Random Expected Utility and Certainty Equivalents: Mimicry of Probability Weighting Functions

Wilcox, Nathaniel

Chapman University

11 August 2016

Random Expected Utility and Certainty Equivalents: Mimicry of Probability Weighting Functions

by

Nathaniel T. Wilcox
Economic Science Institute
Chapman University
Orange, CA 92866
nwilcox@chapman.edu

August 11, 2016
(version 1.1)

Abstract: For simple prospects of the kind routinely used for certainty equivalent elicitation, random expected utility preferences imply a conditional expectation function that can mimic deterministic rank dependent preferences. That is, an agent with random expected utility preferences can have mean certainty equivalents that look exactly like rank dependent probability weighting functions of the inverse-s shape discussed by Quiggin (1982) and later advocated by Tversky and Kahneman (1992) and other scholars. It seems that certainty equivalents cannot nonparametrically identify preferences, at least not in every relevant sense, since their conditional expectation depends on assumptions concerning the source and nature of their variability.

I am grateful to the Economic Science Institute and Chapman University for their ongoing support. Jose Apesteguia, Matias Del Campo Axelrod, Miguel A. Ballester, Mark DeSantis, Glenn W. Harrison, John P. Nolan and Mark Schneider provided help or commentary, though none are responsible for remaining errors.

Elicitation of certainty equivalents has become routine in laboratory measurement of preferences under risk and uncertainty (Tversky and Kahneman 1992; Tversky and Fox 1995; Wu and Gonzales 1999; Gonzales and Wu 1999; Abdellaoui 2000; Abdellaoui, Bleichrodt and Paraschiv 2007; Halevy 2007; Abdellaoui, Bleichrodt, and L'Haridon 2008; Bruhin, Fehr-Duda and Epper 2010; Vieider et al. 2015). While elicitation methods vary across such studies, the formal empirical interpretation of elicited certainty equivalents—or computed quantities interpreted to be certainty equivalents—is overwhelmingly the same: The subject is assumed to have a unique and fixed preference order, implying (under unchanged conditions of background wealth, risk and so forth) a unique and fixed certainty equivalent C_L for each prospect L. One may then interpret an elicited certainty equivalent c as C_L plus some error ε of harmless origin with standard properties. When elicitation is repeated for exactly the same prospect, elicited certainty equivalents are variable within subjects (e.g. Tversky and Kahneman 1992, p. 307; Gonzalez and Wu 1999, pp. 144-146) and other studies also suggest inherent variability of elicited certainty equivalents (Butler and Loomes 2007; Loomes and Pogrebna 2014). Empirical interpretation needs to take a position on this variability, and adding mean zero error to an otherwise deterministic model of certainty equivalents is one option. I call this the *standard model* of an elicited certainty equivalent.

Alternatively, one might assume that the individual subject's preference order is a random variable, and that any one certainty equivalent elicited from that subject is fully determined by a single realization of that random variable. I call this a *random (preference) model* of an elicited certainty equivalent. Interest in random preference models is long-standing (Becker, DeGroot and Marschak 1963; Eliashberg and Hauser 1985; Hilton 1989; Loomes and Sugden 1995, 1998; Gul and Pesendorfer 2006; Apesteguia and Ballester 2016), particularly in the realm of discrete choice. Here, I examine implications of this model for elicited certainty equivalents c and find a significant complication of their empirical interpretation.

Specifically, random model expected utility preferences (or more simply *random EU* as termed by Gul and Pesendorfer 2006) imply a conditional expectation function for c that can *mimic* standard model rank-dependent preferences (or more simply *standard RDU*). That is, a random EU agent can have mean certainty equivalents that appear to reveal rank dependent probability weighting functions of the inverse-s shape discussed by Quiggin (1982) and advocated by Tversky and Kahneman (1992) and other scholars. I believe Hilton (1989) first showed that certainty equivalents have some unexpected properties under random EU; additionally, recent work by Navarro-Martinez et al. (2015) contains a strong suggestion of my direction here. In my conclusion I discuss implications of this finding more thoroughly, but it seems that elicited certainty equivalents cannot nonparametrically identify preferences, at least not in the mainstream econometric sense of the phrase "nonparametric identification," since their conditional expectation depends on the source and nature of their variability.

Years ago Karni and Safra (1987) observed that, from the perspective of rank-dependent preference theory, incentive-compatible elicitation of certainty equivalents might not be possible; at the same time, many behavioral economists doubt that incentive compatibility matters much for preference elicitation (Camerer and Hogarth 1999; Loewenstein 1999), though for many kinds of value elicitation there is overwhelming evidence to the contrary (Harrison and Rutström 2008). Let me firmly distinguish my findings from the concerns of Karni and Safra, and also stipulate (for argument's sake) what the behavioral mainstream believes about methods. My formal results require only general assumptions about elicitation methods, and say nothing about incentive compatibility. The results come first, followed by some intuition behind the results, and I close with some caveats and implications of the results.

1. Formal results.

Consider *simple prospects* $L = (W, p)$ with money outcomes $z = W > 0$ with probability p and $z = 0$ with probability $1 - p$. Simple prospects figure prominently in theoretical and empirical discussions of rank-dependent utility (RDU) and cumulative prospect theory (CPT) because their certainty equivalents are thought to reveal the *probability weighting function* of the rank-dependent family when the utility or value of outcomes is linear (Tversky and Kahneman 1992; Prelec 1998). To see this, let the utility or value of outcomes have the power form $v(z) = z^{1/x}$, where I write the power as $1/x$ for convenience. Given any specific $x \in (0, \infty)$, the rank-dependent utility or RDU of L will be

$V_L = \pi(p|\omega)v(W) + \big(1 - \pi(p|\omega)\big)v(0) = \pi(p,\omega)W^{1/x}$, where $\pi(p|\omega)$ is a probability weighting function depending on preference parameters ω. The certainty equivalent of $L = (W, p)$, given x, is then $W\pi(p|\omega)^x$, but divide these by W to free them of dependence on W and let $C_s^{rd}(p|x,\omega) \equiv \pi(p|\omega)^x$ be the RDU *normalized certainty equivalent* of any simple prospect (given specific x and ω). Notice that when $x = 1$ (that is for a linear value of outcomes), one has $C_s^{rd}(p|1,\omega) \equiv \pi(p|\omega)$ so that these normalized certainty equivalents of simple prospects are thought to reveal the RDU weighting function when $x = 1$. Expected utility or EU is the special case where $\pi(p|\omega) \equiv p$, so also define $C_s^{eu}(p|x) \equiv p^x$ as the EU normalized certainty equivalent of any simple prospect (given specific x).

Let c be an observed certainty equivalent for $L = (W, p)$, elicited from some subject and divided by W to normalize it. Very commonly, the empirical specification for these observed certainty equivalents is $c = E(c|p) + \varepsilon$. Here $E(c|p)$ is the *conditional expectation function* or c.e.f. of c, usually derived from EU, RDU, CPT or another preference functional (as done above); and ε is an error term, usually thought to arise (for instance) from "carelessness, hurrying, or inattentiveness" (Bruhin, Fehr-Duda and Epper 2010) and assumed to satisfy conventional assumptions ($E(\varepsilon) = E(\varepsilon|p) = 0$). Estimation of $E(c|p)$ can then proceed using nonlinear least squares or another estimator such as maximum likelihood. This is the essence of the standard model approach.

Standard model RDU assumes that $E(c|p) = C_s^{rd}(p|x,\omega) \equiv \pi(p|\omega)^x$, where x is a nonrandom parameter to be estimated. Bruhin, Fehr-Duda and Epper (2010) use maximum likelihood this way, while Tversky and Kahneman (1992) use nonlinear least squares. In other words, the conditional expectation of c is taken to be $\pi(p|\omega)^x$, and the data analyst's job is to estimate unique and fixed parameters x and ω from multiple observations of c observed under experimentally varied values of p. Obviously $c \in [0,1]$, so a few other modeling assumptions appropriate to limited dependent variables are needed: The distribution of ε cannot be wholly independent of p unless it is degenerate. But this can be done while keeping $E(\varepsilon|p) = 0$ in straightforward ways (see e.g. Gonzalez and Wu 1999; Bruhin, Fehr-Duda and Epper 2010). Frequently, estimations also involve some (or even wholesale) pooling across different subjects. Random preference models of individuals can also be models of the expected behavior of groups under suitable conditions (McFadden

1974), so my results may apply with added force to such pooled estimation. My appendix is a very brief illustrative Monte Carlo look at many such estimation techniques in the presence of random preference variation.

For many common weighting functions $\pi(p|\omega)$ it will not be possible to identify both x and all of ω solely from normalized certainty equivalents of simple prospects. For instance suppose $\pi(p|\omega)$ is the well-known 2-parameter Prelec (1998) weighting function $\exp(-\beta[-\ln(p)]^{\alpha})$ where α and β are strictly positive parameters. Then the normalized certainty equivalent will be $C_s^{rd}(p|x,\alpha,\beta) = \exp(-x\beta[-\ln(p)]^{\alpha})$, and only α and the product $x\beta$ will be estimable. This is very well-known, so experimental designs meant to separately estimate all three parameters always contain some prospects that are not simple prospects as defined here. I focus on simple prospects because of their tractability and their simple interpretation: Definitionally, $C_s^{rd}(p|1,\omega) \equiv \pi(p|\omega)$, so the normalized certainty equivalents of simple prospects are identical to weighting functions (under standard RDU) at linear $v(z)$ (Tversky and Kahneman 1992; Prelec 1998).

In general, a *random (preference) model* might take both x and ω to be realizations of nondegenerate random variables X and Ω within an individual. However, as far as I am aware, existing random preference estimations treat any weighting function parameters ω as fixed within any individual (e.g. Loomes, Moffatt and Sugden 2002; Wilcox 2008, 2011), and contemporary random preference theory seems to be confined to treatment of X as random only (e.g. Gul and Pesendorfer 2006; Apesteguia and Ballester 2016). Therefore, all of my random preference analysis treats only x as the realization of a random variable X, taking any weighting function parameters ω as fixed within the subject.

I confine my analysis of the random model to the random EU special case. In any given trial of any elicitation, an ordinary random EU assumes that a new realization x of X occurs and determines the normalized certainty equivalent $C_s^{eu}(p|x) = p^x$. Assume that a probability density function $f(x|\psi)$ of X with support $(0,\infty)$ lies within an individual, depending on parameters ψ. Then define $C_r^{eu}(p|\psi)$ as

(1) $$C_r^{eu}(p|\psi) \equiv E_X[C_s^{eu}(p|x)] = \int_0^\infty p^x f(x|\psi)dx = \int_0^\infty \exp(-xt) f(x|\psi)dx,$$

where the final integral (which becomes useful shortly) lets $t = -\ln(p)$ and rewrites p^x as $\exp(-xt)$. The function $C_r^{eu}(p|\psi)$ is the mean normalized certainty equivalent of a random EU agent for simple prospects (W, p), given her underlying p.d.f. $f(x|\psi)$.

Allow a brief digression on elicitation methods. There is another way of thinking about this p.d.f. $f(x|\psi)$ of X. Suppose an experimenter uses some method M to elicit a normalized certainty equivalent c from a subject, and suppose that $x(c,p)$ solves $p^x = c$; that is, let $x(c,p) \equiv \ln(c)/\ln(p)$. Suppose that in repeated elicitations using method M, the empirical c.d.f. of $x(c,p)$ is observed to be $\hat{F}_M(x)$; and suppose that $\hat{F}_M(x)$ converges to $F_M(x)$ as the sample of observations grows. If $F_M(x)$ is independent of p, the variability of the normalized certainty equivalents observed by the experimenter could be interpreted as arising from a random EU model of the kind assumed here, where the p.d.f. $f(x|\psi)$ is derived from $F_M(x)$. This suggests ways in which one might test a random EU model (or a random RDU model, and later I will return to this), but additionally indicates that the results below do not depend much on the specific laboratory method used to elicit a certainty equivalent. The two key assumptions about any elicitation method are (1) that repeated trials using the method yield variability in elicited certainty equivalents, and (2) this variability is consistent with the assumptions of a random EU model—namely, that $F_M(x)$ is independent of p. Note that neither of those assumptions rule out any dependence of $f(x|\psi)$ on the elicitation method M.

Under random EU, we again have an empirical model of the form $c = E(c|p) + \xi$, but the c.e.f. is now $E(c|p) = C_r^{eu}(p|\psi)$ as given by eq. (1) and errors ξ are $p^x - C_r^{eu}(p|\psi)$. By the eq. 1 definition, the new errors ξ also satisfy the usual properties ($E(\xi) = E(\xi|p) = 0$), so one may estimate both the random EU model $c = C_r^{eu}(p|\psi) + \xi$ and the standard RDU model $c = C_s^{rd}(p|x,\omega) + \varepsilon$ using the same variety of estimators.

The close resemblance of these two models suggests two possible types of *mimicry*. First, since $C_s^{rd}(p|1,\omega) \equiv \pi(p|\omega)$ in standard RDU, it will be troubling if $C_r^{eu}(p|\psi)$ can "look like" a stereotypical $\pi(p|\omega)$, that is, can have properties like those that scholars believe are empirically characteristic of RDU weighting functions. I will refer to this as *weak mimicry* (of standard RDU by random EU). Second, it may happen that for some well-known and specific $\pi(p|\omega)$, there exists a specific $f(x|\psi)$ such that $C_r^{eu}(p|\psi)$ is a *re-parameterization* of $C_s^{rd}(p|x,\omega)$. Let D be the set of possible parameter vectors (x,ω), and let Ψ be the set of possible parameter vectors ψ; and suppose that, for some $f(x|\psi)$, there exists a function $H_f: D \rightarrow \Psi$ such that $C_r^{eu}[p, H_f(x,\omega)] \equiv C_s^{rd}(p|x,\omega) \equiv \pi(p|\omega)^x$: Then one may say there is *strong mimicry* (of standard RDU by random EU) for $f(x|\psi)$. Notice that strong mimicry implies weak mimicry but not vice versa.

Since $-\ln(p) > 0\ \forall\ p \in (0,1)$, so that $t > 0$ too, the final integral in eq. 1 is the one-sided Laplace transform $\mathcal{L}\{f\}(t)$ of the p.d.f. $f(x|\psi)$—provided it exists; and below I only use p.d.f.s for which the existence and form of $\mathcal{L}\{f\}(t)$ have been demonstrated and derived by others. In such instances, these known Laplace transforms $\mathcal{L}\{f\}(t)$ of a p.d.f. $f(x|\psi)$ make it simple to derive various examples of $C_r^{eu}(p|\psi)$, using the relationship $C_r^{eu}(p|\psi) = \mathcal{L}\{f\}[-\ln(p)]$. Some examples follow.

Example 1. Suppose X has the Gamma p.d.f

$$f(x|k,\theta) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-(x/\theta)} \text{ for } x, k, \text{ and } \theta \in (0,\infty).$$

It's widely known that this has the Laplace transform $\mathcal{L}\{f\}(t) = (1+\theta t)^{-k}$, implying that

(2) $$C_r^{eu}(p|k,\theta) = (1-\theta\ln(p))^{-k}.$$

Figure 1-A shows this Gamma c.e.f. for $k = 0.75$ and θ =2.79. At these parameter choices, it has the "inverse-s" shape many believe is characteristic of weighting functions $\pi(p|\omega)$ and the fixed point $p \approx e^{-1}$ which is also characteristic of Prelec's (1998) 1-parameter weighting function; so this is an instance of weak mimicry. One needs to say that this Gamma c.e.f. *can* (not *must*) weakly mimic this characteristic shape. Figure 1-B shows the Gamma c.e.f. for $k = 0.75$ and θ =0.9: Here we see the "Optimist" shape discussed by Quiggin (1982), and also the plurality shape of individually estimated weighting functions in Wilcox (2015). Such shape flexibility is also characteristic of 2-parameter weighting functions found in the literature on RDU and CPT, where this flexibility is usually regarded as a feature rather than a weakness.

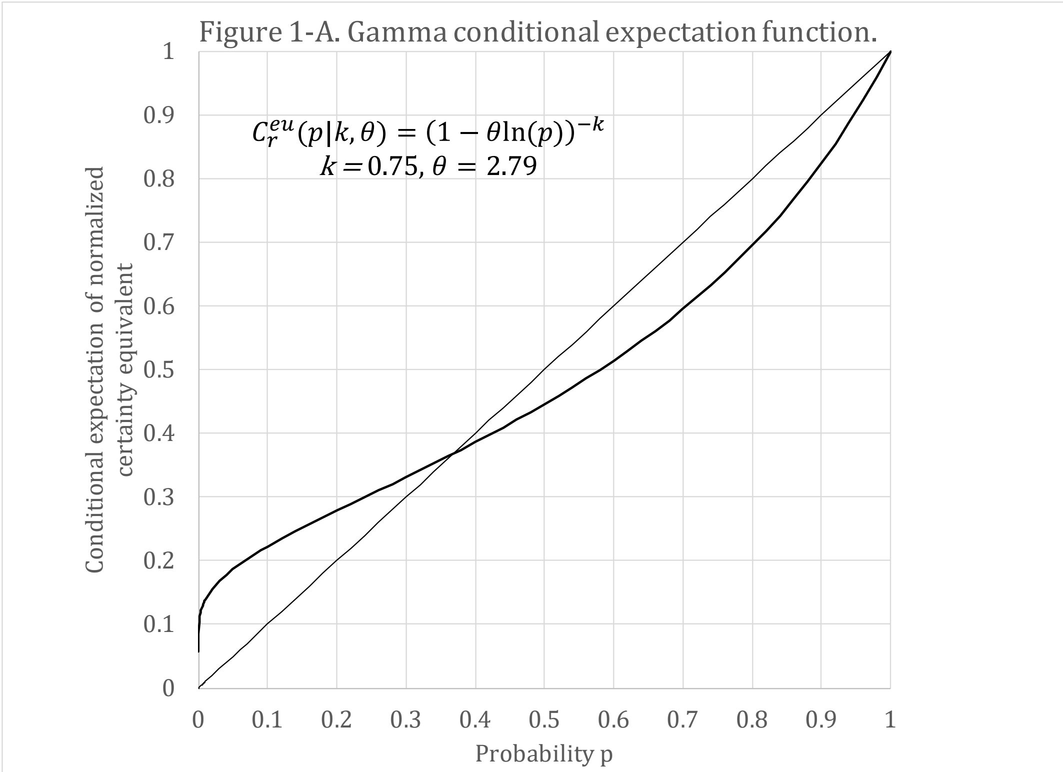


Figure 1-A. Gamma conditional expectation function.
$C_r^{eu}(p|k,\theta) = (1-\theta\ln(p))^{-k}$
$k = 0.75$, $\theta = 2.79$
Conditional expectation of normalized certainty equivalent
Probability p
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

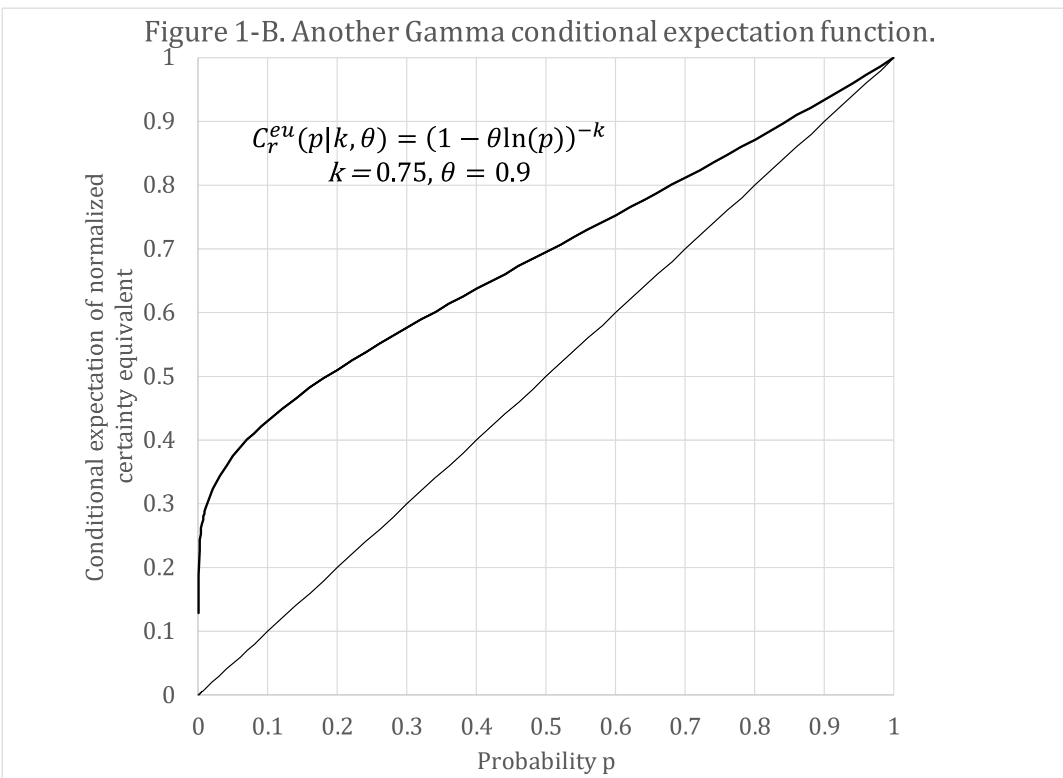


Figure 1-B. Another Gamma conditional expectation function.
$C_r^{eu}(p|k,\theta) = (1-\theta\ln(p))^{-k}$
$k = 0.75$, $\theta = 0.9$
Conditional expectation of normalized certainty equivalent
Probability p
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Example 2. Suppose X has the Inverse Gaussian (Wald) p.d.f.

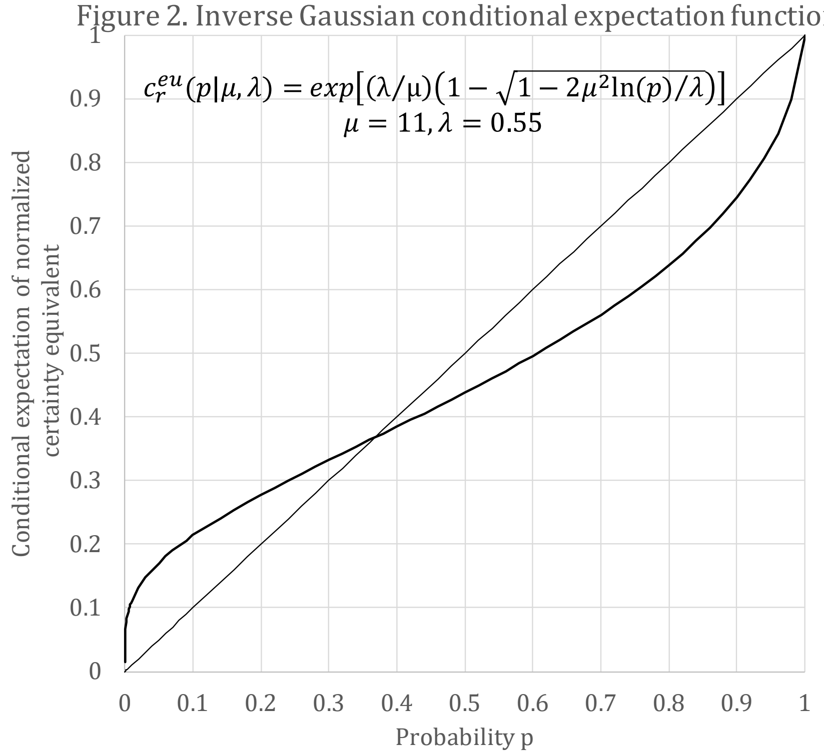
$$f(x|\mu,\lambda) = \sqrt{\frac{\lambda}{2\pi}}\, x^{-3/2} exp\left[\frac{-\lambda(x-\mu)^2}{2\mu^2 x}\right] \text{ for } x, \mu \text{ and } \lambda \in (0,\infty).$$

This has the Laplace transform $\mathcal{L}\{f\}(t) = exp\left[(\lambda/\mu)\left(1-\sqrt{1+2\mu^2 t/\lambda}\right)\right]$ (Seshadri 1993 p. 41), implying that

(3) $$C_r^{eu}(p|\mu,\lambda) = exp\left[(\lambda/\mu)\left(1-\sqrt{1-2\mu^2\ln(p)/\lambda}\right)\right].$$

Figure 2 shows this c.e.f. for $\mu = 11$ and $\lambda = 0.55$. This also has the inverse-s shape and is also an instance of weak mimicry. Again, different parameter values will produce a wide variety of different shapes of this Inverse Gaussian c.e.f.

Figure 2. Inverse Gaussian conditional expectation function.



Example 3. Suppose X has the (unshifted) Lévy p.d.f.

$$f(x|\delta) = \frac{\delta}{2\sqrt{\pi}} x^{-3/2} exp\left(\frac{-\delta^2}{4x}\right) \text{ for } x \text{ and } \delta \in (0, \infty).$$

This has the Laplace transform $\mathcal{L}\{f\}(t) = exp\left(-\delta t^{1/2}\right)$ (González-Velasco 1995, p. 537), implying that

(4) $$C_r^{eu}(p|\delta) = exp\left(-\delta[-\ln(p)]^{1/2}\right).$$

Earlier I noted that in the case of the Prelec (1998) 2-parameter weighting function, $c_s^{rd}(p|x,\alpha,\beta) = \exp(-x\beta[-\ln(p)]^{\alpha})$. Clearly, this is identical to eq. 4 if we set $\delta = x\beta$ and require that $\alpha = 1/2$. This is very close to being a case of strong mimicry, but not quite, since eq. 4 can only mimic the Prelec weighting function when α just happens to be 1/2. Empirically, estimates of α have a wider range than a small neighborhood of 1/2.

However, this result for the Lévy distribution provides a strong and fruitful hint. The Lévy distribution is but one specific instance of the Lévy Alpha-Stable distributions, also known more simply as the Stable distributions. Except for special cases (e.g. Normal, Cauchy and Lévy), Stable random variables X usually have no p.d.f. expressible in terms of elementary functions: Instead, they are generally expressed by their characteristic function $\varphi(t) \equiv E[\exp(itX)]$. For Stable random variables with support $(0, \infty)$, one parameterization (Nolan 2017, pp. 8-12) is $\varphi(t) = \exp\left\{-\gamma^a|t|^a\left[1 - i\left(tan\frac{a\pi}{2}\right)\right]\right\}$, where $\gamma > 0$ is a scale parameter and $a \in (0,1)$ is called the *index of stability* or *characteristic exponent*. Nolan (2017, p. 109) also shows that the Laplace transform of these particular Stable distributions on $(0, \infty)$ exists and is $\mathcal{L}\{f\}(t) = \exp\left\{-\gamma^a\left(sec\frac{a\pi}{2}\right)t^a\right\}$, so for these Stable distributions we have

(5) $$C_r^{eu}(p|a,\gamma) = \exp\left\{-\gamma^a\left(sec\frac{a\pi}{2}\right)[-\ln(p)]^a\right\},$$

Eq. 5 is identical to $C_s^{rd}(p|x,\alpha,\beta) = \exp(-x\beta[-\ln(p)]^{\alpha})$ when we set $a = \alpha$ and $\gamma = \left[\beta x \left(cos\frac{\alpha\pi}{2}\right)\right]^{1/\alpha}$, so we have strong mimicry of the 2-parameter Prelec (1998) function, provided that $\alpha < 1$. Since this is both characteristic of most empirical estimates of α and indeed yields the characteristic inverse-s shape, this is strong mimicry of a well-known and widely used probability weighting function in the relevant part of the parameter space.

2. Intuition.

To see the intuition behind the formal results, it helps to return to the more usual representation of the power utility function, that is $u(z) = z^{\sigma}$, thinking now of σ as having a distribution with support $\Sigma \subseteq (0,\infty)$ under the random preference model. Then under EU the normalized certainty equivalent of a simple prospect (given any value of σ) will be $p^{1/\sigma}$, whose second derivative with respect to σ is

(7) $$\frac{\partial^2}{\partial\sigma^2}\, p^{1/\sigma} = \frac{-\ln(p)p^{1/\sigma}}{\sigma^4}[-\ln(p) - 2\sigma] > 0 \;\text{ for all }\; \sigma < -\frac{1}{2}\ln(p).$$

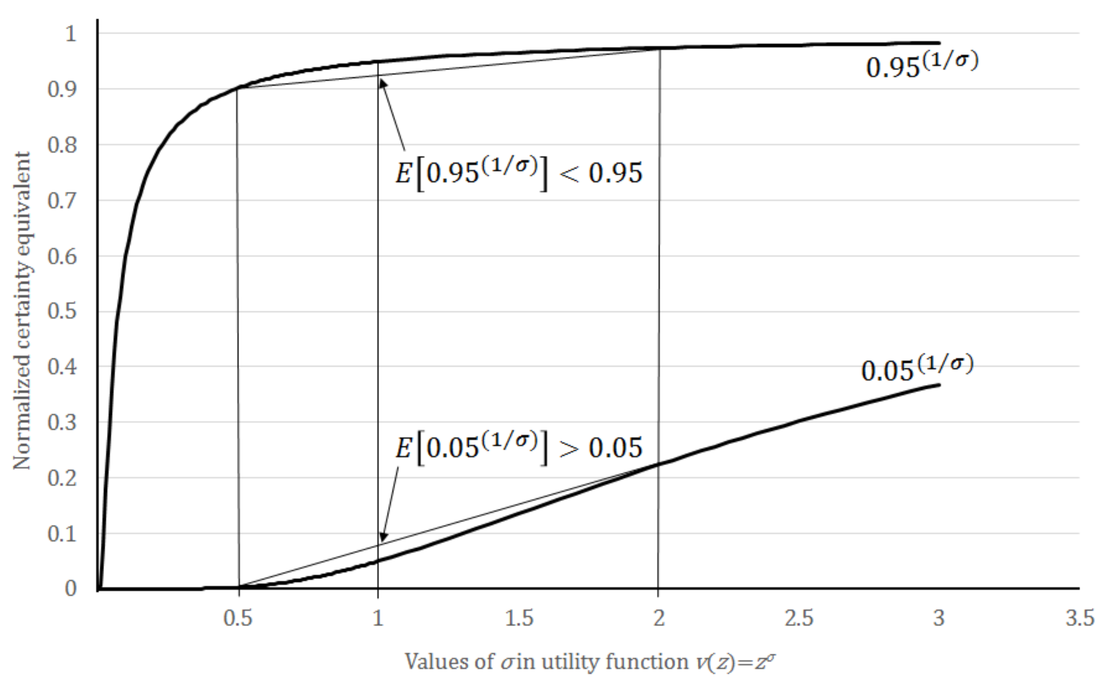
As p approaches zero, eq. 7 shows that the normalized certainty equivalent $p^{1/\sigma}$ approaches being convex in σ at all $\sigma \in (0,\infty)$. If Σ is in fact bounded above, there must be sufficiently small p such that $p^{1/\sigma}$ is convex $\forall\ \sigma \in \Sigma$: In that event, Jensen's Inequality implies that $E\left(p^{1/\sigma}\right) > p^{1/E(\sigma)}$. Assuming that $E(\sigma) \geq 1$, then, we have $E\left(p^{1/\sigma}\right) > p$ for sufficiently small p. That is, mean normalized certainty equivalents will exceed p when p is small enough: We have apparent overweighting of small (enough) probabilities if the support of σ is bounded above.

As p approaches 1, on the other hand, eq. 7 shows that $p^{1/\sigma}$ becomes concave at almost all σ, and the argument above flips around: If Σ is bounded below away from zero, there will be p sufficiently close to 1 such that $p^{1/\sigma}$ is concave $\forall\ \sigma \in \Sigma$, and Jensen's Inequality then implies that $E\left(p^{1/\sigma}\right) < p^{1/E(\sigma)}$. Again setting $E(\sigma) = 1$, we have $E\left(p^{1/\sigma}\right) < p$ for p sufficiently close to one. That is, mean normalized certainty equivalents will fall short of p

when p is high enough: We have apparent underweighting of high (enough) probabilities if the support of σ is bounded below away from zero.

Figure 3 illustrates this intuition. Assume that the agent has a binomial distribution on σ such that $E(\sigma) = 1$: Specifically she has $\sigma = 0.5$ with probability $2/3$ and $\sigma = 2$ with probability $1/3$. Figure 3 shows the function $p^{1/\sigma}$ for $\sigma \in (0,3]$ for two values of p. The upper heavy curve is for $p = 0.95$ and, as can be seen, this curve is overwhelmingly and strongly concave: In this case, $E\left[0.95^{(1/\sigma)}\right] < 0.95$, so this agent appears to underweight high probabilities. The lower heavy curve is for $p = 0.05$ and, as can be seen, this curve is first convex and, for σ beyond about 1.5, very gently concave: Here, $E\left[0.05^{(1/\sigma)}\right] > 0.05$, so this agent appears to overweight low probabilities.

Figure 3. Behavior of normalized certainty equivalent in σ at $p = 0.05$ and $p = 0.95$



This story does not completely explain the formal results: All of the examples in section 1 involve p.d.f.s with support $(0, \infty)$, so this story (which is told by appealing to a support bounded above and bounded below away from zero) is only an aid to intuition, not any sort of demonstration—which is another way of saying the formal results are necessary. The intuition does, however, explain why one may easily derive the characteristic s-shape from many p.d.f.s $f(x|\psi)$ underlying a random EU model.

3. Discussion and Conclusions

My results complicate empirical interpretation of elicited certainty equivalents. However, I say 'complicate' rather than 'undermine' for several reasons. First, I have not shown that random preference EU and standard RDU are indistinguishable. My formal results are entirely about conditional expectations and say nothing about conditional medians or conditional variances and other moments; and one might test both random EU and random RDU on the basis of these other characteristics. For instance, recall that $C_s^{rd}(p,x,\omega) \equiv \pi(p|\omega)^x$ and suppose we now assume that x is a realization of a random variable X, giving a random RDU model. Let CV denote coefficient of variation; then under random RDU, assuming that any weighting function parameters are fixed (not themselves random variables), we have

$$(8) \qquad CV(-\ln(c)) \equiv \frac{\sqrt{V(-\ln(c))}}{E(-\ln(c))} = \frac{\sqrt{V(X)[-\ln(\pi(p|\omega))]^2}}{E(X)[-\ln(\pi(p|\omega))]} = \frac{\sqrt{V(X)}}{E(X)} = CV(X)$$

This says that for any given individual, the coefficient of variation of $-\ln(c)$ will be equivalent to the coefficient of variation of X and, moreover, independent of the particular W and p of any simple prospect (W,p), regardless of whether the weighting function is an identity function (EU) or not (RDU). This immediately suggests a test of both random preference EU and RDU based on multiple (more than two) certainty equivalent elicitation trials for several different simple prospects. To my knowledge, such data are scarce but more could be gathered with appropriate experimental designs. The key point, however, is that for certainty equivalents, the random preference hypothesis can make strong refutable predictions about higher moments that are independent of the form or even the presence of any rank-dependent weighting function. Whether the same can be said of the standard model is unclear to me.

Second, discrete choice experiments already suggest that random EU cannot be a complete model of decision behavior (e.g. Becker, DeGroot and Marschak 1963; Loomes and Sugden 1998). Under the random preference hypothesis, much of what EU predicts concerning pairs of related discrete choice problems remains unchanged relative to what EU predicts in its deterministic form (Loomes and Sugden 1995; Gul and Pesendorfer 2006; Wilcox 2008). This implies that many well-known discrete choice violations of EU also violate random EU. Here I showed once more (see Hilton 1989) that certainty equivalents are a different matter: Under random EU, the expected values of certainty equivalents can mimic predictions of standard RDU and CPT. The upshot of this fact is that when one estimates risk models from certainty equivalents, part of the estimates (perhaps substantial parts) may reflect random preference heterogeneity as well as any underlying mean preference.

Third, my formal analysis only complicates estimation based on conditional expectation functions. While this is the overwhelmingly common basis for *estimation*, some of the empirical literature on RDU and CPT uses conditional medians of certainty equivalents for *description* (see e.g. Tversky and Kahneman 1992, pp. 309-311). It may be that conditional median estimation (that is, least absolute deviation or LAD estimators) can sidestep the issue uncovered here. Recall the key role played by Jensen's Inequality in Section 2 where I discussed intuition: There is no counterpart of Jensen's Inequality for medians. As far as I am aware, there are no conditional median estimations (based on elicited certainty equivalents) of either RDU or CPT models. My appendix takes a look at a LAD estimator and finds encouraging results for random EU data, but *not* for standard model EU data. I do not know whether any estimator exists that would correctly identify weighting functions regardless of the true error model generating the data; finding such an estimator would be a nice contribution to decision research.

However, meaningful preference measurement may not be possible without strong assumptions concerning the random part of decision behavior (Wilcox 2008; Blavatskyy and Pogrebna 2010; Wilcox 2011; Apesteguia and Ballester 2016). Many scholars say that elicited certainty equivalents, or quantities that are argued to be estimated certainty equivalents, permit "nonparametric" identification and estimation of preferences (Gonzales and Wu 1999; Abdellaoui 2000; Abdellaoui, Bleichrodt and Paraschiv 2007). The word “nonparametric” gets used in many different ways, but many authors divide discussion of models in two parts: (1) a conditional expectation function, or perhaps a conditional median function, and (2) the error, the random part that remains once such a function has been removed in a way that makes the expectation of the error zero. Generally, in the preference measurement literature, scholars who say their estimation is “nonparametric” mean that they are making relatively few assumptions about the *form* of preference entities (utilities or values, and probability weights, and so forth) that appear in a c.e.f. However, they routinely make a strong assumption about the random part of observed behavior, and I showed that this assumption is critically consequential.

The essence of this strong but implicit assumption is that the c.e.f. has an obvious interpretation—the intended interpretation being that of algebraic (deterministic) decision theory. Hendry and Morgan (2005, p. 23) argue that when we speak of model identification, we have things in mind beyond the original Cowles Foundation meaning—including “correspondence to the desired entity” and “satisfying the assumed interpretation (usually of a theory model).” Estimation of preferences from elicited certainty equivalents is now complicated in just these senses. We cannot take the standard model for granted, and under a random preference model, the c.e.f. in part reflects the underlying distribution of preferences within the individual, in ways that can mimic the “desired entity”—in the case discussed here, the preference entity called the probability weighting function. I do not know whether certainty equivalents can nonparametrically identify such entities: This question needs a good answer. However, at this time there is certainly no clear presumption that elicited certainty equivalents dominate elicited discrete choices as a basis for estimation of preference. Both kinds of data seem to require strong assumptions for meaningful estimation of preference entities.

References

Abdellaoui, M., 2000, Parameter-free elicitation of utility and probability weighting functions. Management Science 46, 1497-1512.

Abdellaoui, M., H. Bleichrodt and C. Paraschiv, 2007, Loss aversion under prospect theory: A parameter-free measurement. Management Science 53, 1659-1674.

Abdellaoui, M., H. Bleichrodt, and O. L'Haridon, 2008, A tractable method to measure utility and loss aversion under Prospect Theory." Journal of Risk and Uncertainty 36, 245-66.

Apesteguia, and M. Ballester, 2016, Monotone stochastic choice models: The case of risk and time preferences. Journal of Political Economy (forthcoming).

Becker, G. M., M. H. DeGroot and J. Marschak, 1963, Stochastic models of choice behaviour. Behavioral Science 8, 41–55.

Blavatskyy, P. and G. Pogrebna, 2010, Models of stochastic choice and decision theories: Why both are important for analyzing decisions. Journal of Applied Econometrics 25, 963-986.

Bruhin, A., H. Fehr-Duda and T. Epper, 2010, Risk and rationality: Uncovering heterogeneity in probability distortion. Econometrica 78:1375-1412.

Butler, D. and G. Loomes, 2007, Imprecision as an account of the preference reversal phenomenon. American Economic Review 97, 277-297.

Camerer, C. and R. Hogarth, 1999, The effects of financial incentives in experiments: A review and capital-labor-production framework. Journal of Risk and Uncertainty 19, 7-42.

Eliashberg, J. and J.R. Hauser, 1985, A measurement error approach for modeling consumer risk preference. Management Science 31, 1-25.

González-Velasco, E. A., 1995, Fourier analysis and boundary value problems. San Diego: Academic Press.

Gonzalez, R. and G. Wu, 1999, On the shape of the probability weighting function. Cognitive Psychology 38, 129-166.

Gul, F. and W. Pesendorfer, 2006, Random expected utility. Econometrica 74,121-146.

Halevy, Y., 2007, Ellsberg revisited: An experimental study. Econometrica 75, pp. 503-536.

Harrison, G. W. and E. E. Rutström, 2008, Experimental evidence on the existence of hypothetical bias in value elicitation methods. In C. R. Plott and V. L. Smith, eds., Handbook of experimental economics results 1, 752-767. Amsterdam: North-Holland.

Hendry, D. F. and M. S. Morgan, 2005, The foundations of econometric analysis. Cambridge: Cambridge University Press.

Hilton, R.W., 1989. Risk attitude under random utility. Journal of Mathematical Psychology 33, 206–222.

Karni, E. and Z. Safra, 1987, "Preference reversal" and the observability of preferences by experimental methods. Econometrica 55, 675-685.

Loewenstein, G., 1999, Experimental economics from the vantage-point of behavioural economics. Economic Journal 109, F25-F34.

Loomes, G., P. Moffatt, R. Sugden, 2002, A microeconometric test of alternative stochastic theories of risky choice. Journal of Risk and Uncertainty 24, 103-130.

Loomes, G. and G. Pogrebna, 2014, Measuring individual risk attitudes when preferences are imprecise. Economic Journal 124, 569-593.

Loomes, G. and R. Sugden, 1995, Incorporating a stochastic element into decision theories. European Economic Review 39, 641-648.

Loomes, G. and R. Sugden, 1998, Testing different stochastic specifications of risky choice. Economica 65, 581-598.

McFadden, D., 1974, Conditional logit analysis of qualitative choice behavior. In P. Zarembka, ed., Frontiers in Econometrics, 105-142. New York: Academic Press.

Navarro-Martinez, D., G. Loomes, A. Isoni and D. Butler, 2014, Boundedly rational expected utility theory. University of Warwick working paper.

Nelder, J. A. and R. Mead, 1965, "A simplex method for function minimization." Computer Journal 7, 308–313.

Nolan, J. P., 2017 (forthcoming), Stable distributions—Models for heavy tailed data. Boston: Birkhauser.

Powell, M. J. D., 1992, A direct search optimization method that models the objective and constraint functions by linear interpolation. Technical Report DAMTP 1992/NA5, Department of Applied Mathematics and Theoretical Physics, University of Cambridge.

Prelec, D., 1998, The probability weighting function. Econometrica 66, 497-527.

Quiggin, J, 1982, A theory of anticipated utility. Journal of Economic Behavior and Organization 3, 323-343.

Seshadri, V., 1993, The inverse Gaussian distribution—A case study in exponential families. Oxford: Clarendon Press.

Tversky, A. and D. Kahneman, 1992, Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty 5, 297–323.

Vieider, F. M., M. Lefebvre, R. Bouchouicha, T. Chmura, R. Hakimov, M. Krawczyk and P. Martinsson, 2015, Common components of risk and uncertainty attitudes across contexts and domains: Evidence from 30 countries. Journal of the European Economic Association 13, 421-452.

Wilcox, N., 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 pp. 197-292. Bingley, UK: Emerald.

Wilcox, N., 2011, 'Stochastically more risk averse:' A contextual theory of stochastic discrete choice under risk. Journal of Econometrics 162, 89-104.

Wilcox, N., 2015. Unusual estimates of probability weighting functions. Chapman University, Economic Science Institute Working Paper #15-10.

Wu, G. and R. Gonzalez, 1999, Nonlinear decision weights in choice under uncertainty. Management Science 45, 74-85.

Appendix: A brief Monte Carlo illustration of the problem (and a possible solution).

Simulated data sets for this brief Monte Carlo analysis of several estimation methods are based on the experimental design of Gonzalez and Wu (1999). Certainty equivalents were elicited from their subjects for $t = 1, 2, \ldots, 165$ distinct two-outcome prospects $L_t = (p_t, h_t; 1 - p_t, l_t)$. These were constructed by fully crossing 15 distinct pairs of high and low outcomes (h_t, l_t) with 11 distinct probabilities p_t of receiving the high outcome h_t (and corresponding probabilities $1 - p_t$ of receiving the low outcome l_t). The probabilities are $p_t \in (.01, .05, .10, .25, .40, .50, .60, .75, .90, .95, .99)$; the high and low outcome pairs are $(h_t, l_t) \in \{(25{,}0), (50{,}0), (75{,}0), (100{,}0), (150{,}0), (200{,}0), (400{,}0), (800{,}0), (50{,}25), (75{,}50), (100{,}50), (150{,}50), (150{,}100), (200{,}100), (200{,}150)\}$. These same 165 prospects are the input to the simulated subjects I create for estimation.

Each simulated subject $s = 1, 2, \ldots 1000$ is given random EU preferences. Each subject s is endowed with parameters k^s and θ^s of the Gamma distribution p.d.f. as given in Example 1 of Section 1. The parameter k^s is drawn *once* for each subject from a Lognormal distribution with mean $E(k) = 0.75$ and variance $\mathrm{V}(k) \approx 0.16$ to provide some subject heterogeneity; the parameter θ^s is then set to the value (given the drawn k^s) such that $C_r^{eu}(e^{-1}|k^s, \theta^s) = (1 + \theta^s)^{-k^s} = e^{-1}$. This endows each random EU agent with a c.e.f. having the fixed point e^{-1} as is characteristic of the 1-parameter Prelec (1998) weighting function, but also creates heterogeneity in the degree of curvature of the subject c.e.f.s. Then, for each subject s, for each prospect t, a new x_t^s is independently drawn from the Gamma distribution with that subject's parameters k^s and θ^s, and these create simulated elicited certainty equivalents $C_t^s = \left[p_t {h_t}^{1/x_t^s} + (1 - p_t) {l_t}^{1/x_t^s}\right]^{x_t^s}$ for Monte Carlo study. This is the "Random EU data."

For comparison it is also useful to have simulated "Standard EU data." To construct this data, each simulated subject s is endowed with a *fixed* value x^s, drawn *once* for each subject from a Gamma distribution with *fixed* parameters $k = 0.75$ and $\theta = 2.79$. The drawn x^s creates expected certainty equivalents $E(C_t^s) = \left[p_t {h_t}^{1/x^s} + (1 - p_t) {l_t}^{1/x^s}\right]^{x^s}$. One may then express this as a proportion of the interval $[l_t, h_t]$, that is as $\Delta_t^s = (E(C_t^s) - l_t)/(h_t - l_t)$, and define parameters of a beta distribution $\alpha_t^s = \varpi\Delta_t^s$ and $\beta_t^s = \varpi(1-\Delta_t^s)$. Draw a beta variate y_t^s using these parameters, and simulated certainty equivalents are $C_t^s = l_t + (h_t - l_t)y_t^s$. I chose $\varpi = 6$ to give the resulting simulated certainty equivalents C_t^s in this Standard EU data conditional variances resembling those found in the Random EU data.

I consider several estimation methods. The first two methods assume a standard RDU model of the c.e.f. of the C_t^s, $E(C_t^s|v^s, w^s) = (v^s)^{-1}\left[w^s(p_t)v^s(h_t) + \left(1 - w^s(p_t)\right)v^s(l_t)\right]$; the corresponding empirical model is then $C_t^s = E(C_t^s|v^s, w^s) + \varepsilon_t^s$. I'll make the standard assumptions about the error, those being $E(\varepsilon_t^s) = E(\varepsilon_t^s|p_t, h_t, l_t) = 0$, but also adopt an assumption of Bruhin, Fehr-Duda and Epper (2010) that $Var(\varepsilon_t^s)$ is proportional to $(h_t - l_t)^2$ for each subject. This all implies that a weighted error ϵ_t^s may be written as $\epsilon_t^s = [C_t^s - E(C_t^s|v^s, w^s)]/(h_t - l_t)$, and the first two estimation methods amount to optimizing some function of these weighted errors.

The first method combines a nonlinear least squares estimator with lean 1-parameter versions of the functions v^s and w^s, $v^s(z) = z^{\sigma^s}$ and $w^s(q) = q^{\gamma^s}/\left[q^{\gamma^s} + (1-q)^{\gamma^s}\right]^{1/\gamma^s}$. This is the estimation method of Tversky and Kahneman (1992): I'll call it NLS-M-L (for "nonlinear least squares, money errors, lean parameterization"). The second method combines a maximum likelihood estimator with the same $v^s(z) = z^{\sigma^s}$ as above and a more expansive 2-parameter weighting function $w^s(q) = \delta^s q^{\gamma^s}/\left[\delta^s q^{\gamma^s} + (1-q)^{\gamma^s}\right]$. The weighted error ϵ_t^s is assumed to have a Normal distribution with zero mean and constant variance. This estimation method is inspired by Bruhin, Fehr-Duda and Epper (2010), but I will always estimate at the individual subject level whereas they estimated finite mixture models of the subject population and included prospect-specific error variance terms (which cannot be done in the case of individual estimation). I'll call this method ML-M-C (for "maximum likelihood, money errors, common parameterization"). The power utility function, combined with some 2-parameter weighting function, is quite common in the literature on risk preference estimation.

The third method writes an estimating equation in utility rather than money terms, and the parameterizations of v^s and w^s are maximally expansive. There are nine distinct outcomes in the experiment, so there are nine distinct values of $v^s(z)$. Since the RDU value function is an interval scale, though, one can choose $v^s(0) = 0$ and $v^s(800) = 1$, leaving seven unique and distinct values of $v^s(z)$ as seven parameters to estimate. Similarly, the eleven distinct probabilities in the experiment become eleven distinct parameters $w^s(p_t)$ to estimate. Now linearly interpolate $v^s(C_t^s)$ from the parameters $v^s(z)$ in the following manner. Let $lub(C_t^s)$ and $glb(C_t^s)$ be the least upper bound and greatest lower bound (among the nine outcomes in the experiment) on C_t^s, and let their values be given by the parameter values $v^s(lub(C_t^s))$ and $v^s(glb(C_t^s))$. Then $\tilde{v}^s(C_t^s) = \frac{[lub(C_t^s)-C_t^s]v^s\left(glb(C_t^s)\right)+[C_t^s-glb(C_t^s)]v^s\left(lub(C_t^s)\right)}{lub(C_t^s)-glb(C_t^s)}$ is a linear interpolation of $v^s(C_t^s)$. This method then assumes that the c.e.f. of $\tilde{v}^s(C_t^s)$ is the RDU of prospect t, that is $E(\tilde{v}^s(C_t^s)|v^s,w^s) = w^s(p_t)v^s(h_t) + \left(1-w^s(p_t)\right)v^s(l_t)$, and one may then think of $\tilde{v}^s(C_t^s) - E(\tilde{v}^s(C_t^s)|v^s,w^s)$ as a "utility error." Following Wilcox (2011), assume the variance of these utility errors is proportional to $[v^s(h_t) - v^s(l_t)\,]^2$. Then $\zeta_t^s = [\tilde{v}^s(C_t^s) - E(\tilde{v}^s(C_t^s)|v^s,w^s)]/[v^s(h_t) - v^s(l_t)\,]$ is a weighted utility error that becomes the object of nonlinear least squares estimation. I call this the NLS-U-E estimation (for "nonlinear least squares, utility errors, expansive parameterization"). It is inspired by Gonzalez and Wu's (1999) estimation method, though there are several differences between their method and this one (see Gonzalez and Wu 199, pp.146-148, for details).

Finally I consider one estimation method that may sidestep the issue identified in the text. Rather than taking $(v^s)^{-1}\left[w^s(p_t)v^s(h_t) + \left(1-w^s(p_t)\right)v^s(l_t)\right]$ to be the conditional mean of C_t^s, this last estimation method takes this to be the conditional median of C_t^s: That is, let $Med(C_t^s|v^s,w^s) = (v^s)^{-1}\left[w^s(p_t)v^s(h_t) + \left(1-w^s(p_t)\right)v^s(l_t)\right]$, and let weighted money errors be $\epsilon_t^s = [C_t^s - Med(C_t^s|v^s,w^s)]/(h_t - l_t)$. Although these errors have exactly the same form as the errors in the first two methods, the fact that we wish to estimate a conditional median function (rather than a c.e.f.) implies that least squares is not the appropriate estimator: Rather, we want a least absolute deviation or LAD estimator. Combined with the same lean parameterization used for the first method, I call this the LAD-M-L estimation (for "least absolute deviation, money errors, lean parameterization").

With the exception of the NLS-U-E estimation method, the well-known simplex algorithm of Nelder and Mead (1965) was used to optimize objective functions. For the NLS-U-E estimation method, I imposed monotonicity constraints on the estimated $v^s(z)$ and $w^s(p_t)$ (one difference versus Gonzalez and Wu 1999) and this requires a different optimization algorithm: Powell's (1992) COBYLA algorithm is used for this estimation instead. All estimations were performed using the SAS procedure "NLP" (nonlinear program) in the 9.4 version of the SAS/OR software.

Rather than providing tabular results of these four estimation methods as applied to the two data sets, I provide a sequence of eight figures. The features of each figure are identical. Estimated weighting functions for the first 250 subjects in each data set are plotted as quite thin, light greyscale lines on a black background: This has the effect of representing the behavior of each method as a light cloud of lines. A heavy light grey identity line shows the (linear, identity) weighting function of an EU agent; deviations from this line represent both sampling variability and possible bias in the estimations. Finally, a heavy dashed white line plots the mean estimated probability weight (across all 1000 subjects in each simulated data set) at each of the eleven values of p_t in the experimental design, illustrating the bias of each estimation method in each data set.

The figures come in pairs on each page that follows. Each page presents the results for one estimation method, with the top and bottom figures showing results for the Standard EU and Random EU data sets, respectively. The pair of Figures A1-a and A1-b show results for the NLS-M-L estimation method; Figures A2-a and A2-b show results for the ML-M-C method; Figures A3-a and A3-b show results for the NLS-U-E method; and Figures A4-a and A4-b show results for the LAD-M-L method.

None of these four estimation methods are bias-free for both the Standard EU and Random EU data sets, and this is the primary finding of this appendix. The method NLS-U-E is biased towards finding inverse-s probability weighting for *both* data sets: In the case of the Standard EU data I suspect this is because this method is just too parametrically expansive for the sample size. By contrast, the NLS-M-L and ML-M-C methods are virtually unbiased for Standard EU data, while they show the predicted bias when applied to the Random EU data. As speculated, the LAD-M-L method provides unbiased (and astonishingly tight) estimates for the Random EU data, but displays a pronounced bias in the Standard EU data in a direction opposite to inverse-s probability weighting. In sum, none of these four estimation methods are robust to the underlying source of randomness in the data generating process.

Figure A1-a: NLS-M-L Weighting Estimates, Standard EU Data

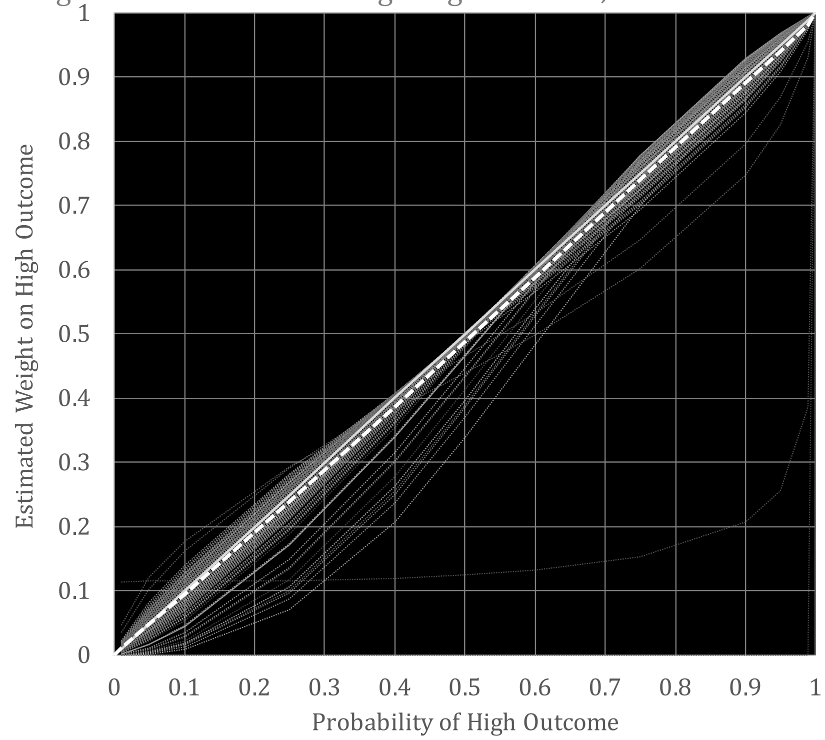


Figure A1-b: NLS-M-L Weighting Estimates, Random EU Data

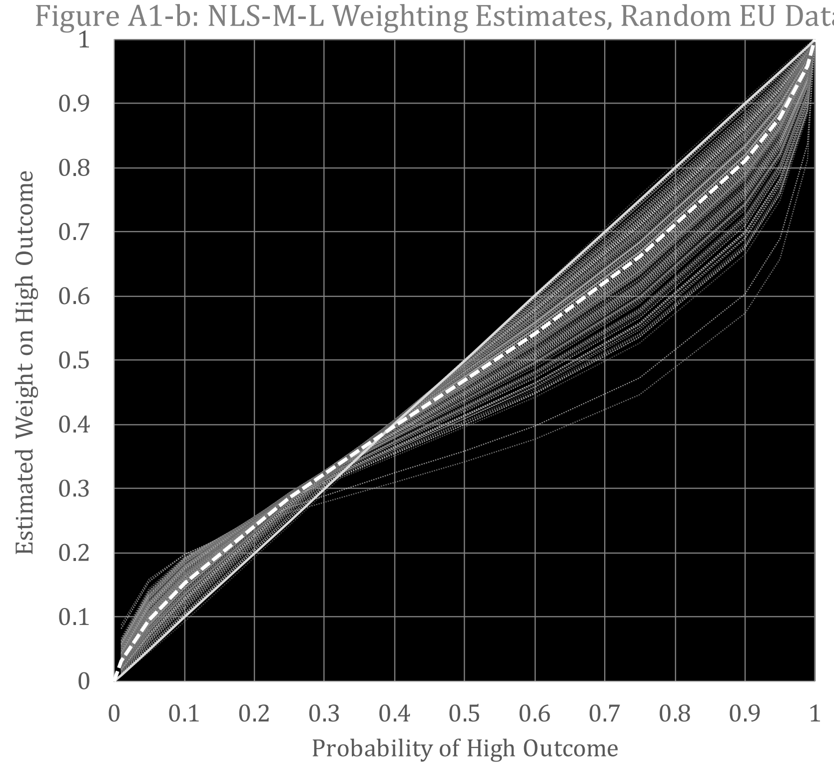


Figure A2-a: ML-M-C Weighting Estimates, Standard EU Data

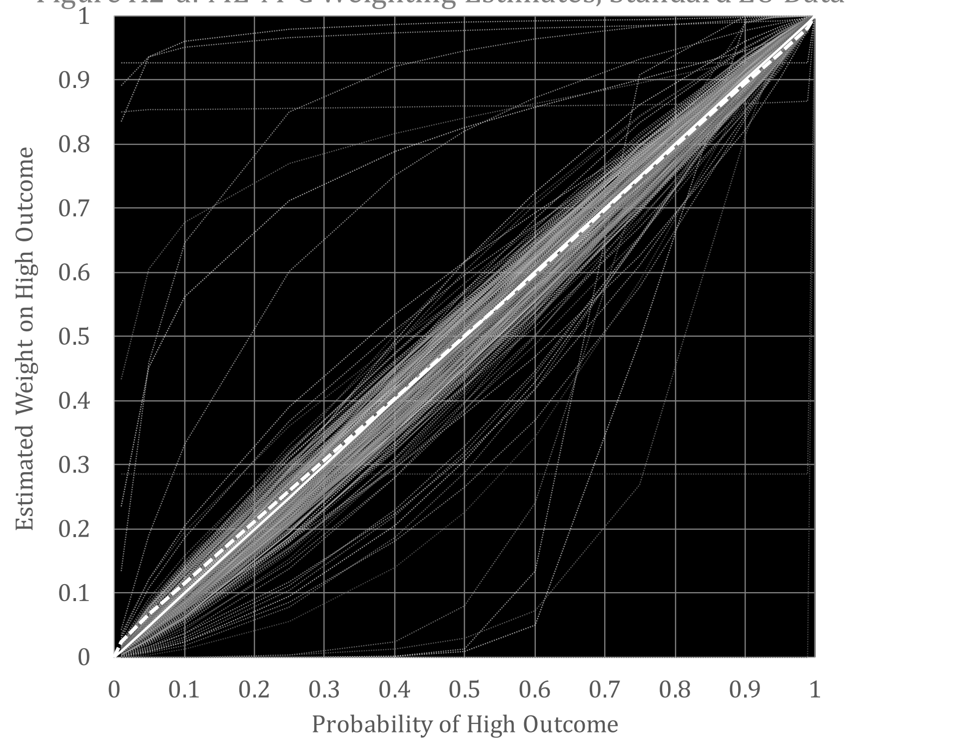
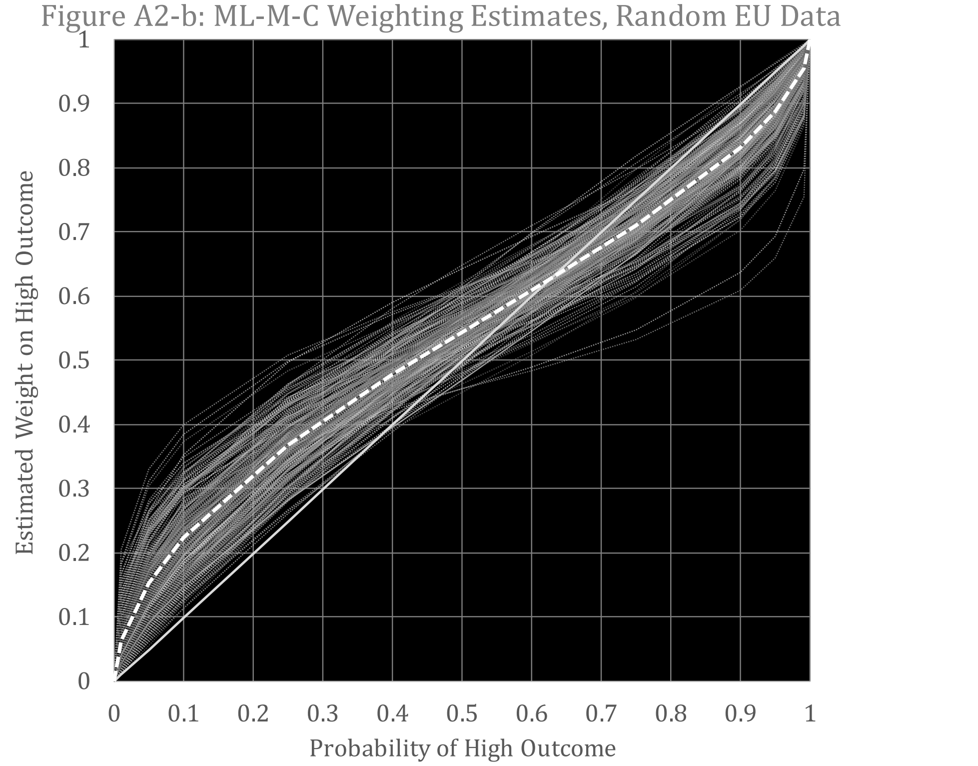


Figure A2-b: ML-M-C Weighting Estimates, Random EU Data



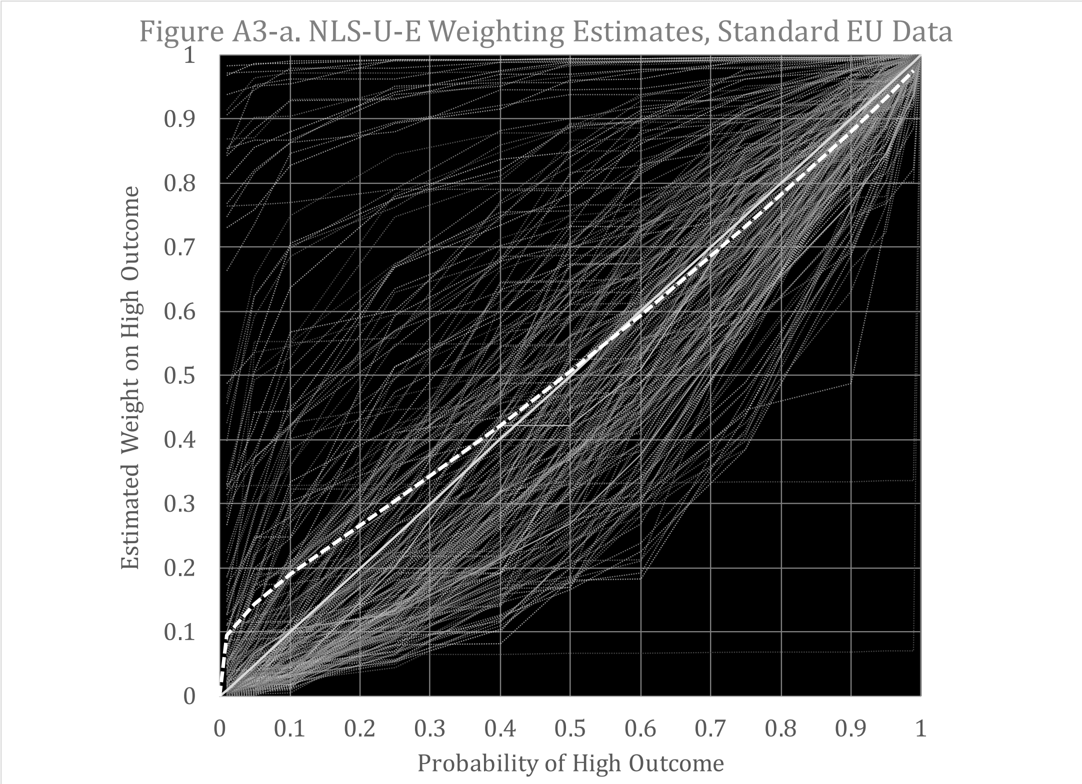


Figure A3-a. NLS-U-E Weighting Estimates, Standard EU Data
Estimated Weight on High Outcome
Probability of High Outcome
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

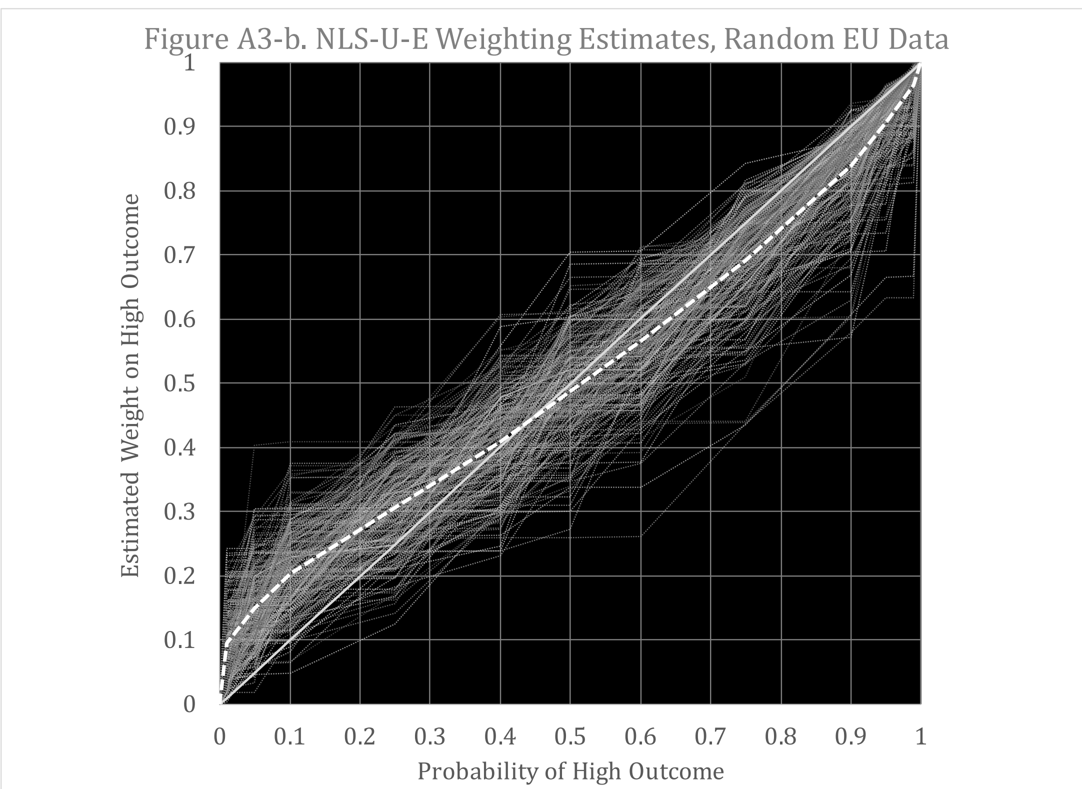


Figure A3-b. NLS-U-E Weighting Estimates, Random EU Data
Estimated Weight on High Outcome
Probability of High Outcome
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Figure A4-a: LAD-M-L Weighting Estimates, Standard EU Data

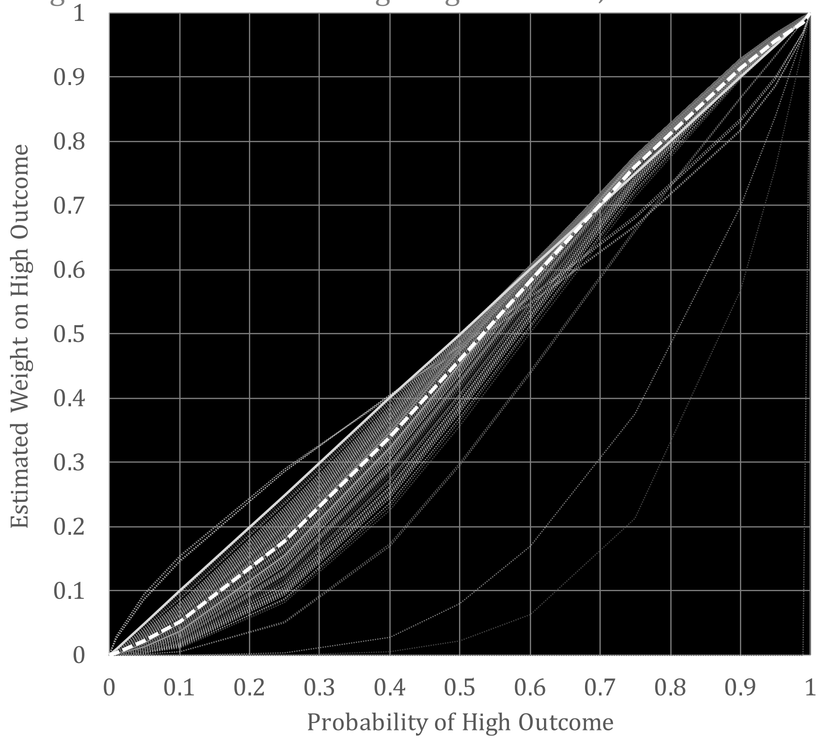


Figure A4-b: LAD-M-L Weighting Estimates, Random EU Data