Evaluating the Normative Coherence of Stochastic Models

Brian Albert Monroe*

February 20, 2023

Abstract

Stochastic models are commonly used to estimate the risk preferences of experimental subjects when the subjects' choices show apparent violations of Expected Utility Theory. While the descriptive properties of these models have been the subject of much investigation, the normative consequences of these models have not been the focus of much discussion. I discuss the normative properties of two popular stochastic models of choice, the Random Utility model and Random Preference model with respect to the evaluation of expected consumer surplus. A thought experiment resembling a classical money pump is proposed to evaluate the two models. Random Utility models make intuitively satisfying normative statements about choice behavior in this thought experiment, while Random Preference models do not. I show that some efforts to alleviate this shortcoming lead to normative under–identification, while others reduce the Random Preference model to a heteroskedastic Random Utility model.

JEL: C59, D60, D81
*Research Unit in behavioral Economics and Neuroeconomics, University of Cape Town.
Email: brianalbertmonroe@gmail.com
I would like to thank Glenn Harrison, Andre Hofmeyr, and Don Ross for providing comments on multiple drafts of this article.

1 Introduction

When presented with the same set of alternatives on multiple occasions, some subjects in incentivized economic experiments do not always choose the same option. In such cases, the weak axiom of revealed preference (WARP) (Samuelson 1938) requires us to infer that subjects must be indifferent between the two options.

When applying WARP, an observation of a choice of A over B is said to reveal that the agent values A at least as much as B. When observing choices from multiple tasks, the same utility function rationalizes all choices. Thus, when choices appear to be inconsistent with a strict preference ordering, one must infer that the agent is indifferent to the options presented to her if WARP is maintained.

For example, when presented with the set of alternatives $X = \{A, B\}$ twice, and we observe the first choice is A and the second is B, we must infer $A \sim B$ under WARP. If two subjects, Anne and Bob, are presented with the same set twenty times and we observe that Anne chooses A for the first nineteen times, and B the twentieth, and Bob chooses B the first nineteen times and A the twentieth, the same inference is made of both Anne and Bob: they are both indifferent between A and B. Stochastic models relax the "interpretive straitjacket" (Wilcox 2008, pp. 198–199) that would require us to infer that Anne and Bob's preferences over A and B are identical.

The ability of various stochastic models to adequately describe choice data has been the focus of much investigation. Wilcox (2008) provides a detailed summary of many of these investigations, and technical descriptions of the properties associated with each of the models. Wilcox (2008) also describes many "intuitively satisfying" *descriptive* predictions of different classes of models, but warns: "If you are waiting for a stochastic model that is intuitively satisfying in every way, you are waiting for Godot. Every stochastic model mutilates your structural intuition in some distinctive way: There is no escape from this."

My goal is to interrogate the extent to which the RU and RP models "mutilate" *normative* intuitions about these models' relationship with the welfare of agents, as opposed to the many *descriptive* peculiarities discussed by Wilcox (2008) and others. I begin by discussing metrics to assess welfare as expected consumer surplus followed by more definitions and descriptions of the two stochastic models. Section 4 discusses how these stochastic models interact with this definition of welfare using a stylized choice problem, and the normative significance of these interactions. I find that the RU model makes coherent normative statements about welfare, but RP models do not. Section 5 discusses how the RP model might be reinterpreted to allow for useful normative statements. This re-interpretation requires that the median, not the mean, of the preference distribution be used for normative analysis, and effectively transforms the RP model into a heteroskedastic RU model.

2 Utility and Welfare

Stochastic models interact with EUT to predict the probability that an option in a set of alternatives is chosen by an agent. The EUT of a lottery A is defined as:

$$\mathrm{EUT}(A) = \sum_{i=1}^{I} p(x_i) \times u(x_i) \tag{1}$$

where i indexes outcomes, x_i, from $\{1, \ldots, I\}$, $u(\cdot)$ is a utility function with the standard properties, and $p(x_i)$ is the probability of outcome x_i.

The utility function $u(\cdot)$ can take many functional forms, and can be normalized in various ways.[1] Unless otherwise specified, however, the constant relative risk aversion (CRRA) function will be assumed throughout the following for clarity:

$$u(x) = \frac{x^{(1-r)}}{(1-r)} \tag{2}$$

where r is the coefficient of relative risk aversion (Pratt 1964).

A single choice scenario or task, represented by t, is a discrete set of mutually exclusive options from which an agent chooses one to consume. I focus on the binary lottery choice problem, where subjects are presented with two lotteries, $\{A, B\}$, and asked to select one for payment. When presented with such a task, a subject's choices are rationalized under WARP by assuming a utility function such that if A is chosen over B, then the utility of A is therefore at least as great as that of B:

$$y_t = A \;\Leftrightarrow\; A \succcurlyeq B \;\Leftrightarrow\; U(A) \geq U(B) \tag{3}$$

where $y_t = A$ is a variable that records which option A is chosen in task t.

For any lottery A, and any r, the certainty equivalent is the certain outcome, CE_A, such that an agent is indifferent between the lottery and the certain amount:

$$A \sim CE_A \;\Leftrightarrow\; U(A) = U(CE_A) \tag{4}$$

[1] For example, Hey and Orme (1994) normalize the utility of the highest outcome across all lottery pairs presented to subjects to be equal to 1, the utility of the lowest outcome to 0, and directly estimated the utility of the interior outcomes.

Combining (1) with the utility function defined in (2), the CE is:

$$CE = \left(\sum_{i=1}^{I} p(x_i) x_i^{1-r}\right)^{1/(1-r)} \tag{5}$$

Thus, if we assume some value of r we can calculate the CE of lottery A for the CRRA utility function.[2] CEs of options in a task usefully allow utility to be normalized to the units of the outcomes. This CE approach is similar to the "money-metric utility" function employed by Samuelson (1974) to calculate welfare. The "money-metric" utility function is used as a normalized utility function by, for instance, Diewert (1983) and King (1983).

The CEs of lotteries can be used to define metrics characterizing the welfare of an agent. If an agent is observed to choose lottery A over lottery B, the difference in value of the chosen lottery and the lottery not chosen can be represented as $CE_A - CE_B$. This is the opportunity cost of choosing CE_A over CE_B and is a measure of the change in welfare:

$$\Delta W = CE_A - CE_B \tag{6}$$

This welfare metric is analogous to the notion of compensating variation in standard consumer theory. If equation (6) is positive, it calculates the minimum amount of money an agent would need as compensation in order to change her choice. If negative, it calculates the maximum an agent should be willing to pay in order to change her choice. Changes in welfare over multiple tasks can be

[2]In general, the CE of any lottery can be calculated with numerical methods even if an analytical solution does not exist because the CE must lie in the interval between the lowest outcome and the highest outcome. Numerically, one can just iterate through this interval until (5) is satisfied, or employ an optimization routine to look for the CE directly.

characterized by summing these compensating variations:

$$\Delta W_T = \sum_{t=1}^{T} (CE_{At} - CE_{Bt}) \tag{7}$$

This method is employed to measure the welfare gains of subjects in experimental settings by Harrison and Ng (2016), Harrison and Ross (2018), Harrison (2019), Harrison, Morsink and Schneider (2020), and Gao, Harrison and Tchernis (2022).

Evaluating the welfare consequences of a set of choices ultimately requires a theory to recommend what an agent *ought* to do, rather than rationalizing what an agent was actually observed doing. What economic agents *ought* to be doing is a subject of intense debate in economics. Most economic theory follows the lead of Samuelson in accepting EUT as the prescription for what agents ought to do, if sometimes begrudgingly, and almost always with the understanding that observed behavior frequently violates EUT.[3]

There are other economic theories of what a rational agent ought to do, many of which are collectively subsumed under the label of "bounded rationality," largely stemming from the work of Simon (1955). Some of these alternatives provide guidance as to how the welfare of agents should be evaluated in their respective, boundedly rational, contexts: see, for example, Salant and Rubinstein (2008), Bernheim and Rangel (2009), and Manzini and Mariotti (2014). Regret Theory (Bell 1982; Loomes and Sugden 1982), provides an alternative to EUT that descriptively accommodates much of the aberrant choice phenomena described

[3]Marschak (1950, p. 127) when detailing axioms of rational choice comparable to the EUT axioms of Von Neumann and Morgenstern (1944), describes a hypothetical agent, "Mr. Smith", who violates the theorems discussed by Marschak (1950). Marschak notes that "[...]the procedure of testing whether Smith is 'tolerably rational' [...] is a statistical one." Marschak maintains the need for a model of deterministic rational choice while recognizing that observed behavior may, even regularly, violate the prescribed model.

by Kahneman and Tversky (1979), and which Loomes and Sugden (1982, pp. 820–822) argue is a normatively coherent, and descriptively attractive alternative to EUT. Other alternatives to EUT intended to accommodate violations of EUT, such as Prospect Theory (Kahneman and Tversky 1979; Tversky and Kahneman 1992), Rank Dependent Utility theory (Quiggin 1982), and the Dual Theory of choice under risk (Yaari 1987) accommodate deviations from EUT while still accepting EUT as normatively correct.[4]

Economists have generally avoided suggesting that agents ought to make choices stochastically,[5] but examples do exist. Fishburn (1978, p. 633) describes the "incremental EU advantage model", which he suggests "can be thought of as a model for 'rational' probabilistic binary choices in risky situations." Machina (1985) avoids labelling his proposed model of stochastic choice from deterministic preferences as being a normative model of choice, but does defend it with regard to supposed normative criteria, such as adherence to stochastic transitivity (p. 582) and path independence (p. 583). More recently, Cerreia-Vioglio, Dillenberger, Ortoleva and Riella (2019) axiomatize a model of "deliberately stochastic" choice which aims to capture and formalize the intuition of Machina (1985). They show that their model explains particular violations of EUT, and prescribes an optimal mixing over available options.

Apesteguia and Ballester (2015) define the "swaps index" to characterize the minimal degree to which observed choices deviate from those predicted by a rational preference relation, the "swaps preference." In this formulation,

[4]Kahneman and Tversky (1979, p. 277) note that the departures from EUT predicted by Prospect Theory, "[...] must lead to normatively unacceptable consequences."

[5]Machina (1985, p. 582), noting this, suggests that economists may be "[...] less confident of our moral certitudes in a world of stochastic behavior."

the swaps preference relation defines explicitly what agents ought to do, and the index defines the degree to which they've deviated from this prescription. The swaps index is therefore taken as a measure of welfare for their stochastic preference relation.

Though the requirements for this preference relation are weaker than those imposed by EUT, in particular the independence axiom is not required, they show (p. 1281) that the swaps preference can easily be extended to characterize EUT, and the swaps index would then reflect the extent of deviations from an EUT preference relation. Apesteguia and Ballester (2015, pp. 1288–1289) additionally show how the RU and Tremble stochastic models can be derived from the swaps preference. Notably absent however, is the ability to characterize the RP stochastic model with the swaps preference.

The extent of the differences between traditional rational choice models and alternative models of rational choice for agents, and the validity of the arguments comparing and contrasting these theories, is beyond the scope of the arguments I present here. The RU and RP models discussed subsequently are typically used in the "neo–Samuelsonian" (Ross 2014, 2021) framework of rationality in which EUT is taken as the normative reference, so no adjudication of competing theories of rationality is necessary for the following discussion. When stochastic models are intended to enhance the descriptive validity of the neo–Samuelsonian framework which characterizes EUT, they must make normatively consistent and coherent prescriptions within this framework. I analyze the extent to which this occurs with the RU and RP models.

3 Stochastic Models

Mosteller and Nogee (1951) note that randomness in observed choices is a well-known empirical regularly in the psychology literature; see, for example, Edwards (1954). They propose (p. 374) a probabilistic interpretation of indifference as the selection of alternatives with equal probability. This early accommodation of randomness in choice data was a precursor to more formal treatment of the phenomena in economics by Quandt (1956), Luce (1958, 1959), and Becker, DeGroot and Marschak (1963) (BDM).

BDM discuss three versions of stochastic choice models: the Random Utility model, the Luce or Strict Utility model, and the Fechner or Strong Utility model. The Luce utility model is derived from the work of Luce (1959) while the Fechner utility model is related to the psychophysics work of Fechner (1966).[6] The Luce and Fechner models closely resemble latent index models well known to economics. What BDM call the Random Utility model closely resembles what is sometimes called a "mixed logit" (Andersen et al. 2012), or "random parameters" model.

There is some confusion when discussing naming conventions of stochastic models. I follow the semantic distinction used by Wilcox (2008), and Apesteguia and Ballester (2018), and refer to the Fechner class of models as Random Utility (RU) models, and the successors of BDM's random utility model as Random Preferences (RP) models. This convention is *not* followed by Cerreia-Vioglio, Dillenberger, Ortoleva and Riella (2019, p. 2426) who discuss random utility/preferences as referring to the RP class of models, and RU models as models of "bounded rationality" or "mistakes," and Gul and Pesendorfer (2006) who refer to their

[6] The interpretation of stochastic choice by Mosteller and Nogee (1951) follows this psychophysical tradition.

model of probability measures over utility functions as a random utility model. RU models are *not* models of bounded rationality in the way that term is used by proponents of the work of Simon (1955), but they *are* models of mistakes, as will be discussed subsequently.

RU and RP models have been used extensively in the economics literature to accommodate randomness in data from choice under risk experiments. The RU model is ubiquitous in this literature; see Hey and Orme (1994), Hey (1995, 2001), Holt and Laury (2002), and Harrison and Rutström (2008) for a more general review of choice under risk. Loomes and Sugden (1995) reformulated and generalized BDM's Random Utility model and reintroduced it as the RP model, and versions of it have been employed by Loomes, Moffatt and Sugden (2002) and Wilcox (2008). Additionally Gul and Pesendorfer (2006) propose a version of the RP model that they call the Random Expected Utility model.

3.1 Random Preference Models

The RP model characterizes each observed choice made by an agent as conforming to a deterministic preference relation which is drawn at random from a set of relations whenever the agent is confronted with a choice task. While the set of preference relations can have a discrete distribution, the RP model is most commonly discussed in terms of utility functions with the relevant parameters being continuously distributed according to some cumulative distribution function (cdf), $F(r \mid \alpha)$, where α represents sufficient parameters defining the cdf. Thus, the probability of choice in a RP model is simply the probability that a value of r is drawn from the distribution $F(r \mid \alpha)$ that would deterministically satisfy that choice:

$$\Pr(y = A) = \Pr\left(r \mid U(A \mid r) \geq U(B \mid r)\right) \tag{8}$$

Letting $\Omega \equiv \{r \mid U(A|r) \geq U(B|r)\}$, and the density function $f\,(r \mid \alpha)$:

$$\Pr(y = A) = \int_{r\in\Omega} f\,(r \mid \alpha)\, dr \tag{9}$$

An alternative formulation for binary lottery pairs with 3 outcomes is discussed by Loomes, Moffatt and Sugden (2002, pp. 113–114) and again by Wilcox (2008, pp. 214–215).[7] Consider two lotteries, A and B, each with three possible outcomes such that $u_1 > u_2 > u_3$, $U(A) = \sum_{i=\{1,2,3\}} p_i u_i$, and $U(B) = \sum_{i=\{1,2,3\}} q_i u_i$, where p_i is the probability of outcome i in lottery A, and q_i is the probability of outcome i in lottery B. Equation (8) can then be reformulated as:

$$\Pr(y = A) = \Pr\left(u_1, u_2, u_3 \,\middle|\, \sum_{i=1,2,3} p_i u_i - \sum_{i=1,2,3} q_i u_i \geq 0\right) \tag{10}$$

Assuming a strictly increasing u, and setting $v = ({u_1 - u_2}/{u_2 - u_3})$, this expression can be written as:

$$\Pr(y = A) = \Pr\left(u_1, u_2, u_3 \,\middle|\, -\left(\frac{(p_2 - q_2) + (p_1 - q_1)}{(p_1 - q_1)}\right) \geq v\right) \tag{11}$$

$$\Pr(y = A) = F\left(-\left[\frac{(p_2 - q_2) + (p_1 - q_1)}{(p_1 - q_1)}\right] \,\middle|\, \alpha\right) \tag{12}$$

where $F(\cdot)$ is some cdf of v with support over $(0, \infty)$ and α is the vector of sufficient statistics for the distribution. The random variable v contains all utility–relevant information.[8] The formulation in (12) is useful because well-

[7] Loomes, Moffatt and Sugden (2002) and Wilcox (2008) discuss this formulation in terms of Rank Dependent Utility (Quiggin 1982), and both restrict the probability weighting parameters to be non-random.

[8] This can be seen by noting that u_1 and u_3 can be arbitrarily set to 1 and 0, respectively, since utility is unique up to a positive affine transformation. The random variable v then

known distributions, such as the log-normal or gamma distributions, can characterize the probability of a choice as a function only of the probabilities of the outcomes in the lotteries.[9] It is limited in that it only applies to lottery pairs with 3 outcomes, but this is not particularly unusual; Hey and Orme (1994), Hey (2001), and Harrison and Rutström (2009) all use 3 outcome lottery pairs.

3.2 Random Utility Models

RU models are commonly formulated as having utilities of different options in a set of alternatives determined by some "core" utility theory, and disturbed by a median 0 random error term, ϵ. A choice is characterized as having incorporated this error. Assuming a binary choice scenario, the choice of option A is characterized as:

$$\begin{aligned} U(A) + \epsilon_A \ &\geq \ U(B) + \epsilon_B \\ [U(A) + \epsilon_A] \ - \ [U(B) + \epsilon_B] &\geq 0 \end{aligned} \tag{13}$$

Setting $\epsilon_A - \epsilon_B = \epsilon\lambda$, where λ is proportional to the standard deviation of ϵ, we can rewrite (13) as:

$$\begin{aligned} U(A) - U(B) + \epsilon\lambda \geq 0 \\ \epsilon \geq \frac{1}{\lambda}\left[U(B) - U(A)\right] \end{aligned} \tag{14}$$

Thus for RU models, the probability option A is chosen is given by:

$$\begin{aligned} \Pr(y = A) &= \Pr\Big(\epsilon \geq \frac{1}{\lambda}\left[U(B) - U(A)\right]\Big) \\ &= 1 - F\left(\frac{U(B) - U(A)}{\lambda}\right) \end{aligned} \tag{15}$$

collapses to ${}^{1-u_2}/_{u_2}$ with $u_2 \in (0, 1)$.

[9] Loomes, Moffatt and Sugden (2002) use the log-normal distribution to model v.

where F is a cdf with median at 0. As λ approaches 0, choice probabilities approach 0 or 1, while as λ approaches ∞, choice probabilities approach 0.5. Usually $F(\cdot)$ is taken to be either the normal or logistic function, but this is not necessary.

Several derivatives of the RU model further adjust the λ parameter to introduce heteroscedasticity to the error term. For example, the Contextual Utility model of Wilcox (2011) multiplies λ by the difference between the utility of the maximum and minimum outcomes in $\{A, B\}$. Stronger Utility as proposed by Blavatskyy (2007, 2014) is an extension of the "incremental EU advantage" model by Fishburn (1978), which adjusts the λ term such that lotteries that first order stochastically dominant (FOSD) all other lotteries in the set of alternatives are chosen with a probability of 1.[10]

3.3 Descriptive Comparisons

The discussion of stochastic models makes it clear that they operationalize stochasticity in different ways. These differences can amount to identifying restrictions (Ballinger and Wilcox 1997; Wilcox 2008) on the stochastic model or the underlying utility structure defining the risk preferences model. I briefly note some of the descriptive consequences of the differences in the RU and RP models; see Wilcox (2008, pp. 208–212) for more details.

[10]Lottery A is said to FOSD lottery B *iff*:

$$\forall\, u_i,\ \sum_i^I p_i \ \geq\ \sum_i^I q_i \quad \textit{and} \quad \exists\, x_i,\ \sum_i^I p_i \ >\ \sum_i^I q_i$$

where p_i represents to probability of outcome i in lottery A, q_i represents to probability of *the same* outcome i in lottery B, and the index i ranks the outcomes of lotteries A and B from lowest to highest. All deterministic theories of utility require the dominant option to be chosen over the dominated option.

Stochasticity is introduced in the RP model by the random draws of preference relations from a preference distribution, $F(\cdot)$. An obvious consequence of this is that even though choices *across* tasks can fail to conform to individual preference relations, the choice *within* a task must conform to some preference relation. The RP model therefore does not descriptively allow for violations of FOSD. Should A FOSD B, there is no EUT compatible preference relation that will allow $U(B) > U(A)$. Thus, if we observe a choice of B in such a scenario, the RP assigns a probability of 0 to the choice.

Violations of *transparent* FOSD are relatively rare, but they do occur. For example, the lottery pairs used by Loomes and Sugden (1998, p. 591) have ten choice problems where one option FOSD the other, and several violations of FOSD are observed. Likewise, Hey (2001, p. 14) includes 5 choice problems where one lottery dominates the other, and observes several violations of FOSD. Rather than discard these data, the RP model is generally combined with the Tremble model (Harless and Camerer 1994) to allow for violations of FOSD. By contrast, the RU model predicts violations of FOSD with positive probabilities.

Another difference in descriptive predictions between the RP and RU model is their treatment of “common ratio effects” (Allais 1953; Kahneman and Tversky 1979). Consider two lotteries, A and B, comprised of outcomes j, k, and l, such that:

$$\begin{aligned} EU(A) &= (1-\tau r)\times u_j + \tau r \times u_k \\ EU(B) &= (1-\tau s)\times u_j + \tau s \times u_l \\ & j<k<l \\ & 0 \le r < s \le 1 \\ & 0 \le \tau \le 1 \end{aligned} \tag{16}$$

It must be the case that either $EU(A) \le EU(B)$ or $EU(B) \le EU(A)$ for all values of τ. This is easy to see by taking the difference between $EU(A)$ and $EU(B)$ and noting that this difference is sign-independent of τ:

$$\begin{aligned} EU(A) - EU(B) = & \\ &=(1-\tau r)\times u_j + \tau r \times u_k - (1-\tau s)\times u_j - \tau s \times u_l \\ &=\tau\left[(s-r)\times u_j + r\times u_k - s\times u_l\right] \end{aligned} \tag{17}$$

The value of τ does not change the sign of the difference in utilities. The RP model therefore requires that $\Pr(A)$ must be the same for all values of τ, and for any set of utilities u_j, u_k, and u_l. This is not the case for the RU model. For a given value of λ, lower values of τ would result in lower $\Pr(A)$.

The RP model has other descriptive qualities. Apesteguia and Ballester (2018) note that the RP model conforms to a notion of "stochastic monotonicity," by which they mean that as the density of the preference distribution shifts to ranges of parameter values that imply greater risk aversion,[11] the probability of a "risky" option being selected over a "safe" option monotonically decreases. RU models, in general, do not share this property. However, the Contextual

[11]For the CRRA parameter in (2), and maintaining EUT, this means greater density on higher values of r.

Utility model of Wilcox (2011) *does* share this property for *certain kinds* of lottery pairs. Wilcox (2011) calls this property the "stochastically more risk averse than" relation.

4 The Stochastic Money Pump

With the definitions of welfare and stochastic models in hand, I begin the discussion of the welfare implications of these models by first introducing a decision problem which resembles the money pump argument against intransitive structures in deterministic choice theory. For this "stochastic money pump", assume two binary lottery pairs, $A = \{X, Z\}$ and $B = \{Z, Y\}$, over outcomes $\{\$5, \$18, \$19, \$20\}$. The probability distributions over these outcomes are shown in Table 1.[12]

Table 1: Probability Distributions for Lotteries

	Probability of outcomes			
	\$5	\$18	\$19	\$20
X	0	0	1	0
Y	0	1	0	0
Z	0.1	0	0	0.9

Table 2: Possible Choice Patterns and Implied Preference Relations

Choice Set	Preference Ordering
$\{X, Z\}$	$X \succeq Z \succeq Y$
$\{X, Y\}$	$X \succeq Y \succeq Z$
$\{Z, Z\}$	$Z \succeq X \succeq Y$
$\{Z, Y\}$	$Y \sim Z \sim X$

There are four possible choice patterns that can be observed from these two lottery pairs: $\{X, Z\}$, $\{X, Y\}$, $\{Z, Z\}$, and $\{Z, Y\}$. Under WARP, these choice

[12]These lottery pairs are taken from from the insurance task of Harrison and Ng (2016, p. 100).

patterns imply the preference orderings shown in Table 2. For the choice pattern $\{X, Y\}$ we might also infer that it is possible that $Y \succeq X \succeq Z$, and likewise for choice pattern $\{Z, Z\}$ that the ordering $Z \succeq Y \succeq X$ is possible. If X and Y represented arbitrary lotteries, this would be the case, however, X and Y represent degenerate lotteries and so $Y \succeq X$ would violate the monotonicity axiom on $\succeq$ unless $X \sim Y \sim Z$.

The first three choice patterns listed in Table 2 can all be rationalized without resorting to declaring the agent indifferent to all options. Of interest here is the fourth choice pattern, $\{Z, Y\}$. Given X FOSD Y, under WARP we must infer $X \sim Y \sim Z$. This is deeply unsatisfying as it implies \$19 $\sim$ \$18.[13] As with the traditional money pump thought experiment, any agent that repeatedly trades a certain amount of money for a lottery, and then trades that lottery for a smaller certain amount of money will face economic elimination. Thus this choice pattern is dominated by the alternative patterns.

Suppose we observe an agent choosing Z over X 38% of the time in pair A, and choosing Y over Z 39% of the time in pair B. For any given set of choices, there is a 15% chance that this agent will have lost \$19 $-$ \$18 = \$1 in opportunity cost by choosing Z over X and then choosing Y over Z. These observed choice probabilities can be accommodated well by both the RU and RP models, while having different implications for the welfare of the agent. The RU stochastic model in (14) with a CRRA parameter of 0.03, and the RP model with the CRRA parameter following a normal distribution with mean -0.14 and standard

[13]Note that in experimental contexts, this kind of revealed indifference may be explained as a failure to satisfy the dominance precept of Smith (1982). See Harrison (1989, 1992) for a discussion of the dominance precept and its importance for drawing valid inferences from experimental data.

deviation 2.52, produce identical choice probabilities.[14] We can calculate the choice probabilities for the RP model as follows:

$$
\begin{aligned}
A_{ZX} &= \{r \mid U(Z|r) \geq U(X|r)\} \\
&= \{r \mid r \leq -0.90\} \\
B_{YZ} &= \{r \mid U(Y|r) \geq U(Z|r)\} \\
&= \{r \mid r \geq 0.55\} \\
\Pr(y_A = Z) &= \int_{r \in A_{ZX}} f\left(r \,\middle|\, \mu, \sigma^2\right) dr \\
&= \phi\left(-0.90 \mid -0.14, 2.52\right) \\
&\approx 0.38 \\
\Pr(y_B = Y) &= \int_{r \in B_{YZ}} f\left(r \,\middle|\, \mu, \sigma^2\right) dr \\
&= 1 - \phi\left(0.55 \mid -0.14, 2.52\right) \\
&\approx 0.39 \\
\Pr(\{Z, Y\}) &\approx 0.15
\end{aligned}
\tag{18}
$$

where ϕ is the cumulative normal distribution. The value of the CRRA parameter that makes the agent indifferent between Z and X is -0.90, and likewise the CRRA parameter that makes the agent indifferent between Y and Z is 0.55. Any CRRA value below -0.90 would induce a choice of Z over X, and any value above 0.55 would induce a choice of Y over Z. Thus the choice probabilities in

[14] In (18) and (19), the parameter values and resulting choice probabilities are rounded for exposition. Taking any r value for the RU model as given, one can calculate the choice probabilities for pairs A and B. For the RP model, given the nature of the choice problem, there exists only one value of r which would render an agent indifferent to the lotteries in A (r_A) and likewise for pair B (r_B). A continuous distribution with two sufficient statistics can be found such that the cumulative densities below the r_A and above r_B are equal to the choice probabilities produced by the RU model. Thus, for this two-choice problem, RU and RP models can always be defined to be descriptively equivalent.

(18) are determined by the cumulative density of the agent's preference distribution below and above these values, respectively.

We can likewise calculate the choice probabilities for the RU model:

$$
\begin{aligned}
\Pr(y_A = Z) &= \frac{\exp(U(Z|r=0.03))}{\sum_{W=\{Z,X\}} \exp\left(U(W|r=0.03)\right)} \\
&\approx 0.38 \\
\Pr(y_B = Y) &= \frac{\exp(U(Y|r=0.03))}{\sum_{W=\{Y,Z\}} \exp\left(U(W|r=0.03)\right)} \\
&\approx 0.39 \\
\Pr(\{Z,Y\}) &\approx 0.15
\end{aligned} \tag{19}
$$

While the observed choice behavior is identical for both models, the welfare implications are not. Using the metric defined in (6), the change in expected consumer surplus of the observed choices can be characterized using the RU model as follows:

$$
\begin{aligned}
\Delta W_A &= CE_Z - CE_X = \$18.48 - \$19 = -\$0.52 \\
\Delta W_B &= CE_Y - CE_Z = \$18 - \$18.48 = -\$0.48 \\
\Delta W_A + \Delta W_B &= -\$1.00
\end{aligned} \tag{20}
$$

In this case, the welfare implications of these choices are clear: with a 0.15 probability, the agent makes 2 consecutive "mistakes" or "choice errors" which results in a loss of \$1.00 in expected consumer surplus to the agent. The RU model provides a coherent normative interpretation of the two choices: the choice of a dominated choice pattern reduces the welfare of the agent, and she is worse off for having made these two choices. It also provides a useful prescription for an agent willing to correct her choices: she ought to choose $\{X, Z\}$.

The RP model, however, does not provide a similarly intuitive understanding of the welfare implications of this decision problem. The RP model requires every choice made by the agent to be characterized by a deterministic preference relation drawn at random from a distribution. Thus, the welfare change of each choice made by the agent must be weakly positive:

$$\begin{aligned} \Delta W_A &= CE_Z - CE_X \geq 0 \\ \Delta W_B &= CE_Y - CE_Z \geq 0 \end{aligned} \tag{21}$$

According to the metric defined in (6) and the decision process for the RP model defined in (18), the ΔW welfare evaluations in (21) must be weak inequalities. However, with a probability approaching 1, the RP model suggests that the subject has strictly improved her welfare by making a pattern of choices that is strictly dominated by other patterns. Indeed the RP model suggests the expected welfare change is:

$$\begin{aligned} E\left(\Delta W_A\right) &= \int_{r \in A_{ZX}} \left(CE_Z - CE_X\right) f\Big(r \;\Big|\; \mu, \sigma^2\Big)\, dr = \$0.66 \\ E\left(\Delta W_B\right) &= \int_{r \in A_{YZ}} \left(CE_Y - CE_Z\right) f\Big(r \;\Big|\; \mu, \sigma^2\Big)\, dr = \$1.86 \\ E\left(\Delta W_A\right) + E\left(\Delta W_B\right) &= \$2.52 \end{aligned} \tag{22}$$

On average, the agent has an expected consumer surplus *increase* equivalent to \$2.52, and would need to be *paid* at least \$0.66 to change her choices. The RP model therefore does not lead to a coherent interpretation of the welfare of an agent that has chosen a dominated choice pattern.

5 Random Preferences as Random Parameters

We may consider the problem of the RP model implying welfare gain from dominated choices simply an issue of how we frame the RP model. The RP model used as an example here has a well-defined distribution, and therefore we may consider the “core” preference of the subject to be some representative point on this distribution, e.g. the mean, median, or mode. Draws from the preference distribution are reformulated as perturbations of this point by n random variable, ϵ, with a median of 0 and a standard deviation equal to the standard deviation of the preference distribution. The ϵ deviations from the representative point would not be used when making inferences about the welfare of the agent, just as the error is discarded during welfare evaluation for the RU model.

This is essentially how Apesteguia and Ballester (2018, p. 78) discuss the RP model, although their focus is on particular econometric and descriptive differences of the RP and RU models, not normative differences. This recasts the Random Preference model as a Random Parameters model, well known to economists.[15]

This recasting of the RP model could provide a solution to the problem in (21) by evaluating the welfare of both choices using the mean of the RP distribution, $r = -0.14$. Under this formulation, welfare changes for each individual choice made by the agent would be modestly different from those given by the RU model in (20), but the joint welfare change would be an identical loss of \$1.00.

[15] Helpfully, we can still refer to this as the RP model.

However, the choice of which point on the preference distribution to represent the "core" preference is not obvious. Each choice implies some restriction on the estimation of preferences to accommodate normative inference. I discuss some issues and surprising consequences of choosing a point on the preference distribution as a "core preference".

5.1 Normative Variability to Transformations in Parameters

Using the mean or mode of the RP distribution to evaluate the welfare of an agent results in variability of normative prescriptions to transformation of the *parameters* of the utility function. Thus, this method yields normative prescriptions that are under–defined.

First, note that the CRRA utility function in (2), $u(x|r)$, can be harmlessly rewritten as:

$$h(x|s) = \frac{x^{1-g^{-1}(s)}}{1-g^{-1}(s)} \tag{23}$$

where $g(\cdot)$ is some continuous, invertible function and $g^{-1}(\cdot)$ is its inverse. For all $g(\cdot)$, r, x, (23) implies $h(x|g(r)) = u(x|r)$; $h(x|g(r))$ is an isomorphic transformation of $u(x|r)$. Using $h(x|g(r))$ in place of $u(x|r)$ clearly makes no difference for the RU model either when calculating choice probabilities, or when calculating welfare in (19). Both functions produce *identical* cardinal utilities for any $g(\cdot)$, r, and x.

For the RP model, the distribution of r must be transformed using the function $g(\cdot)$. If X is a random variable with density f_x, and $Y = g(X)$, then the density of Y is:

$$f_Y(y) = \sum_{x \in g^{-1}(y)} \frac{f_X(x)}{|g'(x)|} \tag{24}$$

where $g^{-1}(y)$ is the inverse function of $g(\cdot)$ when $g(\cdot)$ is monotone, in which case $g^{-1}(y)$ is a singleton, and is the correspondence of the inverse of $g(\cdot)$ when $g(\cdot)$ is non-monotone, in which case $g^{-1}(y)$ is a set.[16] The mean and variance of the transformed distribution are:

$$\begin{aligned} \omega &= E\left(g(X)\right) = \int g(x) f(x) dx \\ \varsigma^2 &= E\left(\left(g(X) - \omega\right)^2\right) = \int \left(g(x) - \omega\right)^2 f(x) dx \end{aligned} \tag{25}$$

Descriptively, the transformation of the preference distribution does not influence predicted choice probabilities for *any* choice problem. However, even with the choice of a globally monotone $g(\cdot)$, $g^{-1}(\omega)$ does not always equal the mean of the untransformed distribution, μ. For example, let $g(r) = \exp(r)$, and therefore $g^{-1}(r) = \ln(r)$. This choice of $g(\cdot)$ transforms the preference distribution for the RP model in (18) from a normal distribution with mean $\mu = -0.14$ and standard deviation $\sigma = 2.52$ to a log–normal distribution with mean $\omega = 20.89$ and standard deviation $\varsigma = 503.75$.

Table 3: Welfare Using the Mean of the Preference Distribution

Utility Function	ΔW_A	ΔW_B	ΔW_{Total}
$u(x\|r = -0.14)$	$-\$0.52$	-$0.48	$1.00
$h(x\|s = 20.89)$	$-\$6.46$	$5.46	$1.00

The ΔW of decision A using $u(x|r = \mu)$, is -$0.52, and the ΔW of the A decision using $h(x|s = \omega)$, is -6.46. Likewise, the ΔW of decision B using $u(x|r = \mu)$ is -$0.48, and the ΔW of the B decision using $h(x|s = \omega)$ is 5.46. Both $u(\cdot)$ and $h(\cdot)$ produce $\Delta W_A + \Delta W_B = -\$1$, which usefully describes the

[16] The pedagogic example of this relationship is to assume $X \sim N(\mu, \sigma)$, and $y = x^2$. In this case, $g(x)$ is non-monotone, with $g^{-1}(y) = \pm\sqrt{y}$. The summation in (24) is over both the positive and negative values of $\sqrt{y}$. This results in Y being χ^2 distributed.

dominated choice pattern as resulting in the same strict loss of welfare, but the choice recommendations for the two utility functions are not the same.

This transformation of the utility function produces a RP model that is *descriptively identical*, but recommends a different choice correction. If isomorphic transformations of the *parameters* in the model result in different normative prescriptions, it is possible to arbitrarily transform the utility function to get any desired prescription for a given choice, thus the *normative* content of this model is under–identified.

However, using the median of the preference distribution as the core preference does not result in the under–identification issue discussed here. The median of an untransformed distribution will always map directly to the median of the transformed distribution. This direct mapping results in consistent normative prescriptions regardless of the choice of $g(\cdot)$.

5.2 Duality of Random Parameters and Random Utility

An implication of choosing a point from the preference distribution to represent the core preference is the effectual transformation of the RP model into an RU model with a heteroskedastic error term. The duality of random parameters models and random utility models is discussed by McFadden and Train (2000) who discuss the regularity conditions under which any RU model can be expressed as a RP model.[17] We can likewise transform an RP model into a (heteroskedastic) RU model.

As a simple example assume a 3-outcome lottery pair, such as decision A in the previous stochastic money pump example. Because utility is unique up to an

[17]McFadden and Train (2000) discuss what is commonly called Random Parameters as random coefficients.

affine transformation, we can arbitrarily set the utility of the smallest outcome equal to 0, the greatest outcome equal to 1, and concern ourselves only with the utility of the intermediate outcome. Since this intermediate utility must be bound between 0 and 1, assume that it is logit-normal distributed.[18]

$$
\begin{aligned}
u_1(\$5) &= 0 \\
u_2(\$19) &\sim \text{Logit-Normal}(\mu, \sigma) \\
u_3(\$20) &= 1
\end{aligned}
$$

Each value of u_2 drawn from the preference distribution corresponds to a difference in the expected utility of the lottery ticket and the sell price in decision A:

$$
\delta = g(u_2) = \sum_{i \in I} (p_i - q_i) u_i \tag{26}
$$

$$
\begin{aligned}
&= (p_1 - q_1) \times u_1 + (p_2 - q_2) \times u_2 + (p_3 - q_3) \times u_3 \\
&= (p_3 - q_3) + (p_2 - q_3) \times u_2
\end{aligned} \tag{27}
$$

$$
\begin{aligned}
&= (0.5 - 0) \times 0 + (0 - 1) \times u_2 + (0.5 - 0) \times 1 \\
&= 0.5 - u_2
\end{aligned} \tag{28}
$$

where p_i and q_i are the probabilities associated with each outcome in decision A and u_i are the associated utilities. Equation (27) shows the general result for a 3-outcome lottery pair, and (28) shows the specific result for this example. If we take the median of the preference distribution to be the "core" preference,[19]

[18]Note that these parameters define the mean and standard deviation of the underlying Normal distribution.

[19]The discussion on variability to transformation precludes the use of the mean as the core preference.

then we further adjust these equations by the median to get the resulting error function for the RU model:

$$\begin{aligned} \epsilon &= (p_3 - q_3) + (p_2 - q_3) \times u_2 - M \\ &= 0.5 - u_2 - M \end{aligned} \tag{29}$$

where M is the median of the preference distribution. In the case of the Logit-Normal distribution, $M = 1/(1 + \exp(\mu))$. Making use of equation (24), the density function of $f(\epsilon) = f_{u_2}(g^{-1}(\epsilon))$, where $g^{-1}(\epsilon) = (\epsilon - p_3 + q_3 + M)/(p_2 - q_3)$ in general and $g^{-1}(\epsilon) = 0.5 - \epsilon - M$ for the A decision.

The duality of the RU and RP models shown by the above exercise demonstrates an underlying consequence of correcting the RP model such that it is normatively coherent: the correction implies a heteroskedastic RU model with no *random* preferences involved. This puts the RP model on the same footing as other heteroskedastic RU models like Contextual Utility (Wilcox 2011), which also enforces stochastic monotonicity, and Stronger Utility (Blavatskyy 2007) which enforces zero probability FOSD.

6 Discussion

Stochastic models of economic choice have received an increasing amount of attention. Two classes of models in particular have been the focus of much scrutiny: the RU model and the RP model. The primary purpose behind the development of these models was to descriptively account for choices which were inconsistent with a deterministic interpretation of EUT, without having to resort to declaring an economic agent indifferent to every alternative. To further this descriptive objective, these models were formulated in such a way as to make specific predictions about observed choice probabilities in particular choice scenarios.

My purpose here, however, is not to investigate the descriptive veracity of these models, but to draw attention to their normative implications and shortcomings. I define metrics to evaluate the welfare consequences of choices following the existing literature, and propose a thought experiment involving an example decision problem, the "stochastic money pump," to evaluate the normative coherence of the RU and RP models. The stochastic money pump is structured in a way to demonstrate the possibility of an agent entering into a series of decision problems with the possibility of selecting a strictly dominated pattern of choices, much like the traditional money pump thought experiment. The RU and RP models can be parameterized in such a way that they produce *exactly* the same choice probabilities for this decision problem. This descriptive equality, however, does not imply that the welfare implications of these models are always equivalent, or even coherent.

The RP model, taken at face value, suggests that a choice pattern that is dominated by alternative choice patterns provides the agent with a positive expected consumer surplus. This characterization of the expected consumer surplus of dominated options is economically incoherent. There must be an opportunity cost of selecting a dominated option. The RP model also fails to provide normative guidance: since every choice is said to reflect the result of a rational preference relation, the agent is therefore always doing what they *ought* to be doing. The same cannot be said of the RU model, which characterizes a dominated choice pattern as resulting in loss of expected consumer surplus, and provides a coherent normative prescription to avoid such losses.

The inability of the RP model to provide a coherent normative interpretation of choice patterns is because of its lack of a consistent, “core” preference across decision tasks. This, however, is by design. When reintroducing the RP model, Loomes and Sugden (1995, p. 643) note that there is no “single true utility function” to characterise an individual:

> The [random preference] model is consistent with a radically different interpretation [of behavior compared to the random utility model]. Here there is no single true utility function which is imperfectly processed: rather, the stochastic element derives from the inherent variability or imprecision of the individual’s preferences, whereby the individual does not always know exactly what he or she prefers.

Such an interpretation of the RP model is at odds with using the median of a preference distribution, or any other point on the distribution, as representative of the “core” preference of the individual.

Given this result concerning the normative incoherence of the RP model, I discuss the possibility of characterizing some point of the preference distribution of the RP model as the “core” preference, and using it to determine the welfare consequences of the choices made by the agent.[20] This line of reasoning, however, is not without its own limitations. If the mean (or mode) is used as the core preference, different parameterizations of the utility function can lead to different normative prescriptions despite being descriptively identical across all choice problems, leaving normative inferences under–defined. Using the median of the distribution as the core preference, however, side–steps this issue.

The use of a core preference also highlights the *econometric* duality of the RP and RU models. Any RP model in which the median is taken to be the

[20] This is the approach of Apesteguia and Ballester (2018) when describing the descriptive properties of the RP model, but they make no normative claims.

core preference can be reformulated as an RU model with that core preference. This does not, however, imply that the RP and RU models are *economically* dual to each other. It is only when the RP model is reformulated as its dual RU counterpart, which discards any interpretation of behavior as being derived from variability in *preferences*, but instead describes it as being due to variability in an *error term*, that it is able to make coherent normative statements.

While the usefulness of various descriptive properties of stochastic models is still being debated, a constructive approach to future stochastic model research is to start first with the understanding that to be *economically* useful, stochastic models cannot only determine the probability of observing choices, but must also allow for certain choices to be declared mistakes or errors. In particular, dominated choice patterns ought to be correctly described as having an opportunity cost. To this end, the notion of random *preferences* should be discarded. The RP model as envisaged by Loomes and Sugden (1995) does not allow for choices to be made in error and therefore results in an incoherent normative interpretation of choices. Reformulated as a random *parameters* model, as discussed by Apesteguia and Ballester (2018) however, the RP model can be used to provide useful normative inferences provided the *median* is used as a core preference, and proper attention is paid to the *error* distribution when discussing distributions of consumer surplus.

References

Allais, M. (Oct. 1953). “Le Comportement de l’Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l’Ecole Americaine.” *Econometrica* 21.4, pp. 503–546.

Andersen, Steffen, Glenn W. Harrison, Arne Risa Hole, Morten Lau and E. Elisabet Rutström (2012). “Non-linear mixed logit.” *Theory and Decision* 73, pp. 77–96.

Apesteguia, Jose and Miguel A. Ballester (2015). “A measure of rationality and welfare.” *Journal of Political Economy* 123.6, pp. 1278–1310.

— (2018). “Monotone stochastic choice models: The case of risk and time preferences.” *Journal of Political Economy* 126.1, pp. 74–106.

Ballinger, T. Parker and Nathaniel T. Wilcox (1997). “Decisions, Error and Heterogeneity.” *Economic Journal* 107.443, pp. 1090–1105.

Becker, G. M., M. H. DeGroot and Jacob Marschak (1963). “Stochastic models of choice behavior.” *Behavioral Science* 8.1, pp. 41–55.

Bell, David E. (1982). “Regret in Decision Making under Uncertainty.” *Operations Research* 30.5, pp. 961–981.

Bernheim, Douglas B. and Antonio Rangel (2009). “Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics.” *Quarterly Journal of Economics* February, pp. 51–104.

Blavatskyy, Pavlo R. (May 2007). “Stochastic expected utility theory.” *Journal of Risk and Uncertainty* 34.3, pp. 259–286.

— (2014). “Stronger utility.” *Theory and Decision* 76.2, pp. 265–286.

Cerreia-Vioglio, Simone, David Dillenberger, Pietro Ortoleva and Gil Riella (2019). “Deliberately Stochastic.” *American Economic Review* 109.7, pp. 2425–2445.

Diewert, W.E. (1983). “Cost-benefit analysis and project evaluation.” *Journal of Public Economics* 22.3, pp. 265–302.

Edwards, Ward (1954). “The theory of decision making.” *Psychological Bulletin* 51.4, pp. 380–417.

Fechner, Gustav (1966). *Elements of Psychophysics. Vol. I.* Ed. by Davis Humphrey Howes and Edwin Garrigues Boring. New York: Holt, Rinehart and Winston, p. 286.

Fishburn, Peter C. (1978). “A Probabilistic Expected Utility Theory of Risky Binary Choices.” *International Economic Review* 19.3, pp. 633–646.

Gao, Xiaoxue Sherry, Glenn W. Harrison and Rusty Tchernis (2022). “Behavioral Welfare Economics and Risk Preferences: A Bayesian Approach.” *Experimental Economics*.

Gul, Faruk and Wolfgang Pesendorfer (2006). “Random expected utility.” *Econometrica* 74.1, pp. 121–146.

Harless, David W. and Colin F. Camerer (1994). “The Predictive Utility of Generalized Expected Utility Theories.” *Econometrica* 62.6, pp. 1251–1289.

Harrison, Glenn, Karlijn Morsink and Mark Schneider (2020). “Literacy and the Quality of Index Insurance Decisions.” *The Geneva Risk and Insurance Review*.

Harrison, Glenn W. (1989). “Theory and Misbehavior of First-Price Auctions.” *American Economic Review* 79.4, pp. 749–762.

— (1992). “Theory and misbehavior of first price auctions: Reply.” *American Economic Review* 79.4, pp. 1426–1443.

— (2019). “The behavioral welfare economics of insurance.” *Geneva Risk and Insurance Review* 44.2, pp. 137–175.

Harrison, Glenn W. and Jia Min Ng (2016). “Evaluating the Expected Welfare Gain From Insurance.” *Journal of Risk and Insurance* 83.1, pp. 91–120.

Harrison, Glenn W. and Don Ross (2018). “Varieties of paternalism and the heterogeneity of utility structures.” *Journal of Economic Methodology* 25.1, pp. 42–67.

Harrison, Glenn W. and E. Elisabet Rutström (2008). “Risk Aversion in the Laboratory.” *Research in Experimental Economics*. Ed. by James C Cox and Glenn W Harrison. Vol. 12. Bingley: Emerald Group Publishing Limited, pp. 41–196.

— (June 2009). “Expected utility theory and prospect theory: one wedding and a decent funeral.” *Experimental Economics* 12.2, pp. 133–158.

Hey, John D. (1995). “Experimental Investigations of Errors in Decision Making Under Risk.” *European Economic Review* 39, pp. 633–640.

— (2001). “Does Repetition Improve Consistency?” *Experimental Economics* 4, pp. 5–54.

Hey, John D. and Chris Orme (1994). “Investigating Generalizations of Expected Utility Theory using Experimental Data.” *Econometrica* 62.6, pp. 1291–1326.

Holt, Charles A. and Susan K. Laury (2002). “Risk Aversion and Incentive Effects.” *American Economic Review* 92.5, pp. 1644–1655.

Kahneman, Daniel and Amos Tversky (1979). “Prospect theory: An analysis of decision under risk.” *Econometrica* 47.2, pp. 263–292.

King, Mervyn A. (1983). “Welfare analysis of tax reforms using household data.” *Journal of Public Economics* 21.2, pp. 183–214.

Loomes, Graham, Peter G. Moffatt and Robert Sugden (2002). “A microeconometric test of alternative stochastic theories of risky choice.” *Journal of Risk and Uncertainty* 24.2, pp. 103–130.

Loomes, Graham and Robert Sugden (1982). “Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty.” *Economic Journal* 92.368, pp. 805–824.

— (1995). “Incorporating a stochastic element into decision theories.” *European Economic Review* 39.3-4, pp. 641–648.

— (1998). “Testing different stochastic specifications of risky choice.” *Economica* 65, pp. 581–598.

Luce, Robert Duncan (1958). “A Probabilistic Theory of Utility.” *Econometrica* 26.2, pp. 193–224.

— (1959). *Individual Choice Behavior: A Theoretical Analysis*. New York: Wiley.

Machina, M. J. (1985). “Stochastic choice functions generated from deterministic preferences over lotteries.” *Economic Journal* 95, pp. 575–594.

Manzini, Paola and Marco Mariotti (2014). “Welfare economics and bounded rationality: the case for model-based approaches.” *Journal of Economic Methodology* 21.4, pp. 343–360.

Marschak, Jacob (1950). “Rational Behavior, Uncertain Prospects, and Measurable Utility.” *Econometrica* 18.2, pp. 111–141.

McFadden, Daniel and Kenneth Train (2000). “Mixed MNL Models for Discrete Response.” *Journal of Applied Econometrics* 15, pp. 447–470.

Mosteller, Frederick and Philip Nogee (1951). “An Experimental Measurement of Utility.” *Journal of Political Economy* 59.5, pp. 371–404.

Pratt, John W. (1964). “Risk Aversion in the Small and in the Large.” *Econometrica* 32.1/2, pp. 122–136.

Quandt, Richard E. (1956). “A Probabilistic Theory of Consumer Behavior.” *Quarterly Journal of Economics* 70.4, pp. 507–536.

Quiggin, John (1982). “A Theory of Anticipated Utility.” *Journal of Economic Behavior & Organization* 3, pp. 323–343.

Ross, Don (2014). “Psychological versus economic models of bounded rationality.” *Journal of Economic Methodology* 21.4, pp. 411–427.

— (2021). “Neo-Samuelsonian methodology, normative economics, and the quantitative intentional stance.” *Working Paper 2021-04*, pp. 1–29.

Salant, Yuval and Ariel Rubinstein (2008). “(A, f) : Choice with Frames.” *Review of Economic Studies* 75, pp. 1287–1296.

Samuelson, P. A. (1938). “A Note on the Pure Theory of Consumer’s Behaviour.” *Economica* 5.17, pp. 61–71.

Samuelson, Paul A. (1974). “Complementarity: An Essay on The 40th Anniversary of the Hicks-Allen Revolution in Demand Theory.” *Journal of Economic Literature* 12.4, pp. 1255–1289.

Simon, Herbert A. (1955). “A Behavioral Model of Rational Choice.” *Quarterly Journal of Economics* 69.1, pp. 99–118.

Smith, Vernon L. (1982). “Microeconomic systems as an experimental science.” *American Economic Review* 72.5, pp. 923–955.

Tversky, Amos and Daniel Kahneman (1992). “Advances in prospect theory: Cumulative representation of uncertainty.” *Journal of Risk and Uncertainty* 5.4, pp. 297–323.

Von Neumann, J. and Oskar Morgenstern (1944). *Theory of Games and Economic Behavior*. Vol. 2. Princeton, New Jersey: Princeton University Press.

Wilcox, Nathaniel T. (2008). “Stochastic models for binary discrete choice under risk: A critical primer and econometric comparison.” *Research in Experimental Economics*. Ed. by James C. Cox and Glenn W. Harrison. Vol. 12. Bingley, U.K.: Emerald Group Publishing Limited, pp. 197–292.

— (2011). “’Stochastically More Risk Averse:’ A Contextual Theory of Stochastic Discrete Choice Under Risk.” *Journal of Econometrics* 162.1, pp. 89–104.

Yaari, Menahem E . (1987). “The Dual Theory of Choice under Risk.” *Econometrica* 55.1, pp. 95–115.