

WELL-DEFINED STOCHASTIC CHOICE MODELS: THE CASE OF RISK AND TIME PREFERENCES*

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

ABSTRACT. We analyze the validity of stochastic choice models contingent on a preference parameter. We show the possibility of a fundamental problem arising in the standard application of random utility models. Specifically, we find that, given two gambles, the probability of selecting the riskier gamble may be greater at higher levels of risk aversion. Similarly, given two monetary streams, the probability of selecting the stream providing the more delayed payoffs may be greater at higher levels of delay aversion. We show that the alternative random preference models, in contrast, are well-defined.

Keywords: Stochastic Choice; Preference Parameters; Risk Aversion; Delay Aversion; Random Utility Models; Random Preference Models.

JEL classification numbers: C25; D81.

1. INTRODUCTION

Early developments in stochastic choice theory by mathematical psychologists have had a profound impact both on the micro-econometric analysis of choice behavior and on choice theory.¹ The crucial exercise in a stochastic choice model is to establish an accurate link between the preferences and choice probabilities of the individual. The risk attitude of an individual, for example, should characterize her choice behavior when

Date: January, 2015.

* We thank Pavlo Blavatsky, Jerome Busemeyer, Syngjoo Choi, David Dillenberger, Kfir Eliaz, Holger Gerhardt, Faruk Gul, Glenn Harrison, John Hey, Paola Manzini, Marco Mariotti, Isabel Melguizo, Efe A. Ok, Ariel Rubinstein, Larry Samuelson, Daniel Sánchez, Karl Schlag, Matthew Shum, Andrej Skrzypacz, Tomasz Strzalecki, Hans-Martin Von Gaudecker and Nathaniel Wilcox 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, ECO2008-04756–Grupo Consolidado C), FEDER and MOVE is gratefully acknowledged.

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

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

¹See, e.g., Thurstone (1927) and Luce (1959) for some of the key first contributions.

choosing between different gambles. Similarly, the individual's time preferences should shape her choice probabilities over different time streams. Consider, in particular, a risky versus a safe gamble. One of the minimal requirements for a stochastic model to be well-defined is that the probability of choosing the riskier gamble should be decreasing with the level of risk aversion. Likewise, consider two payoff streams respectively offering more, versus less, delayed monetary payoffs. In this case, we would expect the probability of choosing the more delayed payoff stream to be decreasing with the individual's level of delay aversion. Contrary to this intuition, this paper shows that the most influential methods for the stochastic modelling of risk and time preferences, based on random utility models, are not well-defined.

In random utility models, the utility of the alternatives is subject to i.i.d. errors. This implies that the probabilities of choice between two alternatives depend on the difference between the utilities of the two alternatives. Section 4.1 studies the case of risk preferences. In a risk context, the utility of a gamble is typically assumed to have the expected utility form, with CARA and CRRA specifications often used to model the utility over monetary outcomes. Theorem 1 shows that, for every pair of gambles where one is unambiguously riskier than the other, random utility models using the above utility representations are not well-defined. It shows, furthermore, that there is always a level of risk aversion beyond which the probability of choosing the riskier gamble increases. We further show in a number of subsequent results that the problem extends well-beyond CARA and CRRA, and, indeed, beyond expected utility. Section 4.2 establishes the corresponding results for the case of time preferences. In time contexts the standard approach is to assume discounted utility, and real applications assume particular parametric structures such as the power and, more recently, the $\beta-\delta$ and hyperbolic discount functions. Theorem 2 proves that, for all three parametric discount functions, and for a large class of pairs of payoff streams involving different degrees of delay, the probability of choosing the one with the greater delay starts to increase beyond a certain level of delay aversion.

These are clearly fundamental problems. From a theoretical perspective, they imply a deep internal inconsistency in stochastic models. From a practical perspective, moreover, when these models are used to estimate levels of risk and delay aversion, as is typical practice in the abundant literature using discrete choice techniques, they may lead to flawed conclusions. The reasons for this are twofold. The first is an identification problem arising from the fact that the same choice probabilities may be associated

to two levels of risk or delay aversion. The second is that there is an upper limit to the level of aversion that can be estimated when using maximum likelihood techniques, even for extremely risk- and delay-averse individuals.

Results such as these constitute an alert to exercise caution when constructing stochastic choice models combining random utility models with certain utility representations. In relation to this, we present a result, Lemma 2, which establishes two simple conditions under which a random utility model is not well-defined and which could also prove useful in applications beyond risk and time scenarios. See section 4.3 for some potential settings where our results might naturally apply.

We then turn to the study of a different class of stochastic choice models: random preference models. In these, the choice probabilities are derived from the assumption of a probability distribution over the preference parameter, and hence the key determinant is the mass of preferences for which one option is better than the other. In Lemma 3 we establish that a minimal regularity condition over the preference parameter guarantees that random preference models are well-defined. Then, in Theorem 3 we show that the expected utility and discounted utility specifications are well-defined.

The rest of the paper is organized as follows. Section 2 reviews the relevant literature. In section 3 we lay down the basic definitions. Section 4 is devoted to the study of random utility models. It first establishes general conditions for these models to be well-defined. Then, section 4.1 covers the case of risk aversion, and section 4.2 that of delay aversion. Section 4.3 discusses other potential applications of interest. The case of random preference models is studied in section 5, and section 6 concludes.

2. RELATED LITERATURE

Discrete choice models in general settings not necessarily involving risk or time are surveyed in McFadden (2001). See also Train (2009) for a detailed textbook introduction.

For theoretical papers recommending the use of random utility models in risk settings, see Becker, DeGroot and Marschak (1963) and Busemeyer and Townsend (1993). The literature using random utility models in the estimation of risk aversion is immense, and certainly too large to be exhaustively cited here. Therefore, we cite only some of the most influential pieces of work. Namely, Friedman (1974), Cicchetti and Dubin (1994), Hey and Orme (1994), Holt and Laury (2002), Harrison, List and Towe (2007), Andersen et al. (2008), Post et al. (2008), von Gaudecker, van Soest and Wengstrom

(2011), Toubia et al. (2013), and Noussair, Trautmann and van de Kuilen (2014). The random utility model is also the most commonly used approach in the estimation of time preferences. See, e.g., Andersen et al. (2008), Chabris et al. (2008), Ida and Goto (2009), Tanaka, Camerer and Nguyen (2010), Toubia et al. (2013), and Meier and Sprenger (2014).

The use of random preference models in settings involving gambles has been theoretically discussed in Eliashberg and Hauser (1985), Loomes and Sugden (1995), and recently in Gul and Pesendorfer (2006). For papers estimating risk aversion in line with this method, see Barsky et al. (1997), Fullenkamp, Tenorio and Battalio (2003), and Kimball, Sahm and Shapiro (2008, 2009). Collier and Williams (1999) and Warner and Pleeter (2001) are two examples of the use of this approach in the context of time preferences.

Using examples with the logit model and for some comparisons of risky versus safe gambles, Wilcox (2008, 2011) discusses how CARA and CRRA may present the sort of problems characterized here. In addition, he proposes the use of a novel model, contextual utility, which solves the problem for some specific comparisons of gambles involving the classical three-outcome case. At the end of section 4.1, we discuss his proposal further and show that the problem persists beyond these comparisons. In addition, Blavatsky (2011) shows that, in the case of random utility models based on expected utility differences, there are no two individuals where one always chooses the non-degenerate gamble over the degenerate gamble with a lower probability than the other individual.

3. PRELIMINARIES

Let X be a set of alternatives and consider an interval Ω of the real numbers. The parameter $\alpha \in \Omega$ aims to capture a preference feature of the agent, associated with an inclination towards certain types of alternatives in X . For instance, higher values of α may represent greater risk aversion or delay aversion causing the individual to be less inclined towards risky gambles or monetary streams involving distant payoffs.

Some pairs of alternatives (x, y) are *unambiguously* ordered with respect to the preference feature under evaluation. For instance, in a risk context, a gamble x involving risk is intuitively *riskier* than a certain outcome y . Similarly, a mean-preserving spread x is riskier than the gamble y from which it is constructed. In a context of time choices, a payoff stream x with a later bonus payout *is more delayed* than the stream y with an

earlier bonus payout. To emphasize, a pair of alternatives is judged unambiguous based on the characteristics of the alternatives and the nature of the preference parameter under consideration.

A parametric stochastic choice model is a map $\rho : \Omega \times X \times X$ that describes the probability $\rho_\alpha(x, y)$ of choosing alternative x when confronted with alternative y , when the value of the preference parameter is α . We say that the parametric stochastic model ρ is *well-defined for the unambiguous pair* (x, y) , whenever $\rho_\alpha(x, y)$ is decreasing in α . That is, the larger the value of the parameter α , the lower the probability of choosing alternative x from the unambiguous pair (x, y) , which captures the aforementioned inclination. This is a minimal condition for the internal consistency of the stochastic model with respect to the preference parameter α , and for the proper empirical estimation of the preference parameter.

In the following sections we study in detail whether two influential classes of stochastic models, random utility models and random preference models, are well-defined.

4. RANDOM UTILITY MODELS

In the most standard random utility model (RUM), the individual chooses the alternative that provides maximal utility, which is assumed to be additively composed by two terms: (i) the representative utility, $U_\alpha(x)$, based on the characteristics of the alternative x and the relevant preference attribute α and, (ii) a random i.i.d. term that follows a continuous cumulative distribution Ψ . Given the pair of alternatives (x, y) , the probability assigned to choosing x , $\rho_\alpha^{rum}(x, y)$, is given by the probability that $\lambda U_\alpha(x) + \epsilon$ is greater than $\lambda U_\alpha(y) + \hat{\epsilon}$, where ϵ and $\hat{\epsilon}$ are two independent realizations from Ψ , and λ is a parameter inversely related to the variance of Ψ .² By far the most widely-used error distributions are the extreme type I and the normal distribution, which lead to the logit model and the probit model, respectively. The former, also known as the Luce model, has closed-form probabilities of choosing x over y that are equal to $\frac{e^{\lambda U_\alpha(x)}}{e^{\lambda U_\alpha(x)} + e^{\lambda U_\alpha(y)}}$.

The following lemma provides a necessary and sufficient condition for a RUM to be well-defined for the unambiguous pair (x, y) .

²The parameter λ is sometimes interpreted as the rationality of the individual. Whenever λ goes to zero, choices become completely random, while, when λ goes to infinity, choices become fully deterministic.

Lemma 1. ρ^{rum} is well-defined for the unambiguous pair (x, y) if and only if the function $U_\alpha(x) - U_\alpha(y)$ is decreasing in α .

Proof of Lemma 1: Consider a RUM and an unambiguous pair (x, y) . Notice that the probability of $\lambda U_\alpha(x) + \epsilon$ being greater than $\lambda U_\alpha(y) + \hat{\epsilon}$ can be expressed as the probability that $\hat{\epsilon} - \epsilon$ is smaller than $\lambda U_\alpha(x) - \lambda U_\alpha(y)$. Hence, denoting by Ψ^* the cumulative distribution of the difference between two i.i.d. error terms with distribution Ψ , it is the case that $\rho_\alpha^{rum}(x, y) = \Psi^*(\lambda(U_\alpha(x) - U_\alpha(y)))$. Since Ψ^* is a continuous cumulative distribution, $\rho_\alpha^{rum}(x, y)$ is decreasing in α if and only if $U_\alpha(x) - U_\alpha(y)$ is decreasing in α , as desired. ■

The following observation is an immediate corollary of Lemma 1. Consider an unambiguous pair (x, y) . Suppose that, as aversion becomes maximal in the domain, an idea which we denote by $\alpha \uparrow$, the difference between the representative utilities of x and y vanishes, i.e., $\lim_{\alpha \uparrow} [U_\alpha(x) - U_\alpha(y)] = 0$. Then, either x is always better than y or the RUM is not well-defined for this pair.

Lemma 2. If $\lim_{\alpha \uparrow} [U_\alpha(x) - U_\alpha(y)] = 0$ and there exists $\alpha^* \in \Omega$ such that $U_{\alpha^*}(y) > U_{\alpha^*}(x)$, then ρ^{rum} is not well-defined for the unambiguous pair (x, y) .

Proof of Lemma 2: Consider a RUM and an unambiguous pair (x, y) . Suppose that there exists $\alpha^* \in \Omega$ such that $U_{\alpha^*}(y) > U_{\alpha^*}(x)$. If the RUM is well-defined for (x, y) , we know from Lemma 1 that $U_\alpha(x) - U_\alpha(y)$ must be decreasing in α . Hence, clearly, either $\lim_{\alpha \uparrow} [U_\alpha(x) - U_\alpha(y)]$ does not exist, or it must be the case that $\lim_{\alpha \uparrow} [U_\alpha(x) - U_\alpha(y)] < 0$. In both cases we have a contradiction, proving the result. ■

In the following sections, we show the relevance of these observations when dealing with risk and time preferences. In particular, we show that an immediate application of the above results shows that the RUMs used in these contexts are essentially not well-defined for almost every unambiguous pair of alternatives. By exploiting the structure of the parametric models most commonly used in the literature, moreover, we are able to provide stronger results that characterize the relationship between the choice probabilities and the preference parameter.³

³In some instances, a logarithmic transformation of the representative utilities is used. The results established here, together with the implications on risk and time domains to be shown next, are basically replicable when using the logarithmic transformation. See Appendix A for details.

4.1. Application: Risk Preferences. A gamble $x = [(p_1, \dots, p_N); (x_1, \dots, x_N)]$ is described by a finite vector of probabilities (p_1, \dots, p_N) , with $p_i > 0$ and $\sum_i p_i = 1$, over a finite vector of monetary outcomes (x_1, \dots, x_N) , with $x_i \in \mathbb{R}_+$. In the analysis that follows, we consider three different types of unambiguous pairs of gambles. The standard textbook treatment of risk aversion uses pairs of gambles (x, y) where the latter involves no risk at all, i.e. $y = [1; y_1]$. The monetary value y_1 is sometimes taken to be the expected value of x . For a broader definition that will enable us to consider a larger set of comparisons, we simply assume here that $\min\{x_1, \dots, x_N\} < y_1 < \max\{x_1, \dots, x_N\}$. We refer to these unambiguous pairs of gambles as being of the degenerate type. Another widely-accepted unambiguous comparison involves pairs (x, y) , where x is a mean-preserving spread of y . Gamble x is a mean-preserving spread of gamble y through outcome y_{j^*} and gamble z , if x can be expressed as a compound gamble that replaces outcome y_{j^*} in gamble y with gamble z , which has y_{j^*} as its expected value. We say that x is a mean-preserving spread of y if there is a sequence of such spreads from y to x , and we refer to such pairs as being of the mean-preserving spread type. Finally, experimental works often use simple pairs of gambles where $x = [(p, 1 - p); (x_1, x_2)]$ and $y = [(p, 1 - p); (y_1, y_2)]$, with $x_1 < y_1 < y_2 < x_2$ and $p \in (0, 1)$. As in the first case, although the mean of x is sometimes assumed to be larger than that of y , here, for the purpose of generality, we impose no further assumptions. We refer to these comparisons as nested gambles.

In most standard analysis, utility functions take the form of expected utility $U_\alpha^{eu}(x) = \sum_i p_i u_\alpha(x_i)$, where u_α is a monetary utility function that is strictly increasing and continuous in outcomes. The constant absolute risk aversion (CARA) and the constant relative risk aversion (CRRA) families of monetary utility functions are by far the most widely-used specifications in real applications. The following are standard definitions. CARA utility functions are such that the utility of monetary outcome m is $u_\alpha^{cara}(m) = \frac{1 - e^{-\alpha m}}{\alpha}$ for $\alpha \neq 0$, and $u_0^{cara}(m) = m$, while CRRA utility functions are defined by $u_\alpha^{crra}(m) = \frac{m^{1-\alpha}}{1-\alpha}$ for $\alpha \neq 1$, and $u_1^{crra}(m) = \log m$, and, in both cases, $\Omega = \mathbb{R}$.⁴ We write U_α^{cara} and U_α^{crra} for the corresponding expected utilities, and $\rho^{rum(cara)}$ and $\rho^{rum(crra)}$ for the corresponding RUM choice probabilities.

We now illustrate how this approach has fundamental flaws. Notice, firstly, that Lemma 2 has immediate bite. This is so because, for every gamble x , we have that $\lim_{\alpha \uparrow} U_\alpha^{cara}(x) = \lim_{\alpha \uparrow} U_\alpha^{crra}(x) = 0$, and also because, for every unambiguous pair of

⁴For ease of exposition, assume that $m \geq 1$.

gambles, there are clearly levels of risk aversion for which the safer gamble is preferred. Hence, our Lemma 2 applies, which means that for every unambiguous pair of gambles both $\rho^{rum(cara)}$ and $\rho^{rum(crra)}$ are not well-defined. Furthermore, by exploiting the functional structure of U_α^{cara} and U_α^{crra} , we can establish stronger results. In the next result, we show that, for every RUM and for every unambiguous pair of gambles (x, y) , there always exists a level of risk aversion $\alpha_{(x,y)}$ above which the probability of choosing the riskier gamble x is strictly increasing. Obviously, this is a serious drawback for the internal consistency of the model, since it means that higher levels of risk aversion are associated with greater probabilities of choosing the riskier gamble. Moreover, the practical implications for estimation purposes are immediate. Firstly, there is an identification problem arising from the fact that different levels of risk aversion are compatible with the same probability of choice. Secondly, if a maximum likelihood exercise is used, an intensely risk-averse individual who consistently chooses the safe option y will be associated with the intermediate level of risk aversion $\alpha_{(x,y)}$.

Theorem 1. *For every unambiguous pair of gambles (x, y) , there exists $\alpha_{(x,y)} \in \mathbb{R}$ such that $\rho_\alpha^{rum(cara)}(x, y)$ and $\rho_\alpha^{rum(crra)}(x, y)$ are strictly increasing in α whenever $\alpha \geq \alpha_{(x,y)}$.*

Proof of Theorem 1: Consider an unambiguous pair of gambles (x, y) , with $x = [(p_1, \dots, p_N); (x_1, \dots, x_N)]$ and $y = [(q_1, \dots, q_M); (y_1, \dots, y_M)]$. With reasoning analogous to that used in Lemma 1, we need to show that there exists a risk-aversion level $\alpha_{(x,y)}$ such that the difference between the utility values of x and y is strictly increasing in α above $\alpha_{(x,y)}$. We first consider the case of CARA and focus on $\alpha \neq 0$, where the family is differentiable.⁵ In this domain, $\rho_\alpha^{rum(cara)}(x, y)$ is strictly increasing in α iff $\frac{\partial[U_\alpha^{cara}(x) - U_\alpha^{cara}(y)]}{\partial\alpha} > 0$ which, by expected utility, is equivalent to $\sum p_i \frac{\partial u_\alpha^{cara}(x_i)}{\partial\alpha} - \sum q_j \frac{\partial u_\alpha^{cara}(y_j)}{\partial\alpha} > 0$. Since $-\frac{\partial u_\alpha^{cara}(m)}{\partial\alpha} = -\frac{e^{-\alpha m}(1+\alpha m)-1}{\alpha^2}$ is a strictly increasing and continuous utility function over monetary outcomes, $\rho_\alpha^{rum(cara)}(x, y)$ is strictly increasing in α iff $V_\alpha^{cara}(y) > V_\alpha^{cara}(x)$, where V_α^{cara} is the expected utility using $-\frac{\partial u_\alpha^{cara}(m)}{\partial\alpha}$. Denoting by $CE(x, V_\alpha^{cara})$ and $CE(y, V_\alpha^{cara})$ the certainty equivalents of V_α^{cara} for gambles x and y , it follows immediately that $\rho_\alpha^{rum(cara)}(x, y)$ is strictly increasing in α iff $CE(y, V_\alpha^{cara}) > CE(x, V_\alpha^{cara})$.⁶

Now consider unambiguous pairs of gambles (x, y) of the degenerate type, i.e., where $y = [1, y_1]$ and $\min\{x_1, \dots, x_N\} < y_1$. Notice that the Arrow-Pratt coefficient of

⁵Note that the discontinuity of the CARA family at this point is not relevant for the result.

⁶In general, the certainty equivalent of a gamble x for some utility function U , is the amount of money $CE(x, U)$ such that $U(x) = U([1; CE(x, U)])$.

risk aversion for $-\frac{\partial u_\alpha^{cara}(m)}{\partial \alpha}$ is simply $\alpha - \frac{1}{m}$.⁷ When α grows, the Arrow-Pratt coefficient goes to infinity, thereby guaranteeing that $\lim_{\alpha \uparrow} CE(x, V_\alpha^{cara}) = \min\{x_1, \dots, x_N\}$. Hence, we can find a value, which we denote by $\alpha_{(x,y)}$, such that, for every $\alpha \geq \alpha_{(x,y)}$, $CE(x, V_\alpha^{cara}) < y_1 = CE(y, V_\alpha^{cara})$, which proves the result. If (x, y) are nested gambles, since the minimum outcomes of x and y satisfy $x_1 < y_1$, the logic used for the degenerate type of unambiguous comparisons also applies here. Now, consider an unambiguous pair (x, y) , where x is a mean-preserving spread of y through outcome y_{j^*} and gamble $z = [(r_1, \dots, r_O); (z_1, \dots, z_O)]$. Using the additive nature of expected utility and the assumptions on the gambles, it is the case that $V_\alpha^{cara}(y) > V_\alpha^{cara}(x)$ if and only if $V_\alpha^{cara}(w) > V_\alpha^{cara}(z)$, where $w = [1; y_{j^*}]$. Hence, we can find a value $\alpha_{(z,w)}$ such that $\rho_\alpha^{rum(cara)}(z, w)$ is strictly increasing in α for every $\alpha \geq \alpha_{(z,w)}$, and hence we can take $\alpha_{(x,y)} = \alpha_{(z,w)}$. In general, if x is a mean-preserving spread of y , we can iteratively repeat this process and consider $\alpha_{(x,y)}$ as the maximum of all the corresponding $\alpha_{(z,w)}$ values that appear in the decomposition, and the result follows.

The proof of the CRRA case can be obtained analogously by considering that, for any $\alpha \neq 1$, CRRA utility functions are differentiable, $-\frac{\partial u_\alpha^{crra}(m)}{\partial \alpha} = -\frac{m^{1-\alpha}(1-(1-\alpha)\log m)}{(1-\alpha)^2}$ is a continuous and strictly monotone utility function over monetary outcomes and the corresponding Arrow-Pratt coefficient is $\frac{\alpha \log m - 1}{m \log m}$. ■

The proof of Theorem 1, using the differentiability of CARA and CRRA, shows that the model is well-defined if and only if gamble x has more expected utility than gamble y when using the monetary utility function $-\frac{\partial u_\alpha}{\partial \alpha}$. The proof establishes that there is always a level of risk aversion $\alpha_{(x,y)}$ beyond which this no longer holds, and hence the models are not well-defined. Furthermore, the proof also helps to explain how the critical values $\alpha_{(x,y)}$ vary with the pair of gambles involved and, in particular, it can be used to visualize how very small these critical values can be. For instance, consider two unambiguous pairs of gambles (x, y) and (x, y') , where y stochastically dominates y' . Take these pairs to be, let us say, of the degenerate type and hence $y > y'$. It is evident from the proof of Theorem 1 that $\alpha_{(x,y)} < \alpha_{(x,y')}$. Hence, the better the safer option, the sooner the problems of the stochastic model arise. Moreover, it is easy to see that the problems become more severe as payoffs increase. This is so because, as

⁷The coefficient has a strictly positive derivative with respect to α and thus, from the classic result of Pratt (1964), it follows that the certainty equivalent of a non-degenerate gamble is strictly decreasing in α .

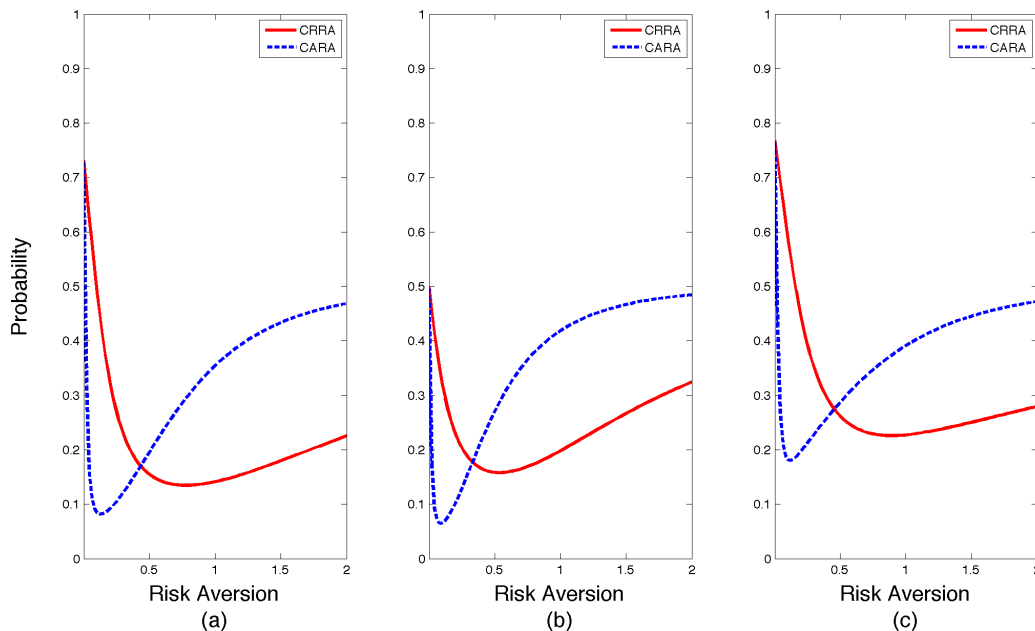


FIGURE 1.—RUM probabilities of choosing gamble x over gamble (a) y , (b) y' , and (c) \bar{y} .

the monetary payoffs increase, the Arrow-Pratt coefficients of $-\frac{\partial u_\alpha}{\partial \alpha}$ tend to α and zero for CARA and CRRA, respectively. Hence, in the case of CARA, the critical values $\alpha_{(x,y)}$ tend towards the point where gamble y starts to be better than x according to U_α^{cara} , and in the case of CRRA it tends to zero. This implies, for instance, that in the case of the mean-preserving spread comparisons, the probability of choosing the mean-preserving spread gamble x tends to start increasing at $\alpha = 0$ for both models. This is very disturbing, since, in an estimation exercise involving large payoffs, it implies that every risk-averse individual would be categorized as being, at the most, risk-neutral.

It is convenient to stress, however, that large payoffs are not necessary in order to obtain small critical values $\alpha_{(x,y)}$. In Figure 1, we illustrate the problems established in Theorem 1 using gambles with small payoffs. The figure has three subfigures, one for each of the three types of unambiguous pairs of gambles. In all three cases, we illustrate using both CARA and CRRA expected utility, with the logit model and $\lambda = 2$. Figure 1a depicts the probability of choosing gamble $x = [(.95, .05); (1, 51)]$ over the degenerate gamble $y = [1; 3]$, Figure 1b the probability of choosing the mean-preserving spread gamble x over $y' = [(.5, .5); (1, 6)]$, and Figure 1c the probability of choosing gamble x over $\bar{y} = [(.95, .05); (2, 20)]$. It can be appreciated how all these probabilities first decrease with the level of risk aversion, then reach a critical level, after which they all

start increasing. Note that, in all six cases, the critical values of α are relatively small. The two main problems discussed earlier are now apparent. Firstly, the maximum risk-aversion parameters that can be estimated using these gambles, even for infinitely risk-averse individuals, are relatively small. Secondly, note that certain ranges of choice probabilities are simultaneously consistent with two risk-averse coefficients. Theorem 1 shows that, for any given error distribution, problems of this sort arise for every unambiguous pair of gambles.

We now comment on several extensions to Theorem 1, which we group into two classes depending on the additive nature of the representative utility function.

Additive Extensions. Theorem 1 works under the assumption of expected utility. Clearly, generalizations of expected utility, such as cumulative prospect theory, rank-dependent expected utility, disappointment aversion, etc, are susceptible to the problems identified above, since they include expected utility as a special case. More importantly, however, the additive nature of these models makes them vulnerable to similar anomalies, even when considering only non-expected utilities. In order to formally establish these reflections, consider a function π that associates to every gamble $x = [(p_1, \dots, p_N); (x_1, \dots, x_N)]$ another gamble $\pi(x) = [(q_1, \dots, q_N); (x_1, \dots, x_N)]$ over the same set of outcomes. We assume that, for any given vector of outcomes, the distortion of probabilities is a one-to-one, continuous monotone function over each argument. Then, the generalized expected utility of gamble x , $U_\alpha^{geu}(x)$, is $U_\alpha^{eu}(\pi(x))$. Accordingly, the generalized CARA expected utility is $U_\alpha^{gcara}(x) = U_\alpha^{cara}(\pi(x))$, while the generalized CRRA expected utility is $U_\alpha^{gcrra}(x) = U_\alpha^{crra}(\pi(x))$, and the corresponding RUM choice probabilities are denoted by $\rho^{\text{rum}(gcara)}$ and $\rho^{\text{rum}(gcrra)}$. It is easy to see that the logic behind Theorem 1 extends immediately to these generalized expected utilities for unambiguous comparisons (x, y) of the degenerate or nested types.⁸

Proposition 1. *For every unambiguous pair of gambles (x, y) of the degenerate or nested type, there exists $\alpha_{(x,y)} \in \mathbb{R}$ such that $\rho_\alpha^{\text{rum}(gcara)}(x, y)$ and $\rho_\alpha^{\text{rum}(gcrra)}(x, y)$ are strictly increasing in α whenever $\alpha \geq \alpha_{(x,y)}$.*

⁸Every unambiguous pair of gambles of the mean-preserving spread type is also problematic, but the nature of the problem depends on the gambles. When the gambles do not share the minimum outcome, the same logic used in the proof of Theorem 1 can be applied to establish the same type of result. Furthermore, when the gambles share the same minimum monetary outcome, Lemma 2 immediately applies, which shows that the model is not well-defined for such gambles.

Proof of Proposition 1: Consider an unambiguous pair of gambles (x, y) of the degenerate or nested type. It is immediate to see that $\rho_\alpha^{rum(gcara)}(x, y)$ is strictly increasing in α iff $CE(\pi(y), V_\alpha^{cara}) > CE(\pi(x), V_\alpha^{cara})$. The certainty equivalents for gambles $\pi(x)$ and $\pi(y)$ converge, with increasing α , towards the corresponding minimum outcomes in $\pi(x)$ and $\pi(y)$. Since the distorted gambles $\pi(x)$ and $\pi(y)$ are defined over the same outcomes as x and y , respectively, the logic followed in the proof of Theorem 1 also applies here, given the nature of the gambles. The case of CRRA utilities is completely analogous and hence omitted. ■

Another important observation is that the use of CARA and CRRA utilities in our main results is essential only in order to characterize the structure of the choice probabilities involved in the problem, and thereby find a global minimum $\alpha_{(x,y)}$, after which the probability of choosing the riskier option increases. We now establish that, even without these functional forms, we can still show that every RUM based on generalized expected utilities using any monetary utility function that is strictly increasing and continuous in outcomes is not well-defined. That is, there is always an unambiguous pair of gambles (x, y) for which the model is not well-defined. We formally establish this result in the following proposition. To this end, let us denote by $\rho^{rum(geu)}$ the RUM probabilities of a generalized expected utility.

Proposition 2. *For every $\rho^{rum(geu)}$ there exists an unambiguous pair of gambles (x, y) , such that $\rho^{rum(geu)}$ is not well-defined for (x, y) .*

Proof of Proposition 2: Suppose, by contradiction, that there exists a generalized expected utility model that is well-defined for every unambiguous pair of gambles. Let π be the function that distorts probabilities, and u the utility function over the monetary outcomes defining the generalized expected utility model. For any unambiguous pair of gambles (x, y) , $\rho_\alpha^{rum(geu)}(x, y) = \rho_\alpha^{rum(eu)}(\pi(x), \pi(y))$, where $\rho_\alpha^{rum(eu)}$ denotes the RUM probabilities of expected utility with monetary utility u .

Let $x = [(p, 1 - p); (x_1, x_2)]$ with $x_1 < x_2$ and $p \in (0, 1)$. Also, let $y = [1; y_1]$ with $x_1 < y_1 < x_2$. From the basic assumptions over the distorted probabilities, it is clearly the case that $\pi(x) = [(p', 1 - p'); (x_1, x_2)]$ with $p' \in (0, 1)$, and $\pi(y) = y$. Since the pair of gambles $(\pi(x), \pi(y))$ is unambiguous, $\rho_\alpha^{rum(eu)}(\pi(x), \pi(y))$ must be decreasing in α . Taking two levels of risk aversion $\alpha_1 < \alpha_2$, it must be the case that $\rho_{\alpha_1}^{rum(eu)}(\pi(x), \pi(y)) \geq \rho_{\alpha_2}^{rum(eu)}(\pi(x), \pi(y))$. By Lemma 1, this is equivalent to $p'u_{\alpha_1}(x_1) + (1 - p')u_{\alpha_1}(x_2) - u_{\alpha_1}(y_1) \geq p'u_{\alpha_2}(x_1) + (1 - p')u_{\alpha_2}(x_2) - u_{\alpha_2}(y_1)$, or, simply,

$p'[u_{\alpha_1}(x_1) - u_{\alpha_2}(x_1)] + (1 - p')[u_{\alpha_1}(x_2) - u_{\alpha_2}(x_2)] \geq u_{\alpha_1}(y_1) - u_{\alpha_2}(y_1)$. By taking the limits of $u_{\alpha_1}(y_1) - u_{\alpha_2}(y_1)$ when y_1 goes to x_1 and x_2 , respectively, we have that $p'[u_{\alpha_1}(x_1) - u_{\alpha_2}(x_1)] + (1 - p')[u_{\alpha_1}(x_2) - u_{\alpha_2}(x_2)] \geq \max\{u_{\alpha_1}(x_1) - u_{\alpha_2}(x_1), u_{\alpha_1}(x_2) - u_{\alpha_2}(x_2)\}$, which ultimately implies that $u_{\alpha_1}(x_1) - u_{\alpha_2}(x_1) = u_{\alpha_1}(x_2) - u_{\alpha_2}(x_2)$.

The above reasoning can be applied to any monetary outcomes x_1 and x_2 , and hence the function $u_{\alpha_1}(m) - u_{\alpha_2}(m)$ must be constant. Consequently, a reasoning analogous to that used to prove Lemma 1 guarantees that $\rho_{\alpha_1}^{rum(eu)}$ coincides with $\rho_{\alpha_2}^{rum(eu)}$ for every pair of gambles, which is a contradiction. ■

Non-Additive Extensions. The certainty equivalent is sometimes used instead of the expected utility as the representative utility. The main intuition behind this approach is that the certainty equivalent is a monetary representation of preferences, where the use of a common scale facilitates interpersonal comparisons. One may entertain the idea that, by creating a common scale, this method could provide a solution to the problem discussed. This is indeed the case in some instances, such as whenever the unambiguous pair (x, y) involves a degenerate gamble. This can be appreciated by noticing that the certainty equivalent of the non-degenerate gamble x decreases with the level of risk aversion, while the certainty equivalent of the degenerate gamble y is constant across risk-aversion levels. Thus, the difference between the certainty equivalents of the two gambles decreases with the level of risk aversion and, by Lemma 1, the probability of choosing the risky gamble decreases, as desired. However, caution is required when using certainty equivalents, because problems may arise with other comparisons. We illustrate this point by considering mean-preserving spread pairs (x, y) , with $\min\{x_1, \dots, x_N\} = \min\{y_1, \dots, y_M\}$. We denote by $\rho^{rum(cecara)}$ and $\rho^{rum(cecrra)}$ the choice probabilities associated with this model, when using the certainty equivalent representation of CARA and CRRA expected utilities, respectively.

Proposition 3. $\rho^{rum(cecara)}$ and $\rho^{rum(cecrra)}$ are not well-defined for every unambiguous pair of gambles (x, y) of the mean-preserving spread type satisfying $\min\{x_1, \dots, x_N\} = \min\{y_1, \dots, y_M\}$.

Proof of Proposition 3: Consider an unambiguous pair of gambles (x, y) of the mean-preserving spread type satisfying $\min\{x_1, \dots, x_N\} = \min\{y_1, \dots, y_M\}$. Since the Arrow-Pratt coefficients of u_{α}^{cara} and u_{α}^{crra} are α and αm , respectively, it follows that $\lim_{\alpha \uparrow} [CE(x, U_{\alpha}^{cara}(x)) - CE(x, U_{\alpha}^{crra}(x))] = \lim_{\alpha \uparrow} [CE(x, U_{\alpha}^{crra}(x)) - CE(x, U_{\alpha}^{crra}(x))] = \min\{x_1, \dots, x_N\} - \min\{y_1, \dots, y_M\} = 0$. Also, for every $\alpha^* > 0$, it is the case that

$U_{\alpha^*}^{cara}(y) > U_{\alpha^*}^{cara}(x)$ and $U_{\alpha^*}^{crra}(y) > U_{\alpha^*}^{crra}(x)$. Hence, Lemma 2 immediately applies, which proves the claim. ■

Let us now consider mean-variance utilities, which are much used in portfolio theory and macroeconomics. Markowitz (1952) was the first to propose a mean-variance evaluation of risky asset allocations. Roberts and Urban (1988) and Barseghyan et al. (2013) provide examples of the use of mean-variance utilities in a RUM, for the estimation of risk preferences. Formally, given a gamble x , let us denote the expected value and variance of x by $\mu(x) = \sum_i p_i x_i$ and $\sigma^2(x) = \sum_i p_i (x_i - \mu(x))^2$, respectively. Mean-variance utilities are then described by $U_{\alpha}^{mv}(x) = \mu(x) - \alpha\sigma^2(x)$. We now argue that the corresponding RUM choice probabilities $\rho^{rum(mv)}$ are always well-defined.

Proposition 4. $\rho^{rum(mv)}$ is well-defined for every unambiguous pair of gambles (x, y) .

Proof of Proposition 4: Consider an unambiguous pair of gambles (x, y) . Notice that $U_{\alpha}(x) - U_{\alpha}(y) = \mu(x) - \mu(y) - \alpha(\sigma^2(x) - \sigma^2(y))$. For every unambiguous pair (x, y) , it is clearly the case that $\sigma^2(x) > \sigma^2(y)$ and hence, $U_{\alpha}(x) - U_{\alpha}(y)$ is decreasing in α . Thus, we can use Lemma 1 to prove the claim. ■

To conclude, Wilcox (2011) suggests normalizing the utility difference between the gambles by the difference between the utilities of the best and worst of all the outcomes involved in the two gambles under consideration. This variation of a RUM goes under the name of contextual utility. The author shows that the suggested normalization solves the problem for cases in which both gambles in the unambiguous pair (x, y) are defined over the same three outcomes, thus covering the important Marschak-Machina triangles, and are related through the notion of mean-preserving spread. However, this normalization does not solve the problem beyond the case mentioned.⁹

4.2. Application: Time Preferences. A monetary stream $x = (x_0, x_1, \dots, x_T)$ describes the amount of money $x_t \in \mathbb{R}_+$ realized at every time t .¹⁰ We consider unambiguous pairs of streams where there is a unique conflict between obtaining a larger

⁹Consider, for example, the gambles $x = [(1, .4, 0, .4, .1); (0, 10, 50, 90, 100)]$ and $y = [(.05, 0, .9, 0, .05); (0, 10, 50, 90, 100)]$, where x is a mean-preserving spread of y , or the unambiguous pair of the degenerate type $(x, [1; 50])$. It is immediate to see that the RUM probability of choosing x using expected utility with CRRA is lower for the risk-aversion coefficient $\alpha_1 = .7$ than for $\alpha_2 = .9$.

¹⁰Whether streams are finite or infinite is irrelevant to this analysis.

monetary payoff with a shorter delay, or some other monetary payoff with a longer delay. That is, we focus on pairs (x, y) , for which $x_t = y_t$ except for two periods $t_y < t_x$, with $y_{t_y} > x_{t_y}$ and $y_{t_x} < x_{t_x}$.¹¹

The standard approach uses discounted utility of streams $U_\alpha^{du}(x) = \sum_t D_\alpha^{du}(t)u(x_t)$, with discount functions for which $\Omega = \mathbb{R}_+$, $D_\alpha(0) = 1$ and $\lim_{t \rightarrow \infty} D_\alpha(t) = 0$. The utility function over monetary outcomes $u : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is assumed to be strictly increasing, with $u(0) = 0$ and continuous. In order to isolate the time-discounting effect, we assume throughout the section that the curvature of the utility function over monetary outcomes u is fixed. The most commonly-used discount function is the power function where $D_\alpha^{pow}(t) = \frac{1}{(1+\alpha)^t}$.¹² The behavioral literature offers two alternative discount functions that seem to better capture certain behavior patterns.¹³ These are the hyperbolic discounting formula $D_\alpha^{hyp}(t) = \frac{1}{1+\alpha t}$, and the $\beta - \delta$ discounting formula where $D_\alpha^{beta}(0) = 1$ and $D_\alpha^{beta}(t) = \beta D_\alpha^{pow}(t)$ whenever $t > 0$, with $\beta \in (0, 1]$.¹⁴ We write U_α^{pow} , U_α^{hyp} and U_α^{beta} for the corresponding discounted utilities, and $\rho^{rum(pow)}$, $\rho^{rum(hyp)}$ and $\rho^{rum(beta)}$ for the corresponding RUM probabilities.

We start by noting that, whenever the moment at which the earlier option y offers the extra monetary payoff is the present, that is $t_y = 0$, the choice probabilities $\rho^{rum(pow)}$, $\rho^{rum(hyp)}$ and $\rho^{rum(beta)}$ are all well-defined. This follows from the fact that the present is not discounted, and hence, all that matters in these cases is how period t_x is discounted. Since this discounting is decreasing in the delay-aversion coefficient α , Lemma 1 guarantees that choice probabilities are well-defined for such pairs (x, y) . Importantly, Chabris et al. (2008) and Tanaka, Camerer and Nguyen (2010) only compare streams in which $t_y = 0$.

However, it is immediate to see that, for any other unambiguous pair of streams (x, y) , the difference between the parametric discounted utilities of x and y converges to 0 as α increases, and also that there is always a level of delay aversion α for which stream

¹¹Benoît and Ok (2007) characterize the notion of more delay aversion. In their study, these streams are key. See also Horowitz (1992).

¹² It is clear that the exponential function $D_{\hat{\alpha}}^{exp}(t) = \exp^{-\hat{\alpha}t}$ is equivalent to the power function, if we consider that $\hat{\alpha} = \log(1 + \alpha)$. This alternative representation is therefore omitted.

¹³See Ainslie (91), Loewenstein and Prelec (1992), Laibson (1997), and O'Donoghue and Rabin (1999).

¹⁴The alternative representation based on the exponential function is sometimes called quasi-hyperbolic. That is, $D_\alpha^{qh}(0) = 1$ and $D_\alpha^{qh}(t) = \beta D_\alpha^{exp}(t)$ whenever $t > 0$, with $\beta \in (0, 1]$. We can omit this functional form for reasons given in the footnote 12.

y is preferred to x . Hence, Lemma 2 applies to all three parametric versions $\rho^{rum(pow)}$, $\rho^{rum(hyp)}$ and $\rho^{rum(beta)}$, showing that all these models are problematic. Analogously to Theorem 1, the next result exploits the functional structure of U^{pow} to reach similar results.

Theorem 2.

- (1) For every unambiguous pair of streams (x, y) with $t_y > 0$, there exists $\alpha_{(x,y)} \in \mathbb{R}_+$ such that $\rho_\alpha^{rum(pow)}(x, y)$ and $\rho_\alpha^{rum(beta)}(x, y)$ are strictly increasing in α whenever $\alpha \geq \alpha_{(x,y)}$.
- (2) For every unambiguous pair of streams (x, y) with $\frac{u(y_{t_y}) - u(x_{t_y})}{u(x_{t_x}) - u(y_{t_x})} > \frac{t_y}{t_x} > 0$, there exists $\alpha_{(x,y)} \in \mathbb{R}_+$ such that $\rho_\alpha^{rum(hyp)}(x, y)$ is strictly increasing in α whenever $\alpha \geq \alpha_{(x,y)}$.

Proof of Theorem 2: Let (x, y) be an unambiguous pair of streams. We first consider the power discount function. By Lemma 2 and the differentiability of this function, $\rho_\alpha^{rum(pow)}(x, y)$ is strictly increasing in α if and only if $\frac{\partial[U_\alpha^{pow}(x) - U_\alpha^{pow}(y)]}{\partial\alpha} = \sum \frac{\partial D_\alpha^{pow}(t)}{\partial\alpha}(u(x_t) - u(y_t)) > 0$. Given that the pair is unambiguous, $u(x_t) - u(y_t) = 0$ except for t_y and t_x , and hence the latter inequality is equivalent to $\frac{\partial D_\alpha^{pow}(t_y)}{\partial\alpha}(u(x_{t_y}) - u(y_{t_y})) + \frac{\partial D_\alpha^{pow}(t_x)}{\partial\alpha}(u(x_{t_x}) - u(y_{t_x})) > 0$. We have, by assumption, that $t_y > 0$ and hence both $\frac{\partial D_\alpha^{pow}(t_y)}{\partial\alpha}$ and $\frac{\partial D_\alpha^{pow}(t_x)}{\partial\alpha}$ are strictly negative. Therefore, the latter inequality is equivalent to $\frac{\frac{\partial D_\alpha^{pow}(t_x)}{\partial\alpha}}{\frac{\partial D_\alpha^{pow}(t_y)}{\partial\alpha}} = \frac{t_x}{t_y}(1 + \alpha)^{t_y - t_x} < \frac{u(y_{t_y}) - u(x_{t_y})}{u(x_{t_x}) - u(y_{t_x})}$. Notice that $\frac{t_x}{t_y}(1 + \alpha)^{t_y - t_x}$ is continuous and strictly decreasing for all values of α , and its limit, when the delay-aversion coefficient α tends to infinity, is zero. By the monotonicity of u , $\frac{u(y_{t_y}) - u(x_{t_y})}{u(x_{t_x}) - u(y_{t_x})}$ is strictly positive, and hence there exists $\alpha_{(x,y)}$ such that $\frac{t_x}{t_y}(1 + \alpha)^{t_y - t_x} < \frac{u(y_{t_y}) - u(x_{t_y})}{u(x_{t_x}) - u(y_{t_x})}$ for every $\alpha \geq \alpha_{(x,y)}$, as desired.

The proof of the $\beta - \delta$ case is completely analogous, since $\frac{\frac{\partial D_\alpha^{beta}(t_x)}{\partial\alpha}}{\frac{\partial D_\alpha^{beta}(t_y)}{\partial\alpha}} = \frac{\frac{\partial D_\alpha^{pow}(t_x)}{\partial\alpha}}{\frac{\partial D_\alpha^{pow}(t_y)}{\partial\alpha}}$ whenever $t_y > 0$. For the hyperbolic case, note that the limit of $\frac{\frac{\partial D_\alpha^{hyp}(t_x)}{\partial\alpha}}{\frac{\partial D_\alpha^{hyp}(t_y)}{\partial\alpha}} = \frac{t_x}{t_y} \left(\frac{1+t_y\alpha}{1+t_x\alpha} \right)^2$ is $\frac{t_y}{t_x}$. Since, by assumption, $\frac{u(y_{t_y}) - u(x_{t_y})}{u(x_{t_x}) - u(y_{t_x})} > \frac{t_y}{t_x}$, the result follows. ■

The proof of Theorem 2, using the differentiability of the standard discount functions, shows that these models are well-defined if and only if $\frac{\frac{\partial D_\alpha(t_x)}{\partial\alpha}}{\frac{\partial D_\alpha(t_y)}{\partial\alpha}} \geq \frac{u(y_{t_y}) - u(x_{t_y})}{u(x_{t_x}) - u(y_{t_x})}$. The proof states that, for all the families studied, there is always a level of delay aversion $\alpha_{(x,y)}$ after which this inequality no longer holds; hence, the models are not well-defined. The proof also explains how the critical values $\alpha_{(x,y)}$ vary with the pair of streams involved.

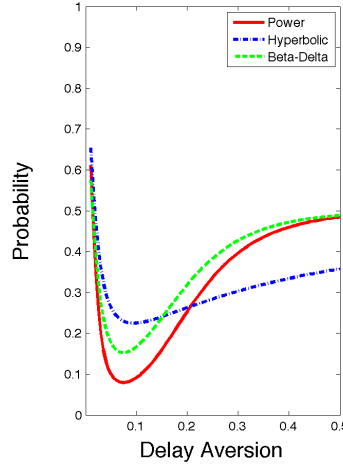


FIGURE 2.—RUM probabilities of choosing stream x over stream y .

The problems become more severe as the right hand side of the inequality increases, in direct relation to the quality of the less delayed reward y_{t_y} , and thus also to the quality of the less delayed stream y , and in inverse relation to the quality of the more delayed reward x_{t_x} , and thus also to that of x . This pattern is similar to that observed in the risk-aversion case.¹⁵ Notice that the right hand side of the above inequality can, indeed, take any real value, thus creating problems at very low levels of delay aversion.

Figure 2 illustrates the problems characterized in Theorem 2. Let us consider two streams offering 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 y offers a bonus payout of 1 with a 14-day delay, while stream x offers a bonus payout of 1.1 with a 21-day delay. That is, $y_0 = y_7 = y_{21} = 1$, $y_{14} = 2$, $x_0 = x_7 = x_{14} = 1$, $x_{21} = 2.1$, with all other payoffs set at 0. We use the logit model with $\lambda = 20$. Figure 2 plots the probability of choosing option x as a function of the discount factor α , for the power discount function, the $\beta - \delta$ discount function with $\beta = .7$, and the hyperbolic discount function, with $u(x) = x$. The probabilities of choosing x decrease with increasing α up to a certain point, after which they start to increase. Notice that the corresponding critical values of the discount factor $\alpha_{(x,y)}$ are low, so the probabilities of choosing x soon start to increase. Hence, the use of RUMs implies that the maximum discount

¹⁵A similar pattern emerges when considering the time periods. Take, for instance, the cases of the power function or $\beta - \delta$ preferences. More severe problems appear when the less delayed option is improved by reducing the delay by one period or the more delayed option is worsened by extending the delay by one period, as long as the rewards at both times are above $\frac{1}{\log(1+\alpha)}$.

factor that can be estimated in this case is low, even for individuals who are so delay averse 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 discount factors.

We close the treatment of time preferences by establishing that this problem pervades beyond the usual parametric functions used in the literature. We show next that every discounted utility RUM is problematic, since there is always an unambiguous pair of streams for which the model is not well-defined. We formally state the result by denoting the RUM probabilities of a discounted utility by $\rho^{rum(du)}$.

Proposition 5. *For every $\rho^{rum(du)}$ there exists an unambiguous pair of streams (x, y) such that $\rho^{rum(du)}$ is not well-defined for (x, y) .*

Proof of Proposition 5: Suppose, by contradiction, that $\rho^{rum(du)}$ is well-defined for every unambiguous pair of streams. We first claim that $D_{\alpha_1}(t) - D_{\alpha_2}(t)$ must be increasing in t , whenever $\alpha_1 < \alpha_2$. Otherwise, there exists a $t_2 > t_1$ such that $[D_{\alpha_1}(t_1) - D_{\alpha_2}(t_1)] > [D_{\alpha_1}(t_2) - D_{\alpha_2}(t_2)]$. Since $u(0) = 0$ and u is continuous and monotone, we can find two monetary outcomes, a and b , such that $[D_{\alpha_1}(t_1) - D_{\alpha_2}(t_1)]u(a) > [D_{\alpha_1}(t_2) - D_{\alpha_2}(t_2)]u(b)$, or, equivalently, $D_{\alpha_1}(t_2)u(b) - D_{\alpha_1}(t_1)u(a) < D_{\alpha_2}(t_2)u(b) - D_{\alpha_2}(t_1)u(a)$. Defining the stream x by $x_{t_2} = a$ and $x_t = 0$ otherwise, and the stream y by $y_{t_1} = b$ and $y_t = 0$ otherwise, they are unambiguous, and the latter inequality, together with Lemma 1, leads to $\rho_{\alpha_1}^{rum(du)}(x, y) > \rho_{\alpha_2}^{rum(du)}(x, y)$, a contradiction which proves that $D_{\alpha_1}(t) - D_{\alpha_2}(t)$ is increasing in t . Now, notice that $D_{\alpha_1}(0) - D_{\alpha_2}(0) = 0$ and $\lim_{t \rightarrow \infty} [D_{\alpha_1}(t) - D_{\alpha_2}(t)] = 0$. Since in addition we have shown that $D_{\alpha_1}(t) - D_{\alpha_2}(t)$ is increasing in t , $D_{\alpha_1}(t) - D_{\alpha_2}(t)$ must be constantly null in t , which is a contradiction that proves the proposition. ■

4.3. Other Applications. In this section we comment on other applications where our findings on the use of RUMs may also be of importance.

Our results relating to risk preferences immediately extend to situations in which the individual operates under conditions of strategic uncertainty, and, consequently, beliefs replace objective probabilities. A prominent example of this approach in game theory is the quantal response equilibrium of McKelvey and Palfrey (1995), which assumes a RUM using (subjective) expected utility. Hence, for given beliefs and according to our results, it may be the case that the more risk-averse individual has a larger probability of choosing the riskier action.

It is clear, moreover, that the problems identified in the static setting studied here are immediately inherited by dynamic discrete choice models, which are frequently used to address a variety of economic problems. Starting with the seminal papers by Wolpin (1984) and Rust (1987), dynamic discrete choice models have been used to tackle 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).¹⁶ The potential problems in this literature are twofold. Firstly, some of these settings involve risk and are modeled by means of RUMs with errors over expected utility, and hence the results in our section 4.1 are directly applicable. Secondly, the dynamic nature of the setting makes the results in our section 4.2 relevant here also.

Beyond risk and time, another preference parameter of interest is the one governing the degree of complementarity between two different inputs. These inputs may be the monetary payoffs to oneself and to another subject, as in a distributive problem with social preferences (see, e.g., Andreoni and Miller, 2002). Another case of interest in this respect is when the inputs refer to present consumption and future consumption, as in the influential Epstein and Zin (1989) preferences.¹⁷ Yet, another example is when the inputs refer to different consumption goods in general, as in a standard CES utility function. Our results advise caution when the complementarity parameter enters non-linearly into the utility function in a RUM estimation framework.

5. RANDOM PREFERENCE MODELS

In a random preference model (RPM), a set of utility functions $\{U_\omega\}_{\omega \in \Omega}$ is considered. At the moment of choice, the value of ω is drawn randomly according to a continuous cumulative distribution function Φ_α , that depends on the preference parameter α .¹⁸ We assume that whenever $\alpha_1 \leq \alpha_2$, Φ_{α_2} first order stochastically dominates Φ_{α_1} .¹⁹ Then, the individual selects the alternative that provides maximal utility according to

¹⁶The literature using dynamic discrete choice models is vast; see Aguirregabiria and Mira (2010) for a survey.

¹⁷These preferences also introduce risk attitudes and time preferences, so sections 4.1 and 4.2 are of interest here too.

¹⁸Again, the distribution function can depend on further parameters. A dispersion parameter, for instance, could be related to the rationality of the individual.

¹⁹Formally, Φ_{α_2} first order stochastically dominates Φ_{α_1} whenever $\Phi_{\alpha_2}(\omega) \leq \Phi_{\alpha_1}(\omega)$ for every $\omega \in \Omega$.

the realized utility U_ω . Given the pair of alternatives (x, y) , choice stochasticity is generated from the fact that, for certain utility realizations option x is preferred to y , and for other realizations the reverse ranking holds.²⁰ We denote the set of ω -realizations for which $U_\omega(x) \geq U_\omega(y)$ by $\Omega_{(x,y)}$, and its probability by $\rho_\alpha^{rpm}(x, y)$.

We can now establish a sufficient condition for a RPM to be well-defined for an unambiguous pair (x, y) .

Lemma 3. *ρ^{rpm} is well-defined for the unambiguous pair (x, y) if $\omega' \leq \omega''$ for every $\omega' \in \Omega_{(x,y)}$ and $\omega'' \in \Omega \setminus \Omega_{(x,y)}$.*

Proof of Lemma 3: Consider the unambiguous pair (x, y) . Let us start by supposing that $\Omega_{(x,y)} = \Omega$ (respectively, $\Omega_{(x,y)} = \emptyset$). Then, $\rho_\alpha^{rpm}(x, y) = 1$ (respectively, $\rho_\alpha^{rpm}(x, y) = 0$) for every α . Otherwise, there exists ω_{xy} within Ω , such that $U_\omega(x) > U_\omega(y)$ whenever $\omega < \omega_{xy}$ and $U_\omega(y) > U_\omega(x)$ whenever $\omega > \omega_{xy}$. In this case, $\rho_\alpha^{rpm}(x, y) = \Phi_\alpha(\omega_{xy})$. Hence, whenever $\alpha_1 \leq \alpha_2$, we know by assumption that Φ_{α_2} first order stochastically dominates Φ_{α_1} and consequently, $\Phi_{\alpha_2}(\omega_{xy}) \leq \Phi_{\alpha_1}(\omega_{xy})$, or, equivalently, $\rho_{\alpha_2}^{rpm}(x, y) \leq \rho_{\alpha_1}^{rpm}(x, y)$.²¹ This concludes the proof. ■

The condition in Lemma 3 can be restated imposing the non-existence of values $\omega' \in \Omega_{(x,y)}$ and $\omega'' \in \Omega \setminus \Omega_{(x,y)}$, such that $\omega' > \omega''$. Lemma 3 establishes that this condition is sufficient for a RPM to be well-defined. Indeed, the condition is also virtually necessary. To see this, notice the necessity of the following condition: there do not exist sets $A \subseteq \Omega_{(x,y)}$ and $B \subseteq \Omega \setminus \Omega_{(x,y)}$ with measure different from zero, such that $a > b$ for every $a \in A$ and $b \in B$.

The next corollary states that the use of expected utilities in risk settings, or discounted utilities in time settings, in conjunction with RPMs, leads to well-defined stochastic models for every unambiguous pair (x, y) , provided that the family of utilities $\{U_\omega\}_{\omega \in \Omega}$ is well-ordered in terms of risk aversion or delay aversion, respectively. For expected utility, $\{U_\omega\}_{\omega \in \Omega}$ is well-ordered in terms of risk aversion whenever the family is ordered in the Arrow-Pratt sense. This is, if $\omega' < \omega''$, the utility function $u_{\omega''}$ is a concave transformation of the utility function $u_{\omega'}$. We assume, moreover, that there exists a value of ω , which, for the sake of simplicity, we take to be zero, such that

²⁰We assume that the probability of a tie is equal to zero, by imposing that $U_\omega(x) = U_\omega(y)$ holds for no more than one value of $\omega \in \Omega$.

²¹Notice that, whenever Φ_α has full support on Ω and stochastic dominance is strict, the probability $\rho_\alpha^{rpm}(x, y)$ is strictly decreasing in α .

u_0 represents a risk-neutral individual. For discounted utility, $\{U_\omega\}_{\omega \in \Omega}$ is well-ordered in terms of delay aversion if, for every $t > t'$, $\frac{D_\omega(t)}{D_\omega(t')}$ is decreasing in ω . This property is shown by Benoît and Ok (2007) to be the analogue of the Arrow-Pratt measure in a time-preference context. Notice that these conditions are mild and natural, and satisfied by every parametric family of expected utility and discounted utility considered in the paper. In order to formally establish the result, denote by $\rho^{rpm(eu)}$ and $\rho^{rpm(du)}$ the RPM built upon such well-ordered families of utilities.

Theorem 3. $\rho^{rpm(eu)}$ (respectively, $\rho^{rpm(du)}$) is well-defined for every unambiguous pair of gambles (respectively, streams) (x, y) .

Proof of Theorem 3: Let us begin with the case of $\rho^{rpm(eu)}$. Consider an unambiguous pair of gambles (x, y) of the degenerate or nested type. It is immediate that, in these pairs of gambles, gamble y has the single-crossing property with respect to gamble x .²² Suppose that for a given realization ω , $U_\omega(y) \geq U_\omega(x)$. Since, by assumption, whenever $\omega' > \omega$, utility $u_{\omega'}$ is a concave transformation of u_ω , it is a well-known fact that the single-crossing property implies that $U_{\omega'}(y) \geq U_{\omega'}(x)$.²³ Thus, our Lemma 3 applies, showing that the model is well-defined for these pairs of gambles. Now consider an unambiguous pair of gambles (x, y) , where x is a mean-preserving spread of y . Given the assumptions, it clearly follows that $U_\omega(x) \geq U_\omega(y)$ if and only if $\omega \leq 0$ and hence, Lemma 3 again applies. Let us now consider $\rho^{rpm(du)}$ and an unambiguous pair of streams (x, y) . From Theorem 1 and Corollary 2 in Benoît and Ok (2007), our assumption over the family of discounted utilities again guarantees that, whenever $U_\omega(y) \geq U_\omega(x)$ and $\omega' > \omega$, it is also the case that $U_{\omega'}(y) \geq U_{\omega'}(x)$, and, once more, Lemma 3 can be used to conclude the proof. ■

Theorem 3 shows that RPMs, along with large classes of utility functions, are free from the problems identified in the RUMs. Notice that Theorem 3 establishes the result for the most general representation of time preferences considered in the paper, and also for expected utility in the case of risk preferences. We now argue that, in fact, the model also works well beyond expected utility. Consider first the case of generalized expected utility. Whenever the unambiguous pair of gambles (x, y) is transformed by

²²Formally speaking, there exists a payoff \bar{m} such that $F_x(m) - F_y(m) \leq 0$ whenever $m \leq \bar{m}$ and $F_x(m) - F_y(m) \geq 0$ whenever $m \geq \bar{m}$, where F_x and F_y are the cumulative distributions over payoffs determined by gambles x and y , respectively.

²³See for instance, Lemma 1 in Grossman and Hart (1983).

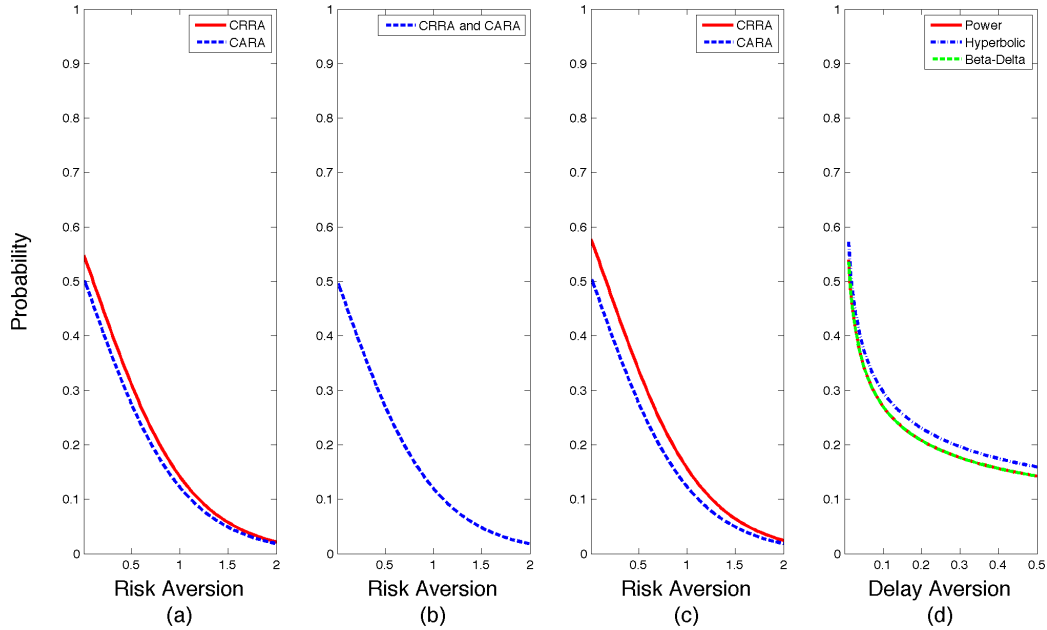


FIGURE 3.—RPM probabilities of choosing (a) gamble x over gamble y , (b) gamble x over gamble y' , (c) gamble x over gamble \bar{y} , and (d) stream x over stream y .

the generalized expected utility into another unambiguous pair of gambles $(\pi(x), \pi(y))$, the logic of Theorem 3 immediately applies, thus establishing that the model is well-defined for (x, y) . This is the case when comparing gambles of the degenerate type, since these pairs are always transformed into other pairs of gambles of the degenerate type. As for the nested type of gambles, notice that they are always transformed into pairs of gambles satisfying the single-crossing property, and hence the same procedure as in Theorem 3 would prove that claim. Notice that the mean-preserving spread case is more convoluted because, clearly, given the wide and unstructured nature of our class of generalized expected utility, transformations of two gambles related by a mean-preserving spread tend not to be related by mean-preserving spreads, or to satisfy the single-crossing property. However, the specification of particular classes of generalized expected utilities may be perfectly compatible with the use of RPMs when studying gambles of the mean-preserving spread type. Lemma 3 provides a simple condition to check whether this is the case. Finally, it is immediate to see that mean-variance utilities satisfy the condition in Lemma 3 and hence also work well with RPMs.

Figure 3 illustrates Theorem 3 by plotting the RPM probabilities of choosing the gambles and streams considered in Figures 1 and 2. Figures 3a, 3b, and 3c use the

logistic distribution while, since we have assumed that the discount factor takes values in the positive reals, Figure 3d uses the log-logistic distribution.

5.1. Remarks. In Lemma 3 we see that the key element for a well-defined RPM is the existence of a cutting point ω_{xy} , such that individuals with $\omega \geq \omega_{xy}$ prefer y over x , and vice-versa. This suggests that we can easily relax the assumption of the existence of a parametric family of utility functions $\{U_\omega\}_{\omega \in \Omega}$. Consider the case of risk preferences, for instance. Under the appropriate assumptions, one can consider the Taylor expansions of utility functions and work directly on the Arrow-Pratt risk-aversion coefficient. In this case, the consideration of a probability distribution over the Arrow-Pratt coefficient would lead to the same positive results provided by Theorem 3. This is basically the approach adopted in Cohen and Einav (2007).

A distinguishing feature of RPMs is that when, for a given pair of options (x, y) , every utility function regards one option as better than the other, then the probability of choosing the former is one. This is sometimes seen as a limitation of the model, as in the case of gambles related by stochastic dominance, for instance, where the observed probability of choosing the dominated gamble is typically above zero. One way to deal with this in the context of RPMs is to add a trembling stage, in the spirit of the trembling hand approach used in game theory. This would work as follows. After a particular utility has been realized, with a large probability $1 - \kappa$ the choice is made according to the realized utility, and with probability κ there is a tremble and the reverse choice is made. It is easy to see that such a model is also well-defined. See Harless and Camerer (1994) for an application of the trembling hand approach to the estimation of risk aversion, which can be interpreted as a RPM with a degenerate probability distribution Φ_α plus trembling.

6. FINAL REMARKS

We have introduced here the notion of a well-defined stochastic choice model with respect to a preference parameter. Namely, consider a pair of alternatives (x, y) unambiguously ordered with respect to the preference parameter under consideration, where larger values of the parameter reflect more aversion to selecting x over y . We then say that a minimal property for a stochastic model to be well-defined is that the probability of selecting x should decrease as the aversion to choosing x increases.

We have focused on two popular stochastic models, random utility models and random preference models. After establishing the conditions for these models to be well-defined, we have focused on the particular cases of risk and delay aversion. We have shown that the standard application of random utility models to risk or time settings is subject to serious inconsistencies. In the main results we have shown that there is a level of risk aversion (respectively, of delay aversion) beyond which the probability of choosing the riskier gamble (respectively, the more delayed stream) increases with the level of risk aversion (respectively, of delay aversion). We have then established that random preference models are free from all these inconsistencies, and hence are well-defined. These findings should constitute an alert to exercise caution when directly applying sound stochastic choice models to settings other than those originally contemplated.

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 incongruities in environments involving risk or time.

APPENDIX A. RUM WITH LOGARITHMIC TRANSFORMATION OF THE REPRESENTATIVE UTILITY

This approach starts by assuming that the representative utility of every option is strictly positive. Then, the probability of selecting option x over option y is simply $P(\lambda \log(U_\alpha(x)) + \epsilon \geq \lambda \log(U_\alpha(y)) + \hat{\epsilon}) = \Psi^*(\lambda \log(U_\alpha(x)) - \lambda \log(U_\alpha(y))) = \Psi^*(\lambda \log(\frac{U_\alpha(x)}{U_\alpha(y)}))$, where Ψ^* is the distribution function of the difference of the i.i.d. errors. Paralleling Lemma 1, it is now immediate that a RUM based on the logs of the utilities, LRUM, is well-defined if and only if the ratio $\frac{U_\alpha(x)}{U_\alpha(y)}$ is strictly decreasing in α . Lemma 2 can also be directly reproduced considering $\lim_{\alpha \uparrow} [\log(U_\alpha(x)) - \log(U_\alpha(y))] = 0$. In order to establish the main results in the applications, denote by $\rho^{lrum(cara)}$, $\rho^{lrum(pow)}$, $\rho^{lrum(beta)}$ and $\rho^{lrum(hyp)}$ the corresponding probabilities for CARA, power, $\beta - \delta$ and hyperbolic utilities.²⁴

²⁴Notice that CRRA functions are not entirely suitable in this context. This is because, for values of α above 1, utilities become negative, which is incompatible with the use of log-transformations. The function $x^{1-\alpha}$, without the normalization $\frac{1}{1-\alpha}$, is positive for values of $\alpha > 1$, but is not in this case monotone in outcomes and thus is also problematic.

Theorem 4.

- (1) For every unambiguous pair of gambles (x, y) , there exists $\alpha_{(x,y)} \in \mathbb{R}$ such that $\rho_{\alpha}^{lrum(cara)}(x, y) > \rho_{\alpha_{(x,y)}}^{lrum(cara)}(x, y)$ whenever $\alpha > \alpha_{(x,y)}$.
- (2) For every unambiguous pair of streams (x, y) with $t_y > 0$ and $y_t > 0$ for some $t < t_y$, there exists $\alpha_{(x,y)} \in \mathbb{R}_+$ such that $\rho_{\alpha}^{lrum(pow)}(x, y)$ and $\rho_{\alpha}^{lrum(beta)}(x, y)$ are strictly increasing in α whenever $\alpha \geq \alpha_{(x,y)}$.
- (3) For every unambiguous pair of streams (x, y) with $\frac{u(y_{t_y}) - u(x_{t_y})}{u(x_{t_x}) - u(y_{t_x})} > \frac{t_y}{t_x} > 0$ and $y_0 > 0$, there exists $\alpha_{(x,y)} \in \mathbb{R}_+$ such that $\rho_{\alpha}^{lrum(hyp)}(x, y)$ is strictly increasing in α whenever $\alpha \geq \alpha_{(x,y)}$.

Proof of Theorem 4: We begin with the risk-aversion case. Consider any unambiguous pair of gambles (x, y) . Given the structure of gambles and the CARA family, there always exists $\hat{\alpha} > 0$ such that $U_{\hat{\alpha}}^{cara}(x) < U_{\hat{\alpha}}^{cara}(y)$, or $\log(U_{\hat{\alpha}}^{cara}(x)) < \log(U_{\hat{\alpha}}^{cara}(y))$. It is also immediate to see that the limits of $\log(\alpha U_{\alpha}^{cara}(x))$ and $\log(\alpha U_{\alpha}^{cara}(y))$ as α increases are both 0 and hence $\lim_{\alpha \uparrow} [\log(U_{\alpha}(x)) - \log(U_{\alpha}(y))] = \lim_{\alpha \uparrow} [\log(\alpha U_{\alpha}(x)) - \log(\alpha U_{\alpha}(y))] = 0$. Therefore, there exists $\tilde{\alpha} > \hat{\alpha}$ such that $\log(U_{\alpha}^{cara}(x)) - \log(U_{\alpha}^{cara}(y)) > \log(U_{\hat{\alpha}}^{cara}(x)) - \log(U_{\hat{\alpha}}^{cara}(y))$ for every $\alpha \geq \tilde{\alpha}$. This, together with the logarithmic counterpart of Lemma 1, implies that $\rho_{\alpha}^{lrum(cara)}(x, y) > \rho_{\tilde{\alpha}}^{lrum(cara)}(x, y)$ for every $\alpha \geq \tilde{\alpha}$. Now, the function $\rho_{\alpha}^{lrum(cara)}(x, y)$ is continuous on $[\hat{\alpha}, \infty)$ and, hence, achieves a minimum α^* in the closed interval $[\hat{\alpha}, \tilde{\alpha}]$, and, by the above reasoning, we know that α^* is also a minimum in $[\hat{\alpha}, \infty)$. Given continuity, we only need to consider $\alpha_{(x,y)}$ to be the largest value of α for which $\rho_{\alpha_{(x,y)}}^{lrum(cara)}(x, y) = \rho_{\alpha^*}^{lrum(cara)}(x, y)$ and the result follows.

We now analyze the discounting case, starting with the power function. From the logarithmic version of Lemma 1 and the differentiability of the discounted utility, monotonicity of $\rho_{\alpha}^{lrum(pow)}(x, y)$ depends on the derivative of $\frac{\sum_t D_{\alpha}^{pow}(t)u(x_t)}{\sum_t D_{\alpha}^{pow}(t)u(y_t)}$ with respect to α . It is not difficult to see that the sign of this derivative is the same as that of $\sum_t \frac{\partial D_{\alpha}^{pow}(t)}{\partial \alpha} u(x_t) [\sum_t D_{\alpha}^{pow}(t)u(y_t)] - \sum_t \frac{\partial D_{\alpha}^{pow}(t)}{\partial \alpha} u(y_t) [\sum_t D_{\alpha}^{pow}(t)u(x_t)]$. Now notice that $\frac{\partial D_{\alpha}^{pow}(0)}{\partial \alpha} = 0$, and hence, the sign of the previous expression is equivalent to the sign of $-\sum_t t D_{\alpha}^{pow}(t)u(x_t) [\sum_t D_{\alpha}^{pow}(t)u(y_t)] + \sum_t t D_{\alpha}^{pow}(t)u(y_t) [\sum_t D_{\alpha}^{pow}(t)u(x_t)]$. This is $\sum_r \sum_s (s - r) D_{\alpha}^{pow}(r) D_{\alpha}^{pow}(s) u(x_r) u(y_s)$ or, simply, $\sum_r \sum_s (s - r) D_{\alpha}^{pow}(r + s) u(x_r) u(y_s)$. To analyze the sign of the previous expression when α is sufficiently large, we only need to consider the smallest integer m for which the term $\sum_r \sum_{s:r+s=m} (s - r) D_{\alpha}^{pow}(m) u(x_r) u(y_s)$ is different from zero. Now, let t^* be the smallest integer such that $t^* < t_y$ with $y_{t^*} > 0$, which exists by assumption. Any sum where $m < t^* + t_y$ is equal to zero, while the sum $\sum_r \sum_{s:r+s=t^*+t_y} (s - r) u(x_r) u(y_s)$ is equal to

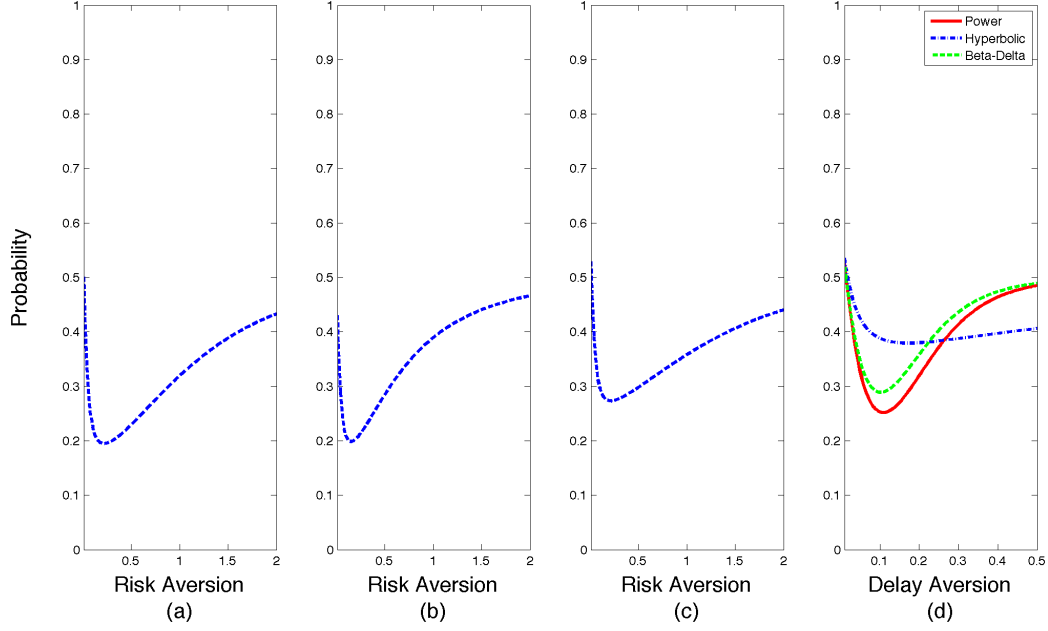


FIGURE 4.—LRUM probabilities of choosing (a) gamble x over gamble y , (b) gamble x over gamble y' , (c) gamble x over gamble \bar{y} , and (d) stream x over stream y .

$(t_y - t^*)u(y_{t^*})(u(y_{t_y}) - u(x_{t_y}))$, which is strictly positive by the assumptions on x and y . This makes the desired derivative strictly positive above certain value $\alpha_{(x,y)}$ and hence $\rho_\alpha^{lrum(pow)}(x, y)$ is strictly increasing above $\alpha_{(x,y)}$.

Let us now consider the case of $\beta - \delta$ preferences. As in the case of the power function, it is immediate to see that we need to analyze the sign of $\beta \sum_s s D_\alpha^{pow}(s) u(x_0) u(y_s) - \beta \sum_r r D_\alpha^{pow}(r) u(x_r) u(y_0) + \beta^2 \sum_{r>0} \sum_{s>0} (s-r) D_\alpha^{pow}(r+s) u(x_r) u(y_s)$. Since $\beta > 0$, for sufficiently high values of α , the sign is equivalent to the sign of $(t_y - t^*)u(y_{t^*})(u(y_{t_y}) - u(x_{t_y}))$. Since $t_y > t^*$ and $y_{t_y} > x_{t_y}$, this term is strictly positive and the result follows.

The same reasoning can be used in the hyperbolic case, where it leads to the analysis of the sign of $\sum_r \sum_s (s-r) [D_\alpha^{hyp}(r) D_\alpha^{hyp}(s)]^2 u(x_r) u(y_s)$. When α 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 forms $s [D_\alpha^{hyp}(s)]^2 u(x_0) u(y_s)$ and $-r [D_\alpha^{hyp}(r)]^2 u(x_r) u(y_0)$. To study the sign of their sum, simply notice that, as α increases, the limit of $\frac{D_\alpha^{hyp}(a)}{D_\alpha^{hyp}(b)}$ is b/a . Hence, the determining expression is $\sum_s \frac{1}{s} u(x_0) u(y_s) - \sum_r \frac{1}{r} u(x_r) u(y_0)$, which is equal to $\frac{1}{t_y} (u(x_0) u(y_{t_y}) - u(x_{t_y}) u(y_0)) + \frac{1}{t_x} (u(x_0) u(y_{t_x}) - u(x_{t_x}) u(y_0))$, with $x_0 = y_0 > 0$. Clearly, this value is strictly positive and the result follows. ■

Figure 4 is analogous to Figures 1 and 2, but it uses the corresponding LRUM. In Figures 4a, 4b and 4c, and in line with footnote 24, we only report on the CARA utilities. It is apparent that the same sort of problems illustrated in Figures 1 and 2 also show up in this case, and hence basically the same conclusions as reached in the previous subsections also apply here.

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] Andreoni, J., and J. Miller (2002) “Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism,” *Econometrica*, 70:737–53.
- [6] Barseghyan, L., F. Molinari, T. O’Donoghue, and J. C. Teitelbaum (2013), “The Nature of Risk Preferences: Evidence from Insurance Choices,” *American Economic Review*, 103.6:2499–2529.
- [7] Barsky, R. B., M. S. Kimball, F. T. Juster, and M. D. Shapiro (1997), “Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study,” *Quarterly Journal of Economics*, 112.2:537–579.
- [8] Becker, G. M., M. H. DeGroot, and J. Marschak (1963), “Stochastic Models of Choice Behavior,” *Behavioral Science*, 8.1:41–55.
- [9] Benoît, J-P. and E. A. Ok (2007), “Delay Aversion,” *Theoretical Economics*, 2:71–113.
- [10] Berkovec, J. and S. Stern (1991), “Job Exit Behavior of Older Men,” *Econometrica*, 59.1:189–210.
- [11] Blavatskyy, Pavlo R. (2011), “Probabilistic Risk Aversion with an Arbitrary Outcome Set,” *Economics Letters*, 1112:34–37.
- [12] Busemeyer, J. R. and J. T. Townsend (1993), “Decision Field Theory: A Dynamic-Cognitive Approach to Decision Making in an Uncertain Environment,” *Psychological Review*, 100.3:432–459.
- [13] 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.
- [14] Cicchetti, C. J. and J. A. Dubin (1994), “A Microeconomic Analysis of Risk Aversion and the Decision to Self-Insure,” *Journal of Political Economy*, 102.1:169–186.
- [15] Cohen, A. and L. Einav (2007), “Estimating Risk Preferences from Deductible Choice,” *American Economic Review*, 97.3:745–788.

- [16] Coller, M. and M. B. Williams (1999), “Eliciting Individual Discount Rates,” *Experimental Economics*, 2.2:107–127.
- [17] Crawford, G. S. and M. Shum (2005), “Uncertainty and Learning in Pharmaceutical Demand,” *Econometrica*, 73.4:1137–1173.
- [18] Diermeier, D., M. Keane and A. Merlo (2005), “A Political Economy Model of Congressional Careers,” *American Economic Review*, 95.1:347–373.
- [19] Eliashberg, J. and J. R. Hauser (1985), “A Measurement Error Approach for Modeling Consumer Risk Preference,” *Management Science*, 31.1:1–25.
- [20] Epstein, L. G. and S. E. Zin (1989), “Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework,” *Econometrica* 57(4):937–969.
- [21] Friedman, B. (1974), “Risk Aversion and the Consumer Choice of Health Insurance Option,” *Review of Economics and Statistics*, 56.2:209–214.
- [22] Fullenkamp, C., R. Tenorio, and R. Battalio (2003), “Assessing Individual Risk Attitudes Using Field Data from Lottery Games,” *Review of Economics and Statistics*, 85.1:218–226.
- [23] Gilleskie, D. B. (1998), “A Dynamic Stochastic Model of Medical Care Use and Work Absence,” *Econometrica*, 66.1:1–45.
- [24] Grossman, S. J. and O. Hart (1983), “Implicit Contracts under Asymmetric Information,” *Quarterly Journal of Economics*, 98:123–156.
- [25] Gul, F., P. Natenzon, and W. Pesendorfer (2014), “Random Choice as Behavioral Optimization,” *Econometrica*, forthcoming.
- [26] Gul, F. and W. Pesendorfer (2006), “Random Expected Utility,” *Econometrica*, 74.1:121–146.
- [27] Harless, D. W. and C. F. Camerer (1994), “The Predictive Utility of Generalized Expected Utility Theories,” *Econometrica*, 62.6:1251–1289.
- [28] Harrison, G. W., J. A. List, and C. Towe (2007), “Naturally Occurring Preferences and Exogenous Laboratory Experiments: A Case Study of Risk Aversion,” *Econometrica*, 75.2:433–458.
- [29] Hey, J. D. and C. Orme (1994), “Investigating Generalizations of Expected Utility Theory Using Experimental Data,” *Econometrica*, 62.6:1291–1326.
- [30] Holt, C. A. and S. K. Laury (2002), “Risk Aversion and Incentive Effects,” *American Economic Review*, 92.5:1644–1655.
- [31] Horowitz, J. K. (1992), “Comparative Impatience,” *Economics Letters*, 38:25–29.
- [32] 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.
- [33] Kimball, M. S., C. R. Sahm, and M. D. Shapiro (2008), “Imputing Risk Tolerance From Survey Responses,” *Journal of the American Statistical Association*, 103.483 :1028–1038.
- [34] Kimball, M. S., C. R. Sahm, and M. D. Shapiro (2009), “Risk Preferences in the PSID: Individual Imputations and Family Covariation,” *American Economic Review Papers and Proceedings*, 99.2:363–368.
- [35] Laibson, D. (1997), “Golden Eggs and Hyperbolic Discounting,” *Quarterly Journal of Economics*, 112.2:443–477.

- [36] Loewenstein, G. and D. Prelec (1992), “Anomalies in Intertemporal Choice: Evidence and an Interpretation,” *Quarterly Journal of Economics*, 107.2, 573–597.
- [37] Loomes, G. and R. Sugden (1995), “Incorporating a Stochastic Element Into Decision Theories,” *European Economic Review*, 39.3:641–648.
- [38] Luce, R.D. (1959), *Individual Choice Behavior: A Theoretical Analysis*, Wiley, New York.
- [39] Manzini P. and M. Mariotti (2014), “Stochastic Choice and Consideration Sets,” *Econometrica*, forthcoming.
- [40] Markowitz, H. (1952), “Portfolio Selection,” *Journal of Finance*, 7.1:77–91.
- [41] McFadden, D. (2001), “Economic Choices,” *American Economic Review*, 91.3:351–378.
- [42] McKelvey, R. D. and T. R. Palfrey (1995), “Quantal Response Equilibria for Normal Form Games,” *Games and Economic Behavior*, 10.1:6–38.
- [43] Meier, S. and C. D. Sprenger (2014), “Temporal Stability of Time Preferences,” *Review of Economics and Statistics*, forthcoming.
- [44] Noussair, C. N., S. T. Trautmann, and G. Van de Kuilen (2014), “Higher Order Risk Attitudes, Demographics, and Financial Decisions,” *Review of Economic Studies*, 81.1:325–355.
- [45] O’Donoghue, T. and M. Rabin (1999), “Doing It Now or Later,” *American Economic Review*, 89.1:103–124.
- [46] Post, T., M. J. Van den Assem, G. Baltussen, and R. H. Thaler (2008), “Deal or No Deal? Decision Making under Risk in a Large-Payoff Game Show,” *American Economic Review*, 98.1:38–71.
- [47] Pratt, J. W. (1964), “Risk Aversion in the Small and in the Large,” *Econometrica*, 32.1/2:122–136.
- [48] Roberts, J. H. and G. L. Urban (1988), “Modeling Multiattribute Utility, Risk, and Belief Dynamics for New Consumer Durable Brand Choice,” *Management Science*, 34.2:167–185.
- [49] Rust, J. (1987), “Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher,” *Econometrica*, 55.5:999–1033.
- [50] Rust, J. and C. Phelan (1997), “How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets,” *Econometrica*, 65:781832.
- [51] 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.
- [52] Thurstone, L.L. (1927), “A Law of Comparative Judgment,” *Psychological Review* 34:273–286.
- [53] 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.
- [54] Train, K. E. (2009), “Discrete Choice Methods with Simulation,” *Cambridge University Press*
- [55] Von Gaudecker, H. M., A. Van Soest, and E. Wengström (2011), “Heterogeneity in Risky Choice Behavior in a Broad Population,” *American Economic Review*, 101.2:664–694.
- [56] Warner, J. T. and S. Pleeter (2001), “The Personal Discount Rate: Evidence from Military Downsizing Programs,” *American Economic Review*, 91.1:33–53.

- [57] Wilcox, N.T. (2008), “Stochastic Models for Binary Discrete Choice Under Risk: A Critical Primer and Econometric Comparison,” in J. C. Cox and G. W. Harrison, eds., *Research in Experimental Economics Vol. 12: Risk Aversion in Experiments*, 197–292. Bingley, UK: Emerald.
- [58] Wilcox, N.T. (2011), “Stochastically More Risk Averse: A Contextual Theory of Stochastic Discrete Choice Under Risk,” *Journal of Econometrics*, 162.1:89–104.
- [59] Wolpin, K. I. (1984), “An Estimable Dynamic Stochastic Model of Fertility and Child Mortality,” *Journal of Political Economy*, 92.5:852–874.