Heterogeneous Preferences for Risk and Betting Market Equilibrium

Amit Gandhi
University of Chicago
Graduate School of Business

Abstract

We present a new model of price equilibrium in betting markets, most notably the odds market at horserace tracks. A well known empirical regularity arising in race track markets is the "favorite long-shot bias", which finds that the odds paid on long-shots are too low relative their actual chances of winning, and the odds paid on favorites are too high. By assuming a representative bettor, previous studies have found that departures from expected utility theory, namely those characterized by cumulative prospect theory, better explain this bias. While these studies mark important first steps in using aggregate market data to discriminate between models of individual decision making, we argue that they do not constitute a sufficient test of expected utility theory. In particular, by assuming a representative bettor, these past studies fail to model one of the basic functions of financial markets, which is to coordinate trade between individuals with different tastes for risk. In this paper, we explicitly account for heterogeneity of risk preferences across bettors, and develop a new model of price equilibrium in betting markets. Under mild continuity conditions on the distribution of preferences, we prove that there exists unique equilibrium odds that clear the betting market. Heterogeneity of risk preferences in turn provides a natural expected utility theoretic explanation of the favorite long-shot bias : risk lovers "overbet" long-shots in order to finance the incentive of risk averters to bet on favorites. The fundamental question is whether heterogeneous expected utility maximizers are capable of making up for the empirical shortcomings of a representative expected utility maximizer in explaining the favorite long-shot bias, and betting market prices more generally. In order to set the stage for the empirical analysis, we uncover the key econometric properties of our betting markets model. First, we find that from any observed odds in a market, we can uniquely solve for the underlying beliefs shared by the bettors about the chances of each horse winning. Along with the actual observed winner in each race, this inverse pricing function provides the key to estimating the model. Finally, we show that by reducing heterogeneity to a single dimensional type, the model of preference heterogeneity is non-parametrically identified under a single crossing condition that is widely satisfied in both expected and non-expected utility theories.

1 Introduction

A fundamental assumption about human behavior used in modern economic modelling is the expected utility hypothesis (EUH). In its most basic form, the EUH is a hypothesis about the nature of individual preferences for risky propsects, i.e. lotteries over a set of possible monetary outcomes. The EUH maintains that probability enters linearly into an economic agent's preference for risk, leading the agent to acts so as to maximize an expected utility of wealth when faced with a choice among lotteries. Today, the EUH is the defacto standard for modelling individual choice under risk in economic settings ranging from games to asset markets.

However, expected utility theory (EUT) has also undergone an immense amount of scrutinty and criticism as a description of how individuals actually behave when faced with a risky decision. For example, the assumption that probability enters linearly into the calculus of comparing lottieries has been shown to be routinely violated by subjects in laboratory settings. An especially vivid demonstration of this phenomena is the well known Allais paradox.[1]

These negative experimental findings have motivated a number of attempts to relax expected utility theory in ways that better fit the experimental evidence (for a review of nonexpected utility theories, see Starmer [2000]). The most prominent of these attempts is cumulative prospect theory (CPT) [Tversky and Kahneman, 1992]. In addition to allowing probability to nonlinearly enter preferences, CPT relaxes EUT a step further by allowing an economic agent to treat losses and gains relative to a reference point asymmetrically. The empirical superiority of CPT over EUT in explaining choice patterns uncovered in the laboratory have led to loud cries from the experimental community to abandon expected utility theory in applied economic modelling (e.g., Rabin and Thaler [2001]).

Despite these developments, the neoclassical defense that the "true" test of an economic theory lies in its capacity to explain the actual choices of economic actors in naturally occuring institutions, not the hypothetical choices of experimental subjects in the laboratory, has provided a safehaven for the EUH [Cubitt et al., 2001]. Thus a fundamental empirical question is whether market data rather than experimental data can be brought to bear on the debate between expected and non-expected utility theory.

Betting markets, and the racetrack for horses in the particular,[2] offer an especially attractive "real world" setting in which to study individual decision making under risk. At the race track, bettors must choose which horse in a race to bet money. If the chosen horse wins the race, the bettor recieves his original bet plus a

---

[1] For a review of the experimental evidence against EUT, see Starmer [2000].

[2] For the sake of aiding understanding, we shall state our results in the language of "racetracks", although the analysis applies to so called "parimutuel" or non-bookmaker betting markets more generally.

profit equal to the original bet times the market determined odds on the horse. If the chosen horse does not win, the bettor loses his original bet. Thus the racetrack presents a menu of simple lotteries from which bettors must choose, providing a natural laboratory to test theories of individual decision making under risk.[3]

However the central problem in using betting markets to estimate and test models of individual choice behavior is that individual bets are not observed. Instead, only aggregate bets in the form of prices (odds) on each horse can be observed. In order to fill this gap, an economic model of equilibrium pricing in the betting market is needed, with empirically distinct consequences for the pricing of bets arising from distinct assumptions about the nature of preferences. The goal of the present paper is to develop just such a model. The key innovation in our model of equilibrium is that it is the first to explicitly account for heterogeneity across bettors in their preferences for risk. Heterogeneity of risk preference, we argue, offers a new and natural neoclassical explanation for the main pricing puzzle that arises in racetrack markets, the "favorite longshot bias" [Thaler and Ziemba, 1988].

The "favorite longshot bias" is a reference to the empirical finding that the expected rate of return from betting on horses with smaller odds (i.e., favorites) is greater than the expected rate of return from betting on horses with larger odds (i.e., longshots). That is, favorites appear to be priced to low, or underbet relative to the frequency with which they actually win. Likewise, longshots appear to be priced to high, or overbet relative to the frequency with which they actually win.

In order to explain the favorite longshot bias and model price equilibrium in betting markets more generally, the literature to date has largely relied upon a representative bettor assumption. A representative bettor assumes that all bettors in the market have the same beliefs as to the chance of each horse winning, and in addition, assumes that all bettors also have the same preferences for risk. In equilibrium, the odds on each horse are such that the representative bettor is indifferent between betting on each horse in the race.

Clearly, the favorite longshot bias is inconsistent with a risk neutral representative bettor, since such a bettor would strictly prefer betting on the horse with the highest expected return (which under the favorite longshot bias is the horse most favored to win). Thus in order to neoclassicaly explain the favorite longshot bias with a representative bettor, this bettor must be a risk loving expected utility maximizer, who is willing to accept the lower expected return from betting on longshots because of the higher volatility these bets offer [Weitzman, 1965].[4]

---

[3] See Jullien and Salanie [2002] for a review of the advantages of betting markets for investigating empirical issues in financial markets more generally.

[4] In a similar result, [Quandt, 1986] shows that the favorite longshot bias is a neccesary consequence of risk loving, mean-variance expected utility maximizing behavior among bettors (not neccessarily a represenative bettor).

In the seminal paper to compare expected and nonexpected utility theories against betting market prices, Jullien and Salanie [2000] find that a CPT representative agent empirically outperforms a risk loving representative agent. Using a different empirical methodology, but maintaining the representative bettor model, Snowberg and Wolfers [2005] arrive at a similar conclusion as to the empirical superiorty of CPT relative to EUT in explaining the pattern of prices at the racetrack. Thus the few studies to date that compare expected and non-expected utility theories using market data point in the direction of rejecting expected utility theory, a fact that has been duly emphasized by behavioral economists Camerer [2000].

In this paper, we present a new model of price equilibrium in betting markets that dispenses with the representative bettor assumption. Allowing for *heterogeneity* of risk preferences in the betting population, we argue, offers a much more natural neoclassical explanation for the favorite longshot bias. Instead of risk love, the favorite longshot bias arises in our analysis from trade between expected utility maximizing bettors with *different* tolerances for risk. Essentially, in equilibrium, risk lovers overbet longshots in order to finance the incentive, in terms of a higher expected return, for risk averters to bet on favorites. Moreover, if heterogeneity of risk preference in the betting population is what actually drives the favorite longshot bias, then a representative expected utility maximizer produces a biased estimate (in the direction of risk loving) of the average behavior in the population. This bias in turn offers an explanation for the empirical shortcomings of the representative expected utility maximizer used in Jullien and Salanie [2000] and Snowberg and Wolfers [2005].

A simple example will make the point more clearly. Suppose there are two horses, with one horse having a .75 chance of winning the race (the favorite), and the other horse having a .25 chance of winning (the longshot). Assume that there are three expected utility maximizing bettors having constant relative risk aversion (CRRA) utility functions of wealth $u(w) = w^{\alpha}$, two of whom are risk averse with $\alpha = .25$, and one of whom is risk loving with $\alpha = 2.0$. If each bettor has a dollar in his pocket, then a natural trade between bettors is for the risk lover to give his dollar to be split between the risk averters if the favorite wins, and in exchange get both risk averter dollars if the longshot wins. Implicitly in this exchange, the risk averters are each betting their dollars on the favorite, which pays them each 1 dollar 50 cents with 75 percent chance and 0 dollars otherwise (i.e. the favorite pays odds of 1 to 2). Likewise, the risk lover is implicitly betting his dollar on the longshot, which pays him 3 dollars with 10 percent chance and 0 dollars otherwise (i.e. the longshot pays odds of 2 to 1).

The risk lover strictly prefers this small chance for a large gain over the large chance for a small gain recieved by the risk averters (i.e. $(.25)3.0^{2.0} > (.75)1.5^{2.0}$), and likewise, each risk averter strictly prefers his large chance for a small gain over the small chance for a large gain received by the risk lover (i.e. $(.75)1.5^{.25} > (.25)3^{.25}$). Thus the prices implicit in this exchange represents an equilibrium, i.e. given that the favorite pays odds of 1 to 2, and the longshot pays 2 to 1, the risk averters strictly prefer to bet their dollars on the favorite and the risk lover strictly prefers to bet his dollar on the longshot, and these choices in turn reproduce the betting odds. But clearly, this equilibrium generates the favorite longshot bias, with only 2/3 of the total bet pool going towards the 75 percent favorite, which implies that a bet on the favorite earns a higher expected rate of return than a bet on the 25 percent longshot (i.e., $(.75).5 + (.25)(-1) > (.25)(2.0) + (.75)(-1)$).[5]

Thus we see that when bettors have heterogeneous tastes for risk, the favorite longshot bias can naturally arise under the EUH from risk lovers overbetting longshots in order to give risk averters the incentive (in terms of higher expected returns) to bet on favorites. Of course, the magnitude of the bias, and even its shape,[6] will depend upon the particular distribution of preferences for gambles in the betting population. The first main result of the paper is to show that under very general conditions on the distribution of preferences, which do not require us to assume any particular utility theoretic representation (such as EUT or CPT), there exists unique equilibrium odds (prices) on each horse that clears the market. More precisely, assuming bettors have a common prior over the chances of each horse winning, we show that any distribution of preferences satisfying two mild continuity conditions implies the existence of unique equilibrium odds on each bet in a "parimutuel" gambling system, such as the betting pool from the horse race example above.

In light of the natural role played by heterogeneity in the betting market, the empirical shortcoming of an expected utility maximizing representative agent becomes understandable. It is often argued that a representative agent is put forth to capture the "average" behavior of a population of economic agents. However it is clear from our example, if there truly exists heterogeneity of preferences in the betting population, then a representative bettor produces a biased estimate of the average behavior. For instance, in our example, the average CRRA coefficient in the population is less than 1, corresponding to risk averse behavior. However if we fit a representative bettor to the equilibrium prices, then we will estimate a risk loving representative bettor, which is thus a biased "explanation" of behavior in the market.

---

[5] Notice that in this example, we do not explain why bettors gamble, Rather, assuming that bettors are gambling, risk preferences in our example explain which horse a bettor prefers to wager. In the general model presented later, the individual rationality of gambling is addressed. It is also worthwhile to note that in this example, we have treated the bettors as "price takers", although each bettor's choice of horse clearly affects the betting odds since there are only three bettors in the population. In the analysis to follow, the betting population will be a continuum, and thus the price taking assumption will be well founded.

[6] It is possible, for instance, that expected returns will be the highest for horses with medium ranged odds. More precisely, in equilibrium, extreme risk averters could opt to have lower expected returns in order to bet on extreme favorites, and extreme risk lovers could opt to have lower expected returns in order to bet on extreme longshots, leaving more risk neutral bettors to gain higher expected returns by betting on horses with moderate odds.

Thus the key empirical question is to ask what happens if we adjust the analysis of Jullien and Salanie [2000] to account for heterogeneity. Does a representative CPT maximizer also empirically outperform a population of heterogeneous expected utility maximizer, or does the failure to account heterogeneity explain the empirical shortcomings of the representative expected utility maximizer in Jullien and Salanie [2000]? Towards this end, we present our second main result : For any distribution of preferences satisfying our basic continuity conditions, and for any observed equilibrium odds in a race, we can solve uniquely for the underlying common prior beliefs held by bettors about the chances of each horse winning. Along with observing the winning horse in each race, this inverse equilibrium pricing function provides the basis for estimating the model and comparing the relative performances of EUT and CPT functional forms.

Finally, we present results on the nonparametric identification of our model of heterogeneity in the betting market. If we reduce heterogeneity across bettors to a single dimensional type, then under a single crossing condition widely satisfied by both expected and non-expected utility theories, the distribution of preferences can be nonparametrically identified up to an increasing transformation.

2 Parimutuel Betting and Horse Races

We will assume that in a race with n horses running, each bettor in the market has the same prior beliefs as to the chances of each horse winning, and bets the same amount of money (say M dollars). Each horse i has a price R_i, namely its betting odds, which equals the net rate of return from betting on the horse in the event that the horse wins. The net rate of return in the event the horse loses is of course -1 because the bettor simply loses his bet. Thus a bet on horse i can be viewed as an Arrow-Debreu security whose price is $\frac{1}{1+R_i}$, and pays 1 dollar in the event horse i wins and zero dollars otherwise. This Arrow-Debreu price is well defined so long as $R_i > -1$.[7]

From the point of view of a bettor, a horse in a race can be represented by a simple gamble (R, p) where R is the odds paid if the horse wins and p is the bettor's belief of the probability that the horse wins the race. Given the odds R_i for $i = 1, \ldots, n$ in a race with n horses, and a bettor with beliefs as to the chance of each horse winning expressed by the probability distribution $(p_1, \ldots, p_n)$, the market from the point of view of this bettor offers a menu of n gambles $G = \{(R_1, p_1), \ldots, (R_n, p_n)\}$. We shall initially assume that each bettor makes a discrete choice from among this menu G as to which horse to place his M dollars. More generally, as alluded to in footnote 5, we will let the menu G include the option of not betting, which is

---

[7] A negative odds has the same interpretation as a negative interest rate, i.e., the price of obtaining a dollar in the event that horse i wins is greater than a dollar (but finite).

equivalent to a gamble that offers a net rate of returm zero with probability one.

Of course, the aggregate choices made by bettors as to which horse to bet (and whether to bet at all) will impact the actual odds $(R_1^*, \ldots, R_n^*)$ that are offered. In parimutuel betting systems, which are the most common method organizing betting markets, the final odds are determined directly by the amounts bet on each horse, in "parimutuel" fashion. That is, the pool of all dollars bet is split evenly among those bettors who bet on the winning horse.[8] An equilibrium set of prices $(R_1^*, \ldots, R_n^*)$ is one in which where the aggregate choices made by bettors on the basis of these prices combined with the parimutuel mechanism reproduces these prices $(R_1^*, \ldots, R_n^*)$.

Our main results are roughly as follows. We will show that under mild regularity assumptions on the distribution of bettor preferences for simple gambles (R, p), the realized odds $(R_1^*, \ldots, R_n^*)$ are uniquely determined in market equilibrium. Moreover, from the realized odds $(R_1^*, \ldots, R_n^*)$, which are observable, we can uniquely solve for the underlying common beliefs $(p_1^*, \ldots, p_n^*)$ that generated the observed odds in equilibrium. This inverse equilibrium pricing function provides the key to estimating the model. Along the way, we illustrate how the "favorite-longshot" bias results naturally under the EUH from trade between risk lovers and averters, even in a symmetric population of expected utility maximizers centered around risk neutrality. Thus a representative bettor, which must be risk loving in order to explain the faviorite longshot bias, would lead to a biased estimate of the average behavior in the population. Finally, we present results on the nonparametric identification of our model when heterogeneity is reduced to a single dimensional type.

3 The Model

We now develop a demand model for horse gambles based on discrete choice behavior amongst bettors with heterogeneous preferences for gambles. Consider a continnum of consumers T. Each consumer $t \in T$ has a complete, continuous, transitive, and strictly monotonic preference relation $\succsim_t$ over gambles $(R, p) \in \mathbf{R} \times [0, 1]$, where $\mathbf{R} = (-1, \infty) \subset \mathbb{R}$ is the permissible range for odds on a horse bet, as discussed in the previous section. Thus each consumer t's preference relation can be represented by a continuous utility function $V_t : \mathbf{R} \times [0, 1] \to \mathbb{R}$ that is strictly increasing in a gamble's net rate of return from winning R (the first argument of V_t) and probability of winning p (the second argument of V_t). We further impose the restriction that each consumer t's utility function is strictly minimized whenever $p = 0$, i.e., $V_t(0, R) = V_t(0, R')$ and $V_t(0, R) < V(p, R')$ for any returns R, R' and any probability $p > 0$. Thus the worst gamble for any consumer

---

[8] Recall we are assuming each bettor bets the same amoint M. Otherwise, the parimutel rules would distribute the betting pool to the winning bettors in proportion to the amount bet.

is one that has no probability of winning, regardless of the return from winning (since this return is never realized).

Let $\mathbf{V} \subset \mathbb{R}^{\mathbf{R}\times[0,1]}$ be the set of all such utility functions. We endow $\mathbf{V}$ with the relative product topology, otherwise known as the topology of pointwise convergence. Let the measurable sets in $\mathbf{V}$ be the Borel subsets (the σ-algebra of subsets generated by the open sets) in this topology. Our population T gives rise to a probability measure $\mathbf{P}_V$ over the space $\mathbf{V}$. The probability measure $\mathbf{P}_V$ describes the distribution of consumer preferences for gambles (R, p).

Now consider any finite set of n gambles

$$G = \{(R_1, p_1), \ldots, (R_n, p_n)\} \subset \mathbf{R} \times [0, 1].$$

The subset of the population T that prefers the i^{th} gamble from the set G is denoted

$$S_i = \{V \in \mathbf{V} : V(R_i, p_i) \geq V(R_j, p_j) \text{ for all } j \neq i\}.$$

The share of the population T that prefers the i^{th} gamble from the set G is thus

$$q_i(G) = q_i(R_1, \ldots, R_n; p_1, \ldots, p_n) = \mathbf{P}_V(S_i).^{9} \tag{1}$$

We refer to q_i as the *market share* of the i^{th} gamble from G (i.e., if the market offered a choice of gambles from the set G, then q_i is the share of the population T that chooses the i^{th} gamble). Notice that for any n-tuple of distinct probabilities $(p_1, \ldots, p_n)$, q_i is a well defined function over $(R_1, \ldots, R_n) \in \mathbf{R}^n$. Likewise, for any n-tuple of returns $(R_1, \ldots, R_n)$, q_i is a well defined function over $(p_1, \ldots, p_n) \in [0, 1]^n$. In our use of the market share functions, we shall only consider variation in the n-tuple of probabilities $(p_1, \ldots, p_n)$ over the n dimensional unit simplex Δ^{n-1}.

We make two mild regularity assumptions on the distribution of preferences $\mathbf{P}_V$, which we refer to as *continuity* and *desireability*. Continuity requires that the probability measure $\mathbf{P}_V$ be sufficiently continuous, or atomless, so as to not permit a positive mass of consumers to be indifferent between two distinct gambles. Desireability requires that for any gamble $g = (p_g, R_g)$ with nonzero probability of winning $p_g > 0$, and for any finite set of gambles G with $g \in G$, it is always possible to induce some positive mass of the population to prefer g from G by making g's return from winning R_g sufficiently large. We now more formally define

---

[9] S_i is measurable based on the topology on $\mathbf{V}$.

continuity and desireability, and examine their consequences for the market share functions q_i.

4 Continuity

Continuity The probability measure $\mathbf{P}_V$ is continuous if for any two distinct gambles (R_i, p_i) and (R_j, p_j) with p_i or p_j greater than 0 (or both), the number of consumers indifferent between gamble i and j has a probability measure of zero. More precisely, if $p_i > 0$ or $p_j > 0$ then

$$\mathbf{P}_V(\{V \in \mathbf{V} : V(R_i, p_i) = V(R_j, p_j)\}) = 0.$$

Lemma 4.1 *If $\mathbf{P}_V$ is continuous, then for any finite set of gambles G with at least one gamble having nonzero probability of winning,*

$$\mathbf{P}_V(S_i \cap S_j) = 0 \textit{ for every } i \neq j.$$

Proof If $p_i = 0$, then $S_i = \emptyset$ because each $V \in \mathbf{V}$ is strictly minimized at $p = 0$. Otherwise, $p_i > 0$, and for $i \neq j$, $S_i \cap S_j$ is a subset of

$$\{V \in \mathbf{V} : V(R_i, p_i) = V(R_j, p_j)\},$$

which by continuity has measure 0 under $\mathbf{P}_V$. ▮

Theorem 4.2 *If $\mathbf{P}_V$ is continuous, then for any finite set of distinct gambles G with at least one gamble in G having nonzero probability of winning, the sum of the market shares equals 1, i.e.,*

$$\sum_{i=1}^{n} q_i(R_1, \ldots, R_n; p_1, \ldots, p_n) = 1.$$

Proof Recall that

$$\sum_{i=1}^{n} q_i(R_1, \ldots, R_n; p_1, \ldots, p_n) = \sum_{i=1}^{n} \mathbf{P}_V(S_i).$$

Moreover, since G is finite, each $V \in \mathbf{V}$ attains a maximum over G, and thus

$$\bigcup_{i=1}^{n} S_i = \mathbf{V}.$$

However by Lemma 4.1, we have that for all $i \neq j$,

$$\mathbf{P}_V(S_i \cap S_j) = 0.$$

Thus

$$\sum_{i=1}^{n} \mathbf{P}_V(S_i) = \mathbf{P}_V\left(\bigcup_{i=1}^{n} S_i\right) = 1. \quad \blacksquare$$

We now establish the following central result.

Theorem 4.3 *If $\mathbf{P}_V$ is continuous, then for any n-tuple of distinct probabilities $(p_1, \ldots, p_n)$,*

$$q_i(R_1, \ldots, R_n; p_1, \ldots, p_n)$$

is a continuous function in $(R_1, \ldots, R_n) \in \mathbf{R}^n$. Furthermore, for any n-tuple of distinct returns $(R_1, \ldots, R_n)$,

$$q_i(R_1, \ldots, R_n; p_1, \ldots, p_n)$$

is a continuous function in $(p_1, \ldots, p_n) \in \Delta^{n-1}$.

Proof We prove the first part of the theorem (continuity in $(R_1, \ldots, R_n)$). The second part (continuity in $(p_1, \ldots, p_n)$) follows similarly to the first.

Let us fix any n-tuple of distinct probabilities $(p_1, \ldots, p_n)$. Now consider any n-tuple of returns $(R_1, \ldots, R_n)$. Define a function $F : \mathbf{V} \to \{0, 1\}$ as

$$F(V) = \prod_{j \neq i} \mathbf{1}\left[V(R_i, p_i) \geq V(R_j, p_j)\right],$$

where $\mathbf{1}(\cdot)$ is the indicator function. Then clearly

$$q_i(R_1, \ldots, R_n; p_1, \ldots, p_n) = \int F(V) \mathbf{P}_V(dV).$$

Now consider any sequence of n-tuples of returns $\{(R_1^t, \ldots, R_n^t)\}_{t \in \mathbb{N}_+}$ that converges to $(R_1, \ldots, R_n)$. For each $t \in \mathbb{N}_+$, define $F^t : \mathbf{V} \to \{0, 1\}$ as

$$F^t(V) = \prod_{j \neq i} \mathbf{1}\left[V(R_i^t, p_i) \geq V(R_j^t, p_j)\right],$$

and thus,

$$q_i(R_1^t, \ldots, R_n^t; p_1, \ldots, p_n) = \int F^t(V) \mathbf{P}_V(dV).$$

We need to establish that $q_i(R_1^t, \ldots, R_n^t; p_1, \ldots, p_n) \xrightarrow{t} q_i(R_1, \ldots, R_n; p_1, \ldots, p_n)$.

For every $V \in \mathbf{V} - S_i$ we have $F(V) = 0$. Thus for every $V \in \mathbf{V} - S_i$,

$$V(R_i, p_i) < V(R_j, p_j) \text{ for some } j \neq i.$$

By continuity of every utility function $V \in \mathbf{V}$, we have that for every $V \in \mathbf{V} - S_i$,

$$F^t(V) \xrightarrow{t} F(V).$$

On the other hand, for every $V \in S_i$, $F(V) = 1$. Since at least one $p_i > 0$, then by Lemma 4.1, for almost every $V \in S_i$,[10]

$$V(R_i, p_i) > V(R_j, p_j) \text{ for every } j \neq i.^{11}$$

Once again, by continuity of every utility function $V \in \mathbf{V}$, we have that for almost every $V \in S_i$,

$$F^t(V) \xrightarrow{t} F(V).$$

Thus for almost every $V \in \mathbf{V}$, $F^t(V)$ converges pointwise to $F(V)$. By Lesbesgue's dominated convergence theorem,

$$q_i(R_1^t, \ldots, R_n^t; p_1, \ldots, p_n) \xrightarrow{t} q_i(R_1, \ldots, R_n; p_1, \ldots, p_n).$$

The proof of continuity in $(p_1, \ldots, p_n)$ follows similarly. The requirement in the theorem that $(p_1, \ldots, p_n)$ range over Δ^{n-1} is overly restrictive. it is only used to ensure that $(p_1, \ldots, p_n) \in \Delta^{n-1}$ implies at least one $p_i > 0$, which allows all the steps from continuity in returns to be repeated. ∎

The last implication of the continuity assumption that we consider concerns the monitinicity of the market share functions q_i. We first define the relevant meaning of monitinicity, and then show that it is satisfied under continity of the distribution of preferences.

Monotinicity For any n-tuple of distinct probabilities $(p_1, \ldots, p_n)$, any n-tuple of returns $(R_1, \ldots, R_n)$, and any strict subset $\mathbf{I} \subset \{1, \ldots, n\}$, consider a change to the returns from winning appearing in the choice set G that weakly increases the returns of the gambles indexed by $\mathbf{I}$ and weakly decreases the returns of the

---

[10] The statement "for almost every $V \in S_i$" means "for all $V \in S_i$ except possibly in a subset $S \subset S_i$ with $\mathbf{P}_V(\mathrm{S}) = 0$".

[11] $S_i = \{V \in \mathbf{V} : V(R_i, p_i) > V(R_j, p_j) \text{ for every } j \neq i\} \cup_{j \neq i} (S_i \cap S_j)$.

remaining gambles. This change leads to a new choice set $G^* = \{(R_i^*, p_i)\}_{i \in \{1,\ldots,n\}}$ with

$$R_i^* \geq R_i \text{ for } i \in \mathbf{I}$$

and

$$R_i^* \leq R_i \text{ for } i \notin \mathbf{I}.$$

We say that the distribution of consumer preferences $\mathbf{P}_V$ satisfies *monotinicity in return* if the sum of the shares of the gambles indexed by $\mathbf{I}$ weakly increase as a result of the change in returns, and the sum of the shares of the gambles indexed by $\{1,\ldots,n\} - \mathbf{I}$ weakly decrease as a result of the change in returns. That is

$$\sum_{i \in \mathbf{I}} q_i^* \geq \sum_{i \in \mathbf{I}} q_i$$

and

$$\sum_{i \notin \mathbf{I}} q_i^* \leq \sum_{i \notin \mathbf{I}} q_i.$$

Remark The definition of *monotinicity in probability* is stated similarly, except for any n-tuple of distinct returns $(R_1,\ldots,R_n)$, and any n-tuple of probabilities $(p_1,\ldots,p_n) \in \Delta^{n-1}$[12], we consider a weak increase of the probabilities of winning of the gambles for a strict subset of the gambles, and a weak decreases of the probabilities of winning for the remaining gambles, producing a new n-tuple of probabilities $(p_1^*,\ldots,p_n^*) \in \Delta^{n-1}$. The distribution $\mathbf{P}_V$ satisfies *monotinicity in probability* if such a change results in an increase in the of sum the shares of the gambles for which the probabilities increased, and a decrease in the sum of the shares of the gambles for which the probabilities decreased.

Theorem 4.4 *If $\mathbf{P}_V$ is continuous, then it satisfies monitinicity in return and monitinicity in probability.*

Proof We prove the theorem for monitincity in returns. A similar argument follows for monitinicity in probability. Let $(p_1,\ldots,p_n)$, $(R_1,\ldots,R_n)$, and $(R_1^*,\ldots,R_n^*)$ be the n-tuples described in the definition of monitinicity. Similarly to the proof of Theorem 4.3, define

$$F_i(V) = \prod_{j \neq i} \mathbf{1}\left[V(R_i,p_i) \geq V(R_j,p_j)\right] \quad \text{and} \quad F_i^*(V) = \prod_{j \neq i} \mathbf{1}\left[V(R_i^*,p_i) \geq V(R_j^*,p_j)\right]$$

---

[12] We restrict the domain of n-tuple of probabilities to the simplex so as to ensure at least one probability is nonzero.

By the linearity of the integral operation

$$\sum_{i \in \mathbf{I}} q_i(R_1, \ldots, R_n; p_1, \ldots, p_n) = \int \sum_{i \in \mathbf{I}} F_i(V)\mathbf{P}_V(dV) \quad \mathit{and} \quad \sum_{i \in \mathbf{I}} q_i(R_1^*, \ldots, R_n^*; p_1, \ldots, p_n) = \int \sum_{i \in \mathbf{I}} F_i^*(V)\mathbf{P}_V(dV)$$

Since at least one $p_i > 0$, then by Lemma 4.1, for almost every $V \in \mathbf{V}$, $\sum_{i \in \mathbf{I}} F_i(V)$ equals 0 or 1.[13] However by the monitinicicty of each $V \in \mathbf{V}$, $\sum_{i \in \mathbf{I}} F_i(V) = 1$ implies $\sum_{i \in \mathbf{I}} F_i^*(V) \geq 1$. Thus for almost every $V \in \mathbf{V}$,

$$\sum_{i \in \mathbf{I}} F_i^*(V) \geq \sum_{i \in \mathbf{I}} F_i(V)$$

and since the integral is an increasing linear operation,

$$\sum_{i \in \mathbf{I}} q_i(R_1^*, \ldots, R_n^*; p_1, \ldots, p_n) \geq \sum_{i \in \mathbf{I}} q_i(R_1, \ldots, R_n; p_1, \ldots, p_n). \qquad \blacksquare$$

Thus to summarize,

$$q_i(R_1, \ldots, R_n, p_1, \ldots, p_n)$$

is the share of population T that chooses the i^th gamble from the choice set of distinct gambles $G = \{(R_1, p_1), \ldots, (R_n, p_n)\}$. The continuity assumption on the probability distribution $\mathbf{P}_V$ of consumer preferences carried three important consequences :

1. For any such choice set G, so long as there is some gamble with $p_i > 0$, the market shares sum to 1, i.e.,
$$\sum_{i=1}^{n} q_i(R_1, \ldots, R_n; p_1, \ldots, p_n) = 1.$$

2. For an n-tuple of distinct probabilities $(p_1, \ldots, p_n)$,
$$q_i(R_1, \ldots, R_n; p_1, \ldots, p_n)$$
is continuous in $(R_1, \ldots, R_n)$ over $\mathbf{R}^n$. Likewise, for any n-tuple of distinct returns, $(R_1, \ldots, R_n)$,
$$q_i(R_1, \ldots, R_n; p_1, \ldots, p_n)$$
is continuous in $(p_1, \ldots, p_n)$ over Δ^{n-1}.

---

[13] The set of V for which $\sum_{i \in \mathbf{I}} F_i(V) > 0$ equals $\cup_{i \neq j; i,j \in \mathbf{I}}(S_i \cap S_j)$

3. For any distinct n-tuple of distinct probabilities $(p_1, \ldots, p_n)$, and any n-tuple of returns $(R_1, \ldots, R_n)$, and any strict subset $\mathbf{I} \subset \{1, \ldots, n\}$,

$$R_i^* \geq R_i \text{ for } i \in \mathbf{I} \quad \text{and} \quad R_i^* \leq R_i \text{ for } i \notin \mathbf{I}$$

implies $\sum_{i \in \mathbf{I}} q_i^* \geq \sum_{i \in \mathbf{I}} q_i$ (i.e., monitinicicty in return). In addition, monitinicity in probability also holds true.

5 Desireability

The final assumption that we wish to place on the distribution of preferences is a formalization of the idea that when the return from winning offered by a gamble is made large enough, then it is always possible to induce some positive mass of consumers to prefer this gamble over all other gambles being offered by the market. This is an essential requirement for equilibrium. Otherwise, it would be possible that the return offered by a bet on a horse with low probability of winning could not be made high enough in order to compensate any positive mass of consumers to bet on the horse. Thus any finite return on such a gamble would would be inconsistent with an equilibrium, for as we shall see, equilibrium in the betting market requires some positive mass of bettors to gamble on each horse in a race.

Desireability A continuous distribution of preferences $\mathbf{P}_V$ satisfies *desireability* if for any n-tuple of distinct probabilities $(p_1, \ldots, p_n)$, and any n-tuple of returns $(R_1, \ldots, R_n)$, $p_i > 0$ implies that for any nondecreasing sequence of returns $\{R_i^t\}_{t \in \mathbb{N}}$ with $\lim R_i^t = \infty$,

$$\lim_{t \to \infty} q_i(R_1, \ldots, R_i^t, \ldots, R_n; p_1, \ldots, p_n) > 0.^{14}$$

From desireability we can deduce the following useful lemma.

Lemma 5.1 *Consider any n-tuple of distinct, nonzero probabilities $(p_1, \ldots, p_n)$, and any subset $\mathbf{I} \subset \{1, \ldots, n\}$. If $\{(R_1^t, \ldots, R_n^t)\}_{t \in \mathbb{N}}$ is a sequence of n-tuples of returns with $\{R_i^t\}_{t \in \mathbb{N}}$ for $i \in \mathbf{I}$ nondecreasing and converging to ∞, and $\{R_i^t\}_{t \in \mathbb{N}}$ for $i \notin \mathbf{I}$ converging to $\bar{R}_i$, then then there exists a positive integer M such that for all $t > M$,*

$$\sum_{i \in \mathbf{I}} q_i^t > 0.$$

---

[14] Notice the limit in this case neccesarily exists because the market share of the i^{th} gamble is nondecreasing in own return R_i by monitinicity in return, and thus q_i^t is a monotone sequence in $[0, 1]$.

That is, at least one of the gambles indexed by $\mathbf{I}$ *has a market share greater than* 0 *for all* n*-tuples of returns far along enough in the sequence.*

Proof Consider fixing $R_i = \bar{R}_i$ for $i \notin \mathbf{I}$, and let the returns for the gambles indexed by $i \in \mathbf{I}$ follow the sequence $\{R_i^t\}_{t \in \mathbb{N}}$. The resulting sequence of market shares, which we denote as $\bar{q}_i^t$ can be shown by desireability and monitinicity in return to satisfy

$$\lim_{t \to \infty} \sum_{i \in \mathbf{I}} \bar{q}_i^t > 0.^{15}$$

Thus there exists a positive integer N such that for all $t \geq N$,

$$\sum_{i \in \mathbf{I}} \bar{q}_i^t > 0.$$

In particular then,

$$\sum_{i \in \mathbf{I}} \bar{q}_i^N > 0.$$

Since the q_i functions are continuous in the n-tuple of returns, we can find an $\epsilon > 0$ such that $|\hat{R}_i - \bar{R}_i| < \epsilon$ for all $i \notin \mathbf{I}$ implies

$$\sum_{i \in \mathbf{I}} \hat{q}_i^N > 0.$$

By assumption we can find an N' such that $t > N'$ implies $|R_i^t - \bar{R}_i| < \epsilon$ for all $i \notin \mathbf{I}$. Taking $M = \max\{N, N'\}$ thus ensures that $t > M$ implies

$$\sum_{i \in \mathbf{I}} q_i^t > 0.^{16} \quad \blacksquare$$

6 Partimutuel Market Equilibrium

In a parimutuel betting market, such as a horse race, the return from betting on horse i in the event horse i wins (i.e. the Arrow-Debreu price on horse i as explained in Section 2) is determined by a market clearing rule, namely the "parimutuel" rule. Thus equilibrium in the market requires that the market clearing returns paid by gambles are consistent with the optimal betting behavior by our population of bettors. We have already established the structure of optimal betting behavior in the form of the market share functions q_i over a finite menu of gambles G. We now explain how the behavior of the market share functions allows us

---

[15] Once again we know by monitinicity in return that the limit exists.

[16] More precisely, $t > M$ implies $\sum_{i \in \mathbf{I}} q_i(R_{\mathbf{I}}^t, R_{-\mathbf{I}}^t) \geq \sum_{i \in \mathbf{I}} q_i(R_{\mathbf{I}}^N, R_{-\mathbf{I}}^t) > 0$

to make several sharp predictions concerning price equilibrium in the betting market.

Consider a parimutuel betting market, which for concreteness we shall refer to as a "horse race", but the model applies to betting markets more generally. The market consists of a set of n states $\Omega = \{\omega_1, \ldots, \omega_n\}$, and an exogenous probabilitiy distribution $\mathbf{p}(\omega_i) = p_i$ over the states. Under the usual horse race interpretation, each state ω_i corresponds to the event that a particular horse i wins the race, and each probability p_i corresponds to the probability that horse i wins. We will assume that the p_i are distinct (i.e. the horses are different), and nonzero (each horse has a chance).

For each state ω_i, there is a market determined odds R_i, which is the net return per dollar bet on horse i that is recieved if horse i wins the race. These market odds $(R_1, \ldots, R_n)$ and the exogenous probabilities $(p_1, \ldots, p_n)$ together give rise to a set of n gambles $G = \{(R_1, p_1), \ldots, (R_n, p_n)\}$. The consumer population T of bettors in turn gives rise to market shares q_i for $i = 1, \ldots, n$ for the gambles in G. These market shares in turn feedback to determine new market clearing odds by way of the parimutuel mechanism, and thus equilibrium in the market is attained at odds that are consistent with the market shares they produce.

More precisely, let a_i be the share of the total bet pool that is placed on the bet that horse i wins the race. Then the parimutuel market mechanism determine the net return per dollar bet on horse i as

$$R_i = \frac{(1-\tau)}{a_i} - 1,$$

where $\tau \in (0, 1]$ is the track take.Thus the set of gambles available in the market has the form $G = \{(R(a_i), p_i)\}_{i \in \{1,\ldots,n\}}$.

Assuming that each consumer $t \in T$ bets on a single horse, and all consumers bet the same monetary amount M dollars, then the share of the bet pool placed on horse i is simply equal to the share of the population that chooses to bet on horse i. That is, $a_i = s_i$, where s_i is the share of the population T that chooses to bet on horse i. Thus the set of gambles available in the market has the form $G = \{(R(s_i), p_i)\}_{i \in \{1,\ldots,n\}}$. Note that if the the market share $s_i = 0$, i.e., a zero mass of bettors bet on the i^{th} horse, then since our population is a continuum, the partimutuel mechansim rewards the entire bet pool to this zero mass of bettors, and thus $R(s_i) = \infty$. Of course, in our model, preferences are only defined over gambles having finite returns, which is a difficulty that we handle in the analysis to follow.

Given such a set G, the market share of the i^{th} gamble, as we have already examined, is given by

$$q_i(R(s_1), \ldots, R(s_n); p_1, \ldots, p_n).$$

Thus the market is in equilibrium when, for some market shares $(s_1^*, \dots, s_n^*)$,

$$s_i^* = q_i(R(s_1^*), \dots, R(s_n^*); p_1, \dots, p_n) \quad \text{for} \quad i = 1 \dots, n. \tag{2}$$

In words, the market is in equilibrium when bettors choose among the gambles in proportions that sustain the odds of the gambles.

We now come to a central result of the paper.

Theorem 6.1 *If the probabilities $(p_1, \dots, p_n) \in \Delta^{n-1}$ are distinct and nonzero, and the distribution of consumer preferences satisfies continuity and desireability, then the parimutuel market has unique equilibrium odds $(R(s_1^*), \dots, R(s_n^*))$, with market shares $(s_1^*, \dots, s_n^*) \in \Delta^{n-1}$ distinct and nonzero.*

Proof We prove the result in three steps. In the first step, we introduce an upper bound $\bar{R}$ on the odds payable by a gamble in the market, and show that an equilibrium exists under this restriction by Brouwer's fixed point theorem. This follows from continuity of $\mathbf{P}_V$ which drives the continuity of the q_i. In the second step, we show that it is always possible to raise the upper bound $\bar{R}$ high enough such that it is not binding in equilibrium, and thus an equilibrium of the form (2) exists. This result is driven by the desireability assumption. Lastly we show that the equilibrium is unique, which is driven by monotinicity in return (a consequence of continuity).

We shall assume there is an upper bound $\bar{R}$ on the net returns payable by a gamble. Under this "restriction" to the parimutuel mechanism, the the market shares s_i determine the market returns $\bar{R}_i$ through

$$\bar{R}_i(s_i) = \min\left(\frac{1-\tau}{s_i} - 1, \bar{R}\right).$$

Thus whenever $s_i \leq (1-\tau)/(1+\bar{R})$, the restriction $\bar{R}$ on the odds is binding.

Since the return vector $(\bar{R}(s_1), \dots, \bar{R}(s_n))$ is clearly a continuous function of the market shares $(s_1, \dots, s_n)$, and since the market share function q_i are continuous in returns by Theorem 4.3, we have $f : \Delta^{n-1} \to \Delta^{n-1}$ given by

$$f_i(s_1, \dots, s_n) = q_i(\bar{R}(s_1), \dots, \bar{R}(s_n); p_1, \dots, p_n) \quad \text{for} \quad i = 1 \dots, n,$$

is a continuous function. By the Brouwer fixed point theorem, the map f has a fixed point $(\bar{s}_1, \dots, \bar{s}_n)$, which is thus an equilibrium of the restricted parimutuel market.

If the upper bound $\bar{R}$ is not binding for any of the $\bar{s}_i$, then clearly the fixed point satisfies the property (2) of being an equilibrium in the unrestricted parimutuel market. We now show that it is possible to raise the bar $\bar{R}$ sufficiently high so that it is not binding for the corresponding restricted equilibrium $(\bar{s}_1 \ldots, \bar{s}_n)$.

Suppose that this was not true. Then there exists a sequence of uppper bounds $\{\bar{R}^t\}$ monotonically converging to ∞ with a corresponding sequence of equilibria $\{(\bar{s}_1^t, \ldots, \bar{s}_n^t)\} \subset \Delta^{n-1}$ where for each t the upper bound $\bar{R}^t$ is binding for at least one $\bar{s}_i$. Since this sequence of market shares lives in a compact space, we can find a convergent subsequence $\{(\bar{s}_1^{t_k}, \ldots, \bar{s}_n^{t_k})\}$ converging to $(\bar{s}_1, \ldots, \bar{s}_n)$, with $\bar{s}_i = 0$ for at least one i (which follows from the fact that the restriction is binding for each t_k).

Let $\mathbf{I} \subset \{1, \ldots, n\}$ be the strict subset of indices i for which $\bar{s}_i = 0$. Then for each $i \in \mathbf{I}$, the sequence $\{\bar{R}(\bar{s}_i^{t_k})\}$ converges to ∞,[17] and without loss of generality we can say it converges to ∞ monotonically.[18]

However by desireability, this situation is not possible. It would be mean that there is a sequence of n-tuples of returns $\{(R_1^m, \ldots, R_n^m\}$ with $\{R_i^m\}$ monotonically converging to ∞ for $i \in \mathbf{I}$ and $\{R_i^m\}$ converging to finite R_i for $i \notin \mathbf{I}$, and

$$\lim_{m \to \infty} \sum_{i \in \mathbf{I}} q_i(R_1^m, \ldots, R_n^m; p_1, \ldots, p_n) = 0,$$

which contradicts lemma 5.1. Thus it must be the case that we can find a large enough upper bound $\bar{R}$ such that $\bar{R}$ is not binding for the equilibrium $(\bar{s}_1, \ldots, \bar{s}_n)$. These market shares thus satisfy the condition for $(s_1^*, \ldots, s_n^*)$ in (2). Moreover, these equilibrium market shares must also be located in the interior of Δ^{n-1} (because $\bar{R}$ is nonbinding), i.e., $\bar{s}_i > 0$ for all i. It also follows that since the probability distribution $(p_1, \ldots, p_n)$ involved distinct probabilities, the equilibrium market shares $(\bar{s}_1, \ldots, \bar{s}_n)$ must be distinct, since otherwise one gamble in the market would dominate another, thereby causing the latter to have zero market share, which would contradict the fact the equilibrium shares lies in the interior of the simplex.

We now address uniqueness. Suppose that there exist two n-tuples of market shares $(\bar{s}_1, \ldots, \bar{s}_n)$ and $(s_1^*, \ldots, s_n^*)$ that satisfy the equilibrium condition (2). Then for $i = 1, \ldots, n$,[19]

$$\bar{s}_i = q_i(R(\bar{s}_1), \ldots, R(\bar{s}_n)) \quad \text{and} \quad s_i^* = q_i(R(s_1^*), \ldots, R(s_n^*)).$$

Since both equilibrium tuples are located in the simplex, it must be the case that for some nonempty strict subset $\mathbf{I} \subset \{1, \ldots, n\}$, $s_i^* \le \bar{s}_i$ for all $i \in \mathbf{I}$ with a strict inequality for at least one $i \in \mathbf{I}$, and and $s_i^* \ge \bar{s}_i$ for

---

[17] Since $\min\left(\frac{1-\tau}{s_i^{t_k}} - 1, \bar{R}^{t_k}\right)$ goes to ∞.

[18] We can always take a subsequence to assure monotonic convergence

[19] We supress the probabilities $(p_1, \ldots, p_n)$ in the following notation because they remain the same.

all $i \notin \mathbf{I}$ with a strict inequality for at least one $i \notin \mathbf{I}$. This imples that

$$\sum_{i \in \mathbf{I}} s_i^* < \sum_{i \in \mathbf{I}} \bar{s}_i,$$

However we also have that $R(s_i^*) \geq R(\bar{s}_i)$ for $i \in \mathbf{I}$ and $R(s_i) \leq R(\bar{s}_i)$ for $i \notin \mathbf{I}$, which by monotinicity in return of $\mathbf{P}_V$, implies that

$$\sum_{i \in \mathbf{I}} s_i^* \geq \sum_{i \in \mathbf{I}} \bar{s}_i.$$

This is a contradiction, and thus the equilibrium is unique. ▮

7 Recovering Beliefs from Odds

We have established that for any n-tuple of distinct and nonzero probabilities $(p_1, \dots, p_n) \in \Delta^{n-1}$, which represent the common prior distribuition over horses (or states of the world more generally) held by bettors, there exists a unique n-tuple of equilibrium odds $(R(s_1^*), \dots, R(s_n^*))$ with distinct and nonzero market shares $(s_1^*, \dots, s_n^*) \in \Delta^{n-1}$.

In empirical practice, we do not observe the probability distribution $(p_1, \dots, p_n)$ known by the bettors in a horse race. Rather we only observe the equilibrium odds $(R(s_1^*), \dots, R(s_n^*))$, and implicitly we also observe the equilibrium market shares $(s_1^*, \dots, s_n^*)$. The key questions is thus the following : knowing the equilibrium odds and market shares, and knowing the distribution of consumer preferences $\mathbf{P}_V$, if all bettors in the market hold common beliefs $(p_1, \dots, p_n)$ about the chances of each horse winning, then can we uniquely recover the $(p_1, \dots, p_n)$? That is, for any observed n-tuple of distinct equilibrium odds $(R(s_1^*), \dots, R(s_n^*))$, with the implicitly observed n-tuple of distnct and nonzero market shares $(s_1^*, \dots, s_n^*) \in \Delta^{n-1}$, can we uniquely solve the system of equations in $(p_1, \dots, p_n) \in \Delta^{n-1}$,

$$s_i^* = q_i(R(s_1^*), \dots, R(s_n^*); p_1, \dots, p_n). \quad \text{for} \quad i = 1, \dots, n \tag{3}$$

Theorem 7.1 *If the odds $(R(s_1^*), \dots, R(s_n^*))$ are distinct, and the distribution of consumer preferences satisfies continuity, then there exists a unique probability distribution $(p_1^*, \dots, p_n^*) \in \Delta^{n-1}$, consisting of distinct and nonzero probabilities, that solves* (3).

Proof Since the R_i^* are assumed distinct, and $\mathbf{P}_V$ satisfies continuity, then we that for any $(p_1, \dots, p_n) \in$

Δ^{n-1},

$$\sum_{i=1}^{n} s_i^* - q_i(R_1^*, \ldots, R_n^*; p_1, \ldots, p_n) = 0.^{20} \tag{4}$$

Furthermore, by continuity once again, $q_i(R_1^*, \ldots, R_n^*; p_1, \ldots, p_n)$ is continuous in $(p_1, \ldots, p_n)$ over Δ^{n-1}. Now consider the following continous self map over Δ^{n-1} (where for simplicity we write $R^* = (R_1^*, \ldots, R_n^*)$). For $(p_1, \ldots, p_n) \in \Delta^{n-1}$,

$$p_i \mapsto \frac{p_i + \max(0, s_i^* - q_i(R^*; p_1, \ldots, p_n))}{\sum_{j=1}^{n}(p_j + \max(0, s_j^* - q_j(R^*; p_1, \ldots, p_n)))} \quad \text{for} \quad i = 1, \ldots, n. \tag{5}$$

By the Brouwer fixed point theorem, this map must have a fixed point $(p_1^*, \ldots, p_n^*) \in \Delta^{n-1}$.

Moreover, this fixed point must satisfy $q_i(R^*, p_1^*, \ldots, p_n^*) = s_i^*$ for $i = 1, \ldots, n$. If these are equalities are not satisfied, then by (4) we have that for at least one i we have $s_i^* > q_i(R^*, p_1^*, \ldots, p_n^*)$, and for at least one j we have $s_j^* < q_j(R^*, p_1^*, \ldots, p_n^*)$. Thus

$$\sum_{i=1}^{n} p_i^* + \max(0, s_i^* - q_i(R^*, p_1^*, \ldots, p_n^*)) > 1,$$

and under the the mapping (5), p_j^* must get sent to a strictly smaller number, which violates the fact that $(p_1^*, \ldots, p_n^*)$ is a fixed point. Thus $(p_1^*, \ldots, p_n^*)$ solves (3)).

Since each s_i^* is nonzero and distinct, it must be the case that each p_i^* is nonzero and distinct. Otherwisem some gamble, indexed by i say, would be dominated (either it has zero probability or it has the same probability as another gamble but lower return), and thus $q_i(R^*, p_1^*, \ldots, p_n^*) = 0$, which is inconsistent with the fact the market share of the i^{th} gamble is $s_i^* > 0$.

The uniquness of $(p_1^*, \ldots, p_n^*)$ follows from monotinicity in a manner parallel to that used in proving the uniquness in Theorem 6.1, except exploiting monitinicity in probability instead of monitinicity in return (both of which recall follows from continuity). ▌

8 The Invertible Pricing Function and Estimation

Thus by way of Theorems 6.1 and 7.1, we have established that for any horse race or parimutuel betting market with $n \geq 2$ states of the world, and a distribution of consumer preferences $\mathbf{P}_V$ over gambles satisfying continuity, there exists an *invertible equilibrium pricing function* $R_i^*(p_1, \ldots, p_n)$ for $i = 1, \ldots, n$. This pricing

---

[20] Note that $\sum s_i^* = 1$ because the s_i are (observed) market shares.

function maps any distinct and nonzero n-tuple of probabilities $(p_1, \ldots, p_n) \in \Delta^{n-1}$ to a n-tuple of returns $(R(s_1^*, \ldots, R(s_n^*))$, with distinct and nonzero market shares $(s_1^*, \ldots, s_n^*) \in \Delta^{n-1}$, that solves the equilibrium condition (2). We express the inverse equilibrim pricing function as

$$p_i^*(R(s_1^*), \ldots, R(s_n^*)) \quad \text{for} \quad i = 1, \ldots, n,$$

which maps any n-tuple of returns $(R(s_1^*), \ldots, R(s_n^*))$ with distinct and nonzero market shares $(s_1^*, \ldots, s_n^*) \in \Delta^{n-1}$ to a distinct and nonzero n-tuple of probabilities $(p_1, \ldots, p_n) \in \Delta^{n-1}$. It is straightforward to show that this inverse pricing function satisfies the symmetry condition

$$p_i^*(R_i, R_{-i}) = p_i^*(R_i, Q_{-i}),$$

where Q_{-i} is a permutation of the elements in R_{-i}.

This inverse equilibrium pricing function provides the key to empirically analyzing the model. Our data consist of a sample of races where the winning horse and the odds on each horse and the track take can be observed. One approach to the analysis is to estimate the inverse equilibrium pricing function nonparametrically from the winners and odds data. Then for any parametric assumption about the distribution of preferences, we can estimate the parameters so as to minimizes the distance between observed shares and predicted shares.

Alternatively, we can mimic the empirical strategy of Jullien and Salanie [2000] and numerically solve for the inverse probabilities for any given value of the model parameters and the observed odds. Using these probabilities for each race along with the the observed winner of each race allows us to construct a likelihood function over model parameters. Further work on the empirically implementing the model is currently underway.

9 Numerical Example : The Favorite Longshot Bias

We now present a numerical illustration of the model of price equilibrium in the betting market that we have developed. In particular, we illustrate how heterogeneous expected utility maximizers naturally produce the favorite longshot bias in equilibrium by way of trade between bettors with different risk tolerances. Essentially, in equilibrium, the most risk loving bettors select ino the horse with the highest odds, and the next most risk loving bettors select into the horse with the second highest odds, and so forth until the we reach the most risk averse bettors, who select into the horse with the lowest odds. In order for this assortment to be an equilibrium, the more risk loving bettors overbet longshots, thereby driving down the expected return on longshots and driving up the expected return on favorite, which gives the more risk averse bettors the incentive to bet on favorties.

Let us illustrate this point using the constant absolute risk aversion (CARA) utility function for wealth, whose funtional form is

$$u(x,\theta) = \frac{1-\exp(-\theta x)}{\theta},$$

where θ is the CARA risk aversion coefficient that is allowed to be heterogeneous across bettors. We assume θ is distributed normally in the population with mean 0, i.e., the average behavior in the population is risk neutrality. It is straightforward to show that this model of preference heterogeneity satisfies continuity and desireability (see Section 10), and thus unique equilibrium prices exist. We take the track take to be 0, and consider a race with 4 horses having probabilities of winning $\mathbf{p} = (.6, .25, .1, .05)$. Thus horse 1 is the extreme favorite and horse 4 is the extreme longshot.

We compute the equilibrium prices in the market for different levels of variance of the distribution of the CARA risk aversion coefficent. In Figure 1, we plot the equilibrium expected returns for each horse using three different levels of variance. Notice that the magnitude of the favorite longshot bias is made larger as the variance gets larger. Essentially, as a greater number of extreme risk lovers become more probable with the larger variance of the distribution, a greater number of extreme risk lovers sort into longshots, and thus magnify the bias. Notice also that an increasingly risk loving representative bettor would be required to explain the equilibrium prices in Figure 1 as the variance gets larger. Thus the bias of the representative bettor model is made larger as the variance of the distribution is increased (because the average behavior in the population is always risk neutrality).

10 Nonparametric Identification

In this section, we reduce heterogeneity to a single dimensional type that varies across bettors, and ask to what extent can our preference model be nonparametrically identfified. Assuming that we can exogenously vary gambles, we present an initial result in this direction. By reducing heterogeneity to a single dimensional type, we prove nonparametric identification under a regularity assumption on the utility function, namely a single crossing condition.

Our true population model consists of two components :

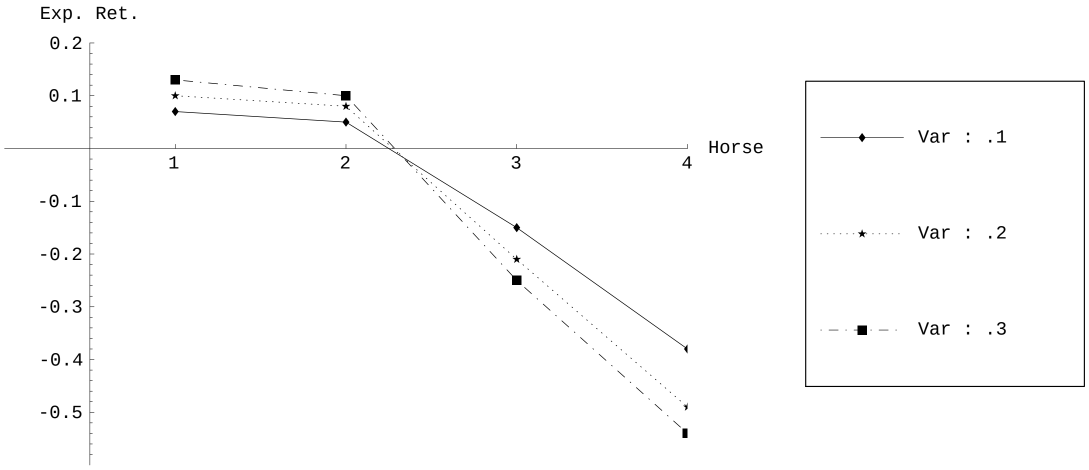

Figure 1: Equilibrium Expected Returns

- We have a parameterized utility function over gambles

$$W : [0,1] \times \mathbf{R} \times \mathbb{R} \to \mathbb{R},$$

  where $W(p, R, \theta)$ is the utility that a person of type θ recieves from consuming the gamble (p, R). We restrict W to be continuous in (p, R, θ) and strictly increasing in p and R.

- The type θ is distributed according to a strictly increasing cdf F over the support of θ with stricly increasing inverse cdf F^{-1}.

We can reparameterize the model in terms of $\tau = F(\theta)$, which is distributed uniformly over $(0, 1)$. A person of type τ recieves utility $V(p, R, \tau)$ from consuming the gamble (p, R), where

$$V(p, R, \tau) = W(p, R, F^{-1}(\tau)).$$

Clearly, V is continuous in (p, R, τ) and strictly increasing in p and R.

We are interested in the following identification problem. Suppose we can exogenously vary a finite number of n gambles $\{(p_1, R_1), \ldots, (p_n, R_n)\}$ for some $n \geq 2$, and we can observe $q_i(R_1, \ldots, R_n; p_1, \ldots, p_n)$ where

$$\begin{aligned} q_i(R_1, \ldots, R_n; p_1, \ldots, p_n) &= \mathbf{P}_\theta \left[W(p_i, R_i, \theta) = \max_j W(p_j, R_j, \theta) \right] \\ &= \mathbf{P}_\tau \left[V(p_i, R_i, \tau) = \max_j V(p_j, R_j, \tau) \right]. \end{aligned}$$

What feature of our population model can we nonparametrically identify? In this note, we show that by using an identifying assumption on W, namely a single crossing condition, we can nonparametrically identify V up to an increasing transformation.

Single Crossing The function $W(p, R, \theta)$ satisfies the single crossing condition if the following holds true : Supposing that for any two gambles (p_1, R_1) and (p_2, R_2) with $p_1 > p_2$, and for some θ, we have that

$$W(p_1, R_1, \theta) \geq W(p_2, R_2, \theta),$$

then $\theta' > \theta$ implies that

$$W(p_1, R_1, \theta') > W(p_2, R_2, \theta').$$

In words, the single crossing condition says that if a person of type θ is willing to make the tradeoff between return and probability in moving from (p_2, R_2) to (p_1, R_1), then a person of type $\theta' > \theta$ is also willing to make the tradeoff. This single crossing condition is satisfied in all the standard parametric forms of W.

Theorem 10.1 (Main Result) *If W satisfies the single crossing condition, then V is nonparametrically identified up to an increasing transformation.*

We approach the proof in four steps. In the first step, we show that if W satisfies the single crossing condition (in θ), then V satisfies the single crossing condition (in τ). Secondly, we show that the single crossing implies that our population model satisfies the "No Ties" property. Then we show that for any set of n distinct gambles, $(0, 1)$ can be divided into adjacent intervals, with the subpopulation of τ's belonging to the i^{th} interval preferring the i^{th} gamble. Finally, the nonparametric identification of V up to an increasing transformation follows.

Lemma 10.2 *If W satisfies the single crossing condtition, then V satisfies the single crossing condition.*

Proof The result follows from the definitions. Suppose that for gambles (p_1, R_1) and (p_2, R_2) with $p_1 > p_2$, and for some τ, we have that

$$W(p_1, R_1, F^{-1}(\tau)) = V(p_1, R_1, \tau) \geq V(p_2, R_2, \tau) = W(p_2, R_2, F^{-1}(\tau)).$$

Now consider $\tau' > \tau$. Then since W satisfies the single crossing condition and F is strictly increasing, then

$$V(p_1, R_1, \tau') = W(p_1, R_1, F^{-1}(\tau')) > W(p_2, R_2, F^{-1}(\tau')) = V(p_2, R_2, \tau'). \quad \blacksquare$$

Next we establish that the single crossing properties ensures that our population model satisfies the "No Ties" property.

Lemma 10.3 (No Ties) *For any two gambles (p_1, R_1) and (p_2, R_2) with $p_1 > p_2$,*

$$\mathbf{P}_\tau \left[V(p_1, R_1, \tau) = V(p_2, R_2, \tau) \right] = 0.$$

Proof Consider the set

$$A = \{\tau : V(p_1, R_1, \tau) = V(p_2, R_2, \tau)\}.$$

The largest A can be is a singleton. To see why, suppose A is larger than a singleton, i.e, a set with at least two elements $\tau_1 < \tau_2$. Then by the single crossing property, we must have that $V(p_1, R_1, \tau_2) > V(p_2, R_2, \tau_2)$, which is a contradiction. A singleton set has probability measure zero under the uniform distribution. ∎

Consider a finite set of distinct gambles

$$G_n = \{(p_1, R_1), \ldots, (p_n, R_n)\},$$

with $p_1 < \cdots < p_n$. By the "No Ties" lemma. we have that

$$q_i(R_1, \ldots, R_n; p_1, \ldots, p_n) = \mathbf{P}_\tau(I_i),$$

where

$$I_i = \{\tau : V(p_i, R_i, \tau) > V(p_j, R_j, \tau) \text{ for all } j \neq i\}.$$

Lemma 10.4 *For each $i = 1, \ldots, n$, $I_i = \emptyset$ or $I_i = (a_i, b_i)$ for $a_i < b_i$.*

Proof Suppose I_i is nonempty. We shall show that I_i must be both open and convex, and hence must take the form of an interval (a_i, b_i) with $a_i < b_i$.

The fact that I_i is open follows simply from the continuity of V in τ. Consider now convexity. Suppose $\tau_1, \tau_2 \in I_i$, $\tau_i < \tau_2$. Then for any τ such that $\tau_1 < \tau < \tau_2$, it follows directly from single crossing that

$$V(p_i, R_i, \tau) > V(p_j, R_j, \tau) \text{ for all } j < i.$$

It also follows from single crossing that *it cannot be the case* that

$$V(p_j, R_j, \tau) \geq V(p_i, R_i, \tau) \text{ for all } j > i.$$

Thus the only possibility is that

$$V(p_i, R_i, \tau) > V(p_j, R_j, \tau) \text{ for all } j \neq i,$$

i.e., $\tau \in I_i$. ∎

Remark Let $\{i_1, \ldots, i_m\}$ be the subset of $\{1, \ldots, n\}$ for which $I_{i_k} \neq \emptyset$. Arrange $\{i_1, \ldots, i_m\}$ in order of increasing probability, i.e., $i_k > i_j$ implies $p_{i_k} > p_{i_j}$. Then it follows from single crossing and lemma 10.4 that

$$I_{i_1} = (a_0, a_1), I_{i_2} = (a_1, a_2), \ldots, I_{i_m} = (a_{i_{m-1}}, a_{i_m}),$$

with

$$0 = a_0 < a_1 < \cdots < a_m = 1.$$

Thus $q_{i_j}(R_1, \ldots, R_n; p_1, \ldots, p_n) = (a_j - a_{j-1})$.

Now we are ready to demonstrate nonparametric identification. Suppose we have a V' function that is *not* related to the true V function by an increasing transformation, but satisfies all of our same assumptions on V. Then there exists some τ and some two gambles (p, R) and (p^*, R^*) with $p^* > p$ such that

$$V(p, R, \tau) < V(p^*, R^*, \tau)$$

and

$$V'(p, R, \tau) > V'(p^*, R^*, \tau).$$

That is, for person τ, the tradeoff between probability and return in moving from (p, R) to (p^*, R^*) is acceptable under V but not acceptable under V'.

Now consider a finite set of gambles G_n, arranged in order of increasing probability, where the gambles (p, R) and (p^*, R^*) are two highest probability gambles in G_n. Thus (p, R) is the $(n-1)^{st}$ gamble in G_n and (p^*, R^*) is the n^{th} gamble in G_n. Then as a consequence of the remark, it must be the case that

$$q_n(R_1, \ldots, R_n; p_1, \ldots, p_n) > q'_n(R_1, \ldots, R_n; p_1, \ldots, p_n).$$

Thus we are nonparametrically identified up to an increasing transformation of V. ∎

11 Conclusion

We have presented a new model of equilibrium in parimutuel betting markets that serves as the basis for empirically testing expected utility theory against cumulative prospect theory using market data. Under very weak conditions on the distributions of preferences, we has shown that there exists unique equilibrium odds on each “horse” that clears the market. Moreover, from the observed odds, we can solve for the common prior held by bettors as to the chances of each horse winning. Using a parametric specification of the distribution of preferences, this allows us to construct a likelihood function using the observed winner in each race. Finally, we present an intitial result on the nonparametric identification of our model that is applicable when we reduce heterogeneity to a single dimensional type that varies across the population. Future work is aimed at empirically implementing the model and carrying out comparisons under different paramteric assumptions on the distributions of preferences. The key comparisons of interest include hetorogeneous expected utility maximizers versus a representative cumulative prospect theoretic agent, and heterogeneous expected utility maximizers versus heterogeneous cumulative prospect theoretic agents.

References

Colin Camerer. Prospect theory in the wild. In Daniel Kahneman and Amos Tversky, editors, *Choices, Values, and Frames*. Cambridge University Press, 2000.

Robin P. Cubitt, Chris Starmer, and Robert Sugden. Discovered preferences and the experimental evidence of violations of expected utility theory. *Journal of Economic Methodology*, 8(3):385–414, 2001.

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

Bruno Jullien and Bernard Salanie. Empirical evidence on the preferences of racetrack bettors, 2002. URL `http://www.crest.fr/pageperso/lei/salanie/HandbookJullienSalanie.pdf`. mimeo.

Richard E Quandt. Betting and equilibrium. *The Quarterly Journal of Economics*, 101(1):201–07, 1986.

Matthew Rabin and Richard H. Thaler. Anomalies: Risk aversion. *Journal of Economic Perspectives*, 15 (1):219–232, 2001.

Eric Snowberg and Justin Wolfers. Explaining the favorite longshot bias : Is it risk-love or misperceptions, 2005. URL `http://www.crest.fr/pageperso/lei/salanie/HandbookJullienSalanie.pdf`. mimeo.

Chris Starmer. Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature*, 38(2):332–382, 2000.

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

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

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