whenever the moment at which the earlier option $\mathbf{e}$ offers the extra monetary payoff is the present, that is

[8] The parameter λ corresponds to the inverse of the variance of the initial distribution and is typically interpreted as a rationality parameter. The larger λ, the more rational the individual. Whenever λ goes to zero, choices become completely random, while when λ goes to infinity, choices become deterministic.

[9] Notice that this analysis implicitly assumes that $U(\mathbf{x}) > 0$ for every option $\mathbf{x}$. This is the case for every pair of streams $\mathbf{e}$ and $\mathbf{l}$ considered in this paper.

$t_e = 0$, the probability of choosing the latter option $\mathbf{l}$ over the earlier one $\mathbf{e}$ is lower for the more impatient discounted utility U_1. This result follows from the condition established in Lemma 1. Notice that, since the present is not discounted, all that matters in this case are the discounted utilities of period t_l. Now, in Corollary 1 we show that, whenever U_1 is more impatient than U_2, it follows that $D_{1-2}(t_l) \leq 0$ and, since $u(l_{t_l}) - u(e_{t_l}) > 0$, the condition of Lemma 1 holds. Importantly, Chabris et al. (2008) and Tanaka, Camerer and Nguyen (2010) use RUM-DDU for comparing $\mathbf{e}$ and $\mathbf{l}$ options with $t_e = 0$. Corollary 1 shows that there is no estimation problem in these cases.

Corollary 1. *Consider any two discounted utilities such that U_1 is more impatient than U_2. Consider any distribution Ψ, and any $\lambda > 0$. If $t_e = 0$, then $f^{U_1}_{\Psi[\lambda]}(\mathbf{l}, \mathbf{e}) \leq f^{U_2}_{\Psi[\lambda]}(\mathbf{l}, \mathbf{e})$.*

Proof of Corollary 1: We start by proving that, when U_1 is more impatient than U_2, it must be that $D_1(t) \leq D_2(t)$ for all $t > 0$. To see this, suppose, by contradiction, that there exists $t^* > 0$ such that $1 \geq D_1(t^*) > D_2(t^*)$. Since $u(0) = 0$ and u is continuous, we can find two monetary outcomes x and y such that $D_1(t^*) > \frac{u(x)}{u(y)} > D_2(t^*)$. Define $\mathbf{e} = (x, 0, 0, \dots)$ and $\mathbf{l}$ as $l_{t^*} = y$ and $l_t = 0$ otherwise. Hence, since $D_1(0) = D_2(0) = 1$, we have that $U_1(\mathbf{l}) = D_1(t^*)u(y) > u(x) = U_1(\mathbf{e})$, but, at the same time, $U_2(\mathbf{l}) = D_2(t^*)u(y) < u(x) = U_2(\mathbf{e})$, contradicting the fact that U_1 is more impatient than U_2.

Now, $t_e = 0$ implies that $D_{1-2}(t_e) = 0$ and hence $D_{1-2}(t_e)(u(l_{t_e}) - u(e_{t_e})) + D_{1-2}(t_l)(u(l_{t_l}) - u(e_{t_l})) = D_{1-2}(t_l)(u(l_{t_l}) - u(e_{t_l}))$. We know that $D_{1-2}(t_l) \leq 0$ and, since u is strictly increasing, we have that $u(l_{t_l}) - u(e_{t_l}) > 0$. Hence, $D_{1-2}(t_e)(u(l_{t_e}) - u(e_{t_e})) + D_{1-2}(t_l)(u(l_{t_l}) - u(e_{t_l})) \leq 0$ and Lemma 1 guarantees that $f^{U_1}_{\Psi[\lambda]}(\mathbf{l}, \mathbf{e}) \leq f^{U_2}_{\Psi[\lambda]}(\mathbf{l}, \mathbf{e})$, as desired.■

Corollary 2 is bad news, since it shows that for any two discounted utilities such that U_1 is more impatient than U_2, and for any distribution of errors, there are always pairs of options, $\mathbf{l}$ and $\mathbf{e}$, such that the later option $\mathbf{l}$ is selected with a higher probability by the more impatient utility. The proof of the result uses a simple pair of options that quite naturally sort patient from impatient individuals. It uses options in which there is a reward for willingness to wait, that is, $e_{t_e} - l_{t_e} < l_{t_l} - e_{t_l}$. The intuition of this result goes as follows. In general, when the higher payoff of option $\mathbf{l}$ is sufficiently delayed, it is valued almost equally by both utility functions. Hence, the only relevant factor is the numerical evaluation of option $\mathbf{e}$, which will be smaller when discounting is higher. This causes the later option $\mathbf{l}$ to be chosen with higher probability by the more impatient utility function.

Corollary 2. *Consider any two different discounted utilities such that U_1 is more impatient than U_2. Consider any distribution Ψ, and any $\lambda > 0$. There exist streams* $\mathbf{e}$ *and* $\mathbf{l}$*, such that* $f^{U_1}_{\Psi[\lambda]}(\mathbf{l}, \mathbf{e}) > f^{U_2}_{\Psi[\lambda]}(\mathbf{l}, \mathbf{e})$.

Proof of Corollary 2: Since U_1 is more impatient than U_2, we know from the proof in Corollary 1 that $D_{1-2}(t) \leq 0$. Since $U_1 \neq U_2$, there exists $t_1 > 0$ such that $D_{1-2}(t_1) < 0$. Since $\lim_{t\to\infty} D_1(t) = \lim_{t\to\infty} D_2(t) = 0$, it is obvious that $\lim_{t\to\infty} D_{1-2}(t) = 0$. There must exist $t_2 > t_1$ such that $D_{1-2}(t_1) < D_{1-2}(t_2) \leq 0$. Since $u(0) = 0$ and u is continuous, we can find two monetary outcomes x and y, with $x < y$, such that $D_{1-2}(t_1)u(x) < D_{1-2}(t_2)u(y)$. Define $\mathbf{e}$ by $e_{t_1} = x$ and $e_t = 0$ otherwise, and define $\mathbf{l}$ by $l_{t_2} = y$ and $l_t = 0$ otherwise. Notice that $D_{1-2}(t_e)(u(l_{t_e}) - u(e_{t_e})) + D_{1-2}(t_l)(u(l_{t_l}) - u(e_{t_l})) = -D_{1-2}(t_1)u(x) + D_{1-2}(t_2)u(y) > 0$. We can use Lemma 1 to conclude that $f^{U_1}_{\Psi[\lambda]}(\mathbf{l}, \mathbf{e}) > f^{U_2}_{\Psi[\lambda]}(\mathbf{l}, \mathbf{e})$, as desired.■

We now exploit the parametric structure of the families of discount functions defined in Section 2, in order to get further results. These families are convenient because the level of impatience is well-ordered parametrically. We show that the problem is not specific to particular streams, but almost generic. For a very large class of pairs of streams $\mathbf{e}$ and $\mathbf{l}$ there exists a level of impatience, δ^*, such that the probability of choosing $\mathbf{l}$ increases above δ^*. This is obviously a serious conceptual problem. It implies moreover, that there is a maximum level of discounting, δ^*, that can be estimated no matter how impatient the individual, and also that some choice probabilities may be associated to two different discount factors. The result establishes that, for the power discount function and the $\beta - \delta$ discount function, these problems involve every single pair of streams, $\mathbf{e}$ and $\mathbf{l}$, where the extra payoff of the earlier option $\mathbf{e}$ takes place at any point in time other than the present, $t_e > 0$. In the case of the hyperbolic discount function, the condition involves further restrictions on the valuation of the extra payoffs, namely $\frac{u(e_{t_e})-u(l_{t_e})}{u(l_{t_l})-u(e_{t_l})} > \frac{t_e}{t_l} > 0$. As we illustrate below, however, it is easy to find examples satisfying the condition.

Theorem 1. *Consider any continuous distribution Ψ, and any $\lambda > 0$. Let $\omega \in \{pow, beta\}$ (respectively, $\omega = hyp$). If $t_e > 0$ (respectively, $\frac{u(e_{t_e})-u(l_{t_e})}{u(l_{t_l})-u(e_{t_l})} > \frac{t_e}{t_l} > 0$), there exists δ^*_ω such that $f^{U^\omega_\delta}_{\Psi[\lambda]}(\mathbf{l}, \mathbf{e})$ is strictly increasing in $[\delta^*_\omega, +\infty)$.*

Proof of Theorem 1: Notice that we can reason locally, as in Lemma 1, to conclude that $f^{U^\omega_\delta}_{\Psi[\lambda]}(\mathbf{l}, \mathbf{e})$ is strictly increasing at a given value δ if and only if the derivative of

$D^\omega_\delta(t_e)(u(l_{t_e}) - u(e_{t_e})) + D^\omega_\delta(t_l)(u(l_{t_l}) - u(e_{t_l}))$ with respect to the discount factor is strictly positive at that value of δ. We denote $\frac{\partial D^\omega_\delta(t)}{\partial \delta}$ by $d^\omega_\delta(t)$. We need to analyze the sign of $d^\omega_\delta(t_e)(u(l_{t_e}) - u(e_{t_e})) + d^\omega_\delta(t_l)(u(l_{t_l}) - u(e_{t_l}))$. It is not difficult to see that, whenever $t_e > 0$, $d^\omega_\delta(t_e)$ is always strictly negative, and hence $f^{U^\omega_\delta}_{\Psi[\lambda]}(\mathbf{l}, \mathbf{e})$ is strictly increasing at δ if and only if $\frac{u(e_{t_e})-u(l_{t_e})}{u(l_{t_l})-u(e_{t_l})} > \frac{d^\omega_\delta(t_l)}{d^\omega_\delta(t_e)}$. Now, we can compute:

$$\frac{d^\omega_\delta(t_l)}{d^\omega_\delta(t_e)} = \begin{cases} \frac{t_l}{t_e}(1+\delta)^{t_e - t_l} \text{ if } \omega \in \{pow, beta\} \\ \frac{t_l}{t_e}\left(\frac{1+t_e\delta}{1+t_l\delta}\right)^2 \text{ if } \omega = hyp \end{cases}$$

Notice that, for any family of discount functions, $\frac{d^\omega_\delta(t_l)}{d^\omega_\delta(t_e)}$ is continuous and strictly decreasing for all values of δ. Also, notice that:

$$\lim_{\delta \to \infty} \frac{d^\omega_\delta(t_l)}{d^\omega_\delta(t_e)} = \begin{cases} 0 \text{ if } \omega \in \{pow, beta\} \\ \frac{t_e}{t_l} \text{ if } \omega \in \{hyp\} \end{cases}$$

Hence, if $\frac{u(e_{t_e})-u(l_{t_e})}{u(l_{t_l})-u(e_{t_l})} \geq \frac{t_l}{t_e}$, define $\delta^*_\omega = 0$. If $\frac{u(e_{t_e})-u(l_{t_e})}{u(l_{t_l})-u(e_{t_l})} < \frac{t_l}{t_e}$, define δ^*_ω as the unique value such that $\frac{u(e_{t_e})-u(l_{t_e})}{u(l_{t_l})-u(e_{t_l})} = \frac{d^\omega_{\delta^*_\omega}(t_l)}{d^\omega_{\delta^*_\omega}(t_e)}$. Notice that, given the computed limits, this value always exists for $\omega \in \{pow, beta\}$ and, provided that $\frac{u(e_{t_e})-u(l_{t_e})}{u(l_{t_l})-u(e_{t_l})} > \frac{t_e}{t_l}$, it also exists for the hyperbolic discount function.■

The intuition for Theorem 1 is the following. When the discount factor of the individual increases sufficiently, the utility of both options is discounted to the degree that the difference between the two discounted utilities starts to decrease and becomes almost null. Hence, larger discount factors lead to less discrimination between the superior and the inferior options, such that the later option $\mathbf{l}$ is chosen with higher probability.

Figure 1 illustrates the problems characterized in Theorem 1. Let us consider two streams realizing payoffs with different degrees of delay: no delay, a 7-day delay, a 14-day delay, and a 21-day delay. In both streams the regular payoff is 1. Now, stream $\mathbf{e}$ offers 1 extra payoff with a 14-day delay, while stream $\mathbf{l}$ offers 1.1 extra payoff with a 21-day delay. That is, $e_0 = e_7 = e_{21} = 1$, $e_{14} = 2$, $l_0 = l_7 = l_{14} = 1$, $l_{21} = 2.1$, and set all other payoffs equal to 0. Assume a logistic probability distribution and that $\lambda = 20$. Figure 1 plots the probability of choosing option $\mathbf{l}$ as a function of the discount factor δ, for the power discount function, the $\beta - \delta$ discount function with $\beta = .7$, and the hyperbolic discount function. In all three cases, we should expect the probabilities

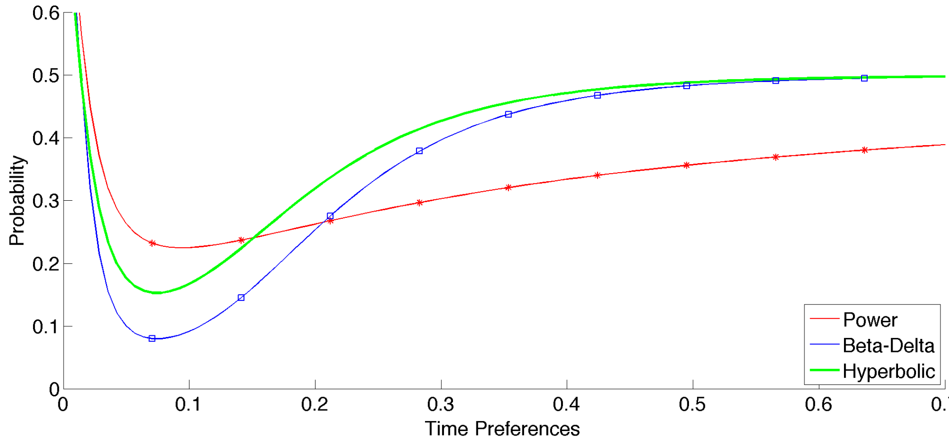

FIGURE 1. RUM-DDU probabilities of choosing $\mathbf{e}$ versus $\mathbf{l}$

of choosing $\mathbf{l}$ over $\mathbf{e}$ to decrease as the individual becomes more impatient, i.e. as δ grows. This, however, is not what happens. In all three cases, the probabilities of taking $\mathbf{l}$ as the choice decrease with δ to a certain point, and then start to increase. Notice that the corresponding critical values of the discount factor, δ^*_ω, are low, so the RUM-DDU probabilities of choosing $\mathbf{l}$ soon start increasing.[10] Hence, the use of RUM-DDU implies that the maximum discount factor that can be estimated in this case, δ^*_ω, is low, even for individuals who are so impatient as to completely disregard future payoffs. Furthermore, the figure also makes it clear that for relatively large ranges of choice probabilities there are two compatible discounting factors.

3.2. **Ratios of Discounted Utilities.** As in the previous subsection, we start by characterizing the case where the probability of choosing option $\mathbf{l}$ over $\mathbf{e}$ is lower for one discounted utility U_1 than for another U_2. This time, due to the ratio form, the condition identified involves the utility valuation of all the payoff delay options, and not only the two in which the streams differ.

Lemma 2. *Consider any two discounted utilities U_1, U_2. Consider any distribution Ψ, and any $\lambda > 0$. Then, $f^{\log(U_1)}_{\Psi[\lambda]}(\mathbf{l},\mathbf{e}) \leq f^{\log(U_2)}_{\Psi[\lambda]}(\mathbf{l},\mathbf{e})$ if and only if $\frac{\sum_t D_1(t)u(l_t)}{\sum_t D_1(t)u(e_t)} \leq \frac{\sum_t D_2(t)u(l_t)}{\sum_t D_2(t)u(e_t)}$.*

Proof of Lemma 2: Note that, by definition, $f^{\log(U_1)}_{\Psi[\lambda]}(\mathbf{l},\mathbf{e}) \leq f^{\log(U_2)}_{\Psi[\lambda]}(\mathbf{l},\mathbf{e})$ if and only if $\Psi^*(\lambda\log(\frac{U_1(\mathbf{l})}{U_1(\mathbf{e})})) \leq \Psi^*(\lambda\log(\frac{U_2(\mathbf{l})}{U_2(\mathbf{e})}))$ which holds if and only if $\frac{U_1(\mathbf{l})}{U_1(\mathbf{e})} \leq \frac{U_2(\mathbf{l})}{U_2(\mathbf{e})}$. Given the

[10] $\delta^*_\omega = .074$ for $\omega \in \{pow, beta\}$ and $\delta^*_{hyp} = .094$.

assumptions on the utilities, this can be written simply as $\frac{\sum_t D_1(t)u(l_t)}{\sum_t D_1(t)u(e_t)} \leq \frac{\sum_t D_2(t)u(l_t)}{\sum_t D_2(t)u(e_t)}$.■

Corollary 3 shows that, in general, the property identified in the above lemma is violated for every pair of discounted utilities where one is more impatient than the other, whatever the probability distribution of errors. As in the case of Corollary 2, the proof of the result emphasizes that the violation of the property always involves pairs of streams that naturally sort patient from impatient individuals. Notice, furthermore, that the proof uses streams such that $t_e = 0$, thus showing that the positive news given by Corollary 1 no longer holds.

Corollary 3. *Consider any two different discounted utilities such that U_1 is more impatient than U_2. Consider any distribution Ψ, and any $\lambda > 0$. If $D_{1/2}$ is not constant for all $t > 0$, there exist streams $\mathbf{e}$ and $\mathbf{l}$, such that $f^{U_1}_{\Psi[\lambda]}(\mathbf{l}, \mathbf{e}) > f^{U_2}_{\Psi[\lambda]}(\mathbf{l}, \mathbf{e})$.*

Proof of Corollary 3: Given the assumption on $D_{1/2}$, there exist delay periods $r, s > 0$ such that $D_{1/2}(s) > D_{1/2}(r)$, or, equivalently, $D_1(s)D_2(r) - D_1(r)D_2(s) > 0$. Since $u(1) > 0$, it must be that $[D_1(s)D_2(r) - D_1(r)D_2(s)][u(1)]^2 > 0$. We can choose x small enough to guarantee that $[D_1(s)D_2(r) - D_1(r)D_2(s)][u(1)]^2 + D_{1-2}(r)u(1)u(x) + D_{1-2}(s)u(1)u(x) > 0$. Now, define the stream $\mathbf{e}$ by $e_0 = x$, $e_r = 1$, and $e_t = 0$ otherwise. Define the stream $\mathbf{l}$ by $e_r = e_s = 1$ and $e_t = 0$ otherwise. It is easy to see that $[D_1(s)D_2(r) - D_1(r)D_2(s)][u(1)]^2 + D_{1-2}(r)u(1)u(x) + D_{1-2}(s)u(1)u(x) > 0$ implies that $\frac{\sum_t D_1(t)u(l_t)}{\sum_t D_1(t)u(e_t)} = \frac{D_1(r)u(1)+D_1(s)u(1)}{u(x)+D_1(r)u(1)} > \frac{D_2(r)u(1)+D_2(s)u(1)}{u(x)+D_2(r)u(1)} = \frac{\sum_t D_2(t)u(l_t)}{\sum_t D_2(t)u(e_t)}$. We can simply use Lemma 2 to conclude that $f^{\log(U_1)}_{\Psi[\lambda]}(\mathbf{l}, \mathbf{e}) > f^{\log(U_2)}_{\Psi[\lambda]}(\mathbf{l}, \mathbf{e})$, as desired.■

Now, turning our focus to parametric discount functions, we begin by establishing some positive news. We show that all the parametric families behave well when comparing certain options. Specifically, consider the class of streams of payoffs where the only positive payoffs are e_{t_e} and l_{t_l}. Denote these streams as $\mathbf{e}^0$ and $\mathbf{l}^0$. Admittedly, these are rather especial streams, since they impose 0 wealth for all delay periods other than t_e and t_l.[11]

Corollary 4. *Consider any two discounted utilities such that $U^{\omega}_{\delta_1}$ is more impatient than $U^{\omega}_{\delta_2}$. Consider any distribution Ψ, and any $\lambda > 0$. Let $\mathbf{e}^0$ and $\mathbf{l}^0$, then $f^{U^{\omega}_{\delta_1}}_{\Psi[\lambda]}(\mathbf{l}^0, \mathbf{e}^0) \leq f^{U^{\omega}_{\delta_2}}_{\Psi[\lambda]}(\mathbf{l}^0, \mathbf{e}^0)$.*

[11] Notice that these streams would be more relevant when, as in prospect theory, all that matters are variations in wealth.

Proof of Corollary 4: Given the stream structure, the characterizing condition in Lemma 2 can be written as $\frac{D^{\omega}_{\delta_1}(t_l)u(l_{t_l})}{D^{\omega}_{\delta_1}(t_e)u(e_{t_e})} \leq \frac{D^{\omega}_{\delta_2}(t_l)u(l_{t_l})}{D^{\omega}_{\delta_2}(t_e)u(e_{t_e})}$, which is simply $\frac{D^{\omega}_{\delta_1}(t_l)}{D^{\omega}_{\delta_1}(t_e)} \leq \frac{D^{\omega}_{\delta_2}(t_l)}{D^{\omega}_{\delta_2}(t_e)}$. We only need to compute

$$\frac{D^{\omega}_{\delta}(t_l)}{D^{\omega}_{\delta}(t_e)} = \begin{cases} (1+\delta)^{t_e - t_l} \text{ if } \omega \in \{pow, beta\} \\ \\ \frac{1+t_e\delta}{1+t_l\delta} \text{ if } \omega = hyp \end{cases}$$

to observe that these ratios are decreasing in δ. Hence, the result follows.■

We close this section with an analysis analogous to that of Theorem 1. We first prove the existence of a discount value δ^*, above which the probability of selecting $\mathbf{l}$ over $\mathbf{e}$ increases. The conditions are very similar to those in Theorem 1. Next, we show that new problems can now appear when impatience vanishes, i.e., for impatience levels close to $\delta = 0$.

Theorem 2. *Consider any continuous distribution Ψ, and any $\lambda > 0$.*

(1) *Let $\omega \in \{pow, beta\}$ (respectively, $\omega = hyp$). If $e_t > 0$ for some $t < t_e$ (respectively, $e_0 > 0$ and $\frac{u(e_{t_e})-u(l_{t_e})}{u(l_{t_l})-u(e_{t_l})} > \frac{t_e}{t_l}$), there exists δ^*_{ω} such that $f^{U^{\omega}_{\delta}}_{\Psi[\lambda]}(\mathbf{l},\mathbf{e})$ is strictly increasing in $[\delta^*_{\omega}, +\infty)$.*

(2) *Let $\omega \in \{pow, hyp\}$ (respectively, $\omega = beta$). If $\sum_r \sum_s (s-r)u(l_r)u(e_s) > 0$ (respectively, $\sum_s su(l_0)u(e_s) - \sum_r ru(l_r)u(e_0) + \beta \sum_{r>0}\sum_{s>0}(s-r)u(l_r)u(e_s) > 0$), there exists $\bar{\delta}_{\omega}$ such that $f^{U^{\omega}_{\delta}}_{\Psi[\lambda]}(\mathbf{l},\mathbf{e})$ is strictly increasing in $[0, \bar{\delta}_{\omega}]$.*

Proof of Theorem 2: We can reason locally, as in Lemma 2, to conclude that the monotonicity of $f^{U^{\omega}_{\delta}}_{\Psi[\lambda]}(\mathbf{l},\mathbf{e})$ depends on the derivative of $\frac{\sum_t D^{\omega}_{\delta}(t)u(l_t)}{\sum_t D^{\omega}_{\delta}(t)u(e_t)}$ with respect to δ. Clearly, the sign of this derivative is the same as that of $\sum_t d^{\omega}_{\delta}(t)u(l_t)[\sum_t D^{\omega}_{\delta}(t)u(e_t)] - \sum_t d^{\omega}_{\delta}(t)u(e_t)[\sum_t D^{\omega}_{\delta}(t)u(l_t)]$.

We begin by analyzing the case of $\omega = pow$. Since $d^{pow}_{\delta}(0) = 0$, the sign of the above expression is the same as that of $-\sum_t tD^{pow}_{\delta}(t)u(l_t)[\sum_t D^{pow}_{\delta}(t)u(e_t)] + \sum_t tD^{pow}_{\delta}(t)u(e_t)[\sum_t D^{pow}_{\delta}(t)u(l_t)]$. This is $\sum_r \sum_s (s-r)D^{pow}_{\delta}(r)D^{pow}_{\delta}(s)u(l_r)u(e_s)$ or more succinctly, $\sum_r \sum_s (s-r)D^{pow}_{\delta}(r+s)u(l_r)u(e_s)$. We start by analyzing the behavior of this expression when δ goes to infinity. Notice, first, that the expression converges to zero. To determine whether the sign is positive or negative above a certain value, consider all the sums of the form $\sum_r \sum_{s:r+s=m}(s-r)D^{pow}_{\delta}(m)u(l_r)u(e_s)$. Clearly, the one having the smallest integer m among those with a value different from zero determines the sign of the derivative when δ approaches ∞. Now, let t^* be the smallest integer such that $t^* < t_e$ with $e_{t^*} > 0$, which exists by assumption. Any sum where $m < t^* + t_e$ is equal to zero, while the sum $\sum_r \sum_{s:r+s=t^*+t_e}(s-r)u(l_r)u(e_s)$ is equal to $(t_e - t^*)u(e_{t^*})(u(e_{t_e}) - u(l_{t_e}))$, which is strictly positive by the assumptions on $\mathbf{e}$ and $\mathbf{l}$. Hence, we can find δ^*_{pow}, such that $f^{U^{pow}_\delta}_{\Psi[\lambda]}(\mathbf{l}, \mathbf{e})$ is strictly increasing in $[\delta^*_{pow}, +\infty)$. For the second claim, simply notice that, when δ goes to zero, the expression in question converges to $\sum_r \sum_s (s-r)u(l_r)u(e_s)$. Hence, if this value is strictly positive, there exists $\bar{\delta}_{pow}$ such that $f^{U^{pow}_\delta}_{\Psi[\lambda]}(\mathbf{l}, \mathbf{e})$ is strictly increasing in $[0, \bar{\delta}_{pow}]$.

Let us now consider the case of $\beta{-}\delta$ preferences. By reasoning analogous to that used in the case of the power function, the determining sign is that of $\beta \sum_s sD^{pow}_\delta(s)u(l_0)u(e_s) - \beta \sum_r rD^{pow}_\delta(r)u(l_r)u(e_0) + \beta^2 \sum_{r>0}\sum_{s>0}(s-r)D^{pow}_\delta(r+s)u(l_r)u(e_s)$. When δ approaches ∞, since $\beta > 0$, the sign is determined by the following expression $(t_e - t^*)u(e_{t^*})(u(e_{t_e}) - u(l_{t_e}))$. Since $t_e < t^*$ and $e_{t_e} > l_{t_e}$, this term is strictly positive and the result follows. When δ approaches zero, the expression in question converges to $\beta[\sum_r ru(l_0)u(e_r) - \sum_s su(l_s)u(e_0) + \beta \sum_{r>0}\sum_{s>0}(s-r)u(l_r)u(e_s) > 0]$. Positivity of β leads to the desired conclusion.

For the hyperbolic case, the same reasoning can be used to analyze the sign of $\sum_r \sum_s (s-r)[D^{hyp}_\delta(r)D^{hyp}_\delta(s)]^2 u(l_r)u(e_s)$. When δ goes to infinity, the expression converges to zero and the dominant terms are all terms in which either r or s is zero, i.e., those of the form $s[D^{hyp}_\delta(s)]^2u(l_0)u(e_s)$ and $-r[D^{hyp}_\delta(r)]^2u(l_r)u(e_0)$. To study the sign of their sum, simply notice that the limit of $\frac{D^{hyp}_\delta(a)}{D^{hyp}_\delta(b)}$, as δ grows, is b/a. Hence, the determining expression is $\sum_s \frac{1}{s}u(l_0)u(e_s) - \sum_r \frac{1}{r}u(l_r)u(e_0)$, which is equal to $\frac{1}{t_e}(u(l_0)u(e_{t_e}) - u(l_{t_e})u(e_0)) + \frac{1}{t_l}(u(l_0)u(e_{t_l}) - u(l_{t_l})u(e_0))$. Given the assumptions, this value is strictly positive, hence there exists δ^*_{hyp} such that $f^{U^{hyp}_\delta}_{\Psi[\lambda]}(\mathbf{l}, \mathbf{e})$ is strictly increasing in $[\delta^*_{hyp}, +\infty)$. When δ goes to zero, we have the same limit as in the power discounting case, and the result follows.■

Figure 2 is analogous to Figure 1. It uses the same two streams, the same probability distribution, the same value of parameter λ and the same discount functions to plot the RUM-RDU probabilities of choosing $\mathbf{l}$ over $\mathbf{e}$. It is apparent that the same sort of problems identified in Figure 1 show up in this case, and hence basically the same conclusions as reached in the previous subsection also apply here.[12]

[12] A simple example of the second case covered in Theorem 2, namely that in which the probability of choosing $\mathbf{l}$ over $\mathbf{e}$ is increasing in a range $[0, \delta^*_\omega]$, is $e_0 = 1, e_7 = 0, e_{14} = 2, e_{21} = 3, l_0 = .5, l_7 = l_{14} = 2, l_{21} = 3$ and all other payoffs are equal to 0. It can be checked that, taking the logistic

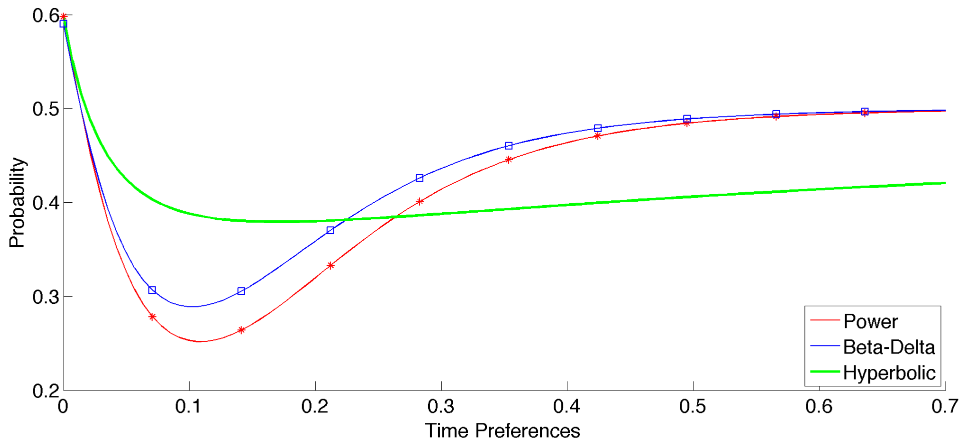

FIGURE 2. RUM-RDU probabilities of choosing $\mathbf{e}$ versus $\mathbf{l}$

4. RANDOM PREFERENCE MODELS

Random preference models (RPM) consider a parametric family $\{U_\delta\}$ of discounted utilities, where higher values of δ signify greater impatience. At the moment of choice, the discount factor or, equivalently, one of the utilities U_δ, is drawn randomly from a continuous cumulative distribution function Φ with mean $\theta \geq 0$ and variance σ^2, and the preferred option is chosen accordingly.[13] Then, the probability of selecting option $\mathbf{l}$ over option $\mathbf{e}$ is given by $f_{\Phi[\theta,\sigma^2]}(\mathbf{l},\mathbf{e}) = P(U_\delta(\mathbf{l}) \geq U_\delta(\mathbf{e})|[\theta,\sigma^2])$, and standard maximum likelihood techniques can then be used to determine the parameters θ and σ^2 that best fit the data. The θ is interpreted as the discounting factor used by the individual, while the σ^2 is interpreted as the inverse of her rationality.[14]

In the next result we establish that this method is free from the problems identified in the previous section.

distribution, for example, when δ is low, the probabilities of choosing $\mathbf{l}$ are increasing up to values of approximately $.059, .026$ and $.035$ for the power, the $\beta - \delta$ with a $\beta = .7$, and the hyperbolic discount functions, respectively. From those points on, the respective probabilities are always decreasing with δ, approaching 0 in the limit.

[13]As is customary, the (parametric) distribution Φ is such that, for a given σ^2, $\Phi[\theta_1,\sigma^2]$ first-order stochastically dominates $\Phi[\theta_2,\sigma^2]$ whenever $\theta_1 \geq \theta_2$.

[14]Notice that random preference models can be understood as non-independent random utility models. That is, any random preference model could be presented alternatively as utility values subject to non-independent errors for the different options involved. The joint distribution of errors on the options can be computed from the distribution Φ.

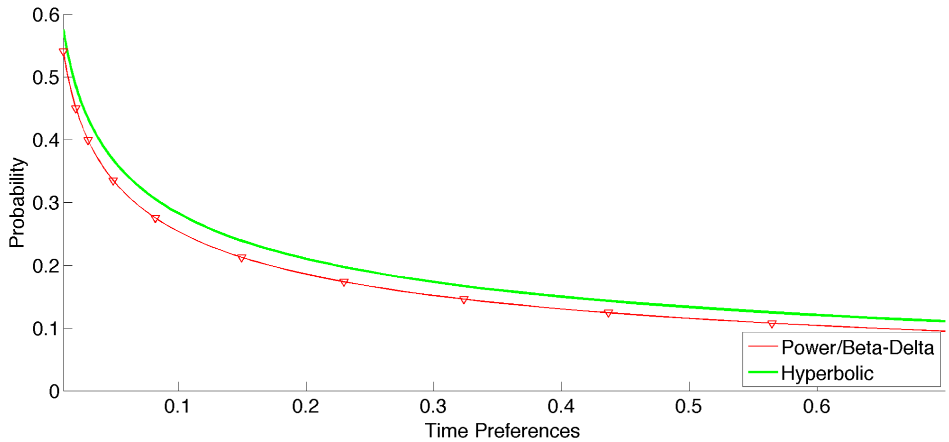

FIGURE 3. RPM probabilities of choosing $\mathbf{e}$ versus $\mathbf{l}$

Theorem 3. *Consider the parametric family of expected utility functions $\{U_\delta\}$. Consider any distribution Φ, and any $\sigma^2 > 0$. Let $\theta_1 \geq \theta_2 \geq 0$. For any pair of streams, $\mathbf{l}$ and $\mathbf{e}$, $f_{\Phi[\theta_1,\sigma^2]}(\mathbf{l},\mathbf{e}) \leq f_{\Phi[\theta_2,\sigma^2]}(\mathbf{l},\mathbf{e})$.*

Proof of Theorem 3: Consider a pair of streams, $\mathbf{e}$ and $\mathbf{l}$. Notice that, whenever $\mathbf{e}$ is better than $\mathbf{l}$ for all U_δ (respectively, $\mathbf{l}$ is better than $\mathbf{e}$ for all U_δ), the probabilities $P(U_\delta(\mathbf{l}) \geq U_\delta(\mathbf{e})|[\theta_1,\sigma^2])$ and $P(U_\delta(\mathbf{l}) \geq U_\delta(\mathbf{e})|[\theta_2,\sigma^2])$, and hence, $f_{\Phi[\theta_1,\lambda]}(\mathbf{l},\mathbf{e})$ and $f_{\Phi[\theta_2,\lambda]}(\mathbf{l},\mathbf{e})$, are equal, since they are both equal to 0 (respectively, to 1). Now suppose that there are some utilities for which $\mathbf{e}$ is better than $\mathbf{l}$ and others for which $\mathbf{l}$ is better than $\mathbf{e}$. Since $\{U_\delta\}$ is ordered by levels of impatience, there must exist $\hat{\delta}(\mathbf{l},\mathbf{e})$ such that $U_\delta(\mathbf{l}) > (<)\ U_\delta(\mathbf{e})$ whenever $\delta < (>)\ \hat{\delta}(\mathbf{l},\mathbf{e})$. Since $\theta_1 \geq \theta_2$, the distribution of parameters associated to θ_1 first-order stochastically dominates the distribution associated to θ_2. Thus, for all k, it is $P(\delta < \ k|[\theta_1,\sigma^2]) \leq P(\delta < \ k|[\theta_2,\sigma^2])$. In particular, it must be that $\Phi[\theta_1,\sigma^2](\hat{\delta}(\mathbf{l},\mathbf{e})) = P(U_\delta(\mathbf{l}) \geq U_\delta(\mathbf{e})|[\theta_1,\sigma^2]) \leq P(U_\delta(\mathbf{l}) \geq U_\delta(\mathbf{e})|[\theta_2,\sigma^2]) = \Phi[\theta_1,\sigma^2](\hat{\delta}(\mathbf{l},\mathbf{e}))$, as desired.■

Figure 3 is analogous to Figures 1 and 2, but, this time, the probability of choosing option $\mathbf{l}$ is modeled using the method of errors on the discount parameter studied in this section. Figure 3 uses the same two streams and discount functions as in the previous figures and, this time, since the error is on the discount factor and this takes values in the positive reals, we use the log normal probability distribution with a $\lambda = 3$.

The figure clearly exemplifies that the problems identified in the previous two cases vanish with the use of this method.

5. Final Remarks

We have shown here, and in our companion paper, that some popular discrete choice methods have serious logical inconsistencies when used to estimate time and risk preferences. These findings should alert attention to the direct application of sound microeconometric techniques to settings other than those originally contemplated. In future research, we intend to re-estimate the data of a number of influential papers dealing with time and risk preferences, and thereby elucidate the specific impact on the parameter estimates of the gambles and streams used therein.

As a final note, we should mention the resurgence of interest in stochastic models that has appeared in the choice theoretical literature (see, e.g., Gul, Natenzon and Pesendorfer 2014, and Manzini and Mariotti 2014). Our results should be informative for the further development of this field, and contribute to the effective handling of potential internal incongruencies in environments involving risk or time.

References

[1] Aguirregabiria, V. and P. Mira (2010), "Dynamic Discrete Choice Structural Models: A Survey," *Journal of Econometrics*, 156:38–67.

[2] Ahn, N. (1995), "Measuring the Value of Children by Sex and Age Using a Dynamic Programming Model," *Review of Economic Studies*, 62.3:361–379.

[3] Ainslie, G. (1991), "Derivation of "Rational" Economic Behavior from Hyperbolic Discount Curves," *American Economic Review*, 81.2:334–340.

[4] Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström (2008), "Eliciting Risk and Time Preferences," *Econometrica*, 76.3:583–618.

[5] Apesteguia, J. and M. A. Ballester (2014), "Discrete Choice Estimation of Risk Aversion," mimeo.

[6] Benoît, J-P. and E. A. Ok (2007), "Delay Aversion," *Theoretical Economics*, 2:71–113.

[7] Berkovec, J. and S. Stern (1991), "Job Exit Behavior of Older Men," *Econometrica*, 59.1:189–210.

[8] Chabris, C. F., D. Laibson, C. L. Morris, J. P. Schuldt and D. Taubinsky (2008), "Individual Laboratory-Measured Discount Rates Predict Field Behavior," *Journal of Risk and Uncertainty*, 37.2/3:237–269.

[9] Coller, M. and M. B. Williams (1999), "Eliciting Individual Discount Rates," *Experimental Economics*, 2.2:107–127.

[10] Crawford, G. S. and M. Shum (2005), "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 73.4:1137–1173.

[11] Diermeier, D., M. Keane and A. Merlo (2005), "A Political Economy Model of Congressional Careers," *American Economic Review*, 95.1:347–373.

[12] Gilleskie, D. B. (1998), "A Dynamic Stochastic Model of Medical Care Use and Work Absence," *Econometrica*, 66.1:1–45.

[13] Gul, F., P. Natenzon, and W. Pesendorfer (2014), "Random Choice as Behavioral Optimization," *Econometrica*, forthcoming.

[14] Horowitz, J. K. (1992), "Comparative Impatience," *Economics Letters*, 38:25–29.

[15] Ida, T. and R. Goto (2009), "Simultaneous Measurement Of Time And Risk Preferences: Stated Preference Discrete Choice Modeling Analysis Depending On Smoking Behavior," *International Economic Review*, 50.4:1169–1182.

[16] Laibson, D. (1997), "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112.2:443–477.

[17] Loewenstein, G. and D. Prelec (1992), "Anomalies in Intertemporal Choice: Evidence and an Interpretation," *Quarterly Journal of Economics*, 107.2, 573–597.

[18] Manzini P. and M. Mariotti (2014), "Stochastic Choice and Consideration Sets," *Econometrica*, forthcoming.

[19] McFadden, D. (2001), "Economic Choices," *American Economic Review*, 91.3:351–378.

[20] Meier, S. and C. D. Sprenger (2014), "Temporal Stability of Time Preferences," *Review of Economics and Statistics*, forthcoming.

[21] O'Donoghue, T. and M. Rabin (1999), "Doing It Now or Later," *American Economic Review*, 89.1:103–124.

[22] Rust, J. (1987), "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55.5:999–1033.

[23] Rust, J. and C. Phelan (1997), "How Social Security and Medicare Affect Retirement Behavior In a World of Incomplete Markets," *Econometrica*, 65.4:781–831.

[24] Tanaka, T., C. F. Camerer, and Q. Nguyen (2010), "Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam," *American Economic Review*, 100.1:557–571.

[25] Toubia, O., E. Johnson, T. Evgeniou, and P. Delquié (2013), "Dynamic Experiments for Estimating Preferences: An Adaptive Method of Eliciting Time and Risk Parameters," *Management Science*, 59.3:613–640.

[26] Train, K. E. (2009), "Discrete Choice Methods with Simulation," *Cambridge University Press*

[27] Warner, J. T. and S. Pleeter (2001), "The Personal Discount Rate: Evidence from Military Downsizing Programs," *American Economic Review*, 91.1:33–53.

[28] Wolpin, K. I. (1984), "An Estimable Dynamic Stochastic Model of Fertility and Child Mortality," *Journal of Political Economy*, 92.5:852–874.