Discrete Choice Estimation of Risk Aversion

José Apesteguía
Miguel Angel Ballester

September 2014

Barcelona GSE Working Paper Series

Working Paper nº 788

DISCRETE CHOICE ESTIMATION OF RISK AVERSION*

JOSE APESTEGUIA† AND MIGUEL A. BALLESTER‡

Abstract. We analyze the use of discrete choice models for the estimation of risk aversion and show a fundamental flaw in the standard random utility model which is commonly used in the literature. Specifically, we find that given two gambles, the probability of selecting the riskier gamble may be larger for larger levels of risk aversion. We characterize when this occurs. By contrast, we show that the alternative random preference approach is free of such problems.

Keywords: Discrete Choice; Structural Estimation; Risk Aversion; Random Utility Models; Random Preference Models.
JEL classification numbers: C25; D81.

1. Introduction

Our understanding of labor, finance, development, health or insurance issues relies critically on an adequate comprehension of individual risk attitudes. Many of the decisions involved in these settings are of a categorical nature: whether or not to enter the labor market, which pension plan to adopt, which crop to sow on a certain plot of land, which drug to prescribe for a particular illness, or whether to accept a deductible in an insurance contract are all discrete decisions involving an element of risk. It is unsurprising, therefore, that there is a large stream of literature attempting to estimate risk aversion through discrete choice analysis.

Date: September, 2014.
* We thank Pavlo Blavatskyy, Jerome Busemeyer, Syngjoo Choi, Kfir Eliaz, Holger Gerhardt, Faruk Gul, Glenn Harrison, Paola Manzini, Marco Mariotti, Isabel Melguizo, Efe A. Ok, Ariel Rubinstein, Larry Samuelson, Karl Schlag, Matthew Shum, 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, 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`.

The discrete choice approach to the estimation of risk aversion starts by specifying the choice probabilities associated with the different gambles. This is basically done by combining the valuation of the gambles with some sort of error. Two broad classes of discrete choice models are used in the literature. We refer to them here as random utility models and random preference models. In the former, the error is assumed to affect the utility of each gamble independently. In the latter, it affects the risk aversion parameter, which ultimately determines the valuation of the gambles. In this paper, we show that the use of random utility models generally leads to a fundamental problem for the estimation of risk aversion. Namely, utility functions with greater risk aversion may with a larger probability choose the riskier gamble.[1] Thus, given two individuals where one is more risk averse than the other, the random utility model estimation may give a lower risk aversion coefficient to the first individual than to the second, thereby seriously disputing the validity of the estimation exercise. Fortunately, we are able to show that the estimation using random preference models is free of this problem.

To be more specific, let us study the comparison of two gambles, where one is clearly riskier than the other. We consider the two well-accepted cases: (i) a risky gamble and a degenerate gamble giving a monetary payoff with certainty, and (ii) two gambles, where one is a mean-preserving spread of the other. We are now in a position to detail the main results of this paper.

Random Utility Models. The extent of the problem depends on the utility representation used in the valuation of gambles. The most direct, and naturally, the most often used, is the one using the expected utility form. In Theorem 1 we show that the problem affects every pair of expected utility functions where one is more risk averse than the other. We then characterize the range of pairs of gambles in which the problem arises. We illustrate further by using the main parametric functional forms used in the literature, namely CARA and CRRA. We show in Theorem 2 that, for every pair of gambles where one is riskier than the other, there exists a level of risk aversion above which the probability of choosing the riskier gamble increases. Hence, this establishes

[1] In this introduction we focus on expected utility, where more risk aversion is equivalent to more curvature of the utility function over monetary payoffs. The implications for utility models beyond expected utility are discussed in subsequent sections.

that there is an upper bound of the risk aversion parameter that can be estimated, however risk averse the individual may be.[2]

The intuition of these results goes as follows. Random utility models derive probabilities from the cardinal difference in the valuation of the gambles. Essentially, a larger difference implies less probability of choosing the inferior option. When using the expected utility representation, greater risk aversion is associated with higher curvature of the utility function of the monetary payoff. As the curvature increases, the cardinal difference between the valuations of the gambles is reduced, and hence greater risk aversion may imply that the riskier gamble is chosen with a larger probability.

Alternatively, the literature on occasions uses the certainty equivalent as a utility representation. We show that, in this case, the results depend heavily on the gambles that are being compared. When the comparison involves a degenerate gamble, Theorem 3 shows that there is no problem whatsoever. This is so because the certainty equivalent of the degenerate gamble is the same for all utility functions, and hence the difference between the certainty equivalent of a risky gamble and a degenerate one depends exclusively on the value of the former, which is obviously decreasing with risk aversion. Unfortunately, however, since the anchor created by the degenerate gamble disappears as soon as non-degenerate gambles are considered, Theorems 4 and 5 show that the problem reappears in the mean-preserving spread case.

Random Preference Models. Estimation by random preference models starts by considering a parametric family of utilities, and assumes that the individual has a risk aversion level subject to error. Theorem 6 establishes that random preference models are free from the problem discussed above. The intuition is as follows. In random preference models, the choice probability of selecting one gamble is determined by the probability that the error on the risk aversion level causes the individual to rank this gamble as the best. Now, consider a higher level of risk aversion. Then, notice that for any given realization of the error, if the riskier gamble is preferred under the higher risk aversion level, the same necessarily holds for the lower risk aversion level. As a consequence, the mass of errors for which the riskier gamble is preferred under more risk aversion is lower, and hence the probability of choosing it is also lower, as desired.

The paper is organized as follows. Section 2 briefly reviews the related literature, while Section 3 establishes the notation and gives the basic definitions. In Sections 4

[2]Theorems 7 and 8 in Appendix A establish the corresponding analogous results for the use of the logarithmic transformation of expected utility, which is often used.

and 5, we present the results regarding random utility models and random preference models, respectively. Section 6 concludes. All the proofs are contained in an appendix.

2. Related Literature

Discrete choice models in general settings, not necessarily involving risk, 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 risky 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. Papers using errors on the utility include Cicchetti and Dubin (1994), Hey and Orme (1994), Harrison, List and Towe (2007), Post et al. (2008), Toubia et al. (2013), and Noussair, Trautmann and van de Kuilen (2014). For papers using errors on the log of the utility, see Donkers, Melenberg, and van Soest (2001), Holt and Laury (2002) and Andersen et al. (2008). Finally, Friedman (1974), Hey, Morone and Schmidt (2009), Coble and Lusk (2010), and von Gaudecker, van Soest and Wengstrom (2011) use errors on the certainty equivalent.

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).

Wilcox (2008, 2011) first inquires into the problem analyzed in this paper. Using examples, with the logistic error structure over utilities or over the logarithmic transformation of utilities, and for comparisons of gambles involving mean-preserving spreads, he discusses how CRRA and CARA 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 case of three outcomes. In Subsection 4.3, we further discuss his proposal, and show that the problem persists beyond these comparisons. In addition, Blavatskyy (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. Our Theorem 1 completely characterizes when these anomalies occur.

3. Preliminaries

A gamble $g = [\mathbf{p}, \mathbf{x}]$ is described by a finite vector of probabilities $\mathbf{p} = (p_1, \ldots, p_N)$, with $p_i > 0$ and $\sum_i p_i = 1$, over a finite vector of monetary outcomes $\mathbf{x} = (x_1, \ldots, x_N)$, with $x_i \in X$. Unless explicitly noted, we assume that $X = \mathbb{R}$. We denote by δ_x the degenerate gamble yielding x with certainty.

Let U be a utility function over gambles. A certainty equivalent of a gamble g for U, is a monetary payoff $CE(g, U)$ such that $U(g) = U(\delta_{CE(g,U)})$. The utility function U is an expected utility function if there exists a utility function over outcomes $u : X \to \mathbb{R}$, that is strictly increasing and continuous, such that for every gamble $g = [\mathbf{p}, \mathbf{x}]$, $U(g) = \sum_i p_i u(x_i)$. It is well-known that for expected utility functions every gamble has a unique certainty equivalent, and also that U_1 is (strictly) more risk averse than U_2 if and only if u_1 is a (strict) concave transformation of u_2.

Throughout the paper we study the attitude of individuals towards pairs of gambles where one is riskier than the other. The standard textbook comparison considers a risky gamble g and a certain monetary payoff $h = \delta_x$. Another widely accepted comparison involves a mean-preserving spread g of a gamble h. Gamble g is a mean-preserving spread of gamble h through outcome y and gamble g' if g can be expressed as a compound gamble that replaces outcome y of h with gamble g', which has y as its expected value. We say that g is a mean-preserving spread of h if there is a sequence of such spreads from h to g. In both cases, gamble g is clearly riskier than gamble h.

The constant absolute risk aversion (CARA) and constant relative risk aversion (CRRA) families of monetary utility functions are by far the most widely used specifications in applications. These functional forms have the useful property of being ordered in terms of risk aversion by a single parameter r. Thus, a (strictly) larger value of r implies a (strictly) greater level of risk aversion. The following are standard definitions. CARA utility functions are defined by $u^r_{CARA}(x) = \frac{1-e^{-rx}}{r}$ for $r \neq 0$, and $u^0_{CARA}(x) = x$. CRRA utility functions are defined by $u^r_{CRRA}(x) = \frac{x^{1-r}}{1-r}$ for $r \neq 1$, and $u^1_{CRRA}(x) = \log x$. For ease of exposition, assume that $x \geq 0$ for the case of CARA and $x \geq 1$ for the case of CRRA.[3] We write U^r_{CARA} and U^r_{CRRA} when referring to either the CARA or CRRA expected utility families, and U^r_ω when referring to both.

[3]These assumptions simplify the analysis of CARA and CRRA, as explained in footnote 14.

4. Random Utility Models

The first risk aversion estimation method to be analyzed here represents a natural application of the standard techniques in the micro-econometrics of discrete choice analysis, and, as such, has been widely used in the literature. In a nutshell, the valuation of an option comprises two parts, one that it is assumed to be observed and is modeled using a utility function from a pre-specified family of utility functions, and another unobserved random part often referred to as errors. It is assumed that the individual selects the option that provides the greatest total value. Thus, each option is selected with a certain probability, which can be expressed as a function of the utility difference between the two. Finally, standard maximum likelihood techniques are used to select a specific utility function from the available family. We refer to these as random utility models.

4.1. **Expected Utility Differences.** This is by far the most widely used specification. Formally, consider an expected utility U.[4] The valuation of gamble g is given by the additive consideration of $U(g)$ and a random i.i.d. unobserved term $\epsilon(g)$, which follows a continuous cumulative distribution Ψ. The probability of selecting gamble g over gamble h is given by $f^U_{\Psi[\lambda]}(g,h) = P(\lambda U(g) + \epsilon(g) \geq \lambda U(h) + \epsilon(h)) = P(\epsilon(h) - \epsilon(g) \leq \lambda(U(g) - U(h))) = \Psi^*(\lambda(U(g) - U(h)))$, where Ψ^* is the distribution function of the difference of errors, and hence has mean zero.[5] By far the most widely used probability distributions of the unobserved terms are the extreme type I and the normal, which involve the logistic model (also known as the Luce model) and the probit model, respectively. The former has a closed-form expression for the probability of selecting gamble g over gamble h equal to $\frac{e^{\lambda U(g)}}{e^{\lambda U(g)} + e^{\lambda U(h)}}$.

We now show that this approach has fundamental flaws. Consider any pair of expected utility functions U_1 and U_2, where U_1 is more risk averse than U_2, and any error distribution. We prove that there are pairs of gambles, g and h, where either $h = \delta_x$, or g is a mean-preserving spread of h, for which the more risk averse utility U_1 chooses the riskier gamble with a larger probability than the less risk averse utility U_2. Theorem 1 characterizes the pairs of gambles for which the anomalies occur, and shows that they

[4]In Remark 1, we discuss the related and influential case where the log of expected utility is used, and in Section 4.3 we discuss generalizations of expected utility.

[5]Parameter λ is inversely related to 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.

form a large set.[6] It does so by using the function $U_{1-2} = U_1 - U_2$, which expresses how differently the two utility functions value the gambles.[7]

Theorem 1. *Consider any two expected utility functions U_1, U_2 such that U_1 is strictly more risk averse than U_2. Consider any continuous distribution Ψ, and any $\lambda > 0$. Every gamble has either one or two certainty equivalents for U_{1-2}, with at least one belonging to the interval of possible outcomes of the gamble. Further, for any pair of gambles g and h such that g is non-degenerate and $h = \delta_x$ (resp. g is a mean-preserving spread of h through outcome y and non-degenerate gamble g'), $f^{U_1}_{\Psi[\lambda]}(g,h) > f^{U_2}_{\Psi[\lambda]}(g,h)$ if and only if:*

- *U_{1-2} is increasing at the unique certainty equivalent of g (resp. g') and $x < CE(g, U_{1-2})$ (resp. $y < CE(g', U_{1-2})$), or*
- *U_{1-2} is decreasing at the unique certainty equivalent of g (resp. g') and $x > CE(g, U_{1-2})$ (resp. $y > CE(g', U_{1-2})$), or*
- *U_{1-2} has two certainty equivalents for g (resp. g') and x is not in $[CE(g, U_{1-2}), CE^*(g, U_{1-2})]$ (resp. y is not in $[CE(g', U_{1-2}), CE^*(g', U_{1-2})]$).*

There are several points worth stressing in relation to Theorem 1. First and foremost, the theorem establishes that, for every error distribution, problems arise in every pair of expected utility functions where one is strictly more risk averse than the other. To see this, notice that, in the case involving degenerate gambles, the result shows that problems arise for any possible gamble g, and for a large range of outcomes x which intersects with the range of monetary payoffs of the gamble g. In the case of the mean-preserving spreads, notice that not every pair of utility functions is affected. Consider, for example, the case of a risk averse utility U_1 and a risk-neutral utility U_2. Clearly, U_2 values the two gambles equally, and hence assigns equal choice probabilities to them both, whereas U_1 dislikes the mean-preserving spread gamble g, and thus selects it with less probability, as desired. Theorem 1 characterizes when the problem arises, namely whenever the gamble g' is better than the certainty of obtaining the expected value of g', namely y, according to U_{1-2}. It is immediate, then, that for every risk averse utility

[6]In the comparisons involving mean-preserving spreads, Theorem 1 characterizes the case in which g is directly obtained from h through outcome y and gamble g'. Obviously, indirect spreads can be characterized by applying the theorem iteratively.

[7]For ease of exposition, we say that U_{1-2} is increasing (resp. decreasing) when evaluated at degenerate gambles, that is, whenever the corresponding $u_{1-2} = u_1 - u_2$ is increasing (resp. decreasing).

function U_2, and for every gamble g', one can find a more risk averse utility function U_1, such that $U_{1-2}(g') > U_{1-2}(y)$.

Let us now provide the intuition for the negative results involving every pair of utility functions. As per the definition of random utility models, U_1 selects the riskier gamble g less often if $U_1(g) - U_1(h) < U_2(g) - U_2(h)$, which is equivalent to $U_{1-2}(g) < U_{1-2}(h)$. Given g, linearity of U_{1-2} makes the latter inequality hold for every h only if the function is constant, which is incompatible with U_1 being strictly more risk averse than U_2. Theorem 1 builds on this idea to characterize the ranges within which the problem arises.

To further illustrate the extent of the identification problem, we now consider two well-known parametric families of utility functions, CARA and CRRA, and study, for every g and h and every probability distribution of errors, the properties of the probability of choosing the riskier gamble g, as the parameter of risk aversion increases. Clearly, the probability should decrease with risk aversion, but we now show that this is not in fact the case. In particular, we show that there is always a level of risk aversion above which the probability of choosing the riskier gamble increases. This implies that there is an upper limit in the possible estimation of the risk aversion coefficient, however risk averse the individual may be. Moreover, different levels of risk aversion may be compatible with the same behavior, thus also creating an identification problem.

Theorem 2. *Consider any continuous distribution Ψ, and any $\lambda > 0$. For every pair of distinct gambles, g and h, where either h is degenerate with monetary outcome strictly within the interval of possible outcomes of g, or g is a mean-preserving spread of h, there exists r^* such that $f^{U^{r_1}_\omega}_{\Psi[\lambda]}(g,h) > f^{U^{r_2}_\omega}_{\Psi[\lambda]}(g,h)$ whenever $r_1 > r_2 \geq r^*$.*

The main step of the proof of Theorem 2 builds on the following observation. Given a pair of gambles, when the individual is sufficiently risk averse, the curvature of the utility function makes the utilities of the possible outcomes associated with the gambles become sufficiently close. Hence, the utility difference between the gambles may be reduced to the point that the individual is less able to discriminate between the two gambles. Thus, in a random utility model, the probability of selecting the riskier gamble g starts to increase.

Theorem 2 is best illustrated in Figure 1, which uses CARA expected utility and the logistic model with $\lambda = 2$. Figure 1a depicts the probability of choosing gamble $g = [(1, 21), (.9, .1)]$ over the degenerate gamble δ_2 for different values of the risk aversion coefficient r, and Figure 1b the probability of choosing the mean-preserving

Figure 1. Differences in Expected Utility

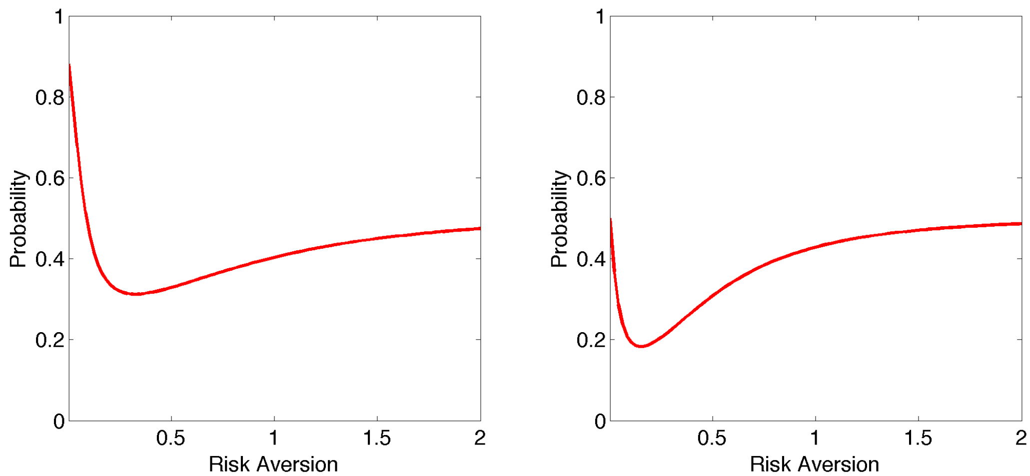

Figure 1a: g versus δ_2 Figure 1b: g versus h

spread gamble g over $h = [(1, 5), (.5, .5)]$. It can be appreciated how this probability decreases with the level of risk aversion as far as risk aversion levels .32 and .15, respectively, after which it starts increasing. 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 .32 and .15, approximately. Secondly, note that there are two risk averse coefficients consistent with choice probabilities in the intervals $(.33, .5)$ and $(.19, .5)$, which generates an identification problem. Theorem 2 shows that, for any given error distribution, problems of this sort arise for every pair of gambles.

Remark 1. The literature contains influential papers in which a logarithmic transformation of the utilities is used instead of actual utility values. Thus, the probability of selecting one gamble over the other depends on the utility ratio, rather than the utility difference. The results established in this section can basically be replicated for the case of the utility ratio. See Appendix A for details.

4.2. **Certainty Equivalent Differences.** In this approach, the risk preferences estimation process is analogous to that studied in the previous section, except that gambles are evaluated in terms of their certainty equivalents, instead of their expected utility. The main intuition for this approach is that the certainty equivalent is a monetary representation of preferences, which facilitates interpersonal comparisons by using a common scale.

Formally, given a utility function U, the evaluation of a gamble g is now given by the additive consideration of its certainty equivalent $CE(g,U)$, and a random i.i.d. unobserved term $\epsilon(g)$, which follows a continuous cumulative distribution Ψ. The probability of selecting gamble g over gamble h is then given by $f^{CE(\cdot,U)}_{\Psi[\lambda]}(g,h) = P(\lambda CE(g,U) + \epsilon(g) \geq \lambda CE(h,U) + \epsilon(h)) = \Psi^*(\lambda(CE(g,U) - CE(h,U)))$, where Ψ^* is the distribution function of the difference of errors. In this case, the probability of selecting gamble g over gamble h for the logistic model is $\frac{e^{\lambda CE(g,U)}}{e^{\lambda CE(g,U)}+e^{\lambda CE(h,U)}}$.

One may entertain the idea that, by creating a common scale, this method could provide a solution to the problem discussed in the previous section. We show that, in some instances, this is in fact the case. Take any two utility functions, where one is more risk averse than the other, a gamble g, and a degenerate gamble δ_x. Clearly, the certainty equivalent of the degenerate gamble is the monetary outcome x for both utility functions, whereas the certainty equivalent of gamble g is obviously lower for the risk averse one. Thus, the difference between the certainty equivalents of the two gambles is greater for the more risk averse utility function, and consequently, its probability of choosing gamble g is lower. This intuition is formally established in the following result.

Theorem 3. *Consider any two expected utility functions U_1, U_2 such that U_1 is more risk averse than U_2. Consider any continuous distribution Ψ, and any $\lambda > 0$. For any non-degenerate gamble g and for any degenerate gamble δ_x, $f^{CE(\cdot,U_1)}_{\Psi[\lambda]}(g,\delta_x) \leq f^{CE(\cdot,U_2)}_{\Psi[\lambda]}(g,\delta_x)$.*

We now show that, unfortunately, this method of estimation may be problematic when non-degenerate gambles are involved. The following results establish that some pairs of gambles related by the mean-preserving spread notion suffer from the sort of anomalies identified in Theorems 1 and 2.

Theorem 4. *Consider any two expected utility functions U_1, U_2 such that U_1 is strictly more risk averse than U_2 and $\lim_{x\to\infty} \frac{U_i(\delta_x)}{x} = U_i(\delta_0) = 0$, $i = 1, 2$. Consider any continuous distribution Ψ, and any $\lambda > 0$. There exist gambles, g, and h, such that g is a mean-preserving spread of h through outcome y and non-degenerate gamble g' with $f^{CE(\cdot,U_1)}_{\Psi[\lambda]}(g,h) > f^{CE(\cdot,U_2)}_{\Psi[\lambda]}(g,h)$.*

The property $\lim_{x\to\infty} \frac{U_i(\delta_x)}{x} = U_i(\delta_0) = 0$ simply imposes that, as x grows, the utility of gamble $[(\frac{x-1}{x}, \frac{1}{x}), (0, x)]$ approaches the utility of obtaining a certain 0 payoff, which is normalized to 0. For the case of risk averse individuals, this is very intuitive since the

FIGURE 2. Differences in Certainty Equivalent

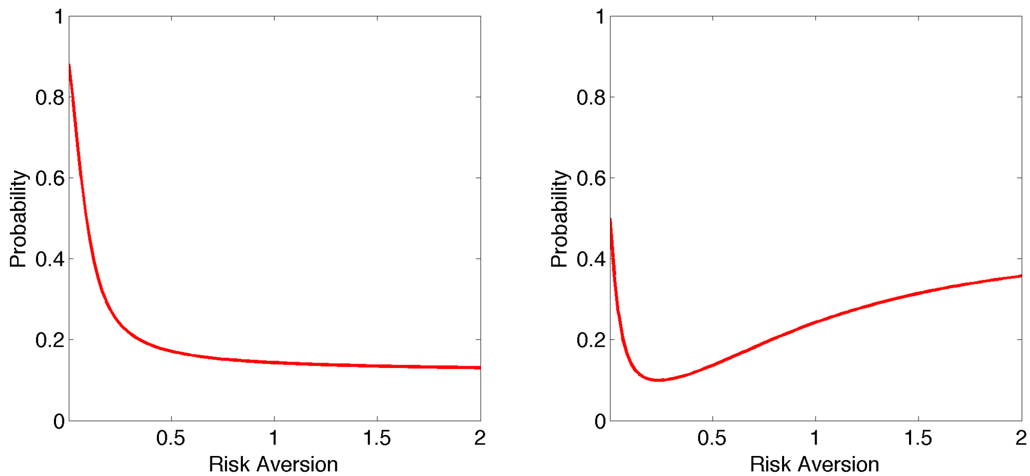

Figure 2a: g versus δ_2 Figure 2b: g versus h

gamble clearly becomes less and less attractive as x increases, and is satisfied by many expected utility functions, such as risk averse CARA and CRRA utility functions.

Theorem 4 establishes that problems may reappear when the anchor established by degenerate gambles is abandoned. Consider two non-degenerate gambles, where one is a mean-preserving spread of the other, and such that the minimum payoff is the same in both. Notice that, in this case, as the level of risk aversion increases, the certainty equivalents of both gambles approach that minimum payoff, and hence the difference goes to 0. The same argument used in the intuition of the above results then applies. Theorem 5 uses this type of gambles to establish the existence of a global minimum for CARA and CRRA.

Theorem 5. *Consider any continuous distribution Ψ, and any $\lambda > 0$. For every pair of distinct non-degenerate gambles, g and h, such that the minimum payoff from both gambles is the same, and g is a mean-preserving spread of h, there exists r^* such that $f_{\Psi[\lambda]}^{U_\omega^r}(g,h) \geq f_{\Psi[\lambda]}^{U_\omega^{r^*}}(g,h)$ whenever $r \geq r^*$.*

Figure 2 uses the same utility functions, probability distribution, and gambles as those used in Figure 1, but now using differences in the certainty equivalents. It is apparent that the probability of choosing g over δ_2 is now well-behaved, whereas the problem persists with respect to the comparison between g and h.

To conclude, we have shown that the use of certainty equivalents in a random utility model may involve problems, but their extent crucially depends on the specific gambles being considered.

4.3. **Other Related Models.** The previous two sections work under the assumption of expected utility. Clearly, any generalization of expected utility, such as cumulative prospect theory, rank-dependent expected utility, disappointment aversion, etc, are susceptible to the problems identified above, as they include expected utility as a special case. Not only this, but the additive nature of these models makes them vulnerable to the same anomaly even when only non-expected utilities are considered. Furthermore, they also share the positive results identified in Theorem 3, involving the certainty equivalent representation in conjunction with comparisons using a degenerate gamble. This is so because these utility theories also have a unique certainty equivalent for every gamble, and hence exactly the same arguments apply as in Theorem 3.

Wilcox (2011) suggests a novel utility representation, namely, contextual utility. When two gambles are defined over the same outcomes, contextual utility normalizes the utility difference of the gambles by the difference between the utility of the best and worst of these outcomes. This normalization solves the problem for cases in which the two gambles are defined over the same three outcomes, and for the mean-preserving spread comparisons. Problems in line with those identified so far arise when using pairs of gambles where one is degenerate or where the gambles involve more than three outcomes with the mean-preserving spread notion.[8]

We close this section by studying 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) are examples of the use of mean-variance utilities in a random utility model, for the estimation of risk preferences. Formally, given a gamble g, denote the expected value and variance of g by $\mu(g)=\sum_i p_i x_i$ and $\sigma^2(g)=\sum_i p_i(x_i-\mu(g))^2$, respectively. A utility function U is a mean-variance utility if $U(g)=\mu(g)-r\sigma^2(g)$,

[8]By using differences over expected utility with CRRA and risk aversion coefficients of $r_1=.9$ and $r_2=.5$, the following gambles illustrate these points: $g=[(0,5,10,15,20),(.1,.3,.2,.3,.1)]$ and δ_{15}, and $g=[(0,5,10,15,20),(.1,.3,.2,.3,.1)]$ and $h=[(0,5,10,15,20),(.1,.2,.4,.2,.1)]$. If one is willing to compare gambles dominated by stochastic dominance, something that Wilcox discarded, then $g'=[(5,10,15),(.3,.4,.3)]$ and δ_{15} is an example of the problems that may arise using the standard notion of risk aversion in the classical three outcomes world.

where r is the risk aversion parameter. We now argue that the use of random utility models in conjunction with mean-variance utility is fortunately free from any of the problems identified so far. First consider a gamble g, and a degenerate gamble δ_x. Clearly, the more risk averse utility is the one that more intensely dislikes gamble g, and hence $U_{1-2}(g) < 0 = U_{1-2}(\delta_x)$ which, as we have argued in relation to Theorem 1, guarantees that the more risk averse utility selects gamble g with less probability. Moreover, when g is a mean-preserving spread of h, simply notice that g has more variance than h and since $r_1 > r_2$, $U_{1-2}(g) = (r_2 - r_1)\sigma^2(g) < (r_2 - r_1)\sigma^2(h) = U_{1-2}(h)$, again as desired.

5. Random Preference Models

In random preference models, the individual is assumed to have a probability distribution over a set of utility functions, of which certain parameters are unknown to the researcher. At the moment of choice, one of the utilities is drawn from the distribution and the individual ends up selecting the best option according to that utility function. This process generates a probability distribution over the choice set and standard maximum likelihood techniques can then be used to determine the unknown parameters.

5.1. **Errors on the Risk Aversion Parameter.** Consider a parametric family of expected utility functions $\{U^r\}$, such that (strictly) larger values of r are associated with (strictly) greater risk aversion. The individual is assumed to have a given risk aversion level that is subject to an error μ, with μ following a symmetric continuous cumulative distribution ϕ with mean zero. As a consequence, the probability of selecting gamble g over gamble h is given by $f^{U^r}_{\phi[\lambda]}(g,h) = P(U^{r+\mu}(g) \geq U^{r+\mu}(h))$, where λ is inversely related to the variance of the error distribution, representing the rationality parameter.[9] To illustrate, whenever ϕ is the logistic, the closed-form probability of selecting gamble g over h is $\frac{1}{1+e^{-\lambda(\hat{r}(g,h)-r)}}$, where $\hat{r}(g,h)$ is the parameter that makes $U^{\hat{r}(g,h)}(g) = U^{\hat{r}(g,h)}(h)$, whenever it exists.[10]

---

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

[10]The proof of Theorem 6 analyzes the conditions required for such a risk aversion parameter to exist.

Figure 3. Random Preference Models

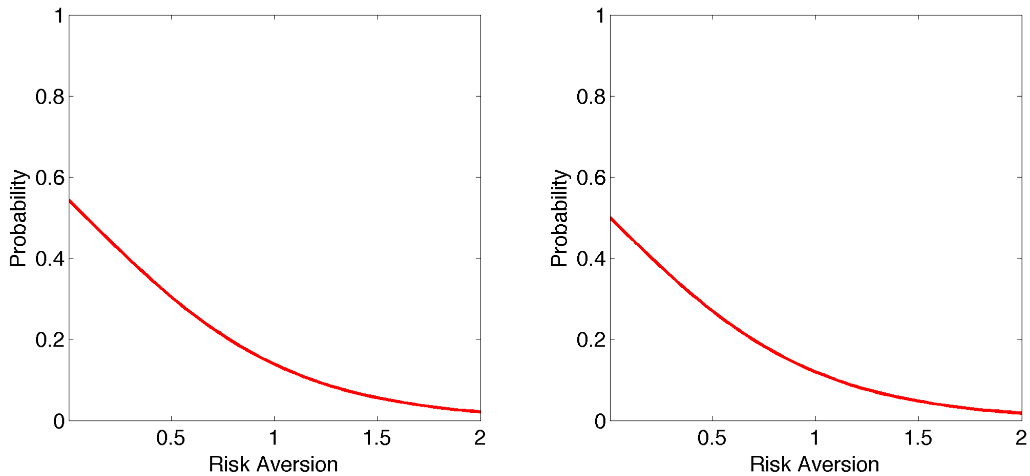

Figure 3a: g versus δ_2 Figure 3b: g versus h

The following result establishes that this method is free from any of the problems identified in the previous sections.

Theorem 6. *Consider the parametric family of expected utility functions $\{U^r\}$, and any two utilities such that $r_1 \geq r_2$. Consider any continuous symmetric distribution ϕ with mean zero, and any $\lambda > 0$. For any pair of gambles, g and h, such that g is non-degenerate and $h = \delta_x$, or g is a mean-preserving spread of h, $f^{U^{r_1}}_{\phi[\lambda]}(g,h) \leq f^{U^{r_2}}_{\phi[\lambda]}(g,h)$.*

Figure 3 illustrates the same case as the previous figures, but the probability of choosing the riskier gamble is now modeled using the errors on the risk aversion parameter, as studied in this section. The figure exemplifies that the above-identified problems vanish when this method is used.

5.2. **Other Related Models.** The assumption on a family of utility functions with a single risk aversion parameter can be extended to the consideration of families with a vector $\mathbf{r}$ of parameters, where an utility is more risk averse than another if all the vector components are larger. In this case, the estimation exercise would involve consideration of independent errors on each of the parameters, but follows the same intuition otherwise. Thus, the parameter space would be divided into two disjoint areas, one in which gamble g is strictly better than gamble h, and another in which the reverse is the case. In line with the definition of $\hat{r}(g,h)$, the family of utility functions must be such that, whenever a vector $\mathbf{r}$ belongs to the former area, and $\mathbf{r} \geq \mathbf{r}'$, the vector $\mathbf{r}'$

also belongs to it. Similarly, whenever a vector $\mathbf{r}$ belongs to the latter area, and $\mathbf{r}' \geq \mathbf{r}$, the vector $\mathbf{r}'$ also belongs to it. Consequently, given two utility functions such that $\mathbf{r}_1 \geq \mathbf{r}_2$, the model implies that the probability of choosing the riskier gamble is lower for the first of these utilities.

It is important to notice that exactly the same logic applies beyond expected utility, provided that there exists the sort of separation threshold described above between the utility functions that prefer g over h and those that prefer the reverse. This is generally the case for rank-dependent expected utility, disappointment aversion or cumulative prospect theory. It is immediate from this that the positive news provided by Theorem 6 applies well beyond expected utility.

Another option is to take a non-parametric approach, by which, under the appropriate assumptions, one can consider the Taylor expansions of utility functions and work directly on the Arrow-Pratt coefficient of risk aversion. In this case, the consideration of errors over the Arrow-Pratt coefficient would lead to the same positive results provided by Theorem 6. This is basically the approach adopted in Cohen and Einav (2007).

Finally, one can entertain a very particular distribution over the set of utility functions, in which only a utility function U and the opposite utility $-U$ have positive mass. Then, in line with the trembling hand approach in game theory, the individual uses the utility function U with probability $1 - \epsilon$, and the utility function $-U$ with probability ϵ. See Harless and Camerer (1994) for an application of this model to the estimation of risk aversion. It is immediate to see that the ordinal ranking of gambles g and h is the only determinant of the choice probabilities, and hence a more risk averse individual will always select the riskier gamble with a lower probability, as desired.[11]

6. Discussion

We have focused on the issue of risk aversion estimation, finding that, in standard random utility models, individuals that are more risk averse may choose the riskier gamble with a larger probability. We note that this problem logically extends from the individual static problem analyzed here to more general strategic or dynamic settings. When the individual operates under conditions of strategic uncertainty, beliefs replace objective probabilities. A prominent example of this approach in game theory

[11]This specification assumes that there are only two gambles. With more gambles, more utilities have to be taken into consideration, but the conclusion would be the same.

is the quantal response equilibrium of McKelvey and Palfrey (1995), which assumes independent errors over expected utility. Hence, for given beliefs and according to our results, it may be the case that the more risk averse individual may choose the riskier action with a larger probability. 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. Some of these settings involve risk and are modeled by means of random utility models with errors over expected utility. In this line, see, for instance, Rust and Phelan (1997) for a retirement model, and Crawford and Shum (2005) for one involving experience goods markets. Hence, our paper shows that it may occur, for example, that the more risk-averse individual selects the riskier retirement or drug option. Finally, we have a companion paper where we identify analogous problems in the estimation of time preferences (Apesteguia and Ballester, 2014).

Appendix A. Expected Utility Ratios

This approach starts by assuming that the utility of every gamble is strictly positive. The probability of selecting gamble g over gamble h is $f^{\log(U)}_{\Psi[\lambda]}(g,h) = P(\lambda \log(U(g)) + \epsilon(g) \geq \lambda \log(U(h)) + \epsilon(h)) = \Psi^*(\lambda \log(U(g)) - \lambda \log(U(h))) = \Psi^*(\lambda \log(\frac{U(g)}{U(h)}))$, where Ψ^* is the distribution function of the difference of errors. In this case, the logistic model gives the closed-form probability of selecting gamble g over gamble h as being equal to $\frac{U(g)^\lambda}{U(g)^\lambda + U(h)^\lambda}$.

Theorem 7 reproduces the results of Theorem 1. Given the logarithmic transformation, it uses the ratio of utilities $U_{1/2} = \frac{U_1}{U_2}$, instead of the difference U_{1-2}.

Theorem 7. *Consider any two expected utility functions U_1, U_2 such that U_1 is strictly more risk averse than U_2. Consider any distribution Ψ, and any $\lambda > 0$. Every gamble has either one or two certainty equivalents for $U_{1/2}$, with at least one belonging to the interval of possible outcomes of the gamble. Further, let us consider that, for any pair of gambles, g and h, such that g is non-degenerate and $h = \delta_x$ (resp. g is a mean-preserving spread of h through outcome y and non-degenerate gamble g' with $U_2(\delta_y) \geq U_2(g')$ and $\frac{U_1(\delta_y)}{U_2(\delta_y)} \geq \frac{U_1(h)}{U_2(h)}$)), $f^{U_1}_{\Psi[\lambda]}(g,h) > f^{U_2}_{\Psi[\lambda]}(g,h)$ if and only if (resp. whenever):*

- *$U_{1/2}$ is increasing at the unique certainty equivalent of g (resp. g') and $x < CE(g, U_{1/2})$ (resp. $y < CE(g', U_{1/2})$), or*

FIGURE 4. Differences in Log Expected Utility

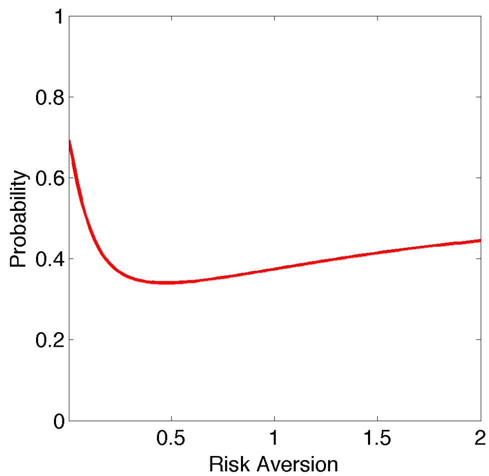

Figure 4a: g versus δ_2

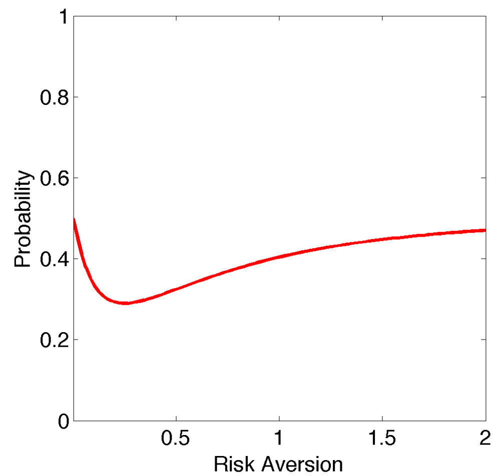

Figure 4b: g versus h

- *$U_{1/2}$ is decreasing at the unique certainty equivalent of g (resp. g') and $x > CE(g, U_{1/2})$ (resp. $y > CE(g', U_{1/2})$), or*
- *$U_{1/2}$ has two certainty equivalents for g (resp. g') and x is not in $[CE(g, U_{1/2}), CE^*(g, U_{1/2})]$ (resp. y is not in $[CE(g', U_{1/2}), CE^*(g', U_{1/2})]$).*

In the next result we reproduce the parametric anomaly for the CARA case. Importantly, notice that CRRA functions are not entirely suitable in this context. This is because for values of r above 1, utilities become negative, and this is incompatible with the use of log-transformations.[12]

Theorem 8. *Consider any continuous distribution Ψ, and any $\lambda > 0$. For every pair of distinct gambles, g and h, where either h is degenerate with monetary outcome strictly within the interval of outcomes of g, or g is a mean-preserving spread of h, there exists r^* such that $f_{\Psi[\lambda]}^{\log(U_{CARA}^{r})}(g, h) \geq f_{\Psi[\lambda]}^{\log(U_{CARA}^{r^*})}(g, h)$ whenever $r \geq r^*$.*

Figure 4 illustrates the same case as the previous figures for the logarithmic transformation.

APPENDIX B. PROOFS

Proof of Theorem 1: We divide the proof into 4 steps.

[12]Notice that the function x^{1-r}, without the normalization $\frac{1}{1-r}$, is positive for values of $r > 1$, but is not in this case monotone in outcomes and thus, is also problematic.

Step 1. We start by proving that every gamble has either one or two certainty equivalents for U_{1-2}, at least one of which has a value in the interval between the smallest and largest monetary outcomes of the gamble. First, the continuity of u_1 and u_2 guarantees that the function $u_{1-2} = u_1 - u_2$ is also continuous. Second, we show that u_{1-2} is strictly quasi-concave. To see this, consider three distinct outcomes $x < y < z$. By the continuity of u_2, there exists $p \in (0, 1)$ such that $u_2(y) = pu_2(x) + (1 - p)u_2(z)$. Given that U_1 is strictly more risk averse than U_2, it must be that $u_1(y) > pu_1(x) + (1 - p)u_1(z)$. Hence, $u_{1-2}(y) > pu_{1-2}(x) + (1 - p)u_{1-2}(z) \geq \min\{u_{1-2}(x), u_{1-2}(z)\}$, showing that u_{1-2} is strictly quasi-concave.

Now consider any gamble $g = [\mathbf{p}, \mathbf{x}]$. To establish the existence of at least one certainty equivalent within the interval of the monetary outcomes of g, notice that $U_{1-2}(g) = \sum_i p_i u_{1-2}(x_i) \in [\min_i\{u_{1-2}(x_i)\}, \max_i\{u_{1-2}(x_i)\}]$. Continuity guarantees that u_{1-2} will achieve the value $U_{1-2}(g)$ for at least one monetary outcome between the outcomes $\arg\min_i\{u_{1-2}(x_i)\}$ and $\arg\max_i\{u_{1-2}(x_i)\}$. Clearly, this monetary payoff defines a certainty equivalent, that trivially belongs to the interval $[\min_i\{x_i\}, \max_i\{x_i\}]$. Finally, notice that the strict quasi-concavity of u_{1-2} guarantees that there are at most two different certainty equivalents of g.

Step 2. We prove that for any two gambles g and h, $f^{U_1}_{\Psi[\lambda]}(g, h) > f^{U_2}_{\Psi[\lambda]}(g, h)$ if and only if $U_{1-2}(g) > U_{1-2}(h)$. To see this, simply notice that $f^{U_1}_{\Psi[\lambda]}(g, h) > f^{U_2}_{\Psi[\lambda]}(g, h)$ if and only if $\Psi^*(\lambda(U_1(g) - U_1(h))) > \Psi^*(\lambda(U_2(g) - U_2(h)))$ if and only if $U_1(g) - U_1(h) > U_2(g) - U_2(h)$ if and only if $U_{1-2}(g) > U_{1-2}(h)$.

Step 3. We now conclude the proof for pairs of gambles g and h, where g is non-degenerate and $h = \delta_x$. By step 1, g has either one or two certainty equivalents for U_{1-2}. If there is only one certainty equivalent $CE(g, U_{1-2})$, notice that the non-degeneracy of g precludes its being the maximizer of u_{1-2}. Since u_{1-2} is strictly quasi-concave, it can only be strictly increasing or strictly decreasing at $CE(g, U_{1-2})$. If it is increasing (resp. decreasing), clearly $x < (>)\ CE(g, U_{1-2})$ if and only if $U_{1-2}(g) > U_{1-2}(\delta_x)$. By step 2, the latter holds if and only if $f^{U_1}_{\Psi[\lambda]}(g, \delta_x) > f^{U_2}_{\Psi[\lambda]}(g, \delta_x)$. If there are two certainty equivalents, $CE(g, U_{1-2})$ and $CE^*(g, U_{1-2})$, strict quasi-concavity guarantees that $x \notin [CE(g, U_{1-2}), CE^*(g, U_{1-2})]$ if and only if $U_{1-2}(g) > U_{1-2}(\delta_x)$ and again, by step 2, the result follows.

Step 4. Finally, we prove the case in which g is a mean-preserving spread of h through outcome y and the non-degenerate gamble g'. Using the additive nature of expected utility and the assumptions on the gambles, it is $U_{1-2}(g) > U_{1-2}(h)$ if and only if

$pU_{1-2}(g') > pU_{1-2}(\delta_y)$, where p is the probability of gamble g' and outcome y in gambles g and h respectively. The same reasoning used in step 3 applies and the result follows.$\blacksquare$

Proof of Theorem 2: The proof is divided into three steps.

Step 1. The marginal analogue of u_{1-2} is $v^r_\omega = \frac{\partial u^r_\omega}{\partial r}$, which is $\frac{e^{-rx}(1+rx)-1}{r^2}$ in the case of CARA and $\frac{x^{1-r}(1-(1-r)\log x)}{(1-r)^2}$ in the case of CRRA.[13] These functions are continuous and strictly decreasing in monetary payoffs, given the domain assumptions on the outcomes. Let $-V^r_\omega$ be the expected utility associated with the continuous and strictly increasing utility function $-v^r_\omega$. It follows immediately from standard arguments that $CE(g, -V^r_\omega) = CE(g, V^r_\omega)$ always exists and is unique.[14] We can now compute the Arrow-Pratt coefficient of risk aversion for $-v^r_\omega$. This is $\frac{rx-1}{x}$ for CARA and $\frac{r\log x-1}{x\log x}$ for CRRA. These coefficients have a strictly positive derivative with respect to r and thus, from the classic result of Pratt (1964), it follows that, for any non-degenerate gamble g, the derivative of $CE(g, -V^r_\omega) = CE(g, V^r_\omega)$ with respect to r is strictly negative.[15]

Step 2. We now prove the degenerate gamble case. Let $g = [\mathbf{p}, \mathbf{x}]$ and $h = \delta_x$ such that $\max_i\{x_i\} > x > \min_i\{x_i\}$. Consider the gamble $\hat{g}$ that assigns exactly the same probability p to outcome $\min_i\{x_i\}$ in g, as assigned by g, and assigns to outcome $\max_i\{x_i\}$ in g the probability $1-p$. Since $\hat{g}$ first order stochastically dominates g, it is the case that $-V^r_\omega(\hat{g}) \geq -V^r_\omega(g)$. We now show that, for sufficiently large $r > 0$, it is the case that $-V^r_\omega(\delta_x) > -V^r_\omega(\hat{g})$. To see this, notice that $-V^r_\omega(\delta_x) > -V^r_\omega(\hat{g})$ if and only if $-v^r_\omega(x) > -pv^r_\omega(\min_i\{x_i\}) - (1-p)v^r_\omega(\max_i\{x_i\})$, which is equivalent to $p[-v^r_\omega(x) + v^r_\omega(\min_i\{x_i\})] > (1-p)[-v^r_\omega(\max_i\{x_i\}) + v^r_\omega(x)]$. Since $[-v^r_\omega(x) + v^r_\omega(\min_i\{x_i\})]$ is positive, this is $\frac{p}{1-p} > \frac{v^r_\omega(x) - v^r_\omega(\max_i\{x_i\})}{v^r_\omega(\min_i\{x_i\}) - v^r_\omega(x)}$. Since the curvature of v^r_ω can be made as large as desired by making r sufficiently large, the right hand side goes to zero as r increases and hence, we can find r^* sufficiently large that the inequality holds. Hence, $-V^{r^*}_\omega(\delta_x) > -V^{r^*}_\omega(\hat{g}) \geq -V^{r^*}_\omega(g)$, and therefore $V^{r^*}_\omega(\delta_x) < V^{r^*}_\omega(g)$ or equivalently, $CE(g, V^{r^*}_\omega) < x$. Given the decreasing nature of V in outcomes, this matches the

[13]These derivatives are well-defined for $r \neq 0$ in the case of CARA, and $r \neq 1$ in the case of CRRA.

[14] Notice that the decreasing nature of v^r_ω, provided by the domain assumptions, guarantees the uniqueness of this certainty equivalent. Without these assumptions, we would need to perform a separate analysis of the remaining cases in which two certainty equivalents exist.

[15]Notice that we cannot say that $CE(g, V^r_\omega)$ is strictly decreasing with r, since this function can be discontinuous at $r = 0$ for CARA or $r = 1$ for CRRA. We can, however, say that $CE(g, V^r_\omega)$ is strictly decreasing both whenever $r < 0$ and whenever $r > 0$.

second case in Theorem 1. Also, for $r \geq r^* > 0$, given that the certainty equivalent is strictly decreasing in r, we can apply Theorem 1 locally at every such level of risk aversion and hence, the probability of selecting g over δ_x is strictly increasing above r^*.

Step 3. We now prove the mean-preserving spread case. If g is a mean-preserving spread of h through outcome y and gamble g', we can apply the argument in step 4 of Theorem 1 and apply the previous step over gambles g' and δ_y, obtaining a value $r^\diamond$ at which $V_\omega^{r^\diamond}(\delta_y) < V_\omega^{r^\diamond}(g')$. If g is a mean-preserving spread of h, we can iteratively repeat the process to reach h. We then only need to consider r^* as the maximum of all the corresponding values $r^\diamond$ in order to obtain the desired result.■

Proof of Theorem 3: We can reproduce step 2 in the proof of Theorem 1 to conclude that $f_{\Psi[\lambda]}^{CE(\cdot),U_1}(g, \delta_x) \leq f_{\Psi[\lambda]}^{CE(\cdot,U_2)}(g, \delta_x)$ if and only if $\Psi^*(\lambda(CE(g, U_1) - CE(h, U_1))) \leq \Psi^*(\lambda(CE(g, U_2) - CE(h, U_2)))$ if and only if $CE(g, U_1) - CE(g, U_2) \leq CE(h, U_1) - CE(h, U_2)$. Now, consider a non-degenerate gamble g and a degenerate gamble δ_x. If U_1 is more risk averse than U_2, it follows that $CE(g, U_1) - CE(g, U_2) \leq 0 = x - x = CE(\delta_x, U_1) - CE(\delta_x, U_2)$, which concludes the proof.■

Proof of Theorem 4: Consider, for any strictly positive integer n, the gamble $g_n = [(1 - \frac{1}{2n}, \frac{1}{2n}), (0, 2n)]$. Since U_1 is strictly more risk averse than U_2, we know that $CE(g_1, U_2) - CE(g_1, U_1) > 0$. Since $\lim_{x\to\infty} \frac{U_i(\delta_x)}{x} = U_i(\delta_0) = 0$, it is the case that $\lim_{n\to\infty} CE(g_n, U_i) = 0$, and hence, there exists $\bar{n} > 1$ such that $CE(g_{\bar{n}}, U_2) - CE(g_{\bar{n}}, U_1) < CE(g_1, U_2) - CE(g_1, U_1)$. Thus $CE(g_{\bar{n}}, U_1) - CE(g_1, U_1) > CE(g_{\bar{n}}, U_2) - CE(g_1, U_2)$, which is equivalent to $f_{\Psi[\lambda]}^{CE(\cdot,U_1)}(g_{\bar{n}}, g_1) > f_{\Psi[\lambda]}^{CE(\cdot,U_2)}(g_{\bar{n}}, g_1)$. Simply notice that $g_{\bar{n}}$ is a mean-preserving spread of g_1 through outcome 2 and non-degenerate gamble $[(1 - \frac{1}{\bar{n}}, \frac{1}{\bar{n}}), (0, 2\bar{n})]$, which concludes the proof.■

Proof of Theorem 5: Consider two distinct non-degenerate gambles g and h such that x_m is the minimum outcome for both and g is a mean-preserving spread of h. Consider any $\hat{r} > 0$. Clearly, $CE(g, U_\omega^{\hat{r}}) < CE(h, U_\omega^{\hat{r}})$. Given that $\omega \in \{CARA, CRRA\}$ and the fact that both gambles are non-degenerate, we know that the limit of both $CE(g, U_\omega^r)$ and $CE(h, U_\omega^r)$ when r increases is x_m. Hence, there exists $\tilde{r} > \hat{r}$ such that $x_m < CE(g, U_\omega^r) < CE(h, U_\omega^r) < x_m + CE(h, U_\omega^{\hat{r}}) - CE(g, U_\omega^{\hat{r}})$, or equivalently, $CE(h, U_\omega^r) - CE(g, U_\omega^r) < CE(h, U_\omega^{\hat{r}}) - CE(g, U_\omega^{\hat{r}})$, for all $r \geq \tilde{r}$. This means

$f_{\Psi[\lambda]}^{CE(\cdot,U_\omega^r)}(g,h) > f_{\Psi[\lambda]}^{CE(\cdot,U_\omega^{\hat{r}})}(g,h)$ for all $r \geq \tilde{r}$. We can now define r^* as the value that minimizes the probability of choosing g over h in the closed interval $[\hat{r},\tilde{r}]$, and the result follows.■

Proof of Theorem 6: We start with the case where g is a non-degenerate gamble and $h = \delta_x$. Suppose first that $x \leq \lim_{r\to+\infty} CE(g,U^r)$. Then, for all μ, it is the case that $U^{r_i+\mu}(g) \geq U^{r_i+\mu}(\delta_x)$, and hence, $f_{\phi[\lambda]}^{U^{r_1}}(g,\delta_x) = f_{\phi[\lambda]}^{U^{r_2}}(g,\delta_x) = 1$. Similarly, if $x \geq \lim_{r\to-\infty} CE(g,U^r)$, it is the case that $f_{\phi[\lambda]}^{U^{r_1}}(g,\delta_x) = f_{\phi[\lambda]}^{U^{r_2}}(g,\delta_x) = 0$. Finally, consider the intermediate case where $x \in (\lim_{r\to-\infty} CE(g,U^r), \lim_{r\to+\infty} CE(g,U^r))$. Clearly, since the family of utility functions is assumed to be strictly ordered by the risk aversion parameter, there exists $\hat{r}(g,\delta_x)$ such that $CE(g,U^r) < (>)\ x$ whenever $r > (<)\ \hat{r}(g,\delta_x)$, and hence, since $r_1 \geq r_2$, $f_{\phi[\lambda]}^{U^{r_1}}(g,\delta_x) = \phi(\hat{r}(g,\delta_x) - r_1) \leq \phi(\hat{r}(g,\delta_x) - r_2) = f_{\phi[\lambda]}^{U^{r_2}}(g,\delta_x)$, as desired.

For the mean-preserving spread case, let g be a mean-preserving spread of h through outcome y (with probability p) and non-degenerate gamble g'. Then, for all μ, it is the case that $U^{r_i+\mu}(g) \geq U^{r_i+\mu}(h)$ if and only if $pU^{r_i+\mu}(g') \geq pU^{r_i+\mu}(\delta_y)$. Hence, we only need to apply the reasoning used in the previous paragraph to conclude that $f_{\phi[\lambda]}^{U^{r_1}}(g,h) = \phi(\hat{r}(g',\delta_y) - r_1) \leq \phi(\hat{r}(g',\delta_y) - r_2) = f_{\phi[\lambda]}^{U^{r_2}}(g,h)$, as desired. ■

Proof of Theorem 7: The proof follows analogous arguments to those in Theorem 1. To reproduce step 1, notice that $\frac{pu_1(x)+(1-p)u_1(z)}{pu_2(x)+(1-p)u_2(z)} = \frac{pu_2(x)}{pu_2(x)+(1-p)u_2(z)}u_{1/2}(x) + \frac{(1-p)u_2(z)}{pu_2(x)+(1-p)u_2(z)}u_{1/2}(z) \geq \min\{u_{1/2}(x), u_{1/2}(z)\}$ where $u_{1/2} = \frac{u_1}{u_2}$. Strict quasi-concavity of $u_{1/2}$ follows immediately. The strict positivity of u_2 can be used to prove continuity of this function and the rest follows. Now, to reproduce step 2, simply notice that $f_{\Psi[\lambda]}^{\log(U_1)}(g,h) > f_{\Psi[\lambda]}^{\log(U_2)}(g,h)$ if and only if $\Psi^*(\lambda(\log(U_1(g)) - \log(U_1(h)))) > \Psi^*(\lambda(\log(U_2(g)) - \log(U_2(h))))$ if and only if $\log(\frac{U_1(g)}{U_1(h)}) > \log(\frac{U_2(g)}{U_2(h)})$, or equivalently, $U_{1/2}(g) > U_{1/2}(h)$. Step 3 is completely analogous to that of Theorem 1. Step 4 requires us to notice that we can write $U_{1/2}(h) = BU_{1/2}(h_{-y}) + (1-B)U_{1/2}(\delta_y)$, where $B = \frac{(1-p)U_2(h_{-y})}{(1-p)U_2(h_{-y})+pU_2(\delta_y)}$ and h_{-y} is the gamble that appears when eliminating outcome y from gamble h and normalizing the corresponding probabilities. Since $U_2(\delta_y) \geq U_2(g')$, we have that $B \leq A = \frac{(1-p)U_2(h_{-y})}{(1-p)U_2(h_{-y})+pU_2(g')}$. Since $U_{1/2}(\delta_y) \geq U_{1/2}(h)$, it must also be the case that $U_{1/2}(\delta_y) \geq U_{1/2}(h_{-y})$ and thus, $U_{1/2}(h) = BU_{1/2}(h_{-y}) + (1-B)U_{1/2}(\delta_y) \leq AU_{1/2}(h_{-y}) + (1-A)U_{1/2}(\delta_y)$. Then, whenever $U_{1/2}(g') > U_{1/2}(\delta_y)$, we have that $U_{1/2}(h) < AU_{1/2}(h_{-y}) + (1-A)U_{1/2}(g') = U_{1/2}(g)$ and the result for mean-preserving spreads holds.■

Proof of Theorem 8: Consider two distinct gambles, $g = [\mathbf{p}, \mathbf{x}]$ and h, where either h is a degenerate gamble with monetary outcome in $(\min_i\{x_i\}, \max_i\{x_i\})$, or g is a mean-preserving spread of h. There exists $\hat{r} > 0$ such that $U^{\hat{r}}_{CARA}(g) < U^{\hat{r}}_{CARA}(h)$, and hence $\log(U^{\hat{r}}_{CARA}(g)) < \log(U^{\hat{r}}_{CARA}(h))$. Notice that, for the mean-preserving spread case, any value of r above 0 is valid, while, for the degenerate gamble case, such an $\hat{r}$ depends on the specific g and the monetary outcome.

Since the limits of $\log(rU^{r}_{CARA}(h))$ and $\log(rU^{r}_{CARA}(g))$ as r increases are both 0, there exists $\tilde{r} > \hat{r}$ such that $\log(rU^{r}_{CARA}(h)) - \log(rU^{r}_{CARA}(g)) < \log(U^{\hat{r}}_{CARA}(h)) - \log(U^{\hat{r}}_{CARA}(g))$. Given that $\log(U^{r}_{CARA}(h)) - \log(U^{r}_{CARA}(g))$ is equal to $\log(rU^{r}_{CARA}(h)) - \log(rU^{r}_{CARA}(g))$, then $\log(U^{r}_{CARA}(h)) - \log(U^{r}_{CARA}(g)) < \log(U^{\hat{r}}_{CARA}(h)) - \log(U^{\hat{r}}_{CARA}(g))$, or, equivalently $\log(\frac{U^{r}_{CARA}(h)}{U^{\hat{r}}_{CARA}(h)}) < \log(\frac{U^{r}_{CARA}(g)}{U^{\hat{r}}_{CARA}(g)})$ for all $r \geq \tilde{r}$. This, as shown in Theorem 7, is equivalent to $f^{\log(U^{r}_{CARA})}_{\Psi[\lambda]}(g, h) > f^{\log(U^{\hat{r}}_{CARA})}_{\Psi[\lambda]}(g, h)$ for all $r \geq \tilde{r}$. Now, the function $f^{\log(U^{r}_{CARA})}_{\Psi[\lambda]}(g, h)$ is continuous on $[\hat{r}, \infty)$ and, hence, achieves a minimum r^* in the closed interval $[\hat{r}, \tilde{r}]$. Since we have proved that $f^{\log(U^{r}_{CARA})}_{\Psi[\lambda]}(g, h) > f^{\log(U^{\hat{r}}_{CARA})}_{\Psi[\lambda]}(g, h)$ for all $r \geq \tilde{r}$, we know that r^* is also a minimum in $[\hat{r}, \infty)$, which concludes the proof.■

References

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

[2] Apesteguia, J. and M.A. Ballester (2014), 'Discrete Choice Estimation of Time Preferences," mimeo.

[3] 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.

[4] 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.

[5] Becker, G. M., M. H. DeGroot, and J. Marschak (1963), "Stochastic Models of Choice Behavior," *Behavioral Science*, 8.1:41–55.

[6] Blavatskyy, Pavlo R. (2011),"Probabilistic Risk Aversion with an Arbitrary Outcome Set," *Economics Letters*, 1112:34–37.

[7] 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.

[8] Cicchetti, C. J. and J. A. Dubin (1994), “A Microeconometric Analysis of Risk Aversion and the Decision to Self-Insure,” *Journal of Political Economy*, 102.1:169–186.

[9] Coble, K. H. and J. L. Lusk (2010), “At the Nexus of Risk and Time Preferences: An Experimental Investigation,” *Journal of Risk and Uncertainty*, 41.1:67–79.

[10] Cohen, A. and L. Einav (2007), “Estimating Risk Preferences from Deductible Choice,” *American Economic Review*, 97.3:745–788.

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

[12] Donkers, B., B. Melenberg, and A. Van Soest (2001), “Estimating Risk Attitudes Using Lotteries: A Large Sample Approach,” *Journal of Risk and Uncertainty*, 22.2:165–195.

[13] Eliashberg, J. and J. R. Hauser (1985), “A Measurement Error Approach for Modeling Consumer Risk Preference,” *Management Science*, 31.1:1–25.

[14] Friedman, B. (1974), “Risk Aversion and the Consumer Choice of Health Insurance Option,” *Review of Economics and Statistics*, 56.2:209–214.

[15] 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.

[16] Gul, F. and W. Pesendorfer (2006), “Random Expected Utility,” *Econometrica*, 74.1:121–146.

[17] Harless, D. W. and C. F. Camerer (1994), “The Predictive Utility of Generalized Expected Utility Theories,” *Econometrica*, 62.6:1251–1289.

[18] 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.

[19] Hey, J. D., A. Morone, and U. Schmidt (2009), “Noise and Bias in Eliciting Preferences,” *Journal of Risk and Uncertainty* , 39.3:213–235.

[20] Hey, J. D. and C. Orme (1994), “Investigating Generalizations of Expected Utility Theory Using Experimental Data,” *Econometrica*, 62.6:1291–1326.

[21] Holt, C. A. and S. K. Laury (2002), “Risk Aversion and Incentive Effects,” *American Economic Review*, 92.5:1644–1655.

[22] 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.

[23] 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.

[24] Loomes, G. and R. Sugden (1995), “Incorporating a Stochastic Element Into Decision Theories,” *European Economic Review*, 39.3:641–648.

[25] Markowitz, H. (1952), “Portfolio Selection,” *Journal of Finance*, 7.1:77–91.

[26] McFadden, D. (2001), “Economic Choices,” *American Economic Review*, 91.3:351–378.

[27] McKelvey, R. D. and T. R. Palfrey (1995), “Quantal Response Equilibria for Normal Form Games,” *Games and Economic Behavior*, 10.1:6–38.

[28] 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.

[29] 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.

[30] Pratt, J. W. (1964), "Risk Aversion in the Small and in the Large," *Econometrica*, 32.1/2:122–136.

[31] 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.

[32] Rust, J. and C. Phelan (1997), "How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets," *Econometrica*, 65:781832.

[33] 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.

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

[35] 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.

[36] Wilcox, N.T. (2011), "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.

[37] Wilcox, N.T. (2011), "Stochastically More Risk Averse: A Contextual Theory of Stochastic Discrete Choice Under Risk," *Journal of Econometrics*, 162.1:89–104.