CESifo WORKING PAPERS

7158
2018

July 2018

Quantum Decision Theory and the Ellsberg Paradox

Ali al-Nowaihi, Sanjit Dhami, Mengxing Wei

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7158*
Category 13: Behavioural Economics

Quantum Decision Theory and the Ellsberg Paradox

Abstract

We formulate a simple quantum decision model of the Ellsberg paradox. We report the results of an experiment we performed to test the matching probabilities predicted by this model using an incentive compatible method. We find that the theoretical predictions of the model are in conformity with our experimental results. We compare the predictions of our quantum model with those of probably the most successful non-quantum model of ambiguity, namely, the source dependent model. The predictions of our quantum model are not statistically significantly different from those of the source dependent model. The source dependent model requires the specification of probability weighting functions in order to fit the evidence. On the other hand, our quantum model makes no recourse to probability weighting functions. This suggests that much of what is normally attributed to probability weighting may actually be due to quantum probability.

JEL-Codes: D030.

Keywords: quantum probability, the Ellsberg paradox, the source dependent model, the law of total probability, the law of reciprocity, the Feynman rules, projective expected utility, bounded rationality, Diebold-Mariano forecasting tests.

*Ali al-Nowaihi**
Division of Economics, School of Business
University of Leicester, University Road
United Kingdom – Leicester, LE1 7RH
aa10@le.ac.uk

Sanjit Dhami
Division of Economics, School of Business
University of Leicester, University Road
United Kingdom – Leicester, LE1 7RH
sd106@le.ac.uk

Mengxing Wei
International Academy of Business and Economics, Tianjin University of Finance and Economics, 25 Zhujiang Road
P.R. China – Tianjin, 300222, Hexi District
mengxing.wei@hotmail.com

*corresponding author

July 14, 2018

Highlights

- We formulate a simple quantum decision theory model that is parameter free, explains the Ellsberg paradox and is in conformity with the evidence.
- The forecasts of our quantum model are not statistically significantly different from the forecasts of the source dependent model.
- We provide a simple introduction to quantum decision theory.

1 Introduction

Situations of ambiguity are pervasive in decision making. In this paper we investigate the potential of *quantum decision theory* (QDT) to provide an explanation. We concentrate on the canonical example of ambiguity, namely, the Ellsberg paradox (Keynes, 1921; Ellsberg 1961, 2001). The Ellsberg paradox has proved to be a particularly useful vehicle for testing models of ambiguity. In addition, there are many real-world situations that appear similar to the Ellsberg paradox. One example is that of *home-bias* in investment (French and Poterba, 1991, Obstfeld and Rogoff, 2000). Investors are often observed to prefer investing in a domestic asset over a foreign asset with the same return and the same riskiness.

A simple quantum model of the Ellsberg paradox was formulated by al-Nowaihi & Dhami (2017). Their derivation of quantum probabilities is parameter-free. Thus, their explanation of the Ellsberg paradox is more parsimonious, hence more refutable, than all the other explanations. Their predicted matching probabilities, based on their quantum model, are close to those empirically observed by Dimmock et al. (2015). The question then arises "is this agreement an accident?"[1] To test this, we performed a new experiment using a very different data set from Dimmock et al. (2015) and a different methodology.[2] To compare with the evidence reported in Dimmock et al. (2015), we chose the same probabilities as they did: $p = 0.1$, 0.5 and 0.9. Another reason for choosing these probabilities is that the greatest difference between the predictions of the various models of the Ellsberg paradox lie in the tails of the distribution. Hence, we have concentrated our resources (over 250 subjects) on the region that is most likely to refute our model.

We found the predictions of our quantum model to be in agreement with the evidence. This is the first main contribution of this paper.

The second main contribution of our paper is to compare the predictive power of our quantum model with that of probably the most successful non-quantum model of ambiguity, namely, the *source dependence* (Abdellaoui et al., 2011; Kothiyal et al., 2014; Dimmock et al., 2015).[3] For this purpose we performed three Diebold-Mariano forecasting tests (Diebold & Mariano, 1995, Diebold, 2014).[4] For each test, the difference between the predictive performance of the source dependent model and our quantum model was not statistically significant. However, the source dependent model requires the specification of probability weighting functions. In fact, the source dependent model can fit any data set depending on the choice of probability weighting functions. On the other hand, our quantum model makes no recourse to probability weighting functions. This suggests that much of what has been attributed to probability weighting might actually be due to quantum probability.

This paper incorporates the results of al-Nowaihi and Dhami (2017). We take the opportunity here to clarify the role played by the quantum law of reciprocity in deriving quantum probabilities (Proposition 7, below). This was left implicit in al-Nowaihi and Dhami (2017).

The rest of the paper is organized as follows. Our model is formulated is section 2. Section 3 reviews standard (Kolmogorov) probability theory. Sections 4-7 briefly review decision theories that are based on standard (Kolmogorov) probability theory, with particular reference to the Ellsberg paradox. Section 8 reviews the elements of quantum probability theory needed

[1]We are grateful to Jerome Busemeyer for raising this issue.

[2]Dimmock et al. (2015) use 666 Dutch subjects from the non-student population. Our subjects were 295 undergraduate students from Qingdao Agricultural University in China.

[3]We are grateful to Stefan Trautmann for suggesting this to us.

[4]We are grateful to Wojciech Charemza for suggesting this test to us.

for this paper. Our quantum analysis of the Ellsberg paradox is given in section 9. Section 10 explains how we elicit certainty equivalents of lotteries and from these infer matching probabilities. Section 11 gives our experimental design. Our experimental results are in section 12. Section 13 compares the forecasting performance of our quantum model with that of the source dependent model. Section 14 summarizes. Our experimental instructions are given in Appendix A. Appendix B gives our post-experimental questionnaire.

2 Formulation

2.1 An Ellsberg experiment and the Ellsberg paradox

We consider the following version of the Ellsberg experiment (Dimmock et al., 2015). This involves two urns: The known urn (K) contains nk balls of n different colors and k balls of each color. The unknown urn (U) also contains nk balls of the same n colors as urn K but in unknown proportions. The subject is presented with the following bet. Suppose l of the n colors are chosen to be winning colors (hence, urn K contains lk balls of the winning colors). The subject wins a prize if a randomly drawn ball from an urn is of the winning color. The question we address is "which urn would a subject choose?" The classical answer is that a subject should be indifferent between the two urns, they should exhibit *ambiguity neutrality*. However, the evidence is that subjects prefer the known urn (K) for medium and high probabilities but prefer the unknown urn (U) for low probability. Thus, the evidence is that subjects exhibit *ambiguity aversion* for medium and high probabilities but *ambiguity seeking* for low probabilities. This behavior is known as *insensitivity*. Thus, classical theory predicts ambiguity neutrality while the evidence reveals insensitivity. This is known as the *Ellsberg paradox*.

2.2 Matching probabilities

Consider subject i. Let p be the probability with which subject i draws a ball of a winning color from urn K. Keep the contents of urn U fixed, but construct a new known urn, K_i, with a known number, M_i, of balls of the winning colors such that subject i is indifferent between urns K_i and U. Let $m_i(p)$ be the probability with which that subject draws a ball of a winning color from urn K_i. Then $m_i(p)$ is the *matching probability* of p for subject i.

2.3 Utility

Let u_i be the utility of subject i, assumed to be strictly increasing and normalized so that $u_i(0) = 0$.

2.4 Quantum probability theory (Assumption Q)

Our main assumption (Q) is that subjects' behavior is determined by quantum probability theory (von Neumann, 1955, original German, 1932), rather than standard Kolmogorov probability theory (Kolmogorov, 1950, original German, 1933). See section 9. However, we shall need further auxiliary assumptions, which we now introduce.

2.5 A behavioral assumption on how urn U is constructed in a subject's mind (Assumption B)

The framing of information is vital in choices. Subjects often simplify complex problems before solving them (Dhami, 2016). For Ellsberg experiments, subjects are typically told that urn U contains the same number of balls of the same colors as urn K, but in unknown proportions. However, the term "unknown proportions" is not defined any further, which raises the question of how subjects perceive this term. There is strong evidence that this is too cognitively challenging for subjects and that subjects do not consider all possible distributions of balls in urn U (Pulford & Colman, 2008).

We conjecture (Assumption B) that subjects model "unknown proportions" in a simple way (al-Nowaihi & Dhami, 2017) as described below.

1. We replace colors by numerals. Furthermore, we consider only two numerals: 1 and 2. The known urn K contains kn balls, kl of which are labeled "1" and $kn-kl$ are labeled "2". Ball 1 is drawn from K with probability p and ball 2 is drawn from K with probability $1-p$. How p is related to l and n will be discussed in subsection 2.7, below. This transformation is only for analytic convenience. In our experiments subjects are always presented with colored balls whose ratios match the probabilities.

2. Point 1 allows us to consider urn K as having just two balls. One of the balls, the winning ball, labeled "1", is drawn with probability p. The other ball, labeled "2", is drawn with probability $1-p$. Likewise

urn U will also have two balls labeled 1 and 2 but the proportions will be unknown, as the following construction shows.

3. Urn K has two balls, labeled 1 and 2, while urn U is initially empty. We conjecture that in the mind of a subject urn U is constructed as follows. In two successive and independent rounds, a ball is drawn at random from urn K and placed in urn U without revealing the labels, 1 or 2, to the subject. At the end of each of the two rounds, the ball that was drawn from urn K is replaced with an identically labeled ball. At the end of the two rounds, urn U contains two balls. The possibilities are that both could be labeled 1, both could be labeled 2, or one could be labeled 1 and the other labeled 2.

4. A ball is drawn at random from whichever urn the subject chooses (K or U). The subject wins a monetary prize $v > 0$ if ball 1 is drawn but wins nothing if ball 2 is drawn.

Based on the above construction, we may define the following states of urn U:

1. $\mathbf{s}_1$ is the state where ball 1 is drawn in each of the two rounds (each with probability p).

2. $\mathbf{s}_2$ is the state where ball 1 is drawn in round one (probability p), then ball 2 is drawn in round two (probability $1 - p$).

3. $\mathbf{s}_3$ is the state where ball 2 is drawn in round one (probability $1 - p$), then ball 1 is drawn in round two (probability p).

4. $\mathbf{s}_4$ is the state where ball 2 is drawn in each of the two rounds (each with probability $1 - p$).

Concerning the rationality of assumption (B), we can take one of two positions. We could view this as an error on behalf of the subjects or we could view this as a consequence of their bounded rationality. We prefer the latter. As an example, consider engineers. They use a finite decimal expansion of π. From a logical point of view, this is an error, because π is irrational. However, we do not view this as caused by irrationality of engineers (although they do make errors) but as a simplification necessitated by their bounded rationality.

2.6 Projective expected utility (L)

We need to relate the quantum probabilities predicted by our model (Proposition 7 of section 9) to matching probabilities (subsection 2.2, above, and Proposition 8 of section 9, below). The appropriate decision theory is projective expected utility (La Mura, 2009). This is essentially expected utility theory but with quantum probabilities replacing Kolmogorov probabilities. Since all the lotteries we consider have only two outcomes: a zero outcome and a positive outcome, projective expected utility appears entirely adequate for our purposes.

2.7 The heuristic of insufficient reason (Assumption I)

In both classical (Kolmogorov) probability theory and quantum probability theory any probabilities (provided they are non-negative and sum to 1) can be assigned to the elementary events. To make a theory predictive, some heuristic rule is needed to assign a' priori probabilities. The heuristic commonly used is that of *insufficient reason* or *equal a' priori probabilities* (we call this a heuristic because it does not follow from either classical or quantum probability theory).

As an application of the heuristic of insufficient reason, consider the Ellsberg experiment of subsection 2.1, above. The known urn (K) contains nk balls of n different colors and k balls of each color. Suppose l of the n colors are chosen to be winning colors, hence, urn K contains lk balls of the winning colors. Since subject i has no reason to think that one color is more likely than another, subject i should assign the probability $p = \frac{lk}{nk} = \frac{l}{n}$ to drawing a ball of a winning color from urn K.

To be sure, this heuristic is not without problems. See, for example, Gnedenko (1968) sections 5 and 6, pp. 37-52. This heuristic is crucial in deriving the Maxwell-Boltzmann distribution in classical statistical mechanics and the Bose-Einstein and Fermi-Dirac distributions in quantum statistical mechanics. See Tolman (1938) section 23, pp. 59-62, for a good early discussion.

2.8 A power function form for the utility function (Assumption P)

The mechanism used by Dimmock et al. (2015) is not incentive compatible. Specifically, Dimmock et al. (2015) constructed urn K_i as follows. The ratio of the colors (whatever they are) in U were kept fixed. However, the ratio in K_i was varied until subject i declared indifference between K_i and U. It turns out that in this method of eliciting matching probabilities subjects have the incentive to declare a preference for U over K_i, even when the reverse is true. However Dimmock et al. (2015) found no evidence in their data that this occurred. Dimmock et al. (2015, pp. 26): "In chained questions, where answers to some questions determine subsequent questions, subjects may answer strategically (Harrison, 1986). In our experiment, this is unlikely. First, our subjects are less sophisticated than students. Second, it would primarily have happened in the end (only after discovery), at the 0.9 probability event, where it would increase ambiguity seeking. However, here we found strong ambiguity aversion".

In this paper, we use the incentive compatible mechanism of Fox & Tversky (1995), study 2. Specifically, we elicit certainty equivalents of lotteries, then infer the corresponding matching probabilities (Proposition 9 of section 10, below). However, his method requires the specification of a utility function, u_i, for each subject, i. This is not required by the Dimmock et al. (2015) mechanism. We use the power function form[5]

$$u_i(x) = x^{\sigma_i}, x \geq 0, \sigma_i > 0. \tag{1}$$

This introduces a free parameter, σ_i. Note, however, that σ_i is only used to give a parsimonious description of the behavior of subjects. In particular, σ_i is not chosen to make the predictions of the theory fit the evidence. The matching probabilities predicted by the theory are parameter-free and are based on assumptions (Q), (B) and (L) only (see Propositions 7 and 8 of section 9, below).

2.9 Discussion

The first assumption (Q) is the main assumption. However, no mathematical structure on its own will yield empirically testable predictions; auxiliary

[5]The power form of the utility function is a popular choice. See Tversky Kahneman (1992), Prelec (1998) and Vincent et al. (2017), to mention just three examples.

assumptions are need. For example, in Newtonian mechanics, in addition to Newton's second law of motion and law of gravity, we need initial conditions and simplifying assumptions. Calculus on its own will not yield empirically testable predictions. In quantum mechanics we need, for example, the momentum operator to be $p_x = -i\frac{h}{2\pi}\frac{\partial}{\partial x}$ and we need to specify a Hamiltonian for the system. Hilbert space on its own is insufficient. In this respect this paper is no exception. In addition to the assumptions of quantum probability theory, we employ the two auxiliary assumptions (B) and (L). These three assumptions (Q&B&L) are sufficient to theoretically derive the matching probabilities (Propositions 7 and 8 of section 9).

Testing any theory requires further assumptions. For example, to test Newton's prediction of the orbits of the planets we need to make assumptions about the human eye, the telescope and the atmosphere. We have added assumptions (I) and (P) for the purpose of testing the theory. Thus, our test is a test of the conjunction Q&B&L&I&P. If we reject this conjunction, then this is a rejection of, at least, one of them, but we would not know which. On the other hand, since Q&B&L&I&P is true if, and only if, all of these are true, then a confirmation is a confirmation of each one of them. However, a confirmation is not a proof. It is merely a failure to reject. Hence further tests may lead to a rejection. No number of confirmations, however large, can prove a theory. The most we can say about a theory, any theory, is that it has so far survived the tests.

3 Standard (Kolmogorov) probability theory

In this section we give a brief review of standard probability theory, also known as Kolmogorov probability theory after Kolmogorov (1950), original German (1933). We do this for two reasons. First, because it is fundamental to all decision theories. Second, to make clear the similarities and differences with quantum probability (section 8). Probabilities can be either *objective*, in the sense that they are the same for all decision makers, or they can be *subjective* in the sense that they can differ across decision makers. In the latter case, they can be elicited from a decision maker's observed choices, given the decision theory under consideration (Wakker, 2010).

3.1 Sample space

In the standard approach we have a non-empty set, Ω, called the *sample space*, and a σ-*algebra*, S, of subsets of Ω. The elements of S are called *events*. S has the following properties: $\emptyset \in S$, $X \in S \Rightarrow \Omega - X \in S$ (hence, $\Omega \in S$), $\{X_i\}_{i=1}^{\infty} \subset S \Rightarrow \cup_{i=1}^{\infty} X_i \in S$ (hence, $\cap_{i=1}^{\infty} X_i \in S$). Note that the distributive laws hold: $X \cap \left(\cup_{j=1}^{\infty} Y_j\right) = \cup_{j=1}^{\infty} \left(X \cap Y_j\right)$ and $X \cup \left(\cap_{j=1}^{\infty} Y_j\right) = \cap_{j=1}^{\infty} \left(X \cup Y_j\right)$.

3.2 Probability measures

A probability measure is then defined as a function, $P : S \to [0,1]$ with the properties that $P(\emptyset) = 0$, $P(\Omega) = 1$ and if $X_i \cap X_j = \emptyset$, $i \neq j$, then $P\left(\cup_{i=1}^{\infty} X_i\right) = \sum_{i=1}^{\infty} P(X_i)$.

3.3 Conditional probabilities, Bayes' law and the law of total probability

Let $X, Y \in S$. Define $P(X|Y) = \frac{P(X \cap Y)}{P(Y)}$, if $P(Y) \neq 0$ and $P(X|Y) = 0$, if $P(Y) = 0$. $P(X|Y)$ is called the probability of X *conditional* on Y. From this we can derive *Bayes law*: $P(X|Y) = \frac{P(Y|X) P(X)}{P(Y)}$, for $P(Y) \neq 0$, and its other equivalent forms. Let $Y \in S$, then $P(X|Y)$ is a probability measure on the set $\{X \in S : X = Z \cap Y, \text{ for some } Z \in S\}$. Importantly, the *law of total probability* holds: Let $X \in S$ and let $\{Y_i\}_{i=1}^{n}$ be a partition of Ω, so $Y_i \in S$, $Y_i \neq \emptyset$, $\cup_{i=1}^{n} Y_i = \Omega$, $Y_i \cap Y_j = \emptyset$ for $i \neq j$; then $P(X) = \sum_{i=1}^{n} P(X|Y_i) P(Y_i)$.

3.4 Random variables

A *random variable* is a mapping, $f : \Omega \to \mathbb{R}$ satisfying: For each $r \in \mathbb{R}$, $\{x \in \Omega : f(x) \leq r\} \in S$. A random variable, f, is *non-negative* if $f(x) \geq 0$ for each $x \in \Omega$. For two random variable, f, g, we write $f \leq g$ if $f(x) \leq g(x)$ for each $x \in \Omega$. A random variable, f, is *simple* if its range is finite. For any random variable, f, and any $x \in \Omega$, let $f^+(x) = \max\{0, f(x)\}$ and $f^-(x) = -\min\{0, f(x)\}$. Then, clearly, f^+ and f^- are both non-negative random variables and $f(x) = f^+(x) - f^-(x)$, for each $x \in \Omega$. We write this as $f = f^+ - f^-$.

Let f be a simple random variable with range $\{f_1, f_2, ..., f_n\}$. Let $X_i = \{x \in \Omega : f(x) = f_i\}$. Then, $X_i \in S$, $X_i \cap X_j = \emptyset$ for $i \neq j$ and $U_{i=1}^{n} X_i = \Omega$. The *expected value* of the simple random variable, f, is $E(f) = \sum_{i=1}^{n} f_i P(X_i)$.

The *expected value* of the non-negative random variable, g, is $E(g) = \sup\{E(f) : f \text{ is a simple random variable and } f \le g\}$. Note that $E(g)$ may be infinite. If $f = f^+ - f^-$ is an arbitrary random variable such that not both $E(f^+)$ and $E(f^-)$ are infinite, then the expected value of f is $E(f) = E(f^+) - E(f^-)$. Note that $E(f)$ can be $-\infty$, finite or ∞. However, if both $E(f^+)$ and $E(f^-)$ are infinite then $E(f)$ is undefined (because $\infty - \infty$ is undefined).

3.5 The Ellsberg paradox under Kolmogorov probability theory

Our behavioral assumption, Assumption B (subsection 2.5, above), about how a subject mentally constructs urn A will play an essential role in our quantum explanation of the Ellsberg paradox. The question then arises whether this behavioral assumption can also explain the Ellsberg paradox when combined with classical (Kolmogorov) probability theory. Proposition 1, below, establishes that this is not the case.

Proposition 1 *: Assume (B). If the probability of drawing a winning ball from the known urn K is p, then the classical probability of drawing a winning ball from the ambiguous urn A is also p.*

Proof of Proposition 1: Let X be the event where a winning ball (ball 1) is drawn from urn A. Let Y_i be the event that urn A is in state $\mathbf{s}_i$, $i = 1, 2, 3, 4$, defined in subsection 2.5, above. By the law of total probability, we then have:

$$P(X) = P(X|Y_1)\,P(Y_1) + P(X|Y_2)\,P(Y_2) + P(X|Y_3)\,P(Y_3) + P(X|Y_4)\,P(Y_4)\,. \tag{2}$$

We have $P(Y_1) = p^2$, $P(Y_2) = p(1-p)$, $P(Y_3) = (1-p)\,p$, $P(Y_4) = (1-p)^2$, $P(X|Y_1) = 1$, $P(X|Y_2) = \frac{1}{2}$, $P(X|Y_3) = \frac{1}{2}$, $P(X|Y_4) = 0$. Hence, from (2), we get:

$$P(X) = p. \tag{3}$$

Hence, if the probability of drawing ball 1 from the known urn K is p, then the classical probability of drawing ball 1 from the ambiguous urn A is also p. ■

Thus, even with our behavioral assumption (B), the classical (Kolmogorov) probability treatment gives the same probability, p, of winning whether a subject chooses urn K or urn A. Hence, a subject has no reason to prefer U over K or K over U on probabilistic grounds.

Keynes (1921) pointed out that there is a difference in the strength or quality of the evidence. Subjects may reason that, although the assignment of the same probability to each color is sound, they are more confident in the correctness of this judgement in the case of K than in the case of U. Hence, they prefer K to U. Thus their preference works through the utility channel rather than the probability channel. However, this explanation appears to be contradicted by the evidence of Dimmock et al. (2015) that subjects are ambiguity seeking for low probabilities.

4 Expected utility theory (EU)

It will be sufficient for our purposes to consider a partition of Ω into a finite set of exhaustive and mutually exclusive events: $\Omega = \cup_{i=1}^{n} X_i$, $X_i \neq \emptyset$, $X_i \cap X_j = \emptyset$ for $i \neq j$, $i = 1, 2, ..., n$. A decision maker can take an *action* $a \in \Lambda$ that results in *outcome* $o_i(a) \in O$ and *utility* $u(o_i(a))$ if the event X_i occurs, where $u : O \to \mathbb{R}$. The decision maker chooses an action, $a \in \Lambda$, before knowing which event, X_i, will occur or has occurred. Let p_i be the probability with which event X_i occurs. Then the decision maker's *expected utility* from choosing the action $a \in A$ is $Eu(a) = \sum_{i=1}^{n} p_i u(o_i(a))$. The decision maker prefers action $a \in \Lambda$ over action $b \in \Lambda$ if $Eu(a) \geq Eu(b)$. The preference is strict if $Eu(a) > Eu(b)$. The decision maker is indifferent between a and b if $Eu(a) = Eu(b)$. The probabilities p_i, $i = 1, 2, ..., n$, can either be *objective* (the same for all decision makers, von Neumann and Morgenstern, 1947) or *subjective* (possibly different for different decision makers, Savage, 1954). In the latter case, it follows from Savage's axioms that these probabilities can be uniquely elicited from the decision maker's behavior. Note that the action $a \in \Lambda$ results in the lottery $(o_1(a), X_1; o_2(a), X_2; ...; o_n(a), X_n)$, i.e., the lottery that results in outcome $o_i(a)$ if the event X_i occurs. In terms of probabilities this lottery can be written as $(o_1(a), p_1; o_2(a), p_2; ...; o_n(a), p_n)$, i.e., the lottery that results in outcome $o_i(a)$ with probability p_i. Sometimes it is more convenient to write the lottery explicitly rather than the action that gave rise to it.

4.1 The Ellsberg paradox under expected utility theory

We now apply expected utility theory to the Ellsberg experiment.

Proposition 2 : *Under, expected utility theory a subject should be ambiguity neutral.*

Proof of Proposition 2: The subject can choose either K or A. From Proposition 1, we see that either action results in the outcome $v > 0$ with probability p or 0 with probability $1 - p$. Recalling that $u(0) = 0$, the subject's expected utility, in either case, is $Eu(K) = Eu(A) = pu(v)$. Hence, the subject is ambiguity neutral. ■

However, the evidence indicates that subjects exhibit insensitivity.

5 The smooth ambiguity model (SM)

The smooth ambiguity model (Klibanoff et al., 2005) is currently the most popular theory in economics for modelling ambiguity. It encompasses several earlier theories as special limiting cases. These include von Neumann and Morgenstern (1947), Hurwicz (1951), Savage (1954), Luce and Raiffa (1957), Gilboa and Schmeidler (1989) and Ghirardato et al. (2004). Conte and Hey (2013) find it provides the most satisfactory account of ambiguity[6].

For our purposes, it will be sufficient to consider the following special case of the smooth model. Recall that under expected utility theory (section 4), a decision maker chooses an action $a \in \Lambda$ that results in the outcome, $o_j(a)$, with probability, q_j, $j = 1, 2, ..., n$. The outcome, $o_j(a)$, yields the utility $u(o_j(a))$ to the decision maker. Hence, her expected utility is $Eu(a) = \sum_{j=1}^{n} q_j u(o_j(a))$. Now suppose that the decision maker is unsure of the probability distribution $(q_1, q_2, ..., q_n)$. Furthermore, she believes that the distribution $(q_{i1}, q_{i2}, ..., q_{in})$ will occur with probability p_i, $i = 1, 2, ..., m$. To characterize the decision maker's attitude to ambiguity, a new function, $\varphi : \mathbb{R} \to \mathbb{R}$, is introduced and is assumed to be increasing. Then the decision maker's *expected utility* under the smooth model that results from choosing the action $a \in \Lambda$ is $SU(a) = \sum_{i=1}^{m} p_i \varphi\left(\sum_{j=1}^{n} q_{ij} u(o_j(a))\right)$. Thus, the sequence of moves is as follows. First, the decision maker chooses the action

[6] However, Kothiyal et al. (2014) disagree, see below.

$a \in \Lambda$, then nature chooses the distribution $(q_{i1}, q_{i2}, ..., q_{in})$ with probability p_i. Finally, nature moves again and chooses the out come $o_j(a)$ with probability q_{ij}.

The smooth model reduces to expected utility theory in the following two cases: (1), $m = 1$, so there is no ambiguity, (2), φ is positive affine (i.e., $\varphi(x) = \beta x$, $\beta > 0$, $x \in \mathbb{R}$).

Suppose $m > 1$, so we do have genuine ambiguity. If φ is strictly concave, then the smooth model can explain ambiguity aversion. It can explain ambiguity seeking, if φ is strictly convex. But it cannot explain insensitivity (i.e., ambiguity seeking for low probabilities and ambiguity aversion for high probabilities) because φ cannot be both strictly concave and strictly convex.

6 Rank dependent expected utility theory (RDU)

The considerable refutations of EU have motivated many developments. One of the most popular of these is *rank dependent expected utility theory* (RDU). Recall that in EU (section 4) probabilities enter the objected function, $Eu(a) = \sum_{i=1}^{n} p_i u(o_i(a))$, linearly. However, in RDU, probabilities enter the objective function in a non-linear, though precise, way. We start with a *probability weighting function*, which is a strictly increasing function $w : [0,1] \overset{onto}{\to} [0,1]$, hence $w(0) = 0$ and $w(1) = 1$. Typically, low probabilities are overweighted and high probabilities are underweighted. The probability weighting function is applied to the cumulative probability distribution. Hence, it transforms it into another cumulative probability distribution. Hence, we may view RDU as EU applied to the transformed probability distribution. The attraction of this is that the full machinery of risk analysis developed for EU can be utilized by RDU (Quiggin, 1982, 1993). We now give the details.

Consider a decision maker who can take an action, $a \in \Lambda$, that results in outcome, $o_i(a) \in O$, with probability p_i, $i = 1, 2, ..., n$, $p_i \geq 0$, $\sum_{i=1}^{n} p_i = 1$. The decision maker has a utility function, $u : O \to \mathbb{R}$. The decision maker has to choose her action before the outcome is realized. Order outcomes in increasing magnitude. Assuming an increasing utility function, this gives: $u(o_1(a)) \leq u(o_2(a)) \leq ... \leq u(o_n(a))$. Define *decision weights*, π_i, $i = 1, 2, ..., n$, as follows. $\pi_n = w(p_n)$, $\pi_i = w\left(\sum_{j=i}^{n} p_j\right) - w\left(\sum_{j=i+1}^{n} p_j\right)$, $i = 1, 2, ..., n-1$. The decision maker's rank dependent expected utility is then

$RDU\left(a\right)=\sum_{i=1}^{n}\pi_i u\left(o_i\left(a\right)\right)$. Expected utility theory (EU) is obtained by taking $w\left(p\right)=p$. Empirical evidence shows that typically $w\left(p\right)$ is inverse-S shaped, so low probabilities are overweighted but high probabilities are underweighted. Probabilities in the middle range are much less affected. It is important to note that this need not be because decision makers misperceive probabilities (although that does happen). Rather, the weights people assign to utilities are much more sensitive to probability changes near 0 and near 1 compared to probability changes in the the middle range.[7]

If a subject chooses urn K, then her rank dependent expected utility is $RDU\left(K\right)=w\left(p\right)u\left(v\right)$. If she chooses urn A, then her rank dependent expected utility is $RDU\left(A\right)=w\left(p\right)u\left(v\right)=RDU\left(K\right)$. Hence, a decision maker obeying RDU will exhibit ambiguity neutrality. Thus, just like EU, RDU is not consistent with insensitivity.

Two important extensions of RDU that we do not review here are cumulative prospect theory (Tversky and Kahneman, 1992) and Choquet expected utility (Gilboa 1987, 2009, Schmeidler, 1989). Cumulative prospect theory extends RDU by including reference dependence and loss aversion from Kahneman and Tversky (1979). Choquet expected utility extends RDU by replacing probability weighting functions with more general capacities (Choquet, 1953-1954). Like a probability measure, a capacity is defined on a σ-algebra of subsets of a set. However, unlike a probability measure, a capacity need not be additive. By contrast, the quantum probability measure is an additive measure but defined on the lattice of closed subspaces of a Hilbert space, rather than a σ-algebra of subsets of a set. Further extensions of both are reviewed in Wakker (2010). Despite their importance, these extensions are not immediately relevant to the results of this paper.

7 The source dependent model (SDM)

The source dependent model (SDM) is probably the most successful classical (i.e., non-quantum) model of ambiguity (Abdellaoui et al., 2011; Kothiyal et al., 2014; Dimmock et al., 2015). It requires, for each subject i, the specification of two probability weighting functions, w_{iK} and w_{iU}, one for urn K and one for urn U. A *probability weighting function* is a strictly increasing function, $w:[0,1]\overset{onto}{\to}[0,1]$. The resulting *source dependent expected utility*

[7] This feature enables RDU to account for the Allais paradox.

from urns K and U are, respectively (recall that $u_i(0) = 0$),

$$Eu_i(K) = w_{iK}(p)\, u_i(v), \text{ for urn } K, \tag{4}$$

$$Eu_i(U) = w_{iU}(p)\, u_i(v), \text{ for urn } U. \tag{5}$$

Proposition 3 *: Let p be the probability of drawing a winning ball from Urn K. Let $m_i^*(p)$ be the matching probability predicted by the source dependent model. Let w_{iK}, w_{iU} be the probability weighting functions for urns K and U, respectively, then*
(a) $w_{iK}(m_i^*(p)) = w_{iU}(p)$.
(b) Subject i is ambiguity averse, ambiguity neutral or ambiguity seeking according to
(i) $Eu_i(K) \gtreqless Eu_i(U)$,
(ii) $w_{iK}(p)\, u_i(v) \gtreqless w_{iU}(p)\, u_i(v)$,
(iii) $w_{iK}(p) \gtreqless w_{iU}(p)$,
(iv) $w_{iK}(p) \gtreqless w_{iK}(m_i^*(p))$,
(v) $p \gtreqless m_i^*(p)$.

Proof of Proposition 3: From Proposition 1 it follows that p is also the probability of drawing a winning ball from Urn U. From (4), (5) and the definition of matching probability (subsection 2.2), it follows that $w_{iK}(m_i^*(p))\, u_i(v) = w_{iU}(p)\, u_i(v)$. Since $u_i(v) > 0$, it follows that $w_{iK}(m_i^*(p)) = w_{iU}(p)$. This establishes (a). Part (i) of (b) follows from the definitions of ambiguity aversion, ambiguity neutrality and ambiguity seeking. Part (ii) then follows from (4) and (5). Part (iii) follows since $u_i(v) > 0$. Part (iv) follows from parts (a) and (iii). Part (v) follows because w_{iK} is strictly increasing. ■

From Proposition 3 b(iii) it follows that if $w_{iK} = w_{iU}$, then subject i is ambiguity neutral for all p, contrary to the evidence. Hence, for the source dependent model to explain the Ellsberg paradox, we must have $w_{iK} \neq w_{iU}$.

For the Prelec (1998) probability weighting functions, we have:

$$w_{iK}(p) = e^{-\beta_{iK}(-\ln p)^{\alpha_{iK}}},\ \alpha_{iK} > 0,\ \beta_{iK} > 0,\ p \in (0,1), \tag{6}$$

$$w_{iU}(p) = e^{-\beta_{iU}(-\ln p)^{\alpha_{iU}}},\ \alpha_{iU} > 0,\ \beta_{iU} > 0,\ p \in (0,1), \tag{7}$$

Proposition 4 *: Let p be the probability of drawing a winning from urn U and $m_i^*(p)$ the matching probability predicted by the source dependent model.*

Let w_{iK} and w_{iU} be the Prelec probability weighting functions for urns K and U, respectively ((6) and (7)). Then

$$-\ln\left(-\ln m_i^*\left(p\right)\right) = \frac{\ln\beta_{iK} - \ln\beta_{iU}}{\alpha_{iK}} + \frac{\alpha_{iU}}{\alpha_{iK}}\left(-\ln\left(-\ln p\right)\right).$$

Proof of Proposition 4: Follows from (4), (5) and Proposition 3(a). ■

Proposition 5 *(Attitudes to ambiguity): Let p be the probability of drawing a winning ball from urn K. Let w_{iK} and w_{iU} be the Prelec probability weighting functions for urns K and U, respectively ((6) and (7)).*
(a) Suppose $\alpha_{iU} = \alpha_{iK}$. Then subject i is

$$\begin{aligned} &\textit{universally ambiguity averse, if } \beta_{iK} < \beta_{iU}, \\ &\textit{universally ambiguity neutral, if } \beta_{iK} = \beta_{iU}, \\ &\textit{universally ambiguity seeking, if } \beta_{iK} > \beta_{iU}. \end{aligned}$$

(b) Suppose $\alpha_{iU} < \alpha_{iK}$. Then subject i is

$$\begin{aligned} \textit{ambiguity averse for } -\ln\left(-\ln p\right) &> \frac{\ln\beta_{iK} - \ln\beta_{iU}}{\alpha_{iK} - \alpha_{iU}}, \\ \textit{ambiguity neutral for } -\ln\left(-\ln p\right) &= \frac{\ln\beta_{iK} - \ln\beta_{iU}}{\alpha_{iK} - \alpha_{iU}}, \\ \textit{ambiguity seeking for } -\ln\left(-\ln p\right) &< \frac{\ln\beta_{iK} - \ln\beta_{iU}}{\alpha_{iK} - \alpha_{iU}}. \end{aligned}$$

(c) Suppose $\alpha_{iU} > \alpha_{iK}$. Then subject i is

$$\begin{aligned} \textit{ambiguity averse for } -\ln\left(-\ln p\right) &< \frac{\ln\beta_{iK} - \ln\beta_{iU}}{\alpha_{iK} - \alpha_{iU}}, \\ \textit{ambiguity neutral for } -\ln\left(-\ln p\right) &= \frac{\ln\beta_{iK} - \ln\beta_{iU}}{\alpha_{iK} - \alpha_{iU}}, \\ \textit{ambiguity seeking for } -\ln\left(-\ln p\right) &> \frac{\ln\beta_{iK} - \ln\beta_{iU}}{\alpha_{iK} - \alpha_{iU}}. \end{aligned}$$

Proof of Proposition 5: Follows from part b(v) of Proposition 3 and Proposition 4. ■

From Proposition 5, it is clear that, for the Prelec probability weighting functions (6) and (7), the source dependent model is consistent with universal ambiguity aversion, universal ambiguity neutrality, universal ambiguity seeking, ambiguity aversion for low probabilities and ambiguity seeking for high probabilities or the reverse. It all depends on the values of the parameters α_{iK}, β_{iK}, α_{iU}, β_{iU}, which have to be estimated from the data the model seeks to explain.

Next, we wish to relate the matching probabilities, $m_i^*(p)$, predicted by the source dependent model, to the certainty equivalents we inferred under Assumption P (subsection 2.8).

Proposition 6 *: Let p be the probability of drawing a winning from urn U and $m_i^*(p)$ the matching probability predicted by the source dependent model. Let w_{iK} and w_{iU} be the Prelec probability weighting functions for urns K and U, respectively ((6) and (7)). Let v be the monetary payment to a subject if a winning ball is drawn. Let p be the probability of selecting a winning ball from the known urn (K). The monetary valuation of the known urn (K) to subject i is denoted by v_{iK} and the monetary valuation of the unknown urn (U) to subject i is denoted by v_{iU}. In addition, assume (P). Then*

$$-\ln\left(-\ln m_i^*(p)\right) = -\ln\left(-\ln p\right) + \frac{1}{\alpha_{iK}} \ln\left(\frac{\ln v - \ln v_{iK}}{\ln v - \ln v_{iU}}\right). \tag{8}$$

Proof of Proposition 6:

Firstly, for the known urn (K), we have

$$(v_{iK})^{\sigma_i} = w_{iK}(p)(v)^{\sigma_i}. \tag{9}$$

Solve (9) for σ_i, to get

$$\sigma_i = \frac{-\ln w_{iK}(p)}{\ln v - \ln v_{iK}}. \tag{10}$$

By definition of matching probability, we have

$$(v_{iU})^{\sigma_i} = w_{iK}(m_i^*(p))(v)^{\sigma_i}. \tag{11}$$

Solve (11) for $w_{iK}(m_i^*(p))$, to get

$$w_{iK}(m_i^*(p)) = \left(\frac{v_{iU}}{v}\right)^{\sigma_i}. \tag{12}$$

Substitute from Equation (10) into Equation (12) to get

$$w_{iK}(m_i^*(p)) = \left(\frac{v_{iU}}{v}\right)^{\frac{-\ln w_{iK}(p)}{\ln v - \ln v_{iK}}}. \tag{13}$$

Taking logs of (13), and some rearranging, we get

$$\ln w_{iK}\left(m_i^*\left(p\right)\right) = \frac{\ln w_{iK}\left(p\right)}{\ln v - \ln v_{iK}}(\ln v - \ln v_{iU}). \tag{14}$$

From the Prelec function (6), we get

$$\ln w_{iK}\left(p\right) = -\beta_{iK}\left(-\ln p\right)^{\alpha_{iK}}, \tag{15}$$
$$\ln w_{iK}\left(m_i^*\left(p\right)\right) = -\beta_{iK}\left(-\ln m_i^*\left(p\right)\right)^{\alpha_{iK}}. \tag{16}$$

From (14)-(16), and some rearranging, we get (8). ■

8 Elements of Quantum Probability Theory

8.1 Preamble

Expected utility theory (EU) is probably still the most popular decision theory in economics. On the other hand, Luce and Raiffa (1957, p35) stated that "reported preferences almost never satisfy the axioms" and, on p37, stated that the evidence against EU is "now bolstered by a staggering amount of empirical data". Since then, the evidence against EU has multiplied several fold. Hence, the hunt is on for a decision theory more in accord with the evidence.

However, the (non-quantum) alternatives that have been proposed are obtained from expected utility theory by relaxing one, or more, of its assumptions. For example, Segal (1990) proposed dropping the *reduction axiom.* The *smooth ambiguity model* relaxes the assumption of linearity of the utility function in probabilities by introducing the function φ, which has to be determined from the data (section 5). *Rank dependent expected utility theory* (section 6) and *source dependent probability theory* (section 7) relax the assumption of linearity of the utility function in probabilities using probability weighting functions. But neither of these theories determine the probability weighting functions they use, which have to be chosen to fit the empirical evidence. Unfortunately, such weakening of EU produces incomplete theories, they introduce greater flexibility at the cost of reducing predictive power.

Quantum decision theory (QDT) originated with Aerts and Aerts (1994) who noticed similarities between paradoxes of human behavior (e.g., those empirical observations that contradict the predictions of expected utility theory) and paradoxes of quantum mechanics (i.e., those empirical observations that contradict the predictions of classical mechanics).

The paradoxes of quantum mechanics led von Neumann (1955, original German 1932) to devise a new mathematical structure in which quantum mechanics can be given a consistent formulation, *Hilbert Space* and *quantum probability*. Events are vector subspaces of Hilbert space, and quantum probability is an additive (though not distributive) measure on these.

In quantum decision theory (QDT), unlike all other decision theories, events are not distributive, and this is the main difference between the two. Thus, in QDT the event "X *and* (Y *or* Z)" need not be equivalent to the event "(X *and* Y) *or* (X *and* Z)". On the other hand, in all other decision theories, these two events are equivalent. This non-distributive nature of QDT is the key to its success in explaining paradoxes of behavior that other decision theories find difficult to explain. For example, *order effects*, the *Linda paradox*, the *disjunction fallacy*, the *conjunction fallacy* and the failure of the *sure-thing principle*. See Busemeyer & Bruza (2012); in particular, their sections 1.2, 4.1-4.3, 5.2 and 10.2.3. As a result of the non-distributive nature of QDT, the *law of total probability* does not generally hold. Instead, we use the *Feynman rules* and the *law of reciprocity*. See Busemeyer & Bruza (2012), pp. 5, 13, 39.

Quantum probability theory is complete in the following sense. Once probabilities are assigned to the elementary events (by, say, the heuristic of insufficient reason, or by the observation of relative frequencies) quantum theory then uniquely determines the probabilities of all events.

In fact, more can be said. Just as the Kolmogorov probability measure is the unique additive measure on subsets of a set (Billingsley, 1995, Theorem 3.1, p.36), so the quantum probability measure is the unique additive measure on subspaces of a vector space (Gleason, 1957).

Thus, we have a choice of two probability measures: Kolmogorov or quantum. The latter is more general in the sense that any phenomenon that can be explained by the former can also be explained by the latter, but the reverse is not true. So, maybe, the need to use probability weighting functions is just a symptom that we should be using quantum probability theory rather than Kolmogorov probability theory. One can take either of the following two positions:

1. Rational beings should follow Kolmogorov probability theory. The more general quantum probability would then give a systematic account of

irrational human behavior.

2. Kolmogorov probability theory is simply inadequate to describe human behavior (just as it is not adequate to explain behavior of material objects; we do not say that material objects are irrational because they disobey classical probability theory). We need a more general probability theory, such as quantum probability or Choquet capacity (Choquet, 1953-1954).

We prefer the second position.

A number of quantum models of the Ellsberg paradox have been developed.

Busemeyer and Bruza (2012, section 9.1.2) applied projective expected utility theory (Subsection 2.6) to explain the Ellsberg paradox. Their model has a free parameter, a. If $a > 0$ we get ambiguity aversion, if $a = 0$, we get ambiguity neutrality, and if $a < 0$ we get ambiguity seeking. However, it cannot explain the simultaneous occurrence in the same subject of ambiguity seeking (for low probabilities), ambiguity neutrality and ambiguity aversion (for medium and high probabilities), because a cannot be simultaneously negative, zero and positive.

Aerts et al. (2014) formulate and study a quantum decision theory (QDT) model of the Ellsberg paradox. They consider one of the standard versions of the Ellsberg paradox. They consider a single urn with 30 red balls and 60 balls that are either yellow or black, the latter in unknown proportions. They use the heuristic of insufficient reason (subsection 2.7) for the known distribution (red) but not for the unknown distribution (yellow or black). They prove that in their mode, the Ellsberg paradox reemerges if they use the heuristic of insufficient reason for the unknown distribution. They, therefore, abandon this heuristic. They choose the ratio of yellow to black to fit the evidence from their subjects. However, other theories can explain the Ellsberg paradox if we abandon *insufficient reason*. Thus, the explanation of Aerts et al. (2014) is not specifically quantum, although it is expressed in that language.

Khrennikov and Haven (2009) provide a general quantum-like framework for situations where Savage's sure-thing principle (Savage, 1954) is violated; one of these being the Ellsberg paradox. Their *quantum-like* or *contextual probabilistic* (*Växjö*) *model* is much more general than either the classical Kolmogorov model or the standard quantum model (see Khrennikov, 2010, and Haven and Khrennikov, 2013). By contrast, our approach is located strictly within standard quantum theory. Furthermore, in their formulation, the Ellsberg paradox reemerges if one adopts (as we do) the heuristic of insufficient reason.[8] On the other hand, although abandoning the heuristic of insufficient reason gives models extra flexibility, it also reduces their predictive power.

Thus, some of the quantum models that have been proposed do explain the Ellsberg paradox, but at the cost of introducing a considerable degree of flexibility. However, when non-quantum models are granted the same degree of flexibility, they too can explain the Ellsberg paradox. Busemeyer & Bruza (2012) section 9.1.2 conclude "In short, quantum models of decision making can accommodate the Allais and Ellsberg paradoxes. But so can non-additive weighted utility models, and so these paradoxes do not point to any unique advantage for the quantum model". By contrast, when we replace quantum probability by Kolmogorov probability in our model, then the Ellsberg paradox reemerges. This is because our underlying decision theory, projective expected utility (La Mura 2009) reduced to expected utility theory when quantum probabilities are replaced with Kolmogorov probabilities, leading to the emergence of the Ellsberg paradox (recall Propositions 1 and 2). Hence, we make essential use of quantum probability theory.

By contrast, our model (section 9, below) provides a parameter-free derivation of quantum probabilities and can explain the simultaneous occurrence in the same subject of ambiguity seeking (low probabilities), ambiguity neutrality and ambiguity aversion (medium and high probabilities). Its predictions are in good agreement with the empirical evidence in Dimmock et al. (2015). Thus, our application of projective expected utility theory has a clear advantage over all other decision theories. Furthermore, projective expected utility can be extended to include reference dependence and loss aversion, to yield *projective prospect theory*, where decision weights are replaced with quantum probabilities. This would have a clear advantage over all the standard (non-quantum) versions of prospect theory.

For papers examining the limits of standard quantum theory when applied to cognitive psychology, see Khrennikov et al. (2014), Basieva & Khrennikov (2015), and Asano et al. (2016).

[8] Khrennikov and Haven (2009), subsection 4.6, p386.

8.2 Vectors

For our purposes (as we shall show), it is sufficient to use a finite dimensional real vector space $\mathbb{R}^n$ (in fact, with $n = 4$). A vector, $\mathbf{x} \in \mathbb{R}^n$, is represented by an $n \times 1$ matrix (n rows, one column). Its *transpose*, $\mathbf{x}\dagger$, is then the $1 \times n$ matrix (one row, n columns) of the same elements but written as a row.[9]. The *zero vector*, $\mathbf{0}$, is the vector all of whose components are zero. Let $r \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ with components x_i and y_i, respectively. Then $r\mathbf{x}$ is the vector whose components are rx_i and $\mathbf{x} + \mathbf{y}$ is the vector whose components are $x_i + y_i$. $\mathbf{y} \in \mathbb{R}^n$ is a *linear combination* of $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m \in \mathbb{R}^n$ if $\mathbf{y} = \sum_{i=1}^m r_i \mathbf{x}_i$ for some real numbers $r_1, r_2, ..., r_m$. The *inner product* of $\mathbf{x}$ and $\mathbf{y}$ is $\mathbf{x}\dagger\mathbf{y} = \sum_{i=1}^n x_i y_i$, where x_i, y_i are the components of $\mathbf{x}$ and $\mathbf{y}$, respectively.[10] If $\mathbf{x}\dagger\mathbf{y} = 0$, then $\mathbf{x}$ is said to be *orthogonal* to $\mathbf{y}$ and we write $\mathbf{x} \perp \mathbf{y}$. Note that $\mathbf{x} \perp \mathbf{y}$ if, and only if, $\mathbf{y} \perp \mathbf{x}$. The *norm*, or *length*, of $\mathbf{x}$ is $\|\mathbf{x}\| = \sqrt{\mathbf{x}\dagger\mathbf{x}}$. $\mathbf{x}$ is *normalized* if $\|\mathbf{x}\| = 1$.[11] $X \subset \mathbb{R}^n$ is a *vector subspace* (of $\mathbb{R}^n$) if it satisfies: $X \neq \emptyset$, $\mathbf{x}, \mathbf{y} \in X \Rightarrow \mathbf{x} + \mathbf{y} \in X$ and $r \in \mathbb{R}, \mathbf{x} \in X \Rightarrow r\mathbf{x} \in X$. Let $\mathcal{L}$ be the set of all vector subspaces of $\mathbb{R}^n$. Then $\{\mathbf{0}\}, \mathbb{R}^n \in \mathcal{L}$. Let $X, Y \in \mathcal{L}$. Then $X \cap Y \in \mathcal{L}$ and $X + Y = \{\mathbf{x} + \mathbf{y} : \mathbf{x} \in X, \mathbf{y} \in Y\} \in \mathcal{L}$. If $X_1, X_2, ..., X_m \in \mathcal{L}$, then $\sum_{i=1}^m X_i = \{\sum_{i=1}^m \mathbf{x}_i : \mathbf{x}_i \in X_i\} \in \mathcal{L}$. The *orthogonal complement* of $X \in \mathcal{L}$ is $X^\perp = \{\mathbf{y} \in \mathbb{R}^n : \mathbf{y} \perp \mathbf{x} \text{ for each } \mathbf{x} \in X\}$. We have $X^\perp \in \mathcal{L}$, $\left(X^\perp\right)^\perp = X$, $X \cap X^\perp = \{\mathbf{0}\}$, $X + X^\perp = \mathbb{R}^n$. Let $\mathbf{z} \in \mathbb{R}^n$ and $X \in \mathcal{L}$, then there is a unique $\mathbf{x} \in X$ such that $\|\mathbf{z} - \mathbf{x}\| \leq \|\mathbf{z} - \mathbf{y}\|$ for all $\mathbf{y} \in X$. $\mathbf{x}$ is called *the orthogonal projection of* $\mathbf{z}$ *onto* X. Let $\delta_{ii} = 1$ but $\delta_{ij} = 0$ for $i \neq j$. $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_m$ form an *orthonormal basis* for $X \in \mathcal{L}$ if $\mathbf{s}_i\dagger\mathbf{s}_j = \delta_{ij}$ and if any vector $\mathbf{x} \in X$ can be represented as a linear combination of the basis vectors: $\mathbf{x} = \sum_{i=1}^m x_i \mathbf{s}_i$, where the numbers $x_1, x_2, ..., x_m$ are uniquely determined by $\mathbf{x}$ and $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_m$. The choice of an orthonormal basis for a vector space is arbitrary. However, the inner product of two vectors is independent of the orthonormal basis chosen. We shall refer to a normalized vector, $\mathbf{s} \in \mathbb{R}^n$, as a *state vector*. In particular, if $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_n$ form an orthonormal basis for $\mathbb{R}^n$, then we shall refer to these as *eigenstates*. Note that if $\mathbf{s} = \sum_{i=1}^n s_i \mathbf{s}_i$,

[9] More generally, in $\mathbb{C}^n$, $\mathbf{x}\dagger$ is the *adjoint*, of $\mathbf{x}$. For example, in $\mathbb{C}^2$, if $\mathbf{x} = \begin{bmatrix} r_1 e^{i\theta_1} \\ r_2 e^{i\theta_2} \end{bmatrix}$, where r_1, θ_1, r_2, θ_2 are real and $i = \sqrt{-1}$, then $\mathbf{x}\dagger = \begin{bmatrix} r_1 e^{-i\theta_1} & r_2 e^{-i\theta_2} \end{bmatrix}$.

[10] More generally, in $\mathbb{C}^n$, $\mathbf{x}\dagger\mathbf{y} = \sum_{i=1}^n x_i^* y_i$, where, if $x = re^{i\theta}$, $r, \theta \in \mathbb{R}$, then $x^* = re^{-i\theta}$.

[11] In *Dirac notation*, $\mathbf{x} = |x\rangle$, $\mathbf{x}\dagger = \langle x|$, $\mathbf{x}\dagger\mathbf{y} = \langle x|y\rangle$, $\|\mathbf{x}\| = \sqrt{\mathbf{x}\dagger\mathbf{x}} = \sqrt{\langle x|x\rangle}$. Physicists are, of course, very familiar with the Dirac notation. On the other hand, most economists are not. Therefore, we use standard algebraic notation, with whish they are familiar.

then $\mathbf{s}$ is a state vector if, and only if, $\|\mathbf{s}\| = 1$, equivalently, if, and only if, $\mathbf{s}\dagger\mathbf{s} = \sum_{i=1}^{n} s_i s_i = 1$. Let $X \in \mathcal{L}$. Let $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_m$ form an orthonormal basis for X. Extend $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_m$ to an orthonormal basis, $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_m, ..., \mathbf{s}_n$, for $\mathbb{R}^n$ (this can always be done). Then $\mathbf{s}_{m+1}, ..., \mathbf{s}_n$ form an orthonormal basis for the the orthogonal complement, $X^{\perp}$, of X. Let $\mathbf{z} = \sum_{i=1}^{n} z_i \mathbf{s}_i \in \mathbb{R}^n$. Then $\sum_{i=1}^{m} z_i \mathbf{s}_i$ is the orthogonal projection of $\mathbf{z}$ onto X and $\sum_{i=m1}^{n} z_i \mathbf{s}_i$ is the orthogonal projection of $\mathbf{z}$ onto $X^{\perp}$.

We will represent the state of the ambiguous Ellsberg urn (A) by a normalized vector in $\mathbb{R}^4$. We have checked that adopting the complex vector space, $\mathbb{C}^4$, changes none of our results.

8.3 State of a system, events and quantum probability measures

The state of a system (physical, biological or social) is represented by a normalized vector, $\mathbf{s} \in \mathbb{R}^n$, i.e., $\|\mathbf{s}\| = 1$. The set of events is the set, $\mathcal{L}$, of vector subspaces of $\mathbb{R}^n$. $\{\mathbf{0}\}$ is the impossible event and $\mathbb{R}^n$ is the certain event. $X^{\perp} \in \mathcal{L}$ is the complement of the event $X \in \mathcal{L}$. If $X, Y \in \mathcal{L}$ then $X \cap Y$ is the conjunction of the events X and Y; $X + Y$ is the event where either X occurs or Y occurs or both (if $X, Y \in \mathcal{L}$ then, in general, $X \cup Y \notin \mathcal{L}$). Recall that in a σ-algebra of subset of a set, the distributive law: $X \cap (YUZ) = (X \cap Y) \cup (X \cap Z)$, and its dual[12], hold. However, its analogue for $\mathcal{L}$: $X \cap (Y + Z) = (X \cap Y) + (X \cap Z)$, and its dual[13], fail to hold in general. Consequently, the law of total probability also fails to hold in general. The failure of the distributive laws to hold in $\mathcal{L}$ has profound consequences. This non-distributive nature of $\mathcal{L}$ is the key to explaining many paradoxes of human behavior. $F : \mathcal{L} \to [0, 1]$ is *additive* if $F\left(\sum_{i=1}^{m} X_i\right) = \sum_{i=1}^{m} F(X_i)$, where $X_i \in \mathcal{L}$ and $X_i \cap X_j = \{\mathbf{0}\}$ for $i \neq j$. A *quantum probability measure* is an additive measure, $P : \mathcal{L} \to [0, 1]$, such that $P(\{\mathbf{0}\}) = 0$, $P(\mathbb{R}^n) = 1$. If a number can be interpreted as either a classical probability or a quantum probability, then we shall simply refer to it as a probability. Otherwise, we shall refer to it as either a classical probability or a quantum probability, whichever is the case.

[12] $X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$

[13] $X + (Y \cap Z) = (X + Y) \cap (X + Z)$

8.4 Random variables and expected values

Let $\mathcal{L}$ be the set of all vector subspaces of $\mathbb{R}^n$. A *random quantum variable* is a mapping, $f : \mathbb{R}^n \to \mathbb{R}$ satisfying: $\{\varphi \in \mathbb{R}^n : f(\varphi) \leq r\} \in \mathcal{L}$ for each $r \in \mathbb{R}$.

A random quantum variable, f, is *non-negative* if $f(\varphi) \geq 0$ for each $\varphi \in \mathbb{R}^n$. For two random quantum variables, f, g, we write $f \leq g$ if $f(\varphi) \leq g(\varphi)$ for each $\varphi \in \mathbb{R}^n$. A random quantum variable, f, is *simple* if its range is finite. For any random quantum variable, f, and any $\varphi \in \mathbb{R}^n$, let $f^+(\varphi) = \max\{0, f(\varphi)\}$ and $f^-(\varphi) = -\min\{0, f(\varphi)\}$. Then, clearly, f^+ and f^- are both non-negative random quantum variables and $f(\varphi) = f^+(\varphi) - f^-(\varphi)$, for each $\varphi \in \mathbb{R}^n$. We write this as $f = f^+ - f^-$.

Let f be a simple random quantum variable with range $\{f_1, f_2, ..., f_n\}$. Let $X_i = \{\varphi \in \mathbb{R}^n : f(\varphi) = f_i\}$. Then $X_i \in \mathcal{L}$, $X_i \cap X_j = \{\mathbf{0}\}$ for $i \neq j$ and $\sum_{i=1}^{n} X_i = \mathbb{R}^n$. Then the *expected value* of the simple random quantum variable, f, is $E(f) = \sum_{i=1}^{n} f_i P(X_i)$. The *expected value* of the non-negative random quantum variable, g, is
$E(g) = \sup\{E(f) : f \leq g \text{ is a simple random quantum variable}\}$. Note that $E(g)$ may be infinite. If $f = f^+ - f^-$ is an arbitrary random quantum variable such that not both $E(f^+)$ and $E(f^-)$ are infinite, then the expected value of f is $E(f) = E(f^+) - E(f^-)$. Note that $E(f)$ can be $-\infty$, finite or ∞. However, if $E(f^+)$ and $E(f^-)$ are both infinite then $E(f)$ is undefined (because $\infty - \infty$ is undefined).

8.5 Transition amplitudes and probabilities

Suppose $\varphi, \chi \in \mathbb{R}^n$ are two states (thus, they are normalized: $\|\varphi\| = \|\chi\| = 1$). $\varphi \to \chi$ symbolizes the transition from φ to χ. Then, by definition, the *amplitude* of $\varphi \to \chi$ is given by $A(\varphi \to \chi) = \varphi\dagger\chi$. Its quantum probability is $P(\varphi \to \chi) = (\varphi\dagger\chi)^2$.[14]

Consider the state $\varphi \in \mathbb{R}^n$ ($\|\varphi\| = 1$). The occurrence of the event $X \in \mathcal{L}$ causes a transition, $\varphi \to \psi$. The new state, ψ ($\|\psi\| = 1$), can be found as follows. Let π be the orthogonal projection of φ onto X (recall subsection 8.2). Suppose that $\pi \neq \mathbf{0}$ (if $\pi = \mathbf{0}$, then π and X are incompatible, that is, if X occurs then the transition $\varphi \to \psi$ is impossible). Then $\psi = \frac{\pi}{\|\pi\|}$ is the new *state conditional on* X.

[14] In $\mathbb{C}^n$, $P(\varphi \to \chi) = (\varphi\dagger\chi)(\varphi\dagger\chi)^*$. However, as we are working in $\mathbb{R}^n$, $(\varphi\dagger\chi) = (\varphi\dagger\chi)^*$.

8.6 Born's rule

We can now give the empirical interpretation of the state vector. Consider a physical, biological or social system. On measuring a certain observable pertaining to the system, this observable can take the value $v_i \in \mathbb{R}$ with probability $p_i \geq 0$, $\sum_{i=1}^{n} p_i = 1$. To model this situation, let $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_n$ form an orthonormal bases for $\mathbb{R}^n$. Take $\mathbf{s}_i$ to be the state (eigenstate) where the observable takes the value (eigenvalue) v_i for sure. Consider the general state $\mathbf{s} = \sum_{i=1}^{n} s_i \mathbf{s}_i$. If the act of measurement gives the value v_i for the observable, then this implies that the act of measurement has caused the transition $\mathbf{s} \to \mathbf{s}_i$. The probability of the transition $\mathbf{s} \to \mathbf{s}_i$ is $P(\mathbf{s} \to \mathbf{s}_i) = (\mathbf{s} \dagger \mathbf{s}_i)^2 = s_i^2 = p_i$. Thus, in the representation of the state of the system by $\mathbf{s} = \sum_{i=1}^{n} s_i \mathbf{s}_i$, s_i^2 is the probability of obtaining the value v_i on measurement[15]

8.7 Feynman's first rule (single path)

See Busemeyer and Bruza (2012), section 2.2, for the Feynman rules.

Let φ, χ, ψ be three states. $\varphi \to \chi \to \psi$ symbolizes the transition from φ to χ followed by the transition from χ to ψ. The amplitude of $\varphi \to \chi \to \psi$ is then the *product*, $A(\varphi \to \chi \to \psi) = A(\varphi \to \chi) A(\chi \to \psi) = (\varphi \dagger \chi)(\chi \dagger \psi)$, of the amplitudes of $\varphi \to \chi$ and $\chi \to \psi$. The quantum probability of the transition, $\varphi \to \chi \to \psi$, is then $P(\varphi \to \chi \to \psi) = (A(\varphi \to \chi \to \psi))^2 = ((\varphi \dagger \chi)(\chi \dagger \psi))^2 = (\varphi \dagger \chi)^2 (\chi \dagger \psi)^2$, i.e., the product of the respective probabilities. This can be extended to any number of multiple transitions along a single path.

8.8 Feynman's second rule (multiple indistinguishable paths)

Suppose that the transition from φ to ψ can follow any of two paths: $\varphi \to \chi_1 \to \psi$ or $\varphi \to \chi_2 \to \psi$. Furthermore, and this is crucial, assume that which path was followed is *not* observable. First, we calculate the amplitude of $\varphi \to \chi_1 \to \psi$, using Feynman's first rule. We also calculate the amplitude of $\varphi \to \chi_2 \to \psi$, using, again, Feynman's first rule. To find the amplitude of $\varphi \to \psi$ (via χ_1 or χ_2) we *add the two amplitudes*. The amplitude of $\varphi \to \psi$ is then $(\varphi \dagger \chi_1)(\chi_1 \dagger \psi) + (\varphi \dagger \chi_2)(\chi_2 \dagger \psi)$. Finally, the probability of

[15] More generally, if we use $\mathbb{C}^n$, then $s_i^* s_i$, is the probability of obtaining the value v_i on measurement.

the transition $\varphi \to \psi$ (via χ_1 or χ_2) is $((\varphi\dagger\chi_1)(\boldsymbol{\chi_1}\dagger\psi) + (\varphi\dagger\chi_2)(\boldsymbol{\chi_2}\dagger\psi))^2 = (\varphi\dagger\chi_1)^2(\boldsymbol{\chi_1}\dagger\psi)^2 + (\varphi\dagger\chi_2)^2(\boldsymbol{\chi_2}\dagger\psi)^2 + 2((\varphi\dagger\chi_1)(\boldsymbol{\chi_1}\dagger\psi)(\varphi\dagger\chi_2)(\boldsymbol{\chi_2}\dagger\psi))$.

8.9 Feynman's third rule (multiple distinguishable paths)

Suppose that the transition from φ to ψ can follow any of two paths: $\varphi \to \chi_1 \to \psi$ or $\varphi \to \chi_2 \to \psi$. Furthermore, and this is crucial, assume that which path was followed *is* observable (although it might not actually be observed). First, we calculate the quantum probability of $\varphi \to \chi_1 \to \psi$, using Feynman's first rule. We also calculate the quantum probability of $\varphi \to \chi_2 \to \psi$, using, again, Feynman's first rule. To find the total quantum probability of $\varphi \to \psi$ (via χ_1 or χ_2) we *add the two probabilities*. The quantum probability of $\varphi \to \psi$ is then $(\varphi\dagger\chi_1)^2(\boldsymbol{\chi_1}\dagger\psi)^2 + (\varphi\dagger\chi_2)^2(\boldsymbol{\chi_2}\dagger\psi)^2$.

Comparing the last expression with its analogue for Feynman's second rule, we see the absence here of the term $2((\varphi\dagger\chi_1)(\boldsymbol{\chi_1}\dagger\psi)(\varphi\dagger\chi_2)(\boldsymbol{\chi_2}\dagger\psi))$. This is called the *interference term*. Its presence or absence has profound implications in both quantum physics and quantum decision theory.

The Feynman rules play a role in quantum probability theory analogous to the rule played by Bayes' law and the law of total probability in classical theory.

8.10 An illustration

We give a simple example where it is clear which Feynman rule should be used. Consider an Ellsberg urn containing two balls. One ball is marked 1 and the other ball is marked 2. If a ball is drawn at random then, in line with the heuristic of insufficient reason, we assign probability $\frac{1}{2}$ to ball 1 being drawn and probability $\frac{1}{2}$ to ball 2 being drawn. Call this initial state $\mathbf{s}$. Let the state where ball 1 is drawn be $\mathbf{s}_1$ and let $\mathbf{s}_2$ be the state if ball 2 is drawn. Now, suppose a ball is drawn but *returned* to the urn. This should not change the initial state of the urn. Both classical reasoning and quantum reasoning should give a probability of 1 to the transition $\mathbf{s} \to \mathbf{s}$.

8.10.1 Classical treatment

Consider the transition $\mathbf{s} \to \mathbf{s}$. This can occur via one of the two paths: $\mathbf{s} \to \mathbf{s}_1 \to \mathbf{s}$ or $\mathbf{s} \to \mathbf{s}_2 \to \mathbf{s}$: either ball 1 is drawn then returned to the urn or ball 2 is drawn then returned to the urn. The classical treatment gives a probability $\frac{1}{2}$ to the transition $\mathbf{s} \to \mathbf{s}_1$. Since returning ball 1 restores the original state of the urn, the classical probability of the transition $\mathbf{s}_1 \to \mathbf{s}$ is 1. Hence, the classical probability of the transition $\mathbf{s} \to \mathbf{s}_1 \to \mathbf{s}$ is $\frac{1}{2} \times 1 = \frac{1}{2}$. Similarly, the classical probability of the transition $\mathbf{s} \to \mathbf{s}_2 \to \mathbf{s}$ is also $\frac{1}{2}$. Hence, the classical probability of the transition $\mathbf{s} \to \mathbf{s}$ via either paths $\mathbf{s} \to \mathbf{s}_1 \to \mathbf{s}$ or $\mathbf{s} \to \mathbf{s}_2 \to \mathbf{s}$ is $\frac{1}{2} + \frac{1}{2} = 1$.

8.10.2 Quantum treatment

We use $\mathbb{R}^2$. Let $\mathbf{s}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ be the state if ball 1 is drawn and let $\mathbf{s}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ be the state if ball 2 is drawn. Take the initial state of the urn be $\mathbf{s} = \sqrt{\frac{1}{2}}\mathbf{s}_1 + \sqrt{\frac{1}{2}}\mathbf{s}_2 = \begin{bmatrix} \sqrt{\frac{1}{2}} \\ \sqrt{\frac{1}{2}} \end{bmatrix}$. Let us check to see if this is a reasonable assignment. $\mathbf{s}_1$ and $\mathbf{s}_2$ form an orthonormal basis for $\mathbb{R}^2$. $\|\mathbf{s}\| = \sqrt{\mathbf{s}\dagger\mathbf{s}} = 1$. Hence, $\mathbf{s}$ is a state vector. The amplitude of the transition $\mathbf{s} \to \mathbf{s}_1$ is $\mathbf{s}\dagger\mathbf{s}_1 = \sqrt{\frac{1}{2}}$. The amplitude of the transition $\mathbf{s}_1 \to \mathbf{s}$ is $\mathbf{s}_1\dagger\mathbf{s} = \sqrt{\frac{1}{2}}$. Hence, by Feynman's first rule (single path), the amplitude of the transition $\mathbf{s} \to \mathbf{s}_1 \to \mathbf{s}$ is $A(\mathbf{s} \to \mathbf{s}_1 \to \mathbf{s}) = A(\mathbf{s} \to \mathbf{s}_1) A(\mathbf{s}_1 \to \mathbf{s}) = \frac{1}{2}$, in agreement with our intuitive reasoning. Similarly, the amplitude of the transition $\mathbf{s} \to \mathbf{s}_2 \to \mathbf{s}$ is $A(\mathbf{s} \to \mathbf{s}_2 \to \mathbf{s}) = \frac{1}{2}$. We now compare the results from applying Feynman's second rule with the results from applying Feynman's third rule. Since $P(\mathbf{s} \to \mathbf{s}) = (A(\mathbf{s} \to \mathbf{s}))^2 = (\mathbf{s}\dagger\mathbf{s})^2 = (1)^2 = 1$, the correct rule is the one that gives this result.

Feynman's second rule (multiple indistinguishable paths) Here we *add* the amplitudes of the transitions $\mathbf{s} \to \mathbf{s}_1 \to \mathbf{s}$ and $\mathbf{s} \to \mathbf{s}_2 \to \mathbf{s}$ to get the amplitude of the transition $\mathbf{s} \to \mathbf{s}$: $A(\mathbf{s} \to \mathbf{s}) = A(\mathbf{s} \to \mathbf{s}_1 \to \mathbf{s}) + A(\mathbf{s} \to \mathbf{s}_2 \to \mathbf{s}) = \frac{1}{2} + \frac{1}{2} = 1$. Hence, the quantum probability of the transition $\mathbf{s} \to \mathbf{s}$, through all paths, is $P(\mathbf{s} \to \mathbf{s}) = (A(\mathbf{s} \to \mathbf{s}))^2 = (1)^2 = 1$, in agreement with our intuitive analysis.

Feynman's third rule (multiple distinguishable paths) Here we calculate the quantum probabilities of the transitions $\mathbf{s} \to \mathbf{s}_1 \to \mathbf{s}$ and $\mathbf{s} \to \mathbf{s}_2 \to \mathbf{s}$. This gives $P(\mathbf{s} \to \mathbf{s}_1 \to \mathbf{s}) = (A(\mathbf{s} \to \mathbf{s}_1 \to \mathbf{s}))^2 = \left(\frac{1}{2}\right)^2 = \frac{1}{4}$ and

$P(\mathbf{s} \to \mathbf{s}_2 \to \mathbf{s}) = (A(\mathbf{s} \to \mathbf{s}_2 \to \mathbf{s}))^2 = \left(\frac{1}{2}\right)^2 = \frac{1}{4}$. Then we add these quantum probabilities to get $P(\mathbf{s} \to \mathbf{s}) = P(\mathbf{s} \to \mathbf{s}_1 \to \mathbf{s}) + P(\mathbf{s} \to \mathbf{s}_2 \to \mathbf{s}) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$. This is a contradiction, since $P(\mathbf{s} \to \mathbf{s}) = (A(\mathbf{s} \to \mathbf{s}))^2 = (\mathbf{s}\dagger\mathbf{s})^2 = (1)^2 = 1$. Hence Feynman's third rule is not the correct rule to apply. The transition from $\mathbf{s}$ back to $\mathbf{s}$ via $\mathbf{s}_1$ should be treated as indistinguishable from the transition from $\mathbf{s}$ back to $\mathbf{s}$ via $\mathbf{s}_2$.

9 A simple quantum model of the Ellsberg paradox

Urn U (unknown composition) contains two balls both labeled 1 if it is in state $\mathbf{s}_1$ (recall subsection 2.5). It contains one ball labeled 1 and the other labeled 2 if it is either in state $\mathbf{s}_2$ or in state $\mathbf{s}_3$. In state $\mathbf{s}_4$ both balls are labeled 2. We represent these states in $\mathbb{R}^4$ by the orthonormal basis:

$$\mathbf{s}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{s}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{s}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{s}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Let $\mathbf{s}$ give the initial state of urn U. Then *Born's rule* leads to:

$$\mathbf{s} = p\mathbf{s}_1 + \sqrt{p(1-p)}\mathbf{s}_2 + \sqrt{(1-p)p}\mathbf{s}_3 + (1-p)\mathbf{s}_4, \tag{17}$$

where there is a probability p^2 that ball 1 is drawn in each round (state $\mathbf{s}_1$), a probability $p(1-p)$ that ball 1 is drawn in round 1 then ball 2 is drawn in round 2 (state $\mathbf{s}_2$), a probability $(1-p)p$ that ball 2 is drawn in round 1 then ball 1 is drawn in round 2 (state $\mathbf{s}_3$) and, finally, a probability $(1-p)^2$ that ball 2 is drawn in each round (state $\mathbf{s}_4$).

In quantum mechanics, the construction described by (17) is called a *state preparation* and p in (17) is both a classical and a quantum probability. They can be given by the heuristic of insufficient reason or by relative frequencies. In particular, $p^2 + p(1-p) + (1-p)p + (1-p)^2 = 1$. By contrast, the action of extracting a ball from urn K or urn U is known as a *measurement*. The probabilities there are quantum probabilities (which may or may not be numerically equal to classical probabilities, may or may not add up to 1, but are never negative).

Let the event that ball 1 is drawn from urn U be denoted by $\mathbf{t}$. We now calculate the probability of event $\mathbf{t}$.

Proposition 7 *(al-Nowaihi & Dhami, 2017, section 5): Assume (Q) and (B). If the probability of drawing a winning ball (a ball labeled* 1*) from the known urn K is p, then the quantum probability of drawing a winning ball (a ball labeled* 1*) from the unknown urn U is*

$$Q(p) = \frac{5p^3 - 8p^2 + 4p}{2 - p}. \tag{18}$$

Proof of Proposition 7: The role played by the *law of reciprocity* was only implicit in al-Nowaihi & Dhami (2017). Here we make it explicit. We take the opportunity to further clarify other points.

In general, the law of total probability, recall (2) above, is not valid in quantum probability theory. See Busemeyer & Bruza (2012), chapter 1, pp. 5. Instead, we use the *Feynman's rules* (see Busemeyer & Bruza (2012), chapter 1, pp. 13) and the *law of reciprocity* (see Busemeyer & Bruza (2012), chapter 2, pp. 39). In our case, working in the Hilbert space $\mathbb{C}^4$ gives the same results as working in $\mathbb{R}^4$, as can be verified by direct calculation. Hence, for simplicity, we shall work in the Hilbert space $\mathbb{R}^4$. Recall that the state of a quantum system is given by normalized vector, $\mathbf{s}$, in Hilbert space, i.e., $\mathbf{s}\dagger\mathbf{s} = (\mathbf{s}\dagger)\,\mathbf{s} = 1$, where $\mathbf{s}\dagger$ is the *conjugate transpose* of $\mathbf{s}$ (in our case, simply the transpose of $\mathbf{s}$, since we are working in $\mathbb{R}^4$). We give the proof in several stages.

Reciprocity

Let $\mathbf{t}$ be the state where ball 1 is drawn from U. We wish to calculate the probability, $P(\mathbf{s} \to \mathbf{t})$, of the transition $\mathbf{s} \to \mathbf{t}$. By the quantum law of reciprocity, $P(\mathbf{s} \to \mathbf{t}) = P(\mathbf{t} \to \mathbf{s})$, both being equal to $(\mathbf{s}\dagger\mathbf{t})^2$. Recall we are working in a real Hilbert space. For a complex Hilbert space, we would have $P(\mathbf{s} \to \mathbf{t}) = P(\mathbf{t} \to \mathbf{s}) = (\mathbf{s}\dagger\mathbf{t})(\mathbf{s}\dagger\mathbf{t})^*$, where $(\mathbf{s}\dagger\mathbf{t})^*$ is the complex conjugate of $\mathbf{s}\dagger\mathbf{t}$. But $P(\mathbf{t} \to \mathbf{s})$ is the probability of the state of U conditional on drawing ball 1 from U. Let $\mathbf{w}$ be this state. To find $\mathbf{w}$, we first project $\mathbf{s}$ onto the subspace spanned by $\{\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3\}$, then normalize. This gives

$$\mathbf{w} = \sqrt{\frac{p}{2-p}}\mathbf{s}_1 + \sqrt{\frac{1-p}{2-p}}\mathbf{s}_2 + \sqrt{\frac{1-p}{2-p}}\mathbf{s}_3. \tag{19}$$

Feynman's rules

To arrive at the state, $\mathbf{w}$, the state of urn U conditional on ball 1 being drawn, we must follow one of the three paths:

1. $\mathbf{s} \to \mathbf{s}_1 \to \mathbf{w}$,

2. $\mathbf{s} \to \mathbf{s}_2 \to \mathbf{w}$.

3. $\mathbf{s} \to \mathbf{s}_3 \to \mathbf{w}$.

Using Feynman's first rule (single path), $A(\mathbf{s} \to \mathbf{s}_i \to \mathbf{w}) = A(\mathbf{s} \to \mathbf{s}_i)\, A(\mathbf{s}_i \to \mathbf{w})$, the relevant transition amplitudes are:
$A(\mathbf{s} \to \mathbf{s}_1) = \mathbf{s}\dagger\mathbf{s}_1 = p$, $A(\mathbf{s}_1 \to \mathbf{w}) = \mathbf{s}_1\dagger\mathbf{w} = \sqrt{\frac{p}{2-p}}$, $A(\mathbf{s} \to \mathbf{s}_1 \to \mathbf{w}) = \sqrt{\frac{p^3}{2-p}}$.
$A(\mathbf{s} \to \mathbf{s}_2) = \mathbf{s}\dagger\mathbf{s}_2 = \sqrt{p(1-p)}$, $A(\mathbf{s}_2 \to \mathbf{w}) = \mathbf{s}_2\dagger\mathbf{w} = \sqrt{\frac{1-p}{2-p}}$, $A(\mathbf{s} \to \mathbf{s}_2 \to \mathbf{w}) = (1-p)\sqrt{\frac{p}{2-p}}$.
$A(\mathbf{s} \to \mathbf{s}_3) = \mathbf{s}\dagger\mathbf{s}_3 = \sqrt{p(1-p)}$, $A(\mathbf{s}_3 \to \mathbf{w}) = \mathbf{s}_3\dagger\mathbf{w} = \sqrt{\frac{1-p}{2-p}}$, $A(\mathbf{s} \to \mathbf{s}_3 \to \mathbf{w}) = (1-p)\sqrt{\frac{p}{2-p}}$.

We shall treat the paths $\mathbf{s} \to \mathbf{s}_2 \to \mathbf{w}$ and $\mathbf{s} \to \mathbf{s}_3 \to \mathbf{w}$ as indistinguishable from each other but both distinguishable from path $\mathbf{s} \to \mathbf{s}_1 \to \mathbf{w}$. Our argument for this is as follows. The path $\mathbf{s} \to \mathbf{s}_1 \to \mathbf{w}$ results in urn U containing two balls labeled 1. This is clearly distinguishable from paths $\mathbf{s} \to \mathbf{s}_2 \to \mathbf{w}$ and $\mathbf{s} \to \mathbf{s}_3 \to \mathbf{w}$, each of which result in urn U containing one ball labeled 1 and one ball labeled 2. From examining urn U, it is impossible to determine whether this arose by selecting ball 1 first (path $\mathbf{s} \to \mathbf{s}_2 \to \mathbf{w}$), then ball 2 (path $\mathbf{s} \to \mathbf{s}_3 \to \mathbf{w}$), or the other way round.

We apply Feynman's second rule (multiple indistinguishable paths) to find the amplitude of the transition $\mathbf{s} \to \mathbf{w}$, via $\mathbf{s}_2$ or via $\mathbf{s}_3$. We add the amplitudes of these two paths. Thus, $A(\mathbf{s} \to \mathbf{w})$, via $\mathbf{s}_2$ or $\mathbf{s}_3$ is $A(\mathbf{s} \to \mathbf{s}_2 \to \mathbf{w}) + A(\mathbf{s} \to \mathbf{s}_3 \to \mathbf{w}) = 2(1-p)\sqrt{\frac{p}{2-p}}$. The probability of this transition is $\left(2(1-p)\sqrt{\frac{p}{2-p}}\right)^2 = \frac{4p(1-p)^2}{2-p}$. The probability of the transition $\mathbf{s} \to \mathbf{s}_1 \to \mathbf{w}$ is $\left(\sqrt{\frac{p^3}{2-p}}\right)^2 = \frac{p^3}{2-p}$. We apply Feynman's third rule (multiple distinguishable paths) to get the total probability of the transition $\mathbf{s} \to \mathbf{w}$, via all paths. We add the two probabilities. This gives $P(\mathbf{s} \to \mathbf{w}) = \frac{p^3}{2-p} + \frac{4p(1-p)^2}{2-p} = \frac{5p^3-8p^2+4p}{2-p}$.

Quantum probability

Recall that $\mathbf{s}$ is the initial state of urn U, $\mathbf{t}$ is the state in which ball 1 is drawn and $\mathbf{w}$ is the state of urn U conditional on ball 1 having been drawn. We wish to calculate the probability, $P(\mathbf{s} \to \mathbf{t})$, of the transition $\mathbf{s} \to \mathbf{t}$. By the quantum law of reciprocity, $P(\mathbf{s} \to \mathbf{t}) = P(\mathbf{t} \to \mathbf{s})$. But $P(\mathbf{t} \to \mathbf{s})$ is the probability of the state of U conditional on drawing ball 1 from U. We have already calculated this to be $\frac{5p^3-8p^2+4p}{2-p}$.

Thus, if the probability of drawing ball 1 from the known urn K is p, then the quantum probability of drawing ball 1 from the unknown urn U is

$$Q(p) = \frac{5p^3 - 8p^2 + 4p}{2 - p}.$$

This completes the proof of Proposition 7. ■

The next proposition gives a simple result that is, nevertheless, crucial to this paper. Note that (18) of Proposition 7 is not used in deriving Proposition 8, below.

Proposition 8 *: Assume (Q), (B) and (L). Then quantum probabilities are matching probabilities, i.e.,*

$$m_i(p) = Q(p), \tag{20}$$

or, equivalently,

$$-\ln(-\ln m_i(p)) = -\ln(-\ln Q(p)). \tag{21}$$

Proof of Proposition 8: Let p be the probability of drawing a winning ball from the known (risky) urn K (recall that p is both a classical, or Kolmogorov, probability and also a quantum probability). Hence, by Proposition 7, the quantum probability of drawing a winning ball from the unknown urn U is $Q(p)$.

Let u_i be the utility function of subject i participating in the Ellsberg experiment as perceived by that subject (recall subsection 2.3, u_i is strictly increasing and normalized so that $u_i(0) = 0$). The subject wins the sum of money, $v > 0$, if a winning ball (a ball labeled 1) is drawn from the unknown urn U, but zero if a losing ball (a ball labeled 2) is drawn from that same urn. Hence, by our assumption (L), her projective expected utility in the sense of La Mura (2009) is

$$Q(p)\, u_i(v). \tag{22}$$

Suppose the contents of the unknown urn U are kept fixed but a new known (risky) urn, K_i, is constructed so that subject i is indifferent between U and K_i. Then, by definition, the probability of drawing a winning ball from

K_i is the matching probability $m_i(p)$. This gives subject i the projective expected utility

$$m_i(p)\, u_i(v). \tag{23}$$

Since the subject is indifferent between U and K, we must have $Q(p)\, u_i(v) = m_i(p)\, u_i(v)$. Since $u_i(v) > 0$, we get $m_i(p) = Q(p)$. ■

From Equation (18), we get

$$Q(0.1) = 0.17105,\ Q(0.5) = 0.41667,\ Q(0.9) = 0.69545$$

in close agreement with the evidence given by Dimmock et al. (2015) and our own evidence given in section 12, below.

Note that $Q(0.5) + Q(0.5) < 1$. This is because quantum probability theory is more general than classical (Kolmogorov) probability theory. Although quantum probabilities are never negative, they may, or may not, sum to 1. This may give an insight into why quantum probability theory can explain the Ellsberg paradox. Suppose the probability of drawing a ball of a winning color from urn K is $p = 0.5$. Then, by Proposition 1 of section 7, the classical probability of drawing a ball of a winning color from urn U is also $p = 0.5$. Hence, the subject has no reason to prefer K to U or U to K on probabilistic grounds. However, the quantum probability of drawing a ball of a winning color from urn U is $Q(p) = 0.41667$. Thus quantum probability theory correctly predicts that a subject would choose K over U; and the agreement with the evidence is quantitative as well as qualitative.

Recall how source dependent theory (section 7) explained the Ellsberg paradox. For $p = 0.5$, the source dependent model predicted that the probability of drawing a ball of a winning color from K or U is $p = 0.5$. However, in the source dependent model, probabilities enter the utility function *nonlinearly*; with *different* weights given to $p = 0.5$ in the two urns. Unlike the quantum model, these weights have to be determined ex post to fit the evidence.

The following result is easily established from (18) and Proposition 8.

$$p < 0.4 \Rightarrow Q(p) > p \Rightarrow \text{ambiguity seeking,}$$
$$p = 0.4 \Rightarrow Q(p) = p \Rightarrow \text{ambiguity neutrality,}$$
$$p > 0.4 \Rightarrow Q(p) < p \Rightarrow \text{ambiguity aversion.}$$

Thus, our model is in agreement with the empirically observed insensitivity.

Note that $Q(p)$ is *not* a probability weighting function. It should not be interpreted as p being the true porbability while $Q(p)$ is its over or under evaluation by subjects. The correct interpretation is that p is the probability (both classical and quantum) of drawing a winning ball from the known urn, K, p is also the classical probability that a winning ball will be drawn from the unknown urn U (Proposition 1 of section 7) and $Q(p)$ is the quantum probability with which a winning ball is drawn from urn U (Proposition 7, above).

On the other hand, classical explanations of the Ellsberg paradox start with the wrong probability, p, of drawing a winning ball from urn U, then transform it using a probability weighting function, $w(p)$, so as to fit the evidence. By contrast, quantum probability theory predicts the correct probability, $Q(p)$, of drawing a winning ball from urn U, without recourse to the device of a probability weighting function.

10 Inferring matching probabilities from cash equivalents

To find the matching probability from the cash equivalents that we obtained, it is necessary to assume a form for the utility function. This is necessary in the methodology in study 2 of Fox & Tversky (1995); which is incentive compatible. It is not necessary in the methodology of Dimmock et al. (2015). However, the latter is not incentive compatible. We use the power function (Assumption P, subsection 2.8) for the utility of player i,

Note that Assumption P is used in the proof of Proposition 9, below, but has not been used before. On the other hand, Proposition 7 is not used in deriving Proposition 9, below. In particular, the matching probabilities inferred from the observed cash equivalents using Proposition 9 may well reject the matching probabilities predicted by our quantum model (Propositions 7 and 8, above).

Proposition 9 *: Assume (L) and (P). Let v be the monetary payment to a subject if a winning ball is drawn. Let p be the probability of selecting a winning ball from the known urn (K). The monetary valuation of the known urn (K) to subject i is denoted by v_{iK} and the monetary valuation of the* *unknown urn (U) to subject i is denoted by* v_{iU}. *Then*

$$m_i\left(p\right)=p^{\frac{\ln v-\ln v_{iU}}{\ln v-\ln v_{iK}}}, \tag{24}$$

or, equivalently,

$$-\ln\left(-\ln m_i\left(p\right)\right)=\ln\left(\frac{\ln v-\ln v_{iK}}{\ln v-\ln v_{iU}}\right)-\ln\left(-\ln p\right). \tag{25}$$

Proof of Proposition 9:
Firstly, for the known urn (K), we have

$$(v_{iK})^{\sigma_i}=p(v)^{\sigma_i}. \tag{26}$$

Solve (26) for σ_i, to get

$$\sigma_i=\frac{-\ln p}{\ln v-\ln v_{iK}}, \tag{27}$$

where all quantities on the right hand side are known. Therefore, σ_i can be calculated using known quantities. Specifically, in our experiments, v is 10 *Yuan*, $p=0.1$, $p=0.5$ or $p=0.9$ and v_{iK} is the cash equivalent that we determined from the experiment.

Similarly, for the unknown urn (U), we have

$$(v_{iU})^{\sigma_i}=m_i\left(p\right)\left(v\right)^{\sigma_i}. \tag{28}$$

Solve (28) for $m_i\left(p\right)$, to get

$$m_i\left(p\right)=\left(\frac{v_{iU}}{v}\right)^{\sigma_i}. \tag{29}$$

Substitute from Equation (27) into Equation (29) to get

$$m_i\left(p\right)=\left(\frac{v_{iU}}{v}\right)^{\frac{-\ln p}{\ln v-\ln v_{iK}}}, \tag{30}$$

which is equivalent to (24). The latter, in turn, is equivalent to (25). ■

11 Experimental design

Our subjects were 295 undergraduate students from Qingdao Agricultural University in China. They attended 8 sessions; no one participated in more than one session. The experimental instructions are given in the Appendix A. We obtained the ethics approval from University of Leicester: The University Ethics Sub-Committee for Sociology; Politics and IR; Lifelong Learning; Criminology; Economics and the School of Education (Ethics Reference: 7274-mw323-economics).

Our treatment was a paper-based classroom experiment. There were three tasks, *Task 1*, *Task 2* and *Task 3*, that were, respectively, designed to implement the three cases $p = 0.5$, $p = 0.1$, $p = 0.9$. Each task required two tables to be completed. The materials for each task were handed out at the beginning of that task and collected before the next task started.

In each task, there is one known urn (Box K) and one unknown urn (Box U). The composition of the 100 colored balls of k different colors in Box K is known; varying this composition gives us the three cases $p = 0.5$, 0.1, 0.9. Box U contains 100 colored balls of the same colors as in Box K, but in unknown proportions. The composition of Box U is randomly decided at the end of the experiment in the following way. Each ball is equally likely to be drawn. The random draw follows the uniform distribution. For example, in Task 2, there are in total 10 different colors. A priori, each color is equally likely to be drawn. Thus, at each stage of the construction of Box U, each color has a probability 0.1 of being the color of the next ball to be placed in Box U. There can be from 0 to 100 balls of any particular color but subject to the restriction that the total number of balls in Box U is 100 balls. The prize for drawing a winning-color ball is 10 *Yuan* whether it is drawn from Box K or Box U. We now explain the three tasks.

1. In Task 1, there are 50 purple balls and 50 yellow balls in Box K, and *purple* is the winning color ($p = 0.5$, by Assumption I). These are the same colors as chosen by Dimmock et al. (2015). The decision maker is shown two tables. In Table I, the choices are to express a preference to receive a monetary amount x for sure or express a preference for betting that a purple ball will be drawn from Box K. A third choice is also offered, namely, to express indifference between x or betting that a purple ball will be drawn from Box K. The monetary amount is varied from $x = 0$ to $x = 10$ and subjects have to state a choice in each case.

The experiments were conducted in China, so the monetary amount is in units of Chinese Yuan. Box U has 100 balls that are either purple or yellow but the proportions are unknown; as explained above. Table II replaces Box K in Table I with Box U but it is otherwise identical. At the end of the experiment, one of the choices from Task 1 is picked at random to be played for real.

2. In Task 2, there are 10 different colors (including purple) in Box K, and *purple* is the winning color ($p = 0.1$, by Assumption I). Box U has 100 balls of the same 10 colors but in unknown proportions. The remaining procedure is as described in Task 1.

3. In Task 3, there are 10 different colors (including purple) in Box K, and the winning color is any ball that is not purple ($p = 0.9$, by Assumption I). Box U has 100 balls of the same 10 colors but in unknown proportions. The remaining procedure is as described in Task 1.

12 Experimental results

12.1 Aggregate data analysis

Consider a sample of N subjects. Choose a probability, p, for drawing a winning ball from urn K. Find the matching probability, $m_i(p)$, for each subject, i, $i = 1, 2, ..., N$. Let the sample average be $m(p) = \frac{1}{N}\sum_{i=1}^{N} m_i(p)$.

Even if our quantum model is correct, it might not be surprising to see much unsystematic variability in the matching probabilities, $m_i(p)$, across the sample. Specifically, let $m_i(p) = Q(p) + \epsilon_i$, where $E(\epsilon_i) = 0$, $i = 1, 2, ..., N$ and ϵ_i and ϵ_j are identically and independently for $i \neq j$. Let $s^2 = \frac{1}{N-1}\sum_{i=1}^{N} (m_i(p) - m(p))^2$ and $t = \frac{m(p)-Q(p)}{s/\sqrt{N}}$. For sufficiently large N, we would expect t to be approximately normally distributed with mean 0 and variance 1. See, for example, Chapter 5 of Wooldridge (2015). For ease of reference, we give the critical values for each of the conventional significance levels (10%, 5%, 1%) for a two-tailed test for the standard normal distribution in Table 1, below.

We collected in total 19470 ($= 11 \times 2 \times 3 \times 295$) data points. This comes from 11 data points (for the 11 rows of Tables I and II, see Appendix A), two tables corresponding to the known and unknown urns (Tables I and II in Appendix A), 3 tasks (Task 1, Task 2, and Task 3) and 295 subjects in

Table 1: Significance levels and the corresponding critical values.

Significance level	Critical value
10%	± 1.64
5%	± 1.96
1%	± 2.58

the experiment. There were 259, 262 and 263 consistent decision makers for the $p = 0.1$, $p = 0.5$ and $p = 0.9$ cases, respectively. We discarded the inconsistent decision makers from the analysis as follows. We discarded data with the following two patterns: Firstly, choosing more than once in the "Indifference" column in the table. Secondly, choosing back and forth in any two or three columns. For the first case, we cannot identify the unique cash equivalent. In the second case, it seems that the subjects don't have consistent preferences. This left us with over 250 subjects.

We estimated the cash equivalents for the decisions in the two tables in Appendix A in the following way. If there is one unique tick in the "Indifference" column in the table, then the cash equivalent is the corresponding amount of money s/he receives for sure (x); On the other hand, if there is no tick in the "Indifference" column, then the cash equivalent is estimated by the midpoint between the lowest amount of money that is preferred to the uncertain bet, and the highest amount of money for which the bet was preferred; we are following the methodology in study 2 of Fox & Tversky (1995).

Since all quantities on the right hand side of (24) are known, the matching probability can be found (recall v_{iK} and v_{iU} are the cash equivalents and v is the prize). Following this approach, we find the mean matching probabilities, $m(p)$, and standard deviations, which are listed in Table 2, below. The fifth column of Table 2 shows the theoretical predictions for the three matching probabilities, $Q(0.1)$, $Q(0.5)$, and $Q(0.9)$, respectively. The theoretically predicted values are found by substituting the values of p, 0.1, 0.5, 0.9 into (18). One subject chose $v_{iK} = v = 10$, for $p = 0.9$. Since the denominator in $\frac{\ln v - \ln v_{iU}}{\ln v - \ln v_{iK}}$ of (24) would then be zero for these values, we discarded this observation.

Table 2, below, shows that the theoretically predicted matching probabilities are quite close to the mean values we obtained from our experiments.

Our null and alternative hypotheses are: $H_0 : m(p) = Q(p)$ and $H_1 : m(p) \neq Q(p)$. From Tables 1 and 2, 1.4758, 1.4437, 2.3906 are all less than

Table 2: t-test for the means.

Matching probability	Mean	Standard deviation	Sample size	Quantum probability Q	t-stat
$m\,(0.1)$	0.1864	0.1708	259	0.1711	1.4437
$m\,(0.5)$	0.4038	0.1416	262	0.4167	-1.4758
$m\,(0.9)$	0.7258	0.2056	263	0.6955	2.3906

2.58. Thus, our experimental results fail to reject our quantum model at the 1% level of significance. Since $m\,(0.1) > 0.1$, $m\,(0.5) < 0.5$, $m\,(0.9) < 0.9$, we find ambiguity seeking for the low probability but ambiguity aversion for the medium and high probabilities.

12.2 Demographic results

In their answers to question 8 on the post-experimental questionnaire (Appendix B), only 4 out of the 295 subjects reported that color affected their decisions. In their answers to question 6, almost none reported prior experience with similar experiments in the past. In their answers to question 4, Degree of study, all students simply gave "undergraduate", thus giving us no useful information. From the answers to question 3 (Field of study), we obtained the data for economics/non-economics. Not surprisingly, we found high colinearity between year of study and age, so we have not reported the latter.

12.2.1 Mann-Whitney U tests

We used two-sided Mann-Whitney U test (nonparametric test) to examine if the demographic characteristics in Appendix B affected the subjects' reported matching probabilities for $p = 0.1$, $p = 0.5$ and $p = 0.9$ in our treatment. The results are shown in Table 3. At the 1% level, no significant differences were found between any of the two groups (male/female; economics/non-economics students; statistics/non-statistics students).

Note: In Table 3, "No" denotes no significant difference at 1%; "No*" denotes difference significant at 5% but not at 1%; "No**" denotes difference significant at 10% but not at 1% nor 5%.

Table 3: Mann-Whitney U test results.

Group	Matching probability	MWU p-value	Sig diff
Male vs. Female	$m\,(0.1)$	0.9533	No
	$m\,(0.5)$	0.2825	No
	$m\,(0.9)$	0.5205	No
Econ vs. Non-econ	$m\,(0.1)$	0.8941	No
	$m\,(0.5)$	0.7529	No
	$m\,(0.9)$	0.1230	No
Stats vs. Non-stats	$m\,(0.1)$	0.0496	No*
	$m\,(0.5)$	0.2053	No
	$m\,(0.9)$	0.7413	No
Year 1 vs. Year 2	$m\,(0.1)$	0.2944	No
	$m\,(0.5)$	0.3981	No
	$m\,(0.9)$	0.0546	No**
Year 2 vs. Year 3	$m\,(0.1)$	0.6826	No
	$m\,(0.5)$	0.8746	No
	$m\,(0.9)$	0.0245	No*
Year 1 vs. Year 3	$m\,(0.1)$	0.0998	No**
	$m\,(0.5)$	0.2693	No
	$m\,(0.9)$	0.4442	No

12.2.2 t-tests

For each demographic group, we also performed a t-test to see if the average reported matching probability, $m(p)$, differed significantly from the predicted value of the quantum probability, $Q(p)$. We report the results in Table 4, below. The only group that showed a significant difference was the group of students with prior training in statistics.

To keep things in perspective, we report in Table 5 how well the classical prediction fairs against the evidence.

Since the absolute values of the t-statistics in Table 5 are large relative to the critical values (Table 1), it follows that the classical prediction is strongly rejected even for students trained in statistics.

13 Comparing the quantum model with the source dependent model

Here we compare the predictive success of our quantum model (section 9) with that of the source dependent model (section 7) using the Diebold-Mariano (1995) forecasting test (see, also, Diebold, 2014). We divide our data into three subsamples. We use each pair of subsamples to estimate the parameters α and β of the source dependent model (37), below, then we forecast for the third subsample. The quantum model has no parameters to estimate. In each case, we find the predictive performance of the two models to be statistically insignificantly different from each other.

13.1 Data

In total, we have 784 data points. Of these, we had to exclude five data points because their matching probabilities, m_i, were zero and, hence, cannot be used to forecast for the source dependent model; see (32), (37), below. Thus we had a total of 779 usable data points, as given in Table 6, below, where p is the probability of selecting a winning ball from the known urn K.

Let m_i be the matching probabilities inferred from the elicited certainty equivalents (Proposition 9) and let

Table 4: t-test results.

Group	Matching probability	Mean	Standard deviation	Sample size	t-stat	Sig diff
	$m(0.1)$	0.1786	0.1296	23	0.2789	No
Econ	$m(0.5)$	0.4080	0.1706	23	−0.2449	No
	$m(0.9)$	0.7603	0.2315	20	1.2531	No
	$m(0.1)$	0.1871	0.1429	236	1.7275	No
Non-econ	$m(0.5)$	0.4126	0.1712	239	−0.3662	No
	$m(0.9)$	0.7227	0.2036	243	2.0881	No
	$m(0.1)$	0.1829	0.1362	116	0.9371	No
Male	$m(0.5)$	0.4239	0.1738	120	0.4557	No
	$m(0.9)$	0.7319	0.2046	124	1.9838	No
	$m(0.1)$	0.1892	0.1462	143	1.4846	No
Female	$m(0.5)$	0.4023	0.1682	142	−1.0181	No
	$m(0.9)$	0.7199	0.2071	139	1.3919	No
	$m(0.1)$	0.1773	0.1465	171	0.5579	No
Year 1	$m(0.5)$	0.4204	0.1729	173	0.2838	No
	$m(0.9)$	0.7195	0.2136	172	1.4763	No
	$m(0.1)$	0.1817	0.0939	27	0.5893	No
Year 2	$m(0.5)$	0.4070	0.1357	29	−0.3837	No
	$m(0.9)$	0.7820	0.1799	28	2.5457	No
	$m(0.1)$	0.2150	0.1576	60	2.1601	No
Year 3	$m(0.5)$	0.3909	0.1816	59	−1.0900	No
	$m(0.9)$	0.7134	0.1930	62	0.7323	No
	$m(0.1)$	0.2076	0.1555	126	2.638	Yes (1%)
Stat	$m(0.5)$	0.4043	0.1653	128	−0.847	No
	$m(0.9)$	0.7398	0.1710	130	2.957	Yes (1%)
	$m(0.1)$	0.1663	0.1243	133	−0.441	No
Non-stat	$m(0.5)$	0.4198	0.1762	134	0.206	No
	$m(0.9)$	0.7117	0.2344	133	0.800	No

Table 5: Comparison with classical probabilities.

Matching probability	Mean	Standard deviation	Sample size	Classical probability	t-stat	Sig diff
$m(0.1)$	0.2076	0.1555	126	0.1	7.7672	Yes
$m(0.5)$	0.4043	0.1653	128	0.5	−6.55	Yes
$m(0.9)$	0.7398	0.1710	130	0.9	−10.682	Yes

Table 6: Task information.

$i = 1 - 257$	$p = 0.1$	Task 2
$i = 258 - 517$	$p = 0.5$	Task 1
$i = 518 - 779$	$p = 0.9$	Task 3

$$X_i = (-\ln(-\ln p_i)), \; i = 1, 2, ..., 779, \quad (31)$$

$$Y_i = (-\ln(-\ln m_i)), \; i = 1, 2, ..., 779. \quad (32)$$

13.2 The quantum model

Let $Q(p_i)$ be the quantum probability of drawing a winning ball from urn U (Proposition 7). Let $\widetilde{m}_i$ the matching probability predicted by the quantum model. By Proposition 8, $\widetilde{m}_i = Q(p_i)$. Hence, $-\ln(-\ln \widetilde{m}_i) = -\ln(-\ln Q(p_i))$. Letting $\widetilde{Y}_i = -\ln(-\ln \widetilde{m}_i)$, the quantum model can be written as

$$\widetilde{Y}_i = -\ln(-\ln Q(p_i)), \; i = 1 - 779, \quad (33)$$

and the squares of the forecast errors (recall (32)) are

$$\widetilde{e}_i^2 = \left(Y_i - \widetilde{Y}_i\right)^2, \; i = 1 - 779. \quad (34)$$

13.3 The source dependent model

The definition of matching probabilities (subsection 2.2) is operational, and does not depend on the underlying decision theory assumed. However, eliciting matching probabilities does require assumptions (as is the case with all observations in science, recall the discussion in subsection 2.9). For example, Dimmock et al. (2015) used an incentive incompatible method but assumed that their subjects did not exploit this (and their evidence supported this). We used the incentive compatible method of Fox and Tversky (1995). This was carried out in two steps. First we elicited certainty equivalents (section 11 and Appendix A). Then, from these, we inferred matching probabilities (section 10).

In particular, $m_i(p)$ was derived under the assumption $(v_{iU})^{\sigma_i} = m_i(p)(v)^{\sigma_i}$ (recall (28) of the proof of Proposition 9). On the other hand, $m_i^*(p)$ was derived under the assumption $(v_{iU})^{\sigma_i} = w_{iK}(m_i^*(p))(v)^{\sigma_i}$ (recall (11) of the proof of Proposition 6). Hence, $m_i^*(p)$ need not be the same as $m_i(p)$. To compare the forecasting performance of two models, we have to compare them on the same data set. In particular, we have to compare their forecasts for the same dependent variable. Proposition 10, immediately below, derives the prediction of the source dependent model (section 7) for the matching probability, $m_i(p)$.

Proposition 10 *: Let $m_i(p)$ be the matching probability. Let X_i and Y_i be as in (31), (32). Let*

$$\alpha_i = \ln \beta_{iK} - \ln \beta_{iU}, \tag{35}$$

$$\beta_i = 1 + \alpha_{iU} - \alpha_{iK}, \tag{36}$$

then, under Assumption (P), the source dependent model (subsection 7) implies that

$$Y_i = \alpha_i + \beta_i X_i. \tag{37}$$

Proof of Proposition 10: From Propositions 4 and 6, we get

$$\frac{\ln \beta_{iK} - \ln \beta_{iU}}{\alpha_{iK}} + \frac{\alpha_{iU}}{\alpha_{iK}}(-\ln(-\ln p)) = -\ln(-\ln p) + \frac{1}{\alpha_{iK}} \ln\left(\frac{\ln v - \ln v_{iK}}{\ln v - \ln v_{iU}}\right). \tag{38}$$

Substituting from (25) of Proposition 9 into (38), and rearranging, gives

$$-\ln(-\ln m_i(p)) = \ln \beta_{iK} - \ln \beta_{iU} + (1 + \alpha_{iU} - \alpha_{iK})(-\ln(-\ln p)). \tag{39}$$

Substituting from (35) and (36) into (39) gives (37). ■

Let $\widehat{m}_i$ be the matching probability predicted by the source dependent model and let $\widehat{Y}_i = -\ln(-\ln \widehat{m}_i)$. Let $\widehat{e}_i^2 = \left(Y_i - \widehat{Y}_i\right)^2$, $i = 1 - 779$. From Proposition 10, we get

$$\widehat{Y}_i = \widehat{\alpha}(1) + \widehat{\beta}(1) X_i,\ i = 1 - 257,$$
$$\widehat{Y}_i = \widehat{\alpha}(2) + \widehat{\beta}(2) X_i,\ i = 258 - 517,$$
$$\widehat{Y}_i = \widehat{\alpha}(3) + \widehat{\beta}(3) X_i,\ i = 518 - 779,$$

where $\widehat{\alpha}(1)$, $\widehat{\beta}(1)$ are estimated from the data for $i = 258 - 779$; $\widehat{\alpha}(2)$, $\widehat{\beta}(2)$ are estimated from the data for $i = 1 - 257$, $i = 518 - 779$ and $\widehat{\alpha}(3)$, $\widehat{\beta}(3)$ are estimated from the data for $i = 1 - 517$.

13.4 The Diebold-Mariano forecasting test (DM)

Let $d_i = \widetilde{e}_i^2 - \widehat{e}_i^2$, $i = 1 - 779$ and let

$$
\begin{aligned}
\overline{d}\,(1) &= \frac{1}{257}\sum_{i=1}^{257} d_i,\ \widehat{\sigma}^2\,(1) = \frac{1}{257}\sum_{i=1}^{257}\left(d_i - \overline{d}\,(1)\right)^2; \\
\overline{d}\,(2) &= \frac{1}{260}\sum_{i=258}^{517} d_i,\ \widehat{\sigma}^2\,(2) = \frac{1}{260}\sum_{i=258}^{517}\left(d_i - \overline{d}\,(2)\right)^2; \\
\overline{d}\,(3) &= \frac{1}{262}\sum_{i=518}^{779} d_i,\ \widehat{\sigma}^2\,(3) = \frac{1}{262}\sum_{i=518}^{779}\left(d_i - \overline{d}\,(3)\right)^2.
\end{aligned}
$$

Thus, $\widehat{e}_i^2$ is the square of the forecast error for observation i for the source dependent model and $\widetilde{e}_i^2$ is the square of the forecast error for observation i for the quantum model, d_i is the difference between these two errors, $\overline{d}$ is the sample average of these differences and $\widehat{\sigma}^2$ is a consistent estimator of the variance of the difference between the two forecasts.

If $\overline{d} > 0$ then, on average for the sample considered, the quantum model forecasts better than the source dependent model (and, conversely, if $\overline{d} < 0$). However, we would like to find out more. We would like to know, is this difference significant? In particular, we would like to know if this difference is significantly different from zero, i.e., we want to carry out the following hypothesis test:

$$H_0 : E\,(d_i) = 0,\ H_1 : E\,(d_i) \neq 0.$$

Assuming covariance stationarity, Diebold and Mariano (1995) proved that under H_0, $\overline{d}$ has an asymptotically normal distribution. More precisely, the distribution of $\frac{\overline{d}}{\widehat{\sigma}}$ converges, in distribution, to the standard normal as the sample size goes to ∞: $z = \frac{\overline{d}}{\widehat{\sigma}} \xrightarrow{D} N\,(0, 1)$.

13.5 Results

Table 7, below, summarizes our results. $\widehat{\alpha}$, $\widehat{\beta}$ are the estimated values of α and β for the source dependent model (7). The last column gives the values of the z-statistic.

From the above table, we see that the quantum model forecasts better than the source dependent model for the two subsamples $i = 1 - 257$ and $i = 258 - 517$ but worse for the subsample $i = 518 - 779$. However, more importantly, the difference is statistically insignificant for each of the three

Table 7: DM results.

i	p	Task	$\widehat{\alpha}$	$\widehat{\beta}$	z-statistic
$1-257$	0.1	2	-1.006	1.336	-0.606
$258-517$	0.5	1	0.581	0.226	-0.430
$518-779$	0.9	3	-0.103	0.585	0.273

subsamples. In particular, the z-statistic, in each case, is well within the acceptance region $[-1.96, 1.96]$ at the 5% level.

In other words, our data show that the quantum model and the source dependent model have no statistically significant difference in predictive accuracy.

14 Summary and suggestions for further research

In this paper we considered a simple quantum decision model of the Ellsberg paradox. We reported the results of an experiment we performed to test the matching probabilities predicted by this model using an incentive compatible method. We found that the theoretical predictions of the model were in conformity with our experimental results.

The source dependent model is probably the most successful non-quantum model of ambiguity. Our forecasting tests showed that there were no statistically significant differences between the predictions of the source dependent model and our quantum model. However, and unlike our quantum model, the source dependent model requires the specification of probability weighting functions in order to fit the evidence. This suggests that much of what is normally attributed to probability weighting may actually be due to quantum probability.

Immediately below, we mention three of the directions in which the work of this paper can be extended.

Our assumption of how the unknown (or ambiguous urn) U is constructed in the minds of the subjects (assumption B) is consistent with the evidence reported in Pulford & Colman (2008). However, it deserves further independent testing.

We elicited the certainty equivalents of risky and unknown lotteries (as in, for example, Fox and Tversky, 1995). Then we inferred the matching probabilities using a simple specification of the utility function. An alternative is for each unknown lottery to elicit the equivalent risky lottery, as in Kocher et al. (2018).[16] The Kocher et al. (2018) method is incentive compatible and does not require the specification of a utility function. Nevertheless, it is interesting that three different methodologies applied to three different data sets (Dimmock et al., 2015, Kocher et al., 2018 and this paper) yielded similar matching probabilities.

We considered only lotteries with zero or positive outcomes. Future work can consider lotteries with gains, lotteries with losses and mixed lotteries with gains and losses (as in Kocher et al., 2018). This would allow us to test what we called projective prospect theory. The latter introduces reference dependence and loss aversion into projective expected utility (La Mura, 2009).

ACKNOWLEDGEMENTS

We are grateful for valuable and critical comments from Jerome Busemeyer, Wojciech Charemza, Andrew Colman, Stephen Dimmock, Martin Dufwinberg, Ehtibar Dzhafarov, Emmanuel Haven, Andrei Khrennikov, Jesse Matheson, Briony Pulford, David Rojo-Arjona, Sandro Sozzo, Stefan Trautmann and Peter Wakker. However, all errors are our own.

Declarations of interest: none.

Funding source: University of Leicester PhD scholarship for Mengxing Wei.

References

[1] Abdellaoui M. A., Baillon A., Placido L. & Wakker P. P. (2011). The rich domain of uncertainty: source functions and their experimental implementation. *American Economic Review*, 101(2). 695–723.

[2] Aerts, D. and Aerts, S. (1994). Applications of quantum statistics in psychological studies of decision processes. *Foundations of Science*, 1, 85-97.

[16] We are grateful to Stefan Trautmann for drawing our attention to this.

[3] al-Nowaihi A., & Dhami, S. (2017). The Ellsberg paradox: A challenge to quantum decision theory? *Journal of Mathematical Psychology*, 78, 40-50. Available online (2016) http://dx.doi.org/10.1016/j.jmp.2016.08.003.

[4] Asano, M., Basieva, I., Khrennikov, A., Ohya, M., & Tanaka, Y. (2016). A quantum-like model of selection behavior. *Journal of Mathematical Psychology.* http://dx.doi.org/10.1016/j.jmp.2016.07.006.

[5] Basieva I. & Khrennikov A. (2015). On the possibility to combine the order effect with sequential reproducibility for quantum measurements. *Foundations of Physics*, 45(10), 1379–1393.

[6] Billingsley, P. (1995). *Probability and Measure*, third edition. John Wiley & Sons, Inc.

[7] Burnham, K. P., Anderson, D. R. 2002. *Model Selection and Multiple Inference: a practical information-theoretic approach.* Second edition. Springer-Verlag, New York.

[8] Busemeyer J. R., & Bruza P. D. (2012). *Quantum Models of Cognition and Decision.* Cambridge University Press.

[9] Camerer C. (2003). *Behavioral Game Theory.* Princeton University Press.

[10] Choquet G. (1953-1954). Theory of Capacities. *Annales de le'Institut Fourier*, **5** (Grenoble): 131-295.

[11] Conte A. and Hey, J.D. (2013). Assessing multiple prior models of behavior under ambiguity. *Journal of Risk and Uncertainty* **46**(**2**). 113–132.

[12] Dhami, S. (2016). Foundations of behavioral economic analysis. Oxford University Press: Oxford.

[13] Diebold, F. X. & Mariano R.S. (1995). Comparing Predictive Accuracy. Journal of Business and Economic Statistics, 13, 253-263.

[14] Diebold, F. X. (2014). Comparing Predictive Accuracy, twenty years later: A personal perspective on the use and abuse of Diebold-Mariano Tests. *Journal of Business and Economic Statistics* Invited Lecture, Allied Social Science Association Meetings, Philadelphia.

[15] Dimmock S. G., Kouwenberg R., & Wakker, P. P. (2015). Ambiguity attitudes in a large representative sample. *Management Science*. Article in Advance. Published Online: November 2, 2015. 10.1287/mnsc.2015.2198

[16] Ellsberg D. (1961). Risk, ambiguity and the Savage axioms. *Quarterly Journal of Economics*, 75, 643–669.

[17] Ellsberg D. (2001). *Risk, Ambiguity and Decision*. Garland Publishers, New York. Original Ph.D. dissertation: Ellsberg D. 1962. Risk, Ambiguity and Decision. Harvard, Cambridge, MA.

[18] Fox C. R., & Tversky A. (1995). Ambiguity aversion and comparative ignorance. *Quarterly Journal of Economics*, 110(3), 585–603.

[19] French K. R. and Poterba J. M. (1991). Investor diversification and international equity markets. *American Economic Review* **81(2)**. 222-226.

[20] Ghirardato P., Maccheroni F. and Marinacci M. (2004). Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory* **118(2)**. 133–173.

[21] Gilboa I. (1987). Expected utility with purely subjective non-additive probabilities. *Journal of Mathematical Economics* **16**. 65-88.

[22] Gilboa I. (2009). *Theory of Decision under Uncertainty*. Cambridge: Cambridge University Press.

[23] Gilboa I. and Schmeidler. (1989). Maximin Expected Utility with a Non-Unique Prior. *Journal of Mathematical Economics* **18**. 141-153.

[24] Gleason, A. M. (1957). Measures on the closed subspaces of a Hilbert space. *Journal of Mathematical Mechanics*, 6, 885-893.

[25] Gnedenko B. V. (1968). *The theory of probability*. Forth edition. Chelsea Publishing Company, New York, N. Y.

[26] Harrison, G. W. (1986). An experimental test for risk aversion. *Economics Letters*, 21, 7–11.

[27] Haven E. and Khrennikov A. (2013). *Quantum Social Science*. Cambridge University Press.

[28] Hey J. D., Lotito, G. and Maffioletti A. (2010). The descriptive and predictive adequacy of theories of decision making under uncertainty/ambiguity. *Journal of Risk and Uncertainty* **41**(**2**). 81–111.

[29] Hurwicz L. (1951). Some specification problems and applications to econometric models. *Econometrica* **19**. 343–344.

[30] Kahneman D. and Tversky A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica* **47**. 263-291.

[31] Keynes J. M. (1921). *A treatise on probability*. London: Macmillan Co.

[32] Khrennikov A. (2010). *Ubiquitous Quantum Structure: From Psychology to Finance*. Springer.

[33] Khrennikov A. and Haven E. (2009). Quantum mechanics and violations of the sure-thing principle: The use of probability interference and other concepts. *Journal of Mathematical Psychology* **53**. 378-388.

[34] Khrennikov A., Basieva I., Dzhafarov E. N. and Busemeyer J. R. (2014). Quantum models for psychological measurement: An unsolved problem. PLoS ONE **9**(**10**) e110909.

[35] Klibanoff P., Marinacci M. and Mukerji S. (2005). A smooth model of decision making under ambiguity. *Econometrica* **73**(**6**). 1849–1892.

[36] Kocher M. G., Lahno A. M., & Trautmann S. T. (2018). Ambiguity aversion is not universal. *European Economic Review*, 101, 268-283.

[37] Kolmogorov A. N. (1950). *Foundations of the Theory of Probability*, New York: Chelsea Publishing Co. Original German, 1933.

[38] Kothiyal A., Spinu, V., & Wakker P. P. (2014). An experimental test of prospect theory for predicting choice under ambiguity. *Journal of Risk and Uncertainty*, 48(1), 1–17.

[39] La Mura P. (2009). Projective expected utility. *Journal of Mathematical Psychology*, 53(5), 408–414.

[40] Luce R. D. and Raiffa H. (1957). *Games and Decisions*. New York: Wiley.

[41] Obstfeld M. and Rogoff K. (2000). The six major puzzles in international economics: Is there a common cause? *NBER Macroeconomic Annual* **15(1)**. 339-390.

[42] Prelec D. (1998). The probability weighting function. *Econometrica*, 60, 497–524.

[43] Pulford B. D., & Colman A. M. (2008). Ambiguity aversion in Ellsberg urns with few balls. *Experimental Psychology*, 55(1), 31–37.

[44] Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization* **3**. 323-343.

[45] Quiggin, J. (1993). *Generalized Expected Utility Theory - The Rank-Dependent Model.* Kulwar Academic Publishers, Boston/Dordrecht/London.

[46] Savage L. J. (1954). *The Foundations of Statistics*. New York: Wiley and Sons.

[47] Schmeidler D. (1989). Subjective probability and expected utility without additivity. *Econometrica* **57**. 571-587.

[48] Segal U. (1990). Two-stage lotteries without the reduction axiom. *Econometrica* **58**(**2**). 349–377.

[49] Thaler R. H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making* **12**. 183-206.

[50] Tolman R. C. (1938). *The principles of statistical mechanics*. Oxford University Press, Oxford.

[51] Tversky A. & Kahneman D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 297-323.

[52] Vincent S., Kovalenko T., Yukalov V. I., & Sornette D. (2017). Calibration of Quantum Decision Theory, aversion to large losses and predictability of probabilistic choices. ETH Zurich. http://ssrn.com/abstract=2275279.

[53] von Neumann J. (1955). *Mathematical Foundations of Quantum Theory.* Princeton Press. Original German, 1932.

[54] von Neumann J. and Morgenstern O. (1947). *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.

[55] Wakker P. P. (2010). *Prospect Theory for Risk and Ambiguity.* Cambridge: Cambridge University Press.

[56] Wooldridge J. M. (2015). *Introductory econometrics: A modern approach*. Nelson Education.

15 Appendix A: Experimental Instruction (translation from Chinese instruction)

General information on the experiment

You are now participating in an economic experiment. If you read the following explanations carefully, you may be able to earn some money depending on your decisions. You will receive 5 Yuan for participation. This is irrespective of your decisions in the experiment. During the experiment you are not allowed to communicate with other participants in any way. If you have questions, please raise your hand, and the experimenter will come to your desk. The experiment will be carried out only once.

This experiment is paper based. there are three tasks: Task 1, Task 2 and Task 3. In each task, there are two boxes- Box K and Box U, and each box contains 100 colored balls. The composition of the balls is known for Box K but unknown for Box U. After you complete a task, the experimenter will collect the materials for that task and you will receive the materials for the next task.

Task 1:

There are 50 *purple* balls and 50 *yellow* balls in Box K. For each of the eleven rows in Table I, tick *exactly one* of the following boxes: "Receive x Yuan for sure", "Indifferent" or "Play Box K".

Box U contains 100 balls (purple or yellow) but in *unknown* proportions. Thus Box U can contain any number of purple balls from 0 to 100 and any number of yellow balls from 0 to 100 provided the sum of balls (purple plus yellow) is 100. The composition of Box U will be randomly decided at the end of the experiment. For each of the eleven rows in Table II, tick *exactly* *one* of the following boxes: “Receive x Yuan for sure”, “Indifferent” or “Play Box U”.

In each table, if you believe that you are indifferent between the choice in the left column and the right column, you may tick the box under the middle column “Indifferent”.

At the end of the experiment, one of the eleven rows of Table I or one of the eleven rows of Table II will be selected at random and played for real money. In Table I, you will receive x Yuan for sure if you have ticked the box under “Receive x Yuan for sure” or, if you have ticked the box under “Play Box K”, you will win 10 *Yuan* if a *purple* ball is drawn from Box K (otherwise you win nothing). In Table II, you will receive x Yuan for sure if you have ticked the box under “Receive x Yuan for sure” or, if you have ticked the box under “Play Box U” you will win 10 Yuan if a *purple* ball is drawn from Box U (otherwise you win nothing). In each table, if you have ticked “Indifferent” in the randomly selected row, then one of the left or right cells in this selected row will be randomly chosen to play for real.

Table I		
Receive x Yuan for sure	Indifferent	Play Box K
$x = 10$ ☐	☐	☐
$x = 9$ ☐	☐	☐
$x = 8$ ☐	☐	☐
$x = 7$ ☐	☐	☐
$x = 6$ ☐	☐	☐
$x = 5$ ☐	☐	☐
$x = 4$ ☐	☐	☐
$x = 3$ ☐	☐	☐
$x = 2$ ☐	☐	☐
$x = 1$ ☐	☐	☐
$x = 0$ ☐	☐	☐

After you complete Task 1, the experimenter will collect the materials for Task 1 and you will receive the materials for Task 2.

Task 2:

There are 100 balls of 10 different colors (including *purple*) in Box K. There are exactly 10 balls of each color. For each of the eleven rows in Table I, tick *exactly one* of the following boxes: “Receive x Yuan for sure”,

Table II		
Receive x Yuan for sure	Indifferent	Play Box U
$x = 10$ ☐	☐	☐
$x = 9$ ☐	☐	☐
$x = 8$ ☐	☐	☐
$x = 7$ ☐	☐	☐
$x = 6$ ☐	☐	☐
$x = 5$ ☐	☐	☐
$x = 4$ ☐	☐	☐
$x = 3$ ☐	☐	☐
$x = 2$ ☐	☐	☐
$x = 1$ ☐	☐	☐
$x = 0$ ☐	☐	☐

"Indifferent" or "Play Box K".

Box U contains 100 balls of the same colors as in Box K but in *unknown* proportions. Thus, Box U could contain any number of purple balls from 0 to 100. And similarly for each of the other 9 colors (provided the sum of balls of all colors is 100). The composition of Box U will be randomly decided at the end of the experiment. For each of the eleven rows in Table II, tick *exactly one* of the following boxes: "Receive x Yuan for sure", "Indifferent" or "Play Box U".

In each table, if you believe that you are indifferent between the choice in the left column and the right column, you may tick the box under the middle column "Indifferent".

At the end of the experiment, one of the eleven rows of Table I or one of the eleven rows of Table II will be selected at random and played for real money. In Table I, you will receive x Yuan for sure if you have ticked the box under "Receive x Yuan for sure". However, if you have ticked "Play Box K", then you shall win 10 *Yuan* if a *purple* ball is drawn from Box K (otherwise you win nothing). In Table II, you will receive x Yuan for sure if you have ticked the box "Receive x Yuan for sure". However, if you have ticked the box "Play Box U" then you win 10 *Yuan* if a *purple* ball is drawn from Box U (otherwise you win nothing). In each table, suppose that you ticked "Indifferent" in the randomly selected row, then one of the left or right cells in this selected row will be randomly chosen to play for real.

After you complete Task 2, the experimenter will collect the materials for Task 2 and you will receive the materials for Task 3.

Task 3:

As in task 2, there are 100 balls in Box K of 10 different colors (including *purple*). There are exactly 10 balls of each color. For each of the eleven rows in Table I, tick *exactly one* of the following boxes: The box "Receive x Yuan for sure", "Indifferent" or "Play Box K".

As with task 2, Box U contains 100 balls of the same colors as in Box K but in *unknown* proportions. For each of the eleven rows in Table II, tick *exactly one* of the following boxes: The box "Receive x Yuan for sure", "Indifferent" or "Play Box U".

In each table, if you believe that you are indifferent between the choice in the left column and the right column, you may tick the box under the middle column "Indifferent".

At the end of the experiment, one of the eleven rows of Table I or one of the eleven rows of Table II will be selected at random and played for real money. In Table I, you will receive x Yuan for sure if you tick the box under "Receive x Yuan for sure". However, now if you have ticked "Play Box K", then you shall win 10 *Yuan* if a *non-purple* ball is drawn from Box K (otherwise you win nothing). In Table II, you will receive x Yuan for sure if you tick the box "Receive x Yuan for sure". However, if you have ticked the box "Play Box U" then you win 10 *Yuan* if a *non-purple* ball is drawn from Box U (otherwise you win nothing). In each table, suppose that you tick "Indifferent" in the randomly selected row, then one of the left or right cells in this selected row will be randomly chosen to play for real.

After you have completed Task 3, the experimenter will collect the materials for Task 3 and the experiment will terminate.

16 Appendix B: Post-experimental Questionnaire

1. Age: ____ years old
2. Gender: (female/male)
3. Field of study: ______
4. Degree of study: ______

5. Year of study: ______
6. Have you participated in similar experiments in the past? (Yes/No)
7. Did you have statistics course(s) before? (Yes/No)
8. Does your preference of some particular color(s) affect your decisions?

A. No. B. Yes. Please specify how your preference of some particular color(s) affected your decisions below.