Do Lottery Gamblers Love Risk or Overweight Small Odds?*

S. Borağan Aruoba
University of Maryland

Melissa Schettini Kearney
University of Maryland

March 6, 2011

Abstract

This paper considers whether lottery betting is best explained by risk-love or an alternative to the expected utility model, namely, the overweighting of small odds. To the best of our knowledge, we are the first to use the common setting of state lottery betting to investigate the fit of expected utility theory against an alternative model of non-linear probability weighting. Our results show that both expected utility with risk-love and risk-neutrality with a nonlinear weighting scheme are consistent with observed data from state lottery gambling. Interestingly, the data prefer a specification that models the agent as having risk-averse preferences *and* overweighting the probability of winning the jackpot using a nonlinear weighting function.

Keywords : Risk preferences, generalized method of moments, gambling, Cumulative prospect theory, expected utility theory, nonlinear weighting, Prelec, behavioral economics, structural modeling

JEL Classification : D03, D12, D81

*We are grateful to John Chao, Emel Filiz Ozbay, Erkut Ozbay and Ingmar Prucha for helpful conversations and to Gabriel Lara-Ibarra for his very capable research assistance. This paper originated as a chapter of Kearney's MIT Ph.D. dissertation (2002). Kearney remains indebted to the following individuals for helpful conversations about that project: Daron Acemoglu, Jonathan Gruber, Botond Koszegi, Sean May, Whitney Newey, Jeff Wilder, and Chris Hansen.

1 Introduction

This paper considers whether lottery betting is best explained by risk-love or an alternative to the expected utility model, namely, the overweighting of small odds. State lottery betting constitutes a real-world setting in which a large population of individuals make repeated decisions involving actual financial consequences. We take the observed allocation of bets across lottery games as an indication of preferences, and we use that to investigate the explanatory power of competing models of decision making under risk. Our paper is most closely related to a small set of papers that evaluate the predictive power of these competing models in the context of horse betting, including Jullien and Salanie (2000) and Snowberg and Wolfers (2010). To the best of our knowledge, we are the first to use the common setting of state lottery betting to investigate the fit of expected utility theory against an alternative model of nonlinear probability weighting. In addition, we consider a model that allows for nonlinear probability weighting and risk-averse preferences. The advantage of considering such a model is that we do not require the data – or the researcher – to choose between risk-aversion and non-linear probability weighting. Indeed, this "compromise" model appears to be well-suited to explaining observed lottery bets.

State lottery gambling almost always involves expected losses, and as such is something that requires explanation within the paradigm of neoclassical economics. Many casual observers simply describe it as a foolish act, driven my misinformation or confusion on the part of bettors. But lottery gambling is not necessarily inconsistent with rational, informed decision making under risk. Perhaps the first and most obvious explanation, within a neoclassical paradigm of expected utility theory, is risk-love. A second plausible explanation, also within the expected utility paradigm, is that lottery gambling includes an entertainment component. Behavioral economics – *Cumulative Prospect Theory*, in particular – offers a third explanation: consumers "overweight" the likelihood of winning the big payout. We examine the statistical fit of each of these various models using data from actual state lottery bets.

Our empirical analysis is based on a sample of pairings of U.S. state lottery games offered over the period during 1992 to 2005 in a set of seven states. We examine weekly sales data on actively selling rolling jackpot lottery game pairs (big jackpot games offered concurrently). Any lottery player entering a lottery vendor in a given state would have the option to place bets on any of the actively selling lotteries. We model the allocation decision made by state lottery players of dollars wagered between games as a function of the win probabilities and jackpot amounts, as well as the structural parameters governing the optimization problem. The choice facing the bettor varies on a weekly basis given the rollover nature of the jackpots.

For example, consider a representative lottery player in the state of Georgia in a given week deciding how to allocate his lottery wagers between Powerball and Lotto South. The pair of gamble characteristics facing this representative lottery player depends on the amount of money that has accumulated in the respective jackpots due to the random chance that the winning numbers have, or have not, been selected in the previous weeks, thereby causing a jackpot to either reset or accumulate. The identification of the parameters in our model comes from the indifference of the players in allocating the marginal dollar between gambles with different statistical properties.

Our investigation proceeds as follows. First, we describe the problem facing a lottery player who is allocating his bets across two concurrently offered games. We work within the established paradigm of this literature in that we (a) assume a representative agent for a subset of bettors and (b) do not endogenize the extensive margin of betting. Second, we make specific assumptions about the players utility, and estimate the governing parameters, under the framework of six alternative models: (i) expected utility model with constant relative risk aversion (CRRA) preferences; (ii) expected utility model with constant absolute risk aversion (CARA) preferences; (iii) expected utility model with risk-neutral preferences and an ad-hoc entertainment component; (iv) nonlinear probability weighting with risk neutral preferences; (v) nonlinear probability weighting with natural logarithm preferences; and (vi) nonlinear probability weighting and general risk averse preferences. Third, we compare the fit of these competing models by applying a novel comparison method based on the Singleton method (1985) of comparing non-nested models and the application of a Condorcet method for ranking the models.

Our results show that both expected utility with risk-love and risk-neutrality with non-linear weighting are consistent with the observed behavior of lottery gambling. In estimating the latter, we obtain levels of overweighting strikingly close to the estimates Snowberg and Wolfers (2010) obtain in the context of horse racing. Perhaps not surprisingly, a specification that allows risk-neutral players to get entertainment value from playing lotteries, over and above the possible expected return from winning the jackpot, performs better than either of the former two alternatives. However, while intuitively quite appealing, the addition of an entertainment function is necessarily ad hoc, and thus for many will be unsatisfying as a modeling device. The alternative that performs the best among those considered, is one that models the agent as having (mildly) risk-averse preferences *and* overweighting the probability of winning the jackpot using a nonlinear weighting function. We suspect many will find this alternative particularly appealing, since it maintains the assumption of risk-aversion – so fundamental to neoclassical economics – while at the same time allowing for nonlinear probability weighting, which has been compellingly put forth by behavioral economics.

There are multiple advantages of the state lottery setting for studying choices under risk. First, lottery game odds are objective and fixed, and potentially knowable by all consumers in advance of their allocation decisions. This contrasts to the frequently-studied case of horse race betting, where win probabilities depend on realized bets and odds are neither fixed nor knowable. Second, lottery gambling is often a repeated activity by lottery players. We can therefore reasonably consider that choices observed reflect established patterns of behavior. This surmounts one concern with the experimental literature, which is that observed deviations from standard theory might in part reflect a lack of experience (see, for example, List, 2003, 2004). Third, the population of lottery players is arguably representative of the general population, as a majority of U.S. adults participate in lottery gambling (see Kearney, 2005). Fourth, we have access to actual (aggregate) sales data and use it to inform our estimation whereas similar studies with horce race betting do not have sales data. Unfortunately, even the state lottery context is not ideal. State lottery sales data are not available for individual bettors. This paper therefore has the limitation of relying on a representative agent assumption among a subgroup of bettors.

The paper proceeds as follows. Section 2 describes the relevant literature. Section 3 presents the problem of a representative lottery player and derives empirical predictions. Section 4 describes the data. Section 5 presents the results of our estimation. Section 6 provides a final discussion.

2 A Brief Review of Relevant Literature

Economists have long noted that certain behaviors do not conform to the joint neoclassical hypothesis of expected utility maximization and diminishing marginal utility. In their classic paper, Friedman and Savage (1948) address the phenomena of people both buying insurance and participating in gambling. They propose that this seemingly-contradictory behavior could be rationalized by placing a convex segment in the middle range of an otherwise concave utility function. This implies that gamblers place a high value on the chance to increase their wealth greatly and thereby move into a new class of wealth. Markowitz (1952) augments the model by placing the convex segment at current wealth, thereby treating gambling as an exploitation of local risk preference. Another tweak on the neoclassical model that can rationalization gambling is to simply allow for an entertainment value of gambling (for example, Conlisk, 1993; Kearney, 2005). As we note below, allowing for entertainment in a non-restrictive way is necessarily *ad hoc*, and thus not completely satisfying as a model.

Tversky and Kahneman's (1992) *Cumulative Prospect Theory* (CPT) – a modified version of their earlier *Prospect Theory* (Kahneman and Tversky, 1979) – offers an alternative class of explanations for documented deviations from the predictions of the neoclassical expected utility hypothesis. CPT is defined by four main properties: (i) reference dependence; (ii) loss aversion; (iii) diminishing sensitivity; and (iv) probability weighting. These four features are designed to explain observed patterns of risky choices, including risk-aversion over gains, risk-seeking over losses, and contemporaneous preference for insurance and gambling. To elaborate on CPT slightly, the concept of "reference dependence" refers to the assignment of value to gains and losses, rather than final assets. (See Rabin, 1996 and Koszegi and Rabin, 2006 for a discussion of *reference-dependent preferences.*) The theory also proposes a *framing effect*, whereby choices made under uncertainty are affected by a shift in the "status quo". The notion of "loss aversion" is captured by a value function that is concave for gains and convex for losses, and generally steeper for losses than gains, resulting in a kink in the utility function at the reference point. In addition, allowing for concavity above the kink and convexity below reflects the observed diminishing sensitivity of preferences. That is, the marginal change in perceived well-being is greater for changes that are closer to one's reference level than those that are further away.

The fourth feature of CPT – and the one considered in this paper – is that it replaces the objective probabilities of the expected utility representation with decision weights, or nonlinear weighting functions.[1] Decision weights are assumed to be lower than the corresponding probabilities in the case of high probabilities and higher than corresponding low probabilities. The weighting function w is a strictly increasing function from $[0, 1]$ to$[0, 1]$ with $w(0) = 0$. Kahneman and Tversky (1979) speculate that overweighting of low probabilities may contribute to the attractiveness of both insurance and gambling. As others have explained, this overweighting of low probabilities is conceptually distinct from a bias in beliefs; rather, it captures preferences, or what might be described as optimism (in the case of gain outcomes) and pessimism (in the case of loss outcomes). For example, an agent evaluating the gamble of $(\$5,000, 0.001)$ understands that he will only receive the \$5000 payout with probability 0.001. The overweighting of the 0.001 probability in the CPT framework is a modeling device that captures the agent's preferences for this gamble over a certain \$5.[2]

Our paper is most closely related to a small set of papers that evaluate the predictive

[1]Unfortunately, the losses associated with each individual lottery purchase are too small and have too little variation to allow us to empirically distinguish between a utility function over wins and losses. We therefore focus on the CPT specification of nonlinear probability weights and do not attempt to test loss aversion in this paper.

[2]State lotteries are a risky, not an uncertain, proposition, as the odds of winning a lotto game are fixed and knowable. However, it is possible that lottery players erroneously attribute some subjective – or biased – element to the risk associated with winning the lottery. For example, Guryan and Kearney (2008) document a belief in a "lucky store" among lottery gamblers. Such a finding suggests that lottery gamblers believe the objective probability to be mutable or in some sense subjective. We consider phenomena such as this to be a bias in beliefs about probabilities, and something conceptually distinct from decision weights.

power of these competing models in the context of horse betting. Jullien and Salanie (2000) explore various alternative models of choice under risk using win data from British horse races between 1986 and 1995. The authors abstract from the extensive decision of whether to bet and the intensive decision of how much to bet, and instead focus on the decision of which horse to bet. In other words, similar to the approach we take in the current paper, the consumer choice problem they model is about the allocation decision, as opposed to the betting decision per se. However, in contrast to our model, they have to assume (for data reasons) that bettors do not spread bets across horses. They assume there is a group of identical bettors who, in equilibrium, are indifferent across horses. The set of equality conditions across the values associated with betting on each horse in the race – determined by the odds and pay-outs – provides the basis for their estimation strategy. The results of their estimation suggest that in the context of betting at British horse races, cumulative prospect theory has higher explanatory power than expected utility theory. In particular, their data reject the assumption of linear probability weighting.

Snowberg and Wolfers (2010) and Gandhi (2010) focus on explaining and exploiting the so-called "long-shot bias" observed in horse race betting. This refers to the well-documented phenomenon that longshots are overbet and favorites are underbet, yielding differential returns (see, for example, Griffith, 1949; Ali, 1977; Weitzman, 1965; and Sauer, 1998). In particular, horses that are less favored to win attract more bets, thereby being both more risky and offering a lower expected return. This reverses the typical mean-variance trade-off that characterizes the typical demand for financial assets. Snowberg and Wolfers (2010) consider the competing explanations of risk-love (with linear probability weights) and over-weighting of small win probabilities (with risk-neutral preferences). They consider complex bets in a way that allows them to observationally distinguish between the predictions of the two models.[3] They utilize data on all 6.4 million horse race starts in the United States from 1992 to 2001. Their empirical examination of bettor choices lends support to a model with decision weights, as opposed to a risk-love model specified under the linear probability assumption of Expected Utility Theory. Gandhi (2010) demonstrates theoretically that the neoclassical hypothesis of risk aversion and rational expectations is consistent with the long-shot bias, once one allows for heterogeneity in consumer preferences. He estimates a random utility model on a sample of North American horse races from 2003-2006 and finds support for this proposition.

[3] To be clear, on this dimension, our paper is more similar to Julien and Salanie (2000) in that our set up does not yield different predictions for gambling choices under the various specifications of the consumer choice problem. Our empirical examination amounts to estimating model parameters under various specifications of the consumer choice problem and observing the resulting parameter estimates and more generally, comparing the fit across models.

We build on this literature in multiple ways. First, to the best of our knowledge, we are the first to use the common setting of state lottery betting to investigate consumer choice under risk in this way. As described elsewhere, the state lottery context offers multiple advantages. One particular advantage as compared to horserace betting is that lottery games involve fixed and objective odds that are potentially knowable to all consumers at the time of their allocation decision. Second, among the class of models we consider, we posit a model that allows for nonlinear probability weighting and risk-averse preferences. An attractive feature of this proposed model is that it does not require one to reject risk-aversion or overweighting in favor of the other.

3 Model Setup

We begin by presenting optimization problem faced by a representative lottery player. We obtain the GMM moment conditions used in our estimation and discuss the various assumptions we must make about preferences. We also provide an overview of the estimation and model comparison methodologies we use.

3.1 General Framework

We focus on the the choice of a representative lottery player between two rolling-jackpot lotto games offered concurrently in a particular week. A one dollar lottery ticket purchase is tantamount to a one dollar bet. In the data, we observe this choice for a particular pairing of games over a series of many weeks as the lottery jackpots accumulate and “reset.”

We make four simplifying assumptions to facilitate our analysis. First, we assume that we can separate the decision of allocating a given amount of money between the two lotteries from all the other decisions the player makes, including the extensive margin (whether or not to bet) and the intensive margin (how much to bet). Lottery players tend to play multiple games, so the allocation decision is a relevant one. A modeling advantage of empirically studying this allocation decision, as opposed to the extensive or the intensive margins, is that we do not have to explicitly account for factors unrelated to the lottery games that drive total lottery sales. Such factors are essentially “differenced out” of the consumer decision-making problem. For example, in a given week in a particular state, total lottery gambling might be depressed because of bad weather or because people are betting instead on a big college basketball game. We implicitly assume that these factors do not affect the allocation of the bets of the player between the two lotteries.[4] Second, we assume that the

[4] To put differently, we do not solve a bigger problem where the player decides how much to spend on lotteries along with other expenditures.

problem and allocation decisions of a representative lottery player for a state can accurately describe the state-level data. This assumption is necessitated by data limitations.[5] Third, we assume that for the purposes of the allocation problem, it is sufficient to focus on the lotto jackpot and not model utility lower prize tiers. This is a reasonable simplification given that the jackpot tends to be orders of magnitude greater than lower prize tiers. Anectodal evidence suggests that lower prize tiers are important to keep people betting and interested in the game. Since we abstract from the participation decision these prizes, it is reasonable to abstract from the lower prize tiers. Fourth, we assume that at most one person can win the jackpot in any given week and one player never wins more than one lottery.

The lotto games in our sample take the form of selecting a set of numbers from a larger field. The probability that a lottery player correctly guesses the draw for lottery i in week t is denoted by p_t^i. These probabilities are readility knowable and reported by state lottery agencies. If a player buys x_t^i tickets for lottery i in week t, then his probability of winning is $p_t^i x_t^i$. The probability p_t^i is typically fixed over time and only changes when the rules of the game are changed. In a rolling-jackpot game, when a jackpot is won in week t, which is determined by the random event of the winning numbers being chosen by a player, the jackpot resets.

If the jackpot is not won in week t, that jackpot money "rolls over" and the *advertised* jackpot for week $t+1$ is the jackpot from week t plus some systematic prediction by the lottery agency. If the jackpot is won in week $t+1$, the winner claims the *actual* jackpot, which is the jackpot of week t, incremented by some fixed fraction of realized sales in week $t+1$. We use R_t^i to denote the present discounted value of the actual jackpot for lottery i in week t. Every week, prior to the lottery draw, the lottery agency announces the advertised jackpot, which we denote by $\tilde{R}_t^i$. For some lotteries the advertised and actual jackpots are identical, which indicates that the lottery agency commits to a certain jackpot and does not adjust it based on sales. For other lotteries, the actual and the advertised jackpots differ. We denote

$$\varepsilon_t^i \equiv R_t^i - \tilde{R}_t^i \tag{1}$$

the *forecast error* of the lottery agency.

[5] To the best of our knowledge, there is no individual level sales data available on lottery gambling in the United States. The 1998 NORC survey is a nationally-representative survey data on gambling behavior in the United States, but it only contains information about the extent of lottery gambling participation in the past year and some very limited information about game type and average amounts wagered.

3.2 Problem of a Lottery Player

As we said, we consider the problem of a lottery player allocating his bets between two concurrently offered lotteries. We normalize the amount the player wants to spend on the lotteries to unity and denote by $s_t \in [0,1]$ the fraction of this amount spent on lottery 1. The consumer is endowed with preferences over risk given by the functional $\mathcal{U}(s_t, M_t, \mathbf{F}_t; \mathbf{\Psi})$, where M_t is the income of the player in week t, $\mathbf{F}_t$ is a vector of the parameters characterizing the distribution of returns of the two lotteries, and $\mathbf{\Psi}$ is a parameter vector describing the player's preferences. $\mathbf{F}_t$ contains $(p_t^1, R_t^1, p_t^2, R_t^2)$ where R_t^i is a random variable at the time the player makes his decisions.[6] The random component comes from the fact that the actual jackpot awarded will depend on the realized sales, which are unknown at the time the player is making his allocation decision.[7]

At the time of making his decision, the player knows the advertised jackpots, $\tilde{R}_t^1$ and $\tilde{R}_t^2$, and the win probabilities, p_t^1 and p_t^2. His eventual payoff, on the other hand, depends on the actual jackpots which will not be observed until the lottery is drawn. The player understands that the difference between the actual and advertized jackpot is ε_t^i as specified above. As we explain in detail below, our empirical strategy relies on linking the player's behavior to his expectations over the jackpots. In order for this strategy to work, these expectations need to satisfy some simple statistical properties: ε_t^i needs to have a zero mean and it needs to be uncorrelated with the current and previous advertized jackpots.[8] We call these the *rational-forecast conditions*, which can be summarized by

$$E\left(\varepsilon_t^1\right) = E\left(\varepsilon_t^2\right) = cov\left[\tilde{R}_{t-s}^1, \varepsilon_t^1\right] = cov\left[\tilde{R}_{t-s}^2, \varepsilon_t^2\right] = 0 \text{ for } s \geq 0 \tag{2}$$

As we explain below, we restrict our empirical analysis to lotteries that satisfy these properties.

The generic problem of the lottery player is to choose s_t maximizing expected utility. Specifically he solves

$$\max_{s_t} E\left[\mathcal{U}(s_t, M_t, \mathbf{F}_t; \mathbf{\Psi}) | I_t\right] \tag{3}$$

where the expectation is taken with respect to the information set at the beginning of the week. This information set includes all the past jackpots on games 1 and 2, as well as the advertized jackpot for the current week.

[6] All moments of a lottery can be computed using (p, R). See Appendix B for details.

[7] In games that require actual jackpot to equal advertized jackpot, there is no random component of R_t^i. This is not a problem for our empirical identification strategy since it requires at least one of the two lotteries to have uncertain jackpots.

[8] If, for example, the mean error is not zero this means the lottery administration systematically mispredicts the jackpot. If, on the other hand, the forecast error is correlated with the advertized jackpot, or any information related to previous weeks, it means that the forecast is not the best that could be made given the information at the time.

Below we solve specific versions of this problem. For each version we obtain a first order condition of this problem, which characterizes any interior solution $s_t \in (0,1)$, in the form of

$$E\left[\xi\left(s_t, M_t, \mathbf{F}_t; \mathbf{\Psi}\right)|I_t\right] = 0 \tag{4}$$

where s_t, M_t and all elements of $\mathbf{F}_t$ are observable to us. We focus on this equation and do not consider corner solutions since in our data, we observe positive sales for every game in every week.

Suppressing the arguments of the $\xi(.)$ function and using $E_t(.)$ to denote the conditional expectation, the first order condition (FOC) is

$$E_t\left(\xi_t\right) = 0 \tag{5}$$

where we label ξ_t as the *FOC error.* This condition states that the deviations of the player's decisions from the solution to the problem above should not be consistently in one direction. In other words, on average the model accurately predicts allocation choices observed in the data.

In addition, since the decision of the player is made using I_t, the FOC error has to be orthogonal to the variables in this information set. Using this insight we can generate additional moment conditions in the form of

$$E\left(\xi_t Z_t\right) = 0 \tag{6}$$

where Z_t is a generic *instrument.* Candidate instruments include any variable that is observed by the player in week t before he makes his decision. In our implementation, we use a set of previous advertised jackpots as instruments.

In what follows, we solve various versions of the player's problem, under different representations of his preferences. First, we consider the standard expected-utility representation of the choice problem. Next, we augment the expected utility representation with an allowance for entertainment. Then we consider a nonstandard representation of the lottery player's problem that allows for nonlinear probability weighting. Finally we consider a risk-averse player who employs a nonlinear probability weighting.

3.2.1 Model 1: Standard Expected Utility

Under the assumption of expected utility, the utility of having x units of cash is given by $u(x;\mathbf{\Psi})$ with parameters $\mathbf{\Psi}$. The lottery player's problem is

$$\begin{aligned}\max_{s_t} E_t\left[\mathcal{U}(s_t, M_t, \mathbf{F}_t; \mathbf{\Psi})\right] &= s_t E_t\left[p_t^1 u\left(R_t^1 + M_t; \mathbf{\Psi}\right) + \left(1-p_t^1\right) u\left(M_t; \mathbf{\Psi}\right)\right] \\ &\quad + \left(1-s_t\right) E_t\left[p_t^2 u\left(R_t^2 + M_t; \mathbf{\Psi}\right) + \left(1-p_t^2\right) u\left(M_t; \mathbf{\Psi}\right)\right]\end{aligned} \tag{7}$$

The first order condition of this optimization problem is

$$\begin{aligned} & E_t\left[p_t^1 u\left(R_t^1+M_t ; \boldsymbol{\Psi}\right)+\left(1-p_t^1\right) u\left(M_t ; \boldsymbol{\Psi}\right)\right] \\ = & E_t\left[p_t^2 u\left(R_t^2+M_t ; \boldsymbol{\Psi}\right)+\left(1-p_t^2\right) u\left(M_t ; \boldsymbol{\Psi}\right)\right] \end{aligned} \tag{8}$$

which states that on the margin the player is indifferent between playing the two lotteries. This must be satisfied for s_t to be an interior solution to the problem, that is for the player to bet in both lotteries.[9] Using this expression we can define the FOC error in (5) as

$$\begin{aligned} \xi_t \equiv & \left[p_t^1 u\left(R_t^1+M_t ; \boldsymbol{\Psi}\right)+\left(1-p_t^1\right) u\left(M_t ; \boldsymbol{\Psi}\right)\right] \\ & -\left[p_t^2 u\left(R_t^2+M_t ; \boldsymbol{\Psi}\right)+\left(1-p_t^2\right) u\left(M_t ; \boldsymbol{\Psi}\right)\right] \end{aligned} \tag{9}$$

Model 1A: Constant Relative Risk Aversion Under the assumption that $u(.)$ is from the constant relative risk aversion (CRRA) family

$$u(x;\gamma)=\begin{cases} \frac{x^{1-\gamma}-1}{1-\gamma} & \text{for } \gamma\neq 1 \\ \log(x) & \text{for } \gamma=1 \end{cases} \tag{10}$$

with γ as the constant of relative risk aversion, (9) simplifies to

$$\xi_t=\begin{cases} \frac{1}{1-\gamma}\left\{p_t^1\left[\left(R_t^1+M_t\right)^{1-\gamma}-\left(M_t\right)^{1-\gamma}\right]-p_t^2\left[\left(R_t^2+M_t\right)^{1-\gamma}-\left(M_t\right)^{1-\gamma}\right]\right\} & \text{for } \gamma\neq 1 \\ p_t^1\left[\log\left(R_t^1+M_t\right)-\log\left(M_t\right)\right]-p_t^2\left[\log\left(R_t^2+M_t\right)-\log\left(M_t\right)\right] & \text{for } \gamma=1 \end{cases} \tag{11}$$

and $\boldsymbol{\Psi}=\gamma$ is the only parameter to be estimated.

Model 1B: Constant Absolute Risk Aversion Under the assumption that $u(.)$ is from the constant absolute risk aversion (CARA) family

$$u(x;\theta)=\begin{cases} \frac{1-\exp(-\theta x)}{\theta} & \text{for } \theta\neq 0 \\ x & \text{for } \theta=0 \end{cases} \tag{12}$$

with θ as the constant of absolute risk aversion, (9) simplifies to

$$\xi_t \equiv \frac{-\exp\left(-\theta M_t\right)}{\theta}\left\{p_t^1\left[\exp\left(-\theta R_t^1\right)\right]-p_t^2\left[\exp\left(-\theta R_t^2\right)\right]+\left(p_t^2-p_t^1\right)\right\} \tag{13}$$

for $\theta\neq 0$. When $\theta=0$, CARA utility and CRRA utility with $\gamma=0$ are identical. The only parameter to be estimated is $\boldsymbol{\Psi}=\theta$.

[9] Note that this expression does not include any marginal utility terms, $u'(.)$, since the choice of the player enters linearly in the objective function $\mathcal{U}(.)$.

3.2.2 Model 2: Nonlinear Weighting

As an alternative to expected utility theory, we consider an alternative representation of consumer preferences from non-expected utility theory. In particular we consider a player who is risk-neutral but he weights the probabilities using a non-linear weighting function (cf. Kahneman and Tversky, 1979). We choose risk-neutral preferences, $u(x) = x$, to isolate the effect of the non-linear weighting function. A common weighting function proposed in the literature is the one derived in Prelec (1998)

$$w(p;\alpha) = \exp\left[-(-\ln p)^{\alpha}\right] \text{ with } \alpha \in (0,1) \tag{14}$$

This weighting function is the unique weighting function that satisfies a number of desirable axioms. It also has the property that $w(p;\alpha) > p$ for all $\alpha \in (0,1)$ and when $p < 1/e = 0.37$. Since the probabilities we work with are in the order of 10^{-7}, a Prelec weighting function is tantamount o an overweighting over the range of win probabilities in the our data.[10]

The lottery player's problem is

$$\begin{aligned}\max_{s_t} E_t\left[\mathcal{U}(s_t, M_t, \mathbf{F}_t; \mathbf{\Psi})\right] &= s_t E_t\left\{w\left(p_t^1;\alpha\right)\left(R_t^1 + M_t\right) + \left[1 - w\left(p_t^1;\alpha\right)\right] M_t\right\} \qquad (15)\\ &\quad + (1 - s_t) E_t\left\{w\left(p_t^2;\alpha\right)\left(R_t^2 + M_t\right) + \left[1 - w\left(p_t^2;\alpha\right)\right] M_t\right\}\end{aligned}$$

with the only parameter to be estimated $\mathbf{\Psi} = \alpha$. The first order condition of the players problem is

$$E_t\left\{w\left(p_t^1;\alpha\right)\left(R_t^1 + M_t\right) + \left[1 - w\left(p_t^1;\alpha\right)\right] M_t\right\} \tag{16}$$

$$-E_t\left\{w\left(p_t^2;\alpha\right)\left(R_t^2 + M_t\right) + \left[1 - w\left(p_t^2;\alpha\right)\right] M_t\right\} = 0 \tag{17}$$

and the FOC error becomes

$$\xi_t = \exp\left[\left(-\ln p_t^2\right)^{\alpha} - \left(-\ln p_t^1\right)^{\alpha}\right] - \frac{R_t^2}{R_t^1} \tag{18}$$

3.2.3 Model 3: Expected Utility with Entertainment

As a minimal departure from expected utility theory, we consider the case where the player is risk-neutral in terms of his lottery proceeds but he gets utility from the nonpecuniary characteristics of playing the lottery.[11] This links our results to the reduced-form analysis of

[10] As an example, with $\alpha = 0.85$ and $p = 10^{-7}$, $w(p;\alpha) = 2.4 \times 10^{-5}$.

[11] Conlisk (1993) has proposed various versions of entertainment value that depend on p and R. In our application we see no value added in making functional form assumptions about the entertainment value.

Kearney (2005). We assume that his utility from playing lottery i is a concave function of s_i, the share of his lottery gambles allocated to lottery i. This entertainment value is given by

$$e\left(s_i;\lambda_i\right)=2\lambda_i\sqrt{s_i} \tag{19}$$

where λ_i is a lottery-specific parameter. Intuitively, differences between λ's captures differential entertainment values between games. For example if a player derives extra consumption value from Powerball, simply because he likes playing the national game relative to the state game, for reasons other than the statistical properties of the games, this will show up as a larger λ for Powerball.

The problem in (3) becomes

$$\max_{s_t} s_t\left[p_t^1 E_t\left(R_t^1\right)+M_t\right]+\left(1-s_t\right)\left[p_t^2 E_t\left(R_t^2\right)+M_t\right]+2\left[\lambda_1\sqrt{s_t}+\lambda_2\sqrt{\left(1-s_t\right)}\right] \tag{20}$$

with the first order condition

$$E_t\left(p_t^1 R_t^1-p_t^2 R_t^2\right)+\frac{\lambda_1}{\sqrt{s_t}}-\frac{\lambda_2}{\sqrt{1-s_t}}=0 \tag{21}$$

The FOC error is defined as

$$\xi_t=p_t^1 R_t^1-p_t^2 R_t^2+\frac{\lambda_1}{\sqrt{s_t}}-\frac{\lambda_2}{\sqrt{1-s_t}} \tag{22}$$

and the parameters to be estimated are $\mathbf{\Psi}=\left(\lambda_1,\lambda_2\right)$. We consider this entertainment specification to be necessarily ad hoc. Nonetheless, it has intuitive appeal since the idea of an entertainment value is a popular explanation for lottery gambling.

3.2.4 Model 4: Log Utility with Nonlinear Weighting

In Section 3.2.2 we considered a risk-neutral player who overweights small probabilities. While this is a popular specification in non-expected utility theory, there is nothing that prevents the player from being risk averse. Especially given that decision makers in many interesting economic problems are risk averse, we think a risk-averse player who overweights small probabilities is an important alternative to consider. As a special case, in this section we consider natural logarithm utility. Combining results from Sections 3.2.1 for $\gamma=1$, and 3.2.2, we obtain the FOC error

$$\xi_t=\exp\left[\left(-\log p_t^2\right)^\alpha-\left(-\log p_t^1\right)^\alpha\right]-\frac{\log\left(R_t^2+M_t\right)-\log\left(M_t\right)}{\log\left(R_t^1+M_t\right)-\log\left(M_t\right)} \tag{23}$$

3.2.5 Model 5: CRRA Utility with Nonlinear Weighting

Finally, a more general specification is a player with a CRRA utility who overweights small probabilities using the Prelec weighting function. Once again, combining results from Sections 3.2.1 and 3.2.2, we obtain the FOC error

$$\xi_t = \exp\left[\left(-\log p_t^2\right)^\alpha - \left(-\log p_t^1\right)^\alpha\right] - \frac{\left(R_t^2 + M_t\right)^{1-\gamma} - \left(M_t\right)^{1-\gamma}}{\left(R_t^1 + M_t\right)^{1-\gamma} - \left(M_t\right)^{1-\gamma}} \tag{24}$$

3.3 Empirical Strategy

We use standard generalized method of moments (GMM) tools to estimate and compare the models specified above. We provide a brief overview of these tools in this section.

3.3.1 Identification

The identification of the parameters in the player's problem comes from the idea that any deviation of the player's choices from what is predicted by the problem above must be due to random events that cannot be predicted at the time of his decision, neither by the player nor by the econometrician. Since we require the advertised jackpots to be rational forecasts of the actual jackpot, large deviations of ξ_t from zero or large correlations of ξ_t with Z_t will lead the data to reject our model. Thus, the estimation will choose the parameters such that the player's choices are as consistent with the data as possible, in essence minimizing the deviations of ξ_t from zero and creating as little correlation between ξ_t and Z_t as possible.

We demonstrate this with an example where foor simplicity we assume risk-neutral preferences, i.e. $u(x; \mathbf{\Psi}) = x$ in the expected utility framework. The moment conditions in (5) and (6) are

$$E_t\left(p_t^1 R_t^1 - p_t^2 R_t^2\right) = 0 \tag{25}$$

$$E_t\left[\left(p_t^1 R_t^1 - p_t^2 R_t^2\right)\tilde{R}_{t-s}^1\right] = 0, \tag{26}$$

where R_t^i is the actual jackpot, $\tilde{R}_t^i$ is the advertized jackpot and $\tilde{R}_{t-s}^1$ is a potential instrumental variable. Using the definition of ε_t from (1), $\varepsilon_t \equiv R_t^i - \tilde{R}_t^i$ these simplify to

$$E_t\left[p_t^1\left(\varepsilon_t^1 + \tilde{R}_t^1\right) - p_t^2\left(\varepsilon_t^2 + \tilde{R}_t^2\right)\right] = 0 \tag{27}$$

$$E_t\left[p_t^1\left(\varepsilon_t^1 + \tilde{R}_t^1\right)\tilde{R}_{t-s}^1 - p_t^2\left(\varepsilon_t^2 + \tilde{R}_t^2\right)\tilde{R}_{t-s}^1\right] = 0 \tag{28}$$

Manipulating further, we get

$$p_t^1\tilde{R}_t^1 - p_t^2\tilde{R}_t^2 + p_t^1 E\left(\varepsilon_t^1\right) - p_t^2 E\left(\varepsilon_t^2\right) = 0 \tag{29}$$

$$p_t^1 E\left(\varepsilon_t^1\tilde{R}_{t-s}^1\right) - p_t^2 E\left(\varepsilon_t^2\tilde{R}_{t-s}^1\right) + \tilde{R}_{t-s}^1\left(p_t^1\tilde{R}_t^1 - p_t^2\tilde{R}_t^2\right) = 0 \tag{30}$$

where the third and fourth terms in (29) and the first and second terms in (30) are zero due to (2) which implies $E\left(\varepsilon_t^1\right) = E\left(\varepsilon_t^2\right) = cov\left[\tilde{R}_{t-s}^1, \varepsilon_t^1\right] = 0$ for $s \geq 0$. Moreover the last term in (30) is zero due to (29). Thus the risk-neutral model, combined with the rationality conditions in (2) leads to

$$p_t^1 \tilde{R}_t^1 = p_t^2 \tilde{R}_t^2 \tag{31}$$

This is an intuitive condition that requires the expected jackpots of the two lotteries to be equal, as we would expect given the specifications of this particular model for a risk-neutral player. This simple example shows that as long as the rational-forecast conditions in (2) are met, that is the advertized jackpots are on average equal to the actual jackpots and the forecast error of the lottery administration is not correlated with past advertized jackpots, then the failure of the moment conditions to hold can be interpreted as the data being inconsistent with the model. We must emphasize that this is a joint test of the model and the rational-forecast conditions in (2). A player can be following out model exactly but if the lottery administration's advertized jackpots do not satisfy one of the conditions in (2), then we may find that the data is not consistent with our model.

3.3.2 Estimation

Let ξ_t^i for $t = 1, ..., T$ be the expectational error from model i, $\mathbf{\Psi}^i$ be a $a \times 1$ vector of parameters to be estimated and Z_t^k for $k = 1, ..., r-1$ be the set of instruments, which are common across all models. Let $\mathbf{Y}_t$ collect all observed variables in period t, i.e. the expectational error and the instruments and $\mathbf{Z}_t \equiv \begin{bmatrix} 1 & Z_t^1 & ... & Z_t^{r-1} \end{bmatrix}'$. Then the $r \times 1$ vector of population moment conditions is given by

$$\mathbf{h}_t^i\left(\mathbf{\Psi}_0^i, \mathbf{Y}_t\right) \equiv \xi_t^i \mathbf{Z}_t \tag{32}$$

where $\mathbf{\Psi}_0^i$ is the vector of true parameters. The sample analog of this condition is defined as the $r \times 1$ vector

$$\mathbf{g}^i\left(\mathbf{\Psi}^i\right) \equiv \frac{1}{T}\sum_{t=1}^{T} \mathbf{h}_t^i\left(\mathbf{\Psi}^i, \mathbf{Y}_t\right) \tag{33}$$

The GMM estimate of $\mathbf{\Psi}^i$ is obtained by minimizing the objective function

$$\hat{\mathbf{\Psi}}^i = \arg\min_{\mathbf{\Psi}^i} \left[\mathbf{g}^i\left(\mathbf{\Psi}^i\right)\right]' \mathbf{W}^i \left[\mathbf{g}^i\left(\mathbf{\Psi}^i\right)\right] \tag{34}$$

where $\mathbf{W}^i$ is a $r \times r$ weighting matrix. Hansen (1982) proves that the optimal weighting matrix is the inverse of the variance of the moment conditions where $\sqrt{T}\mathbf{g}^i\left(\mathbf{\Psi}^i\right) \to \mathbf{S}^i$ and $\mathbf{W}^i = \left(\mathbf{S}^i\right)^{-1}$. Defining the sample auto-covariance of $\mathbf{h}^i\left(\mathbf{\Psi}^i, \mathbf{Y}_t\right)$ as

$$\hat{\mathbf{\Gamma}}_v^i \equiv \frac{1}{T}\sum_{t=v+1}^{T} \left[\mathbf{h}^i\left(\hat{\mathbf{\Psi}}^i, \mathbf{Y}_t\right)\right]\left[\mathbf{h}^i\left(\hat{\mathbf{\Psi}}^i, \mathbf{Y}_t\right)\right]' \tag{35}$$

the Newey-West (1982) estimator of S is given by

$$\hat{\mathbf{S}}^i = \hat{\mathbf{\Gamma}}_0^i + \sum_{v=1}^{q} \left(1 - \frac{v}{q+1}\right) \left[\hat{\mathbf{\Gamma}}_v^i + \left(\hat{\mathbf{\Gamma}}_v^i\right)'\right] \tag{36}$$

We use an iterative procedure to estimate $\hat{\mathbf{\Psi}}^i$. First, setting $\mathbf{W}^i$ to the appropriate identity matrix and obtain an estimate of $\hat{\mathbf{\Psi}}^i$ using (34). Next we plug this estimate in (36) and compute $\hat{\mathbf{S}}^i$ and thus a new $\mathbf{W}^i$. Using this new weighting matrix we estimate a new $\hat{\mathbf{\Psi}}^i$. We continue this process until the change in the objective function becomes arbitrarily small. While all estimates in obtained during the iterations are consistent, the last one we obtain is the most efficient one.

3.3.3 Testing

As long as $a < r$, which is always satisfied in our application, there are $r - a$ extra moment conditions and we can use over-identifying restrictions to test the fit of our model. Hansen (1982) shows that the test statistic given by

$$J^i = T \left[\mathbf{g}^i\left(\hat{\mathbf{\Psi}}^i\right)\right]' \left(\hat{\mathbf{S}}^i\right)^{-1} \left[\mathbf{g}^i\left(\hat{\mathbf{\Psi}}^i\right)\right] \tag{37}$$

which is simply T times the minimized value of the objective function, is distributed χ^2 with $r - a$ degrees of freedom. Intuitively, if the objective function is sufficiently larger than zero, the value if the underlying model is true, then this constitutes evidence against the model. It should be noted that this test would be a joint test of the model and the particular instruments used in estimation. To be symmetric in our treatment of models, we use the same instruments across models.

While the J-stat is a useful test of one model, it is not directly useful for comparing multiple models, especially if they are not nested as is the case for our models. To accomplish this, we use the test developed by Singleton (1985).[12] Let the null hypothesis be that model i is true and the alternative hypothesis be that model j is true, which we refer to as testing model i versus model j. Then the test statistic is given by

$$\lambda^{i,j} = \frac{T\left\{\left[\mathbf{g}^{*i}\left(\hat{\mathbf{\Psi}}^i\right) - \mathbf{g}^{*j}\left(\hat{\mathbf{\Psi}}^j\right)\right]' \left(\hat{\mathbf{S}}^{*i}\right)^{-1} \mathbf{g}^{*i}\left(\hat{\mathbf{\Psi}}^i\right)\right\}^2}{\left[\mathbf{g}^{*i}\left(\hat{\mathbf{\Psi}}^i\right) - \mathbf{g}^{*j}\left(\hat{\mathbf{\Psi}}^j\right)\right]' \mathbf{U}^{1/2} \left(\hat{\mathbf{S}}^i\right)^{-1} \hat{\mathbf{\Sigma}} \left(\hat{\mathbf{S}}^i\right)^{-1} \mathbf{U}^{1/2} \left[\mathbf{g}^{*i}\left(\hat{\mathbf{\Psi}}^i\right) - \mathbf{g}^{*j}\left(\hat{\mathbf{\Psi}}^j\right)\right]} \tag{38}$$

[12] Specifically we use the λ^* [*scaled*] test in Singleton (1985). The scaling is done to make sure all moment conditions across models have similar scales. While there has been advances in the field of comparison of non-nested models estimated via GMM that extended the work of Singleton (1985), such as Ghysels and Hall (1990) and Smith (1992), our application meets the conditions of Singleton (1985) (e.g. number of instruments being identical across competing models) and therefore does not require these more general cases.

which is distributed as χ^2 with one degree of freedom.[13] If $\lambda^{i,j}$ is greater than the specified critical value, this means that there is sufficient evidence against model i in the direction of model j.

3.3.4 Model Comparison

The Singleton test, by its very nature, is a pairwise test and it provides a way of comparing two models at a time. We develop a novel methodology using results from the voting literature to combine test results from multiple models and multiple pairs of lotteries to determine an ordering of models.

Specifically, we use the Condorcet method, which we sketch briefly and refer the reader to Black (1986) for details. Let's consider a voting situation with I alternatives and S voters. In the Condorcet method, we compare alternatives one pair at a time and ask each of the voters to choose one alternative from every pairwise comparison. We then aggregate these results across voters to determine the winner in each pairwise comparison. The Condorcet winner is the alternative that wins in all the pairwise comparisons. Once the winner is determined, we remove this alternative from the choice set and repeat the process to find the second-place alternative and so on. It is well known that a Condorcet winner may not exist and a typical alternative is to use the Borda method. Fortunately, in our application we have a winner in most cases. Moreover, and more importantly, as we explain below, our analysis may not yield a clear ordering of alternatives, which is what the Borda method requires.

The Condorcet method requires "votes". In order to interpret the results from our pairwise Singleton tests we employ the following approach. There are four possible outcomes of a pairwise comparison of two models using the Singleton test: (i) data can reject Model 1 and fail to reject the Model 2, (ii) reject Model 2 and fail to reject Model 1, (iii) fail to reject both models and (iv) reject both models. Let P^s_{ij} and P^s_{ji} denote the p-values of testing model i versus model j and model j versus model i using the pair s. We propose two ways of classifying these results as votes. First, if $P^s_{ij} < \mathcal{P}$ and $P^s_{ji} > \mathcal{P}$ for a significance level $\mathcal{P}$ then we count this as a vote for model j. Alternatively, if $P^s_{ij} > \mathcal{P}$ we count this as a vote for model i. We can use a similar rule to infer the ordering of models for a particular pair but due to the nature of the underlying pairwise test, transitivity of preferences (across

[13] Here the variables with stars are their scaled versions and are defined as $\mathbf{g}^{*i}\left(\hat{\mathbf{\Psi}}^i\right) \equiv \mathbf{U}^{-1/2}\mathbf{g}^i\left(\hat{\mathbf{\Psi}}^i\right)$, $\mathbf{g}^{*j}\left(\hat{\mathbf{\Psi}}^j\right) \equiv \mathbf{V}^{-1/2}\mathbf{g}^j\left(\hat{\mathbf{\Psi}}^j\right)$, $\hat{\mathbf{S}}^{*i} \equiv \mathbf{U}^{-1/2}\hat{\mathbf{S}}^i\mathbf{U}^{-1/2}$ where $\mathbf{U}^{-1/2} \equiv \left[diag\left(\hat{\mathbf{S}}^i\right)\right]^{-1/2}$ and $\mathbf{V}^{-1/2} \equiv \left[diag\left(\hat{\mathbf{S}}^j\right)\right]^{-1/2}$. We also use the definition $\mathbf{\Sigma} \equiv \hat{\mathbf{S}}^i - \hat{\mathbf{D}}^i\left[\left(\hat{\mathbf{D}}^i\right)'\left(\hat{\mathbf{S}}^i\right)^{-1}\left(\hat{\mathbf{D}}^i\right)\right]^{-1}\left(\hat{\mathbf{D}}^i\right)'$ where $\hat{\mathbf{D}}^i \equiv (1/T)\sum_{t=1}^{T}\left[\frac{\partial \mathbf{h}^i\left(\hat{\mathbf{\Psi}}^i\right)}{\partial \hat{\mathbf{\Psi}}^i}\right]$.

competing models) is not guaranteed.

We conclude this section with an important caveat. While the Singleton test is a valid statistical test which is useful for comparing two models, the statistical properties of the methodology we develop for ranking models as described above is not yet known. More work utilizing Monte Carlo experiments needs to be done to that end to formalize this methodology. As such, we consider our model comparison results to be mostly descriptive.

4 Data

The empirical analysis of this paper utilizes repeated cross-sectional data on lottery games and sales. Each observation represents a state-specific pairing of two lotto games in a week. First, information on lottery games, including win probabilities, was obtained from state lottery agency websites and directly from state lottery agencies as needed.[14] Second, historical data on advertised jackpot amounts and actual jackpot amounts were obtained directly from the Multi-State Lottery Association and individual state lottery agencies. (In constructing our sample of states and games, we are limited to those for which a state lottery agency was willing and able to provide historical data on non-winning jackpots.) All jackpot amounts are adjusted to year 2002 dollars. Third, we purchased weekly sales data at the level of state and game for the full period 1992 through 2005 from Lefleur's Inc., a private company that collects sales data from state lottery agencies.

An observation is defined at the level of state and week for an actively selling pair of concurrently-offered rolling jackpot games. As noted elsewhere, our estimation approach requires that there be some uncertainty in the model at the time the bettor is making his allocation decision. This comes from a game paying out actual jackpots that are based on realized sales, and hence deviate from the advertised jackpot. We require that this deviation, which we have labeled to be the forecast error, to have mean zero. We also require that the forecast error be orthogonal to the advertised jackpot with at least a five-week lag. Specifically, we require

$$E\left(\varepsilon_t^i\right)=0 \text{ and } cov\left[\tilde{R}_{t-s}^i,\varepsilon_t^i\right]=0 \text{ for } s\geq 5. \tag{39}$$

We start at $s=5$ instead of $s=0$ to make sure any dependence in the short-sample is cleared. In particular, the reset of a lottery is a fairly rare event and in most weeks the advertised jackpot will be a mechanical function of last week's advertised jackpot. For $s<5$ virtually all lotteries fail the rational-forecast restrictions.[15]

[14] The data is an updated subset of the data used in the reduced-form demand analysis of Kearney (2005).

[15] The failure of either of these conditions to hold might suggest that these conditions are not satisfied in

The final analysis dataset is comprised of seven state-game-pairs, observed for differing numbers of weeks between 1992 and 2005. In each state, the pairing includes the multi-state Powerball game. The second game in each of the seven state-game-pairs is as follows: (1) Georgia - Lotto South; (2) Kentucky – Lotto South; (3) Louisiana – Louisiana Lotto; (4) Missouri – Missouri Lotto (5) Montana – Montana Cash; (6) Oregon – Megabucks; and (7) Wisconsin- Megabucks. In the cases of Georgia and Kentucky, the Powerball is sold alongside the multi-state Lotto South game, which sold in Georgia, Kentucky, and Virginia. In the other five cases, Powerball is paired against a proprietary within state game. In our application, the forecast error always comes from the Powerball game – that is, the deviation of the actual jackpot from the advertised amount. The state lotto games in our sample all pay out the advertised amount as the winning jackpot, so there is no forecast error for those games. Table 1 reports summary statistics about the lottery games. Appendix A provides some description about the particular games.

5 Results

We start by looking at some simple statistics from the data, which are reported in Table 2. For each lottery pair we report the average (across weeks) of the expected value and standard deviation of each of the lotteries, where the latter is a measure of the riskiness of the lottery.[16] To account for the variation in these statistics over time we also report the percentage of observations where the first lottery has a higher expected value or standard deviation. Finally, we report average and minimum share of the sales of the two lotteries.

A common pattern emerges from the data. Powerball has a lower expected value (except for pair 1 where the expected values are statistically equal), a higher standard deviation and yet its share in sales is, on average, the same or larger than alternative lotteries.[17] It is clear that this pattern is not consistent with expected utility theory with risk-averse or risk-neutral lottery players: given that the expected return of Powerball is lower and its risk is larger, a risk-averse or risk-neutral lottery player would put all his money in the alternative lottery. This implies that for this behavior to be consistent with expected utility

the population of draws, in which case the state lottery agency (and lottery bettors) could systematically improve their forecast of the actual jackpot amount. These conditions might be violated in our sample of weeks simply due to a small numbers problem. As an example, we discarded the Wildcard game - sold in Idaho, Montana, North Dakota, and South Dakota - from our sample of potential games to be included in the analysis sample because in our sample of weeks, the mean forecast error does not equal zero.

[16] As we show in Appendix B, the expected value of a lottery is given by pR and the standard deviation is given by $\sqrt{p\,(1-p)}R$.

[17] These expected values are based just on the odds of winning the top prize times the jackpot amount. They do not include payouts for lower prize tiers, and thus are not interpretable as the actual expected value of a particular lottery game.

theory, lottery platers need to be risk loving: the large risk of the lottery compansates for the lower expected value. Alternatively, the lottery players may be (at least locally) risk neutral but overweight the very small probability of winning. Since the expected value of the lottery, which is the only thing a risk-neutral player cares about, is simply pR, if the player percieves p to be sufficiently larger than it is, he may end up being indifferent between the two lotteries. This would also lead him to put some of his money to each of the lotteries.

With this motivation, in the following sections, we estimate each of the alternative models we developed in Section 3 and compare them. For each model we report the estimates for the parameters of the model, as well as the Hansen (1982) J-stat which tests the over-identifying restrictions. As we explained above, this is a joint test of the particular model and the instruments used in estimation. Finally, we run the pairwise Singleton (1985) tests and aggregate the results using the Condorcet method in order to obtain a ranking of models.

5.1 Risk-Love vs. Over-weight

As we explained above, perhaps the most obvious candidates to explain the data in Table 2 are assuming risk-love and over-weighting. We use the CRRA and CARA expected utility specifications discussed in Section 3.2.1 for the former and the Prelec (1998) nonlinear weighting discussed in Section 3.2.2 for the latter.

The two panels of Table 3 report the results for the expected utility specifications. With the exceptions of Pairs 1 and 3, results show a mild degree of risk-loving behavior. For Pairs 1 and 3, risk-neutrality ($\gamma = 0$ for CRRA and $\theta = 0$ for CARA) cannot be rejected. Turning to the J-stats, there are no rejections of the over-identifying restrictions for the CRRA specification, while there is only one rejection, Pair 6, for the CARA specification. Overall, it seems like risk-loving behavior with standard expected utility is a reasonable description of the data.

Table 4 reports the results for risk-neutral preferences with Prelec nonlinear weighting. All estimates of α are statistically different from unity, which show that probability weighting is nonlinear. The average of the α estimates is 0.89, which is remarkably close to the estimate in Snowberg and Wolfers (2010). The J-stats indicate that the over-identifying restrictions are not rejected for any of the pairs. We thus conclude that this specification is also a reasonable description of the data.

Results on Table 5 attempt to distinguish the three alternative models (the two expected-utility specifications and the Prelec specification) and possibly rank them. The first panel reports the J-stats of each estimation and the sum of these J-stats at the bottom. As we explained in Section 3.3.3, since our alternative models are not nested, comparing J-stats would be misleading, especially because the differences in J-stats may not be statististically significant. Nevertheless, we think that the comparisons in the first two panels of Table 5 may be somewhat useful.[18] According to panel (a), the CRRA specification of expected utility ranks first, followed by the Prelec specification and the CARA specification of expected utility is ranked last. In panel (b), we count the number of pairs for which a given model has lower J-stat relative to another model and prepare a Condorcet table. In this and similar tables that follows, the number corresponding to the cell (i, j) report the number of "votes", appropriately defined, in favor of model i compared to model j. For example, panel (b) of Table 5 shows that Model 1A has 6 votes in its favor and 1 vote against, compared to Model 1B. Dark shadings reflect the winner in a pairwise comparison. According to the results in this panel, then, the ranking of the three models is CRRA, followed by CARA and Prelec.

The most reliable ordering of the models can be obtained by the Singleton tests. Table A1 in the Appendix report results for all possible pairwise comparisons. In panel (c) of Table 5, we convert these to a Condorcet table, following the procedure described in Section 3.3.4 with significance level $\mathcal{P} = 0.10$. According to this results as well, CRRA is the best alternative, followed by CARA and then Prelec.

To sum up results so far, the CRRA specification of expected utility seems to fit the data best. Moreover, since the Singleton test takes into account the statistical significance of differences between J-stats, perhaps results in panel (c) should receive more weight. As a result, we conclude that expected utility with risk-love is preferred by the data compared to over-weighting.

5.2 Entertainment

While risk-love seems like a reasonable explanation of the large sales share of Powerball, – or at least lottery players' indifference versus other alternatives – for reasons explained in the Introduction this may not be a good description of behavior of economic decision makers. To explore if we can attribute the preference for Powerball to some entertainment value, we consider the ad-hoc entertainment specification with risk-neutral preferences presented in Section 3.2.3.

The first panel of Table 6 shows the estimation results. For four pairs out of seven, Powerball has a significantly larger λ (entertainment value) than the alternative lottery, as evidenced by the Wald test which test $\lambda_1 = \lambda_2$. The J-stats do not indicate any evidence against this model for any of the pairs. Moreover, the sum of the J-stats, our informal major of goodness of fit, is 13.96 which is substantially lower than the other alternatives considered so far. Indeed the second panel shows that using the J-stat criterion, entertainment model

[18] Note that if we had perfectly overlapping samples across pairs, then the total J-stat of individual estimations could have been considered to be the J-stat of a bigger system combining all, with equal weights.

is superior than the other alternatives. Using the Singleton test, the last panel shows that entertainment is a close second after the CRRA specification of the expected utility model, determined by only a single "vote" difference with a Singleton test p-value of 0.04.

Summing up,combining the results above, it seems like the CRRA specification and the entertainment models can be ranked similarly, meaning that there is about the same level of support for risk-love and entertainment which is significantly more than over-weighting.

5.3 Log Utility with Nonlinear Weighting

Our results so far indicate that risk-love and an ad-hoc entertainment value provide the best fit to the data. This result is not very appealing for at least two reasons. First, in many contexts economists think of decision makers as being risk averse and modeling them as risk-loving in the lottery betting context seems inconsistent. Second, while entertainment is certainly a reasonable explanation (both ex ante and ex post), it is a "cheap" one since one can explain any behavior by changing the utility function.

To address all of these concerns, we estimate the model with log utility (equivalently CRRA utility with $\gamma = 1$), combined with the Prelec weighting function, developed in Section 3.2.4. This model makes progress in all of the three points above since it displays risk-averse decision makers that care only about the statistical properties of the return from betting and they possibly overweight the probability of winning the jackpot. We should note that to the best of our knowledge, this specification has not been estimated or tested in the literature before.

Table 7 presents the results. Compared to the results on Table 4, where the lottery players were assumed to be risk-neutral, the degree of overweighting is substantially larger, with an average of $\alpha = 0.58$ for the Prelec parameter. This indicates that the lottery players need to be over-weighting the probability of winning the jackpot by three orders of magnitude[19]. While this is somewhat extreme and the α estimate is smaller than similar estimates in the literature, Prelec (1998) argue that a value of $\alpha = 0.65$ provides a similar degree of over-weighting as in the estimates in Tversky and Kahneman (1992) and some other related studies.

Turning to the J-stats, there are no rejections and the sum of the J-stats across pairs is 16.41, which is only slightly larger than the level achieved by the entertainment model and lower than all other alternatives. According to J-stat Condorcet results in the second

[19]With $p = 10^{-7}$, and $\alpha = 0.58$ we have $w(p;\alpha) = 6.6 \times 10^{-3}$, .while with $\alpha = 0.89$ we have $w(p;\alpha) = 6.99 \times 10^{-6}$.

panel, the entertainment model is still the best alternative but log utility with nonlinear weighting comes ahead of other alternatives. The Singleton Condorcet results presented in the last panel does not show a clear winner. Log utility with nonlinear weighting is tied (separately) with both CRRA specification of expected utility and the entertainment model. It is clear that the entertainment model is not the winner since loses versus the CRRA specification of the expected utility model. It is therefore reasonable to conclude that log utility with overweighting and the CRRA specification of the expected utility model are tied and are better than other alternatives. Inspecting the underlying p-values in Table A1 in the appendix, we see that the former has a slight advantage since the p-values against it are typically larger.

5.4 General Risk Aversion with Nonlinear Weighting

In the previous section we presented evidence that log utility (i.e. CRRA with $\gamma = 1$) along with nonlinear weighting of probabilities can go a long way in explaining why lottery players play Powerball, despite its low expected return and high risk. In this section we pursue this idea further and investigate other levels of risk-aversion using the model in Section 3.2.5, in addition to $\gamma = 1$, as there is nothing magical about this level. In principle we would have liked to estimate this specification, obtaining estimates for γ and α jointly. However, this turned out to be infeasible since many pairs of (γ, α) is consistent with the data as we demonstrate below, i.e. these parameters are not separately identified.

To understand which pairs of (γ, α) is consistent with the data, in Figure 1 we plot the contours of the J-stats, which is directly linked to the value of the GMM objective function. Each of the panels plot these contours for a pair in the range $\gamma \in [-1, 10]$ and $\alpha \in [0, 1]$. Darker (blue) colors represent small numbers and brigter (red) colors represent large numbers. A "valley" running from the top-left to the bottom-right is visible in each of these figures. The solid line shows the value of α that is estimated for each value of γ that is fixed, similar to the exercise in the previous section. In essence this line traces the bottom of the said valley and represent the locus of (γ, α) pairs that are "reasonable". The rightmost red star in each graph is the risk-neutral ($\gamma = 0$) preferences with the overweighting parameter estimated (roughly $\alpha = 0.9$) in Section 5.1. The other red star shows log preferences ($\gamma = 1$) with nonlinear weighting estimated in the previous section with and α of about 0.5. These figures clearly show that most levels of risk-aversion are consistent with the data with an appropriate degree of overweighting.

To find the "trough" of these valleys in the range of $\gamma \in [-1, 10]$, in Figure 2 we plot the J-stats obtained at the (γ, α) pairs on the solid line in Figure 1. In other words, for every value of γ, we fix the level of risk aversion and estimate the CRRA model with nonlinear weighting and report the value of the J-stat from this estimation in Figure 2. The vertical dotted lines in each panel denote the location of $\gamma = 1$, which is the point estimated in the previous section. A few observations stand out. First, quite remarkably for Pairs 1, 2, 3 and 4, $\gamma = 1$ achieves almost the lowest level of J-stat, that is the best fit. Second, except for Pair 6, all pairs display the lowest level of J-stat in the risk-averse range, that is when $\gamma > 0$. For Pair 6 the minimum occurs at $\gamma = -1.5$ (not shown). Third, most pairs, J-stats are decreasing around $\gamma = 10$. This is the source of the identification problem. Our estimation routine continues to search past $\gamma = 10$ and eventually reaches extremely small levels of α which creates numerical problems.

As our last model, we use the levels of γ that minimizes the J-stat in Figure 2 and estimate the degree of nonlinearity in the weighting function. For Pair 6, since the minimum occurs at level consistent with risk-love, we use $\gamma = 1$, i.e. the results from the previous section, to be conservative. Table 8 presents the estimation results. For every pair we report the value of γ we fix, the estimated overweighting parameter α and the J-stat of the estimation. With two exceptions, Pairs 5 and 7, the α estimates are similar to those estimated in the previous section. None of the J-stats indicate a rejection and the sum of J-stats is 12.27, which is lower than all other alternatives and closest to that obtained using entertaiment.

Table 9 compares all alternative models. The new model considered here, Model 5 is quite simply the best alternative: it achieves the lowest sum of J-stats, it is the winner in the Condorcet rankings using both J-stats as well as Singleton tests. Thus, our results indicate that some level of risk aversion, perhaps as low as the one implied by log utility, coupled with nonlinear weighting of probabilities is consistent with the behavior of lottery players in our sample.

6 Final Discussion

This paper considers whether lottery betting is best explained by risk-love or an overweighting of small odds. We consider the allocation problem of an agent faced with a pair of lottery games offering different statistical properties; in particular, we observe the allocation of lottery bets between the Powerball game and a smaller (typically single state) lotto game. The identification of the parameters in our model comes from the indifference of the players in allocating the marginal dollar between gambles with different statistical properties. Our empirical analysis uses data at the state-week level for seven such pairs, for a total of 3,205 observations observed at various points during the period 1992 to 2005. We specify the consumer problem under various alternative models and utilize GMM techniques to estimate the parameters of the competing models.

We add to a small literature of empirical papers that attempt to estimate and compare these models outside of a laboratory setting. To the best of our knowledge, we are the first to use the common setting of state lottery betting for this purpose. Another contribution of our paper, vis a vis previous studies, is that we consider a model that maintains the assumption of risk-aversion but allows for nonlinear probability weighting. However, like previous papers that have used the context of horse race betting, our paper is limited in two important ways: (1) we do not model the extensive decision of whether to bet at all, but rather focus on the allocation problem across games and (2) the absence of individual-level data on bets requires that we make a representative agent assumption among a subgroup of betters. Our results do not yield rejections in terms of model fit. Therefore, to compare and rank models, we apply a novel approach employing the Singleton (1985) method of comparing pairs of non-nested models estimated via GMM and the Condorcet method of aggregating votes.

Our results show that both expected utility with risk-love and risk-neutrality with non-linear weighting are consistent with the observed behavior of lottery gambling. With risk-neutral preferences, we find a Prelec weighting parameter of about 0.9. Perhaps not surprisingly, an entertainment-augmented expected utility representation of the problem, with risk-neutral preferences, performs better than either of the former two alternatives. However, this ad-hoc entertainment specification is arguably unsatisfying as a modeling device. Much more surprising and gratifying (in our view), is that the alternative that performs the best among those considered, is one that models the agent as having risk-averse preferences and overweighting the probability of winning the jackpot using a nonlinear weighting function. The CRRA risk aversion parameters consistent with this specification are typically around unity and the estimated Prelec parameters are around 0.5.[20] This specification maintains the assumption of risk-aversion – so fundamental to neoclassical economics – while at the same time allowing for non-linear probability weighting, which has been compellingly put forth by behavioral economics. This potentially suggests that there need not be a tension. It could be that people are always and everywhere risk averse but use a Prelec-type nonlinear weighting function to evaluate probabilities. For relatively large probabilities, e.g. those greater than 1%, this overweighting plays essentially no role.[21] But for events with probabilities in the tails of the distribution, overweighting might indeed be fundamental to explaining behavior.

[20] Chetty (2006) develops a method of estimating the coefficient of relative risk aversion (gamma) using data on labor supply behavior. In particular, he shows that existing evidence on the effects of wage changes on labor supply imposes a tight upper bound on the curvature of utility over wealth (gamma < 2). Chetty calculates a mean implied value of gamma of 0.71, with a range of 0.15 to 1.78. Except for two states, our levels of risk aversions are within these bounds.

[21] As an example, with $\alpha = 0.85$ and $p = 10^{-7}$, $w(p;\alpha) = 2.4 \times 10^{-5}$ but with $p = 0.10$, $w(p;\alpha) = 0.13$, which is only slightly higher.

References

[1] Ali, M.M. 1977. "Probability and utility estimates for racetrack bettors." Journal of Political Economy 85, 803-815.

[2] Black, Duncan. 1986. The Theory of Committees and Elections. Springer.

[3] Chetty, Raj. 2006. "A New Method of Estimating Risk Aversion," American Economic Review 96(5): 1821-1834, Dec. 2006.

[4] Conlisk, John. 1993. "The Utility of Gambling." Journal of Risk and Uncertainty 6(3): 255-275.

[5] Friedman, Milton and L.J. Savage. 1948. "The Utility Analysis of Choices Involving Risk," Journal of Political Economy 56(4), 279-304.

[6] Ghandi, Amit. 2010. "Estimating Preferences for Risk: Solving the 'Anomaly' of the Longshot Bias," University of Wisconsin-Madison, mimeo.

[7] Ghysels, Eric and A. Hall (1990). "Testing Nonnested Euler Equations with Quadrature-Based Methods of Approximation," Journal of Econometrics, 273-308

[8] Hansen, L. 1982. "Large Sample Properties of Generalized Method of Moments Estimators," Econometrica 50, 1029-1054.

[9] Griffith, R.M. 1949. "Odds Adjustments by American Horse-Race Bettors." The American Journal of Psychology 62 (2):290–294.

[10] Guryan, Jonathan and Melissa S. Kearney. "Gambling at Lucky Stores: Evidence from State Lottery Sales," American Economic Review 98 (1), March 2008, 458-73.

[11] Jullien, B. and B. Salanie. 2000. "Estimating Preferences under Risk: The Case of Racetrack Bettors," Journal of Political Economy 108, 503-530.

[12] Kahneman, Daniel, and Amos Tversky. 1979. "Prospect theory: An Analysis of Decision under Risk," Econometrica 47, 263-291.

[13] Kearney, Melissa Schettini. 2005. "State Lotteries and Consumer Behavior." Journal of Public Economics 89, 2269-2299.

[14] Koszegi, Botond and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences," The Quarterly Journal of Economics 521(4), November: 1133-1165.

[15] List (2003)

[16] List, John. 2004. “Neoclassical Theory versus Prospect Theory: Evidence from the Marketplace,” Econometrica 72(2), March: 615-625.

[17] Newey, W.K. and K.D. West. 1982. “A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix,” Econometrica: 703-708

[18] Prelec, Drazen. 1998. “The Probability Weighting Function.” Econometrica 66(3): 497-527.

[19] Rabin, Matthew. 1996. “Psychology and Economics,” Mimeo, University of California at Berkeley, September.

[20] Sauer, Raymond D. 1998. “The Economics of Wagering Markets.” Journal of Economic Literature 36 (4):2021–2064.

[21] Singleton, K.J. 1985. “Testing specifications of economic agents’ intertemporal optimum problems in the presence of alternative model,” Journal of Econometrics: 391-413

[22] Smith, R.J. 1992. “Testing specifications of economic agents’ intertemporal optimum problems in the presence of alternative model,” Econometrica: 973-980.

[23] Snowberg, Erik and Justin Wolfers. 2010. “Explaining the Favorite-Longshot Bias: Is it Risk-Love or Misperceptions?” Journal of Political Economy.

[24] Tversky, Amos and Daniel Kahneman. 1992. “Advances in Prospect Theory: Cumulative Representation of Uncertainty,” Journal of Risk and Uncertainty 5, 297-323.

[25] Weitzman, Martin. 1965. “Utility Analysis and Group Behavior: An Empirical Study.” The Jour- nal of Political Economy 73 (1):18–26.

A Description of Lottery Games

Powerball is a multi-state lottery game offered. Players select 5 numbers from a field of 59 white balls and one "bonus" number from a field of 39 red balls. Drawings occur twice a week. Powerball's advertised jackpot starts at $20 million (annuity) and can roll into the hundreds of millions over many drawings. The jackpot is won by matching all five white balls in any order and the red Powerball. Payment can be received as either an annuitized prize paid out over 29 years (30 payments) or a lump sum payment. Each ticket costs $1. The odds of winning the top prize are 1 in 195,249,054.

Lotto South is a multi-state lottery game that ran in Georgia, Kentucky, and Virginia between 2001 and 2006. It is a typical "pick 6" type game, in which player picked six distinct numbers from 1 to 49. Lower prize tiers are awarded for matching fewer than six numbers. Typically, three matching numbers wins about four dollars, four matching numbers win $50 and five numbers win anywhere from $600 to $1000. All of the estimated jackpots are calculated by the lottery agencies using pari-mutual estimates, based on the expectation of how many tickets will sell in a given week. The odds of winning any prize in this game are 1 in 54. The odds of matching all six numbers to win the grand prize are 1 in 13,983,816.

Louisiana Lotto is an in-state lottery game that is played by picking six distinct numbers from 1 to 40. Tickets are $1 for one play. The jackpot starts at $250,000 and accumulates as long as no one wins. Smaller prizes are awarded for correctly picking five, four, and three numbers. The odds of winning the jackpot are 1 in 3,838,380.

Missouri Lotto is an in-state lottery game that is played by picking six distinct numbers between 1 and 44. Tickets are $1 for two plays. The jackpot prize starts at $1 million and accumulates until someone wins the top prize by correctly picking all six numbers. Smaller prizes are awarded for correctly picking five, four, and three numbers. The odds of winning the jackpot are 1 in 3,529,526.

Montana Cash is an in-state game played by selecting five distinct numbers from 1 through 37. Tickets are $1 for two plays. Jackpots start at a guaranteed $20,000 and accumulate each time the jackpot is not won. Drawings are held every Wednesday and Saturday night. Smaller prizes of $200 and $5 are offered for selecting four and three numbers, respectively. The odds of winning any prize are 1 in 43. The odds of winning the jackpot are 1 in 217,949.

Oregon Megabucks is an in-state lottery game played by selecting six distinct numbers from a set of 48. Tickets cost $1 for two plays. Drawings are held on Monday, Wednesday, and Saturday evenings. Megabucks jackpots start at $1 million and grow until someone wins. Pari-mutual prizes are offered for correctly picking five and four numbers; a free ticket is awarded for picking two correct numbers. The odds of winning the jackpot are 1 in 6,135,756.

Wisconsin Megabucks is an in-state lottery game played by selecting six distinct numbers from 1 through 49. Tickets are \$1 for two plays. Drawings are held on Wednesday and Saturday evenings. Megabucks jackpots grow until someone wins. Smaller prizes of \$500, \$30, and \$2 are offered for selecting five, four, and three numbers correctly, respectively. The odds of winning the jackpot are 1 in 6,991,908.

B Moments of a Lottery

The first three moments of a lottery can be computed as follows.

Expected value of the lottery :

$$\begin{aligned} \mathcal{E}(p,R) &\equiv pR + (1-p)\,0 && (40) \\ &= pR && (41) \end{aligned}$$

Variance of the lottery :

$$\begin{aligned} \mathcal{V}(p,R) &\equiv p\,(R-pR)^2 + (1-p)\,(0-pR)^2 && (42) \\ &= p\,(1-p)\,R^2 && (43) \end{aligned}$$

Skewness of the lottery :

$$\begin{aligned} \mathcal{S}(p,R) &\equiv \frac{p\,(R-pR)^3 + (1-p)\,(0-pR)^3}{\mathcal{V}(p,R)^{3/2}} && (44) \\ &= [p\,(1-p)]^{-1/2}\,(1-2p) && (45) \end{aligned}$$

Table 1 - Lottery Games: Descriptive Statistics

Pair	Alternative Lottery	State	Multistate	Period (number of weeks)	Mean Advertized Jackpot	Mean Actual Jackpot	Minimum Actual Jackpot	Maximum Actual Jackpot	Odds of Winning 1 in …
	Powerball		Yes	1/1992-12/2005 (724)	21.75	21.81	2.66	173.82	146,107,962
1	Lotto South	GA	Yes	6/1995-8/1996 (62)	2.71	2.71	1.09	6.60	9,345,794
2	Lotto South	KY	Yes	1/1992-12/2005 (621)	5.05	5.05	0.69	25.54	6,761,539
3	Lousiana Lotto	LA	No	3/1995-12/2005 (559)	1.58	1.58	0.24	17.33	3,831,418
4	Missouri Lotto	MO	No	1/1993-12/2005 (674)	3.08	3.08	0.92	19.18	7,042,254
5	Montana Cash	MT	No	1/1992-12/2005 (724)	0.07	0.07	0.02	0.29	217,865
6	Megabucks	OR	No	1/1997-12/2005 (468)	3.63	3.63	0.76	17.04	6,134,969
7	Megabucks	WI	No	7/1992-12/1999 (201)	2.38	2.38	0.57	10.27	6,993,007

Notes: Jackpot amounts are in million US dollars. In the Period column we list the end points of the samples for each lottery in month/year format and list the total number of weekly observations available. In many cases the number of weeks available is less than the total number of weeks in the period due missing data. Odds of winning refers to the odds of winning the jackpot.

Table 2 - Lottery Pairs: Statistical Properties and Sales

Pair	Average EV		% of times EV1 > EV2	Average STDEV		% of times STDEV1 > STDEV2	Share of Sales of Lottery 1	
	Lottery 1	Lottery 2		Lottery 1	Lottery 2		Average	Minimum
1	0.30	0.29	53%	2251	886	84%	56%	37%
2	0.27	0.60	23%	2459	1692	65%	66%	4%
3	0.28	0.41	53%	2571	807	89%	69%	38%
4	0.28	0.44	29%	2510	1160	76%	74%	35%
5	0.28	0.33	39%	2410	154	100%	73%	45%
6	0.27	0.59	24%	2591	1467	72%	47%	18%
7	0.26	0.34	44%	2061	901	80%	67%	40%

Note: Lottery 1 refers to Powerball. Shaded cells indicate higher mean and lower standard deviation.

Table 3 - Estimates of Model 1 (Risk Love)

(a) Model 1A (Constant Relative Risk Aversion)

Pair	gamma	p-value	J-stat	p-value
1	-0.02	0.81	1.3782	0.85
2	**-0.48**	**0.00**	3.2604	0.52
3	-0.14	0.11	1.6895	0.79
4	**-0.22**	**0.00**	1.3388	0.85
5	**-0.04**	**0.00**	2.6806	0.61
6	**-0.39**	**0.00**	6.8880	0.14
7	**-0.16**	**0.01**	3.7956	0.43

(a) Model 1B (Constant Absolute Risk Aversion)

Pair	theta	pvalue	J-stat	p-value
1	0.03	0.83	1.5627	0.82
2	**-0.25**	**0.00**	5.4292	0.25
3	-0.10	0.09	2.1791	0.70
4	**-0.16**	**0.00**	3.6827	0.45
5	**-0.08**	**0.00**	2.5024	0.64
6	**-0.20**	**0.00**	**12.3762**	**0.01**
7	**-0.19**	**0.03**	4.6709	0.32

Notes: Gamma is the estimated CRRA parameter and theta is the estimated CARA parameter. Red boldface indicates statistical significance at 5% level. J-stat is Hansen's (1982) test of over-identifying restrictions. p-value for J-stat tests the over-identifying restrictions using the chi-square distribution with four degrees of freedom.

Table 4 - Estimates of Model 2 (Prelec Weighting)

Pair	alpha	p-value	J-stat	p-value
1	**0.945**	**0.00**	1.8741	0.76
2	**0.782**	**0.00**	4.4025	0.35
3	**0.931**	**0.00**	3.7170	0.45
4	**0.864**	**0.00**	3.2340	0.52
5	**0.965**	**0.00**	5.0374	0.28
6	**0.840**	**0.00**	2.5956	0.63
7	**0.916**	**0.00**	6.4832	0.17

Notes: Alpha is the estimated Prelec weighting function parameter. The third column reports the p-value for testing alpha equals to one. Red boldface indicate statistical significance at 5% level. J-stat is Hansen's (1982) test of over-identifying restrictions. p-value for J-stat tests the over-identifying restrictions using the chi-square distribution with four degrees of freedom.

Table 5 - Model Comparison - Risk Love versus Over-weighting

(a) J-stats

Pair	Model 1A	Model 1B	Model 2
1	1.38	1.56	1.87
2	3.26	5.43	4.40
3	1.69	2.18	3.72
4	1.34	3.68	3.23
5	2.68	2.50	5.04
6	6.89	12.38	2.60
7	3.80	4.67	6.48
Sum	**21.03**	**32.40**	**27.34**

(b) Condorcet Table Based on J-Stats

		Model 1A	Model 1B	Model 2
Winner	**Model 1A**	-	6	6
	Model 1B	1	-	4
	Model 2	1	3	-

(c) Condorcet Table Based on Singleton Tests

		Model 1A	Model 1B	Model 2
Winner	**Model 1A**	-	1	3
	Model 1B	0	-	2
	Model 2	1	1	-

Notes: Panel (a) reports the J-stats from the estimation of each model. The last row sums these J-stats across pairs. Panel (b) reports the number of times the model in rows wins versus the model in columns. For example the entry in the (1,2) cell show that Model 1A has a lower J-stat than model 1B for 6 of the 7 pairs. Gray shading reflects the model that wins the bilateral comparison. Panel (c) reports the number of times the model in the row wins versus the model int the column in the bilateral Singleton tests. Model *i* is considered to win versus model *j* if the p-value in the *i* vs. *j* comparison is greater than 0.10 and the p-value in the *j* vs. *i* comparison is less than 0.10.

Table 6 - Estimates of Model 3 (Entertainment) and its Comparison to Other Models

(a) Model Estimates

Pair	lambda(1)	p-value	lambda(2)	p-value	Wald p-value	J-stat	p-value
1	0.93	**0.00**	0.83	**0.00**	**0.02**	0.1064	0.99
2	0.18	0.17	-0.04	0.64	**0.00**	3.3114	0.35
3	1.21	0.10	0.71	0.12	0.07	0.0735	0.99
4	0.15	0.85	0.01	0.98	0.70	0.8728	0.83
5	0.10	0.54	0.03	0.74	0.33	2.4356	0.49
6	1.66	**0.00**	1.50	**0.00**	**0.00**	5.3363	0.15
7	0.69	**0.01**	0.42	**0.02**	**0.00**	1.8213	0.61
					Sum	13.96	

(b) Condorcet Table Based on J-stats

		Model 1A	Model 1B	Model 2	Model 3
Winner	**Model 1A**	-	6	6	1
	Model 1B	1	-	4	0
	Model 2	1	3	-	1
	Model 3	6	7	6	-

(c) Condorcet Table Based on Singleton Tests

		Model 1A	Model 1B	Model 2	Model 3
Winner	**Model 1A**	-	1	3	1
	Model 1B	0	-	2	0
	Model 2	1	1	-	1
	Model 3	0	3	5	-

Notes: In panel (a), lambda(1) and lambda(2) are the estimated entertainment parameters (multiplied by 10^6) for the first and second lottery, respectively. p-value columns report the p-values for the significance of the respective lambda estimate. Red boldface indicate statistical significance at 5% level. The Wald column tests if the two entertainment parameters are equal to each other. J-stat is Hansen's (1982) test of over-identifying restrictions. p-value for J-stat tests the over-identifying restrictions using the chi-square distribution with three degrees of freedom. For panels (b) and (c) see the notes to Table 5.

Table 7 - Estimates of Model 4 (Log Utility with Prelec Weighting) and its Comparison to Other Models

(a) Model Estimates

Pair	alpha	p-value	J-stat	p-value
1	**0.595**	**0.00**	1.5506	0.82
2	**0.490**	**0.00**	1.5328	0.82
3	**0.615**	**0.00**	1.5286	0.82
4	**0.541**	**0.00**	1.3140	0.86
5	**0.689**	**0.00**	1.1318	0.89
6	**0.518**	**0.00**	5.2875	0.26
7	**0.595**	**0.00**	4.0625	0.40
		Sum	16.41	

(b) Condorcet Table Based on J-stats

		Model 1A	Model 1B	Model 2	Model 3	Model 4
Winner	**Model 1A**	-	6	6	1	2
	Model 1B	1	-	4	0	0
	Model 2	1	3	-	1	1
	Model 3	6	7	6	-	4
	Model 4	5	7	6	3	-

(c) Condorcet Table Based onSingleton Tests

		Model 1A	Model 1B	Model 2	Model 3	Model 4
Winner	**Model 1A**	-	1	3	1	0
	Model 1B	0	-	2	0	0
	Model 2	1	1	-	1	1
	Model 3	0	3	5	-	1
	Model 4	0	2	2	1	-

Notes: In panel (a) Alpha is the estimated Prelec weighting function parameter. The third column reports the p-value for testing if alpha equals to one. Red boldface indicate statistical significance at 5% level. J-stat is Hansen's (1982) test of over-identifying restrictions. p-value for J-stat tests the over-identifying restrictions using the chi-square distribution with four degrees of freedom. See Table 5 for notes for panels (b) and (c). In panel (c), grained shading shows a tie.

Table 8 - Estimates of Model 5
(General CRRA with Prelec)

(a) Model Estimates

Pair	gamma	alpha	p-value	J-stat	p-value
1	0.8	**0.693**	**0.00**	1.5025	0.83
2	1.0	**0.490**	**0.00**	1.5328	0.82
3	1.3	**0.467**	**0.00**	1.1055	0.89
4	0.9	**0.590**	**0.00**	1.2840	0.86
5	7.2	**0.016**	**0.00**	0.4812	0.98
6	1.0	**0.518**	**0.00**	5.2575	0.26
7	6.1	**0.000**	**0.00**	1.0787	0.90
			Sum	12.24	

Notes: Gamma is the fixed (i.e. not estimated) CRRA parameter. Alpha is the estimated Prelec weighting function parameter. The third column reports the p-value for testing if alpha equals to one. Red boldface indicate statistical significance at 5% level. J-stat is Hansen's (1982) test of over-identifying restrictions. p-value for J-stat tests the over-identifying restrictions using the chi-square distribution with four degrees of freedom.

Table 9 - Comparison of All Models

(a) J-stats of Various Models

Pair	Model 1A	Model 1B	Model 2	Model 3	Model 4	Model 5
1	1.38	1.56	1.87	0.11	1.55	1.50
2	3.26	5.43	4.40	3.31	1.53	1.53
3	1.69	2.18	3.72	0.07	1.53	1.11
4	1.34	3.68	3.23	0.87	1.31	1.28
5	2.68	2.50	5.04	2.44	1.13	0.48
6	6.89	12.38	2.60	5.34	5.29	5.29
7	3.80	4.67	6.48	1.82	4.06	1.08
Sum	**21.03**	**32.40**	**27.34**	**13.96**	**16.41**	**12.27**

(b) Condorcet Table Based on J-Stats

		Model 1A	Model 1B	Model 2	Model 3	Model 4	Model 5
Winner	Model 1A	-	6	6	1	2	1
	Model 1B	1	-	4	0	0	0
	Model 2	1	3	-	1	1	1
	Model 3	6	7	6	-	4	3
	Model 4	5	7	6	3	-	1
	Model 5	6	7	6	4	5	-

Condorcet Ranking : Model 5, Model 3, Model 4, Model 1A, Model 1B, Model 2

(c) Condorcet Table Based on Singleton Test

		Model 1A	Model 1B	Model 2	Model 3	Model 4	Model 5
Winner	Model 1A	-	1	3	1	0	0
	Model 1B	0	-	2	0	0	0
	Model 2	1	1	-	1	1	1
	Model 3	0	3	5	-	1	0
	Model 4	0	2	2	1	-	0
	Model 5	1	3	3	1	1	-

Condorcet winner : Model 5

Notes: See the notes to Table 5.

Table Appendix 1 - Model Comparisons - p-values of Singleton Test Statistics

Pair	1A vs 1B	1B vs 1A	1A vs 2	2 vs 1A	1A vs 3	3 vs 1A	1A vs 4	4 vs 1A	1A vs 5	5 vs 1A
1	0.30	0.24	0.71	0.35	0.26	0.76	0.81	0.46	0.79	0.48
2	0.35	0.12	0.42	**0.08**	0.49	0.80	0.30	0.98	0.30	0.98
3	0.93	0.47	0.66	0.26	0.20	0.88	0.41	0.33	0.60	0.76
4	0.61	0.12	0.31	**0.08**	0.46	0.89	0.38	0.32	0.37	0.32
5	0.58	0.91	0.60	**0.09**	0.69	0.80	0.22	0.92	0.21	0.69
6	0.41	**0.02**	**0.04**	0.40	0.13	**0.04**	0.20	0.95	**0.02**	0.53
7	0.81	0.30	0.60	0.13	0.10	0.19	0.54	0.73	**0.09**	0.99

Pair	1B vs 2	2 vs 1B	1B vs 3	3 vs 1B	1B vs 4	4 vs 1B	1B vs 5	5 vs 1B
1	0.73	0.50	0.23	0.78	0.77	0.55	0.76	0.59
2	0.17	**0.07**	0.10	0.47	**0.08**	0.57	**0.08**	0.57
3	0.35	0.21	0.14	0.79	0.57	0.68	0.15	0.72
4	0.16	0.15	**0.08**	0.63	0.15	0.45	0.14	0.43
5	0.59	**0.09**	0.93	0.67	0.25	0.71	0.25	0.72
6	**0.00**	0.69	**0.02**	0.28	**0.01**	0.97	**0.00**	0.73
7	0.40	0.16	**0.06**	0.46	0.34	0.99	**0.05**	0.99

Pair	2 vs 3	3 vs 2	2 vs 4	4 vs 2	2 vs 5	5 vs 2
1	0.17	0.75	0.76	0.96	0.70	0.88
2	**0.07**	0.98	**0.06**	0.87	**0.06**	0.87
3	**0.05**	0.79	0.32	0.52	0.11	0.42
4	**0.09**	0.49	0.16	0.93	0.16	1.00
5	**0.08**	0.15	**0.04**	0.75	**0.04**	0.95
6	0.50	**0.02**	0.32	**0.05**	0.23	0.59
7	**0.02**	0.51	0.10	0.83	**0.02**	0.85

Pair	3 vs 4	4 vs 3	3 vs 5	5 vs 3
1	0.77	0.22	0.76	0.23
2	0.25	0.61	0.25	0.61
3	0.85	0.22	0.79	0.30
4	0.58	0.32	0.55	0.32
5	0.22	0.87	0.23	0.75
6	**0.03**	0.31	**0.02**	0.61
7	0.22	**0.09**	0.20	0.43

Pair	4 vs 5	5 vs 4
1	0.89	0.97
2	0.73	0.73
3	0.51	0.92
4	0.94	0.96
5	0.53	0.87
6	**0.04**	0.61
7	**0.07**	0.93

Notes: The table reports the p-value for testing the first model versus the second model for each pair. Red boldface denotes numbers less than 0.10.

Figure 1

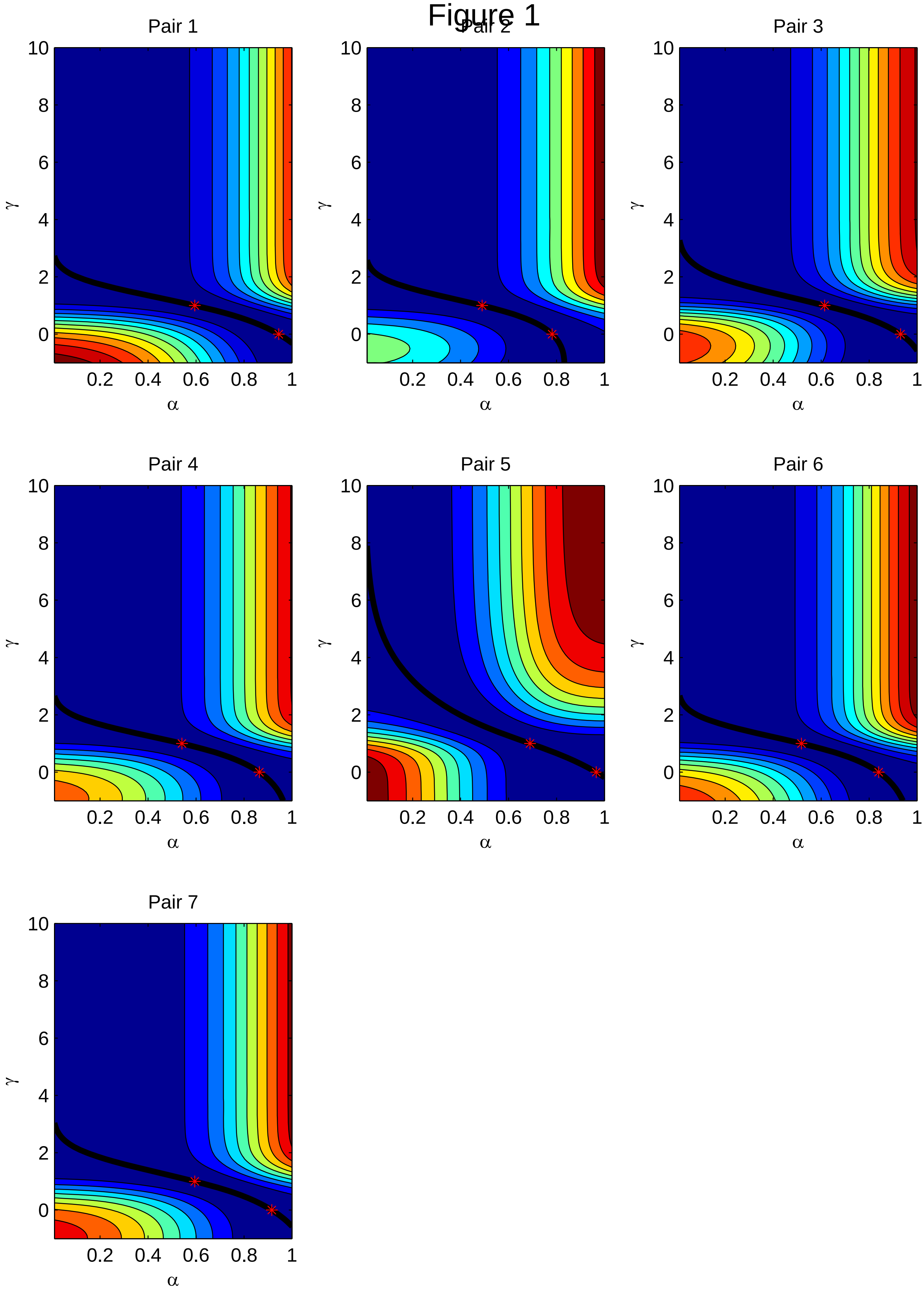
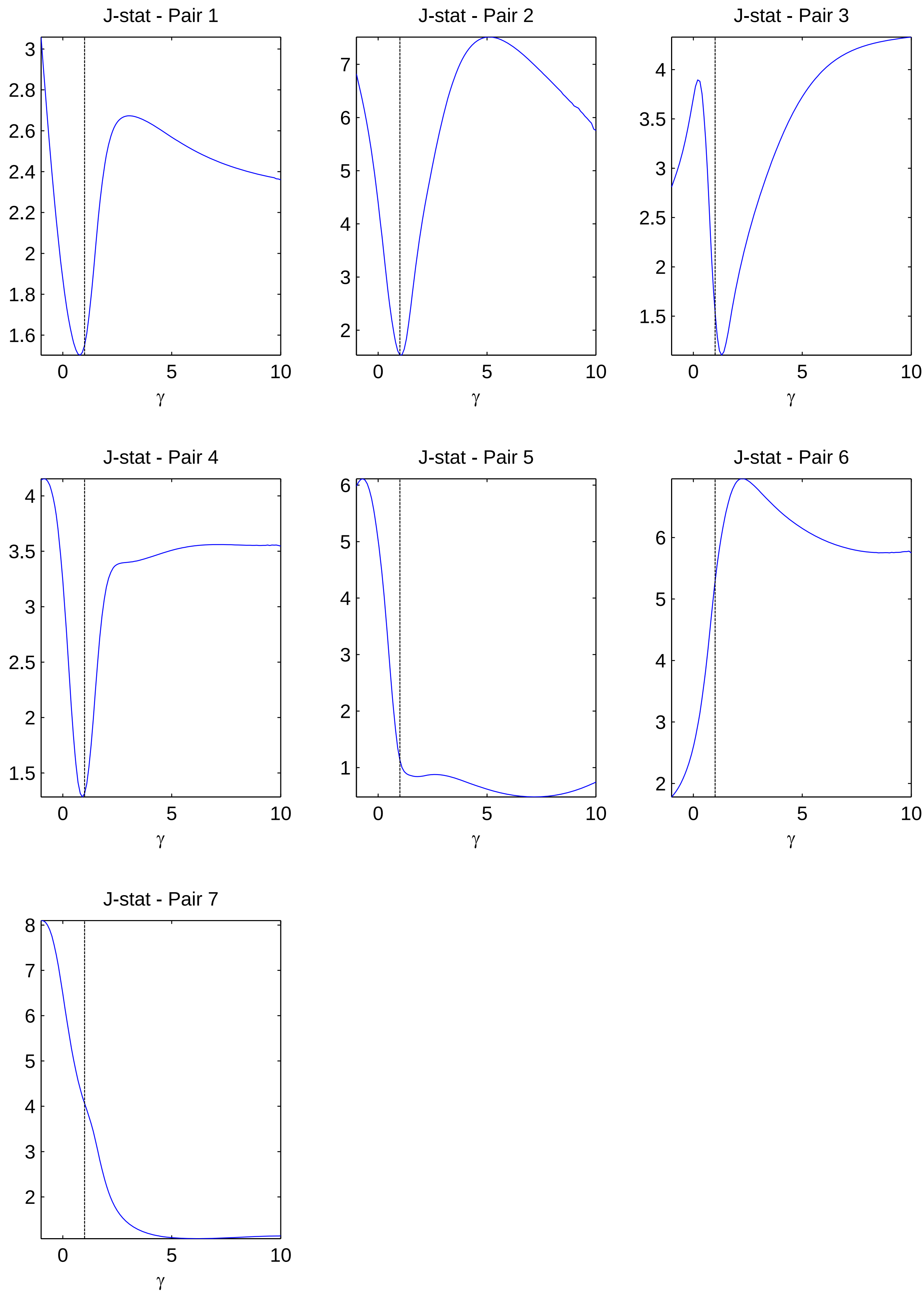

Pair 1
Pair 2
Pair 3
Pair 4
Pair 5
Pair 6
Pair 7
γ
α
10
8
6
4
2
0
0.2
0.4
0.6
0.8
1

Figure 2


J-stat - Pair 1
J-stat - Pair 2
J-stat - Pair 3
J-stat - Pair 4
J-stat - Pair 5
J-stat - Pair 6
J-stat - Pair 7
γ