Estimating Risk Preferences from a Large Panel of Real-World Betting Choices*

Angie Andrikogiannopoulou†
Princeton University

November 2010

Abstract

We examine individual decision-making under risk using a unique panel dataset of consumer betting activity in an online sportsbook. Our econometric model accounts for individual heterogeneity in risk preferences both between and within rational and behavioral theories of choice under risk. Besides being able to estimate individual-level utility parameters with relative accuracy, the panel aspect of our dataset also allows us to test for possible state-dependence in agents' risk-taking behavior which would arise if, for example, subsequent risk-taking is affected by previous bet outcomes. We find that the majority of bettors in our sample depart from full rationality when presented with complex real-life decisions. However, there is significant heterogeneity across individuals regarding the aspect of rational preferences they violate: state-dependence, probability weighting, loss aversion and utility curvature vary across bettors. Methodologically, we use a discrete choice framework and estimate a multinomial mixed logit model using Bayesian econometrics techniques. Our findings have significant implications for micro-founded models in economics and finance.

Keywords: Risk Preferences, Risk Aversion, Prospect Theory, State Dependence, Bayesian Estimation, Discrete Choice, Mixture Models
JEL Classification: D11, D12, D81, G00

*For helpful comments, I would like to thank my advisor, Hyun Shin, as well as Bo Honore, Andriy Norets, Wei Xiong and seminar participants at Princeton University.

†Department of Economics, Fisher Hall 001, Princeton University, Princeton, NJ 08544, e-mail: aandriko@princeton.edu, http://www.princeton.edu/∼ aandriko

1 Introduction

Individual risk preferences are a fundamental building block of any model in economics and finance and has therefore been at the heart of academic research for many decades. On the theoretical front, a wide variety of theories of choice under risk have been advanced but on the empirical side our knowledge about individual decision-making behavior is still very limited. Despite the abundance of experimental evidence in this area, the empirical investigation of people's attitudes towards risk in real-world scenarios that are of practical interest to economics has largely been hindered by data constraints, so fundamental questions remain open. Which theory better explains people's attitudes towards risk, what the magnitudes of different utility parameters are, how heterogenous individual risk preferences are and whether they are stable over time or path-dependent are some of these questions. In this study we examine these issues using for the first time a rich *panel* of *real-life* individual choices. In particular, we utilize a unique dataset that includes all the bets placed by sports gamblers through an online bookmaker over the past 3.5 years.

Sports gambling has lately evolved into a hugely popular activity around the globe with hundreds of bookmakers offering a wide range of betting opportunities on various sporting events and millions of people risking significant amounts of their income on the unknown future outcome of these events. The popularity of this and other forms of gambling, including casino games, lottery tickets, and racetrack betting, has led some people to believe that a taste for gambling is an inherent part of human nature.[1] Having said that, understanding how people behave in this setting will help us uncover common aspects of individual risk preferences and contribute to the totality of empirical evidence in this field.

Our sports betting data have several important advantages over the datasets that have been used so far in the literature of estimating risk preferences. First, contrary to experimental studies, we observe real-world choices of regular market participants who risk varying amounts of their own money. Second, we use data at the individual bettor level instead of aggregate price data that have been previously used in the gambling domain (e.g. Jullien and Salanie (2000), Golec and Tamarkin (1998)). Thus, we can refrain from the representative agent assumption that these studies are forced to make, and we are in fact able to test whether treating all individuals as one representative agent would be a reasonable approach or not. Third, we are the first to use a relatively large panel which allows us to i) estimate individual-level risk preference parameters with relative accuracy, ii) test for the presence of heterogeneity in risk preferences across individuals, and iii) examine possible path-dependence in individual risk preferences which could arise, for example, if people's behavior is

[1] For example, the sociologist William Thomas stated that gambling is an instinct that "is born in all normal persons. It is one expression of a powerful reflex, fixed far back in animal experience".

affected by their previous betting performance. Finally, it is important to note that sports betting decisions share significant similarities with stock trading decisions, but at the same time they have two desirable characteristics that make them better-suited for the empirical analysis of individual risk preferences. In particular, contrary to the stock market where people's investment horizon and beliefs are difficult to observe, in our case it is relatively easy to construct the distribution of the "lottery" that people choose when placing their bets since i) sports bets reach an observable terminal value when the sporting event in question takes place; and ii) the odds quoted by the bookmaker provide a good approximation for the probabilities with which people evaluate their bets. It is therefore clear that sports betting constitutes an idealized laboratory setting for the analysis of the preferences of a group of people who resemble regular stock market participants.

The popularity of gambling and the actual behavior of gamblers has been usually explained by the cumulative prospect theory (CPT) of Tversky and Kahneman (1992). One of the key features of this theory, namely that people systematically distort the probabilities of events by overweighting small probabilities and underweighting large probabilities, can explain why bettors prefer the positively-skewed lotteries that are commonly encountered in a sportsbook. Barberis (2010) shows that prospect theory can also explain the popularity of 50:50 casino bets and that it captures many other features of the observed gambling behavior. Therefore, a natural starting point for our analysis is the characterization of risk preferences within the behavioral choice paradigm of cumulative prospect theory. We estimate the distributions of the utility parameters of the CPT specification, namely the utility curvature, the loss aversion and the probability weighting parameter. Our findings suggest that for the average bettor in our sample i) the value function is mildly concave (convex) in the region of gains (losses) indicating that bettors are mildly risk averse (risk loving) over gains (losses); ii) the value function has a sharp kink at the origin which indicates greater sensitivity to losses than to gains (loss aversion) and matches the empirical observation that the average bettor in our sample tends to reject 50 : 50 bets; and iii) the deviation from linear probability weighting is mild which matches the empirical observation that bettors favor lotteries with medium positive skewness but typically reject the multitude of very positively-skewed lotteries available in the sportsbook.

Nevertheless, a striking feature of the estimated distributions of preference parameters is that substantial heterogeneity is present in all parameters, indicating that economic analysis (e.g., in an asset-pricing model) based on a single estimate might be inappropriate and may lead to biased results. In particular, it seems that there is a continuum of prospect theory agents in the population, whose choices are characterized by different parameter combinations. An examination of the correlations of the utility parameters reveals that the three features of prospect theory tend to move together. This indicates that it might be possible to classify individuals into discrete types: individuals whose choices are fully characterized by prospect theory preferences but also individuals who essentially act as expected value maximizers.

In the next step, we exploit the panel aspect of our dataset to examine the presence of state dependence in (some) bettors' risk preferences. Such state dependence exists, for example, if bettors' behavior is affected by the outcomes of their previous bets as suggested by phenomena such as the "house money effect" or the disposition effect. These two effects imply opposite predictions regarding people's behavior after prior gains versus losses: The "house money effect" suggests that prior gains (losses) increase (decrease) subsequent risk-taking and has been documented so far mainly in experimental studies (e.g., Thaler and Johnson (1990)) and quasi-experimental studies such as game shows (e.g., Post et al. (2008)), while the disposition effect predicts the opposite and has been documented in datasets of individual investor trading activity. Prospect theory offers a natural framework to study these issues since gains and losses are evaluated relative to a reference point and the framing of outcomes can affect this reference point. Indeed, our findings suggest that prior outcomes significantly affect subsequent risk-taking, emphasizing that people integrate the outcomes of successive gambles and that individual loss aversion changes depending on previous betting performance. In particular, our estimation results so far indicate that bettors become effectively more risk loving after losses which is consistent with the disposition effect observed in the stock market.

Although prospect theory offers a coherent framework for analyzing betting choices, researchers have noted that gambling behavior can also be reconciled with the rational choice paradigm of expected utility theory (EUT). Several approaches have been proposed to explain the tendency of risk averse people to accept unfair gambles: Friedman and Savage (1948) suggest that local convexities in an otherwise concave utility function can create a preference for skewness, while Conlisk (1993) explains gambling behavior by introducing a direct utility of gambling that captures the entertainment/excitement associated with casino games. Finally, a stochastic choice behavior model that allows for variation in the preferences of a single individual when faced repeatedly with the same choice situation, can explain why people are not always observed to choose the alternative with the highest utility (here the option not to gamble). On top of that, the availability of near safe lotteries inside the sportsbook further shrinks the utility difference between entering the casino and not. The ability to reconcile gambling with EUT coupled with the evidence that it is a very widespread activity,[2] leads us to question the conventional wisdom that only people with CPT-type preferences gamble, and hence that our sample is drawn exclusively from the population of prospect theory agents. As a result, we also examine individual heterogeneity in risk preferences *between* rational and behavioral theories of choice, by estimating a finite mixture model of the

[2]The Economist, 8th July 2010, mentions that "[i]n 2007, nearly half of America's population and over two-thirds of Britain's bet on something or other", and that "[h]underds of millions of lottery tickets are sold every week."

EUT and the CPT. Allowing for individual heterogeneity across theories rather than assuming that all individuals have the same type of preferences is important in several respects that have been ignored by the literature so far. First, it enables us to estimate the proportion of individuals that behave according to each theory and therefore test the relative strength of competing theories in explaining individual behavior. Second, from an econometric point of view, if there are more than one latent populations from which the observed sample is drawn, estimating a specification that assumes just one population can lead to misleading inferences with respect to the estimated risk parameters. That is, simply allowing for individual heterogeneity within specifications and sequentially estimating different specifications on the whole sample, as it is commonly done, can neither yield valid parameter estimates, nor be used to prove or disprove the validity of a specific utility class. Our results point to the presence of some heterogeneity across individuals in the utility theory that best describes their observed choices with the mean probability of being classified in the population of EUT agents being estimated between 15%-20%.

Taken together, our estimation results are important for testing the validity of the *assumptions* rather than the *predictions* of prominent theoretical models in economics and finance. In the finance domain, for example, our findings could be used to i) test whether people faced with complex real-life decisions actually behave as hypothesized by theoretical models that have been developed to understand various "puzzling" financial phenomena, such as the equity premium puzzle; ii) to confirm that explicitly modeling investor heterogeneity both within and between rational and behavioral theories is important; iii) to calibrate existing research and guide future research involving models with proportions of rational and behavioral agents, time-varying risk preferences, etc..

It is important to note that when estimating the distribution of risk preferences from our panel dataset, play frequency is an important confounding factor. It can be informative of the frequency with which individuals *have* the opportunity to log into the sportsbook, which depends on exogenous factors such as the amount of free time available to them, but it can also be informative of the frequency with which they choose to *accept* the opportunity to bet, which depends on their risk preferences. To estimate risk preferences that take into account the confounding effect of play frequency, we model individual observed and unobserved heterogeneity in both risk preferences and the probability with which individuals get the chance to bet on each day. We are able to identify both effects by exploiting variation in bettors' observed play frequency together with variation in individuals' betting choices conditional on their decision to play.

A standard criticism of estimating preferences from betting choices is that the same choices could be explained with several combinations of risk preferences and subjective beliefs. Thus, the fact that different people make different choices might be either due to the fact that they have different risk preferences or due to the fact that they hold different beliefs regarding the various outcomes' probabilities. Which is the primary reason why people engage in sports and racetrack betting has divided authors over the years and given rise to two strands in the literature that aims to explain people's betting behavior. Several authors such as Weitzman (1965), Ali (1977), Quandt (1986) and Jullien and Salanie (2000), have provided empirical evidence of racetrack gamblers' risk preferences under the assumption that bettors know the objective win probabilities. In contrast, studies like Figlewski (1979), Bassett (1981) and Woodland and Woodland (1991) have argued that sports and racetrack betting markets primarily exist due to a divergence of beliefs among individuals. In our baseline study, we follow the first school of thought, i.e., we use the observed betting choices to estimate individual risk preferences, under the assumption that the probabilities people use to evaluate the gambles available in the sportsbook are equal to those quoted by the bookmaker.[3] In an extension, we also consider the possibility that the heterogeneity in people's observed betting behavior may be partially due to belief heterogeneity. In particular, we allow for a random, but systematic, individual- and event-specific "error" in probability assesments, which could account for possible deviations from the quoted probabilities due to optimism, skill/information, and/or mistakes. Finally, by tracking individual risk-adjusted betting performance, we verify that no bettors in our sample exhibit significantly positive skill in picking their bets while a few bettors exhibit significantly negative bet-picking skill.

Methodologically, our analysis proceeds in three steps. First, we represent the lottery chosen by each bettor on each day, by constructing all possible combinations of payoffs of all bets placed during that day, and calculating their corresponding probabilities. Second, we approximate the choice set from which the observed lotteries are chosen by choosing a set of representative lotteries from the set of all day lotteries chosen at least once by any bettor in our sample. This choice set is then augmented with a safe lottery representing the option to restrain from betting on that day. Finally, we use a multinomial mixed logit framework to model bettors' choices among the available alternatives and develop a finite latent class model to allow the coexistence of agents with different utility types. We estimate the resulting models using Bayesian econometrics by applying Markov Chain Monte Carlo (MCMC) techniques and in particular the Metropolis-Hastings within Gibbs sampling algorithm.

The remainder of the paper is organized as follows. Section 2 discusses the relation to the literature. Section 3 describes the data and the analysis thereof. Section 4 lays out the econometric

[3]Depending on how the bookmaker sets the odds, these could be the probabilities that clear the market, the probabilities that maximize the bookmaker's profits or the probabilities that accurately predict the game outcomes. The price-setting behavior of the bookmaker is beyond the scope of this study (for this, see Levitt (2004) and Paul and Weinbach (2008)). What matters for us is that the quoted probabilities are the probabilities that bettors are confronted with when placing their bets.

specification employed and Section 5 describes its estimation. Section 6 presents the results of our benchmark specification, Section 7 describes the setup and discusses the results of the finite mixture model and Section 8 presents various robustness checks.

2 Relation to Literature

2.1 Estimation of Risk Preferences

The vast majority of the earlier empirical literature on estimating individual risk preferences comes from laboratory experiments (Hey and Orme (1994), Holt and Laury (2002), Tversky and Kahneman (1992), Choi et al. (2007)) and hypothetical survey questions (Donkers et al. (2001)). In addition to these, a large body of quasi experimental empirical studies has used television game shows as a controlled natural experiment to analyze contestants' risk attitudes (Gertner (1993), Metrick (1995), Beetsma and Schotman (2001), Bombardini and Trebbi (2007), Post et al. (2008)). More recently, with access to real-world data becoming more widespread, a few studies examine the observed behavior of regular market participants in a variety of settings: insurance markets (Cicchetti and Dubin (1994), Cohen and Einav (2007), Barseghyan et al. (2010)), labor supply decisions (Chetty (2006)), person-to-person lending portfolio choices (Paravisini et al. (2010)). Closer to our domain, there exists a strand of empirical literature (Golec and Tamarkin (1998), Jullien and Salanie (2000), Kopriva (2009)) that analyzes aggregate price data, mainly from racetrack betting, to infer the preferences of a representative bettor.

Our sports betting data seem to have important advantages over the datasets that have been previously used in this area. First, we are observing real-world choices of normal market participants who risk varying amounts of their own money. In contrast, neither experimental subjects nor game show participants ever experience real losses, casting doubt on the extent to which individuals truthfully reveal information about their risk preferences in these settings. In addition to that, experiments are usually limited to small stakes and thus do not provide evidence about risk preferences towards prospects that could significantly affect one's lifetime wealth.

Second, we are using data at the individual bettor level instead of aggregate odds data that have been widely used in the past. We can therefore refrain from the representative agent assumption that the aforementioned studies are forced to make. Recent research (e.g., Cohen and Einav (2007)) finds significant heterogeneity in risk aversion across subjects, suggesting that treating all individuals as one representative agent can be misleading. In the sports betting setting, for example, aggregation would assign more weight to frequent and/or large bettors biasing the risk aversion estimates towards their behavior.

Third, we have a panel dataset with more than 35 observations for the average individual (and up to 270 for some individuals) which allows us i) to estimate the distribution of risk preferences in the population with some relative accuracy, ii) to examine the presence and extent of heterogeneity in risk preferences across individuals and iii) to examine possible state-dependence in bettors' risk preferences. Recent studies that use microdata from real-life decisions, such as portfolio or insurance choices, usually observe very few data points per subject. They are therefore not able to do this type of analysis at all or accurately enough.

Finally, an attractive feature of our dataset is that individuals make their choices out of a large choice set consisting of all the bets available in the sportsbook at any time, which considerably improves data fitting. To our knowledge, this is the first study that analyzes individual risk preferences using a rich, panel dataset from a real-life setting.

Research in this area should be studied with the following caveats in mind. First, a sample selection bias is present to different extents in most of the existing literature (e.g., experiments, racetrack betting, game shows) and raises caution when attempting to extrapolate the risk preference estimates to the whole population. In our case, it is clear that the sample of people who choose to bet on an online sportsbook might not be representative of the whole population, or even of the population of all bettors including those that visit land-based casinos. Furthermore, generalizations of risk preferences in other settings are difficult to make since, as Rabin and Thaler (2001) suggest, different decisions in life are taken under different conditions and might thus be subject to different utility specifications. People playing in the sportsbook might be primed to be more risk-seeking than they otherwise would be, since they are *gambling*, the same way that insurance buyers might be primed to be more risk averse, since they are buying *insurance*. In addition to these, it should be kept in mind that the sample of individuals under study might also engage in other risky activities that are unobservable to the researcher. In our case, it is impossible to know whether bettors hold accounts with multiple sports bookmakers, or even engage in other forms of gambling like casino games, lottery tickets etc. The implicit assumption here is that choices are made under narrow framing, so unobserved gambling and other types of choices do not affect the observed gamble chosen. Indeed, this would also be the assumption behind any experiment, otherwise one could still postulate that subjects integrate choices in the experiment with choices outside the experiment.

With these caveats in mind, this study takes a step towards better understanding the behavior of online sports gamblers. We hope that information from this domain could help us uncover aspects of preferences that we could not uncover from other drastically different domains like insurance choices. In fact, the insurance domain deals with people's fear of risk rather than their affinity to risk. Since risk is particular to the context in which it is measured, we believe that the current study can be viewed as complimentary to the existing research in the field, contributing to the totality of evidence on risk preferences.

2.2 Relation to Finance Literature

Sports betting markets share significant similarities with financial markets, both in terms of organization and in terms of the characteristics of their participants. In both cases there is a large number of potentially different types of participants, who bet on uncertain future events. The role of the sports bookmaker is similar to that of a market maker for securities: he accepts bets on all possible outcomes of an event and maintains a spread (vigorish) which will ensure a profit regardless of the outcome of the wager. Finally, information about sport events and odds is nearly as widely available as information about companies and stock prices. At the same time, sports betting has several desirable characteristics that make it an idealized laboratory setting for the empirical analysis of individuals' risk preferences. Its main advantages are i) that sports bets reach an observable terminal value when the sporting event in question takes place and ii) that the odds quoted by the bookmaker provide a good approximation for the probabilities with which people evaluate bets. Therefore, contrary to the stock market where it is difficult to know people's investment horizon and beliefs, in our case it is relatively easy to construct the distribution of the "lottery" that people choose when placing their bets.

In the finance literature, a large number of theoretical models have been developed to help us understand various "puzzling" financial phenomena about the aggregate stock market, the cross-section of average stock returns and investors' trading behavior among others. The papers in the asset pricing literature could be roughly separated along two dimensions based on i) whether they adopt a representative agent framework or explicitly model investor heterogeneity and ii) whether they assume individual rationality or not.

Along the first dimension, several papers have attempted to explain the empirically observed features of asset prices by explicitly modeling the interaction among heterogeneous investors. For instance, Chan and Kogan (2002) show that the countercyclicality of the aggregate risk premium can be the result of endogenous cross-sectional redistribution of wealth in an economy with multiple agents with heterogeneous risk aversion. Dumas (1989) and Wang (1996) examine the effect of investor preference heterogeneity on the dynamics and the term structure of interest rates.[4] In other models, heterogeneity in individual risk exposure arises from differences in beliefs rather than

[4] Other papers addressing financial phenomena under risk preference heterogeneity include Bhamra and Uppal (2010), Kogan, Makarov, and Uppal (2007), Ehling and Heyerdahl-Larsen (2010), among others.

differences in risk aversion across investors.[5]

Along the second dimension, a large number of theories that draw heavily on the experimental psychology literature, argue that some financial phenomena can be explained by investors not being fully rational, either in their preferences or in their beliefs. Non-standard preferences can be described by one or more of the following features: i) loss aversion, ii) probability weighting, iii) concavity/convexity of value function over gains/losses and/or iv) time-varying risk preferences. Organized on the basis of the feature of non-standard preferences that drives their results, behavioral finance papers provide explanations for a wide variety of financial "anomalies":

Benartzi and Thaler (1995)and Barberis et al. (2001) show that if investors are loss averse over annual changes in the value of their stock market holdings, they require a high average return on equity in order to hold the market supply, which helps us understand the high historical equity premium. Barberis et al. (2001) also show that if loss aversion changes over time depending on previous outcomes, fluctuations in the stock market caused by economic fundamentals will be amplified, sheding light on the puzzlingly high historical volatility of the stock market. Finally, loss aversion is essential in explaining some aspects of individuals' investing behavior, e.g. stock market non-participation, insufficient diversification, etc. (e.g. Barberis et al. (2006)).

Barberis and Huang (2008) show that in a financial market in which investors overweigh extreme events, assets with positively (negatively) skewed returns are going to be overpriced (underpriced) and will therefore earn low (high) returns, on average. The authors argue that this idea has many potential applications in finance, e.g. it can help us understand why IPO stocks have a poor long-term average return.

Several authors have argued that the concavity (convexity) of the value function over gains (losses) would predict the disposition effect, i.e. the puzzling tendency of individual investors to sell stocks that have risen in value, rather than dropped in value, since purchase. Barberis and Xiong (2009) show that the link between prospect theory and the disposition effect is less obvious if the value function is only mildly concave in the region of gains and investors' loss aversion is also taken into account.

Asset-pricing models use habit-formation preferences to explain the mean and countercyclicality of asset return risk premia found in the data (e.g. Constantinides (1990), Campbell and Cochrane (1999)). Mehra and Prescott (2003) point out that it is not clear whether investors actually have the huge time varying countercyclical variations in risk aversion postulated by these models. Brunnermeier and Nagel (2008) find evidence against time-varying risk preferences by examining microdata

[5]Notable papers that study the determination of asset prices and their dynamics under heterogeneity in beliefs include, but are not limited to, Harrison and Kreps (1978), Varian (1985), De Long et al. (1990), Harris and Raviv (1993), Detemple and Murthy (1994), Zapatero (1998), and Basak (2000).

on how households allocate their wealth between risky and riskless assets.

Since various assumptions about individual behavior have been hypothesized in these studies to explain a number of facts about financial markets, it becomes essential to empirically investigate whether people actually behave according to these assumptions. This is what we attempt to do in this study. We will therefore revisit some of the aforementioned studies in light of our estimation results in the following sections.

3 Data

3.1 Background Information and Data Description

The dataset used in this study was provided to us by a large online sports bookmaker whose operations are based in Europe. It consists of detailed information on all the bets placed by each of 100 gamblers, randomly selected from the company's customers. We track the betting histories of these people in all sports markets over a period of 3.5 years, from October 2005 to May 2009.

Before we proceed with a detailed description of our dataset, we will first provide some background information about online sports gambling. When betting online, a potential bettor must first open and fund an online sports betting account at an online bookmaker. Once he/she has established the account, he can log into it at any time and place bets on a wide variety of future sporting events. The bookmaker studied here offers betting opportunities on a variety of sports, from soccer, baseball, and tennis to golf, boxing and auto racing. The bettor chooses the sport of his/her interest and the event description he/she would like to wager on. There is a wide range of event descriptions associated with each game. For example, in a soccer match one can wager on the final outcome of the match (i.e., home win, away win or draw), the total number of goals scored (e.g., over 2.5, under 2.5), the correct score, the first goalscorer etc. In addition to betting on actual matches, online bettors can also make "proposition bets", i.e., they can bet on anything from how many goals will Lionel Messi score in the World Cup, to whether a fight will go longer than six rounds, or which team will win the Super Bowl.

An event on which bettors may bet, has several possible outcomes, each of which is associated with a price. The price of an outcome is essentially the return of a unit wager on that outcome. Prices are initially set by the bookmaker, and subsequently they vary according to supply and demand. Associated with any given set of prices (or odds) for the possible outcomes of an event, is a set of implied probabilities . These are the respective probabilities with which the outcomes would need to occur to "justify" the odds. That is, if the real probabilities were different than the implied probabilities, then the bet would have a non-zero expected value.

After selecting a number of event descriptions, the bettor then decides on the type of bet(s) he wants to place together with the amount of money he wants to wager on each bet type.[6] There are many different types of bets available in the sportsbook, that involve varying levels of risk. A "single" is the most traditional bet type: you wager on one individual selection and get a return if this selection is successful. However, bettors can also combine more than one selections in the same bet. For example, a "double" is a bet on two selections that yields a return only if both selections are successful. A "treble", "four-fold", "five-fold", etc. follow the same logic. There also exist all different combinations of the aforementioned bet types. For example, a "single*s*", "double*s*", etc. is a combination of a number of "single", "double", etc., respectively, i.e., you choose the number of events you want to wager on, and win if at least one, two, etc. of your selections win.[7] Finally, there are even more complex bet types that enable you to combine all of the above: For example, a "trixie" is a wager on three events consisting of three "doubles" and one "treble" and a minimum of two selections must win to gain a return. After combining the selected events in a number of bet types, the bettor then submits a betting slip that summarizes all information regarding his wager. Figure 1 presents a sample betting slip that includes an example of a "single", a "double", and a "doubles" on a number of sports events.

[Figure 1 about here.]

In this study, we observe the following information for each bet placed by each of the bettors in our sample: i) the date of the bet, ii) the event on which the bet was made (e.g., final outcome of baseball match between teams A and B), iii) the outcome chosen (e.g., team win, total number of goals scored, etc.), iv) the bet amount, v) the bet type (single or exact type of combination bet), vi) the odds at which the bet was placed, and vii) the bet result. In addition to these, our data include information about the gender, age, country of origin and zip code of the bettors as provided by them upon registration in the sportsbook. Using the directly observable demographic information, we approximate each bettor's annual labor income, education level and family situation by matching the gender, age and zip code information with the relevant census data of the bettor's country of origin.

Figure 2 provides summary statistics for the characteristics of bettors and their selected bets. The typical bettor in our sample is a male in his early thirties with annual labor income around €19,000. The majority of bets are placed on standard events (i.e., final match outcome) of soccer matches. People usually combine more than one events in the same bet type and place on average

[6]Generally, there is no upper and lower bound in the amount of money people are allowed to bet in the sportsbook. The bookmaker, however, reserves the right to impose maximum stake limits if he/she detects arbitrage trading.

[7]Often, the only limit to the number of selections included within such a bet is the bookmaker's maximum allowable payout on one bet.

3.7 bet types per bet day. Finally, the majority of bets are placed either on the same or one day before the sporting event in question takes place.

[Figure 2 about here.]

3.2 Data Analysis

3.2.1 Lottery Representation

In this section we will explain how we use the bet information we observe, to represent the risky gambles selected by each individual in our sample.

In the first step, we create a lottery representation for all single bets selected by each bettor. Each bet available in the sportsbook is associated with a probability distribution over a set of possible monetary prizes. In particular, there are two possible prizes associated with each selection: you either win and earn a positive return if the selected outcome occurs, or lose your stake if the outcome does not occur. The net total that will be paid out to the bettor, should he win, is implied by the odds associated with the selected outcome. For example, odds of 4/1 imply that the bettor will make a profit of $400 for each $100 staked. The odds quoted by the bookmaker on all outcomes of an event also imply a probability with which each of these outcomes is expected to occur.[8] For example, odds of 4/1 imply that the outcome has a chance of one in five of occurring and therefore the implied probability is 20%. The sum of the implied probabilities over all outcomes of an event is always greater than 1, with the amount by which the sum exceeds 1 representing the bookmaker's commission.[9] The "sum to one" probability is then obtained by dividing the implied probability of an outcome by the sum of the implied probabilities over all outcomes of the event in question. For example, if the implied probabilities sum to 1.1, then the "sum to one" probability of an outcome whose odds are 4/1 would be 0.18 (i.e., 0.2/1.1).

In the second step, we create a lottery representation for all combination bets selected, i.e. a number of single bets combined under a specific bet type. As explained in the previous section, bet types differ in the number of selections involved and the minimum number of winning selections required to guarantee a positive return. In particular, bets of type "double", "treble", etc. have 2 outcomes, i.e., you either get some positive return if all of your selections win, or lose your stake if at least one selection loses. The odds for these types of bets are calculated by simply multiplying the odds quoted by the bookmaker on the separate events involved. For example, the odds on the "double" presented in Figure 3 is the product of the odds for the two single bets it contains. On the other hand, bets of type "single*s*", "double*s*", "trebles", etc. have 2^n outcomes where n is the number

[8] This implied probability is often quoted by bookmakers together with the odds.

[9] The bookmaker's commission is usually around 7%-10%, though it varies a lot across different sports and events.

of events that people have combined in the same bet type. These types of bets can be decomposed into a number of "single", "double", etc. each of which is evaluated separately. For example, a "doubles" bet on n events is equivalent to $\begin{pmatrix} n \\ 2 \end{pmatrix}$ "double" bets, a "trebles" is equivalent to $\begin{pmatrix} n \\ 3 \end{pmatrix}$ "treble" bets, etc.[10] More complex bet types are similarly treated according to their definition. An important note here is that bettors are not allowed by the bookmaker to place combination bets on events that are considered related (e.g. winner of a soccer match and total number of goals scored, winner of the semi-final and winner of the final, etc.). Figure 3 shows how odds and prizes are calculated for each of the bet types included in the betting slip of Figure 1.

[Figure 3 about here.]

In the third step, we create a lottery representation for all bets selected during a given play session. A play session is defined as the set of all bets placed over a single day by an individual.[11] We thus represent the day lottery by constructing all possible combinations of payoffs of all the bets chosen on that day. One challenge in calculating the day lottery is that the outcomes of some bets placed on the same day might not be independent.[12] We therefore identify the events whose outcomes are not independent and deal with days that involve these events accordingly. In particular, for more than one bets on the same match and event description (e.g., on the final outcome of the same soccer match) we i) compute all possible combinations of event outcomes of all independent matches wagered on that day, ii) find the corresponding probability of each combination and iii) calculate the payoff of the bettor in each combination according to the choices he/she has made. In cases of more than one bets placed on separate related matches (e.g., on the outcome of the semi-final and the outcome of the final) or on different event descriptions of the same match (e.g., on the final outcome of a soccer game and on the total number of goals scored) we are unable to construct the day lottery without further information about the correlation among the win probabilities of the event outcomes involved. We therefore drop the bet days that involve related bets of this type. After these steps, we have constructed all day lotteries chosen by every bettor in our sample. Our final sample contains 3,755 day lotteries chosen by 100 bettors from October 2005 to May 2009.

[10]The amount staked on a "doubles", "trebles", etc. is divided evenly among the "double", "treble", etc. bets involved.

[11]Lacking information on the exact time of bets, we decide to carry out our analysis at the lowest level of aggregation available to us which is the bet day. The choice of this time window is supported by the fact that the majority of the bets in our sample are placed one day before the actual event date.

[12]Although related bets are not allowed within the same bet type, people may bet on them under different bet types placed within the same bet day.

3.2.2 Descriptive Statistics

Figure 4 reports summary statistics of the day lotteries chosen by the bettors in our sample. A typical day lottery contains 2 prizes, has negative expected value (-2.2) and expected return (-0.23), high variance (308) and positive skewness (2.1). The bet amounts range from €0.01 to €1768 with the median bet amount being equal to €10. The maximum prize of the day loterries ranges from €0.24 to €12000 with the median maximum prize equal to €62. The summary statistics in Figure 4 indicate that there is a multitude of bets available in the sportsbook that have actually been chosen by the bettors in our sample. In fact, since bettors are allowed to construct any desired day lottery by combining a wide variety of odds, bet types and stakes, they can essentially choose anything from an almost safe lottery that returns a tiny payoff with probability 0.99 to a highly skewed lottery that returns a very high payoff with probability 0.01. However, as Figure 4 shows, only around 10% of the chosen lotteries have very high positive skewness (over 10) and about the same proportion of lotteries are negatively skewed. The reason why people do not bet on very positively skewed lotteries is that their expected return is very low. In fact, there is a strong negative correlation between the skewness and the expected return of the chosen lotteries (-0.7). The reason for this is that the majority of highly skewed lotteries represent combinations bets, and when bettors combine more than one events, the effect of the bookmaker's commission in each event is compounded to the detriment of the bettor in terms of the financial return. Finally, note that only 4.5% of the chosen lotteries are close to 50:50 bets to win or lose some fixed amount.

[Figure 4 about here.]

It is important to note here that the lotteries considered in this study are quite different from those faced by individuals in related studies. Experimental subjects are typically confronted with a sequential series of choices over lottery pairs. The gambles involved have a small number of outcomes (usually two to four) and the stakes are usually limited to less than \$50. In the most widely studied game show "Deal or Not Deal", contestants are typically asked to decide whether to accept some deterministic cash offer or face a lottery where a list of monetary prizes ranging from \$0.01 to \$1000000 are equally possible. Depending on the game round, the risky lotteries involved can have up to 20 outcomes, all of which are positive. Insurance subjects are asked to choose a contract from a menu of deductible and premium combinations. Choice is typically exercised from a set of two binary lotteries representing unknown future expenditure: a lottery that involves a high premium but provides a high deductible payment in the event of an accident, and a lottery that involves a low premium but provides a low deductible payment. The typical contract in the data of Cohen and Einav (2007) offers an individual to pay an additional premium of \$55 in order to save

\$182 in deductible payments. As the summary statistics in Figure 4 indicate , the lotteries available in a sportsbook span a much wider range of prizes and probabilities than those considered so far.

3.3 Choice Set

In the empirical analysis, we employ discrete choice modeling methods to analyze bettors' lottery choices. The development of discrete choice models is a significant advance in the study of individual choice behavior. They are widely used in economics, marketing, transportation and other fields to represent choice from a set of mutually exclusive alternatives. Therefore, specifying the set of alternatives among which choice is exercised is an essential component of all discrete choice models. In some choice situations[13], however, the actual choice set is either not observable or extremely large so that it becomes essential for the researcher to describe it as best as possible using some reasonable, deterministic or probabilistic rules.

In our setting, the true set of alternatives should ideally be the set of all day lotteries considered by each bettor on each bet day. Obviously, what constitutes a feasible alternative for any particular individual on any particular day is difficult to determine. Therefore, the true choice set is a latent construct to us since nothing is observed about it except for the chosen alternative. Moreover, given the large number of day lotteries that can be constructed from the bets available at any time in the sportsbook, especially if you consider that the bet amount is selected from an infinite set of possible values, the number of elements in the true choice set is immense. We therefore follow the literature in discrete choice analysis and approximate this universal choice set by choosing a set of representative lotteries from the set of all day lotteries chosen at least once by any bettor in our sample.

We therefore proceed in the following steps. First, we reduce the number of alternative lotteries in the choice set by clustering the 3,755 observed day lotteries in 100 clusters. We do so using a hierarchical agglomeration algorithm, according to which, in each step we combine in a cluster the two most similar clusters, until we reach the desired level of clustering. The similarity of two degenerate clusters, i.e., two lotteries, is measured by their Wasserstein distance, which is a metric defined on the space of probability distributions (see Rachev (1991)).[14] The similarity of two non-degenerate clusters is measured by the mean of all the pairwise Wasserstein distances. After

[13]Examples include consumers' residential choice (Weisbrod et al. (1980)), travelers' choice of destination (Daly (1982)) and household choice of telephone service options (Train et al. (1987)) among others.

[14]The Wasserstein distance between two probability measures μ and ν on the real line $\mathbb{R}$, is defined as:

$$W(\mu,\nu) = \left(\frac{1}{n}\sum_i (x_i - y_i)^2\right)^{\frac{1}{2}}$$

,where $\{x_i\}$,$\{y_i\}$ are sorted "observations".

clustering to the desired level is performed, we choose the most representative lottery from each cluster, i.e., the lottery that has the smallest mean Wasserstein distance to the other members of the cluster. We assume that the choice set faced by a bettor on a given day includes the lottery chosen on that day and the 100 lotteries that are most representative of all day lottery choices observed in our sample, the idea being that this reduced choice set reasonably spans the space of all feasible day lotteries available in the sportsbook.[15] Finally, we augment the resulting choice set with a safe lottery that represents the option not to place any bets on that day.

4 The Empirical Model

4.1 Econometric Model

To analyze bettors' choices out of the set of alternative lotteries, we use a multinomial discrete choice model and in particular, a random utility model with random coefficients (mixed logit). The mixed logit model is one of the most flexible and widely used innovations of discrete choice modeling. Unlike standard logit, which allows for only observed heterogeneity in tastes across individuals, mixed logit also accounts for unobserved taste heterogeneity by assuming that individual preference parameters are randomly distributed in the population (McFadden and Train (2000)).[16]

Under the random utility model, the utility U_{njt} that individual n obtains from choosing lottery j on day t can be decomposed into a deterministic and a random component as:

$$U_{njt} = V_{njt} + \varepsilon_{njt}$$

with V_{njt} and ε_{njt} being the observed and unobserved parts of utility respectively.

In this study, V_{njt} will take a non-linear utility functional form $h\left(\rho_n, \mathbf{P}_{njt}\right)$, where $\mathbf{P}_{njt}$ represents the probability distribution over the set of monetary prizes of lottery j chosen by individual n on day t (representative utility) and ρ_n is a vector of parameters representing the tastes of individual n, which is to be estimated from the data. For instance, under the expected utility theory with CRRA, $h\left(\cdot\right)$ will be given by:

$$h\left(\cdot\right) = \sum_k p_k U\left(m_k\right) = \sum_k p_k \left(\frac{m_k^{1-\rho_n}}{1-\rho_n}\right)$$

[15] An alternative approach would be to adopt a probabilistic choice set generation process where the probability that a particular choice set is chosen among the set of all possible choice sets is either a function of exogenous factors (e.g., budget constraints) as Manski (1977) proposes, or a function of the underlying utilities of the alternatives that belong in each choice set as Swait (2001) proposes. In our case, with 100 alternative lotteries the set of all subsets of the alternatives would contain $2^{100} - 1$ elements so we would need to impose further restrictions on it.

[16] Under an alternative interpretation, mixed logit allows for correlation in unobserved errors in choice over time (Revelt and Train (1998)) and/or rich substitution patterns across alternatives.

where ρ_n is the risk aversion coefficient of bettor n and $\mathbf{P}_{njt} := (p_1, m_1; \ldots; p_K, m_K)_{njt}$ where probability p_k is assigned to monetary outcome m_k, for $k = 1, 2, \ldots, K$.

The unobserved utility term ε_{njt} can be interpreted as i) possible mistakes that the bettor might make due to carelessness in evaluating or choosing a bet, ii) utility components that are known to the bettor but unobserved by the econometrician or iii) unobserved variation in the choice behavior of individual n when facing the same observed portion of utility V_{njt} $\forall j, t$. The logit model is obtained by assuming that each of the error terms ε_{njt} is i) extreme value distributed with a location parameter η_n and a scale parameter $\mu_n > 0$ and ii) independently and identically distributed across individuals, alternatives and time. The *mixed* logit model assumes that the unknown utility parameters, labeled ρ_n, vary over decision makers in the population with some density $f(\cdot|\theta)$ to be specified by the researcher.

The probability $L(y_{njt})$ that lottery j is chosen by bettor n on day t out of the set of alternative lotteries C is given by:

$$L(y_{njt}) = \Pr\left(V_{njt} + \varepsilon_{njt} \geq \max_{j \in C}(V_{njt} + \varepsilon_{njt})\right)$$

which under the assumptions of the mixed logit specification becomes:

$$L(y_{njt}|\theta) = \int_{\rho_n} \prod_t L(y_{njt}|\rho_n) f(\rho_n|\theta)\, d\rho_n \tag{1}$$

where

$$L(y_{njt}|\rho_n) = \frac{e^{k_n V_{njt}}}{\sum\limits_{i \in C} e^{k_n V_{nit}}} \tag{2}$$

is the standard logit choice probability. Note that in Equation 2, we have normalized the variance μ_n of the error terms to unity and therefore the scale of the utility k_n can be interpreted as the relative precision of the error. It can be shown that when the cardinality of the choice set is greater than 2, the relative error precision k_n affects the estimation of the utility parameters. We therefore treat k_n as an additional free parameter, randomly distributed across individuals, to be estimated along with the parameters included in ρ_n.

4.2 Cumulative Prospect Theory

Many authors have pointed out the suitability of prospect theory for explaining the popularity of gambling and the actual behavior of gamblers. Tversky and Kahneman (1992) first suggested that prospect theory can explain why people bet on longshot outcomes at actuarially unfair odds (e.g. purchase of lottery tickets). Barberis (2010) shows that prospect theory can also explain the popularity of 50:50 casino bets and that it captures many features of the observed gambling behavior. On the empirical side, Jullien and Salanie (2000) and Snowberg and Wolfers (2010) find that prospect theory has the highest explanatory power for the actual behavior of a representative racetrack gambler.

Cumulative prospect theory (CPT) is a variant of the original prospect theory proposed by Kahneman and Tversky (1979). According to CPT, an agent evaluates a lottery $(p_{-m}, m_{-m}; \ldots; p_0, m_0; \ldots; p_n, m_n)$ where $m_{-m} \leq \ldots \leq m_0 = 0 \leq \ldots \leq m_n$ by assigning it the value:

$$\sum_i \pi_i u(m_i)$$

where

$$\pi_i = \begin{cases} w^+(p_i + \ldots + p_n) - w^+(p_{i+1} + \ldots + p_n) & \text{for } 0 \leq i \leq n \\ w^-(p_{-m} + \ldots + p_i) - w^-(p_{-m} + \ldots + p_{i-1}) & \text{for } -m \leq i < 0 \end{cases} \tag{3}$$

with $w^+(\cdot)$ and $w^-(\cdot)$ being the probability weighting functions for gains and losses, respectively, and $u(\cdot)$ is the value function.[17] The main characteristics of this theory are the following: i) Utility is defined over gains and losses relative to a reference point rather than over final wealth positions; ii) The value function u is concave in the domain of gains and convex in the domain of losses, meaning that people are risk averse over gains, and risk-seeking over losses. It also has a kink at the origin, indicating a greater sensitivity to losses than to gains (loss aversion); iii) The framing of outcomes can affect the reference point used in the evaluation of prospects (mental accounting); and iv) People evaluate gambles using transformed probabilities obtained by applying a probability weighting function on objective probabilities. The probability weighting function has an "inverse S-shape", implying that people tend to overweigh low probabilities and underweigh high probabilities. The difference of CPT from the original version of prospect theory is that probability weighting is applied to the cumulative probability distribution instead of the probability density function.[18]

Here we employ the value function and probability weighting function proposed by Tversky and Kahneman (1992):

$$u(m) = \begin{cases} ((W+m) - RP)^a & \text{for } W+m \geq RP \\ -\lambda (RP - (W+m))^a & \text{for } W+m < RP \end{cases} \tag{4}$$

[17] When $i = n$ and $i = -m$, Equation 3 reduces to $\pi_n = w^+(p_n)$ and $\pi_{-m} = w^-(p_{-m})$, respectively.

[18] This procedure for assigning weights ensures that the utility function satisfies monotonicity. It also has the appealing property that the weights attached to two outcomes with the same objective probability may differ depending on their relative standing. This would imply, for example, that extreme outcomes are assigned particularly high or low weights.

and

$$w^{-}(P) = w^{+}(P) = \frac{P^{\gamma}}{\left(P^{\gamma} + (1-P)^{\gamma}\right)^{\frac{1}{\gamma}}} \tag{5}$$

where $RP = W$ is the reference point that separates losses from gains, $a \in (0, 1)$ measures the curvature of the value function, $\lambda > 1$ is the coefficient of loss aversion, and $\gamma \in (0.28, 1)$ measures the curvature of the probability weighting function.[19][20]

4.2.1 State Dependence

The panel aspect of our dataset is particularly interesting, because it allows us to examine the existence of possible state and time dependence in bettors' risk preferences. One form of state dependence we consider is the dependence of bettors' risk preferences on the outcomes of their previous bets, an idea that is supported by the psychology literature. Thaler and Johnson (1990) study risk-taking behavior in an experimental setting and find that when people are faced with sequential gambles, they are more willing to take risk after prior gains than after prior losses. This behavior is labeled "the house money effect", reflecting the idea that after a prior gain, losses are coded as reductions in this gain, as if losing the "house's money" is less painful than losing one's own cash. Gertner (1993) and Post et al. (2008)obtain similar results when studying the behavior of TV game show contestants. Contrary to these studies in which subjects never lose money out of their own pockets, our study examines this effect from bettors who experience real losses ranging from a few dollars to thousands of dollars.

Prospect theory offers a natural framework to study how prior outcomes could affect people's subsequent choices. Since the framing of outcomes affects the reference point relative to which gains and losses are evaluated, it seems reasonable to allow this reference point to depend on individuals' prior betting performance. In particular, after incurring a gain/loss, the reference point might not update completely to reflect this gain/loss but rather remain sticky to some previous level resulting in prior outcomes being integrated with current payoffs.

Formally, we express this idea by using a variant of the cumulative prospect theory described above. First, we create a measure of previous gains/losses experienced by bettor n on day t, denoted by $CumProfit_{nt}$, as the cumulative payoff of all the bets previously placed by bettor n that were settled from one month before the current bet day up to one day before the current bet day, i.e., from $t-30$ up to $t-1$.[21] Then, we allow the reference point RP_{nt} that bettor n uses to evaluate

[19] To reduce the number of free parameters we restrict the curvature of the probability weighting function to be equal in the domain of gains and losses. Using experimental data, Tversky and Kahneman (1992) estimate the probability weighting parameters for their median subject to be 0.61 and 0.69 for gains and losses respectively.

[20] Ingersoll (2008) notes that the probability weighting parameters need to be above the critical value 0.28 to ensure that decision weights are positive and the probability weighting function is monotone.

[21] Note that this variable is just a *proxy* of bettor n's cumulative gains on day t. First, it is day-specific rather than

the payoffs of the lottery of day t to be a convex combination of bettors' wealth with and without the payoffs of previous bets, i.e.

$$RP_{nt} = \delta_n W_{nt} + (1 - \delta_n)(W_{nt} - CumProfit_{nt})$$

where W_{nt} is the wealth of bettor n in the beginning of day t including all gains/losses incurred during the last month up to day t, and $\delta_n \in [0, 1]$ is an individual-level parameter to be estimated, that measures the stickiness of the reference point for bettor n. When $\delta_n = 1$, the reference point is instantaneously updated to incorporate the outcomes of previous bets as standard CPT suggests.

To illustrate this idea with an example, consider the following situation. Suppose that your wealth in the beginning of day t is \$120 including your last month's cumulative gains which are $-\$20$. According to the standard version of CPT, the reference point relative to which people will evaluate the outcomes of the day lottery they are about to take should be $\$120 - \$20 = \$100$. Here we suggest that this reference point does not update completely to \$100 but rather remains sticky to some number between \$120 and \$100, let's say \$110. Therefore, positive outcomes less than \$10 will now be coded as a loss.

Apart from the dependence of the reference point on previous betting performance, the richness of our dataset allows us to consider an additional, previously unexplored, channel through which prior outcomes could affect bettors' subsequent risk-taking. In particular, we consider an extension of standard CPT in which the loss aversion parameter λ and the curvature of the value function α of bettor n on day t are allowed to be a linear function of earlier bet outcomes, i.e.

$$\begin{aligned} \lambda_{nt} &= \lambda_{n0} + \lambda_{n1} CumProfit_{nt} \\ \alpha_{nt} &= \alpha_{n0} + \alpha_{n1} CumProfit_{nt} \end{aligned}$$

where λ_{n0}, α_{n0} are individual-specific intercepts and λ_{n1},α_{n1} measure the sensitivity of λ and α to previous betting performance.

In addition to the dependence on previous bet outcomes, in future research we are planning to explore more types of possible time dependence in bettors' behavior. For instance, it would be

bet-specific creating a problem if some bets placed on day t were placed after some previous bets placed on the same day have already paid off. Since we only observe the dates on which bets were placed but not their exact times, this problem cannot be resolved. Second, a bet is considered settled when all uncertainty about it has been revealed, i.e., we know the outcome of all the matches it involves. However, the earliest time at which the payoff of a bet is known is not necessarily the conclusion of the last game involved in it, i.e., the bettor might already know his payoff if he knows the outcome of some, not all, of the games that he has combined in the same bet. For example, in the case of a double, if someone loses the bet on the first match, he knows he has lost the bet no matter what happens in the second match or in the case of a doubles on 3 games, he knows he has won at least some amount of money if he wins 2 of the matches.

interesting to allow utility parameters to vary with the GDP of bettors' country of origin representing the effect on individuals' preferences of aggregate macroeconomic conditions. Furthermore, possible time dependence above and beyond the effect of economic growth and business cycle fluctuactions could be captured by a set of time dummies. Finally, we can introduce state dependence in terms of outstanding bets (i.e. bets that have not paid off before the current day lottery is chosen), for example by allowing risk preference parameters to be a linear function of the characteristics of these bets (e.g. mean, variance, skewness).

4.3 Frequency of Betting

In addition to the lotteries chosen by bettors on the days on which they have decided to bet, we also observe the frequency with which individuals place their bets. In our setting, both the play frequency and the lottery choice per play are informative of bettors' risk aversion so the two should not be considered in isolation. For example, failing to consider the play frequency, a bettor that concentrates all of his bets on one day of the week will be mistakenly predicted to have lower risk aversion than a bettor who spreads the same bets on more than one days.

The observed play frequency is determined both by the frequency with which people get the chance to log into the sportsbook depending on their lifestyle (e.g., a single bettor might have more free time than a married bettor, and will therefore play more frequently) but also by the number of bet opportunities people reject depending on their risk preferences. To account for the effect of play frequency in the empirical analysis, we partition bettors' decision-making process into three sequential stages: In the first stage, bettor n gets or does not get the opportunity to log into the sportsbook on day t with probability p_{nt} and $1 - p_{nt}$ respectively.[22] This stage is not a decision that the bettor makes: whether he gets the chance to bet or not on a given day is determined solely by his demographic characteristics (i.e., gender, age, family situation, etc.) and time covariates (i.e., day of the week, etc.). If the bettor gets the chance to log into the sportsbook, in the second stage he decides whether or not to accept this opportunity. If he accepts it, in the third stage he chooses a day lottery from the set of alternative risky lotteries. If he rejects it, he essentially chooses a safe lottery that will pay back his stake with probability 1.

Letting $y_n = (y_{n1}, \ldots, y_{nT_n})'$ be the sequence of daily lottery choices we observe for individual n and $x_n = (x_{n1}, \ldots, x_{nT_n})'$ be a set of dummy variables indicating whether we observed play on a given day or not for individual n, the conditional probability of observing $\{x_n, y_n\}$ is given by:

[22] This probability is allowed to vary with bettors' demographic characteristics as well as day, month and year dummies.

$$L\left(x_n, y_n|\rho_{nt}, p_{nt}\right) = \prod_{t=1}^{T_n} \left\{x_{nt}p_{nt}L\left(y_{nt}|\rho_{nt}\right) + \left(1-x_{nt}\right)\left[p_{nt}L\left(y_{nt}=0|\rho_{nt}\right) + \left(1-p_{nt}\right)\right]\right\} \tag{6}$$

Equation 6 says that if the bettor is offered the opportunity to play (with probability p_{nt}), he may either reject it, i.e., choose a safe lottery, with probability $L\left(y_{nt}=0|\rho_{nt}\right)$, or he may accept it and choose one of the risky lotteries available in the sportsbook with probability $L\left(y_{nt}|\rho_n\right)$. Note that the former is observed as no bet for that day. Probabilities $L\left(y_{nt}=0|\rho_{nt}\right)$ and $L\left(y_{nt}|\rho_{nt}\right)$ are the standard logit choice probabilities given by Equation 2. If the bettor does not have the opportunity to bet (with probability $1-p_{nt}$), then we simply observe no bet for that day. The probability p_{nt} is treated as a latent parameter to be estimated from the data together with the vector of utility parameters ρ_n.

5 Estimation of the Benchmark Model

There are two ways to estimate our extension of the multinomial mixed logit model described in the previous section. The classical approach would be to implement a Maximum Simulated Likelihood (MSL) methodology, which requires the maximization of the log-likelihood function given by the logarithm of Equation 11 above, where the integrals are approximated through simulation.[23] The Bayesian approach suggests starting out with some prior density for the latent parameter vector $\beta_n = \binom{\rho_n}{p_n}$ and then using the observed data to update this prior and form a posterior density. With a large enough sample, both approaches should asymptotically provide similar results. However, given the complexity of the empirical specification we employ, in our case the Bayesian approach is far more convenient in terms of computation time. Moreover, as noted by Train (2001), the Bayesian procedure also has a theoretical advantage over the classical one, which derives from the fact that both estimations require the simulation of an integral that cannot be expressed in closed form: The conditions under which the simulation-based version of the Bayesian estimator becomes consistent, asymptotically normal and efficient are less stringent than those required for the classical estimator.[24] In view of the above, we decided to follow a Bayesian perspective in estimating our

[23]The simulation is performed by taking draws from $f(\cdot)$, calculating the integrand for each draw, and averaging the results.

[24]The Bayesian estimator is consistent and asymptotically normal for a fixed number of simulation draws. If the number of draws rises with the sample size L at any rate, the simulation noise disappears asymptotically and the Bayesian estimator becomes efficient and asymptotically equivalent to the MLE. In contrast, the MSL estimator becomes consistent when the number of draws rises with L and asymptotically normal (and equivalent to the MLE) when the number of draws rises faster than $\sqrt{L}$. As Train (2001) points out, since it is difficult to know in practice how to satisfy the condition that the number of draws rises faster than $\sqrt{L}$, the Bayesian estimator is attractive relative to the MSLE, even though their non-simulated counterparts are equivalent.

empirical model.

In the following sections we present the hierarchical model we employ and discuss how we estimate it using Markov Chain Monte Carlo (MCMC) methods in the Bayesian framework.

5.1 Hierarchical Bayes Model

We would like to have a posterior distribution for the utility parameter vector $\rho = (\rho_1', \ldots, \rho_N')'$ and the exogenous probability of having the opportunity to bet $p = (p_1', \ldots, p_N')'$. The vector of individual-specific parameters associated with the CPT specification that includes path-dependence is $\rho_n = (k_n, \alpha_n, \lambda_n, \gamma_n, \delta_n)'$ where k_n is the scale of the utility, α_n the curvature of the utility function, λ_n the loss aversion coefficient, γ_n the curvature of the probability weighting function and δ_n measures the degree to which the reference point that separates gains from losses is affected by previous bet outcomes.

Selected parameters are allowed to vary with bettors' demographic characteristics and variables that capture possible state dependence on bettors' risk preferences, such as previous betting performance. Therefore, the elements of $\beta_n = \binom{\rho_n}{p_n}$ can be expressed as:

$$\beta_n = \beta_0 + Z_n'\beta_1 + G_{nt}\beta_2$$

where Z_n is a vector of individual specific attributes and G_{nt} is a vector of individual and time varying covariates. For identification purposes, we restrict the effect of the demographics to be fixed, i.e., the same across individuals, and allow the intercept β_0 and the sensitivity to state dependent variables β_2 to vary randomly in the population. The covariates in Z_n include gender, age, education level, family situation and a set of home area dummies. As such, we divide the set of parameters β into two sets; $\hat{\beta}$ represents a part of β containing deterministic parameters and $\tilde{\beta}_n$ is a set of parameters randomly distributed in the population.

The Bayesian estimation methodology we employ involves the following steps. In the first step, we specify the prior distributions for the parameters of interest. For $\tilde{\beta}_n$ we adopt a hierarchical normal prior. In particular, we assume that all individual-level parameters are i.i.d. drawn from a multivariate normal distribution with population mean b and population variance-covariance matrix W, i.e., $\forall\tilde{\beta}_n$

$$\tilde{\beta}_n | b, W \sim N(b, W).$$

Appropriate transformation functions are applied so that all parameters in $\tilde{\beta}_n$ lie in $(-\infty, +\infty)$. In particular, a logarithmic function is used to map to the real line parameters that are allowed to take only positive values and a Johnson's transformation function is used to transform parameters that are bounded by 0 and 1.[25] The population level parameters b and W follow a multivariate normal-inverse Wishart unconditional prior distribution, i.e.:

$$W^{-1} \sim W\left(\underline{\mathrm{v}}, \underline{\mathrm{V}}^{-1}\right)$$
$$b|W \sim N\left(\underline{\mathrm{b}}, W/\underline{\mathrm{k}}\right)$$

In Section 8.1 we discuss our choice of the prior hyperparameters $\underline{\mathrm{v}}$,$\underline{\mathrm{V}}$,$\underline{\mathrm{b}}$ and $\underline{\mathrm{k}}$.

In the second step, we derive the posterior distribution of the parameters of interest, i.e., $\tilde{\beta}_n$, b, W, $\hat{\beta}$, conditional on the observed data $\{X, Y\}$. The joint posterior distribution $K\left(\left\{\tilde{\beta}_n\right\}_{n=1}^{N}, b, W, \hat{\beta}|X, Y\right)$ will be proportional to the product of the likelihood function and the prior densities, i.e.:

$$K\left(\left\{\tilde{\beta}_n\right\}_{n=1}^{N}, b, W, \hat{\beta}|X, Y\right) \propto \prod_n L\left(x_n, y_n|\tilde{\beta}_n, \hat{\beta}\right) \cdot \left(f_N\left(\tilde{\beta}_n|b, W\right) \cdot k\left(b, W\right)\right) \cdot k\left(\hat{\beta}\right) \quad (7)$$

where i) $f_N\left(\tilde{\beta}_n|b, W\right)$ is the normal density with mean b and variance W, ii) $k_\beta\left(b, W\right) = k\left(b\right)k\left(W\right)$ is the prior on b and W, i.e., normal for b and inverted Wishart for W, and v) $k\left(\hat{\beta}\right)$ is the prior for the fixed parameters $\hat{\beta}$. The likelihood $L\left(x_n, y_n|\tilde{\beta}_n, \hat{\beta}\right)$ is given by Equation 6.

Finally, information about the joint posterior $K\left(\tilde{\beta}_n, b, W, \hat{\beta}|X, Y\right)$ is obtained through simulation, that is, by taking draws from the posterior and computing relevant statistics, e.g., moments, over these draws. Since this posterior is not from a known family of distributions and cannot be drawn from directly, we use the Gibbs Sampler algorithm (Gelfand and Smith (1990)) to make sampling possible. In particular, we generate a sequence of draws from the conditional posterior distribution of each parameter given the previous draw of the other parameters and the data. The resulting sequence of draws from the conditional posteriors is a Markov chain with a stationary distribution equal to the joint posterior distribution. An attractive feature of Gibbs sampling is that it allows for data augmentation of latent variables (Tanner and Wong (1987)), which amounts to simulating the individual-level parameters $\tilde{\beta}_n$ conditional on the data and the population level parameters and then treating these simulations as part of the observed data. This technique enables us to obtain individual-level parameter estimates while facilitating the derivation of the posterior distribution. We use joint distribution tests (see Geweke (2004)) to verify that the Gibbs sampler is derived and programmed correctly.

[25] As an example, the utility scale parameters $k_n > 0$ and the probability parameter $0 < p_n < 1$ will be transformed as follows:

$$\tilde{k_n} = \log\left(k_n\right)$$
$$\tilde{p}_n = \log\left(\frac{p_n}{1-p_n}\right)$$

6 Estimation Results

In this section we present the estimation results of the CPT specification with and without the presence of state-dependence in individual risk preferences.

6.1 Cumulative Prospect Theory Without State-Dependence

Figure 5 shows the posterior simulator distribution of the latent parameter vector $\tilde{\beta}_n$ estimated by taking draws from a normal distribution with parameters equal to the posterior mean of the population level parameters.[26]

[Figure 5 about here.]

The posterior means of the CPT preference parameters α and λ imply that for the average bettor in our sample the concavity (convexity) of the value function in the region of gains (losses) is mild ($\alpha = 0.82$), while the kink at the origin is sharp ($\lambda = 2.73$). Also, the posterior mean of the probability weighting parameter, $\gamma = 0.85$, indicates that the average bettor exhibits only moderate deviation from linear probability weighting. The estimated loss aversion parameter matches the empirical observation that the average bettor in our sample tends to reject 50-50 bets The estimated probability weighting parameter ($\gamma = 0.85$) matches the empirical observation that bettors favor lotteries with medium positive skewness but typically reject lotteries with very high positive skewness, which imposes a lower bound in the estimated parameter. Using a different set of lotteries, Tversky and Kahneman (1992) find that their median experimental subject has a value function with a shape similar to the one we estimate here ($a = 0.88$ and $\lambda = 2.25$), albeit with a more pronounced probability distortion ($\gamma = 0.65$).

Figures 6 and 7 show the posterior simulator distribution evaluated not only at the posterior mean (solid line) but also at the lower and upper bound of the 95% Highest Predictive Density Interval (HPDI) (dashed lines) for the population mean and the population variance of the CPT parameters.[27]

[Figure 6 about here.]

[Figure 7 about here.]

[26]The appropriate transformation functions are applied to transform parameters in $\tilde{\beta}_n$ that need to be constrained within a specific range.

[27]The HPDI is the Bayesian analog to the classical 95% confidence interval, i.e. a 95% HPDI delivers a lower and an upper bound such that the resulting interval is the smallest possible to contain 95% of the density mass of a given distribution.

The two panels of Figure 8 show the shape of the value function and the probability weighting function in the CPT specification, where the solid lines correspond to the curves for the posterior population means of parameters a, λ and γ and the dashed lines delimit their respective HPDIs.

[Figure 8 about here.]

Nevertheless, the most striking feature of Figure 5 is the presence of substantial heterogeneity in all utility parameters, which indicates that making inferences based simply on the posterior means is inappropriate and may lead to biased results.[28] In particular, the posterior standard deviation of parameters a, λ and γ are 0.19, 2.2 and 0.11 respectively. We can therefore conclude that there is a continuum of prospect theory types in the population, whose choices are characterized by different parameter triplets (a, λ, γ).

By looking at the correlations among the CPT parameters presented in Figure 9, we observe that there is a very strong negative correlation between parameters γ and λ (-0.87), a less strong positive correlation between parameters α and γ (0.34), and a negative correlation between α and λ (-0.27), all of which are significant at the 95% level. These correlations imply that the three features of prospect theory tend to move together, and more predominantly, that individuals with high sensitivity to losses (high λ) tend to significantly distort probabilities (low γ). This negative correlation between λ and γ seems reasonable if we consider that increasing the loss aversion parameter λ makes the global shape of the value function more concave, which in turn implies that people dislike positively skewed lotteries (i.e. large gains with small probability) than they otherwise would. Since we observe that bettors in our sample often choose gambles with positive skewness, by lowering γ the counterintuitive effect of λ on positively skewed gambles is mitigated.

[Figure 9 about here.]

6.2 Interpretation of Results

In this section we explore how the estimated preference parameters reflect individuals' actual behavior by examining the relationship between the individual-level estimates and the main characteristics of the chosen lotteries.

It is well-known that the parameter α controls the curvature of the value function: A smaller α corresponds to a higher degree of risk-aversion in gains and a higher degree of risk-lovingness in losses. The results in Panel A of Figure 10 show that individuals for whom we estimate a smaller α don't exhibit higher risk-aversion in gains (see Columns 3 and 4), but they do exhibit a statistically significantly higher degree of risk-lovingness in losses (see Columns 5 and 6). What is somewhat

[28] It is important to note that this heterogeneity is not driven by our selection of priors (see Section 8.1).

less well-known about α is that for some distributions of outcomes, a small α is also associated with skewness-aversion. In particular, Barberis and Huang (2008) show that in a financial market with several assets having normally distributed payoffs and an asset having a positively-skewed payoff, the expected excess return on the positively-skewed asset in equilibrium is decreasing in α. In addition, Ågren (2006) shows that in a financial market with assets having normal-inverse Gaussian distributed returns, the derivative of an investor's utility with respect to skewness is increasing in α. These are by no means fully general results, but they are indicative of another channel through which α can affect behavior. Indeed, the results in Panel B of Figure 10 verify that in our data, individuals for whom we estimate a smaller α choose gambles with statistically significantly smaller skewness.

[Figure 10 about here.]

The parameter λ controls the degree of loss aversion of the value function: A larger λ corresponds to a more pronounced kink of the value function at the reference point. This higher degree of local concavity is often used to explain peoples' aversion to 50-50 gambles. Interestingly, Barberis and Huang (2008) and Ågren (2006) also show that a large λ is associated with skewness-aversion. Again, the results in Figure 10 verify this in our data, since individuals for whom we estimate a larger λ choose gambles with statistically significantly smaller skewness. Consistent with this intuition, the results in Panel B of Figure 10 show that a larger λ is also associated with significant variance-aversion.

Finaly, the parameter γ controls the degree of probability weighting: Loosely speaking, a smaller γ corresponds to a more pronounced systematic overweighting of small probabilities. As a result, a smaller γ is associated with a preference for positive skewness and an aversion to negative skewness. The results in Panel B of Figure 10 verify this in our data, since individuals for whom we estimate a smaller γ choose gambles with statistically significantly larger skewness.

6.3 Cumulative Prospect Theory With State-Dependence

6.3.1 Descriptive Statistics

Before proceeding with the results of the structural estimation model described in Section 4.2.1, it would seem natural to start our investigation of the effect of previous betting performance on subsequent risk-taking by conducting some preliminary regression analysis. We therefore estimate two models, one for the decision to play in the sportsbook or not, and one for the characteristics of the lotteries chosen given that the bettor has decided to play. We shall henceforth refer to the former as the participation model and to the latter as the lottery-choice model. We are interested in how both the participation decision and the lotteries chosen are affected by bettors' previous betting performance measured by their cumulative gains over the period of one month ($CumProfit_{nt}$). The participation model is given by:

$$Bet_{nt} = X_{nt}\beta + \varepsilon_{nt}$$

where Bet_{nt} is an indicator variable for whether bettor n has placed at least one bet on day t or not and X_{nt} is a vector of observable covariates including the variable of interest $CumProfit_{nt}$ as well as individual fixed effects. The lottery-choice model is given by:

$$Lottery_{nt} = X_{nt}\alpha + v_{nt}$$

where $Lottery_{nt}$ is the characteristic of the lottery chosen by bettor n on day t, e.g. mean, variance, skewness, etc., and X_{nt} is defined as above. It is important to note here that the variable $Lottery_{nt}$ is observable not only conditional on the bettor having decided to bet in the sportsbook on a given day, but also for all other days of an individual's betting history, where a safe lottery has been selected on days on which we observe no play. Therefore, we do not have to worry about the classic selection bias problem often encountered in this type of analysis and we can estimate the two models independently. We should keep in mind, however, that in these simple regressions it is not easy to account for the confounding factor of play frequency discussed in Section 4.3, which might affect the parameter estimates of the participation model. With this caveat in mind, the results of this section are meant to give us a first indication of the directional effect of previous betting performance on subsequent risk-taking and are by no means conclusive.

Figure 11 presents the results of the pooled probit regression (top panel) and the pooled OLS regression (bottom panel) of the participation model and the lottery-choice model respectively. We observe that previous betting performance significantly affects both the participation decision and the characteristics of the chosen lotteries. In the participation model, we find that the variable $CumProfit_{nt}$ has a significant positive coefficient implying that bettors are more likely continue to play after prior gains than after prior losses. In the lottery choice model, however, we find that the variable $CumProfit_{nt}$ has a significant negative effect on the variance of the chosen lotteries indicating that bettors who have experienced previous gains (losses) subsequently become more risk averse (risk loving). Keeping in mind the caveat associated with the specification of our participation model, we now proceed with the results of our formal estimation model which properly controls for the observed play frequency.

[Figure 11 about here.]

6.3.2 Results

Next, we turn our attention to the estimation of the full econometric model of the CPT specification with state-dependence described in Section 4.2.1. We first employ the Deviance Information Criterion to confirm that the extended model that includes state-dependence is actually to be preferred than the standard CPT specification without state-dependence.

Our estimation results indicate that i) for the majority of bettors the reference point does not update completely to reflect prior outcomes but rather remains sticky to some previous level resulting in prior gains/losses being integrated with current payoffs (the posterior mean of parameter δ is 0.32), ii) the sensitivity of the curvature of the value function on previous betting performance, denoted by α_1, is positive but insignificant, and iii) the sensitivity of the loss aversion parameter on previous betting performance, denoted by λ_1, is positive and significant.

Our finding that the sensitivity to losses increases after prior outcomes implies that bettors become effectively more risk averse (loving) after gains (losses). This finding is consistent with the disposition effect observed in financial markets (e.g. Odean (1998), Zhu and Dhar (2006)), i.e. the tendency of investors to sell stocks that have risen in value, rather than dropped in value, since purchase. At first sight, it also seems to be opposite to the "house money" effect reported in the experiment of Thaler and Johnson (1990) and the game show studies of Gertner (1993) and Post et al. (2008). There is, however, an important caveat to this evidence, which suggests that our finding should not be interpreted as inconsistent with the "house money" effect documented in the aforementioned studies. The "house money" effect focuses primarily on how people behave after previous gains and not after previous losses which is what we usually observe in our gambling data. In fact, by allowing the utility parameters to be just a linear function of previous bet performance, we are not able to capture possible differential effects of previous outcomes on wins versus losses.[29] In fact, given that most bettors in the sportsbook have prior losses, it is possible that our finding is driven by increased risk-taking after prior losses and not necessarily by decreased risk-taking after prior gains. Indeed, Thaler and Johnson (1990) report that prior losses induce risk seeking behavior for gambles that offer bettors the opportunity to break even, and in our setting the possibility of such gambles is abudant.

The behavior documented in our gambling data seems to be inconsistent with some of the assumptions employed by some behavioral finance studies. For instance, Barberis et al. (2001) explain the volatility puzzle assuming that individuals' loss aversion increases after previous losses. In our setting, however, the opposite effect seems to govern gamblers' behavior. Barberis and Xiong (2009)

[29] In the future we are planning to introduce interaction terms that will allow for a different slope in the region of gains and losses.

show that, contrary to conventional wisdom, prospect theory might not explain the disposition effect. The authors argue that a loss averse prospect theory agent is only willing to buy stocks that have a reasonably skewed distribution. They then show that this skewed distribution together with the mild concavity in the value function estimated by Tversky and Kahneman (1992) and the assumption that people integrate the outcomes of successive gambles, implies that the optimal strategy for a prospect theory agent is to increase (decrease) risk after gains (losses). The authors conclude that prospect theory actually predicts the opposite of the disposition effect, i.e. the house money effect. In our setting, we confirm that the average bettor is loss averse with a mildly curved value function and a sticky reference point as standard CPT suggests, but we also find that previous bet outcomes directly affect parameter λ, inducing greater (less) risk aversion after gains (losses). This result indicates that even for skewed distributions a prospect theory agent might in fact display the disposition effect. We are planning to explore this idea further in the future.

7 Extensions

7.1 Mixture of Utilities

Although prospect theory offers a coherent framework for analyzing betting choices, researchers have noted that gambling behavior can also be reconciled with an expected utility model. Several approaches have been proposed to explain the tendency of risk averse people to accept unfair gambles. Friedman and Savage (1948) and Markowitz (1952) suggest that local convexities in an otherwise concave utility function can create a preference for skewness while Golec and Tamarkin (1998) find empirical support for risk averse, skewness loving behavior in the marginal racetrack bettor. Clotfelter and Cook (1989) and Conlisk (1993) explain gambling behavior using a standard expected utility function augmented with a direct utility of gambling that captures the entertainment/excitement associated with casino games. Finally, a stochastic choice behavior that allows for variation in the preferences of a single individual when faced repeatedly with the same choice situation, can explain why people are not always observed to choose the alternative with the highest utility (here the option not to play in the casino) and might as well appear in our sample. On top of that, the availability of near safe lotteries inside the sportsbook further shrinks the utility difference between entering the casino and not.

Since it is not clear that our sample is drawn exclusively from the population of prospect theory agents, in this section we estimate a utility mixture model that allows for differences across population groups in the utility theories that have generated the observed choices. In particular, we assume that our sample is drawn from two latent populations: the population of rational agents whose choices are characterized by the expected utility theory of von Neumann and Morgenstern (1944) and the population of behavioral agents whose choices are characterized by the cumulative prospect theory of Tversky and Kahneman (1992). In this way, we are able to identify and characterize the distribution of risk preferences both across and within the two most widely used theories of choice under risk.

The cumulative prospect theory model is described in Section 4.2. Under the expected utility theory, an agent evaluates a gamble $(p_1, m_1; \ldots; p_n, m_n)$ by assigning it the value:

$$\sum_i p_i u(m_i)$$

where p_i is the probability associated with outcome m_i for $i = 1, \ldots, n$. The functional form for $u(\cdot)$ that we estimate here is a flexible HARA specification, which nests, among other things, both the CRRA and the CARA families:

$$U(m) = \frac{(\rho_0 + W + m)^{1-\rho_1}}{1 - \rho_1} \tag{8}$$

where W represents the level of wealth against which the lottery payoffs are evaluated and ρ_0, ρ_1 are parameters to be estimated. Strictly speaking, under the expected utility framework, the reference wealth level W should be interpreted as individuals' lifetime wealth at the time they place each bet. However, because lifetime wealth is not observable and because some individuals might not integrate their total wealth with the bet payoffs, this measure of W is less suitable in empirical work. We therefore follow Post et al. (2008) and Beetsma and Schotman (2001) and treat W as an individual-specific free parameter that we estimate from the data. Therefore, we can identify parameters $\tilde{\rho_0} = \rho_0 + W$ and ρ_1.[30]

Under the utility mixture model, each of the utility classes is described by a component probability density function, and its mixture weight is the probability that an individual comes from this component[31], i.e. the choices of this individual are best described by the specific utility theory. Mixture modeling has often been applied in Bayesian econometrics to allow for flexible prior distributions of the parameters to be estimated (e.g. Geweke and Keane (1997)) The use of mixture models to compare the validity of theories of choice under risk is a novel approach. There are a few recent experimental studies that use finite mixture models to characterize individual risk taking behavior, but their work differs from ours in several respects. Harrison and Rutström (2009) apply

[30] An alternative would be to approximate lifetime wealth using the census data information on individuals' annual labor income (e.g., Bombardini and Trebbi (2007)). However, one issue with using proxies for income is that their measurement error will be added to the parameter estimates.

[31] This mixing probability can also be related to individuals' observable characteristics.

a utility mixture model by allowing the same subject to behave in accord with one utility class for some choices and in accord with another utility class for other choices. In contrast, in this study we find more reasonable to assume that all choices of the same subject are generated by one *or* the other utility theory. Conte et al. (2009) estimate a mixture model of EUT and rank-dependent EUT using maximum simulated likelihood techniques. In their model, they estimate the population-level parameters of each utility by assuming that the population of each utility type consists of *all* individuals rather than the subset of individuals that are classified to each type. Given their finding that all individuals are clearly associated with one utility class, their estimation procedure might have substantially affected the parameter estimates. Bruhin et al. (2010) estimate a mixture of two prospect theory types using a maximum likelihood-expectation maximization algorithm. In their setup, they allow for heterogeneity only within the prospect theory specification rather than both within and across different utility classes as we do in this study. Furthermore, Bruhin et al. (2010) are not able to identify individual loss aversion since they consider lotteries that contain only positive or only negative prizes. More importantly, since all these studies are based on choices among simple lotteries presented to experimental subjects, it becomes particularly interesting to examine whether the distribution of risk preferences changes when individuals are faced with more complex real-life decision tasks.

Formally, in our utility mixture model, each subject is allocated to one of a set of Q utility classes. Since the true allocation is unknown to the researcher, these are latent classes. The utility U_n that individual n obtains is therefore given by:

$$U_n = \sum_{q=1}^{Q} e_{nq} u_q \left(\rho_{nq}\right),$$

where $u_q(\cdot)$ is the $q-$class utility specification (e.g., prospect theory, etc.) with individual-specific parameter vector ρ_{nq}, and e_{nq} indicates the component utility in the mixture that the $n-$th individual uses, i.e., $e_{nq} = 0$ or 1 and $\sum_{q=1}^{Q} e_{nq} = 1$, $\forall n$.

Letting $\rho_n = (\rho_{n1}, \ldots, \rho_{nQ})'$ be the utility parameter vector, the conditional probability that individual n makes a sequence of choices $\{x_n, y_n\}$ given that he/she belongs to utility class q for $q = 1, \ldots, Q$ can now be written as:

$$L\left(x_n, y_n | \rho_n, e_{nq} = 1, p_n\right) = \prod_{t=1}^{T_n} \left\{x_{nt} p_{nt} L\left(y_{nt} | \rho_{nq}, e_{nq} = 1\right) + \left(1 - x_{nt}\right)\left[p_{nt} L\left(y_{nt} = 0 | \rho_{nq}, e_{nq} = 1\right) + \left(1 - p_{nt}\right)\right]\right\}. \tag{9}$$

In practice, however, it is unknown which component utility the $n-$th individual uses, and thus we let h_q for $q = 1, \ldots, Q$ be the probability of the $q-$th utility being used, i.e., $h_q = P\left(e_{nq} = 1\right)$, $\forall n$

with $0 \leq h_q \leq 1$ and $\sum_q h_q = 1$. Letting $e_n = (e_{n1}, \ldots, e_{nQ})'$, we formally write $e_n|h \sim M(1,h)$, where $M(1,h)$ stands for the multinomial distribution that picks one alternative with probability vector h.

Therefore, the probability of observing the data $\{x_n, y_n\}$ for individual n, conditional on the parameter vector $\beta_n = \binom{\rho_n}{p_n}$ can now be written as:

$$L(x_n, y_n|\beta_n, h) = \sum_{q=1}^{Q} h_q \cdot L(x_n, y_n|\rho_{nq}, e_{nq} = 1, p_n) \tag{10}$$

where $L(x_n, y_n|\rho_{nq}, e_{nq} = 1, p_n)$ is given by Equation 9 above.

The unconditional probability will be the integral of $L(x_n, y_n|\beta_n, h)$ over the latent parameter vector β_n (see Equation 1 above) and as a result the likelihood of observing our data $\{X, Y\}$, where $X = (x'_1, \ldots, x'_N)'$ and $Y = (y'_1, \ldots, y'_N)'$, is given by:

$$L(X, Y|h, \theta) = \prod_n \left(\ln \left(\sum_{q=1}^{Q} h_q \cdot \int_{\beta_n} L(x_n, y_n|\rho_{nq}, e_{nq} = 1, p_n) \cdot f(\beta_n|\theta)\, d\beta_n \right) \right) \tag{11}$$

where $f(\beta_n|\theta)$ is the density function of β_n characterized by parameters θ. In this study, we consider $Q = 2$ utility classes representing the expected utility theory and the prospect theory. To overcome the labeling identification problem that is common in all mixture models, we assume a specific order for the utilities considered, i.e., $q = 1$ corresponds to the HARA specification and $q = 2$ corresponds to the standard cumulative prospect theory specification.

Methodologically, the utility mixture model is estimated using an extension of the hierarchical Bayes model described in Section 5.1. In addition to the hierarchical normal prior used for the parameter vector β_n, we adopt a hierarchical Dirichlet prior for the individual-specific class membership indicator vector e_n, i.e. we assume that $e_n|h \sim M(1,h)$ where $h = (h_1, \ldots, h_Q)$ follows the Dirichlet distribution, i.e., $h \sim D(\underline{h})$. The joint posterior distribution in Equation 7 now becomes:

$$\begin{aligned} K\left(\left\{\tilde{\beta}_n\right\}_{n=1}^{N}, b, W, \hat{\beta}, \{e_n\}_{n=1}^{N}, h|X, Y\right) \quad \propto \\ \prod_n L\left(x_n, y_n|\tilde{\beta}_n, \hat{\beta}, e_n\right) \cdot \left(f_N\left(\tilde{\beta}_n|b, W\right) \cdot k(b, W)\right) \cdot \left(f_M(e_n|h) \cdot k(h)\right) \cdot k\left(\hat{\beta}\right) \end{aligned} \tag{12}$$

where $f_M(e_n|h)$ is the multinomial density that picks one alternative with probability vector h, and $k(h)$ is the Dirichlet prior on h. The likelihood $L\left(x_n, y_n|\tilde{\beta}_n, \hat{\beta}, e_n\right)$ is given by Equation 10 where we have replaced the h_q term with e_{nq} because given e_n we know which component utility bettor n uses. We take draws from the above posterior distribution using the Metropolis-Hastings within Gibbs sampler algorithm analytically described in the Appendix.

7.1.1 Results of Utility Mixture Model

Figure 12 shows the posterior density of the individual mixing probability for the EUT model, i.e. the probability with which a bettor is classified as an EUT agent. We find that the posterior mean of the EUT mixing probability is 14%, the posterior median is 11% and the posterior standard deviation is 0.13.

[Figure 12 about here.]

Figure 13 presents the posterior simulator densities of the utility parameters of the HARA and the CPT specifications. Substantial heterogeneity is present in all utility parameters. The distributions of the CPT utility parameters are similar to the ones we report in our benchmark specification. It should be noted, however, that the CPT estimates include a minority of subjects who are essentially expected value maximizers since they display an almost linear value function ($\alpha > 0.95$), near linear probability weighting function ($\gamma > 0.95$) and no loss aversion ($\lambda < 1.05$). Therefore the proportion of bettors who strictly belong to the EUT class should be somewhat more than the 14% implied by the posterior mean of the mixing probability. This proportion is similar to the one estimated in the experimental studies of Conte et al. (2009) and Bruhin et al. (2010) using a different set of lotteries and different estimation procedures. It is important to note here that by allowing for preference heterogeneity both within and across the two utility theories and given our sample size, we do not expect to find a clear classification of individuals to either utility type. In other words, the fact that almost all mixing probabilities are less than 0.5 does not mean that prospect theory preferences fit the data better than EUT preferences for *all* subjects. On the contrary, the choices of subjects who fall in the right tail of the mixing probability distribution are better explained by the EUT model but the number of observations available in the sample are not enough to move their mixing probabilities too far away from the population mean.

[Figure 13 about here.]

Regarding the parameters of the EUT specification, we observe that the posterior mean of the risk aversion parameter ρ_1 is 2.4 with the posterior median and standard deviation being 2.45 and 2.75 respectively. As mentioned in other parts of the paper, we might have expected an agent with a globally concave utility function to always turn down the negative expected value bets offered in the sportsbook. The positive risk aversion coefficient that we find without imposing any ex ante constraints on its sign is not inconsistent with this observation. First, as Friedman and Savage (1948) suggest, local convexities in bettors' otherwise concave utility functions can create a preference for the skewed lotteries available in the sportsbook. If this is the case, then our estimated positive ρ_1 captures just the general shape of the utility function rather than this exact shape. In addition to this, under the random utility model that we employ, people make their decisions based on a probabilistic ranking of the alternatives available in their choice set. This framework can explain not only why a standard expected utility maximizer will choose to bet in a sportsbook and has therefore not been dropped from our sample but also why we do not estimate a negative ρ_1 for all subjects although that would immediately imply that someone is willing to accept negative expected value bets. To be more specific, a positive ρ_1 under a RUM means that the choices bettors make out of the set of alternative lotteries are more consistent with the rankings implied by a positive ρ_1 than with the rankings implied by a negative ρ_1. To illustrate this idea, consider the following simplified example. Suppose that the choice set among which choice is exercised consists of 3 alternative lotteries: a safe lottery $S := (0, 1)$ that returns payoff 0 with probability 1 and represents no bet, and two risky lotteries that have negative expected value but different variance: a low-variance lottery $R_L := (-1, 0.6; 1, 0.4)$ and a high variance lottery $R_H := (-10, 0.6; 10, 0.4)$[32]. A bettor with a positive ρ_1 would rank these alternatives in the order (S, R_L, R_H) while someone with a negative rho would rank them in the order (R_H, R_L, S). Observing a sequence of choices in which a bettor has chosen the safe lottery $20 - 30\%$ of the time and the low-variance lottery the rest of the time, is more consistent with a positive than with a negative risk aversion coefficient.

8 Robustness Tests

8.1 Prior Sensitivity Analysis

In this section, we perform prior sensitivity analysis, to check whether our results depend crucially on the prior assumptions.

Our baseline priors, for which we have reported results elsewhere, are $\underline{k} = 1$, $\underline{v} = 5$, $\underline{b} = 0 \cdot I$, and $\underline{V} = 1.5I$. Setting $\underline{k}$ and $\underline{v}$ as low as possible (note that $\underline{v}$ needs to be at least as high as the number of parameters we estimate so that it remains proper) makes the priors as weak as possible, and lets the data determine the posteriors. Doing so, renders the choice of $\underline{b}$ immaterial, so we have set it equal to 0 by default. However, the choice of $\underline{V}$ does affect some posterior results, therefore we report below results for different choices of $\underline{V}$. In particular, we report results for $\underline{V} = 15I$, $\underline{V} = 1.5I$, $\underline{V} = 0.015I$, which we refer to as prior 1, prior 2, and prior 3, respectively. Prior 1 assumes almost no variance/heterogeneity in the population, while in the other end of the spectrum, prior 3 assumes

[32]The choice set that we actually use in this study consists of a safe lottery together with a 100 negative expected value lotteries with varying levels of variance.

very large variance/heterogeneity, and prior 2 - our baseline - is somewhere in between.

From Figure 14 below, we see that the posterior population means are largely unaffected by the choice of the prior. For α, the mode (mean) ranges from 0.87 (0.85)to 0.92 (0.90); for λ the mode (mean) ranges from 1.7 (2.1) to 2.2 (2.5); while for γ the mode (mean) ranges from 0.87 (0.87) to 0.89 (0.89). In addition, the variance and skewness of the posterior population mean distributions aren't substantially different, meaning that the HPDIs for the parameters aren't substantially different either.

[Figure 14 about here.]

From Figure 15 below, we see that the posterior population variances are somewhat more affected. More specifically, while the distribution of the posterior population variance for λ is identical for all prior hyperparameter choices, the corresponding distributions for α and particularly for γ are quite affected by the choices of priors. In the case of α, the population variance is estimated to be as low as 4.34 or as high as 6.95. Comparing the prior densities for α with the corresponding posterior densities, we conclude that the true population variance lies very close to 6.95. In the case of γ, the population variance is estimated to be as low as 1.95 or as high as 7.87. Comparing this time the prior densities for γ with the corresponding posterior densities, we can only conclude that the true population variance lies between the bounds of 1.95 and 7.87, which is admittedly a large range. However, we can safely conclude that the variance is bounded far away from 0, and therefore there is significant heterogeneity. It is also comforting to know that despite the substantial effect that the choice of prior has on the posterior estimate of the population variance for γ, all other parameters are largely unaffected.

[Figure 15 about here.]

8.2 Information Errors

In the baseline model presented so far, bettors are assumed to evaluate the gambles available in the sportsbook using the probabilities quoted by the bookmaker. Systematic distortions in people's beliefs with respect to the win probabilities are captured in the cumulative prospect theory specification by the probability transformation function that converts the quoted odds into subjective decision weights. Well-documented biases, such as the favorite-longshot bias revealed in studies of racetrack betting (i.e., the overbetting of longshot outcomes and underbetting of favorite outcomes), have largely been explained by a systematic tendency of individuals to overestimate (underestimate) the chances of low (high) probability outcomes (Thaler and Ziemba (1988), Jullien and Salanie (2000)). However, other factors might as well affect bettors' subjective beliefs with respect to the win probabilities in a non-systematic fashion. In this section we consider an extension of our baseline model in which we allow for idiosyncratic errors in the probabilities people use to evaluate bets.

Strictly speaking, since people's subjective beliefs are not observable, several combinations of risk preferences and beliefs could be used to explain bettors' observed choices, creating an apparent identification problem. Here, following the literature that views sports betting as a primarily risk-taking activity, we choose to explain variation in bettors' choices by variation in their risk preferences, imposing a sensible restriction on the form of their subjective beliefs. This approach is supported by the observation that there is a multitude of bets associated with each event, and therefore even though the event choice might be driven by information reasons, the specific lottery chosen is still informative about bettors' risk preferences. In other words, choosing the specific lotteries that we observe reveals a lot about bettors' risk preferences, even acknowledging that part of their choice might be motivated by information. In particular, we assume that people evaluate bets using the probabilities implied by the posted odds "plus" some individual-match specific error. Letting $1-p_{nut}$ be the quoted win probability for the outcome of event u selected by bettor n on day t, we define bettor n's *perceived* probability of winning this bet as:

$$1 - p_{nut}\varepsilon_{nut}$$

where $0 \leq \varepsilon_{nut} \leq 1$ is an i.i.d. error term drawn from a distribution with support $[0, 1]$. $\varepsilon_{nut} = 1$ implies that the bettor evaluates the gamble using the probability quoted by the bookmaker, while $\varepsilon_{nut} = 0$ implies that the bettor considers the gamble a sure bet. These errors can be interpreted either as individual-specific optimism and skill/information or as individual-match specific probability assessment mistakes.

Since the errors ε_{nut} are associated with individual sporting events but our analysis is carried out at the day level, it is clear that the vector of ε_{nut} is much larger than the number of our datapoints. Therefore, although we are unable to identify ε_{nut}'s for all unit bets chosen, we exploit the idea that these errors are i.i.d. draws from the same population distribution, i.e., $\forall\varepsilon_{nut}$ we have

$$\varepsilon_{nut}|b_\varepsilon, W_\varepsilon \sim N\left(b_\varepsilon, W_\varepsilon\right)$$

where a Johnson's transformation function is used to map the errors to the real line. To the extent that individual-specific optimism and/or skill/information can be linked to observable individual attributes, such as bettors' previous betting performance, we could introduce them in the model by allowing the mean of the error distribution to depend deterministically on them, i.e., $b_\varepsilon = X_n'\gamma$.

Assuming that bettors' beliefs take this specific form gives us the identification power we need to estimate variation in risk preferences from variation in choices across individuals. Having said that, clearly identification in this setting will be weaker than in our baseline case and is expected to manifest itself as higher variance in our preference estimates. Methodologically, we estimate the distribution of the probability errors ε_{nut} by employing the Gibbs sampling techniques with data augmentation described in the Appendix.

8.3 Testing for Skill in Picking Bets

In this section we examine whether there is heterogeneity in skill across bettors in our sample. We calculate the proportion of noise traders who pick bets randomly and the proportion of bettors who exhibit significant positive or negative "skill" in picking their bets. We will refer to the former as "zero-alpha" bettors and to the latter as "skilled" and "unskilled" bettors respectively. Skilled bettors systematically outperform the bookmaker at predicting the outcomes of matches; this could be the case, for example, if some bettors have superior private information about match outcomes while the odds reflect the true win probabilities, if some sophisticated bettors exploit the biases of other bettors that are reflected in the quoted odds, etc.. Unskilled bettors exhibit significantly negative performance; this could be the case, for example, if there exist bettors who suffer from a systematic bias in their beliefs with respect to the win probabilities, e.g. they overestimate the win probability of longshots and thus overbet them, bettors who take on face value the biased odds quoted by the bookmaker, etc.. The identification of both of these types of bettors is important for us since their bet choices might not reflect solely their risk preferences.

The first step in testing for skill heterogeneity across bettors is to create a measure of individual betting performance. Since varying levels of risk are involved in the day lotteries chosen by different bettors and possibly also by the same bettor on different bet days, it may not be appropriate to make comparisons based on absolute performance measures. We therefore create a risk-adjusted betting performance measure, similar to the "alpha" that is widely used in the finance literature. We define "alpha" as the realized daily return of a bettor (i.e. the realized return from all bets placed within the same day) minus the bettor's expected daily return (i.e. the expected return of the day lottery we have constructed by combining all individual bets placed within the same day).

To investigate if and how many individuals in our sample possess true bet-picking skills, we carry out individual-level hypothesis tests by examining whether the average "alpha" of each bettor across all bet days is significantly different from zero.[33] However, when performing a set of hypothesis tests

[33]Since the sample size is small and the normality assumption might fail, we calculate bootstrap p-values for each bettor's alpha coefficient. To do so, we create $B = 1000$ bootstrap samples that satisfy the null hypothesis. That is, in each of these samples, we draw for each bettor and each bet day a possible payoff from the lottery selected on that day, and calculate the corresponding realized return. We then follow the approach of Davidson and MacKinnon

simultaneously (multiple testing), the probability of incorrectly rejecting the null hypothesis (Type I error) is increased. To illustrate, consider, for example, that with one test performed at the 5% significance level, there is only a 5% chance of rejecting the null hypothesis when in fact it is true. Therefore, when performing 100 tests at the 5% level each, we expect 5 of these comparisons to be declared as significant purely due to chance. If the tests are independent, this implies that the probability of at least one incorrect rejection is 99.4%. In order to properly adjust for the problem of "false discoveries" in multiple hypothesis testing, various statistical techniques have been developed. The main idea behind these methods is that a stricter significance threshold is required for the individual tests, in order to compensate for the fact that multiple tests have been performed. In our setting, since the individual tests are independent, we use a variant of the Bonferroni correction (the Šidák correction)[34] according to which the significance level a of individual tests is set such that the probability of at least one false discovery in n tests equals some desired significance level β, i.e.

$$a = 1 - (1-\beta)^{1/n}$$

After making this adjustment with $\beta = 0.05$, the proportion of "skilled" bettors for whom the estimated alpha coefficient is significantly positive is 0 while the proportion of "unskilled" bettors for whom the estimated alpha coefficient is significantly negative is 2.7%. These results indicate that under the odds quoted by the bookmakers, no bettors earn significantly positive excess returns, while there are a few bettors who earn significantly negative excess returns.

8.4 Testing for "Favorite-Team Bias"

One of the concerns associated with sports betting data is that bettors' choices might be driven by their preference towards teams that they support, e.g. home area teams, rather than their risk preferences. Indeed, bettors might choose to bet on their favorite team(s) either because they are getting extra utility from it (for example because they are planning to watch the match live on TV) or because they are emotionally involved and naively think that the probability of their favorite team winning is higher than the one implied by the bookmaker's odds. Both of these cases could

(2004) and compute the p-value as:

$$\hat{p_n} = 2 \cdot min\left(\frac{1}{B}\sum_b I\left\{\hat{t_n^b} > \hat{t_n}\right\}, \frac{1}{B}\sum_b I\left\{\hat{t_n^b} < \hat{t_n}\right\}\right)$$

where $I\left\{\hat{t_n^b} > \hat{t_n}\right\}$ is an indicator function that takes the value 1 if the alpha t-statistic $\hat{t_n^b}$ of indivdual n calculated from the bootstrap sample b, is higher than the corresponding t-statistic, $\hat{t_n}$, calculated from the real sample.

[34] An alternative method which has been recently employed in the finance literature to examine the prevalence of skill in mutual funds and hedge funds (Barras et al. (2010) is the "False Discovery Rate" (FDR) approach. It is a less conservative procedure for comparison, with greater power than the Bonferroni correction, at a cost of increasing the likelihood of obtaining type I errors.

introduce a bias in our risk preference estimates and could be formally captured in our model by allowing the unobserved utility term and the errors in probabilities, respectively, to be a function of observable team characteristics. However, for the sake of reducing the complexity of our model, we present here some features of our data that indicate that the "favorite-team bias" is unlikely to have affected our risk preference estimates.

The first observation is based on the histogram of the fraction of total wagers placed on the favorite team of each individual (see top-left panel of Figure 16). We define the favorite team in two alternative ways. According to the first definition, we identify the team(s) that are based on the bettor's home location: For international matches the favorite team is defined as the national team of the bettors' country of origin and for domestic matches as the team that is based in the bettor's area of residence (according to the bettor's zip code). According to the second definition, for each bettor we identify the team with the highest betting frequency. Under both definitions, we observe that the proportion of unit bets that most bettors have placed on their favorite team seems to be relatively small. In particular, in the bottom two panels of Figure 16 it is clear that the proportion of unit bets placed on the team with the highest betting frequency is comparable to the proportion placed on the team with the second highest betting frequency, indicating that there is no one team that is massively overbet for any bettor. These results imply that, under the odds quoted by the bookmaker, people do not systematically bet on one team, e.g. because they gain extra utility from it, and therefore their choices do not simply reflect the odds usually associated with this team.

[Figure 16 about here.]

The second observation is that people tend to combine more than one bets on different events within the same day (the top-right panel of Figure 16 shows that the proportion of days on which people have bet solely on their favorite team is close to 0 for our median bettor). When combining bets, there is a wide variety of day lotteries that can be constructed, especially if you consider that choosing a day lottery involves a series of sub-decisions that people have to make, i.e. bettors choose the number of matches they want to combine, the type of bet (e.g. single or combination), the event description (e.g. final match outcome, total number of goals, etc.), the odds, the bet amount, etc. Therefore, even though the choice of some or all of these matches might be driven by reasons independent of bettors' risk preferences, we are confident enough that the resulting day lottery can still be informative about bettors' risk preferences.

Finally, we have left for future research the investigation of whether the expected return bettors get from betting on their favorite team is systematically lower than the bookmaker's commission, which would indicate that the particular team has been overbet, i.e. it has been bet more frequently than it is justified by the bookmaker's odds. This could be the case either if bettors hold subjective beliefs with respect to the win probability of their favorite team or if bettors' risk preferences are such that makes them want to bet on the odds usually associated with the particular team.

9 Conclusion

In this paper we have used a unique panel dataset of consumer betting activity in an online sportsbook to examine individual attitudes towards risk. We have developed an econometric model that accounts for individual heterogeneity in risk preferences both *between* and *within* the two dominant theories of choice under risk: the rational choice paradigm of expected utility theory and the behavioral paradigm of prospect theory. The panel aspect of our dataset allowed us i) to estimate individual-level risk preference parameters with relative accuracy, ii) to test for the presence of heterogeneity in risk preferences across individuals and iii) to examine possible state dependence in bettors' risk-taking behavior, which would arise if, for example, subsequent risk-taking is affected by previous bet outcomes. Our findings suggest that the majority of bettors in our sample are likely to depart from full rationality when presented with complex real-life decisions. However, there is significant heterogeneity across individuals regarding the aspect of rational preferences they violate: state-dependence, probability weighting, loss aversion and utility curvature vary across bettors. Methodologically, we have used a discrete choice framework and estimated a multinomial mixed logit model using Bayesian econometrics techniques. Our findings have substantial implications for microfounded models in economics and, in particular, in finance, since gambling decisions share significant similarities with stock trading decisions.

Appendix

A. Estimation Algorithm

A1. Gibbs Sampler for the Utility Mixture Model

In this section, we describe the MCMC methods used for the estimation of the finite mixture model of EUT and CPT presented in Section 7.1. The estimation of our benchmark specification of CPT alone is a direct simplification of the procedure descibed in this section.

In the utility mixture model, the parameters for which we would like to have a posterior distribution are: i) the utility parameter vector $\rho = (\rho_1', \ldots, \rho_N')'$ where $\rho_n = (\rho_{n1}, \ldots, \rho_{nQ})'$, ii) the exogenous probability of having the opportunity to bet $p = (p_1', \ldots, p_N')'$, and iii) the dummy variable vector that indicates the component utility used in the mixture $e = (e_1', \ldots, e_N')'$

where $e_n = (e_{n1}, \ldots, e_{nQ})'$. The vector of individual-specific parameters associated with EUT is $\rho_{n1} = \left(k_n^1, \tilde{\rho_0}, \rho_1\right)'$where k_n^1 is the scale of the utility, $\rho_1$is the risk aversion parameter and $\tilde{\rho_0}$ is a constant plus the level of wealth that bettors integrate with the current bet payoffs, and the vector of individual-specific parameters associated with CPT is $\rho_{n2} = \left(k_n^2, \alpha_n, \lambda_n, \gamma_n, \delta_n\right)'$ where k_n^2 is the scale of the utility, α_n is the curvature of the utility function, λ_n is the loss aversion coefficient, γ_n is the curvature of the probability weighting function and δ_n measures the degree to which the reference point that separates gains from losses is affected by previous bet outcomes.

In each iteration $k+1$, the Gibbs sampler produces sequential draws from the following conditional distributions:

$$K\left(b, W|\tilde{\beta}_n^{(k)}, e_n^{(k)}\right) \tag{13}$$

$$K\left(\tilde{\beta}_n|x_n, y_n, b^{(k+1)}, W^{(k+1)}, \hat{\beta}^{(k)}, e_n^{(k)}\right) \tag{14}$$

$$K\left(\hat{\beta}|x_n, y_n, \tilde{\beta}_n^{(k+1)}, b^{(k+1)}, W^{(k+1)}, e_n^{(k)}\right) \tag{15}$$

$$K\left(h|e_n^{(k)}\right) \tag{16}$$

$$K\left(e_n|x_n, y_n, \tilde{\beta}_n^{(k+1)}, \hat{\beta}^{(k)}, h^{(k+1)}\right) \tag{17}$$

The conditional densities for each block of the Gibbs sampler take the following forms:[35]

1. The conditional density in Equation 13 follows the normal-inverse Wishart distribution, i.e.:

$$K\left(b, W|\tilde{\beta}_n, e_n\right) \sim N - IW\left(\overline{b}, \overline{k}, \overline{v}, \overline{V}\right)$$

where the posterior hyper-parameters for utility class $q \in \{1, 2\}$ are given by:

$$\begin{aligned}
\overline{b_q} &= \frac{\left(\sum_n e_{nq}\right)\left(\sum_n \left(e_{nq}\tilde{\beta}_{nq}\right)\right) + \underline{k_q}\underline{b_q}}{\overline{k_q}} \\
\overline{k_q} &= \underline{k_q} + \sum_n e_{nq} \\
\overline{v_q} &= \underline{v_q} + \sum_n e_{nq} \\
\overline{V_q} &= \underline{V_q} + \sum_n \left(\tilde{\beta}_{nq} - \tilde{\beta_{nq}^*}\right)\left(\tilde{\beta}_{nq} - \tilde{\beta_{nq}^*}\right)' + \frac{\underline{k_q}}{\overline{\overline{k_q}}}\left(\sum_n e_{nq}\right)\left(\tilde{\beta_{nq}^*} - \underline{b_q}\right)\left(\tilde{\beta_{nq}^*} - \underline{b_q}\right)'
\end{aligned}$$

where $\tilde{\beta_{nq}^*} = \sum_n \left(e_{nq}\tilde{\beta}_{nq}\right)$ is the sample mean of $\tilde{\beta}_{nq}$ and $\underline{b_q}$, $\underline{k_q}$,$\underline{v_q}$, $\underline{V_q}$ are the utility-specific

[35] We have suppressed the k iteration superscript for notational convenience.

prior hyper-parameters of (b, W).[36]

2. The conditional posterior in Equation 14 can be expressed as proportional to the product of the likelihood function and the prior density of β_n:

$$K\left(\tilde{\beta}_n|x_n, y_n, b, W, \hat{\beta}, e_n\right) \propto L\left(x_n, y_n|\tilde{\beta}_n, \hat{\beta}, e_n\right) \cdot f_N\left(\tilde{\beta}_n|b, W\right)$$

Since this conditional posterior does not have a convenient shape, to make draws from it we use a Gaussian random-walk Metropolis-Hastings algorithm (see Chib and Greenberg (1995)). To produce a draw from some target density, this MCMC algorithm simulates a candidate draw $\tilde{\beta}_n^* = \tilde{\beta}_n + \sigma L\eta$ where $\sigma L\eta$ is a jumping factor with $\eta \sim N(0, I)$, σ a positive parameter specified by the researcher[37] and L the Cholesky factor of W. This candidate draw is accepted with a probability that depends on the ratio of the posterior density at the proposed and the current draw. If the candidate draw is not accepted, the algorithm retains the current draw and proceeds to the next block of the Gibbs sampler.

3. The conditional posterior in Equation 15 is proportional to the product of the likelihood across all individuals and the prior density of $\hat{\beta}$. Under the assumed flat prior for $\hat{\beta}$ we derive that:

$$K\left(\hat{\beta}|x_n, y_n, \tilde{\beta}_n, b, W, e_n\right) \propto \left(\prod_n L\left(x_n, y_n|\tilde{\beta}_n, \hat{\beta}, e_n\right)\right)$$

Drawing from this posterior density requires another application of the Metropolis-Hastings algorithm on the pooled data this time.

4. The conditional posterior in Equation 16 follows the Dirichlet distribution, i.e.:

$$K(h|e_n) \sim D\left(\bar{h}\right)$$

[36] For the probability parameter p_n which is included in $\tilde{\beta}_n$ but is independent of the utility specification used, the posterior hyper-parameters are given by:

$$\begin{aligned}
\overline{b_p} &= \frac{N\left(\sum_n p_n\right) + \underline{k_p}\underline{b_p}}{\overline{k_p}} \\
\overline{k_p} &= \underline{k_p} + N \\
\overline{v_p} &= \underline{v_p} + N \\
\overline{V_p} &= \underline{V_p} + \sum_n \left(p_n - \sum_n p_n\right)\left(p_n - \sum_n p_n\right)' + \frac{\underline{k_p}}{\overline{\overline{k_p}}} N \left(\sum_n p_n - \underline{b_p}\right)\left(\sum_n p_n - \underline{b_p}\right)'
\end{aligned}$$

[37] As recommended by Gelman et al. (1995) we adjust σ in each iteration of the Gibbs sampler based on the acceptance rate among the N trial draws of $\beta_n, \forall n$ in the previous iteration. In particular, we lower σ if the acceptance rate is below .2 and raise it if the rate is above .2.

where it can be shown that $\bar{h} = \underline{h} + \sum_{i=1}^{N} e_n$.

5. The conditional posterior in Equation 17 can be expressed as proportional to the product of the likelihood function and the prior density of e_n:

$$K\left(e_n | x_n, y_n, \tilde{\beta}_n, \hat{\beta}, h\right) \propto L\left(x_n, y_n | \tilde{\beta}_n, \hat{\beta}, e_n\right) \cdot f_M\left(e_n | h\right)$$

where

$$f_M\left(e_n | h\right) = M\left(1, h\right) = \frac{1}{e_{n1}! \dots e_{nQ}!} h_1^{e_{n1}} \dots h_Q^{e_{nQ}}.$$

If, say, $e_{nq} = 1$ and $e_{n,-q} = 0$, then $p\left(e_{nq} = 1 | h\right) = h_q$ while $L\left(x_n, y_n | \tilde{\beta}_n, \hat{\beta}, e_{nq} = 1\right)$ is given by Equation 9, i.e.:

$$L\left(x_n, y_n | \tilde{\beta}_n, \hat{\beta}, e_{nq} = 1\right) =$$

$$\prod_{t=1}^{T_n} \left\{ x_{nt} p_{nt} L\left(y_{nt} | \rho_{nq}, e_{nq} = 1\right) + \left(1 - x_{nt}\right) \left[p_{nt} L\left(y_{nt} = 0 | \rho_{nq}, e_{nq} = 1\right) + \left(1 - p_{nt}\right) \right] \right\}$$

and so

$$K\left(e_{nq} = 1 | x_n, y_n, \tilde{\beta}_n, \hat{\beta}, h\right) \propto L\left(x_n, y_n | \tilde{\beta}_n, \hat{\beta}, e_{nq} = 1\right) \cdot h_q.$$

So we can write

$$K\left(e_n | x_n, y_n, \tilde{\beta}_n, \hat{\beta}, h\right) \propto \prod_{q=1}^{Q} h_q \cdot \left(L\left(x_n, y_n | \tilde{\beta}_n, \hat{\beta}, e_{nq} = 1\right)\right)^{e_{nq}},$$

which is the kernel of the multinomial distribution

$$M\left(1, \frac{h_1 \cdot L\left(x_n, y_n | \tilde{\beta}_n, \hat{\beta}, e_{n1} = 1\right)}{\sum_{q=1}^{Q} h_q \cdot L\left(x_n, y_n | \tilde{\beta}_n, \hat{\beta}, e_{nq} = 1\right)}, \dots, \frac{h_Q \cdot L\left(x_n, y_n | \tilde{\beta}_n, \hat{\beta}, e_{nQ} = 1\right)}{\sum_{q=1}^{Q} h_q \cdot L\left(x_n, y_n | \tilde{\beta}_n, \hat{\beta}, e_{nq} = 1\right)}\right).$$

A2. Gibbs Sampler for the Model with Information Errors

In this section we describe the Gibbs sampling techniques we employ for the estimation of our extension model that includes the information errors ε_{nut}. We treat ε_{nut} as additional latent variables with population parameters b_ε and W_ε and explore the posterior distribution of all model parameters and latent variables by sequentially drawing from their conditional posteriors. Assuming a hierarchical normal prior for ε_{nut} where the population-level parameters $(b_\varepsilon, W_\varepsilon)$ are distributed according to the normal-inverse Wishart distribution, our Gibbs sampler produces draws from the conditional distributions described in the previous section (now also conditioned on $\varepsilon_{nut}, b_\varepsilon, W_\varepsilon$) and from the following conditional distributions:

1. We draw $(b_\varepsilon, W_\varepsilon)$ from:

$$K\left(b_\varepsilon, W_\varepsilon | \varepsilon_{nut}\right) \sim N - IW\left(\overline{b_\varepsilon}, \overline{k_\varepsilon}, \overline{v_\varepsilon}, \overline{V_\varepsilon}\right)$$

   where the posterior hyper-parameters are given by:

$$\begin{aligned}
\overline{b_\varepsilon} &= \frac{N\left(\sum_n \varepsilon_n\right) + \underline{k_\varepsilon b_\varepsilon}}{\overline{k_\varepsilon}} \\
\overline{k_\varepsilon} &= \underline{k_\varepsilon} + \sum_n \varepsilon_n \\
\overline{v_\varepsilon} &= \underline{v_\varepsilon} + \sum_n \varepsilon_n \\
\overline{V_\varepsilon} &= \underline{V_\varepsilon} + \sum_n \left(\varepsilon_n - \sum_n \varepsilon_n\right)\left(\varepsilon_n - \sum_n \varepsilon_n\right)' + \frac{\underline{k_\varepsilon}}{\overline{k_\varepsilon}}\left(\sum_n e_{nq}\right)\left(\sum_n \varepsilon_n - \underline{b_\varepsilon}\right)\left(\sum_n \varepsilon_n - \underline{b_\varepsilon}\right)'
\end{aligned}$$

   where ε_n is a randomly selected error for individual n and $\underline{b_\varepsilon}$,$\underline{k_\varepsilon}$,$\underline{v_\varepsilon}$,$\underline{V_\varepsilon}$ are the prior hyper-parameters of $(b_\varepsilon, W_\varepsilon)$.

2. For each n,u,t, we draw ε_{nut} from:

$$K\left(\varepsilon_{nut} | x_n, y_n, \tilde{\beta}_n, b, W, \hat{\beta}, e_n, b_\varepsilon, W_\varepsilon\right) \propto L\left(x_n, y_n | \tilde{\beta}_n, \hat{\beta}, e_n, \varepsilon_{nut}\right) \cdot f_N\left(\varepsilon_{nut} | b_\varepsilon, W_\varepsilon\right)$$

   To draw from this posterior we use the Gaussian random-walk Metropolis-Hastings algorithm specified above.

References

Ali, M. M., 1977. Probability and utility estimates for racetrack bettors. The Journal of Political Economy 85 (4), 803–815.

Barberis, N., 2010. A model of casino gambling, mimeo.

Barberis, N., Huang, M., December 2008. Stocks as lotteries: The implications of probability weighting for security prices. American Economic Review 98, 2066–2100.

Barberis, N., Huang, M., Santos, T., February 2001. Prospect theory and asset prices. The Quarterly Journal of Economics 116 (1), 1–53.

Barberis, N., Huang, M., Thaler, R. H., September 2006. Individual preferences, monetary gambles, and stock market participation: A case for narrow framing. American Economic Review 96 (4), 1069–1090.

Barberis, N., Xiong, W., April 2009. What drives the disposition effect? an analysis of a long-standing preference-based explanation. Journal of Finance 64 (2), 751–784.

Barseghyan, L., Molinari, F., O'Donoghue, T., Teitelbaum, J. C., 2010. The nature of risk preferences: Evidence from insurance choices, mimeo.

Bassett, Gilbert W., J., 1981. Point spreads versus odds. The Journal of Political Economy 89 (4), 752–768.

Beetsma, R. M. W. J., Schotman, P. C., 2001. Measuring risk attitudes in a natural experiment: Data from the television game show lingo. The Economic Journal 111 (474), 821–848.

Benartzi, S., Thaler, R. H., February 1995. Myopic loss aversion and the equity premium puzzle. The Quarterly Journal of Economics 110 (1), 73–92.

Bhamra, H., Uppal, R., 2010. Asset prices with heterogeneity in preferences and beliefs, mimeo.

Bombardini, M., Trebbi, F., 2007. Risk aversion and expected utility theory: A field experiment with large and small stakes, mimeo.

Bruhin, A., Fehr-Duda, H., Epper, T., July 2010. Risk and rationality: Uncovering heterogeneity in probability distortion. Econometrica 78 (4), 1375–1412.

Brunnermeier, M. K., Nagel, S., June 2008. Do wealth fluctuations generate time-varying risk aversion? micro-evidence on individuals' asset allocation. American Economic Review 98 (3), 713–736.

Campbell, J. Y., Cochrane, J. H., April 1999. By force of habit: A consumption-based explanation of aggregate stock market behavior. The Journal of Political Economy 107 (2), 205–251.

Chan, Y. L., Kogan, L., December 2002. Catching up with the joneses: Heterogeneous preferences and the dynamics of asset prices. The Journal of Political Economy 110 (6), 1255–1285.

Chetty, R., 2006. A new method of estimating risk aversion. American Economic Review 96 (5), 1821–1834.

Chib, S., Greenberg, E., 1995. Understanding the metropolis-hastings algorithm. The American Statistician 49 (4), 327–335.

Choi, S., Fisman, R., Gale, D., Kariv, S., 2007. Consistency and heterogeneity of individual behavior under uncertainty. The American Economic Review 97 (5), 1921–1938.

Cicchetti, C. J., Dubin, J. A., 1994. A microeconometric analysis of risk aversion and the decision to self-insure. Journal of Political Economy 102 (1), 169–186.

Clotfelter, C. T., Cook, P. J., Apr. 1989. The demand for lottery products. Nber working papers, National Bureau of Economic Research, Inc.

Cohen, A., Einav, L., 2007. Estimating risk preferences from deductible choice. American Economic Review 97 (3), 745–788.

Conlisk, J., June 1993. The utility of gambling. Journal of Risk and Uncertainty 6 (3), 255–75.

Constantinides, G. M., June 1990. Habit formation: A resolution of the equity premium puzzle. Journal of Political Economy 98 (3), 519–43.

Conte, A., Hey, J. D., Moffatt, P. G., 2009. Mixture models of choice under risk. Journal of Econometrics In Press, Corrected Proof.

Daly, A., 1982. Estimating choice models containing attraction variables. Transportation Research 16B, 5–15.

Donkers, B., Melenberg, B., Van Soest, A., 2001. Estimating risk attitudes using lotteries: A large sample approach. Journal of Risk and Uncertainty 22 (2), 165–95.

Dumas, B., 1989. Two-person dynamic equilibrium in the capital market. Review of Financial Studies 2 (2), 157–88.

Figlewski, S., 1979. Subjective information and market efficiency in a betting market. The Journal of Political Economy 87 (1), 75–88.

Friedman, M., Savage, L. J., 1948. The utility analysis of choices involving risk. Journal of Political Economy 56, 279.

Gelfand, A. E., Smith, A. F. M., 1990. Sampling-based approaches to calculating marginal densities. Journal of the American Statistical Association 85 (410), 398–409.

Gertner, R., 1993. Game shows and economic behavior: Risk-taking on "card sharks". Quarterly Journal of Economics 108 (2), 507–521.

Geweke, J., 2004. Getting it right: Joint distribution tests of posterior simulators. Journal of the American Statistical Association 99 (467), 799–804.

Geweke, J. F., Keane, M. P., 1997. Mixture of normals probit models. Tech. rep.

Golec, J., Tamarkin, M., 1998. Bettors love skewness, not risk, at the horse track. The Journal of Political Economy 106 (1), 205–225.

Ågren, M., Oct. 2006. Prospect theory and higher moments. Working Paper Series 2006:24, Uppsala University, Department of Economics.

Harrison, G., Rutström, E., June 2009. Expected utility theory and prospect theory: one wedding and a decent funeral. Experimental Economics 12 (2), 133–158.

Harrison, J. M., Kreps, D. M., May 1978. Speculative investor behavior in a stock market with heterogeneous expectations. The Quarterly Journal of Economics 92 (2), 323–36.

Hey, J. D., Orme, C., 1994. Investigating generalizations of expected utility theory using experimental data. Econometrica 62 (6), 1291–1326.

Holt, C. A., Laury, S. K., 2002. Risk aversion and incentive effects. American Economic Review 92 (5), 1644–1655.

Ingersoll, J., 2008. Non-monotonicity of the tversky-kahneman probability-weighting function: A cautionary note. European Financial Management 14 (3), 385–390.

Jullien, B., Salanie, B., 2000. Estimating preferences under risk: The case of racetrack bettors. Journal of Political Economy 108 (3), 503–530.

Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decisions under risk. Econometrica 47 (2), 263–292.

Kopriva, F., 2009. Risk aversion estimates from betting markets, mimeo.

Manski, C. F., 1977. The structure of random utility models. Theory and Decision 8, 229–254.

Markowitz, H., 1952. The utility of wealth. Journal of Political Economy 60, 151.

McFadden, D., Train, K., 2000. Mixed mnl models for discrete response. Journal of Applied Econometrics 15 (5), 447–470.

Mehra, R., Prescott, E. C., 2003. The equity premium in retrospect. In: Constantinides, G., Harris, M., Stulz, R. M. (Eds.), Handbook of the Economics of Finance. Vol. 1 of Handbook of the Economics of Finance. Elsevier, Ch. 14, pp. 889–938.

Metrick, A., 1995. A natural experiment in "jeopardy!". American Economic Review 85 (1), 240–253.

Odean, T., October 1998. Are investors reluctant to realize their losses? Journal of Finance 53 (5), 1775–1798.

Paravisini, D., Rappoport, V., Ravina, E., 2010. Risk aversion and wealth: Evidence from person-to-person lending portfolios, mimeo.

Post, T., van den Assem, M. J., Baltussen, G., Thaler, R. H., 2008. Deal or no deal? decision making under risk in a large-payoff game show. American Economic Review 98 (1), 38–71.

Quandt, R. E., 1986. Betting and equilibrium. The Quarterly Journal of Economics 101 (1), 201–208.

Rabin, M., Thaler, R. H., 2001. Anomalies: Risk aversion. Journal of Economic Perspectives 15 (1), 219–232.

Rachev, S., 1991. Probability metrics and the stability of stochastic models. John Wiley & Son Ltd.

Revelt, D., Train, K., November 1998. Mixed logit with repeated choices: Households' choices of appliance efficiency level. The Review of Economics and Statistics 80 (4), 647–657.

Snowberg, E., Wolfers, J., May 2010. Explaining the favorite-longshot bias: Is it risk-love or misperceptions?

Swait, J., 2001. Choice set generation within the generalized extreme value family of discrete choice models. Transportation Research Part B: Methodological 35 (7), 643–666.

Tanner, M. A., Wong, W. H., 1987. The calculation of posterior distributions by data augmentation. Journal of The American Statistical Association 82, 528–549.

Thaler, R. H., Johnson, E. J., 1990. Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. Management Science 36 (6), 643–660.

Thaler, R. H., Ziemba, W. T., 1988. Anomalies: Parimutuel betting markets: Racetracks and lotteries. The Journal of Economic Perspectives 2 (2), 161–174.

Train, K., 2001. A comparison of hierarchical bayes and maximum simulated likelihood for mixed logit. Tech. rep.

Train, K. E., McFadden, D. L., Ben-Akiva, M., 1987. The demand for local telephone service: A fully discrete model of residential calling patterns and service choices. The RAND Journal of Economics 18 (1), pp. 109–123.

Tversky, A., Kahneman, D., 1992. Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty 5 (4), 297–323.

von Neumann, J., Morgenstern, O., 1944. The Theory of Games and Economic Behavior. Princeton University Press, Princeton, NJ.

Wang, J., 1996. The term structure of interest rates in a pure exchange economy with heterogeneous investors. Journal of Financial Economics 41 (1), 75 – 110.

Weisbrod, G. E., Lerman, S. R., Ben-Akiva, M. E., 1980. Tradeoffs in residential location decisions: transportation versus other factors. Transport Policy and Decision Making 1, 13–26.

Weitzman, M., 1965. Utility analysis and group behavior: An empirical study. The Journal of Political Economy 73 (1), 18–26.

Woodland, B. M., Woodland, L. M., 1991. The effects of risk aversion on wagering: Point spread versus odds. The Journal of Political Economy 99 (3), 638–653.

Zhu, N., Dhar, R., 2006. Up close and personal: An individual level analysis of the disposition effect. Management Science 52 (5), 726–740.

List of Figures

BETTING SLIP						
Event	**Sport**	**Selection**	**Bet Type**	**Odds**	**Total Stake**	**Potential Winnings**
USA vs Germany (FIBA World Championship) Match Winner	Basketball	Germany	Single	1/3	50	66.67
Man United vs Liverpool (English Premier League) Match Winner	Soccer	Draw	Double	3/1	2	
Roger Federer vs Diego Hartfield (Australian Open 1st Round) 1st set winner	Tennis	Federer	Double	3/5	2	12.80
Man United vs Liverpool (English Premier League) Total Goals Over/Under 2.5	Soccer	Over 2.5	Doubles	5/4	30	
New York Yankees vs Toronto Blue Jays (MLB Regular Season) Total Match Runs Odd/Even	Baseball	Odd	Doubles	4/6	30	
Phil Taylor vs Ronnie Baxter (World Grand Prix 1st rnd) Match Winner	Darts	Ronnie Baxter	Doubles	1/10	30	80.58

Figure 1: Example of a betting slip containing the selections made by a sports bettor. The shaded lines denote related bets and cannot be combined in a multiple bet.

	Variable	Mean	Std. Dev.	Min	Max
Demographics:					
	Female	0.04	0.17	0	1
	Age	32.26	8.61	19	57
	Income	18990	4282	8156	35474
Event Characteristics:					
	Commission	0.09	0.029	0.02	0.35
	Events In Bet Day	9.96	9.25	1	315
	In-running	0.05	0.22	0	1
	Days to event date	0.33	0.75	0	6
	Favorite Team	0.07	0.09	0	1
Bet Type Attributes:					
	Bet types In Bet Day	3.7	2.14	1	25
Bet Day Attributes:					
	Bet Days per Bettor	37.55	46.89	1	274
	Bet days per year	21.23	23.07	1	122
	Weekend	0.45	0.5	0	1
	CumPNL	-2.44	58.81	-1500	1318

Figure 2: Summary Statistics of Lottery Characteristics.

Bet Type	Number of winning selections	Unit Stake	Odds			Prizes		
Single	1	50	1/3			66.67		
Double	2	2	(3/1)*(3/5)			12.80		
Doubles	2	30/3	(5/4)*(4/6)	(5/4)*(1/10	4/6)*(1/10	37.50	24.75	18.33

Figure 3: Calculation of odds and prizes for possible winning outcomes of sample bets in the betting slip in Figure 1.

Variable	Percentiles										
	1	**10**	**20**	**30**	**40**	**50**	**60**	**70**	**80**	**90**	**99**
Number of prizes	2	2	2	2	2	2	4	4	8	24	237.56
Expected Value	-64.25	-16.91	-7.70	-4.68	-3.17	-2.20	-1.48	-0.98	-0.58	-0.29	-0.01
Expected Return	-0.7	-0.41	-0.34	-0.29	-0.26	-0.23	-0.19	-0.15	-0.11	-0.09	-0.05
Variance	0.01	9.31	32.95	80.74	165.39	308.39	571.60	1134.31	2929.75	13594.27	161846.72
Skewness	-1.73	-0.08	0.35	0.88	1.45	2.10	2.86	3.88	5.78	10.48	99.78
Kurtosis	1.00	1.22	2.00	2.75	3.96	6.36	10.36	17.85	38.56	128.59	11588.00
Bet Amount	-310.98	-100.00	-44.00	-26.00	-15.00	-10.00	-7.00	-5.00	-3.00	-1.00	-0.05
Maximum Prize	0.24	5.00	14.25	24.48	40.37	62.20	96.77	160.87	278.27	722.99	12002.00
Minimum Probability	4.59E-10	4.66E-05	0.00	0.01	0.02	0.04	0.08	0.14	0.23	0.38	0.49
Maximum Probability	0.08	0.30	0.50	0.57	0.67	0.75	0.83	0.89	0.94	0.98	1.00

Figure 4: Summary statistics of lottery characteristics. All prizes are in Euros; Bet Amount is the minimum prize in the lottery; and Expected Return is Expected Value divided by Bet Amount.

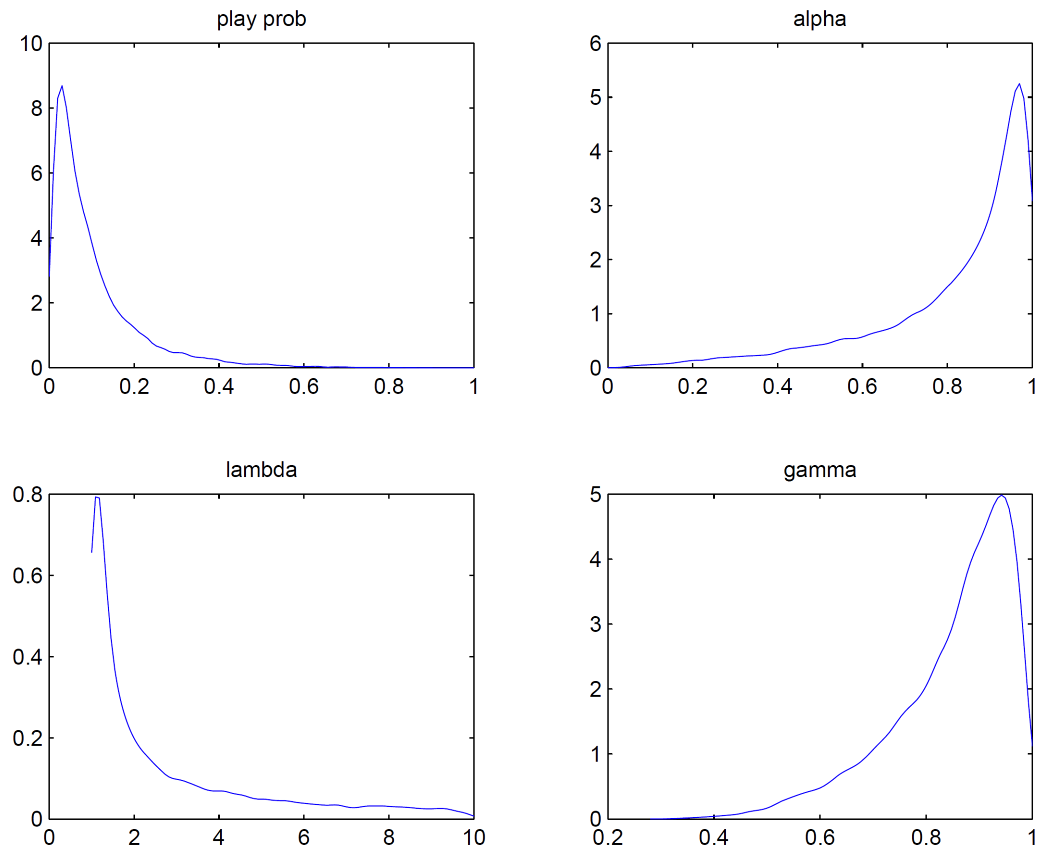

Figure 5: Posterior Simulator Distribution for $\tilde{\beta}_n$.

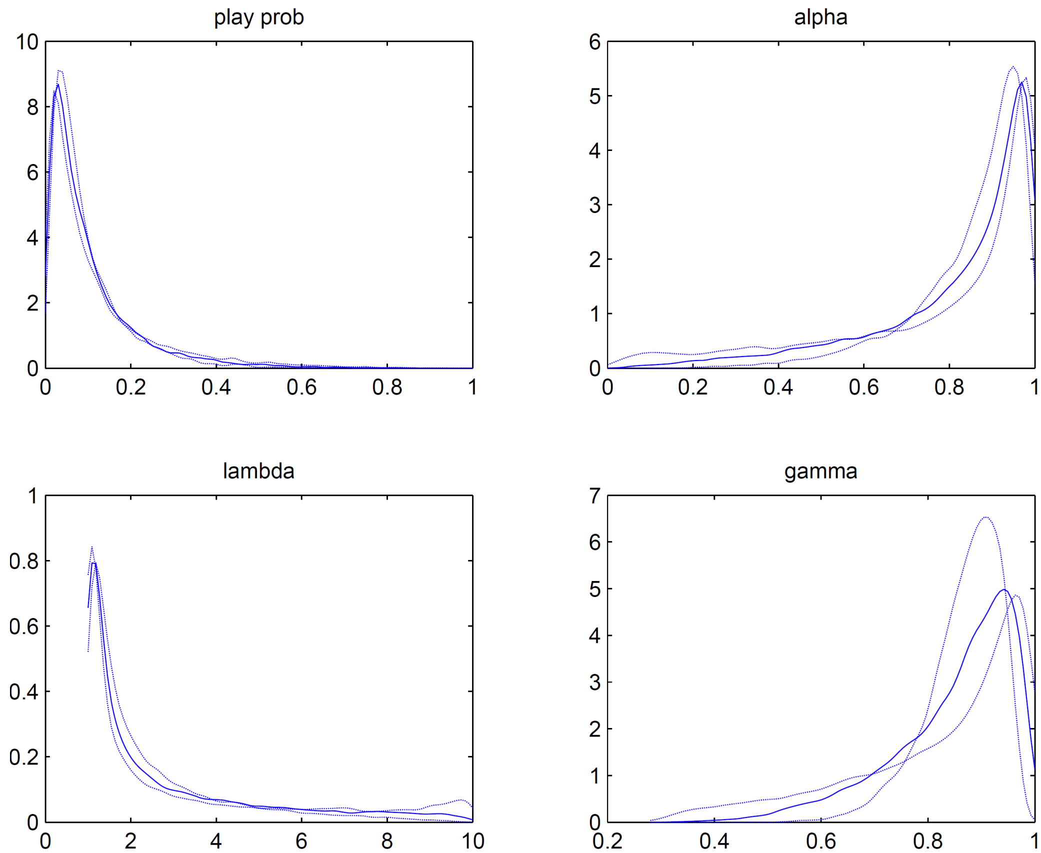

Figure 6: Posterior Simulator Distribution for $\tilde{\beta}_n$, for bounds of HPDIs for Mean.

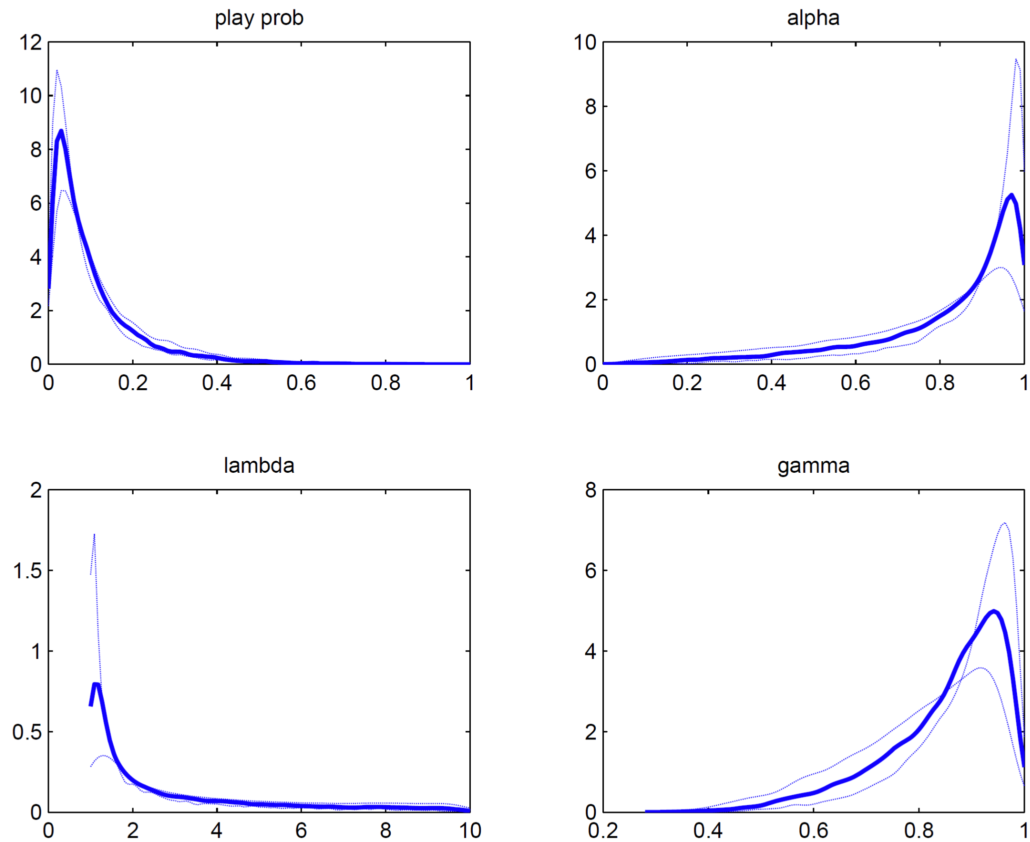

Figure 7: Posterior Simulator Distribution for $\tilde{\beta}_n$, for bounds of HPDIs for Variance.

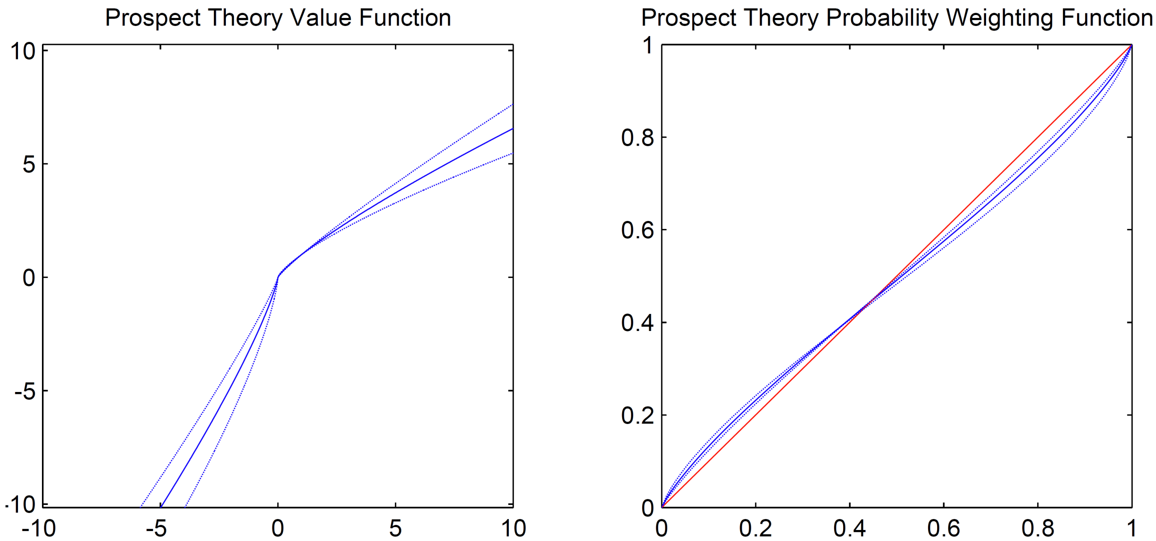

Figure 8: Value and probability weighting functions for the posterior estimates of CPT parameters and for their 95% HPDIs (dotted).

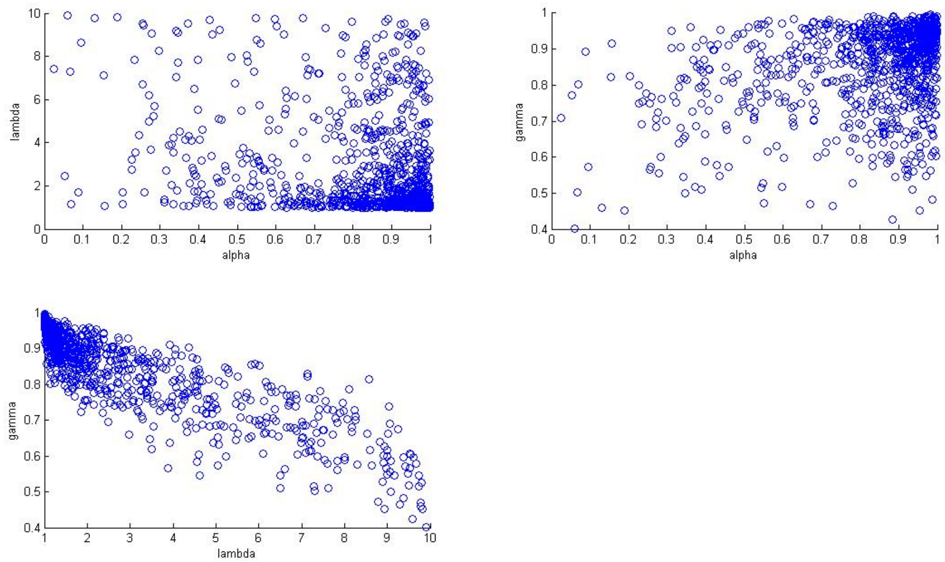

Figure 9: Scatterplot showing correlations among Cumulative Prospect Theory parameters.

Panel A

	Variance	Variance	Variance Pos	Variance Pos	Variance Neg	Variance Neg
Constant	5.846 ***	-0.993	7.061 ***	9.349 *	3.203 ***	-1.809
	14.158	-1.191	7.907	1.833	6.274	-0.675
Alpha	0.187	-1.137 ***	1.961	-1.586	-3.196 ***	-2.583 *
	0.286	-2.633	1.254	-0.600	-3.575	-1.860
Lambda	-0.493	0.353	2.365	0.823	0.684	1.017
	-0.439	0.746	0.974	0.285	0.493	0.669
Gamma	-0.745	1.345	7.944	10.880	15.317 **	10.841
	-0.138	0.535	0.679	0.708	2.290	1.341
Bet Amount		1.924 ***		1.527 ***		1.402 ***
		26.032		3.377		5.898
Expected Return		0.071		-4.160 *		0.312
		0.183		-1.754		0.250
Skewness		1.602 ***		1.024		0.609
		8.224		0.859		0.972
R squared	0.95%	89.83%	0.14%	20.82%	21.90%	45.76%

Panel B

	Skewness	Skewness
Constant	0.778 ***	2.029 ***
	7.641	7.419
Alpha	2.482 ***	1.118 ***
	13.915	7.791
Lambda	-0.732 ***	-0.824 ***
	-2.641	-4.631
Gamma	-10.318 ***	-5.110 ***
	-7.730	-5.693
Bet Amount		-0.582 ***
		-8.234
Expected Return		-0.711 ***
		-4.963
Variance		0.286 ***
		8.224
R squared	75.49%	92.62%

Figure 10: OLS regressions of Lottery Characteristics on Estimated Preference Parameters. In Panel A, the dependent variables are the average Variance (Columns 1 and 2), average Variance among positive (Columns 3 and 4), and average Variance among negative outcomes (Columns 5 and 6), for individuals' chosen lotteries. In Panel B, the dependent variable is average Skewness for individuals' chosen lotteries. The main explanatory variables are the estimated preference parameters: α the curvature, λ the loss aversion, and γ the probability weighting. Alternative specifications including as explanatory variables other lottery characteristics are presented. $t-$statistics are reported below the OLS estimates. */**/*** indicate significance at the 10%/5%/1% levels.

Panel A

	Bet/No Bet Dummy
Constant	-0.9278 ***
	-3.3447
CumProfit	0.0002 ***
	4.3107
Pseudo R-squared	13.58%

Panel B

	Expected Value	Variance	Skewness	Bet Amount
Constant	0.038	-2300.000 **	-29.990 ***	-0.098
	0.212	-2.838	-5.659	-0.091
CumProfit	-0.001 ***	-16.184 ***	0.008	0.013 ***
	-2.380	-6.957	0.510	4.131
R squared	45.45%	31.44%	11.62%	35.60%

Figure 11: Panel A presents a probit regression of the probability that an individual makes a bet on a given day on past profits. Panel B presents OLS regressions of lottery characteristics on the past profits. In particular, in Panel B the dependent variables are the Expected Value, the Variance, the Skewness, and the Bet Amount of the lottery chosen by an individual on a given day. The main explanatory variable is past profits, measured as cumulative profits in the past month. All specifications include individual-specific dummies. $t-$statistics are reported below the OLS estimates. $^{*}/^{**}/^{***}$ indicate significance at the 10%/5%/1% levels.

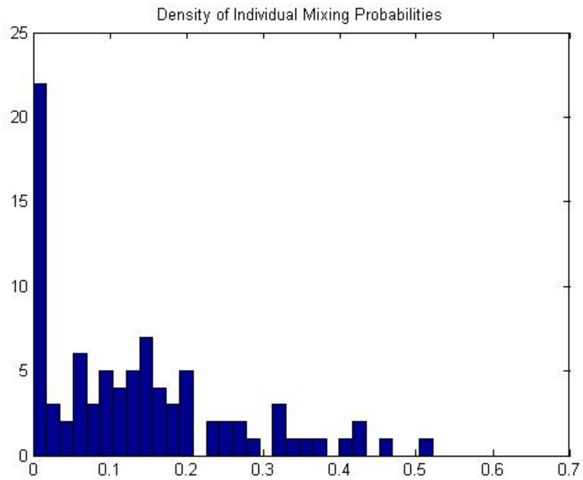

Figure 12: Histogram of individuals' posterior probability of belonging to HARA population.

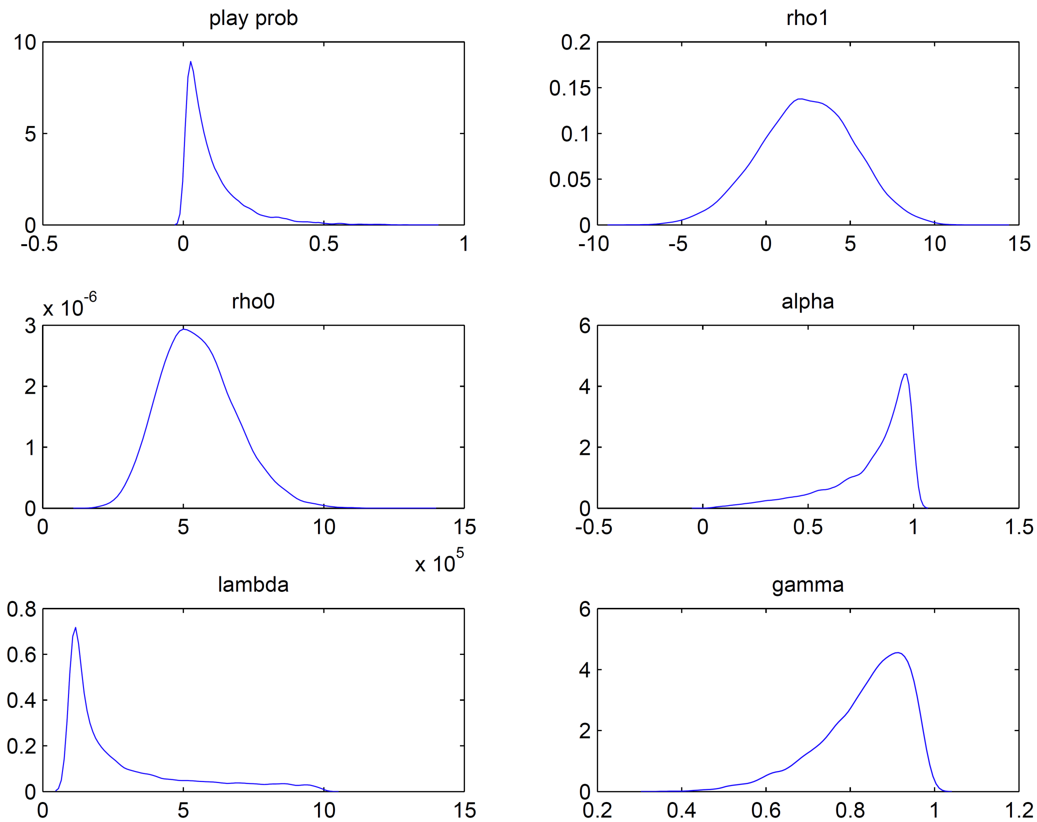

Figure 13: Posterior simulator densities of utility parameters of the HARA and CPT specifications.

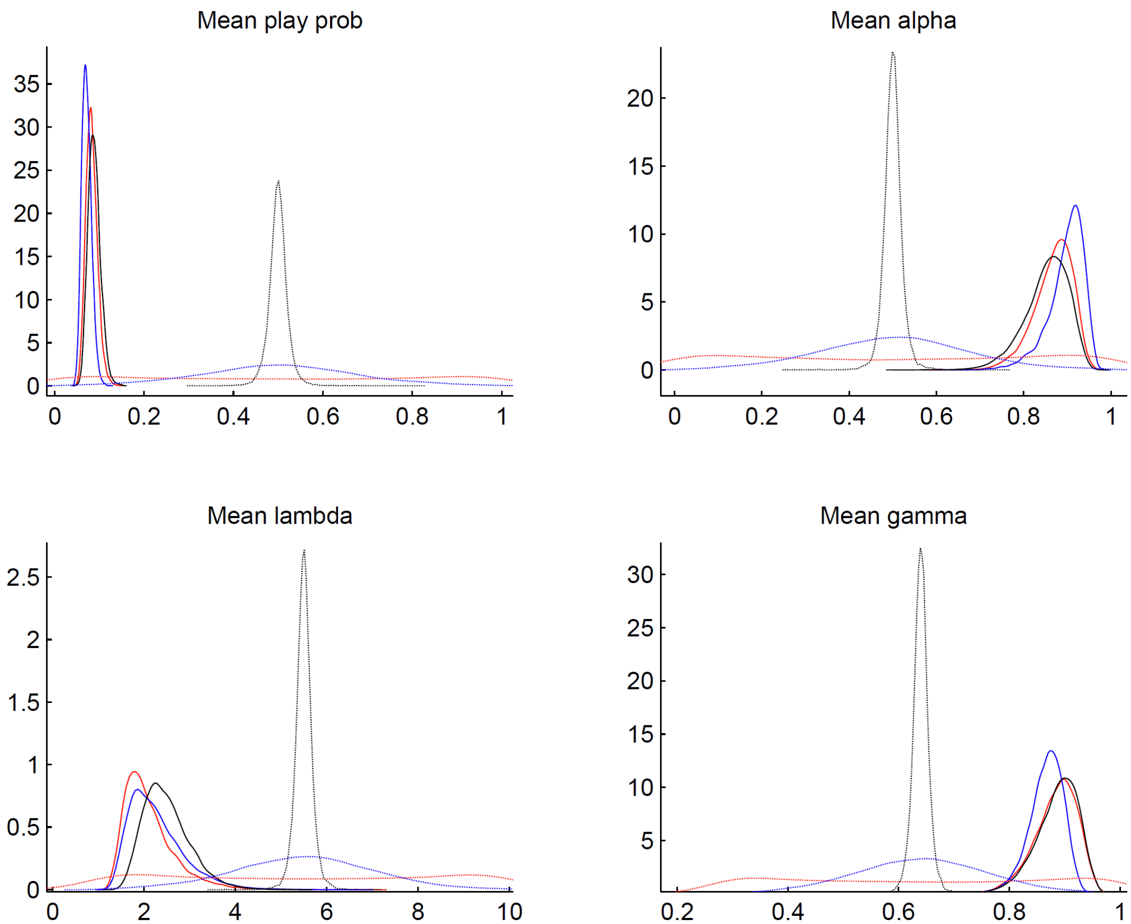

Figure 14: Prior (dotted) and corresponding posterior densities for $\tilde{\beta}_n$, for three priors.

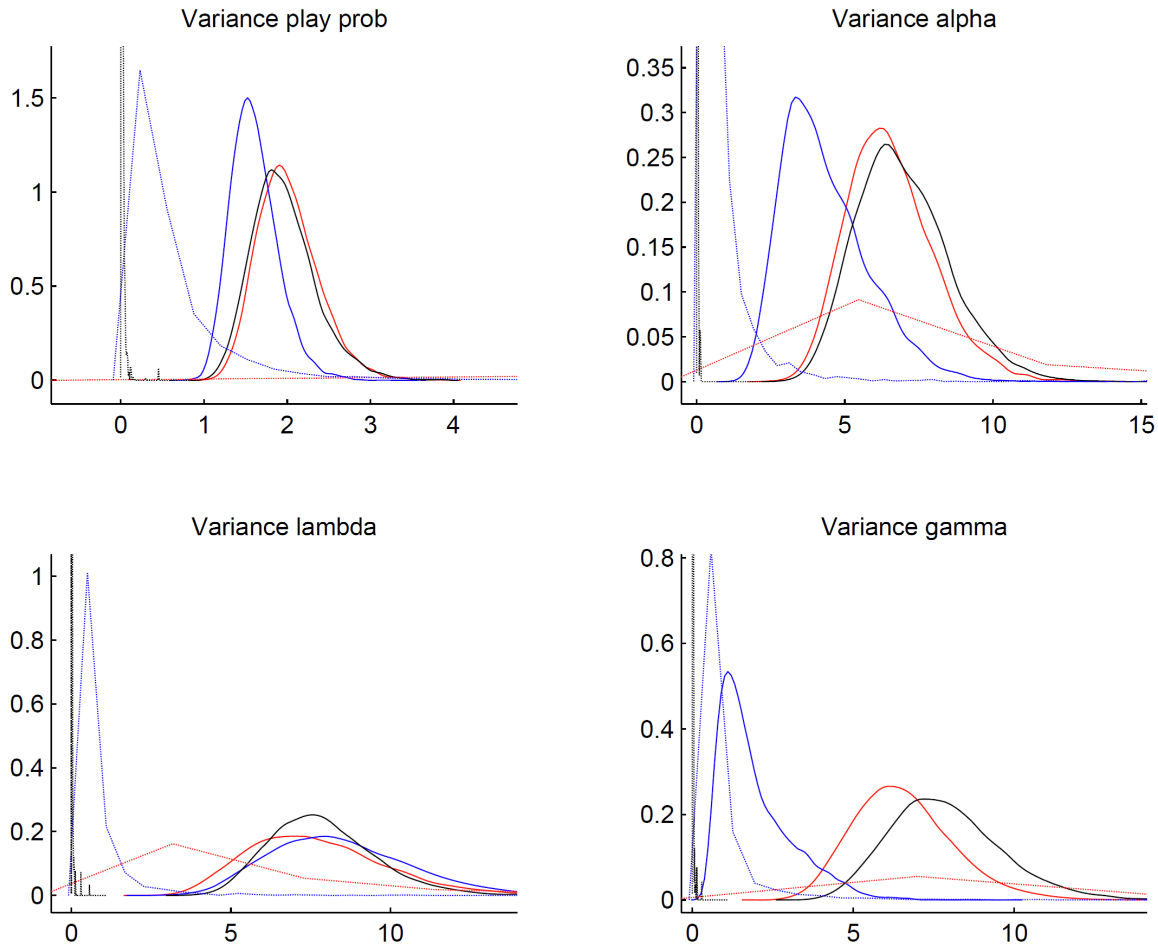

Figure 15: Prior (dotted) and corresponding posterior densities for the population variance, for three priors.

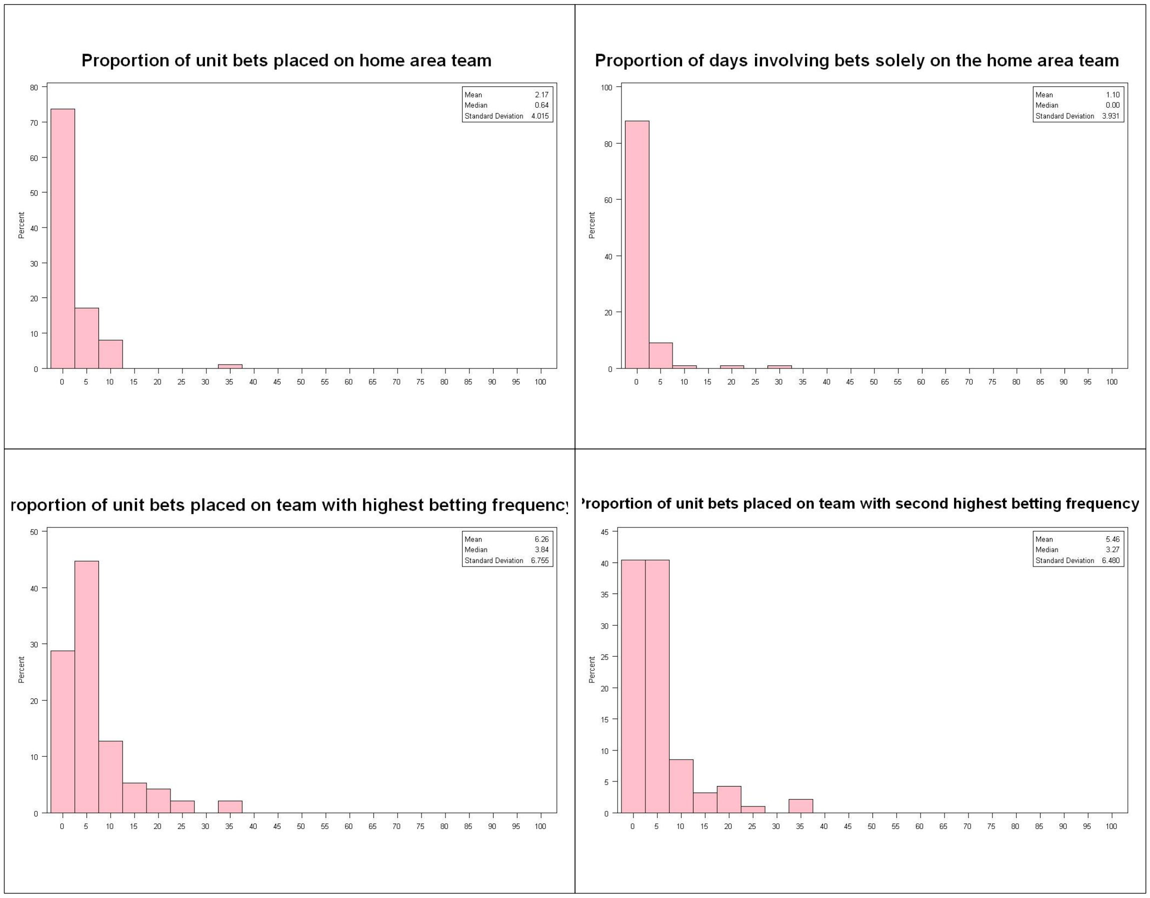

Figure 16: Histograms showing the proportions of unit bets or days (for top-right panel) placed on the home team (for the top two panels) and the most-frequently and second-most-frequently bet teams (bottom two panels).