Do People Maximize Quantiles?*

Luciano de Castro† Antonio F. Galvao‡ Charles N. Noussair§ Liang Qiao¶

May 16, 2020

Abstract

Quantiles have been used for decision making in banking and investment (in the form of Value at Risk) and in the mining, oil and gas industries (in the form of "probabilities of exceeding" a certain level of production). However, it is unknown how common quantile-based decision making actually is among individual decision makers. This paper describes an experiment that aims to (1) compare how common quantile decision making is relative to expected utility maximization, and (2) estimate risk attitude parameters under the assumption of quantile preferences. The experiment has two parts. In the first part, individuals make pairwise choices between risky lotteries, and the competing models are fitted to the choice data. In the second part, we directly elicit a decision rule from a menu of alternatives. The results show that the quantile preference model outperforms expected utility for a considerable minority, 30%–50%, of participants, depending on the metric. The majority of individuals are risk averse, and women are more risk averse than men, under both models.

Keywords: Quantile Preference, Risk Attitude, Experiment.

JEL: D81, C91

*The authors would like to express their appreciation to Julian Romero for helpful comments and discussions. We would also like to thank the Economic Science Laboratory at the University of Arizona for funding.

†Department of Economics, University of Iowa, Iowa City, USA. decastro.luciano@gmail.com

‡Department of Economics, University of Arizona, Tucson, USA. agalvao@email.arizona.edu

§Department of Economics, University of Arizona, Tucson, USA. cnoussair@email.arizona.edu

¶Department of Economics, University of Arizona, Tucson, USA. liangqiao@email.arizona.edu

1 Introduction

Expected Utility (EU) maximization, first axiomatized by von Neumann and Morgenstern (1944) and Savage (1954), is the standard decision model assumed in economic and financial theory for the analysis of choices under risk. The model is elegant, general, and amenable to theoretical modeling. The assumption of maximizing average utility, the average being a simple measure of central tendency, has intuitive appeal as a behavioral postulate. Nevertheless, the EU framework has been subjected to a number of criticisms, mostly arising from experimental evidence that has not been supportive. Early studies suggesting that individuals did not always employ objective probabilities, as well as the robustness of the Allais (1953) paradox, resulted in Prospect Theory (Kahneman and Tversky, 1979), Rank-Dependent Expected Utility Theory (Quiggin, 1982), and Cumulative Prospect Theory (Tversky and Kahneman, 1992).[1] Rabin (2000) criticized EU theory on the basis of some implausible predictions it generates, arguing that EU would require unreasonably large levels of risk aversion to explain the data from some small-stakes laboratory experiments.[2] In response to the critiques it has received, the EU model has been successfully generalized to accommodate a variety of behavioral phenomena, a testament to the flexibility and tractability of the EU structure.[3]

In this paper, we are concerned with a different, though also fundamental, potential source of departure from the EU Theory. This is the assumption that individuals seek to maximize an *average* of some function of their payoffs. To us, it is not obvious why a decision maker would necessarily seek to maximize a mathematical expectation of a distribution of either true or perceived payoffs. An alternative is the possibility that individuals attempt to maximize a *quantile* of their payoff distribution. In other words, they maximize the payoff that they have at most a given probability of failing to attain. The simplest example of such a rule is that of maximizing one's median payoff, the 50th percentile. Another example, familiar in risk management, is Value-at-Risk[4] targeting, in which a firm attempts to maximize a low quantile (typically the 1st or 5th percentile) of the payoff from an investment. In oil and gas extraction, executives are frequently concerned with the P90, P50 and P10 indicators of a prospective well, which correspond to the levels of production that are exceeded with 90%, 50% (median) or 10% probability, respectively.[5]

[1]In particular, compelling empirical evidence has been gathered regarding the violation of independence axiom (e.g. Allais, 1953; Starmer and Sugden, 1989; List and Haigh, 2005).

[2]Cox and Sadiraj (2006) argue that Rabin's critique is only valid if participants in experiments fully integrate the payoffs that they earn in the experiment with their overall wealth. That is, Rabin's critique is valid for agents maximizing the expected utility of their overall wealth, but not of their current income. The fact that experimental procedures typically are intended to achieve isolation from participants' experience outside the laboratory makes this counterargument more compelling for experimental research.

[3]Two of the more well-known generalizations are to include the presence of regret (Bell, 1982) and ambiguity in beliefs about probabilities (Gilboa and Schmeidler, 1989).

[4]See, e.g., Duffie and Pan (1997) and Jorion (2007). The VaR measure is one of the main practical tools for reporting the exposure to risk by financial institutions.

[5]See, e.g., Apiwatcharoenkul et al. (2016) and Fanchi and Christiansen (2017).

Manski (1988) was the first to study the properties of quantile preferences (QP), which were later axiomatized by Chambers (2009) and Rostek (2010). In his article, Manski (1988) also discussed a notion of risk aversion for the quantile model, which was further developed by Rostek (2010), and is related to the idea of quantile-preserving spread introduced by Mendelson (1987). The risk attitude is captured by a single-dimensional parameter, the quantile $\tau \in (0, 1)$, with lower quantiles corresponding to more aversion to risk. QP have several attractive features. An individual's decision is independent of the form of her utility function and thus an optimal choice is relatively easy to compute. The measure of risk aversion is simple and intuitive. The structure is robust and flexible, with a family of preferences indexed by quantiles. Recently, models of QP have been attracting increasing attention.[6] However, to our knowledge, no empirical work has considered the incidence of QP in any population.

While the QP model is theoretically appealing and is known to describe decision making in some specific contexts, its usefulness ultimately depends on the breadth of its empirical relevance. This paper presents the results of an experiment designed to study the incidence of QP maximization, relative to EU maximization, among individuals choosing between risky lotteries.

Our experiment has two parts. In the first, participants engage in 242 binary decisions between risky lotteries, and we test whether their pattern of decisions conforms more closely to the EU or the QP model. We also estimate the participants' risk aversion parameters under both models. Under EU, the risk aversion parameter is a measure of curvature of the utility function, while under QP, it is the quantile being maximized. We employ structural estimation for both models and use a maximum likelihood estimator with a parametric functional form for the underlying latent choice models.

In the second portion of the experiment, we directly elicit individuals' intended choice rules. Six alternative decision rules are presented to participants in each period. Three of the rules involve the maximization of EU and three are quantile maximization rules. We ask subjects to choose one of the rules, with the understanding that a bot will employ the rule on their behalf in a series of lottery choice tasks. Our methodology has the advantages that it elicits the *intended* decision rule of the participant, and that it requires no statistical assumptions to estimate the rule being used, as would be necessary with data from choices between lotteries. To our knowledge, this is the first study that directly elicits a decision rule for binary lottery choice tasks.

Two initial remarks are in order regarding our choice of experimental design and the interpretation of our results. The first is that we view pitting the QP model against EU as the natural starting point in evaluating the QP model, since EU is the standard model employed

[6] Bhattacharya (2009) studies the problem of optimally dividing individuals into peer groups to maximize a quantile of social gains from heterogeneous peer effects. Giovannetti (2013) models a two-period economy with one risky and one risk-free asset, where the agent has QP instead of the standard EU. de Castro and Galvao (2019a) develop a dynamic model of rational behavior under uncertainty, in which the agent maximizes a stream of the future quantile utilities.

in economics. If the QP model performs well against EU, it can be tested against other models in follow-up work. The second regards what is a reasonable standard of performance. We do not expect QP to describe every individual's decisions, or even the decisions for a majority of participants, better than EU. However, if the QP model describes even a substantial minority of individuals at least as well as EU, then it should be taken seriously as a descriptive model of behavior. On the other hand, if very few individuals can be classified as QP maximizers, it would suggest that QP maximization is confined to a few special situations, rather than a widely and commonly employed decision model.

We use statistical classification methods to compare the scope of the two models to describe participants' behavior. We classify individuals as users of the QP or EU model based on the models' ability to predict their decisions given their own estimated parameters. We find that a considerable minority of participants, 30%–50% depending on the classification criteria, behave as QP maximizers rather than EU maximizers. There is some weak evidence that women are more likely to be QP maximizers than men.

The individual risk attitude estimates suggest considerable heterogeneity among participants. The average estimated quantile is 0.42, and a majority of subjects are risk averse. Most individuals, about 80%, have quantile estimates between 0.3 and 0.6.[7] For the EU model, we estimate the Power Utility version of the CRRA utility function, given by $u(x) = x^{\gamma}$.[8] Our average estimate is $\gamma = 0.87$, closely in line with the existing experimental literature (see Harrison and Rutström (2008) for a review).[9]

The design of our experiment allows us to consider whether the estimates differ between genders and are affected by a modest change in stakes. We find that women are significantly more risk averse than men on average under the EU model, a result in agreement with much of the previous literature (see Eckel and Grossman, 2008). We also obtain an analogous result for the QP model, in that women maximize significantly lower quantiles than men on average, a pattern consistent with greater risk aversion on the part of women. As the stakes increase, estimates of the CRRA model under EU show greater risk aversion, which is also consistent with previous work (Holt and Laury, 2002; Harrison et al., 2005), while the estimated quantile maximized is not affected by the modest changes in monetary stakes that are present in our experiment.

[7]Given that a parametric specification is used in the estimation, it is important to consider alternative specifications and functional forms, even if the ones we have chosen are standard. Results reported in the Appendix show that varying the scale parameter underlying the logistic likelihood function does not qualitatively affect the results.

[8]See Wakker (2008) for a discussion of the properties of the CRRA utility function

[9]Tversky and Kahneman (1992) and Abdellaoui et al. (2008) estimate the Power utility specification and obtain median estimates of $\gamma = 0.88$ and $\gamma = 0.86$, respectively. These are very close to ours. Other studies estimate a CRRA utility function of the form $u(x) = \frac{x^{1-\rho}}{1-\rho}$. Some of the resulting estimates are $\rho = .61$ (Hey and Orme (1994), sample median), $0.15 - 0.68$ depending on the stakes (Holt and Laury (2002), sample median), $0.26 - 0.54$ (Harrison et al. (2005), average estimate), and 0.89 (Noussair et al. (2014), pooled estimate for representative individual). These are similar to ours in that they reflect a moderate degree of risk aversion, with estimates lying between risk neutrality $\rho = 0$ and $u(x) = \ln(x)$, (the limit as $\rho \to 1$).

The data from the second part of the experiment is also used to classify subjects as QP or EU maximizers, on the basis of their stated preferred decision rule. Here, no statistical method is required for classification, and a mere count of their choices can be used. The results concur with the estimates from the first part of the experiment, and show that a significant number of subjects, in the range of 30–50% depending on how it is measured, are quantile maximizers rather than EU maximizers.

The results illustrate the empirical relevance of QP. A very substantial minority of our participants exhibit behavior that is more consistent with the QP than the EU model. The strength of the empirical support for QP stands in contrast with the scant attention that quantile models have received in decision science. In our view, the incidence of quantile maximization that we have observed justifies the continued exploration of the properties of QP and their implications for economic and financial decision making.

The remainder of the paper is structured as follows. Section 2 reviews the QP model. Section 3 discusses the design and implementation of the experiment. In Section 4, we describe and discuss the main results. Finally, Section 5 concludes.

2 Quantile Model

This section describes the quantile model. In Section 2.1, we review the definition of quantile preferences (QP). Section 2.2 describes the measure of risk attitude under QP. In Section 2.3, we discuss decisions between pairwise lotteries under QP.

2.1 Quantile Preferences

Consider a random variable X, and let F_X (or simply F) denote its cumulative distribution function (CDF), that is, $F_X(\alpha) \equiv \Pr[X \leqslant \alpha]$. The quantile function $Q : [0,1] \to \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ is the generalized inverse of F:

$$Q(\tau) \equiv \begin{cases} \inf\{\alpha \in \mathbb{R} : F(\alpha) \geqslant \tau\}, & \text{if } \tau \in (0,1] \\ \sup\{\alpha \in \mathbb{R} : F(\alpha) = 0\}, & \text{if } \tau = 0. \end{cases}$$

The definition is different for $\tau = 0$, so that the quantile assumes a value in the support of X. It is clear that if F is invertible, that is, if F is strictly increasing, its generalized inverse coincides with the inverse, that is, $Q(\tau) = F^{-1}(\tau)$. Usually, it will be important to highlight the random variable to which the quantile refers. In those cases, we will denote $Q(\tau)$ by $Q_\tau[X]$. For convenience, throughout the paper we will focus on $\tau \in (0,1)$, unless explicitly stated otherwise. A well-known and useful property of quantiles is "invariance" with respect to monotonic transformations, that is, if $g : \mathbb{R} \to \mathbb{R}$ is a continuous and strictly increasing function, then

$$Q_\tau[g(X)] = g\left(Q_\tau[X]\right). \quad (1)$$

To contextualize the difference between expected utility (EU) and QP, let $\mathcal{R}$ denote a set of random variables (lottery payoff distributions). We say that the functional $\varphi : \mathcal{R} \to \mathbb{R}$ *represents the preference* $\succeq$ over $\mathcal{R}$ if for all $X, Y \in \mathcal{R}$ we have:

$$X \succeq Y \iff \varphi(X) \geqslant \varphi(Y). \tag{2}$$

Under von-Neumann and Morgenstern's EU, $\varphi(X) = E[u(X)]$. Under QP, the functional φ in (2) is given by a quantile function, $\varphi(X) = Q_\tau[u(X)]$, so that:[10]

$$X \succeq Y \iff Q_\tau[u(X)] \geqslant Q_\tau[u(Y)].$$

A very important feature of QP is their invariance with respect to the utility function. Let $u(X)$, where $u : \mathbb{R} \to \mathbb{R}$, be an increasing utility function describing an individual's preferences. Then, for a given quantile $\tau \in (0, 1)$, the optimization problem is

$$\max_{X \in \mathcal{R}^*} Q_\tau[u(X)], \tag{3}$$

where $\mathcal{R}^* \subset \mathcal{R}$ is the subset of random variables (lotteries) available.

Given the invariance property described in (1) it can be directly seen that the maximization argument x^* solves (3) if and only if it solves

$$\max_{X \in \mathcal{R}^*} Q_\tau[X]. \tag{4}$$

Equations (3) and (4) show the the quantile optimization problem (3), using a given utility function, is equivalent to maximizing the quantile obtained directly from the distribution of the random variable in (4). Hence, the optimal choice under QP does *not* depend on any particular specification of the utility function.

2.2 Risk Attitudes under Quantile Preferences

In order to discuss risk attitude under QP, we introduce the concept of quantile-preserving spreads. This is analogous to the familiar notion of mean-preserving spreads of Rothschild and Stiglitz (1970) and similarly captures the notion of "added noise." That is, the intuition that Y is equal to X plus noise can be formalized with either the statement that "Y is a mean-preserving spread of X" or that "Y is a quantile-preserving spread of X." The choice of formalization is a subjective matter. In order to discuss this, we follow Mendelson (1987) and define:[11]

[10]Formal axiomatizations of QP can be found in Chambers (2009), Rostek (2010), and de Castro and Galvao (2019b).

[11]Mendelson (1987) formalizes four other conditions and shows that they are all equivalent to Definition 2.1.

Definition 2.1 (Quantile-preserving spread)**.** *We say that* Y *is a* τ*-quantile-preserving spread of* X *for* $\tau \in (0,1)$ *and* $q \in \mathbb{R}$*, if* $Q_\tau[Y] = Q_\tau[X] = q$ *and the following holds:*

(i) $t < q \implies F_Y(t) \geqslant F_X(t)$*;*

(ii) $t > q \implies F_Y(t) \leqslant F_X(t)$*.*

Y *is a quantile-preserving spread of* X *if it is a* $\bar{\tau}$*-quantile-preserving spread of* X *for some* $\bar{\tau} \in (0,1)$*.*

Figure 1 illustrates the CDFs of two random variables, Y and X, when Y is a $\bar{\tau}$-quantile-preserving spread of X. Notice that the definition of spread captures the notion that Y is riskier than X, since it puts greater weight on more extreme values than X. Manski (1988) uses a different terminology for the same notion, referring to the property of "single crossing from below": F_X crosses F_Y once and from below when Y is a quantile-preserving spread of X.

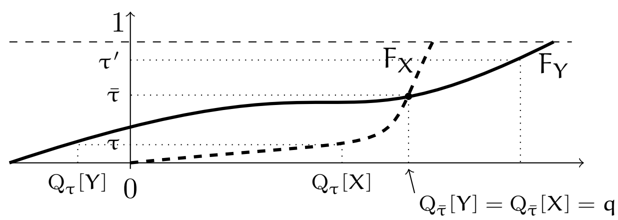


Figure 1: Y is a $\bar{\tau}$-quantile-preserving spread of X.

Note that if $Q_\tau[Y] = q$ and X is equal to q with probability 1, then Y is a τ-quantile-preserving spread of X. In other words, any risky asset Y with τ-quantile q is a quantile-preserving spread of any risk-free asset X with value q.

In addition, Figure 1 suggests that the choice among risky alternatives of a τ-quantile maximizer (a τ-decision maker or τ-DM) depends on whether τ is below or above the quantile $\bar{\tau}$ where the two CDFs cross. That is, on the one hand, when $\tau < \bar{\tau}$ as in Figure 1, a τ-DM prefers the safer asset X: $Q_\tau[X] \geqslant Q_\tau[Y]$. On the other hand, if $\tau > \bar{\tau}$, a τ-DM prefers the riskier asset Y: $Q_\tau[X] \leqslant Q_\tau[Y]$. The following result, first observed by Manski (1988), formalizes this intuition.

Proposition 2.2. *Let* Y *be a* $\bar{\tau}$*-quantile-preserving spread of* X *for* $\bar{\tau} \in (0,1)$*. Then:*

(i) $\tau \leqslant \bar{\tau} \implies Q_\tau[X] \geqslant Q_\tau[Y]$*, that is, a* τ*-DM prefers the less risky asset* X *if* τ *is low;*

(ii) $\tau \geqslant \bar{\tau} \implies Q_\tau[X] \leqslant Q_\tau[Y]$*, that is, a* τ*-DM prefers the riskier asset* Y *if* τ *is high.*

Proof. The proof of this result is in Appendix A. □

We turn now to the problem of comparing the risk attitude of τ-DM with different τ. For this, we consider the definition of "more uncertainty averse than" given by Ghirardato and

Marinacci (2002, Definition 4, p. 263). Following the standard practice, a real number $q \in \mathbb{R}$ will be identified with the random variable that takes value q with probability one and, in this way, a preference can compare random variables and real numbers.

Definition 2.3 (Ghirardato-Marinacci, 2002)**.** *A preference $\succeq'$ is more uncertainty averse than preference $\succeq$ if for any $q \in \mathbb{R}$, and random variable X, $q \succeq X \Rightarrow q \succeq' X$ and $q \succ X \Rightarrow q \succ' X$.*

The intuition for Definition 2.3 is that if a DM with preference $\succeq$ would rather have the certain outcome $q \in \mathbb{R}$ than the risky prospect X, then the more uncertainty averse $\succeq'$ DM prefers it as well. Ghirardato and Marinacci (2002)'s definition is a generalization of the standard notion of risk aversion in the context of risk under expected utility. This notion allows us to provide a suitable characterization for risk attitude for quantile preferences, as t is possible to construct a similar characterization of risk attitude for quantile preferences, as the following result establishes.

Proposition 2.4. *Consider quantile maximizing preferences $\succeq_\tau$ and $\succeq_{\tau'}$. The following statements are equivalent:*

1. $\tau \geqslant \tau'$;
2. $\succeq_{\tau'}$ *is more uncertainty averse than* $\succeq_\tau$;
3. *If* Y *is a quantile-preserving spread of* X *and* $X \succ_\tau Y$, *then* $X \succeq_{\tau'} Y$.[12]
4. *If* Y *is a quantile-preserving spread of* X *and* $Y \succ_{\tau'} X$, *then* $Y \succeq_\tau X$.

Proof. The proof of this result is in Appendix A.[13] □

This result shows that $\succeq^\tau$ is more risk averse than $\succeq^{\tau'}$ if and only if $\tau < \tau'$. This property implies that an agent with a quantile given by τ_1 is more risk preferring than another agent with quantile given by τ_2 if $\tau_1 > \tau_2$, independently of the functional form of utility. Thus, a decision maker that maximizes a lower quantile is more risk averse than one who maximizes a higher quantile. In other words, the risk attitude is defined by the quantile rather than by the concavity of the utility function.

2.3 Decision Between Pairwise Risky Lotteries under Quantile Preferences

In this section, we describe the decision of an agent maximizing a given quantile τ when making a pairwise choice between two risky lotteries.

[12] Notice that we are not specifying what is the quantile $\bar{\tau}$ for which Y is a $\bar{\tau}$-quantile-preserving spread of X. The same observation is valid for the other item.

[13] Elements of Proposition 2.4 can be found in Rostek (2010, Section 6.1) and Manski (1988, Section 5), but not in the form presented here. In particular, they do not use the language of quantile-preserving spreads introduced by Mendelson (1987) nor the notion of "more uncertainty averse than" used by Ghirardato and Marinacci (2002).

Within the class of QP models, an agent's choices are determined by the characteristic quantile. Recall from equations (3) and (4) that QP are invariant with respect to the utility function. Consider a choice between two lotteries A and B. The decision of a QP maximizer is based on evaluating the difference of the quantiles of the two lotteries. If we understand A and B as random variables, this difference is:

$$\Delta Q(\tau) = Q_\tau[B] - Q_\tau[A].$$

The agent maximizing a given quantile τ prefers B to A when, for that τ, the quantile of lottery B is larger than A. That is,

$$\Delta Q(\tau) = Q_\tau[B] - Q_\tau[A] > 0.$$

To illustrate this, we consider an example of a decision taken when employing QP with two lotteries, A and B, each have two possible payoffs, $\{a_1, a_2, b_1, b_2\}$ and corresponding outcome probabilities p, q, so that:

- Lottery A yields $a_1 = 6$ with probability p, and $a_2 = 10$ with probability $1 - p$.
- Lottery B yields $b_1 = 2$ with probability q, and $b_2 = 16$ with probability $1 - q$.

This lottery can be represented graphically, as illustrated in Figure 2.

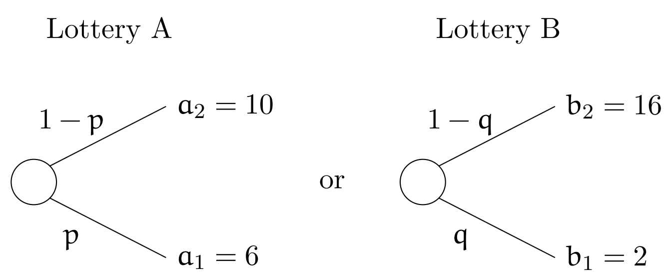


Figure 2: Lottery Choice in Task L1

The calculation of $\Delta Q(\tau) = Q_\tau[B] - Q_\tau[A]$ depends on the quantile τ, the payoffs $\{a_1, a_2, b_1, b_2\}$, and the probabilities (p, q). Specifically,

$$\Delta Q(\tau) = \begin{cases} b_1 - a_1, & \text{if } \tau \leqslant \min\{p, q\} \\ b_1 - a_2, & \text{if } p < \tau \leqslant q \\ b_2 - a_1, & \text{if } q < \tau \leqslant p \\ b_2 - a_2, & \text{if } \tau > \max\{p, q\}. \end{cases} \tag{5}$$

Specializing the example further, fix the quantile at the median, $\tau = 0.5$, and the outcome probabilities at $p = 0.3$ and $q = 0.1$. The quantile functions of lotteries A and B are plotted in the left panel of Figure 3. Lottery A (solid line) pays $a_1 = 6$ with probability 0.3 and $a_2 = 10$ with probability 0.7, and lottery B (dashed line) pays $b_1 = 2$ with probability 0.1, and $b_2 = 16$ with probability 0.9. The solid vertical line at 0.5 represents the quantile of interest. To complete the example of an agent maximizing the median, $\tau = 0.5$, and choosing between lotteries A and B, we compute $\Delta Q(\tau)$. The calculation is simple and only requires one to subtract the quantile of A from that of B. From the left panel in Figure 3, we can see that $\Delta Q(0.5) = Q_{0.5}[B] - Q_{0.5}[A] = 16 - 10 = 6$. Therefore, the agent chooses lottery B.

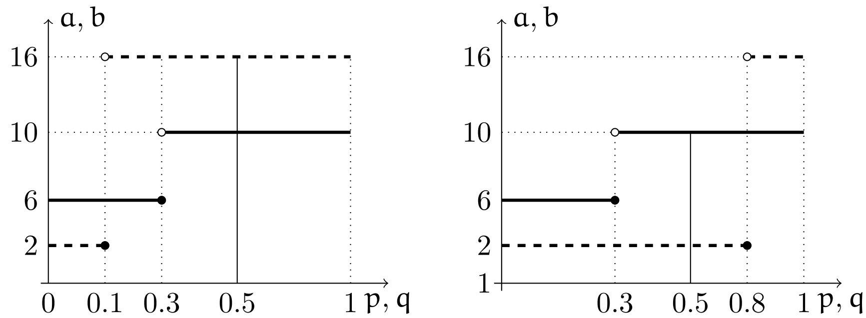


Figure 3: Quantile function of lotteries A (solid line) and B (dashed line). The left plot considers $p = 0.3$ and $q = 0.1$. The right plot considers $p = 0.3$ and $q = 0.8$

Suppose now that we slightly modify the lotteries by changing the probability q of lottery B, so that we have:

- Lottery A: $a_1 = 6$ with probability $p = 0.3$, and $a_2 = 10$ with probability $1 - p = 0.7$;
- Lottery B: $b_1 = 2$ with probability $q = 0.8$, and $b_2 = 16$ with probability $1 - q = 0.2$.

The quantile functions of the new A and B lotteries are displayed in the right panel of Figure 3. Lottery A (solid line) pays $a_1 = 6$ with probability 0.3 and $a_2 = 10$ with probability 0.7, and lottery B (dashed line) pays $b_1 = 2$ with probability 0.8, and $b_2 = 16$ with probability 0.2. In this case, we can see that the calculation of $\Delta Q(0.5)$ for the median (solid vertical line) is $\Delta Q(0.5) = Q_{0.5}[B] - Q_{0.5}[A] = 2 - 10 = -8$, and hence, the agent chooses lottery A.

3 The Experiment

The experiment consisted of 8 sessions conducted at the Economic Science Laboratory at the University of Arizona, in February and March, 2020. All procedures received prior approval from the Institutional Review Board at the University. The 61 participants in the experiment were all undergraduate students at the university, who were recruited from a subject pool maintained by the Laboratory. 52.5% of the subjects were male, and 50.8% of them were majoring in Business or Economics. The average cognitive reflection test (CRT) score (see

Frederick, 2005) was 1.1 out of a maximum of 3. The experiment was computerized and used the ZTree platform.

A session consisted of ten periods. Periods 1 and 2 were referred to as Part 1 of the session. At the outset of the session, the instructions for these two periods were read and the two periods took place. Afterward, the instructions for Part 2 of the session, consisting of Periods 3–10 were read. Participants did not receive any feedback regarding the outcome of their choices until the session ended. The subjects then completed periods 3–10. The instructions are available in Appendix D.

3.1 Periods 1 and 2

In each one of periods 1 and 2, subjects made 121 separate decisions between two lotteries: called A and B. For ease of notation, we denote the tasks used in the two periods of Part 1 of the session as L1 and L2. The payoffs in the two tasks are given in Table 1. In the table, L3–L10 refer to the payoffs in period 3–10, which will be described in Section 3.2. The payoffs from lottery A were $a_1 = 6$ and $a_2 = 10$ in both L1 and L2. However, those from lottery B were $b_1 = 2$ and $b_2 = 16$ in L1, and $b_1 = 2$ and $b_2 = 24$ in L2. In each of the 121 decisions of a period, the participant made a choice between Option A, which yielded a_1 with probability p and a_2 with probability $1 - p$, and Option B, which yielded b_1 with probability q and b_2 with probability $1-q$. The probabilities of the relatively low payoffs under each lottery, p and q, took on values from the set {0,0.1,...,0.9,1}. Each of the 11 possible values of p was paired with each possible value of q in exactly one decision, generating $11 \times 11 = 121$ decisions in each of the first two periods.

To control for any order effect, the sequence of the pairwise choice lottery was counterbalanced. That is, we assigned the order of the lottery pairs L1 and L2 in the first two periods to participants on a random basis, with about half of subjects receiving each order.[14] The 121 decisions within a period were presented in random order on 11 separate computer screens, with 11 decisions displayed on each screen.

As described in Section 2.3, we are able to identify the quantile of interest by keeping the payoffs $\{a_1, a_2, b_1, b_2\}$ constant and varying the associated probabilities (p, q). In addition, the difference in the high payment b_2 in lottery B between the two periods allows us to look at a specific comparative static. This is the effect of changing exactly one of the possible outcomes. If individuals use the same quantile independently of the magnitude of payoffs involved, decisions would be identical in the two periods. However, it is known that on average individuals tend to become more risk averse as the monetary stakes of the gambles they are confronted with increase. An analogous pattern, less risk taking behavior as stakes increase, for quantile decision makers, would mean that lower quantiles would be used at higher stakes.

[14]As a result, 33 out of 61 participants received L1 in period 1 and L2 in period 2. The remaining 28 participants saw L2 in period 1 and L1 in period 2.

Table 1: Lottery Payoffs For All Periods.

Task (payoff vector)	Lottery A a_1(\$)	a_2(\$)	Lottery B b_1(\$)	b_2(\$)	Description
L1	6	10	2	16	Used in period 1 or 2
L2	6	10	2	24	Used in period 1 or 2
L3	6	10	2	16	Same as L1
L4	6	10	2	24	Same as L2, L3's $b_2 + 8$
L5	3	5	1	12	50% of L4
L6	8	10	6	17	L5+5
L7	8	10	0.5	19.25	HL lotteries
L8	8	10	0.5	9.25	L7's $b_2 - 10$
L9	12	15	0.7	28.75	150% of L7
L10	13	15	5.5	24.25	L7+5

3.2 Periods 3–10

In periods 3–10, participants were presented with the payoffs $\{a_1, a_2, b_1, b_2\}$, indicated as L3–L10 in Table 1. In each period, they were offered a menu of six decision rules, and were required to specify a rule from the menu for a computer robot to use on their behalf. The computer then made 121 decisions of the form that participants made in periods 1 and 2, based on the rule participants specified.

The six decision rules were the following:

- Maximize average earnings (EN)
- Maximize the average square root of earnings (EA)
- Maximize the average square of earnings (ES)
- Maximize the amount of money that you have at least a 50% chance of earning (Q50)
- Maximize the amount of money that you have at least a 25% chance of earning (Q75)
- Maximize the amount of money that you have at least a 75% chance of earning (Q25)

EN, EA and ES are rules implementing the expected utility (EU) model. Specifically, EN orders the robot to make those choices that maximize the expected value of a risk-neutral utility function $u(x) = x$. EA commands the bot to maximize the expectation of a risk-averse utility function $u(x) = \sqrt{x}$, and ES does so for a risk-seeking function $u(x) = x^2$.

Q50, Q25 and Q75 specify three different quantile preferences (QP) rules. Q50 orders the robot to maximize the participant's median payoff. Q25 and Q75 instruct the bot to maximize the 75th and 25th percentiles, respectively.[15]

To overcome possible order effects in the way the rules were presented, we used two different versions of the instructions. In the instructions that one-half of the participants received, the EU rules were described first, and for the other half, the QP rules were described first. The position of the six rules on participants' computer screens was randomly reshuffled in each period. This was done in an independent random sequence for each individual.

The potential lottery payoffs in periods 3–10 are indicated as L3–L10 in Table 1. The order in which each payoff vector appeared to subjects was generated randomly and independently, without replacement, for each participant. L3 and L4 are the same lotteries used in periods 1 and 2. The payoffs in L5 are equal to 50% of those in L4. L6 is derived by adding a constant of \$5 to all payoffs in L5. L7 is proportional, in terms of payoffs, to those in the lotteries employed in the classic study of Holt and Laury (2002). L8 is similar to L7 except that b_2 is \$10 lower. L9 has the payoffs of L7 scaled up by 50 percent, and L10 adds a constant of \$5 to all payoff levels of L7.

We illustrate theoretical decisions taken under the EN, EA, ES, Q50, Q25 and Q75 rules in Figure 4. The shaded area of each individual panel indicates the combinations of probabilities p and q for which lottery B is chosen in L1, according to the decision rule depicted. We can see from the figure that, as risk tolerance under either QP and EU increases, the choice of lottery B becomes more likely. Nevertheless, the choices taken under the two models do exhibit substantial differences.

To describe the rules in a manner that participants could be expected to understand, we chose our phrasing very carefully. The instructions are available in the Appendix. To describe a rule that maximized $u(x) = \sqrt{x}$, we wrote *"This rule will take each possible amount of money you could earn, and give it a point score equal to the square root of the payment. It then makes a choice to maximize the average number of points you get. In other words, it calculates the average square root of the amount of money that you would receive under each choice, and always decides for the choice that has a higher average square root of the payment."* Similar language was used to describe maximization of $u(x) = x^2$.

Describing a quantile maximization rule is also not straightforward. For example, we chose the following language to represent a rule where $\tau = 0.25$: *This rule will consider each choice and ask the question: "How much money do I have at least a 75% chance of earning?" It will then decide on the choice that has a higher payment in answer to this question.* The responses on the quiz administered following the instructions indicate that the task was quite well understood.

[15]The decision made under the Q50 rule is slightly different from the one maximizing $\tau = 0.5$, because of the discontinuity of probabilities in the experiment. More precisely, Q50 can be viewed as maximizing the $0.5 + \epsilon$ quantile, or $\tau = 0.5001$. We choose this wording to describe the quantile rules to facilitate the participant's comprehension.

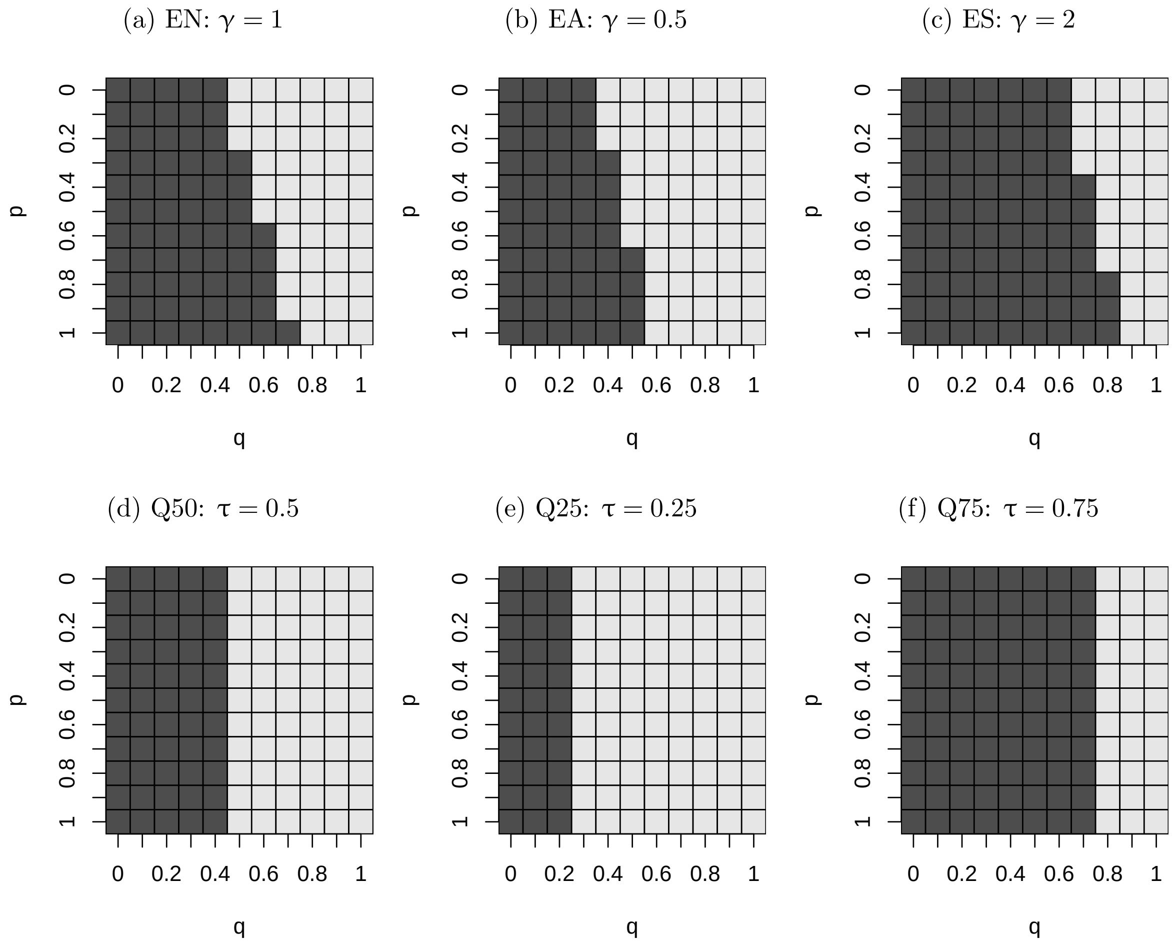


Figure 4: Choices of lottery B over lottery A in L1 based on the decision rule specified. (a) EN with $\gamma = 1$, (b) EA with $\gamma = 0.5$, (c) ES with $\gamma = 2$, (d) Q50 with $\tau = 0.5$, (e) Q25 with $\tau = 0.25$, (f) Q75 with $\tau = 0.75$. The darker color represents choosing lottery B for the cell, while the lighter color indicates choosing lottery A.

3.3 Payment

The final payment to participants equaled the sum of the $5 show-up fee and the earnings from one of the ten periods in the session. The period that counted toward the final payment was determined randomly at the end of the experiment by the computer. In the period that counted, one of the 121 decisions in the period was randomly chosen to count toward earnings.

The instructions for periods 1 and 2 were given at the beginning of the experiment. After participants finished periods 1 and 2, the instructions for periods 3–10 were provided, and were followed by a quiz to check participants' understanding of the decision rules introduced in the new instructions. The experimenter checked the quiz and helped participants to correct wrong answers privately. Then, periods 3–10 were conducted, in which participants selected decision rules. At the end of the session, participants filled in a brief questionnaire and were paid privately. The average payment was US$16.6.

4 Results

In this section we report our findings. The section is organized in the following manner. Section 4.1 reports the risk attitude parameter estimates for both the quantile preferences (QP) and expected utility (EU) models using the data from the first two periods. In Section 4.2, we use statistical classification methods to consider the percentage of individuals that adhere more closely to each model. In Section 4.3 we study how model choice and parameters correlate with gender and the payoffs of the lotteries. Finally, Section 4.4 presents the results from periods 3–10 of the experiment. There, we classify participants by their intended choice of decision rule rather than estimating a rule from their lottery choice decisions.

4.1 Risk Attitude Parameters

4.1.1 Summary Statistics of Periods 1 and 2

In each of the first two periods, every participant makes 121 decision between lotteries A and B. Figure 5 provides heat maps that summarize the observed choices. It shows the proportion of the participants choosing lottery B over A in each decision as a function of the probabilities p and q. The figure divides the data into five proportion intervals: $[0, 0.2)$, $[0.2, 0.4)$, $[0.4, 0.6)$, $[0.6, 0.8)$, $[0.8, 1]$, which are represented with five different color shades. A darker color corresponds to a higher proportion of choices of B. The left panel of the figure shows the combined results for periods 1 and 2. The middle and right panels display separate results for tasks 1 and 2 (L1 and L2), respectively.

A comparison of the observed choice results in the left panel of Figure 5 with the theoretical patterns in Figure 4 give the impression that the representative individual's decisions lie somewhere between the EU and QP models, and display considerable aversion to risk. A contrast between the middle and right panels in Figure 5 shows that there are more choices of lottery B in L2, where b_2 is larger, than in L1. This is consistent with substitution toward lottery B when it has a greater expected value. As we shall see later, however, the difference is smaller than would be predicted if the coefficient of risk aversion γ were the same in the two tasks, and the estimated coefficient is smaller for L2 than for L1, indicating more risk aversion for L2. The relatively low responsiveness to the change in b_2 is also consistent with a fraction of participants employing the QP model, which predicts no difference in choices between the two tasks.

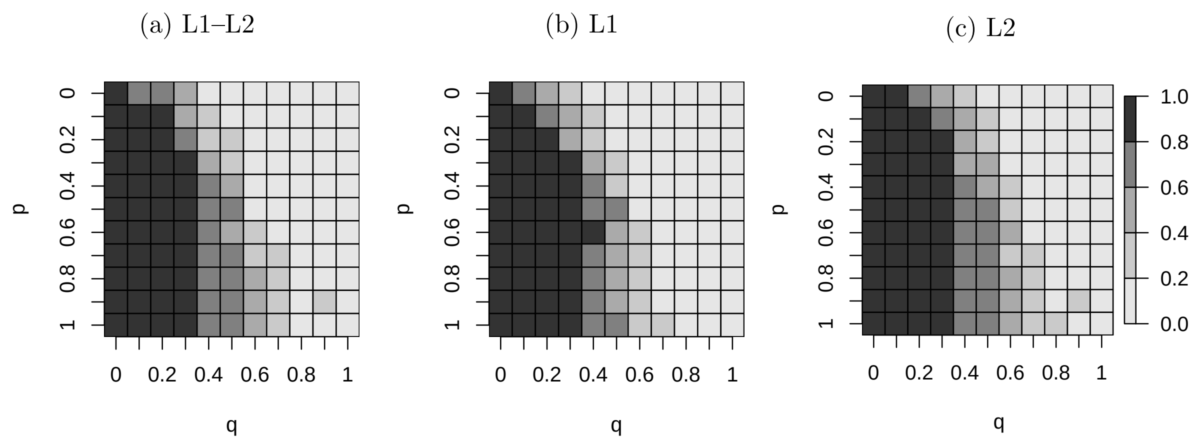


Figure 5: Heat maps for observed choices of lottery B over lottery A. The left panel contains the pooled data for periods 1 and 2. The middle panel shows the data for L1, where $b_2 = \$16$, and the right panel the data for L2, where $b_2 = \$24$.

4.1.2 Estimation Method

We estimate the risk aversion parameters for both the QP and EU models using structural estimation. Under the assumption of EU, a utility function must be specified. We assume constant relative risk aversion (CRRA), with the specific power functional form,

$$u(x) = x^{\gamma}. \tag{6}$$

The variable x is the monetary payoff and γ is the risk aversion parameter to be estimated. We choose the CRRA functional form for two reasons. The first is that it is the most widely employed in economics and this facilitates a comparison of our estimates with others in the literature. The second is that it has one parameter, and as such, can be readily compared in terms of fit to the QP model without necessitating a correction for a different number of parameters in the two models.[16] We follow the literature (see, e.g., Moffatt (2016) and references therein) and use a Maximum Likelihood Estimator (MLE) to estimate the parameter γ.

With regard to the risk attitude parameter for the QP model, the quantile τ, recall from Section 2.1 that QP are invariant with respect to the utility function. Hence, it is not necessary to specify any particular parametric functional form of utility to estimate the quantile. We adopt the following structural estimation strategy for the QP model using MLE.

Let $x_{l,k}$ be the monetary payoff associated with the probability $p_{l,k}$ of the outcome k for

[16] One popular alternative is the expo-power utility function, $u(x) = \theta - \exp\{-\beta x^{\alpha}\}$, introduced by Saha (1993). This utility function has two free parameters (when θ is fixed as a constant), and allows testing of whether decreasing absolute risk aversion and increasing relative risk aversion are present. While this form can typically provide a better fit than CRRA, it would have to be compared to a quantile rule with two parameters, perhaps with one parameter indicating how the quantile being maximized varies by income.

lottery l. The payoff $x_{l,k}$ and probabilities $p_{l,k}$ are induced by the experimenter, so that the cumulative distribution function for each lottery l is given by:

$$F_l(x_{l,k}) = P_l(X_l \leqslant x_{l,k}),$$

for outcomes $k = 1, ..., K$. In our experiments, $k = 2$, and there are two lotteries, so that $l = \{A, B\}$.

To compare the lotteries, we employ the framework described in Section 2.3 above. In particular, we specify two lotteries: B, the "high risk" lottery, associated with CDF F_B, and A, the "low risk" lottery, with CDF F_A. We use the quantile preserving spread concept, and assume that F_B crosses F_A from below at the quantile τ^*, making lottery B riskier than A, as in Figure 1.

Under the QP model, for each lottery pair and τ^* (crossing point of the CDFs), the choice between lotteries A_i and B_i is based on the difference of quantiles that is calculated for a candidate estimate of τ by evaluating the differences index

$$\Delta Q_i(\tau) \equiv Q_\tau[B_i] - Q_\tau[A_i]. \tag{7}$$

Equation (7) indicates that for a fixed quantile τ, an agent chooses B_i over A_i when $Q_\tau[B_i] > Q_\tau[A_i]$. The latent index, based on latent preferences, is then linked to the observed choices using a specified CDF, denoted by $G(\cdot)$. In particular, we follow the literature and use a latent variable model to derive the likelihood function for the quantile model.

We model the latent (unobserved) variable y_i^* for task i, and assume that the observed binary response variable y_i is generated, as

$$y_i^* = \Delta Q_i(\tau) + \varepsilon_i$$
$$y_i = \begin{cases} 1, & \text{if } y_i^* > 0 \text{ (choose lottery B)} \\ 0, & \text{otherwise.} \end{cases}$$

Let the CDF of ε_i be denoted by G. The CDF G, or link function, takes any argument on the real line, in this case the difference $\Delta Q(\tau)$, and transforms it into a number between 0 and 1. Thus, the probability of choosing lottery B, Prob(choose lottery B) $= \mathsf{Prob}(y_i = 1)$, depending on the link function, can be calculated. The model can be written as

$$\mathsf{Prob}(y_i = 1) = P(\varepsilon_i > -\Delta Q_i(\tau)) = 1 - G(-\Delta Q_i(\tau))$$
$$\mathsf{Prob}(y_i = 0) = G(-\Delta Q_i(\tau)).$$

To write the likelihood function, suppose we have a random sample (y_i, w_i) for tasks

$i = 1, ..., n$, with $w_i = \Delta Q_i(\tau)$. For a symmetric link function G, the likelihood is

$$\mathcal{L}(y_i, w_i; \tau) = \prod_{i:y=0} (1 - G(\Delta Q_i(\tau))) \prod_{i:y=1} G(\Delta Q_i(\tau))$$
$$= \prod_{i=1}^{n} G_i^{y_i} (1 - G_i)^{1-y_i} .$$

Then, the log-likelihood function is given by

$$\ell(y_i, w_i; \tau) = \sum_{i=1}^{n} \left[y_i \log G(\Delta Q_i(\tau)) + (1 - y_i) \log(1 - G(\Delta Q_i(\tau)))\right]. \tag{8}$$

The MLE of τ is simply the value of τ that maximizes the log-likelihood function in equation (8). The CDF function $G(\cdot)$ can be specified as, for instance, the Logit (logistic distribution), the Probit (normal distribution), or another distribution.

We remark that, in contrast to the EU case, where the log-likelihood is smooth and differentiable everywhere, the log-likelihood function in equation (8) is not smooth. Nevertheless, there is an existing literature in econometrics establishing the asymptotic properties – consistency, asymptotic normality, and bootstrap inference – for this class of semiparametric estimators (as the MLE), where the criterion function does not obey standard smoothness conditions. The theories allow for non-smooth objective functions of finite-dimensional unknown parameters (e.g., Pakes and Pollard (1989) and Newey and McFadden (1994, Section 7)) and both finite-dimensional and infinite- dimensional parameters (e.g., Chen et al. (2003)). In addition, Chen et al. (2003) show that bootstrapping for these methods provides asymptotically correct confidence regions for finite-dimensional parameters. Throughout this paper we apply bootstrap procedures, with 1,000 bootstrap repetitions, to compute the standard errors of the parameters of interest.

4.1.3 Estimates of Risk Attitude Under Both Models

We specify a Logit model for the estimation of both the QP and EU models, so that the link function is a logistic distribution with location parameter zero and scale two. We choose this scale parameter since we have payoffs ranging from \$2 to \$24, implying a relatively wide range of QP and EU difference indices. A scale that is too small makes the likelihood function insensitive to large values in the index. Nevertheless, in Appendix B we present results for different scale parameters and discuss their effect on the results.

We start by estimating the models for each participant separately. Figure 6 presents a histogram of the distribution of quantile estimates $\widehat{\tau}$ (risk attitude) from tasks 1–2 together (242 decisions). The figure shows that most subjects have an estimated risk attitude below 0.5, indicating that they are more risk averse than a maximizer of the median. About 80%

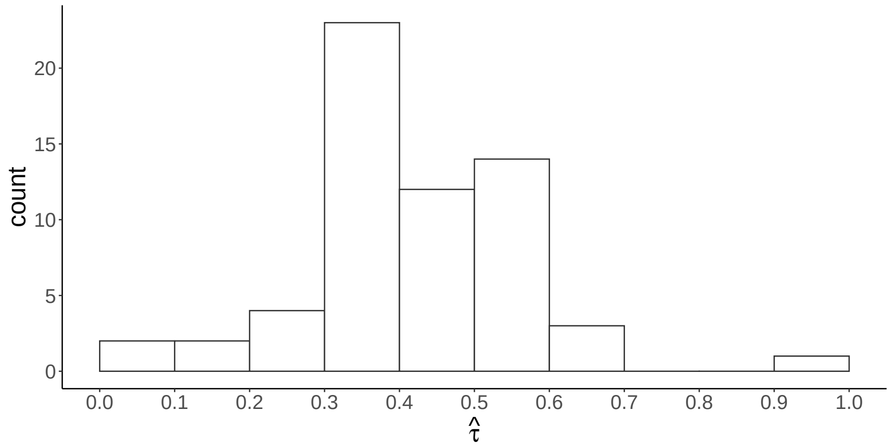


Figure 6: Histogram of individual $\widehat{\tau}$ estimates, all participants (mean=0.42; standard error=0.02).

of subjects have quantile estimates between 0.3 and 0.6. The average, over all the quantile estimates, is 0.42, and the sample standard error is 0.02. As Figure 6 demonstrates, the distribution is skewed to the right, with relatively few estimates exhibiting risk attitude above $\widehat{\tau} = 0.6$.

Figure 7 shows the distribution of CRRA coefficient $\widehat{\gamma}$ estimates in (6) using the analogous Logit MLE for the data from tasks 1–2. The estimates are in line with the existing literature. The average estimate of $\widehat{\gamma}$ is 0.87, with a standard error of 0.02. Moreover, about 79% of subjects present $\widehat{\gamma}$ smaller than 0.975, suggesting that a large majority are risk averse. About 11% of subjects have $\widehat{\gamma}$ larger than 1.025 and thus can be classified as risk seeking. The remaining 10% of $\widehat{\gamma}$ is in the interval of $[0.975, 1.025]$ which is essentially risk neutral. The proportions that are risk averse, risk seeking and risk neutral are comparable to those in previous studies. This suggests that the scale parameter in our estimation is appropriately calibrated for comparison to previous work.

We also compute pooled risk attitude parameter estimates for both QP and EU models, treating the entire sample as one "representative" participant. For risk aversion under the EU model, we obtain a point estimate of $\widehat{\gamma} = 0.85$ with a standard error of 0.02. For the QP model the point estimate is $\widehat{\tau} = 0.35$ with a standard error of 0.01. The standard error is computed by applying the bootstrap-based cluster robust standard errors of Cameron et al. (2008), with the clustering at the level of the individual participant. In other words, assuming that individuals have CRRA utility of the power utility form, the group of 61 subjects behaves like a representative person with a risk aversion coefficient of 0.85. Analogously, under the assumption of QP maximization, the "representative" subject acts as if she is maximizing the 35th percentile of her earnings.

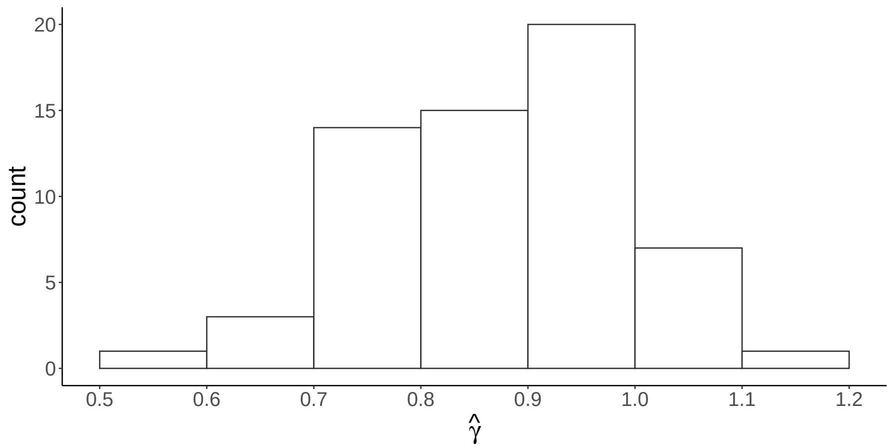


Figure 7: Histogram of individual γ estimates, all participants (mean=0.87; standard error=0.02).

4.2 Comparison of QP and EU Models

We now consider the relative performance of the two models, by classifying participants as QP or EU maximizers, based on the model that better describes their decisions. To accomplish this, we compute the confusion matrix (see, e.g., James et al. (2017)) using the Logit model. A confusion matrix compares the predictions of the model with the actual choices for all the observations in the data set. The first step in computing the matrix is to calculate the predictions for all observations for each subject. This is implemented by computing the probability $P(y_i = 1)$. In our case, choice 1, $y = 1$, is choosing lottery B. The alternative choice, $y = 0$, is choosing lottery A. Once the coefficients have been estimated, it is easy to compute the predicted probability of choosing lottery B for any given difference in payoffs. For example, one computes $\widehat{P}(y_i = 1)$ for every decision i. Nevertheless, since this predicted probability is a number between zero and one, to classify an observation as choice 1, we have to specify a threshold. For instance, if 0.5 is the threshold, then for $\widehat{P}(y_i = 1) \geqslant 0.5$ we assign observation i to class 1, otherwise observation i is assigned to class 0. Finally, these predictions can be compared to the actual choices, and the percentage of correct predictions, the hit rate, can be computed.

By constructing the confusion matrix and computing the hit rate of both the QP and EU models for each subject's data, we can classify individuals into three categories: EU maximizer, QP maximizer, and tie. The methodology to classify a subject into a group is as follows. First, for each individual, we compute the confusion matrices using both the EU and the QP models. Second, for each individual, we compute the hit rate (number of correct classifications) for both EU and QP. Third, we compare the hit rates for both models. If the hit rate over all of the decisions is higher for the QP model than the EU model, the subject is classified as a QP maximizer. When the contrary occurs, the individual is a EU maximizer. A tie happens when the hit rates are equal for the two models. The results for the classification appear in Table 2 for three different choices of the threshold parameter: 0.4, 0.5, and 0.6. The table shows that a majority of subjects are classified as EU in all three cases. However, a significant proportion of the subjects, between 31% (19 out of 61) and 41% (25 out of 61) are classified as QP. This result is a clear indication of the relevance of the QP model.

Table 2: Number of participants classified as QP Maximizers, EU Maximizers, and Ties, classification based on the hit rate of models' predictions.

Threshold probability	# QP Maximizers	# EU Maximizers	# Ties
$p = 0.40$	21	38	2
$p = 0.50$	19	40	2
$p = 0.60$	25	35	1

We also compare the two models using both parametric and non-parametric within-subject tests for the threshold probability $p = 0.5$. We first subtract the hit rate of the QP from the EU model for each individual to compute a within-subject measure of relative performance of the two models. The sample average for the difference is -0.009 with a standard error of 0.006. We then test the null hypothesis that the average difference is equal to zero. The corresponding t-test statistic is -1.45 (p-value 0.15), and we cannot reject the null hypothesis, at standard levels of significance, that the average differences in the hit rates of the QP and EU models is statistically equal to zero. The sample median for the difference is -0.017. We test the null hypothesis that this median is equal to zero. A Wilcoxon signed rank test yields a test statistic equal to 666.5 (p-value 0.10). Moreover, a paired-sample sign test of the null hypothesis that the difference is equally likely to be positive or negative yields a p-value of 0.01. Hence, there is some evidence that the EU model describes decisions better than the QP model for a significant majority of participants.[17]

4.3 Gender and Stakes Effects

In this section we consider gender differences and the effect of stakes on the risk attitude estimates and the relative performance of the two models.[18] As discussed earlier, a substantial fraction of prior studies have found that on average, women are more risk averse than men (see, e.g., Eckel and Grossman (2008) for a review). If a similar pattern exists for QP maximization, one might expect women on average to have lower τ estimates than men.

[17]For the threshold probabilities of $p = 0.4$ and $p = 0.6$, the results for the average and median tests reach the same conclusions as for the $p = 0.5$ case. The results for the paired-sample sign tests have p-values of 0.04 and 0.25, for $p = 0.4$ and $p = 0.6$, respectively.

[18]In Appendix C, we present the results of regressions of QP and EU estimates and classification on gender, program of study and CRT score.

Figures 8 and 9 present histograms of the distribution of the risk attitude parameter estimates for the EU and QP models, for females and males separately. Figure 8 reveals a gender difference in the $\widehat{\gamma}$ estimates. The top histogram shows that overall, females have smaller risk aversion coefficients, and thus are more risk averse, compared to males. The average estimate for γ is 0.83 for women and 0.91 for men. The t-statistic for the equality of average γ for males and females is $t = 2.81$, which rejects the null of equality at the 1% level. The tests are one-sided, since there is prior evidence that women are more risk-averse than men on average. In Figure 9, we observe an analogous pattern for the QP model. The quantile estimates are more concentrated on small values for females. A larger proportion of females, relative to males, has an estimated $\widehat{\tau} < 0.3$, while the opposite is observed for $\widehat{\tau} > 0.6$. The average estimate of τ is 0.38 for women and 0.45 for men. The t-statistic for the null hypothesis of the equality of average τ for males and females is $t = 1.87$, so that we reject the null of equality at the 5% level.[19]

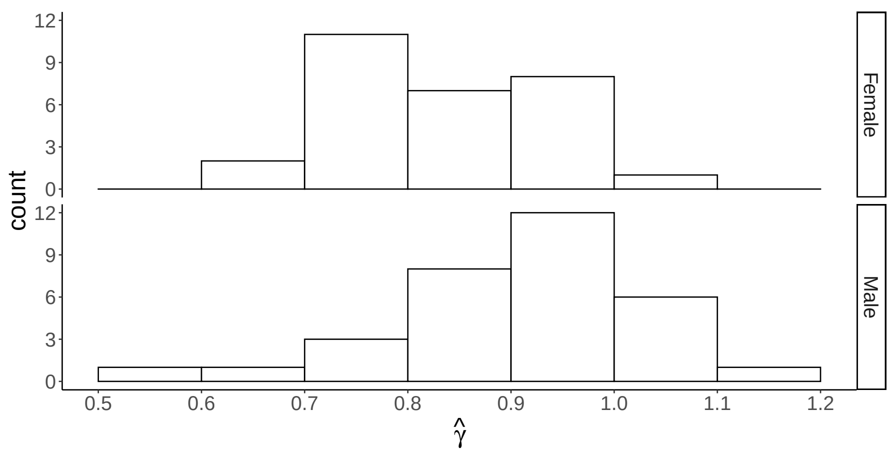


Figure 8: Histograms of individual γ estimates by gender. Female mean= 0.83 and standard error=0.02, male mean=0.91 and standard error=0.02.

[19]We can also conduct tests for each of the two decision problems separately. For the QP model, the t-statistic for the difference in means between males and females, when $b_2 = 16$, is $t = 1.69$. The t-statistic for the difference in means between males and females for $b_2 = 24$ is $t = 1.98$. Moreover, for the EU model, the t-statistic for the difference in means between males and females when $b_2 = 16$ is $t = 2.33$, and for $b_2 = 24$ it is $t = 2.84$. Thus, for both QP and EU models, we reject the null hypothesis that the risk aversion parameter is the same for males and females in all cases at the 5% significance level.

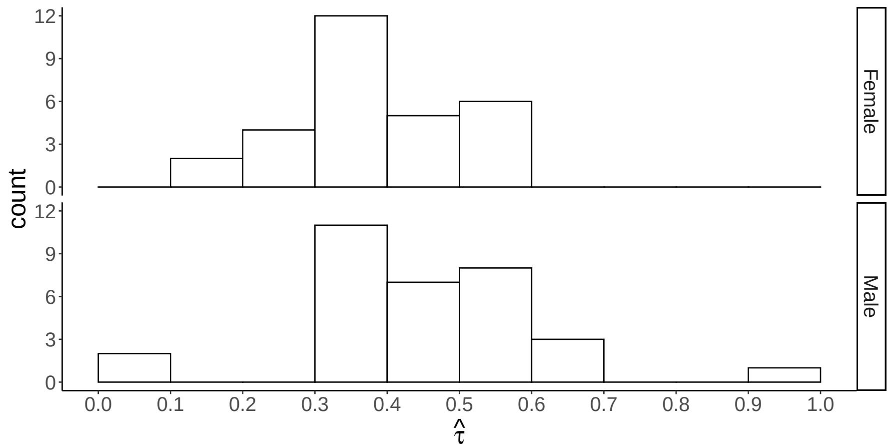


Figure 9: Histograms of individual τ estimates by gender. Female mean=0.38 and standard error=0.02, male mean=0.45 and standard error=0.03.

We now turn our attention to the stakes of the lotteries. Recall that rounds 1 and 2 differ only in the high payoff attainable under lottery B. Hence, they allow us to study the effect of a modest change in monetary stakes. Some prior evidence suggests that many individuals take less risk at higher stakes.[20] If this is the case, then the estimated γ parameters would be lower for task 2 (in which $b_2 = 24$) than task 1 (in which $b_2 = 16$). If a similar pattern of more "conservative" decisions at higher stakes appears under QP, then one would find individuals maximizing lower quantiles at higher stakes. However, the quantile model presented in Section 2 is robust in that changing the b_2 payoff in the manner we have done has no effect on the lottery chosen.

Figures 10 and 11 display the distribution of the risk attitude parameter estimates for each of the two tasks, by gender, for both the QP and EU models. The left (right) hand side plots the estimates for high (low) stakes. Consider first the quantile estimates in Figure 10. The figure shows that the average quantile maximized does not differ for the two stake levels. Of course, the changes in stakes here are modest and perhaps effects would appear for much larger changes in the magnitudes of payoffs.

Figure 11 collects the corresponding results for the EU model. The modal bins of $\widehat{\gamma}$ at $b_2 = 24$ and $b_2 = 16$ for females are 0.8 and 0.9. Those two corresponding modal categories for males are 0.9 and 1.1. This figure corroborates the existing results in the literature (see, e.g., Holt and Laury (2002), and Kachelmeier and Shehata (1992)) that individuals are more risk averse when the lottery stakes are higher.[21]

[20] See for example Holt and Laury (2002) or Harrison et al. (2005).

[21] One might ask how it could be the case that individuals are more risk averse at higher stakes under the EU model, but not under the quantile model. One feature of our two tasks is that the only difference is that $b_1 \neq b_2$. Consider an individual whose decision is not responsive to the value of b, so that her decisions are the same in the two tasks. Her estimated quantile would not differ in the two tasks. However, her estimated γ would change, since for a greater value of b, she would appear to have a stronger tendency to choose a lower-risk lottery with lower expected value.

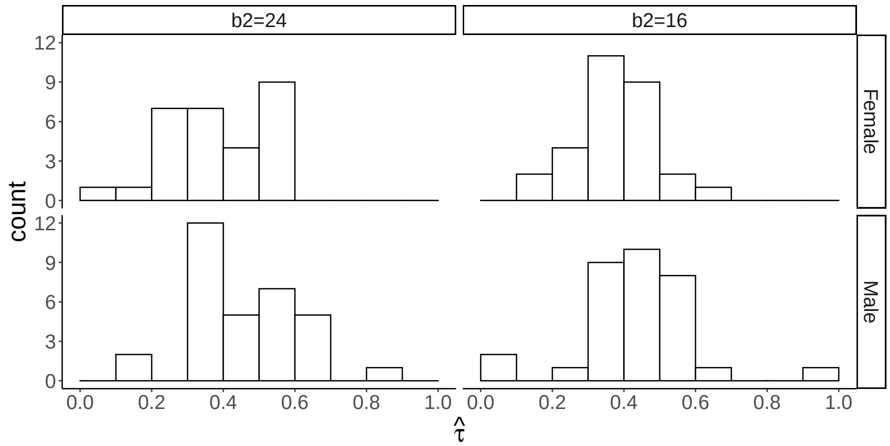


Figure 10: Histograms of individual τ estimates by gender and task. Group means of females at $b_2 = 24$ and $b_2 = 16$ are 0.38 (standard error=0.03) and 0.38 (standard error=0.02), respectively. Group means of males at $b_2 = 24$ and $b_2 = 16$ are 0.46 (standard error=0.03) and 0.44 (standard error=0.03), respectively.

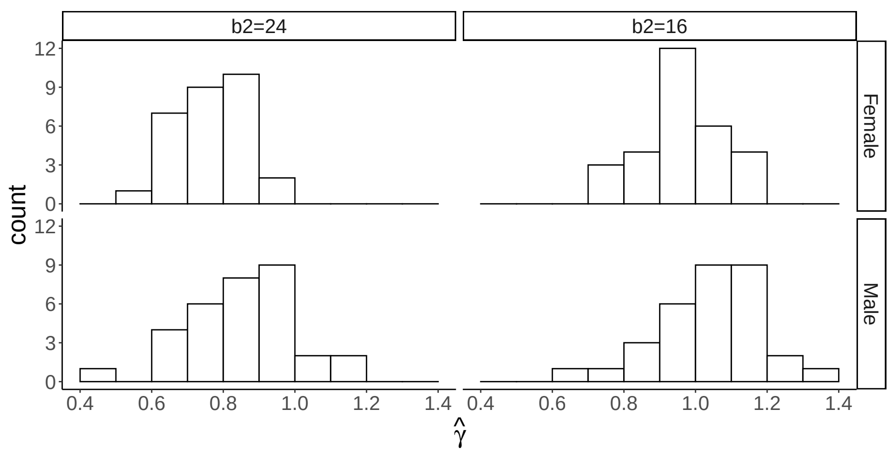


Figure 11: Histograms of individual γ estimates by gender and task. Group means of females at $b_2 = 24$ and $b_2 = 16$ are 0.77 (standard error=0.02) and 0.96 (standard error=0.02), respectively. Group means of males at $b_2 = 24$ and $b_2 = 16$ are 0.86 (standard error=0.02) and 1.04 (standard error=0.03), respectively.

We conduct statistical tests to confirm the patterns suggested from the figures. Consider first the QP model. We test for equality of mean τ estimates between low and high stakes, for males and females separately. For males, the t-statistic for the difference in means between $b_2 = 24$ and $b_2 = 16$ is $t = 0.55$. For females, the comparable t-statistic is $t = 0.20$. Thus, we are not able to reject the null hypothesis that the mean quantiles for both low and high stakes are equal at standard levels of significance, for both males and females. We apply the same tests for the γ estimates. For males, the t-statistic for the difference in means between $b_2 = 24$ and $b_2 = 16$ is $t = -5.11$. For females, the t-statistic for the difference in means between $b_2 = 24$ and $b_2 = 16$ is $t = -6.80$. Therefore, we are able to reject the null of mean equality at 1%, in favor of the alternative that γ is lower at the greater payoff level for each gender.

We now consider whether there are gender differences in the percentage of participants adhering to each model. The results are given in Table 3, for different threshold probabilities. The table shows that the QP model fits well for a greater percentage of females than males. We formally test whether the proportion of adherents are female is the same in the QP and EU groups. The test statistics for threshold probabilities $p = 0.4$ and $p = 0.5$ are -0.76 and -1.29, respectively. Hence, we are not able to reject the null hypothesis that the gender ratio is equal in QP and EU groups. However, the test statistic is -2.27 (p-value < 0.03) for the threshold probability $p = 0.6$. Therefore, there is some weak evidence that the percentage of females adhering to QP is greater than that of males.

Table 3: Number of subjects who are QP Maximizers, EU Maximizers, and Ties, by gender. Classification based on the hit rate.

Threshold probability	# QP Maximizers		# EU Maximizers		# Ties	
	Male	Female	Male	Female	Male	Female
$p = 0.40$	10	11	22	16	0	2
$p = 0.50$	8	11	24	16	0	2
$p = 0.60$	9	16	23	12	0	1

4.4 What Decision Model Do Individuals Intend to Use?

In this section we analyze the data from periods 3–10. Recall from Section 3 that in these periods we directly ask subjects to choose the decision rule they wish to employ from six available alternatives. The goal is to directly measure the intended rule choice of participants. One very important advantage of this new methodology is that there is no requirement of a statistical model to analyze the data. We measure the decisions subjects make directly.

4.4.1 Summary Statistics for Periods 3 to 10

As discussed in Section 3.2 the rules 1, 2, and 3 correspond to decisions within the EU model, whereas rules 4, 5, and 6 are implementations of the QP model. Figure 12 plots the histogram of the decision rule choices of all participants. There are two prominent patterns in the figure. The first is that there is a comparable level of choices of EU and QP rules. The second is that maximizing expected value (EN) and median payoff (Q50) are the most commonly used rules within the EU and QP classes, respectively. These two rules capture risk neutrality in their respective models.

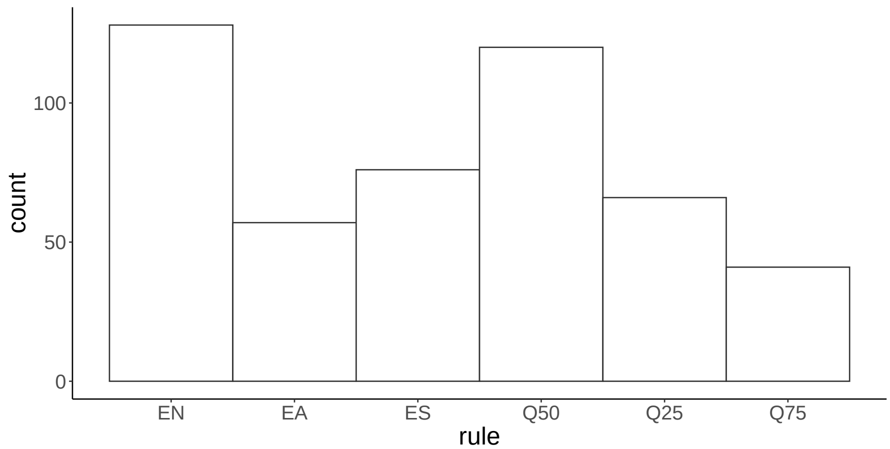


Figure 12: Histogram of observed rule choices in L3–L10, all participants. The horizontal labels are as follows: EN risk neutral; EA risk averse; ES risk seeking; Q50 the median; Q25 the 25th percentile; Q75 the 75th percentile.

Figure 13 displays a histogram of the observed rule choices for females and males separately. The figure shows that more women choose QP rules than men, while men are more likely to opt for the risk-seeking EU rule than are women. We reject the hypothesis that proportion of females in the QP group is equal to the proportion in the EU group ($t = -1.83, p < .1$). Thus, we obtain some more evidence that women tend to use quantile rules more than men. Within the subset of those individuals who choose EU rules, we reject the hypothesis that there is an equal proportion of men and women opting for the ES rule ($t = -1.98, p < .05$).

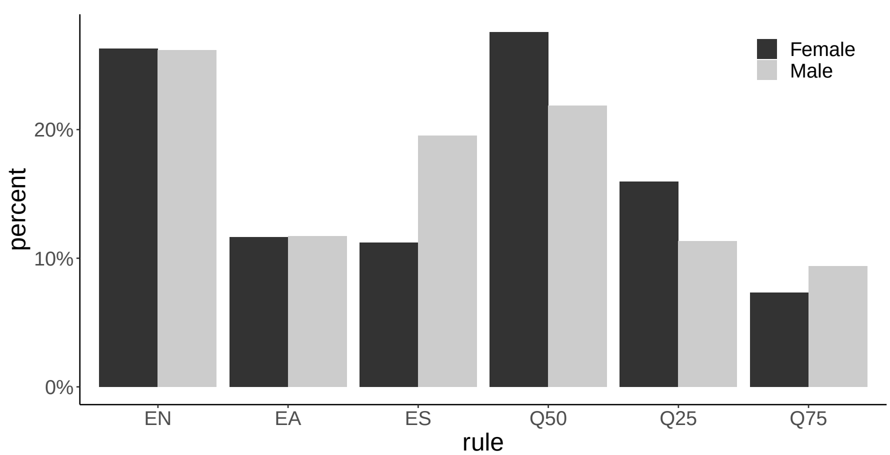


Figure 13: Histogram of observed rule choices in L3–L10 by gender. The horizontal labels are as follows: EN risk neutral; EA risk averse; ES risk seeking; Q50 the median; Q25 the 25th percentile; Q75 the 75th percentile.

Figure 14 depicts the decision rule choices for each of the eight periods L3–L10, separately. Each panel concerns a different payoff vector for each period, and the relationships among

L3–L10 are summarized in Table 1 in Section 3.2 above. The figure shows similar patterns as those in Figure 12. There is a comparable incidence of choices of EU and QP rules, and the most popular rules are EN and Q50. The Q50 rule is the most commonly used in L3, L4, L6, and L9, while EN is most common in L5, L8, and L10. L5 and L8 are the lotteries with the lowest stakes. In addition, task L8 is the only one for which the quantile-preserving spread relationship does not hold for lotteries $\mathbb{A}$ and $\mathbb{B}$. The percentages choosing QP rules in the eight tasks L3–L10 are, respectively, 48%, 46%, 38%, 51%, 46%, 38%, 51%, and 56%.

4.4.2 Classification

We classify subjects as QP or EU maximizers based on the decision rules they choose. There is no need to use an underlying statistical model, and instead we employ a simple count of their choices. We adopt three different criteria to assign participants into three categories: QP maximizer, EU maximizer, and mixture. To satisfy the less strict criterion, a subject has to choose a rule from one model in at least 5 of the 8 total periods. If the criterion is not satisfied (in cases where an individual chooses EU and QP rules in four instances each), the person is classified as employing a mixture. The stricter criteria require a subject to be consistent with one specific model in at least 6 and 7 periods, respectively.

The results are reported in Table 4. The lower part of the table displays the data for each gender separately. The table shows that while a plurality of subjects have the intention to maximize expected utility, a considerable minority choose to maximize a quantile function. Applying the less strict threshold ($\geqslant 5$) indicates that about 31% of the total number of subjects choose a quantile rule, while about 48% of the subjects intend to maximize EU. When we tighten the criterion to $\geqslant 7$, the number of individuals classified as using the QP and EU models becomes very similar, with 10 (16%) and 12 (20%) employing each model, respectively. Overall, these results are similar to those in Table 2 and corroborate the evidence that a relatively large minority of individuals' economic decisions are more consistent with quantile, than with expected utility, maximization.

Table 4: Number of subjects classified as QP Maximizers, EU Maximizers, and Mixture, by gender. Classification based on rules selected in tasks L3–L10.

Number choices	# QP Maximizers		# EU Maximizers		# Mixture	
$\geqslant 5$	19		29		13	
$\geqslant 6$	13		17		31	
$\geqslant 7$	10		12		39	
	Male	Female	Male	Female	Male	Female
$\geqslant 5$	10	9	18	11	4	9
$\geqslant 6$	6	7	11	6	15	16
$\geqslant 7$	3	7	8	4	21	18

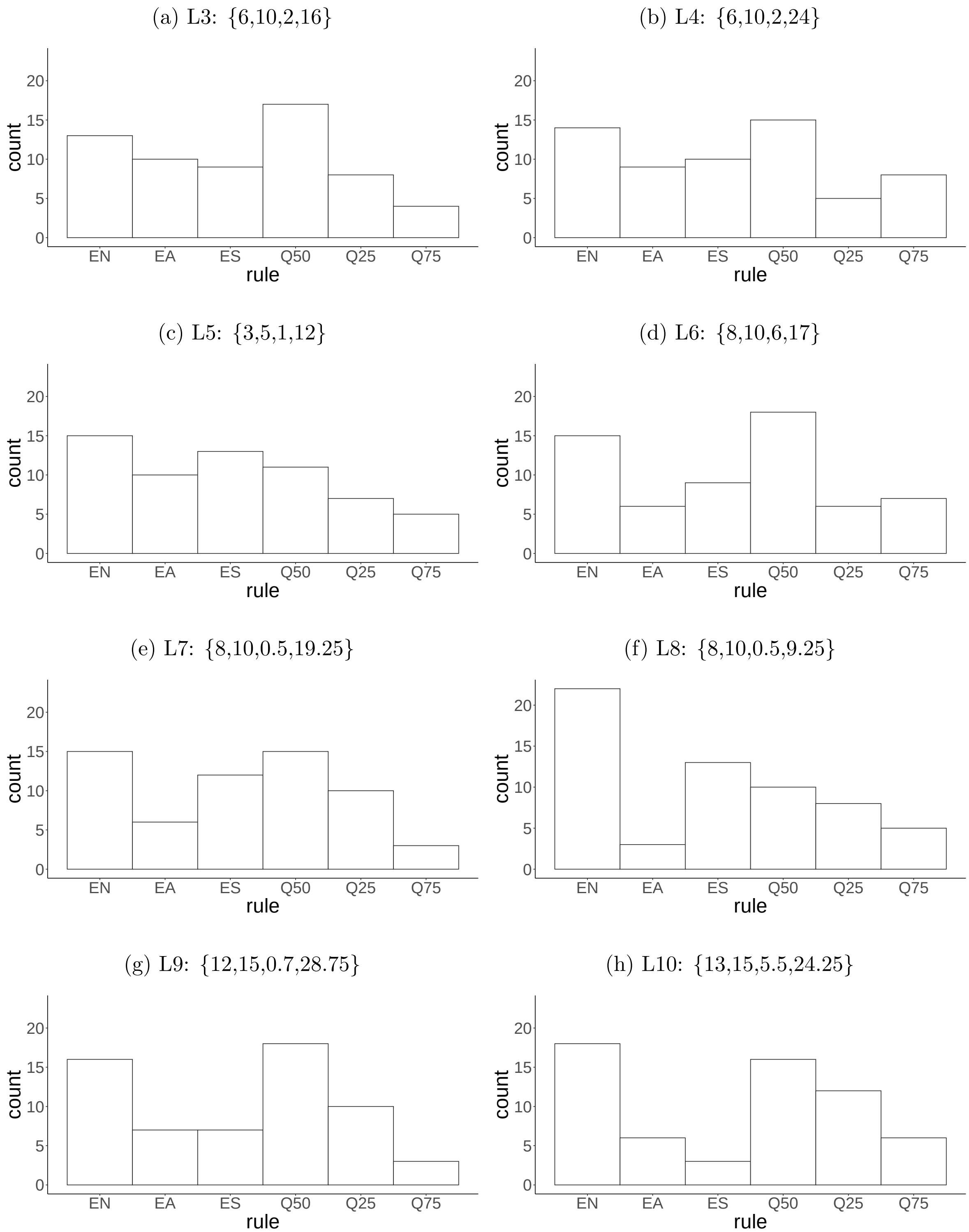


Figure 14: Histograms of observed rule choices in L3–L10. The horizontal labels are as follows: EN risk neutral; EA risk averse; ES risk seeking; Q50 the median; Q25 the 25th percentile; Q75 the 75th percentile. The numbers in curly brackets indicate the payoff vector $\{\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{b}_1, \mathfrak{b}_2\}$.

The results by gender are given in the lower part of Table 4. They show that for those adhering to the QP model, the number of males and females are close to equal for the less strict category. For the stricter category, there is a greater number of females than males. We conduct a formal test of the null hypothesis of equality of proportions of males and females in the QP and EU models. When applying the $\geqslant 5$ criterion, yields a two-sample test statistic is -0.65, such that we are not able to reject the null hypothesis that the models are identical in terms of the proportion of adherents who are female. If apply the $\geqslant 6$ criterion, the test statistic is -1.02. However, when using the $\geqslant 7$ criterion, the test statistic is equal to -1.71, sufficient to conclude at the 10% level of significance that males are less likely to choose QP maximization, despite the low power to reject equality under this strict criterion for what constitutes an observation. This constitutes additional weak evidence that women have a greater tendency to employ the QP model than men.

Table 5: Number of participants classified as QP Maximizers, EU Maximizers, and Mixture categories, all participants pooled and by gender. Classification is based on rules selected in tasks L3–L10 (excluding L8).

Number choices	# QP Maximizers		# EU Maximizers		# Mixture	
$\geqslant 4$	27		34		–	
$\geqslant 5$	16		20		25	
$\geqslant 6$	12		12		37	
	Male	Female	Male	Female	Male	Female
$\geqslant 4$	13	14	19	15	–	–
$\geqslant 5$	8	8	12	8	12	13
$\geqslant 6$	5	7	8	4	19	18

As discussed in the previous section, the quantile-preserving spread relationship does not hold between lotteries A and B in task L8. Hence, we recompute the results in Table 4 excluding period L8. In this case, we have a total of seven tasks and use 4, 5, and 6 as the number of choices consistent with a model as classification criteria. The results are given in Table 5, with the lower part separating the results for gender. When comparing the results in Tables 4 and 5 we see that the number of QP maximizers increases substantially when L8 is removed, and about 44% of the total number of subjects are classified as QP maximizers for the least stringent $\geqslant 4$ rule, while an equal number are classified as EU and QP under the strictest $\geqslant 6$ standard (12 employing each model). The patterns regarding gender in Table 5 are similar to those in Table 4, with females having a stronger tendency to employ QP rules than males. We conduct a two-sample test to assess the equality of proportions of gender between QP maximizers and EU maximizers. The two-sample test statistics are both -0.60 when applying the $\geqslant 4$ and $\geqslant 5$ criteria, and -1.23 under the $\geqslant 6$ criterion. Thus, we are not able to reject the null hypothesis of no gender difference in the use of the QP and EU models in any of the three cases.

5 Conclusion

In this paper, we compared the performance of the standard decision model in economics, expected utility (EU), with a model of quantile preference (QP) maximization. As the EU model requires a choice of utility function, we have assumed the CRRA form, since it is also standard. We considered (1) the fit of the two models to participants' binary lottery decisions, and (2) the percentage of the time agents chose to employ each model from description. The EU model provides a better fit for a small majority of decision makers, while the QP model outperforms for a considerable minority, roughly 30%–50% of our sample, depending on the classification criterion. In our view, the large fraction using quantile rules justifies continued exploration of the properties of quantile models of decision making and their application to economic contexts.[22]

Of course, this study is only a first step in the evaluation of the model. One might argue that the CRRA functional form is not appropriate to employ with EU, and that the mediocre performance of the EU model is due to the assumption of CRRA. Perhaps a utility function with more parameters that is known to provide a better fit to laboratory data, such as expo-power utility, might perform better. However, it is also not obvious how one could describe expo-power utility in terms that participants could understand, as would be necessary for part 2 of our experiment. Expo-power utility has two parameters, and thus is less parsimonious than CRRA. Unless we correct for the different numbers of parameters, it would have to be evaluated against an extension of the quantile model with an additional parameter, perhaps one that would allow the quantile to vary depending on a measure of the stakes of the lottery payoffs. Indeed, one of the advantages of the QP model is that decisions do not depend on the utility function.

In addition to pitting the quantile model against different functional forms under EU, it would be interesting to evaluate it against Cumulative Prospect Theory, or other models that allow probability weighting. Again, however, it is unclear how a protocol like our part 2 could be used, because probability weighting would have to be explained to individuals who are choosing between decision rules. A model with probability weighting and a utility curvature parameter would presumably provide a better fit to the data than EU, which is a special case of the model. However, it would be less parsimonious, with more free parameters, than the QP model.

[22] We do not argue that a model with a single τ parameter would apply at any level of monetary stakes. Clearly, if we were to consider a case where $b_2 = \$1,000,000$ we presumably would not observe many individuals maximizing the same quantile as they would at $b_2 = \$10$. However, the same criticism would also apply to CRRA preferences under EU, whose parameter estimates for given individuals are known to differ considerably as the stakes of gambles change.

Appendix

We include four appendices. Appendix A contains the proofs of the propositions in Section 2. In Appendix B, we consider the implications of changing the scale of the Logit specification used to estimate the risk attitude parameters and to classify participants by model for part 1 of the experiment. Appendix C reports the results of regressions that test for correlates of risk attitude estimates under each of the two decision models. Finally, Appendix D contains the instructions for the experiment.

A Proofs

This appendix collects the proofs for the results in the main text. We will use two properties of quantiles that are easy to verify. First, quantiles are non-decreasing, that is, $\tau \leqslant \bar{\tau}$ implies $Q_\tau[X] \leqslant Q_{\bar{\tau}}[X]$. Second, for any random variable X with CDF F_X,

$$F_X(t) \geqslant \tau \iff Q_\tau[X] \leqslant t. \tag{9}$$

For reader's convenience, we provide below a detailed proof of Proposition 2.2, but the result can also be found in Manski (1988, Proposition 3, p. 95).

Proof of Proposition 2.2

Proof. Let Y be a $\bar{\tau}$-quantile-preserving spread of X and let $q = Q_{\bar{\tau}}[Y] = Q_{\bar{\tau}}[X]$, so that

$$t < q \implies F_Y(t) \geqslant F_X(t); \text{ and} \tag{10}$$

$$t > q \implies F_Y(t) \leqslant F_X(t). \tag{11}$$

To show (i), assume for a contradiction that $\tau \leqslant \bar{\tau}$ and $t \equiv Q_\tau[X] < Q_\tau[Y]$. Since $Q_\tau[Y] > t$, by (9), $F_Y(t) < \tau$. Since $Q_\tau[X] = t$ implies $F_X(t) \geqslant \tau$ again by (9), we conclude that

$$F_Y(t) < \tau \leqslant F_X(t). \tag{12}$$

Since quantiles are non-decreasing, we have $t < Q_\tau[Y] \leqslant Q_{\bar{\tau}}[Y] = q$. But then (12) contradicts (10). The contradiction establishes (i).

Similarly to show (ii) with a contradiction, assume that $\tau \geqslant \bar{\tau}$ and $t \equiv Q_\tau[Y] < Q_\tau[X]$. Since $Q_\tau[X] > t$, by (9), $F_X(t) < \tau$. Since $Q_\tau[Y] = t$ implies $F_Y(t) \geqslant \tau$ again by (9), we conclude that

$$F_X(t) < \tau \leqslant F_Y(t). \tag{13}$$

Since quantiles are non-decreasing, we have $t > Q_\tau[Y] \geqslant Q_{\bar{\tau}}[Y] = q$. But then (13) contradicts (11). The contradiction establishes (ii). □

Proof of Proposition 2.4:

Proof. (1) $\Rightarrow$ (4) : Let $\tau \geqslant \tau'$ and Y be a $\bar{\tau}$-quantile-preserving spread of X and $Y \succ_{\tau'} X \Leftrightarrow Q_{\tau'}[Y] > Q_{\tau'}[X]$. By Proposition 2.2(i), $\tau' > \bar{\tau}$, which implies $\tau > \bar{\tau}$. By Proposition 2.2(ii), $Q_\tau[X] \leqslant Q_\tau[Y] \Rightarrow Y \succeq_\tau X$.

(4) $\Rightarrow$ (3) : Assume that Y be a $\bar{\tau}$-quantile-preserving spread of X and $X \succ_\tau Y$. For a contradiction, assume that $\neg (X \succeq_{\tau'} Y) \Leftrightarrow Y \succ_{\tau'} X$. By (4), this implies that $Y \succeq_\tau X$, which contradicts $X \succ_\tau Y$.

(3) $\Rightarrow$ (1) : For a contradiction, assume that $\tau < \tau'$ and let Y be a $\bar{\tau}$-quantile-preserving spread of X satisfying

$$Q_{\hat{\tau}}[X] = Q_{\hat{\tau}}[Y] \Rightarrow \hat{\tau} = \bar{\tau}, \tag{14}$$

for some fixed $\bar{\tau} \in (\tau, \tau')$, that is, the quantile functions of Y and X only coincide at $\bar{\tau}$. Since $\tau < \bar{\tau}$, Proposition 2.2(i) implies that $Q_\tau[X] \geqslant Q_\tau[Y]$, which must be $Q_\tau[X] > Q_\tau[Y]$ because of (14). By (3), we must have $Q_{\tau'}[X] \geqslant Q_{\tau'}[Y]$. Since $\tau' > \bar{\tau}$, by Proposition 2.2(ii), $Q_{\tau'}[X] \leqslant Q_{\tau'}[Y]$. Therefore, $Q_{\tau'}[X] = Q_{\tau'}[Y]$, which contradicts (14) since $\tau' > \bar{\tau}$.

(1) $\Rightarrow$ (2) : Let $\tau \geqslant \tau'$. Since quantiles are monotonic, $Q_\tau[X] \geqslant Q_{\tau'}[X]$. Therefore,

$$q \succeq_\tau X \iff q \geqslant Q_\tau[X] \implies q \geqslant Q_{\tau'}[X] \iff q \succeq_{\tau'} X; \text{ and}$$
$$q \succ_\tau X \iff q > Q_\tau[X] \implies q > Q_{\tau'}[X] \iff q \succ_{\tau'} X.$$

(2) $\Rightarrow$ (1) : Assume that $\succeq_{\tau'}$ is more uncertainty averse than $\succeq_\tau$ and, for a contradiction, that $\tau < \tau'$. By monotonicity, $Q_\tau[X] \leqslant Q_{\tau'}[X]$ for any X. Let X be such that $Q_\tau[X] < Q_{\tau'}[X]$ and $q \equiv Q_\tau[X] \Rightarrow q \succeq_\tau X$. Since $\succeq_{\tau'}$ is more uncertainty averse than $\succeq_\tau$, this implies $q \succeq_{\tau'} X \Leftrightarrow q = Q_\tau[X] \geqslant Q_{\tau'}[X]$, which contradicts $Q_\tau[X] < Q_{\tau'}[X]$. □

B Robustness to Scaling of Logit Model

In this appendix, we report the risk attitude estimates for the QP and EU models from part 1 of the experiment, and compare their relative performance, using a likelihood function modified from that in Section 4.1.3, to have a different scaling. Specifically, we report results from the estimation of a Logit with scales 1 and 3, rather than 2, as reported in the main text. Although not presented here, results using Probit models for the link function are very similar to those in corresponding Logit cases.

Figure 15 presents histograms of the individual estimates of the risk attitude τ in the QP model when using Logit with scales 1 and 3 as the link function in the log-likelihood. The plots show that the histogram for scale 1 is very close to the results in the main text for scale 2. For the case of scale 3, the distribution slightly shifts to the left. Nevertheless, the average of all estimates and sample standard errors for all the three distributions (scales 1, 2, and 3) are very similar, in particular, the averages are 0.41, 0.42, and 0.43 for scales 1, 2, and 3 respectively, and in all three cases the standard error is equal to 0.02. In addition, the modes are 0.35 for each of the three distributions. Overall, the results in Figure 15 show that the risk attitude parameter τ estimates for the quantile model are not very sensitive to changes in the specification of the underlying likelihood function.

(a) Logit with scale 1 (b) Logit with scale 3

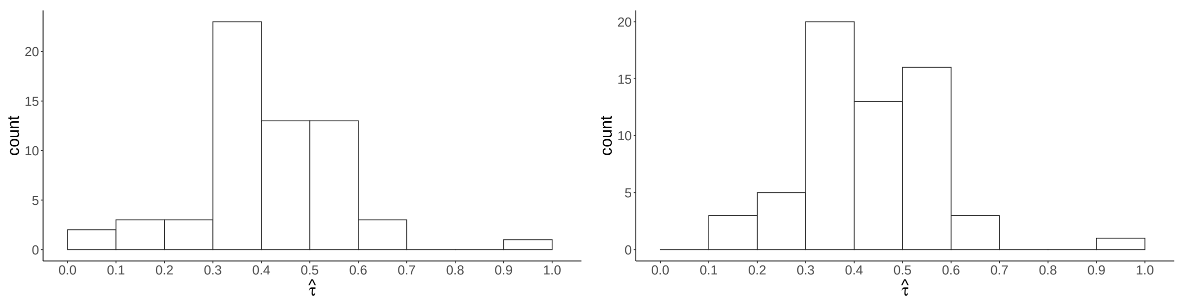


Figure 15: Histograms of individual τ estimates. The left panel is from a Logit with scale 1 (mean=0.41; standard error =0.02). The right panel is from a Logit with scale 3 (mean=0.43; standard error =0.02).

Figure 16 presents histograms of the individual estimates of the risk attitude γ in the EU model when using Logit with scales 1 and 3 as the link function in the log-likelihood. We can see from the figures that as the scale increases from 1 to 3, the distribution shifts to the right. In particular, for scale 1, the average over the individual estimates of $\widehat{\gamma}$ is 0.73, for scale 2 (reported in the text) the average is 0.87, and for scale 3 it is 0.97. The modal bins of the estimates $\widehat{\gamma}$ in Logit models with scale 1, scale 2 and scale 3 are those around 0.6, 0.9, and 1.0, respectively. These results show that the risk aversion parameter under EU/CRRA is sensitive to the specification of the underlying likelihood function.

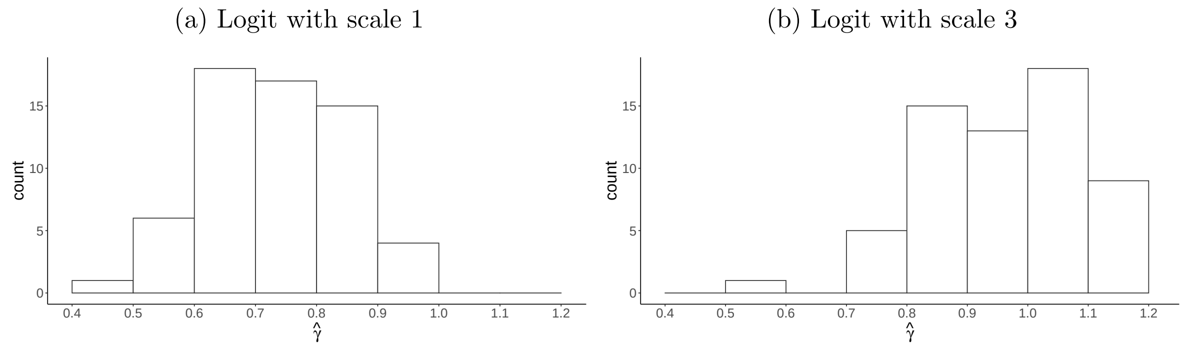


Figure 16: Histograms of individual γ estimates. The left panel contains estimates from a Logit with scale 1 (mean=0.73; standard error =0.01). The right panel are estimates from a Logit with scale 3 (mean=0.97; standard error =0.02).

We now turn to the classification of individuals by decision model when assuming the Logit specification with the different scales. Tables 6 and 7 collect the results for the Logit with scales 1 and 3, respectively. The results are similar to those in Table 2 in the main text. Table 6 indicates that the number of QP maximizers is between 31% and 36% of the total number of participants. When we increase the scale parameter of the logistic distribution, this number increases, as can be seen in Table 7. In particular, for the case $p = 0.60$ in Table 7, the results show that 27 subjects (44%) are QP maximizers and 32 (52%) are EU maximizers. These results are in line with those in the main text and are strong evidence of the importance of the quantile preferences.

Table 6: Number of subjects who are QP Maximizers, EU Maximizers, and Ties, all participants. Classification based on hit rate with Logit scale 1.

Threshold probability	# QP Maximizers	# EU Maximizers	# ties
$p = 0.40$	19	41	1
$p = 0.50$	19	42	0
$p = 0.60$	22	38	1

Table 7: Number of subjects who are QP Maximizers, EU Maximizers, and Ties, all participants. Classification based on hit rate with Logit scale 3.

Threshold probability	# QP Maximizers	# EU Maximizers	# ties
p = 0.40	25	34	2
p = 0.50	22	35	4
p = 0.60	27	32	2

C Correlates of Risk Attitudes Under Both Models

In this section we report a number of regressions results that estimate relationships between decision variables and available covariates. The covariates are: (i) gender; (ii) participants majoring in Business or Economics (BE); and (iii) cognitive reflection test (CRT) scores.

In the first set of results, presented in Table 8, we study the relationship between the estimated risk attitude coefficients and the covariates. Each individual is a unit of observation. The first three columns of Table 8 report linear OLS estimates of correlates of $\widehat{\gamma}$. The last three consider some determinants of $\widehat{\tau}$. The estimates show that in the EU model, males are significantly less risk averse than females. In the QP model, males are less risk averse, but the effect is significant only when controlling majors and CRT scores. Under the EU model, participants that are majoring in Business or Economics (BE) are significantly less risk averse than participants who have other majors. BE major is not significant under the QP model. The CRT score, a measure of ability/willingness to reflect on a problem, does not exhibit a significant correlation with risk attitude under either model.

Table 8: Determinants of $\widehat{\gamma}$ and $\widehat{\tau}$, linear regressions

	$\widehat{\gamma}$			$\widehat{\tau}$		
	(1)	(2)	(3)	(4)	(5)	(6)
Male	0.08^{**}	0.08^{**}	0.09^{**}	0.07	0.07	0.08^{*}
	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
BE		0.06^{*}	0.08^{*}		0.06	0.08
		(0.03)	(0.03)		(0.04)	(0.04)
CRT			-0.02			-0.02
			(0.01)			(0.02)
Constant	0.83^{***}	0.80^{***}	0.80^{***}	0.38^{***}	0.35^{***}	0.35^{***}
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
F-statistic	7.71	6.08	5.43	3.38	3.12	2.60
Adj. R^2	0.10	0.14	0.22	0.04	0.07	0.09
Num. obs.	61	61	49	61	61	49
RMSE	0.12	0.11	0.10	0.15	0.14	0.14

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 9: Determinants of classification as a QP maximizer in Part 1 of sessions, Logit estimation

	(1)	(2)	(3)
Male	-0.72	-0.78	-0.65
	(0.57)	(0.60)	(0.74)
BE		-1.47^{*}	-2.44^{**}
		(0.62)	(0.85)
CRT			0.58
			(0.39)
Constant	-0.37	0.28	-0.05
	(0.39)	(0.49)	(0.62)
AIC	76.49	72.40	54.95
BIC	80.64	78.63	62.35
Log Likelihood	-36.24	-33.20	-23.47
Deviance	72.49	66.40	46.95
Num. obs.	59	59	47

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

D Instructions and Quiz used in the Experiment

This Appendix contains the instructions that were distributed and read to participants during the experiment. There were two versions of the instructions. Approximately 1/2 of participants received the version reprinted here. The other half received a version that was identical, except for the order in which some material in the instructions for periods 3 - 10 appeared. Specifically, in the other version, the order that the QP and EU rules appeared in the graphical display example shown in the instructions was reversed with the EU rules on top and the QP rules on the bottom. The order in which the rules were described afterward reflected the change in order on the screenshot. Recall, however, that during the periods themselves, the position of each of the six rules on the screen was randomly rescrambled in each period.

Instructions for experiment

You are about to participate in an experiment in the economics of decision making. The instructions are simple and if you follow them carefully and make good decisions, you will earn a considerable amount of money that will be paid to you in cash at the end of the experiment.

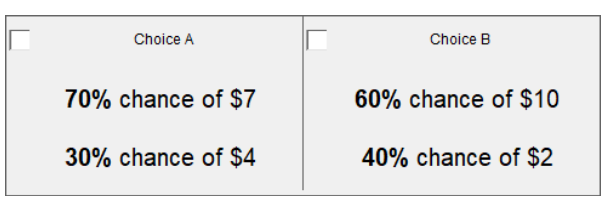
The experiment consists of ten periods. The instructions for periods 1 and 2 will be given now, and those for periods 3–10 will be given later.

Your earnings will equal your show-up fee of $5 plus your earnings from one of the ten periods of the experiment. The period that counts toward your earnings will be determined randomly at the end of the experiment by the computer.

Instructions for periods 1 and 2

In each period you will make 121 separate decisions between two options, called Choice A and Choice B. There will be 11 computer screens in each period, and on each screen, you will have to enter 11 decisions. Each choice can lead to two different payments. The chance that you get either payment is going to change from decision to decision.

An example of a decision is given below. In this example, if you decide for choice A, you can receive $4 or $7. There is a 30% chance of receiving $4 and a 70% chance of receiving $7 if you decide for A. If you decide for B, you can receive $2 or $10. There is a 40% chance that you receive $2 and a 60% chance that you receive $10 if you decide for B. To decide for one of the choices, you can select it by clicking the upper-left corner of the box containing the option.



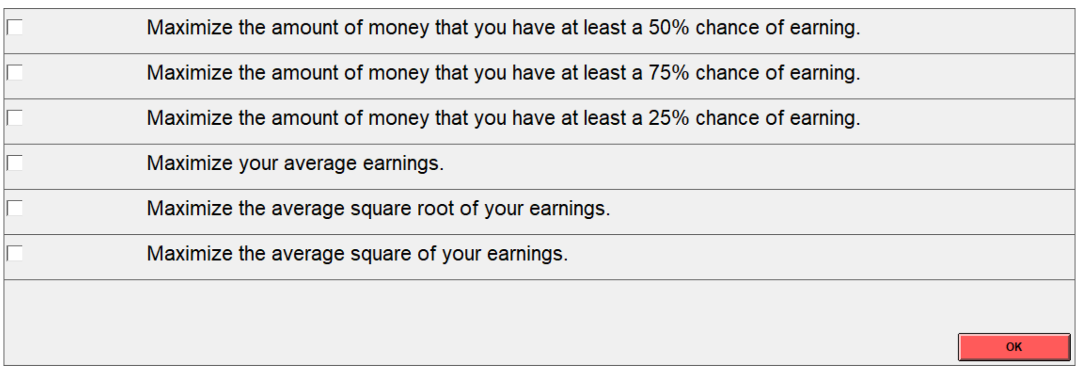
When you have completed the 11 decisions on a screen, select the OK button on the lower right of your screen. You will then see a display of the next 11 decisions.

The possible payments will differ between periods 1 and 2. If either period 1 or period 2 is chosen to count toward your earnings, one of the 121 decisions in the period will be randomly chosen to count toward your payment. This means that if one decision you make counts toward your earnings, then all of the other decisions will not count.

Instructions for periods 3–10

In each period 3–10, a computer robot (a bot) will make 121 decisions between two choices for you. The decisions are similar to the ones that you faced in periods 1 and 2. You will receive the earnings that result from the computer's choices, for the period that is chosen to count toward your earnings. Your role is to choose the rule that the computer will use to make its decision on your behalf.

In each period, you will see a display like the following.



You must select one of the rows. Each row will order the bot to make the 121 choices for the period according to a different rule.

How each rule works:

1. **Maximize the amount of money that you have at least a 50% chance of earning**: This rule will consider each choice and ask the question: "How much money do I have at least a 50% chance of earning?" It will then decide on the choice that has a higher payment in answer to this question. For example, as shown in Table 1 below, for choice A there is a 40% chance of getting \$4 and a 60% of \$9, so that there is at least a 50% chance of at least \$9. For choice B, there is a 70% chance of \$1 and a 30% chance of \$16, so that there is at least a 50% chance of getting at least \$1. Because the payment that you have at least a 50% chance of getting is higher under choice A, the bot will pick A.

Table 1: Example of a decision

Choice A	Choice B
60% chance of $9	**30%** chance of $16
40% chance of $4	**70%** chance of $1

2. **Maximize the amount of money that you have at least a 75% chance of earning**: This rule will consider each choice and ask the question: "How much money do I have at least a 75% chance of earning?" It will then decide on the choice that has a higher payment in answer to this question. For example, for choice A there is a 40% chance of getting $4 and a 60% of $9, so that there is at least a 75% chance of getting at least $4. For choice B, there is a 70% chance of $1 and a 30% chance of $16, so that there is at least a 75% chance of getting at least $1. Because the payment that you have a 75% chance of getting is higher under choice A, the bot will pick A.

3. **Maximize the amount of money that you have at least a 25% chance of earning**: This rule will consider each choice and ask the question: "How much money do I have at least a 25% chance of earning?" It will then decide on the choice that has a higher payment in answer to this question. For example, for choice A there is a 40% chance of getting $4 and a 60% of $9, so that there is at least a 25% chance of getting at least $9. For choice B, there is a 70% chance of $1 and a 30% chance of $16, so that there is at least a 25% chance of getting at least $16. Because the payment that you have at least a 25% chance of getting is higher under choice B, the bot will pick B.

4. **Maximize your average earnings**: This rule will calculate the average amount of money that you would receive under each choice, and always decide for the choice that has a higher average payment. For example, suppose choice A gives you a 40% chance of receiving $4 and a 60% chance of $9, and choice B gives you a 70% chance of $1 and a 30% chance of $16. Then the average payment from choice A is $.4 \times \$4 + .6 \times \$9 = \$7$, and the average payment from choice B is $.7 \times \$1 + .3 \times \$16 = \$5.5$. Because the average payment for choice A is higher, the bot will decide on choice A.

5. **Maximize the average square root of your earnings**: This rule will take each possible amount of money you could earn, and give it a point score equal to the square root of the payment. It then makes a choice to maximize the average number of points you get. In other words, it calculates the average square root of the amount of money that you would receive under each choice, and always decides for the choice that has a higher average square root of the payment. For example, suppose choice A gives you a 40% chance of receiving $4 and a 60% chance of $9, and choice B gives you a 70% chance

of \$1 and a 30% chance of \$16. Then the average square root of the payment from choice A is $.4 \times \sqrt{4} + .6 \times \sqrt{9} = .4 \times 2 + .6 \times 3 = 2.6$ points, and the average square root payment from choice B is $.7 \times \sqrt{1} + .3 \times \sqrt{16} = .7 \times 1 + .3 \times 4 = 1.9$ points. The average square root of payment for choice A is higher, so the bot will decide on choice A rather than B.

Because the bot chooses choice A, you will have a 40% chance of receiving \$4 and a 60% chance of getting \$9. Make sure you realize that the actual payment that you get from each option is the payment amount shown in Table 1; **the point score is only used to determine the bot's choice**. That is, **even though the bot chose A because it has a higher square root of payment, the actual payment you get has a 40% chance of being \$4 and a 60% chance of being \$9**. The difference between rules 4 and 5 is that rule 5 chooses in a way that decreases the variability of your payment at the cost of getting a lower average payment.

6. **Maximize the average square of your earnings**: This rule will take each possible amount of money you could earn, and give it a point score equal to the square of the payment. It then makes a choice to maximize the average number of points you get. This rule will calculate the average square of the amount of money that you would receive under each choice, and always decide for the choice that has a higher average squared payment. For example, suppose choice A gives you a 40% chance of receiving \$4 and a 60% chance of \$9, and choice B gives you a 70% chance of \$1 and a 30% chance of \$16. Then the average square of the payment from choice A is $.4 \times 4^2 + .6 \times 9^2 = .4 \times 16 + .6 \times 81 = 55$, and the average squared payment from choice B is $.7 \times 1^2 + .3 \times 16^2 = .7 \times 1 + .3 \times 256 = 77.5$. Because the average squared payment for choice B is higher, the bot will decide on choice B rather than A.

   Because the bot chooses choice B, you will have a 70% chance of receiving \$1 and a 30% chance of getting \$16. Make sure you realize that the actual payment that you get from each option is the same as shown in Table 1; **the point score is only used to determine the bot's choice**. That is, **even though the bot chose B because it has a higher square of payment, the actual payment you get has a 70% chance of being \$1 and a 30% chance of being \$16**. The difference between rules 4 and 6 is that rule 6 chooses in a way that increases the variability of your payment at the cost of getting a lower average payment.

Quiz

Suppose that there is a choice option which gives you a 60% chance of \$4 and a 40% chance of \$1. Under this choice, what is the

a) Average earnings

b) Average square root of your earnings

c) Average square of your earnings

d) Highest earnings that you have at least a 50% chance of receiving

e) Highest earnings that you have at least a 75% chance of receiving

f) Highest earnings that you have at least a 25% chance of receiving

References

Abdellaoui, M., H. Bleichrodt, and O. L'Haridor (2008): "A Tractable Method to Measure Utility and Loss Aversion under Prospect Theory," *Journal of Risk and Uncertainty*, 36, 245–266.

Allais, M. (1953): "Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine," *Econometrica*, 21, 503–546.

Apiwatcharoenkul, W., L. Lake, and J. Jensen (2016): "Uncertainty in Proved Reserves Estimates by Decline Curve Analysis," in *SPE/IAEE Hydrocarbon Economics and Evaluation Symposium*, Society of Petroleum Engineers.

Bell, D. E. (1982): "Regret in Decision Making Under Uncertainty," *Operations Research*, 30, 961–981.

Bhattacharya, D. (2009): "Inferring Optimal Peer Assignment From Experimental Data," *Journal of the American Statistical Association*, 104, 486–500.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008): "Bootstrap-Based Improvements for Inference With Clustered Errors," *Review of Economics and Statistics*, 90, 414–427.

Chambers, C. P. (2009): "An Axiomatization of Quantiles on the Domain of Distribution Functions," *Mathematical Finance*, 19, 335–342.

Chen, X., O. Linton, and I. van Keilegom (2003): "Estimation of Semiparametric Models When the Criterion Function is not Smooth," *Econometrica*, 71, 1591–1608.

Cox, J. C. and V. Sadiraj (2006): "Small- and Large-Stakes Risk Aversion: Implications of Concavity Calibration for Decision Theory," *Games and Economic Behavior*, 56, 45–60.

de Castro, L. and A. F. Galvao (2019a): "Dynamic Quantile Models of Rational Behavior," *Econometrica*, 87, 1893–1939.

——— (2019b): "Static and Dynamic Quantile Preferences," University of Arizona, mimeo.

Duffie, D. and J. Pan (1997): "An Overview of Value at Risk," *Journal of Derivatives*, 4, 7–49.

Eckel, C. C. and P. J. Grossman (2008): "Men, Women and Risk Aversion: Experimental Evidence," *Handbook of Experimental Economics Results, C. Plott, V. Smith, eds., New York, Elsevier*, 1, 1061–1073.

Fanchi, J. R. and R. L. Christiansen (2017): *Introduction to Petroleum Engineering*, Wiley Online Library.

Frederick, S. (2005): "Cognitive Reflection and Decision Making," *Journal of Economic Perspectives*, 19, 25–42.

Ghirardato, P. and M. Marinacci (2002): "Ambiguity Made Precise: A Comparative Foundation," *Journal of Economic Theory*, 102, 251–289.

Gilboa, I. and D. Schmeidler (1989): "Maxmin Expected Utility with a Non-unique Prior," *Journal of Mathematical Economics*, 18, 141–153.

Giovannetti, B. C. (2013): "Asset Pricing Under Quantile Utility Maximization," *Review of Financial Economics*, 22, 169–179.

Harrison, G. W., E. Johnson, M. M. McInnes, and E. E. Rutström (2005): "Risk Aversion and Incentive Effects: Comment," *American Economic Review*, 95, 897–901.

Harrison, G. W. and E. E. Rutström (2008): "Risk Aversion in the Laboratory," in *Risk Aversion in Experiments (Research in Experimental Economics, Vol. 12)*, ed. by J. Cox and G. Harrison, Emerald Group Publishing Limited, Bingley.

Hey, J. D. and C. Orme (1994): "Investigating Generalizations of Expected Utility Theory Using Experimental Data," *Econometrica*, 62, 1291–1326.

Holt, C. A. and S. K. Laury (2002): "Risk Aversion and Incentive Effects," *American Economic Review*, 92, 1644–1655.

James, G., D. Witten, T. Hastie, and R. Tibshirani (2017): *An Introduction to Statistical Learning with Applications in R*, New York, NY: Springer.

Jorion, P. (2007): *Value at Risk. The New Benchmark for Managing Financial Risk*, vol. 81, McGraw-Hill. 3rd Edition.

Kachelmeier, S. J. and M. Shehata (1992): "Examining Risk Preferences Under High Monetary Incentives: Experimental Evidence from the People's Republic of China," *American Economic Review*, 82, 1120–1141.

Kahneman, D. and A. Tversky (1979): "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47, 263–292.

List, J. A. and M. S. Haigh (2005): "A Simple Test of Expected Utility Theory Using Professional Traders," *Proceedings of the National Academy of Sciences*, 102, 945–948.

Manski, C. (1988): "Ordinal Utility Models of Decision Making under Uncertainty," *Theory and Decision*, 25, 79–104.

Mendelson, H. (1987): "Quantile-Preserving Spread," *Journal of Economic Theory*, 42, 334–351.

Moffatt, P. (2016): *Experimetrics: Econometrics for Experimental Economics*, Red Globe Press. 1st Edition.

Newey, W. K. and D. L. McFadden (1994): "Large Sample Estimation and Hypothesis Testing," in *Handbook of Econometrics, Vol. 4*, ed. by R. F. Engle and D. L. McFadden, North Holland, Elsevier, Amsterdam.

Noussair, C. N., S. Trautmann, and G. van de Kuilen (2014): "Higher Order Risk Attitudes, Demographics, and Financial Decisions," *Review of Economic Studies*, 81, 325–355.

Pakes, A. and D. Pollard (1989): "Simulations and the Asymptotics of Optimization Estimators," *Econometrica*, 57, 1027–1057.

Quiggin, J. (1982): "A Theory of Anticipated Utility," *Journal of Economic Behavior & Organization*, 3, 323–343.

Rabin, M. (2000): "Risk Aversion and Expected-Utility Theory: A Calibration Theorem," *Econometrica*, 68, 1281–1292.

Rostek, M. (2010): "Quantile Maximization in Decision Theory," *Review of Economic Studies*, 77, 339–371.

Rothschild, M. and J. E. Stiglitz (1970): "Increasing Risk: I. A Definition," *Journal of Economic Theory*, 2, 225 – 243.

Saha, A. (1993): "Expo-Power Utility: A 'Flexible' Form for Absolute and Relative Risk Aversion," *American Journal of Agricultural Economics*, 75, 905 – 913.

Savage, L. (1954): *The Foundations of Statistics*, New York: Wiley.

Starmer, C. and R. Sugden (1989): "Violations of the Independence Axiom in Common Ratio Problems: An Experimental Test of Some Competing Hypotheses," *Annals of Operations Research*, 19, 79–102.

Tversky, A. and D. Kahneman (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5, 297–323.

von Neumann, J. and O. Morgenstern (1944): *Theory of Games and Economic Behavior*, Princeton University Press.

Wakker, P. P. (2008): "Explaining the Characteristics of the Power (CRRA) Utility Family," *Health Economics*, 17, 1329–1344.