



Dual choice axiom and probabilistic choice

Pavlo R. Blavatskyy¹

Published online: 26 September 2020

© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

A decision maker chooses in a probabilistic manner if she does not necessarily prefer the same choice alternative when repeatedly presented with the same choice set. Probabilistic choice may occur for a variety of reasons such as unobserved attributes of choice alternatives, imprecision of preferences, or random errors/noise in decisions. The Luce choice model (also known as strict utility or multinomial logit) is derived from the choice axiom (also known as the independence from irrelevant alternatives). This axiom postulates that the relative likelihood of choosing one choice alternative A over another choice alternative B is not affected by the presence or absence of other choice alternatives in the choice set. This paper presents a dual choice axiom: the relative probability of NOT choosing A over the probability of NOT choosing B is independent from irrelevant alternatives. A new model of probabilistic choice is derived from this dual axiom. This model coincides with Luce's choice model only in the case of a binary choice. The new model has similar properties as the Luce choice model: the higher is the utility of a choice alternative, the higher is the probability that a decision maker chooses this alternative and the lower is the probability that he or she chooses any other alternative. The new model differs from the Luce choice model in two aspects: utility of choice alternatives is bounded (from above and below) and choice probabilities are more sensitive to differences in utility of choice alternatives.

Keywords Decision making · Probabilistic choice · Choice axiom · Independence from irrelevant alternatives · Luce choice model · Strict utility

JEL Classifications D01 · C25 · D91

✉ Pavlo R. Blavatskyy
p.blavatskyy@montpellier-bs.com

¹ Montpellier Business School, 2300, Avenue des Moulins, 34185 Montpellier Cedex 4, France

1 Introduction

Economic theory is traditionally based on deterministic preferences¹ leaving little room for probabilistic choice.² Empirical research, however, strongly backs probabilistic choice.³ Probabilistic choice may occur for a variety of reasons such as unobserved attributes of choice alternatives (e.g. McFadden 1976), imprecision of preferences (Falmagne 1985; Butler and Loomes 2007, 2011), random errors/noise in decisions (e.g. Fechner 1860; Hey and Orme 1994). Models of probabilistic choice originated in mathematical psychology and include inter alia random utility (also known as random preference or random parameter) approach (e.g., Falmagne 1985; Loomes and Sugden 1995), the Fechner (1860) model of random errors (or strong utility)⁴ and the Luce (1959) choice model (strict utility or multinomial logit).⁵ These models are often used in econometric estimation on microeconomic data.

Yet, psychological models of probabilistic choice may not always suit economic data, which creates a demand for new models of probabilistic choice designed for economic applications. Random utility/preference/parameter approach can rationalize the preference reversal phenomenon in choice under risk (e.g. Loomes 2005, p.310–311) but it cannot account for rare violations of dominance that are occasionally observed in the data (e.g. Loomes and Sugden 1998, p. 585). Fechner (1860) and Luce (1959) models can rationalize some instances of the common ratio effect (e.g. Loomes 2005, p.305) or violations of the betweenness axiom (e.g., Blavatskyy 2006) in risky choice and some instances of the common difference effect (e.g., Blavatskyy 2017, Section 4, pp. 144–145) in intertemporal choice but they generate too many violations of dominance (Loomes and Sugden 1998, p. 593).⁶

Luce (1959) developed a general theory of probabilistic choice from the choice axiom, which is also known as the independence from irrelevant alternatives. Luce (1959) choice model is often used for eliciting risk preferences (e.g., Camerer and Ho 1994; Wu and Gonzalez 1996; Holt and Laury 2002, Eq. (1), p. 1652) and time preferences (e.g., Andersen et al. 2008, p. 599, Eq. 9; Meier and Sprenger 2015, p. 276, Eq. 1). Luce (1959) choice axiom postulates that the ratio of the probability of choosing one alternative to the probability of choosing another alternative is not

¹ E.g., expected utility (von Neumann and Morgenstern 1947), subjective expected utility (Savage 1954) or discounted utility (Samuelson 1937).

² Machina (1985) and Chew et al. (1991) develop models of probabilistic choice under risk as a result of deliberate randomization by decision makers with (deterministic) quasi-concave preferences. Hey and Carbone (1995) find that conscious randomization cannot rationalize their experimental data but Agranov and Ortoleva (2017) reach the opposite conclusion.

³ E.g., Camerer (1989, p.81), Starmer and Sugden (1989, p.170), Hey and Orme (1994, p.1296), Ballinger and Wilcox (1997, p.1100), Hey et al. (2010), Blavatskyy and Maafi (2018, p. 266).

⁴ The Fechner model is an econometric model of discrete choice with random errors (noise, attention slips, carelessness) additive on the (latent) utility scale (cf. Becker et al. 1963, pp. 44–45; Hey and Orme 1994, p. 1301). Binary Luce (1959) model is a special case of binary Fechner model when such errors are drawn from the logistic distribution (cf. Yellott 1977).

⁵ Harless and Camerer (1994) model probabilistic choice with a constant probability of a pure tremble. Yet, Carbone (1997) and Loomes et al. (2002) find that this constant error model fails to explain their experimental data and it is essentially “inadequate as a *general* theory of stochastic choice”.

⁶ Modified Fechner-type models with heteroscedastic random errors can avoid violations of monotonicity (e.g., the contextual utility model of Wilcox (2008, 2011)) or violations of the first-order stochastic dominance (e.g., Blavatskyy 2014).

affected by the presence (or absence) of other “irrelevant” alternatives in the choice set. This axiom is simple and intuitively appealing in choice situations when there are no substitution/complementary effects among choice alternatives.

In Luce (1959) choice model every alternative is chosen with a strictly positive probability. This property may be undesirable in economic applications where clearly inferior alternatives are never chosen. Luce (1959) addressed this problem via a two-stage procedure: first, dominated alternatives are discarded from a choice set; and then a decision maker chooses in a probabilistic manner among the remaining alternatives. Yet, this two-stage procedure creates a discontinuous choice function. This paper proposes a new model of probabilistic choice in which a decision maker may never choose some alternatives (that are sufficiently undesirable compared to other available options).

This paper demonstrates that one can formulate a dual version of Luce (1959) choice axiom that is arguably as intuitively appealing as its classic sibling. The dual choice axiom proposed in this paper postulates that the ratio of the probability of not choosing (i.e. rejecting) one alternative to the probability of not choosing (rejecting) another alternative is not affected by the presence (or absence) of other “irrelevant” alternatives in the choice set. In case of a binary choice, the dual choice axiom simply mirrors Luce (1959) choice axiom. The choice of alternative A over alternative B in a direct binary choice is equivalent to B being rejected in favor of A . Yet, when the choice set contains three or more alternatives, the dual choice axiom has different implications from Luce (1959) choice axiom. This paper derives a new model of probabilistic choice from the above dual choice axiom. In this model, the likelihood that a decision maker does not choose one alternative from some choice set is proportionate to the ratio of the disutility of this alternative to the sum of disutilities of all alternatives in the choice set.

For example, consider a country that is about to leave the European Union and faces three mutually exclusive scenarios: to leave the union without any agreement (“No Deal Exit”), to leave the union with an agreement (“Deal”) and to stay in the union (“No Exit”). According to the classic choice axiom, the relative likelihood that one country leaves the union without any agreement (“No Deal Exit”) rather than stays in the union (“No Exit”) is not affected by the characteristics of any exiting agreement that the country negotiated with the union for an orderly exit (“Deal”). According to the dual choice axiom, the relative likelihood of either “No Deal Exit” or “Deal” rather than either “No Exit” or “Deal” is not affected by the characteristics of the exiting agreement. In other words, the dual choice axiom postulates that the relative chances that the country leaves the union (with or without agreement) rather than does not leave without an agreement are independent from the agreement itself. When choices are hard, a decision maker can formulate the independence from irrelevant alternatives for a probability of rejecting rather than choosing certain options.

This paper is closest to Marley and Louviere (2005) who consider a model of probabilistic choice where a decision maker selects the worst alternative from the choice set.⁷ Their model is similar to Luce (1959) choice model: the higher is the disutility of a choice alternative the more it is likely to be selected as the worst alternative. Marley and Louviere (2005) effectively impose the independence from

⁷ They also consider a related model where a decision maker selects the best and the worst alternative from the choice set.

irrelevant alternatives on the relative probability of selecting the worst choice alternative whereas the dual choice axiom proposed in this paper imposes the independence from irrelevant alternatives on the relative probability of rejecting a choice alternative. In other words, we consider alternatives that are not chosen (rejected) from a choice set, which does not necessarily imply that these alternatives are the worst available alternatives (considered in Marley and Louviere (2005)). In case of a binary choice, when an alternative that is not chosen is automatically the worst available alternative, the model of probabilistic choice derived in this paper from the dual choice axiom coincides with the model of Marley and Louviere (2005) and a binary Luce (1959) choice model.

The remainder of the paper is organized as follows. Section 2 presents mathematical notation, formulates classic Luce (1959) choice axiom and a new dual choice axiom. Section 3 presents a new model of probabilistic choice derived from the dual choice axiom. Section 4 compares the new model with Luce (1959) choice model using an example of ternary choice. Section 5 concludes with a general discussion.

2 Notation, choice axiom and dual choice axiom

Let S be a choice set with $n \geq 2$ choice alternatives. A choice alternative $A \in S$ can be a consumption bundle, a risky lottery (a probability distribution), an uncertain act (a random variable), a stream of intertemporal outcomes, a behavioral strategy etc. We denote choice alternatives by capital Latin letters A, B, C etc.

A decision maker can be an individual or a group of individuals. Let $P(A|S) \in (0,1)$ denote the probability that a decision maker chooses alternative $A \in S$ from the choice set S . The probability that alternative $A \in S$ is not chosen from the choice set S is given by $1 - P(A|S)$. As Luce (1959), we assume that choice alternatives, which are never chosen, can be deleted from the choice set without any effect on decision making. Thus, without loss of generality, we consider only choice sets that contain alternatives that are chosen with a strictly positive probability. Finally, let $P(A,B) \in (0,1)$ denote the probability that a decision maker chooses alternative A over alternative B in a direct binary choice (i.e. B is not chosen from the set $\{A,B\}$) and let $P(B,A) = 1 - P(A,B)$ denote the probability that a decision maker chooses alternative B over alternative A in a direct binary choice (i.e. A is not chosen from set $\{A,B\}$).

We consider two axioms imposed on choice probabilities. The first axiom is Luce (1959) choice axiom (Eq. 2 below) and the second axiom is a dual choice axiom (Eq. 3 below). To distinguish between these two cases, we use subscript ‘‘Luce’’ to denote choice probabilities that satisfy Luce (1959) choice axiom. For example, $P_{Luce}(A,B)$ denotes the probability that a decision maker chooses alternative A over alternative B in a direct binary choice in Luce choice model.

A decision maker can choose only one choice alternative from the choice set so that

$$\sum_{A \in S} P(A|S) = 1 \quad (1)$$

Luce (1959) *choice axiom* postulates that for any $A, B \in S$, the relative likelihood

$$\frac{P_{Luce}(A|S)}{P_{Luce}(B|S)} = \frac{P_{Luce}(A, B)}{P_{Luce}(B, A)} \quad (2)$$

depends only on the characteristics of choice alternatives A and B (and is independent from the characteristics of other “irrelevant” choice alternatives).⁸ For example, the relative likelihood that an airline purchases a Boeing aircraft rather than an Airbus aircraft depends only on the characteristics of these two aircrafts and is not affected by the characteristics of other aircrafts available on the market.

In his discussion of the independence from irrelevant alternatives, Luce (1959) makes it clear that the choice axiom (2) is only one reasonable version of the independence from irrelevant alternatives in the context of probabilistic choice: “... if one is comparing two alternatives according to some algebraic criterion, say preference, this comparison should be unaffected by the addition of new alternatives or the subtraction of old ones (different from the two under consideration). Exactly what should be taken to be the probabilistic analogue of this idea is not perfectly clear, but one reasonable possibility is the requirement that the ratio of the probability of choosing one alternative to the probability of choosing the other should not depend upon the total set of alternatives available”. In other words, Luce (1959) acknowledged that condition (2) is *a* choice axiom, not *the* choice axiom that it became over the last 60 years. This paper formulates a different probabilistic version of the independence from irrelevant alternatives. It is particularly appealing when decision makers face difficult choices so that it may be intuitive to reason in terms of alternatives that are not chosen.

In this paper we formulate a *dual choice axiom*: the ratio of probability that $A \in S$ is NOT chosen from the set S to probability that $B \in S$ is NOT chosen from the set S depends only on the characteristics of choice alternatives A and B (and is independent from the characteristics of other “irrelevant” choice alternatives). For example, the relative likelihood that a Democratic candidate *does not* win an election compared to the likelihood that a Republican candidate *does not* win an election is not affected by the presence of other independent candidates. While the classic choice axiom establishes independence from irrelevant alternatives for the relative probability of choice, the dual choice axiom does the same for the relative probability of not choosing. Arguably, the dual choice axiom has the same intuitive appeal as the classic choice axiom.

Formally, for any two alternatives $A, B \in S$, our dual choice axiom can be written as

$$\frac{1-P(A|S)}{1-P(B|S)} = \frac{P(B, A)}{P(A, B)} \quad (3)$$

⁸ Luce (1959, Axiom 1, part (i)) formulated his choice axiom in a different form: the probability that a decision maker chooses an alternative from set T that lies in the nonempty set $R \subset T$ is equal to the probability that this decision maker chooses an alternative from set T that lies in the nonempty set $S \subset T$ multiplied by the probability that the decision maker chooses an alternative from set S that lies in set $R \subset S$. Yet, it can be easily shown that this formulation implies the independence from irrelevant alternatives property (2), cf. Luce (1959, Lemma 3).

The dual choice axiom (3) constrains binary choice probabilities in the same way as Luce (1959) choice axiom (2). Specifically, if dual choice axiom (3) holds then binary choice probabilities must satisfy the product rule (cf. Eq. (16) in the proof of proposition 1 in the [Appendix](#)). Likewise, if Luce (1959) choice axiom (2) holds then binary choice probabilities must satisfy the product rule as well. Thus, the constraints implied by dual choice axiom (3) on binary choice probabilities are the same as the constraints implied by Luce choice axiom (2). For non-binary choice probabilities, the constraints implied by (3) are similar (but not the same) as the constraints implied by Luce choice axiom (2). The main difference is that dual choice axiom (3) restricts probabilities of not choosing (which sum up to $n-1$, cf. Eq. (7) below) whereas Luce choice axiom (2) restricts probabilities of choosing (which sum up to one, cf. Eq. (1) above).

3 Model of probabilistic choice

First, we show that dual choice axiom (3) implies the same restriction on binary choice probabilities as the classic choice axiom (2) — a binary Luce choice model, which is also known as binary strict utility.

Proposition 1 If $n \geq 3$ and dual choice axiom (3) holds then

$$P(A, B) = \frac{U(A)}{U(A) + U(B)} \quad (4)$$

where $U: S \rightarrow \mathbb{R}_+$ is utility function unique up to a multiplication by a positive constant.

The proof is presented in the [Appendix](#).

According to proposition 1, our newly proposed dual choice axiom (3) has the same implication for binary probabilistic choice as the classic choice axiom (2). This result is perhaps not surprising given that only two outcomes are possible in binary choice: either A is chosen, or A is not chosen. In other words, in the context of binary choice, dual choice axiom simply mirrors classic choice axiom. Yet, for choice among $n \geq 3$ alternatives, implications of dual choice axiom (3) differ from those of classic choice axiom (2), as demonstrated by the following proposition 2.

Proposition 2 If $n \geq 3$ and dual choice axiom (3) holds then

$$P(A|S) = 1 - \frac{n-1}{\sum_{B \in S} \frac{U(A)}{U(B)}} \quad (5)$$

where $U: S \rightarrow \mathbb{R}_+$ is utility function unique up to a multiplication by a positive constant.

The proof is presented in the [Appendix](#).

Model (5) can be alternatively presented as Eq. (6).

$$1 - P(A|S) = (n-1) \frac{U^{-1}(A)}{\sum_{B \in S} U^{-1}(B)} \quad (6)$$

Equation (6) emphasizes the duality of the new model to classic Luce (1959) choice model. According to formula (6), the probability $1 - P(A|S)$ of not choosing alternative A from the set S is proportionate to the disutility of this alternative $U^{-1}(A)$. If the choice set contains n alternatives then probabilities of not choosing (rejecting) an alternative should sum up to $n-1$.

$$\sum_{A \in S} [1 - P(A|S)] = n-1 \quad (7)$$

Therefore, disutility $U^{-1}(A)$ in formula (6) is divided by the sum of disutilities of all alternatives and multiplied by $n-1$.

Model of probabilistic choice (5) has similar properties as Luce (1959) choice model. The higher is utility of alternative A , the higher is the probability that a decision maker chooses A from the choice set S and the lower is the probability that he or she chooses any other alternative B from the choice set S . From a dual perspective, the higher is utility of alternative A , the lower is the probability $1 - P(A|S)$ that a decision maker *does not* choose A from the choice set S and the higher is that the probability that he or she *does not* choose any other alternative B from the choice set S . If alternative A yields a higher utility than alternative B then according to model (5) a decision maker is more likely to choose alternative A rather than alternative B from any choice set containing both A and B . Model (5) satisfies the regularity condition (e.g., Marley 1965): a decision maker is more likely to choose alternative A from a subset of the choice set S than from the set S itself.

In the degenerate case $n = 1$ Eq. (5) becomes simply $P(A|S) = 1$, i.e. when the choice set contains only one alternative a decision maker always chooses this alternative. For $n = 2$ Eq. (5) becomes Eq. (4), i.e. when the choice set contains two choice alternatives a decision maker behaves as in a binary Luce (1959) choice model (strict utility). If all choice alternatives yield the same utility ($U(A) = U(B)$ for all $A, B \in S$) then Eq. (5) becomes $P(A|S) = 1/n$, i.e. a decision maker chooses at random among all such alternatives.

In Luce (1959) choice model utility of an alternative is bounded to be a positive number. Yet, it is unbounded in relative terms—utility of alternative A can be much smaller than utilities of other available alternatives and a decision maker still chooses A with a strictly positive probability. This can be viewed as an undesirable feature of Luce (1959) choice model. Arguably, when one alternative is much less desirable than other available alternatives (in terms of utility) a choice decision is simple—a clearly inferior alternative is discarded.

Model of probabilistic choice (5) allows this possibility—an alternative that is much less desirable than other alternatives can be chosen with zero probability. Unlike Luce (1959) choice model, model (5) imposes “smart” relative bounds on the utility of every alternative. For an alternative to be chosen with a positive probability, its utility must be sufficiently high compared to the utilities of other available alternatives.

When utility of alternative A decreases, the probability (5) that a decision maker chooses this alternative decreases as well. Yet, so far, nothing restricts probability (5) to remain strictly positive. In other words, utility of alternative A must remain high enough so that a decision maker chooses alternative A with a strictly positive probability, i.e. the denominator of the fraction on the right-hand-side of Eq. (5) should always exceed the nominator. This imposes a lower bound (8) on the utility of alternative A .

$$U(A) > \frac{n-2}{\sum_{\substack{B \in S \\ B \neq A}} \frac{1}{U(B)}} \quad (8)$$

When $n=2$ inequality (8) becomes simply $U(A) > 0$, i.e., like in Luce (1959) choice model, utility must be a positive real number. When $n > 2$ inequality (8) effectively restricts the utility of A not to fall below $(n-2)/(n-1)$ of the harmonic mean of utilities of all other available alternatives.

Second, when utility of alternative A increases, the probabilities, with which a decision maker chooses other alternatives, decrease. Yet, so far, nothing restricts these probabilities to remain strictly positive. Thus, utility of alternative A must remain low enough so that a decision maker chooses other available alternatives with a strictly positive probability. It is sufficient only to check that the alternative that has the lowest utility $\min_{C \in S} U(C)$ is chosen with a strictly positive probability (since any alternative that yields a higher utility is chosen with a higher probability). This imposes inequality (9) on utility of alternative A .

$$\frac{1}{U(A)} > \sum_{\substack{B \in S \\ B \neq A}} \left[\frac{1}{\min_{C \in S} U(C)} - \frac{1}{U(B)} \right] \quad (9)$$

When all available alternatives but A yield the same utility, which is smaller than utility of A , then the right-hand-side of inequality (9) becomes zero, i.e. inequality (9) is satisfied for any positive real valued utility function. This happens, for example, in case of a binary choice when there is only one alternative other than A and this alternative yields a smaller utility than A . If not all available alternatives other than A yield the same utility, the right-hand side of inequality (9) is strictly positive and we can rewrite inequality (9) as an upper bound (10) on utility of alternative A .

$$U(A) < \frac{1}{\sum_{\substack{B \in S \\ B \neq A}} \left[\frac{1}{\min_{C \in S} U(C)} - \frac{1}{U(B)} \right]} \quad (10)$$

Inequalities (8) and (10) can be combined in one condition: the least desirable alternative in the choice set must yield utility greater than $(n-1)/n$ of the harmonic mean utility of all available choice alternatives (i.e., $\min_{C \in S} U(C) > (n-1) / \sum_{B \in S} \frac{1}{U(B)}$). In other

words, in model (5) the least desirable alternative cannot be significantly inferior compared to other alternatives.

4 Example: Ternary choice

Figure 1 illustrates model of probabilistic choice (5) for ternary choice among three alternatives A , B and C . Utilities of alternatives B and C are fixed at $U(B)=2$ and $U(C)=3$. Utility of alternative A is shown on the horizontal axis (in this case, the lower bound (8) is $U(A) > 1.2$ and the upper bound (10) is $U(A) < 6$). Probabilities $P(A|\{A,B,C\})$, $P(B|\{A,B,C\})$ and $P(C|\{A,B,C\})$ are shown as black, grey and white areas correspondingly. Figure 1 shows that as $U(A)$ increases from 1.25 to 6 probability $P(A|\{A,B,C\})$ increases from 0.02 to 0.67; probability $P(B|\{A,B,C\})$ decreases from 0.39 to 0 and probability $P(C|\{A,B,C\})$ decreases from 0.59 to 0.33. Note that the ratio of probability that B is not chosen to probability that C is not chosen $[1 - P(B|\{A,B,C\})]/[1 - P(C|\{A,B,C\})]$ stays constant at $3/2$ for all values $U(A) \in [1.25, 6]$, in accordance with the dual choice axiom (3). In contrast, the ratio of probability that B is chosen to probability that C is chosen diminishes from 0.66 to zero, in violation of the classic choice axiom (2).

For comparison, Fig. 2 illustrates classic Luce (1959) choice model

$$P_{Luce}(A|\{A, B, C\}) = \frac{U(A)}{U(A) + U(B) + U(C)}$$

for the same parameters $U(A) \in [1.25, 6]$, $U(B)=2$ and $U(C)=3$. Figure 2 shows that as $U(A)$ increases from 1.25 to 6 probability $P_{Luce}(A|\{A,B,C\})$ in Luce (1959) choice model increases from 0.2 to 0.55; probability $P_{Luce}(B|\{A,B,C\})$ in Luce (1959) choice model decreases from 0.32 to 0.18 and probability $P_{Luce}(C|\{A,B,C\})$ in Luce (1959)

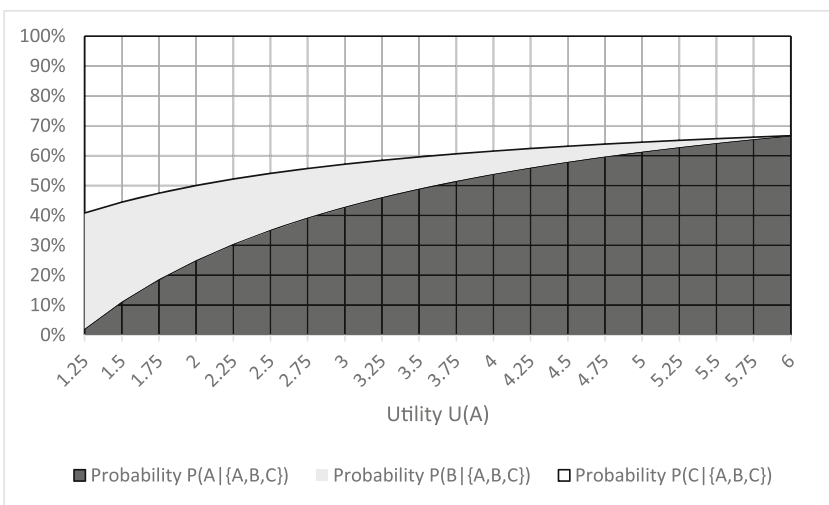


Fig. 1 Model (5) for ternary choice when $U(B)=2$, $U(C)=3$ and different values of $U(A)$ as shown on the horizontal axis

choice model decreases from 0.48 to 0.27. Note that the ratio of probability that *B* is chosen to probability that *C* is chosen in Luce (1959) choice model stays constant at 2/3 for all values $U(A) \in [1.25, 6]$, in accordance with the classic choice axiom (2). In contrast, the ratio of probability that *B* is not chosen to probability that *C* is not chosen in Luce (1959) choice model diminishes from 1.31 to 1.13, in violation of the dual choice axiom (3).

A comparison of Figs. 1 and 2 shows that model (5) is relatively more sensitive to differences in utility than classic Luce (1959) choice model. For example, as $U(A)$ increases from 1.25 to 2, $P(A|\{A,B,C\})$ in model (5) increases from 0.02 to 0.25 and $P(B|\{A,B,C\})$ in model (5) decreases from 0.39 to 0.25. In contrast, in Luce (1959) choice model, $P_{Luce}(A|\{A,B,C\})$ increases from 0.2 to 0.29 and $P_{Luce}(B|\{A,B,C\})$ decreases from 0.32 only to 0.29.

From another perspective, Fig. 3 shows probability $P(A|\{A,B,C\})$ in model (5) as a function of utility ratio $U(A)/U(B)$, shown on the horizontal axis, and utility ratio $U(A)/U(C)$, shown on the vertical axis. Note that for a ternary choice bound (8) becomes

$$\frac{U(A)}{U(B)} + \frac{U(A)}{U(C)} > 1 \tag{11}$$

which is shown as a solid black line on Fig. 3 corresponding to the case $P(A|\{A,B,C\}) = 0$. Similarly, for a ternary choice bound (10) becomes

$$\left| \frac{U(A)}{U(B)} - \frac{U(A)}{U(C)} \right| < 1 \tag{12}$$

which is shown as two parallel dashed black lines on Fig. 3.

For any combination of utility ratios $U(A)/U(B)$ and $U(A)/U(C)$ that are above the solid black line and between two dashed black lines on Fig. 3 there is a well-defined

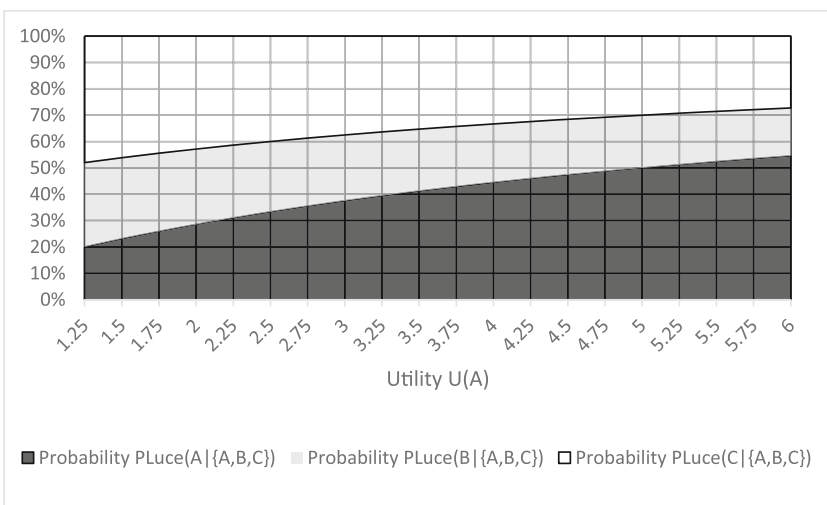


Fig. 2 Luce (1959) choice model when $U(B) = 2$, $U(C) = 3$ and different values of $U(A)$ as shown on the horizontal axis

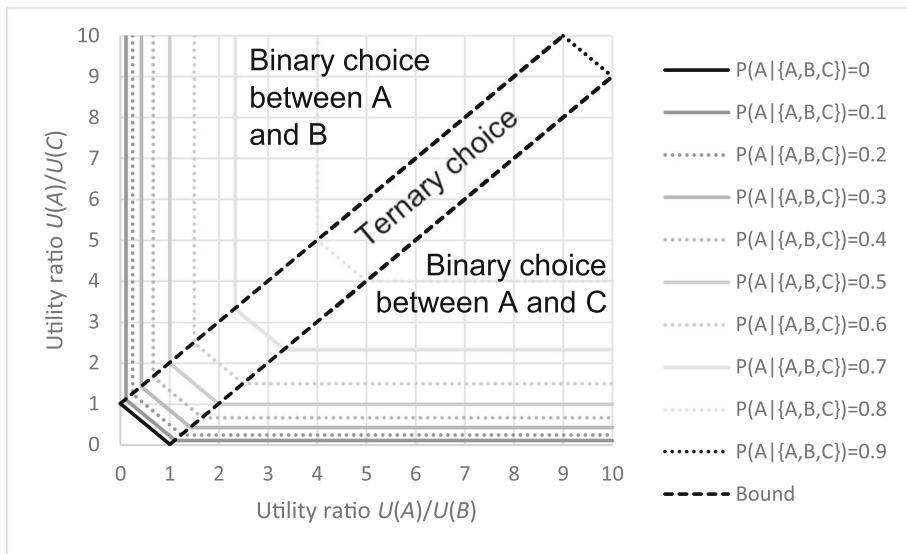


Fig. 3 Probability $P(A|\{A,B,C\})$ in model (5) as a function of utility ratios $U(A)/U(B)$ and $U(A)/U(C)$

probability $P(A|\{A,B,C\})$ in model (5). Figure 3 illustrates the sets of utility ratios $U(A)/U(B)$ and $U(A)/U(C)$ which generate probabilities $P(A|\{A,B,C\}) = p$ for all $p \in \{0.1; 0.2; \dots; 0.9\}$.⁹ For comparison, Fig. 4 does the same for Luce (1959) choice model.¹⁰ Figures 3 and 4 show that both in model (5) and in the classic Luce (1959) choice model a decision maker chooses alternative A with a greater probability when utility ratios $U(A)/U(B)$ and $U(A)/U(C)$ increase. Yet, the same increase in utility ratios has a greater impact on the choice probability in model (5) compared to Luce (1959) choice model. For example, increasing utility ratios $U(A)/U(B)$ and $U(A)/U(C)$ from one to two increases the likelihood of choosing A in model (5) from 1/3 to 0.6 and in Luce (1959) choice model—from 1/3 to 0.5.

If utility ratio $U(A)/U(C)$ is relatively high so that bound (12) is violated (cf. an area above the higher dashed black line on Fig. 3) then alternative C is too inferior (in terms of utility) compared to A and B. In this case C is never chosen with a positive probability and choice becomes binary. The decision maker chooses either A or B. Binary choice probability is given by (4), which depends only on ratio $U(A)/U(B)$. Alternative A is chosen with probability $p \in \{0.1; 0.2; \dots; 0.9\}$ when the ratio $U(A)/U(B)$ equals $p/(1-p)$. Figure 3 illustrates these binary choice probabilities in an area above the higher dashed black line. Similarly, if ratio $U(A)/U(B)$ is relatively high, i.e. bound (12) is violated, then B is too inferior and never chosen with a positive probability. The decision maker chooses either A or C. Binary choice probability now depends only on ratio $U(A)/U(C)$. Alternative A is chosen with probability $p \in \{0.1; 0.2; \dots; 0.9\}$ when the ratio $U(A)/U(C)$ equals $p/(1-p)$. Figure 3 illustrates these binary choice probabilities in an area below the lower dashed black line.

⁹ These sets are defined by equation $\frac{U(A)}{U(B)} + \frac{U(A)}{U(C)} = \frac{1+p}{1-p}$ subject to bounds (11) and (12).

¹⁰ In Luce (1959) choice model $P_{Luce}(A|\{A,B,C\}) = p$ is generated by utility ratios $\frac{U(B)}{U(A)} + \frac{U(C)}{U(A)} = \frac{1-p}{p}$.

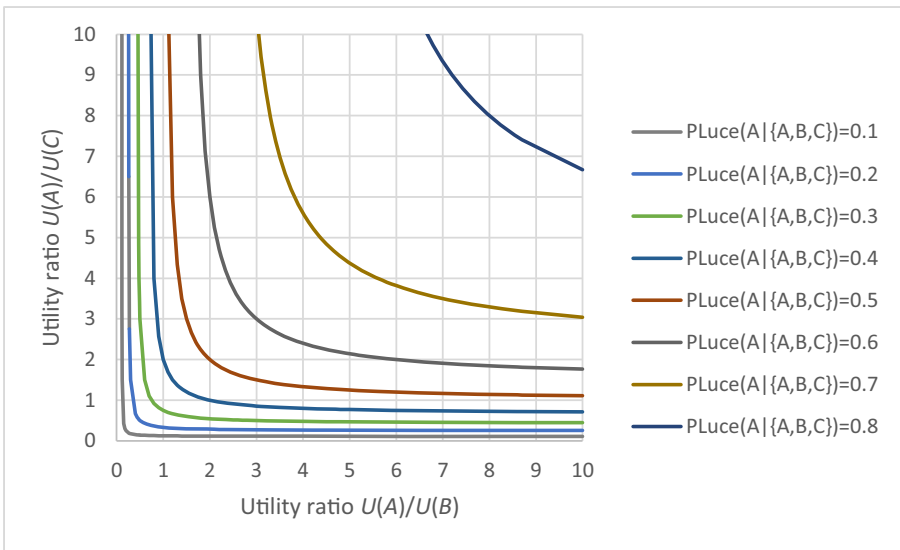


Fig. 4 Probability $P_{Luce}(A|\{A,B,C\})$ in Luce (1959) choice model as a function of utility ratios $U(A)/U(B)$ and $U(A)/U(C)$

5 Conclusion

Luce (1959) choice model has a comparative advantage over other classic models of probabilistic choice. Specifically, random utility/preference approach (e.g., Falmagne 1985; Loomes and Sugden 1995) violates weak stochastic transitivity akin to the Condorcet (1785) paradox in social choice. Yet, decision makers rarely exhibit such violations (e.g., Rieskamp et al. 2006). Moreover, practical applications of random utility/preference often require restrictive parametric assumptions characterizing utility/preference only with one parameter¹¹ to avoid large variance-covariance matrices. Classic Fechner (1860) model of random errors (strong utility) violates dominance.¹² Yet, decision makers rarely violate transparent dominance (for choice under risk e.g., Carbone and Hey 1995; Loomes and Sugden 1998, Table 2, p. 591; Hey 2001, Table 2, p.14; see however, Birnbaum and Navarrete 1998, p. 61; Birnbaum 2005, p.1356). Moreover, Fechner's strong utility applies only to binary choice.¹³

Luce (1959) choice model is derived from an intuitively appealing choice axiom. This paper presents a dual choice axiom. A new model of probabilistic choice is derived from this dual choice axiom. For binary choice, the new model coincides with Luce (1959) choice model. For choice among $n > 3$ alternatives, the new model is qualitatively similar to Luce (1959) choice model (a decision maker is more likely to choose more desirable alternatives) but utility of an alternative is bounded (from above

¹¹ So that this approach is also known as a random parameter model. Examples in the context of intertemporal choice are Collier and Williams (1999, p. 115, Section 4.2), Warner and Pleeter (2001, p. 38, Section III.A) and Harrison et al. (2002, p. 1611, Section III.A).

¹² The first-order stochastic dominance in choice under risk, state-wise dominance in choice under uncertainty/ambiguity or the first-order temporal dominance in choice over time.

¹³ Blavatsky (2018) recently proposed its generalization to choice among $n > 2$ alternatives.

and below) and choice probabilities are relatively more sensitive to differences in utility.

Utility function in the new model of probabilistic choice, as in classic Luce (1959) choice model, maps choice alternatives to positive real numbers. Decision theories often use real valued utility functions (i.e. mapping choice alternatives to zero or negative numbers) that are unique up to a positive affine transformation (e.g., von Neumann and Morgenstern (1947) expected utility function in choice under risk, Samuelson (1937) discounted utility function in intertemporal choice). The newly proposed model of probabilistic choice, like Luce (1959) choice model, can be adapted to such utility functions via transformation $U(A) = \exp(V(A))$, where $V: S \rightarrow \mathbb{R}$ is utility function that is unique up to a positive affine transformation. With this transformation, our proposed model of probabilistic choice becomes

$$P(A|S) = 1 - \frac{n-1}{\sum_{B \in S} e^{b[V(A)-V(B)]}} \quad (13)$$

where $b > 0$ is the scale parameter of utility function $V: S \rightarrow \mathbb{R}$. This scale parameter can be interpreted as the degree of noise/randomness in probabilistic choice. When b is close to zero, a decision maker chooses every alternative with probability close to $1/n$. Model (13) is a special case of a generalized Fechner's strong utility recently proposed by Blavatsky (2018): $P(A|S) = F(V(A) - V(A_2), V(A) - V(A_3), \dots, V(A) - V(A_n))$ where $F: \mathbb{R}^{n-1} \rightarrow [0, 1]$ denotes a symmetric function¹⁴ and $A_2, \dots, A_n \in S \setminus \{A\}$ denote choice alternatives other than A .

Debreu (1960, p. 188) criticized Luce (1959) choice model with the following example. Let A denote the Debussy quartet, B —the 8th symphony of Beethoven and C —the same symphony with a different conductor. In a direct binary choice a (presumably French) decision maker is indifferent between B and C , so that $P(B, C) = 0.5$, but he or she has a slight preference for A over B or C , so that $P(A, B) = 0.6$ and $P(A, C) = 0.6$. Yet, according to Luce (1959) choice model, in a ternary choice among A , B and C this decision maker chooses A only with probability $3/7$, i.e. less than 0.5, revealing now a slight preference for Beethoven. Debreu's critique also applies to model (5) although in a weaker form. According to model (5), the above decision maker chooses A with probability 0.5, B —with probability $1/4$ and C —with probability $1/4$ as well in a ternary choice among A , B and C . The dual choice axiom, as well as the classic choice axiom, are intuitively appealing in choice situations without any significant substitution/complementary effects among available alternatives. Debreu's example clearly does not fall into this category as the two versions of Beethoven's symphony are highly substitutable.

McKelvey and Palfrey (1995) developed the concept of logit quantal response equilibrium for solving strategic games based on Luce (1959) choice model. A promising avenue of future research is to develop an analogous equilibrium solution concept when players choose among strategies according to model (5). Compared to Luce (1959) choice model, in model (5) strategies that fall below (greatly exceed) other

¹⁴ This function must satisfy restriction $F(v_1, v_2, \dots, v_{n-1}) + F(-v_1, v_2 - v_1, \dots, v_{n-1} - v_1) + F(-v_2, v_1 - v_2, v_3 - v_2, \dots, v_{n-1} - v_2) + \dots + F(-v_{n-1}, v_1 - v_{n-1}, \dots, v_{n-2} - v_{n-1}) = 1$ for all $v_1, v_2, \dots, v_{n-1} \in \mathbb{R}$, which implies inter alia $F(0, \dots, 0) = 1/n$.

strategies in terms of expected utility are chosen with probabilities close to zero (one), which can help a decision maker to avoid dominated strategies and to choose dominant strategies more frequently.

Blavatsky (2009, 2012) presents an algorithm how to extend a model of binary probabilistic choice to choice among $n > 2$ alternatives: 1) pick at random two out of $n > 2$ alternatives and choose between them; 2) discard the less preferred alternative back into the choice set; 3) pick at random one out of $n-1$ alternatives and choose between this alternative and the previously chosen (more preferred) alternative; 4) repeat steps 2–3 ad infinitum. The dual approach proposed in this paper can be also applied to this algorithm. Such dual algorithm extends binary choice probabilities to probabilities of choosing *the worst* among $n > 2$ alternatives: 1) pick at random two out of $n > 2$ alternatives and choose between them; 2) discard *the more* preferred alternative back into the choice set; 3) pick at random one out of $n-1$ alternatives and choose between this alternative and the previously retained (less preferred) alternative; 4) repeat steps 2–3 ad infinitum. Asymptotic probability that alternative $A \in S$ is retained as the worst alternative in the choice set S is then given by

$$Q(A|S) = \frac{\sum_{G \in \Gamma(S) | R(G)=A} \prod_{\{B,C\} \in E(G)} P(C, B)}{\sum_{G \in \Gamma(S)} \prod_{\{B,C\} \in E(G)} P(C, B)}$$

where G denotes an arborescence with the vertex set S , $I(S)$ is the set of all arborescences with the vertex set S , $R(G)$ denotes the root of arborescence G and $E(G)$ denotes the edge set of arborescence G (probability $P(C, B)$ for any $\{B, C\} \in E(G)$ is then the probability that a decision maker chooses the terminal vertex (head) C over the initial vertex (tail) B).

Funding Pavlo Blavatsky is a member of the Entrepreneurship and Innovation Chair, which is part of LabEx Entrepreneurship (University of Montpellier, France) and funded by the French government (Labex Entreprenre, ANR-10-Labex-11-01).

Appendix

Proof of Proposition 1

For any two alternatives $A, B \in S$, dual choice axiom implies Eq. (3). Similarly, for any two alternatives $B, C \in S$, dual choice axiom implies Eq. (14).

$$\frac{1 - P(B|S)}{1 - P(C|S)} = \frac{P(C, B)}{P(B, C)} \tag{14}$$

Multiplying Eq. (3) by Eq. (14) yields Eq. (15).

$$\frac{1 - P(A|S)}{1 - P(C|S)} = \frac{P(B, A)}{P(A, B)} \frac{P(C, B)}{P(B, C)} \tag{15}$$

According to dual choice axiom the left-hand side of Eq. (15) is equal to $P(C, A)/P(A, C)$. Using this result we can rewrite Eq. (15) as Eq. (16).

$$P(A, B)P(B, C)P(C, A) = P(A, C)P(C, B)P(B, A) \quad (16)$$

Equation (16) is known as the product rule (e.g., Estes 1960, p. 272; Luce and Suppes 1965, definition 25, p. 341). According to Theorem 48 in Luce and Suppes (1965, p. 350), a binary choice probability function $P: S \times S \rightarrow \mathbb{R}$ satisfies the product rule (16) if and only if it is a binary strict utility (4). *Q.E.D.*

Proof of Proposition 2

Using the result of proposition 1, for any two alternatives $A, B \in S$, we can rewrite dual choice axiom (3) as Eq. (17).

$$\frac{1-P(A|S)}{1-P(B|S)} = \frac{U(B)}{U(A)} \quad (17)$$

Equation (17) can be rearranged as Eq. (18).

$$P(B|S) = 1 - [1 - P(A|S)] \frac{U(A)}{U(B)} \quad (18)$$

Since Eq. (18) holds for any alternative $B \in S$, we can sum it over all $B \in S$ to obtain Eq. (19).

$$\sum_{B \in S} P(B|S) = n - [1 - P(A|S)] \sum_{B \in S} \frac{U(A)}{U(B)} \quad (19)$$

The left-hand side of Eq. (19) is equal to one due to (1). Thus, we can rewrite Eq. (19) as Eq. (20).

$$[1 - P(A|S)] \sum_{B \in S} \frac{U(A)}{U(B)} = n - 1 \quad (20)$$

A simple rearranging of Eq. (20) then yields model of probabilistic choice (5). *Q.E.D.*

References

- Agranov, M., & Ortoleva, P. (2017). Stochastic choice and preferences for randomization. *Journal of Political Economy*, 125(1), 40–68.
- Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica*, 76(3), 583–618.
- Ballinger, P., & Wilcox, N. (1997). Decisions, error and heterogeneity. *Economic Journal*, 107(443), 1090–1105.
- Becker, G. M., DeGroot, M. H., & Marschak, J. (1963). Stochastic models of choice behavior. *Behavioral Science*, 8, 41–55.

- Birnbaum, M. (2005). Three new tests of independence that differentiate models of risky decision making. *Management Science*, 51(9), 1346–1358.
- Birnbaum, M., & Navarrete, J. (1998). Testing descriptive utility theories: Violations of stochastic dominance and cumulative independence. *Journal of Risk Uncertainty*, 17(1), 49–78.
- Blavatsky, P. (2006). Violations of betweenness or random errors? *Economics Letters*, 91(1), 34–38.
- Blavatsky, P. R. (2009). How to extend a model of probabilistic choice from binary choices to choices among more than two alternatives. *Economics Letters*, 105(3), 330–332.
- Blavatsky, P. R. (2012). Probabilistic choice and stochastic dominance. *Economic Theory*, 50(1), 59–83.
- Blavatsky, P. (2014). Stronger utility. *Theory and Decision*, 76(2), 265–286.
- Blavatsky, P. (2017). Probabilistic intertemporal choice. *Journal of Mathematical Economics*, 73, 142–148.
- Blavatsky, P. (2018). Fechner's strong utility model for choice among $n > 2$ alternatives: Risky lotteries, Savage acts, and intertemporal payoffs. *Journal of Mathematical Economics*, 79, 75–82.
- Blavatsky, P., & Maafi, H. (2018). Estimating representations of time preferences and models of probabilistic intertemporal choice on experimental data. *Journal of Risk and Uncertainty*, 56(3), 259–287.
- Butler, D. J., & Loomes, G. C. (2007). Imprecision as an account of the preference reversal phenomenon. *American Economic Review*, 97(1), 277–297.
- Butler, D. J., & Loomes, G. C. (2011). Imprecision as an account of violations of independence and betweenness. *Journal of Economic Behavior and Organization*, 80(3), 511–522.
- Camerer, C. F. (1989). An experimental test of several generalized utility theories. *Journal of Risk and Uncertainty*, 2(1), 61–61.
- Camerer, C., & Ho, T. (1994). Violations of the betweenness axiom and nonlinearity in probability. *Journal of Risk and Uncertainty*, 8(2), 167–196.
- Carbone, E. (1997). Investigation of stochastic preference theory using experimental data. *Economics Letters*, 57(3), 305–311.
- Carbone, E., & Hey, J. (1995). A comparison of the estimates of EU and non-EU preference functionals using data from pairwise choice and complete ranking experiments. *Geneva Papers on Risk and Insurance Theory*, 20(1), 111–133.
- Chew, S., Epstein, L., & Segal, U. (1991). Mixture symmetry and quadratic utility. *Econometrica*, 59(1), 139–163.
- Coller, M., & Williams, M. (1999). Eliciting individual discount rates. *Experimental Economics*, 2(2), 107–127.
- de Condorcet, M. (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Paris: L'Imprimerie Royale.
- Debreu, G. (1960). Individual choice behavior: A theoretical analysis by R. Duncan Luce. *American Economic Review*, 50, 186–188.
- Estes, W. K. (1960). A random-walk model for choice behavior. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), *Mathematical methods in the social sciences* (pp. 265–276). Stanford: Stanford University Press.
- Falmagne, J.-C. (1985). *Elements of psychophysical theory*. New York: Oxford University Press.
- Fechner, G. (1860). *Elements of psychophysics*. New York: Holt, Rinehart and Winston.
- Harless, D., & Camerer, C. (1994). The predictive utility of generalized expected utility theories. *Econometrica*, 62(6), 1251–1289.
- Harrison, G., Lau, M., & Williams, M. (2002). Estimating individual discount rates in Denmark: A field experiment. *American Economic Review*, 92(5), 1606–1617.
- Hey, J. D. (2001). Does repetition improve consistency? *Experimental Economics*, 4(1), 5–54.
- Hey, J., & Carbone, E. (1995). Stochastic choice with deterministic preferences: An experimental investigation. *Economics Letters*, 47, 161–167.
- Hey, J. D., & Orme, C. (1994). Investigating generalisations of expected utility theory using experimental data. *Econometrica*, 62(6), 1291–1326.
- Hey, J. D., Lotito, G., & Maffioletti, A. (2010). The descriptive and predictive accuracy of theories of decision making under uncertainty/ambiguity. *Journal of Risk and Uncertainty*, 41(2), 81–111.
- Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5), 1644–1655.
- Loomes, G. (2005). Modelling the stochastic component of behaviour in experiments: Some issues for the interpretation of data. *Experimental Economics*, 8(4), 301–323.
- Loomes, G., & Sugden, R. (1995). Incorporating a stochastic element into decision theories. *European Economic Review*, 39(3–4), 641–648.
- Loomes, G., & Sugden, R. (1998). Testing different stochastic specifications of risky choice. *Economica*, 65(260), 581–598.

- Loomes, G., Moffatt, P., & Sugden, R. (2002). A microeconomic test of alternative stochastic theories of risky choice. *Journal of Risk and Uncertainty*, 24(2), 103–130.
- Luce, R. D. (1959). *Individual choice behavior*. New York: Wiley.
- Luce, R. D., & Suppes, P. (1965). Preference, utility, and subjective probability. In *Handbook of mathematical psychology* (Vol. III, pp. 249–410). New York: Wiley.
- Machina, M. (1985). Stochastic choice functions generated from deterministic preferences over lotteries. *Economic Journal*, 95(379), 575–594.
- Marley, A. A. J. (1965). The relation between the discard and regularity conditions for choice probabilities. *Journal of Mathematical Psychology*, 2(2), 242–253.
- Marley, A. A. J., & Louviere, J. J. (2005). Some probabilistic models of best, worst, and best-worst choices. *Journal of Mathematical Psychology*, 49(6), 464–480.
- McFadden, D. (1976). Quantal choice analysis: A survey. *Annals of Economic and Social Measurement*, 5, 363–390.
- McKelvey, R., & Palfrey, T. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1), 6–38.
- Meier, S., & Sprenger, C. (2015). Temporal stability of time preferences. *Review of Economics and Statistics*, 97(2), 273–286.
- Rieskamp, J., Busemeyer, J., & Mellers, B. (2006). Extending the bounds of rationality. *Journal of Economic Literature*, 44(3), 631–661.
- Samuelson, P. (1937). A note on measurement of utility. *The Review of Economic Studies*, 4(2), 155–161.
- Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley.
- Starmer, C., & Sugden, R. (1989). Probability and juxtaposition effects: An experimental investigation of the common ratio effect. *Journal of Risk and Uncertainty*, 2(2), 159–178.
- von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior* (Second ed.). Princeton: Princeton University Press.
- Warner, J., & Pleeter, S. (2001). The personal discount rate: Evidence from military downsizing programs. *American Economic Review*, 91(1), 33–53.
- Wilcox, N. (2008). Stochastic models for binary discrete choice under risk: A critical primer and econometric comparison. In J. C. Cox & G. W. Harrison (Eds.), *Research in experimental economics* (Vol. 12, pp. 197–292). Bingley: Emerald.
- Wilcox, N. (2011). Stochastically more risk averse: A contextual theory of stochastic discrete choice under risk. *Journal of Econometrics*, 162(1), 89–104.
- Wu, G., & Gonzalez, R. (1996). Curvature of the probability weighting function. *Management Science*, 42(12), 1676–1690.
- Yellott Jr., J. I. (1977). The relationship between Luce’s choice axiom, Thurstone’s theory of comparative judgement, and the double exponential distribution. *Journal of Mathematical Psychology*, 15(2), 109–144.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.