PARTITION DEPENDENCE IN BINARY OPTION AND PREDICTION MARKETS: FIELD AND LAB EVIDENCE

Ulrich Sonnemann*

Colin F. Camerer**

Craig R. Fox***

Thomas Langer*

JEL classification: D8, G1. Keywords: prediction markets, partition-dependence, framing effects, experimental economics, field experiments, behavioral economics

11/20/09. Comments welcome. We thank audiences at Toronto (JDM meeting 2005), Muenster (July 2007), Mannheim (June 2006), Caltech (May 2007), Rome (ESA 2007), Warsaw (SPUDM 2007), and Chicago (October 2007). Thanks to Sera Linardi and Stephanie Wang for help. This research was supported by the DFG-grant LA1316/3-1, NSF grant SES-00-99209 (CRF), and Moore Foundation, NSF-HSD and HSFP grants (CFC). Affiliations: *University of Muenster, **Caltech, ***UCLA.

Direct correspondence to: Colin Camerer

PARTITION DEPENDENCE IN BINARY OPTION AND PREDICTION MARKETS: FIELD AND LAB EVIDENCE

Abstract

Financial instruments such as binary options, credit-default swaps, and catastrophe bonds offer a fixed payment if a future event occurs. Markets for such instruments are prediction markets in which prices reflect aggregate subjective beliefs of events occurring. Psychology experiments have shown that judged probabilities can be affected by the partition of the state space ("partition-dependence"). We report evidence from field data on macroeconomic derivatives and horse racing markets, and lab data from short-run (two-hour) and long-run (seven-week) experiments on pricing events in two different partitions. All four markets show some evidence of partition-dependence in market prices and judgments.

(100 words)

I. INTRODUCTION

In recent years a number of financial products have been introduced that offer one fixed amount when in-the-money and a different fixed amount when out-of-the money. These binary options are also known as "all-or-nothing options," "digital options" on foreign exchange and interest rate markets, and "fixed return options" on the American Stock Exchange. Similarly, credit-default-swaps typically offer a premium return unless a company defaults on its debt. Catastrophe bonds typically pay a premium return unless a natural disaster occurs (e.g., a hurricane in Florida that leads to more than $1B in claims), in which case the principal is lost entirely.

Binary options have long been available over-the-counter (OTC), mainly as part of more complex structured products. They have recently been introduced as stand-alone products at the Chicago Board Options Exchange (CBOE)[1], the Chicago Board of Trade (CBOT)[2], and EZTrader and HedgeStreet, which both offer binary options on equity indices, commodities, and currencies.

The closing price of a binary option is related to the market's aggregate belief that a foreign exchange rate will close above a particular level price, a company will default on its debt, or this year's hurricane season will be especially malicious. How closely prices are related to beliefs can also depend on the investors' risk tastes. This is a difficult question we discuss further

[1] Binary options traded at the CBOE are based on the S&P 500 and the CBOE Volatility Index (www.cboe.com/binaries).

[2] The CBOT recently created binary options markets on the Federal Funds target rate, a leading indicator of the U.S. economy (Majumder, Diermeier, Rietz, and Amaral (2009)).

below, and which some of our experimental data can clearly address since both individual beliefs and aggregate prices are measured.

Some binary option markets, called “prediction markets” are specifically designed to provide probability forecasts of particular natural events, such as the outcomes of political elections, corporate sales figures, or sporting events. Recent research suggests that it is possible to obtain excellent probability forecasts for future states of the world (e.g., the Dow-Jones year-end close falls within the interval [8200, 8400]) from the market price of a binary option winner-take-all contract that pays when in the money. Remarkably, these predictions are found to be generally more accurate than those derived from opinion polls or expert judgments (Berg and Rietz (2003); Wolfers and Zitzewitz (2004)).

By partitioning the state space into an exclusive and exhaustive set of target events on which winner-take-all contracts are traded, it is possible to obtain a market probability distribution over these events. Occasionally, binary option and prediction markets are defined by canonical partitions over events that are defined by natural categories (e.g., Company X loses or retains its AAA credit rating; a Democrat rather than a Republican or Independent is elected President). More often, they are defined by arbitrary partitions of continuous distributions (e.g., the price of the Turkish Lira will close below $1.20 vs. above $1.20; the Dow-Jones will close in the range [less than 8000], [8000-9000], [more than 9000]). Neoclassical economic theorizing (e.g., subjective expected utility) holds that the subjective probability distribution over events should not be affected by the particular partition of the event space that happens to be selected. This paper tests the validity of this assumption based on earlier psychological evidence, using two new analyses of field data sets and two new experiments. The research question is whether different partitions affect pricing, to what degree and in what types of markets, how such effects in prices are linked to individual judgments.

Recent psychological experiments have shown that the judged probability distribution of a continuous variable, such as the closing price of a stock index, varies systematically with the particular intervals into which the variable's possible values are divided. This phenomenon is known as "partition-dependence." In particular, judged probabilities seem to reflect reliance on a diffuse or "ignorance" prior probability of $1/N$ for each of the N intervals into which the state space is partitioned, plus an adjustment up or down for specific likelihood of each event.[3] The present paper asks to what extent market forces mitigate partition-dependence observed in individual judgments. The question is important for financial markets since some assets (such as credit default swaps) are claims based on events defined using a particular partition. It is possible that partition-dependence influences the prices of these assets.

Bias toward the ignorance prior probability distribution implies that "unpacking" an interval $[I_1, I_2]$ into two separate sub-intervals $[I_1, I_1+x)$ and $[I_1+x, I_2]$ increases the total judged probability from adding the two sub-intervals. For instance, in one study Fox and Clemen (2005) asked expert members of the Decision Analysis Society (DAS) to assess the probabilities that the total number of members of their society would fall into different ranges five years in the future (the current number was 764.) Fifty-eight of 169 contacted members participated (34%) and

[3] Bias toward the ignorance prior can explain a previously documented tendency for probabilities assigned to particular causes of system failure in "fault trees" to diminish when pruned and included implicitly within a residual category (Fischhoff, Slovic, and Lichtenstein, 1978).

were randomly assigned to either a low partition group or a high partition group. The low group was asked to assign likelihoods of membership falling in each of the intervals [0, 400], [401, 600], [601, 800], [801, 1000], [1001+]. The high group was asked the likelihoods for the membership intervals [0, 1000], [1001, 1200], [1201, 1400], [1401, 1600], [1601+]. To rule out the possibility that the presented partition conveys private information about the experimenter's subjective beliefs, both the high and low group subjects were informed about the partitions in both experimental conditions, and subjects were assigned to the two conditions randomly.[4] Nevertheless, the authors found pronounced partition dependence: the median judged probability that the DAS would have more than 1000 members was 10% for those in the "low" partition condition (for whom this was one out of five judgments), but it was 35% for those in the "high" partition condition (for whom this was the sum of four out of five judgments).

Partition-dependence has been replicated in controlled learning environment (See Fox, and Rottenstreich (2006)) with incentive-compatible payoffs (Fox and Rottenstreich (2003); Fox and Levav (2004); Fox and Clemen (2005)), and a related tendency has been found in experimental valuation of hypothetical insurance policies (Johnson et al. (1993)). Partition-dependence has also been shown in allocation of savings to personal investments (Langer and

[4] In the high group, for example, the fact that there are four narrow intervals covering numbers above 1000 could suggest that the experimenters expected the actual number to be low so that many intervals were needed to include the likely values. However, in this study both groups of subjects were told about both sets of partitions and participants randomly assigned themselves to a partition based on the last digit of their primary home telephone number. Thus, the same information was communicated to both groups and subjects knew their own partition assignment was random so that any information conveyed by the partitions presented should affect both groups equally.

Fox (2009); Benartzi and Thaler (2001)) and capital allocations to various businesses in multi-division firms (Bardolet, Fox, and Lovallo, (2009); Scharfstein and Stein (2000)).

Despite the robustness of partition-dependence across domains and to small incentives the question frequently arises whether individual biases shown in surveys will persist in market trading in both laboratory and field settings (e.g., Fehr and Tyran (2005); Camerer and Fehr (2006); Della Vigna (2009)). In this paper we test, for the first time, whether partition-dependence affects prices of assets in binary option markets, both in field and experimental data. If partition-dependence applies to financial markets, this would suggest that asset prices in binary option and prediction markets might be more predictable, from the partition structure alone, and less well-calibrated than has been commonly asserted.

Naturally, if there is a tendency toward partition-dependence in judgment and pricing, institutions may have evolved (or may yet evolve) to take advantage of this tendency. For example, divisional managers may be eager to split a growing division into two new divisions to gain from the apparent bias in capital allocation toward $1/N$. The structure of debt into several subordinated tranches (i.e., groups of similarly-risky debt obligations) could also take advantage of these biases.[5] We do not have any evidence on whether these institutional responses occur, but our results suggest they are possible and could inspire future empirical work.

In order to provide converging evidence of our interpretation we present a range of contexts and methods. We present four studies of naturally occurring and experimental binary

[5] To illustrate, suppose there are 10 tranches ordered by riskiness. If the first nine tranches are small in size, and holders of the safest debt in the large 10th tranche do not account for the difference in tranche sizes, debt holders in that tranche might feel safer than they ought to, thinking there are 9 tranches ahead of them must default before their debt is threatened.

options markets with substantial incentives and opportunities for learning, with substantial payoffs linked to choices and outcomes, and in which trading takes place over time periods lasting from ten minutes to several weeks.

Plan of the Paper and Preview of Results

The next four sections of this paper present evidence of partition-dependence in four market settings. Section II describes data from naturally-occurring "economic derivatives markets." A structural model of these data, which assumes that observed prices reflect partial reliance on a $1/N$ ignorance prior belief, leads to the prediction that some systematic forecasting errors should show up in pricing, an implication of partition dependence that we verify in the data. A brief Section III extends this analysis to explicitly examine the influence of partitions in horse racing betting. We find stronger over-betting in races with fewer horses, which is consistent with a bias toward $1/N$ (since $1/N$ is larger when there are fewer horses). In Section IV we describe short-run experimental asset markets (two 10-minute trading periods) for three naturally-occurring event domains in which we systematically manipulate partitions (strike prices that define binary options) and observe concomitant shifts in equilibrium prices that parallel biases in mean probability judgments for events in different partitions. Finally, Section V extends this observation to a longer-run market experiment conducted on the Web and lasting several weeks, in which participants traded assets based on team wins in the NBA playoffs and goals scored in the FIFA World Cup.

Full instructions for the lab and field experiments, and many technical details, are presented in a set of Appendices.

II. STUDY 1: BINARY OPTION MARKETS FOR ECONOMIC DERIVATIVES

In October 2002, Goldman Sachs and Deutsche Bank launched large-scale markets for bets on the outcomes of macroeconomic indicators. The indicators are the change in U.S. non-farm payrolls (NFP), levels of the Institute for Supply Management's (ISM) purchasing manager index (a measure of business confidence), U.S. initial jobless claims (IJC) (adjusted to reflect seasonal hiring patterns), retail sales (RSX) (excluding automobiles, adjusted for normal seasonal variations), and others. These "economic derivatives" (ED) markets were designed to give professionals such as institutional traders (hedge funds, proprietary traders, pension funds, large banks, etc.) the opportunity to take positions in unexpected fluctuations of macroeconomic risks, and potentially to provide better widespread distributional forecasts of the underlying variables.

For each underlying numerical variable (i.e., the release of a specific numerical macroeconomic indicator) a diverse set of contracts was available for trading, including capped options (capped calls and floored puts), forwards (range forwards), and digital (binary) options (digital calls, digital puts, and digital range options) (see CME (2005)).[6] Contrary to plain-vanilla options, digital range options, for example, have two strike prices (a lower and an upper bound) that define a certain interval of the possible outcomes of the indicator. These options pay

[6] Trading in digital options represents about three quarters of trades, but less than half in terms of transaction volume. Hedgers usually prefer plain-vanilla options which account for much larger volumes; see Beber and Brandt (2009, p. 12).

out a fixed amount ($1) if the released statistic falls within this range and nothing otherwise. Therefore, digital range options in economic derivatives have the same basic structure as the all-or-nothing contracts that are common in standard financial prediction markets.

The market mechanism employed is a parimutuel system (see Appendix A), which is generally used in horse race betting and explored a little in the lab (e.g., Axelrod et al. (2009)). In this mechanism the prices of the instruments are based solely on relative demand for their implied outcomes. Investors who bet on event *A* and win (i.e., event *A* occurs) share the winnings from those who bet on all other ("losing") events. As parimutuel clearing applies to all kinds of derivatives offered in the economic derivatives markets (capped options, forwards, and digital (binary) options), these instruments are decomposed into a combination of several "state contingent claims" (SCC) for valuation. The state contingent claims are in fact digital range options based on available exercise prices, which highlights the relevance of the different strike-price intervals of the state space for the pricing of these derivatives. As in horse betting, the trading system periodically discloses interim prices showing what the payouts would be if no further orders were submitted. These auctions typically take place in the morning before the economic statistic is released and are sometimes preceded by other auctions on the same statistic release one or two days before (e.g., non-farm payrolls auctions are held on the morning the data are released and the two days before). Thus, these markets usually have a very short-term forecast horizon and thereby offer hedging opportunities against so-called event risks. Each auction lasts for 1–2 hours.

Market prices can be used to derive a risk neutral density function of the market's aggregated beliefs about the outcome of every single data release. For instance, Figure 1 shows these probabilities from one set of digital options, for a retail trade statistic announced in April 2005. As illustrated in the Figure, these markets typically entailed 10 to 20 equidistant strike prices for each upcoming data release, dividing the state space in 11 to 21 mutually exclusive

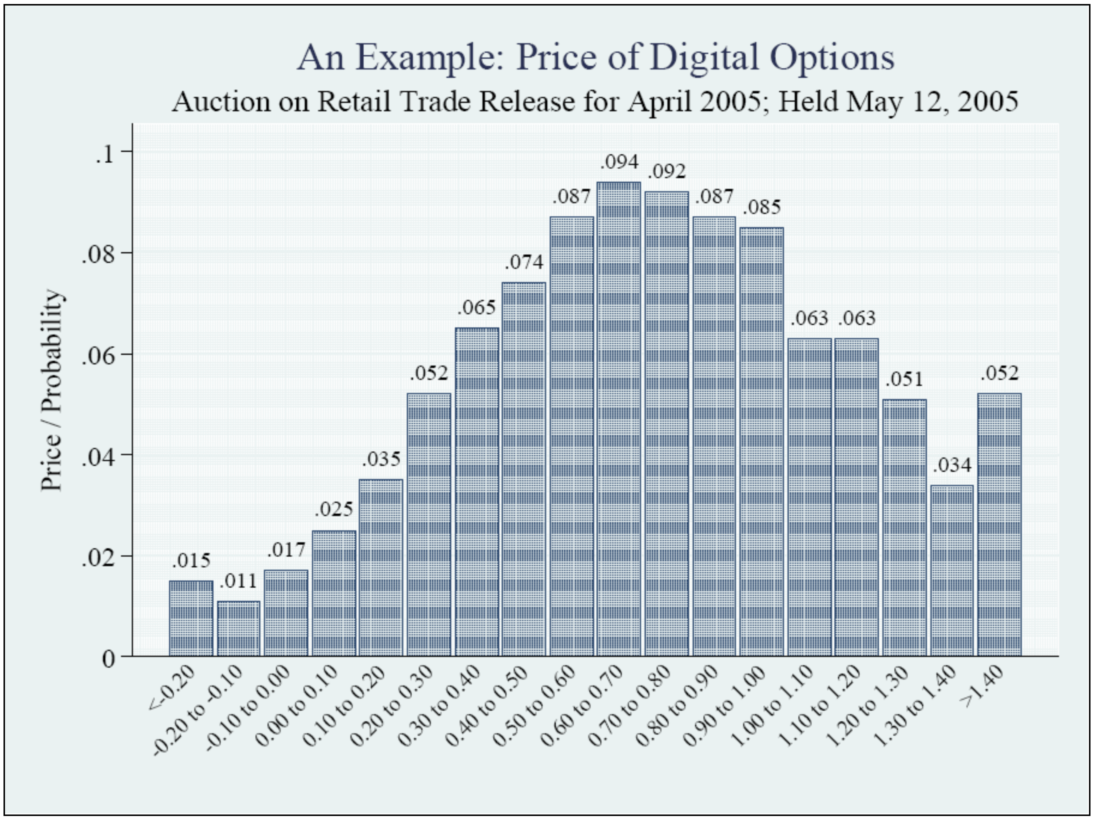

Figure 1. Prices of various digital options in the market for macroeconomic derivatives. Payouts are determined by the retail trade release for April 2005 (taken from Gürkaynak and Wolfers (2006)).

and collectively exhaustive ranges. The ranges are deliberately set to reflect the likely range of the outcomes: strike prices are set to cover at least two to three standard deviations based on historical volatility of the indicator and the scale midpoint is chosen to reflect mean survey expectations.[7] For example, the strike prices for a U.S. non-farm payrolls auction may range from –250,000 (i.e., a decrease in payrolls from the previous level) to +100,000 in increments of 25,000 jobs. This generates a total of sixteen intervals and a scale midpoint of –75,000 jobs (]–∞; –250,000], [–249,999; –225,000], …, [75,001; 100,000], [100,001; +∞[).

Obviously, one cannot know perfectly from the prices alone how well the probabilities implied by these market prices related to the distribution of subjective probabilities among investors. In principle, we could try to mimic the procedure of Bliss and Panigirzoglou (2004) to estimate the risk aversion of the economy and make appropriate adjustments.[8] However, the goal of our analysis is not to derive the true subjective probability distribution for the underlying event, but simply to examine whether partition dependence is reflected in the forecasts that are derived from these prices. Gürkaynak and Wolfers (2006), who report data covering the first 2½ years of these markets, argue that risk attitude plays a minor role in this context and show that "market-generated forecasts" (i.e. the expectations that follow from the market prices if the risk-neutral distribution is interpreted as the true subjective distribution) are more accurate than the "survey forecast" released by Money Market Services (MMS) on the Friday before a data release. (Their finding parallels the finding that the Iowa political market prices are typically more accurate than comparable political polls (Berg and Rietz (2006)). They also repeat many of

---

[7] See Filippov (2005, p. 47), and CME (2005, p. 5).

[8] This would be more complicated in our case though, as these adjustments would not be applied to the distribution of the entire market index but to specific assets that are more or less correlated with the market.

the tests that were provided in earlier literature, asking whether forecast errors are predictable, and did not find any significant predictability. They do not look, however, for partition dependence in these prices.

To do so, first note that each ED market presents a single partition of possible event outcomes to participants (the digital option outcome ranges). We can posit a simple econometric model to estimate the degree of partition-dependence: For each event category x, assume $f_{obs}(x) = (1 - \lambda) \cdot f_{true}(x) + \lambda \cdot f_{1/N}(x)$, where $f_{obs}(x)$ is the probability distribution implied by the observed market prices, $f_{true}(x)$ is the unobserved unbiased probability distribution, $f_{1/N}$ is a distribution assigning equal probability mass to each interval, and λ is the weight on the $1/N$ ignorance prior.

If each event were traded repeatedly, the empirical distribution of realized outcomes could be compared to the distribution of implied probabilities and the $1/N$ distribution to produce a sharp estimate of the apparent weight on the $1/N$ component. However, there is only one observation of implied probabilities for each point of time and each economic statistic. Therefore, we pool the data for the different points of time and across the different statistics.[9] We compute a mean forecast $M_{obs} = \mu(f_{obs})$ for each event by weighting the interval midpoints by the observed probabilities f_{obs} and determine a respective ignorance prior mean $M_{1/N}$ by assigning equal weight to each interval midpoint (and treating the extreme intervals as Gürkaynak and Wolfers (2006) do).

[9] To make the statistics comparable we follow Gürkaynak and Wolfers (2006) and normalize the data by the historical size of the forecast error.

From the definition of $f_{obs}(x)$ it follows that M_{obs} is a linear function $(1-\lambda)\cdot M_{true} + \lambda \cdot M_{1/N}$. Call the actual realization of the economic statistic X. A little algebra shows that the observed forecast error can be written as

$$M_{obs} - X = [M_{true} - X] - \lambda/(1-\lambda) \cdot [M_{obs} - M_{1/N}]. \qquad (1)$$

That is, the observed forecast error $M_{obs} - X$ has two components. The first component is the error term from a de-biased forecast based on $f_{true}(x)$ (which has an expectation of zero). The second component is a negatively-weighted term which reflects the degree of partition-dependence (through the weight λ). Intuitively, suppose the forecast from market data M_{obs} is above the equal-weight forecast $M_{1/N}$. If partition-dependence contaminates $f_{obs}(x)$, then $f_{obs}(x)$ is biased downward (toward $M_{1/N}$) relative to the de-biased ideal forecast $f_{true}(x)$ (which is an unbiased predictor of X). This downward bias means the forecast error $M_{obs} - X$ is likely to be negative. Thus, when $[M_{obs} - M_{1/N}]$ is positive $M_{obs} - X$ is likely to be negative (and vice versa). The negative correlation can be used to estimate $-\lambda/(1-\lambda)$ and infer an implied λ.

Table I summarizes the results of estimating regression (1) for markets for four different statistics.[10] This analysis reveals a negative correlation between forecast errors and the forecast – 1/N gap, which is consistent with bias toward a 1/N prior. Although one of the event domains (initial unemployment claims) shows no bias, the other three domains show substantial bias. When these results are pooled across all domains the results are highly significant. The coefficient estimated from pooling all the event domains, –.77 implies a value of the weight λ=.44 (because –.77 is an estimate of $-\lambda/(1-\lambda)$). Three of the four event domains imply values

[10] These are all the data available from Gürkaynak and Wolfers.

of λ from .39 to .56.

Table I

Results of Regressions of Forecast Errors. Regression equation is

$M_{obs} - X = [M_{true} - X] - \lambda / (1-\lambda) \cdot [M_{obs} - M_{1/N}]$, where M_{obs} and $M_{1/N}$ are mean forecasts calculated by weighting the interval midpoints by the observed probabilities f_{obs} and by assigning equal weight to each interval midpoint, respectively. M_{true} is the mean of the unbiased probability distribution. p-values are for test of the null hypothesis that the regression coefficient is different than zero, one-tailed.

	No. of Events	Regression Results			Implied Weight λ on 1/*N*	
		Coefficient −λ / (1 − λ)	*t*-Statistic	*p*-value (one-tailed)	λ Implied by Regression	Error-minimization
Initial jobless claims (IJC)	64	0.13	0.16	0.44	-0.15	0.06
Business confidence (ISM)	30	-0.64	-1.88	0.04	0.39	0.08
Non-farm payrolls (NFP)	33	-1.29	-1.53	0.07	0.56	0.56
Retail sales (excl. autos) (RSX)	26	-1.01	-1.32	0.10	0.50	0.50
All statistics pooled	153	-0.77	-2.60	0.01	0.44	0.39

A second analysis computes the mean absolute error between the actual realization of the economic statistic, and the λ-weighted combination of the forecast from the observed probability, M_{obs}, and the forecast $M_{1/N}$, for various weights λ. The values of λ that minimize the error from an λ-weighted combination are provided in the rightmost column of Table I. For two of the statistics (unemployment claims and business confidence) the weights are low, but positive. For the other two statistics the weights are close to .50. For all statistics pooled, the error-minimizing λ weight is .39.

These calculations suggest significant partition-dependence in all four ED market prices taken as a whole (a total sample of 153 separate markets), and a substantial degree of partition-dependence for two of the four markets when estimated separately.

III. Study 2: Evidence From Horse Racing Betting

The economic derivatives data provide initial evidence consistent with a bias toward 1/*N* for each of *N* distinct digital options that partition a numerical event space. Although this bias is consistent with our interpretation of a bias toward an ignorance-prior distribution, it is also possible that bettors generally overweight lower-probability events relative to higher-probability events, as predicted by prospect theory (Kahneman and Tversky (1979); Tversky and Kahneman (1992)) and consistent with both market-level and neural evidence (e.g., Barberis and Huang (2008); Hsu et al. (2009)). Bettors could also conceivably infer that the partition intervals are chosen by the market designers to equalize likely interval probabilities. Both overweighting of low probabilities and informational bias could therefore contribute to the observations of partition-dependence reported in Study 1.

Data from horse racing allows us to rule out the necessity of overweighting and informational bias to explain partition-dependence. In horse racing, "longshot" horses with low subjective probabilities of winning, as expressed by overall betting, have high potential payoff per dollar bet, because winners share the amount bet by losers in the pari-mutuel system. Many studies show that these longshots are overbet relative to their actual chances of winning, as measured in large samples (and favorites are relatively underbet). This tendency is called the "favorite-longshot bias" (see Snowberg and Wolfers, 2007). . Explanations that have been advanced include the ideas that the bias is due to market frictions which prevent equilibration (Axelrod et al. (2009)), interactions of information and timing of betting (Ottaviani and Sorensen (in press)) preferences for positive skewness bets, or to subjective weighting of probabilities (Snowberg and Wolfers (2007)).

A simple prediction of the $1/N$ bias account is that the longshot bias will be greater in races with fewer horses. Because N is lower in those races, $1/N$ is higher and so the subjective probabilities of longshots in those races should be higher if betters are adjusting insufficiently from the higher $1/N$ starting point. And in this case, the hypothesis that the $1/N$ bias comes from rational inference about probability due to the deliberate choice of partition (the number of horses entering) makes little sense.[11]

Detecting this effect, if it exists, requires comparing market probabilities (from betting odds) to actual winning frequencies which are estimated precisely both in different probability

[11] Trainers choose to enter horses essentially simultaneously a couple of weeks before a race is held when a "conditions book" is published listing the eligibility requirements (age of horses, gender, etc.) for different races. Usually there is some high limit on the number of horses (such as 14) but the number who enters is generally lower. Once entered, horses typically run. It would only be sensible to infer likely probabilities biased toward $1/N$ if the racetrack secretary could hand-pick groups of horses with highly comparable chances of winning, and then exclude horses who are too good or too poor. But the racing secretary has little control over which horses enter after the race conditions are published.

categories and in races with different numbers of horses. This requires a gigantic sample. Fortunately, Snowberg and Wolfers' (2007) have such a sample of all horse race starts in the United States from 1992 to 2001 (6.3 million horses). They use the post time odds to estimate the crowd's aggregate subjective probability. Horses with similar subjective probabilities are then placed into bins and the actual frequencies of those horses winning is estimated across the large sample. At our suggestion, Snowberg and Wolfers (2007, Appendix B) added an analysis of their data according to the number of horses entered in each race.

If the favorite-longshot bias simply reflects a subjective transformation of objective probability, there should be no difference in that transformation for races with different numbers of horses. However, if there is a bias toward $1/N$ then the subjective probability should be higher in races with fewer horses, and lower in races with many horses.

Figure 2A shows smoothed functions of perceived probability estimated from betting odds (on the y-axis) against "actual probability" (relative frequencies of actual wins in the different odds categories, on the x-axis). The differences are small, but are clearly ordered; the odds are higher for races with fewer horses in the range of actual probability from .005 to about .1 (when the curves bunch together).

Figure 2A.The relation between "actual" (sample relative frequencies) and perceived probability (estimated from aggregate market betting) for races with different numbers of horses. $1/N$ bias predicts the curve from races with fewer horses will lie above the curve from races with more horses. Figure reprinted from Snowberg and Wolfers (2007).

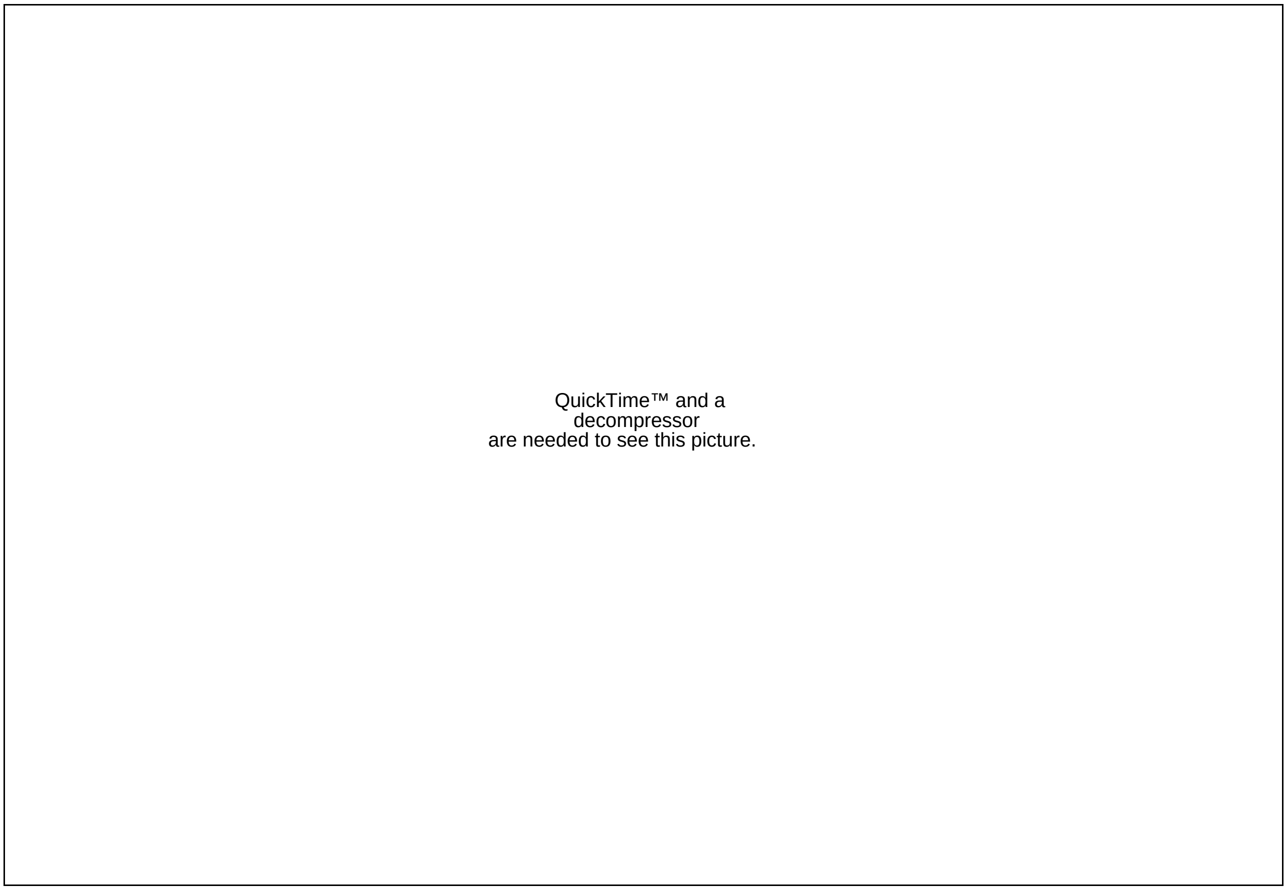

Figure 2B. Perceived probabilities are higher for races with fewer horses. This Figure is a magnification of Figure 2A for the range of actual probabilities between .02 and .05.

Figure 2B shows a zoomed-in closeup of Figure 2A for the range of actual probability (empirical frequencies) from .02 to .05. This closeup shows the separation more clearly; perceived probabilities in lower-N races are clearly higher than those in higher-N races (holding “actual” probability on the x-axis constant). Although this effect is not huge, it is orderly across changes in the race size N, and shows that the favorite-longshot bias cannot be attributed *entirely* to overweighting the chances of low-probability horses winning. Instead, the pattern is consistent with some reliance on the ignorance prior $1/N$ in races that range from N=6 to N=12 horses.

IV. Study 3: A Laboratory Experiment

Studies 1 and 2 provide evidence of partition dependence in naturally occurring binary option markets in which partitions were held constant for each market. To provide further evidence of this phenomenon we created experimental binary option markets in which we could independently manipulate the partitions presented to participants and observe the impact of this manipulation on equilibrium prices. The laboratory setting also allowed us the unique opportunity to elicit judged probabilities of all target events both before and after trading, which we could compare to equilibrium prices. Note that if risk-aversion distorted how subjective probabilities influenced market prices we would expect to observe differences in partition-dependence in judgment and prices; instead we find very similar effects on the two measures.

IV.A. Experimental Design

We conducted twelve two-hour experimental trading sessions in April 2007 with 16 traders in each session, divided into two self-contained groups (markets) with 8 traders in each group. Participants were N=192 undergraduate finance students (134 male, 58 female) from the University of Muenster (in Germany). The sessions spanned one week and took place in a computer lab where participants were separated from each other by dividers during trading periods. Instructions were read aloud to ensure that all information about the experiment was common knowledge (the full instructions can be found in Appendix B).

The essential part of the experiment consisted of several trading rounds in a set of three simple assets that are “all-or-nothing” betting contracts on the occurrence of specific future events. Three mutually exclusive and exhaustive events were defined for each market (e.g., the future closing of the German DAX stock market index). If an event occurred (did not occur), the asset corresponding to that event would pay the owner 100 cents (0 cents) after the uncertainty about the outcome was resolved. Thus, exactly one of the three assets would pay 100 cents while the other two assets would expire worthless.

By construction, since the events are mutually exclusive and exhaustive, a complete set of assets is certain to pay 100 cents. To allow arbitrage when the sum of state space-spanning prices is above or below 100 cents, and to create liquidity, subjects could trade a unit portfolio of all assets with the experimenter at any time for 100 cents.

Our experimental setting included three trading event domains: finance, weather, and sports. Figure 3 shows the event partitions for the German DAX stock index on the day two weeks after the experiments. In partition 1 (the low partition), the events are defined by the DAX index closing in the intervals [0, 7327.99], [7328, 7496.99], or 7497 and above (denoted [7497+]). In partition 2 (the high partition) the events are defined by the intervals [0, 7496.99], [7497, 7646.99] and [7647+]. The weather outcome refers to the maximum temperature in Muenster on May 31, approximately one month after the experiments. The sports outcome is the total number of goals scored by the teams of German “Bundesliga” (Federal League) on the final game day of the current soccer season, 3–4 weeks after the experiments. Subjects were grouped into high- and low-competence groups based on self-reported knowledge on soccer (though as we mention below, competence did not seem to affect prices or measured partition-dependence).

For each event domain, participants in different markets were randomly assigned to trade one of the two different partitions of the state space. *In order to eliminate the possibility that partition dependence is driven by information conveyed by the presented partition, we described both partitions to all participants in the instructions.* To create these partitions, each state space was initially divided into four disjoint and exhaustive intervals (I_1 to I_4), as illustrated in Figure 3. In each partition two of the adjacent intervals were combined to form a single asset. In partition 1 (the low partition) the upper two intervals were combined (forming an asset 3 with interval denoted $I_3 \cup I_4$), and the lower two intervals were traded separately (I_1 and I_2). In partition 2 (the high partition) the lower two intervals were combined (forming an asset 1 with interval denoted $I_1 \cup I_2$), and the upper two intervals were traded separately (I_3 and I_4). Both partitions therefore have three separate events. Note that by construction, asset 1 in partition 2 is a fusion of assets 1 and 2 in partition 1. Asset 3 in partition 1 is a fusion of assets 2 and 3 in partition 2.

For the weather and sports domains the interval boundaries were chosen arbitrarily based on historical outcomes. For the finance DAX event domain, the four intervals were created from historical data: Given the previous DAX closing price, and the recent short-term historical volatility of the DAX, we calculated the expected probability density function (PDF) for the DAX close two weeks in the future. Then we defined the interval boundaries such that each of

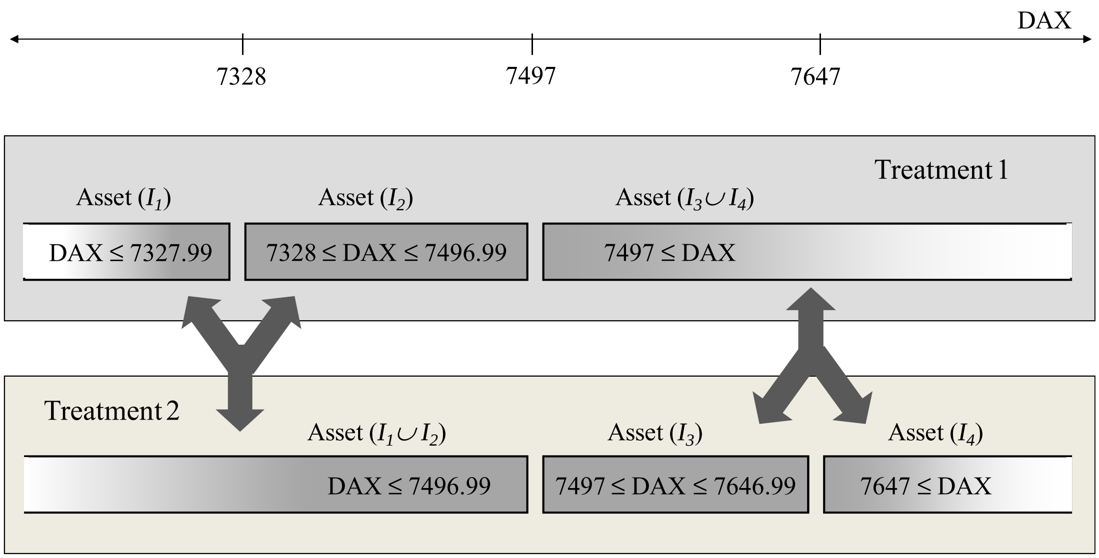

Figure 3. Construction of assets for the two DAX partitions in Study 3. The digital option will pay a fixed amount (1€) only if the DAX closes within the specified interval two weeks in the future.

the four intervals represented a particular percentile of the expected PDF.[12]

For each of the three event domains we ran successively two identical and independent trading rounds, resulting in six trading rounds per participant and experimental group, as depicted in Figure 4. Each trading round lasted ten minutes (with short breaks between rounds). The order in which participants traded assets from the three event domains varied for each experimental session and was perfectly counterbalanced (i.e., for each of six possible event domain orders there were two experimental sessions) to avoid any order effects. In each of the

[12] This was done to allow comparison of the experimental market prices to historical frequencies. We will not discuss this issue in this paper.

six trading rounds participants were initially endowed with a combination of assets (i.e., unit portfolios spanning the set of assets) and cash, with a total value of €20. Participants were told that their compensation would include their final cash and asset holdings for a trading round that would be randomly selected at the end of the experiment. They were told that this would be settled once the relevant uncertainty about the future outcome was resolved and asset payoffs (either 100 or 0 cents) became clear. There was no credit line and no short selling, though traders could use their available cash to buy unit portfolios and then sell the underlying assets. No explicit transaction costs were imposed for trading.

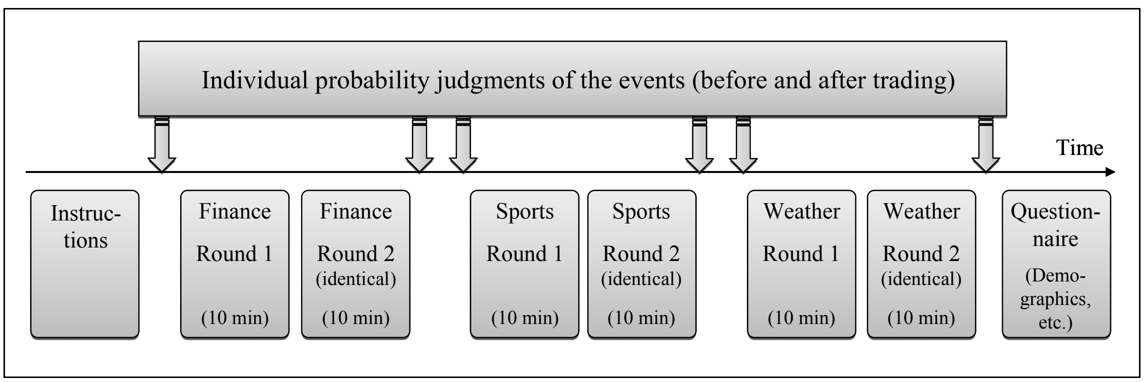

Figure 4. Example of the time course of an experimental session in Study 3.

The trading institution was a multi-unit continuous double auction (CDA) with a hidden order book. Subjects only saw the best bid and ask quotes for each asset (see Appendix C for a screenshot and further information on the trading software). Participants could submit bid and ask quotes for each asset simultaneously, so they could act as effective market makers. Trading took place only among the eight traders that were assigned to the same market; in particular they could not trade across markets with different partitions. During instruction and a practice trading round, participants were explicitly told how to exploit arbitrage opportunities by buying or selling unit portfolios with the experimenter for cash.

Before the first trading round for each event domain, and after the second (and final) trading round, participants were asked to provide individual probability judgments for each of the three events they traded. These judgments were not incentivized.[13] Some earlier studies have compared individual judgments of probabilities (as often elicited or inferred from psychology experiments) with probabilities expressed by market trades (see Camerer (1987); Camerer, Loewenstein, and Weber (1989); Ganguly, Kagel, and Moser (2000)). Like those studies, one question our method can address is whether bias is expressed by individual judgments, and whether it is moderated by the bundle of institutional and learning properties of markets (see also Fehr and Tyran (2005)).

IV.B. Results: Judged Probabilities

We first consider judged probabilities in the different event domains (which are required, by instruction, to sum to 1.0 across the exhaustive set of events). The notation $p(I_1)$ refers to the judged probability of unpacked interval I_1, and $p(I_1 \cup I_2)$ refers to the judged probability of the single packed interval which is the union of intervals I_1 and I_2 (as in partition 2 in Figure 3). The partition-dependence hypothesis is that $p(I_1) + p(I_2) > p(I_1 \cup I_2)$, and similarly for I_3 and I_4 .

[13] Participants rated their own competence in making probability judgments in each domain (from 1 to 7). We grouped participants by their self-rated general knowledge of soccer in order to see whether there would be group-level effects of competence on prices and accuracy, but we found no such effects.

Table II

Average pre-trading individual probability judgments. Probability judgments are surveyed before the first trading round of the finance, weather and sports event domains. PD difference is defined as $p(I_1) + p(I_2) - p(I_1 \cup I_2)$ and $p(I_3) + p(I_4) - p(I_3 \cup I_4)$, respectively.

Treatment	Pre-Trading Individual Judgment	Event Domain		
		Finance (N_1=N_2=96)	Sports (N_1=N_2=96)	Weather (N_1=N_2=95)
		Mean	Mean	Mean
1	$p(I_1)$	0.219	0.279	0.144
1	$p(I_2)$	0.497	0.398	0.333
	$p(I_1)+p(I_2)$	0.717	0.678	0.477
2	$p(I_1 \cup I_2)$	0.405	0.417	0.199
	PD difference	*0.312*	*0.261*	*0.278*
2	$p(I_3)$	0.397	0.378	0.349
2	$p(I_4)$	0.198	0.205	0.451
	$p(I_3)+p(I_4)$	0.595	0.583	0.801
1	$p(I_3 \cup I_4)$	0.283	0.322	0.523
	PD difference	*0.312*	*0.261*	*0.278*

Table II shows the average pre-trading individual probability judgments surveyed *before* the first trading round of the finance, weather and sports event domains (N=96 participants in each of the two partitions). The mean difference between summed probabilities of unpacked intervals and of the packed interval is .312, .261, and .278 for the finance, sports, and weather event domains, respectively. (All reported differences are statistically highly significant, based on a Kruskal-Wallis test ($p<.0001$)).

Note that if participants relied purely on ignorance priors of $1/N$ to assess probabilities of the three target events, the difference between the sum of the segregated-interval judgments and the packed-interval judgments should be one-third (=2/3 – 1/3). Thus, these judgments show a substantial effect of partition-dependence, consistent with the findings of earlier psychology experiments (e.g., Fox and Clemen (2005); See, Fox, and Rottenstreich (2006)).

IV.C. Results: Market Prices

Although mean judged probabilities exhibited pronounced partition-dependence, it is an open question whether this tendency would also be reflected in prices of assets following two 10-minute trading periods. In some respects markets are a dollar-activity-weighted opinion poll that also provide substantial time for reflection and opportunities for learning from others. We next turn to probabilities inferred from market prices.

The average number of trades and shares traded per market (except for unit portfolios) were 43 and 144. Trading was relatively continuous across the 10-minute trading period and similar across all three event domain (for more details on trading volume see Appendix C.)

Recall that buying or selling the unit portfolio could have been used rapidly to exploit arbitrage opportunities. A bid (ask) arbitrage opportunity existed if the bid (ask) quotes summed to more (less) than 100 cents. Actual arbitrage opportunities are typically very small in magnitude (less than 5 cents) and were exploited after a few seconds.[14]

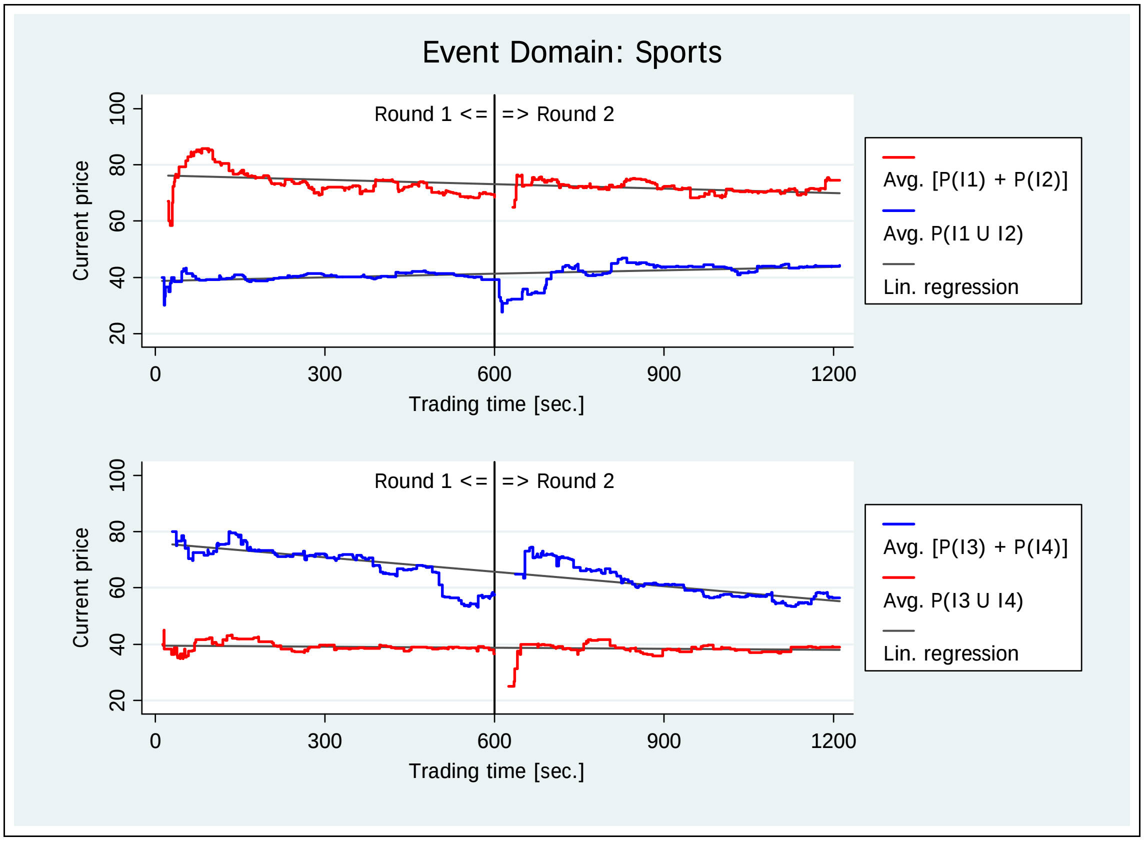

Figure 5. Development of price differences over time for the sports assets in Study 3. Prices are averaged over all twelve (identical) markets. Round 1 and 2 are identical replications.

Figure 5 shows the development of asset prices over time for the sports stimulus, averaging over all twelve (identical) markets. In both charts, the lower path shows the average price for the packed asset and the upper path shows the sum of prices for the corresponding

[14] For ask (bid) arbitrage, 87 (20) of 144 markets had no arbitrage opportunities at all and opportunities that emerged were exploited in a median of 10.83 (12.24) seconds after they first appeared. Appendix D provides more details.

unpacked assets. The gap between the paths shows the magnitude and persistence of partition-dependence. Estimates of linear regression lines for the price paths give a crude measure of convergence. The slope of the time trend for the difference in prices is .0094 (top) and .0159 (bottom) (see Appendix E for results of other markets). These estimates imply that there is some slow convergence. We will examine market prices over longer time periods in Study 4.

For each market, define the "equilibrium market price" $P^*(I_j)$ to be the quantity-weighted average of the last three trade prices (price at which trades were executed, not bids and asks) in the second trading round for the interval I_j asset. The partition-dependence hypothesis for prices is that $P^*(I_1) + P^*(I_2) > P^*(I_1 \cup I_2)$.

The experiment generates twelve equilibrium prices per asset for each partition. Table III displays mean prices (divided by 100 cents to make them comparable to probabilities) for the three assets of each partition. For comparison we also report average judgments after trading and the partition difference (PD) from pre-trading judgments (medians are in Appendix FI).

The difference between summed prices (rescaled) of unpacked assets and packed asset, averaged across the two partitions, is .267, .229, and .149 for finance, sports, and weather, respectively. (All reported differences are statistically highly significantly different from zero, using session-level differences, based on a Kruskal-Wallis test, $p<.0001$). The corresponding differences from pre-trading judged probabilities are .312, .261, and .278, and from post-trading judged probabilities are .257, .256, and .226. Partition-dependence is slightly less pronounced for market prices than probability judgments in the finance and sports domains, and it is much less pronounced for market prices in the weather domain.

Table III: Mean equilibrium prices (2nd round) and individual judgments (post-trading) for the finance, weather and sports event domains.

Treatment		Finance: Equilibrium Market Prices (Round 2)	Finance: Post-Trading Individual Judgment	Sports: Equilibrium Market Prices (Round 2)	Sports: Post-Trading Individual Judgment	Weather: Equilibrium Market Prices (Round 2)	Weather: Post-Trading Individual Judgment
		Mean Judged Probability/Equilibrium Prices					
1	I_1	0.152	0.205	0.230	0.252	0.048	0.116
1	I_2	0.561	0.494	0.490	0.432	0.256	0.307
	$I_1 + I_2$	0.713	0.699	0.720	0.684	0.303	0.422
2	$I_1 \cup I_2$	0.424	0.442	0.439	0.428	0.149	0.196
	PD difference	*0.289*	*0.257*	*0.281*	*0.256*	*0.154*	*0.226*
	(Pre-Trading Judgm.)		*(0.312)*		*(0.261)*		*(0.278)*
2	I_3	0.404	0.382	0.416	0.403	0.354	0.352
2	I_4	0.177	0.176	0.152	0.169	0.496	0.452
	$I_3 + I_4$	0.581	0.558	0.568	0.572	0.850	0.804
1	$I_3 \cup I_4$	0.336	0.301	0.391	0.316	0.707	0.578
	PD difference	*0.245*	*0.257*	*0.177*	*0.256*	*0.143*	*0.226*
	(Pre-Trading Judgm.)		*(0.312)*		*(0.261)*		*(0.278)*

However, conservative tests using session-level data only show a statistically significant reduction for the weather event domain.[15] Market experience also creates a slight reduction in

[15] Significance levels for Wilcoxon matched-pairs signed-rank tests comparing the size of partition-dependence for ex-ante judgments (i.e., $p(I_1) + p(I_2) - p(I_1 \cup I_2)$) with the effect size for market prices (i.e., $P^*(I_1) + P^*(I_2) - P^*(I_1 \cup I_2)$) are .58 for finance, 0.53 for sports, and 0.02 for weather (the respective levels for I_3 , I_4 , and $I_3 \cup I_4$ are 0.35, 0.10, and 0.02).

partition-dependence between pre-trading and post-trading judgments. A simple test statistic is whether judged probability of the packed interval increases after trading for each participant (note that the ignorance prior for the packed interval is 1/3, whereas this interval comprises two of four elementary events). The mean post-trading judgment is larger for 52.1% of participants, negative for 34.4%, and zero for the rest, significant by a sign test ($p<.01$). Thus, twenty minutes of trading appears to have a modest de-biasing effect on individual judgments, though a substantial degree of partition dependence remains after 20 minutes of trading.

IV.D. Alternative explanations

Study 3 provides further support for the notion that market prices for binary options are biased by reliance on ignorance prior probabilities, using a controlled laboratory environment in which partitions could be experimentally manipulated and explicit beliefs could be measured. However, two alternative explanations for these results merit further scrutiny: (1) partitions might convey legitimate information to participants about the expectations of the experimenter, and (2) prices might not reflect mean or median beliefs.

The first explanation is disabled by the fact that we explicitly told *all* subjects about *both* their own event partition *and* the different partition for the same event given to the other subjects. Thus, there might well be an informational effect but it should be the same in both groups (up to sampling error) and therefore cannot explain the between-group differences that are observed.

The second alternative explanation, that prices do not reflect average or median beliefs, requires more careful attention. For example, suppose prices reflect the most optimistic beliefs. Then the sum of prices for events 1 and 2 could be more than the price for event $1 \cup 2$. This explanation is closely related to a theoretical debate about what prediction market prices reveal about moments of the aggregate belief distribution. Manski (2006) noted that whether the market price is close to the mean or median of a distribution of beliefs in a population of traders depends on the traders' risk tastes and wealth constraints, and the correlation of those tastes and constraints with beliefs. (For example, in principle one very wealthy risk-neutral trader who is extremely optimistic could effectively set a market price, which reflects his extreme optimism, if others trade less because of wealth constraints or risk-aversion.) Wolfers and Zitzewitz (2007) replied to Manski by showing sufficient conditions under which the market price *is* the mean of the belief distribution, and showing that under other reasonable conditions prices are likely to be close to the mean belief (and see Gjerstad (2005)).

The fact that participants in our experiment provided individual probability judgments for events, *and* collectively created market prices for the *same* events, affords us a rare opportunity to shed some empirical light on this debate. We can directly compute the quantile of the individual belief distribution to which market prices correspond most closely.

We focus only on the second of the two consecutive markets and compare the quantiles of post-market event beliefs elicited from participants with the volume-weighted average of the last three trade prices. In each market there are eight belief judgments (one for each participant) and a single (averaged) price for all three events that were traded. The estimated quantile is the percentage of those eight beliefs that are below the market price. Separate quantile estimates can be computed for 12 separate experimental sessions for each of the three events.

The mean belief quantiles for prices of the finance, sports and weather events are .507, .512, and .483, respectively (all close to the median belief, the .5 quantile). Furthermore, in 54% of the 36 session-event comparisons, the estimated belief quantile is .375, .500, or .625 (i.e., the price is between the highest and lowest three out of eight beliefs about half the time). Thus, these data support the hypothesis that market prices approximately reflect the median belief.

At first glance, it would appear that prices could conceivably deviate from median beliefs due to heterogeneous optimism. The argument might proceed as follows. Suppose agents differ in their beliefs about the probabilities of events 1 and 2. In markets where separate claims are traded for events 1 and 2, those agents who believe 1 is likely will bid up event 1 shares. Agents who believe event 2 is likely will bid up event 2 shares. The prices for events 1 and 2, traded separately, could then sum to more than the price of a claim on the packed event $1 \cup 2$, creating the appearance of partition-dependence.

This argument starts to unravel when we consider that participants who are pessimistic about events can also sell shares, driving prices down. Because optimists and pessimists can always buy and sell unit portfolios (sets of claims that span all events), there is no reason to think that optimistic buying will exceed pessimistic selling. Unit portfolios were bought and sold, respectively, by 73% and 53% of individual traders. They bought and sold an average of 2.13 and 1.95 times, and traded an average of 9.56 and 7.65 portfolios when they did trade. The fact that unit portfolios are actively traded, and that prices reflect median beliefs, suggests there is no optimistic bias.

In sum, Study 3 was designed to see whether partition-dependence occurs and persists in short-run experimental markets, and to compare effects expressed in probability judgments with effects revealed by market trading prices. Both judgments and prices do show strong effects of partition-dependence across the three event domains that we used. Market prices show a much smaller effect in one of three event domains, and there is a small influence of market experience on post-trading individual judgments. The hypothesis that prices are systematically biased relative to beliefs is refuted because prices are close to the median quantile of measured beliefs.

V. Study 4: An NBA/FIFA Field Experiment

The modest effects of trading experience on partition-dependence seen in Study 3, following twenty minutes of trading, suggest the possibility that with much longer trading spans (and perhaps with more knowledgeable traders), partition-dependence might diminish or disappear. Study 4 tests this hypothesis in a field experiment that lasted several weeks.

V.A. Experimental Design

From April to July 2006 we conducted internet-based prediction markets for outcomes in the NBA basketball playoffs 2005/06 and the FIFA soccer World Cup 2006. Trading markets were open continuously for nine weeks for the NBA markets (April 20 through June 21, 2006) and seven weeks for the FIFA markets (May 24 through July 9, 2006), except for markets that closed when teams were eliminated. We recruited *N*=317 undergraduate finance students from the University of Muenster (in Germany) and *N*=139 students from an experimental economics laboratory e-mail list at UCLA (United States).

Contracts were all-or-nothing contingent claims on intervals defined by the total number of victories for a particular NBA team during the playoffs, and the total number of goals scored by a particular national team during the entire World Cup tournament (excluding shoot-out goals). As in Study 3, for each event domain there were two partitions that combine sub-events differently. For example, in the NBA markets the first partition packed the victory intervals [4, 7] and [8, 11] into a single interval [4, 11], and unpacked the interval [12, 16] into two components of [12, 15] and [16]. In order to control for the possibility that partitions would provide information, all participants were explicitly informed of both partitions and that they would be randomly assigned to trade only one. Every participant could trade assets based on four teams—called "team markets"—using the same partition for each of the four teams.

The NBA intervals correspond to the number of victories needed by a team to advance across the four playoffs rounds, so bets on the various win-total events are equivalent to betting that teams will lose in the first round, the second round, and so forth. The intervals for the number of goals in the FIFA soccer World Cup were not structured to correspond to advancement across rounds, but were chosen such that they all appeared plausible based on the results from the three previous World Cups.

The experimental protocol was similar to Study 3, but was adapted for the Web (see Appendix G). Participants were instructed by e-mail about the composition of assets and markets (including the partitions of assets they could trade, and the alternative partition), how to use the trading system, and some details concerning the NBA playoffs and the FIFA World Cup. They also had Internet access to a homepage with study details, FAQs, and a practice market. The market was open continuously. As in Study 3, the trading mechanism was a multi-unit continuous double auction (CDA) with a hidden order book, so that participants could see only the best bid and ask quotes and the most recent trade price for each asset. Traders could submit bid and ask quotes for each asset simultaneously, acting as market makers. Trading took place only among the twenty participants eligible to trade in each market. There was no credit line or short selling opportunity, except for purchases of the unit portfolio from the experimenter.

Participants were initially endowed with different combinations of cash and unit portfolios totaling €10 in each "team market" of the NBA playoffs and were endowed again in the World Cup markets. At the end of the experiment we randomly drew one out of the four teams for each experimental group to compensate participants based on the sum of the actual asset values in their final portfolio and their cash balance, for an expected payment of €20 (€10 for playoffs and €10 for World Cup markets) per person. We also collected questionnaire data from all participants before and after trading, including individual judgments of the probabilities that outcomes would fall into intervals corresponding to the assets they traded.

It is important to stress that participants only had direct access to their own four-team NBA and World Cup markets. They could not directly observe market data (e.g., prices or quotes) from other groups trading different partitions. For each part of the study—NBA playoffs and soccer World Cup—assets on 16 teams were traded. We can thus compare trading prices from two different partitions for each of 16 teams in each of the two event domains. Due to the large number of participants that were recruited, we could fill two identical experimental settings ("clones") with German students and one identical setting with U.S. students.

V.B. Results

The analysis of results is similar to the analysis from laboratory Study 3 in Section IV. First, we test for partition-dependence in individual probability judgments elicited before the beginning of trade. Next, we look for partition-dependence in bids, asks, and trading prices in the markets. We also test whether probabilities for the lowest-outcome event ([0, 3]), which is the same interval in both partitions, differs across markets with different partitions. The present account predicts that these probabilities will not differ significantly because the ignorance prior probability is 1/4 for this event in both partitions; any difference should reflect sampling error.

Because some participants did not submit probability judgments before the first playoff game was played (we exclude their judgments), there are N=302 (199 German and 103 U.S.) sets of judgments for the NBA teams. For the World Cup, N=263 judgments were submitted by German participants before the opening game was played.[16]

Table IV shows the median differences in judgments (×100) for the two different partitions. As expected, the differences in the commonly partitioned event [0, 3] between partitions (in the first column) are close to zero and not statistically significant. The differences between the summed probabilities of the unpacked events and the probability for the corresponding packed event are positive, almost always highly significantly different from zero, and are comparable in magnitude to the effects reported earlier (approximately a .20 increase in

[16] No results are reported for U.S. participants because there was a large dropout rate for World Cup markets among these participants.

probability when the interval is unpacked).

We next examine in some more detail activity in the two most liquid experimental markets for each event (as measured by the overall number of trades). In the NBA playoffs, the

Table IV: Partition-dependence in pre-trading judgments (median $P(X)+P(\underline{Y})-P(X\cup Y)$) The Table presents differences in medians for interval I_0 and differences in medians for the sum of unpacked events and the packed event per NBA and FIFA team. Number of subjects giving judgments is 30-49 for (1,2) and 30-51 for (3,4). Each subject (N=302 and N=263 resp.) provided judgments for four different teams resulting in 1,208 (1,052) judgments in total. *, **, *** indicate significance at the 10%, 5%, and 1% level (two-tailed) based on a Kruskal-Wallis test for each team.

Team	Δ_{Median}, Whole Population (N=302x4), German and U.S. Subjects (Pooled)			
	Event I_0 Equality	$p(I_1)+p(I_2)-p(I_1\cup I_2)$	$p(I_3)+p(I_4)-p(I_3\cup I_4)$	N_1/N_2
CHI	15.0	19.0 **	22.5 ***	37/36
CLE	2.0	20.0 ***	15.0 ***	37/36
DAL	0.0	24.5 ***	20.0 ***	40/37
DEN	-2.0	17.5 ***	18.0 ***	41/34
DET	2.5	25.0 ***	36.5 ***	41/34
IND	-10.0	25.0 ***	12.0 ***	41/34
LAC	5.0	5.0	10.0 ***	40/37
LAL	-5.0	22.5 ***	10.0 **	40/37
MEM	15.0	25.0 ***	23.0 ***	37/36
MIA	-10.0	20.0 ***	15.0 ***	40/37
MIL	-5.0	20.0 ***	10.0 **	40/37
NJN	-10.0 *	30.0 ***	20.0 ***	40/37
PHX	0.0	30.0 ***	27.5 ***	37/36
SAC	-12.5	25.0 ***	2.0 *	41/34
SAS	0.0	30.0 ***	40.0 ***	40/37
WAS	-10.0	10.0 ***	10.0	40/37

Team	Δ_{Median}, Whole Population (N=263x4), German Subjects			
	Event I_0 Equality	$p(I_1) + p(I_2) - p(I_1 \cup I_2)$	$p(I_3) + p(I_4) - p(I_3 \cup I_4)$	N_1/N_2
ARG	-1.5	37.5 ***	40.0 ***	30/34
AUS	-5.0	0.0	5.0	33/34
BRA	0.0	22.0 ***	20.0 ***	33/34
CIV	6.5	10.0 *	8.5 ***	30/34
CRC	0.0	15.0 *	9.0 ***	35/32
CRO	10.0 **	10.0 **	20.0 ***	33/34
CZE	8.0 *	35.0 ***	37.5 ***	35/30
ECU	10.0	10.0	9.0 ***	35/32
GER	0.0	35.0 ***	37.5 ***	35/32
GHA	-2.5	12.5 **	5.0 *	35/30
ITA	4.0	25.0 ***	27.5 ***	35/30
JPN	5.0 **	0.0	6.0 ***	33/34
NED	0.0	32.5 ***	30.0 ***	30/34
POL	5.0	30.0 ***	30.0 ***	35/32
SCG	10.0	0.5	14.0 ***	30/34
USA	0.0	15.0 ***	7.0 ***	35/30

most liquid markets were the Dallas Mavericks (DAL) and Miami Heat (MIA) (partly because these teams became the two finalists, their assets were traded for the longest span of time). We show the graphs for Dallas in the main text and present the corresponding graphs for Miami moved to Appendix H.

For DAL, there were 129 and 119 trades in partitions 1 and 2, respectively. For MIA, there were 101 trades and 102 trades in partitions 1 and 2, respectively.

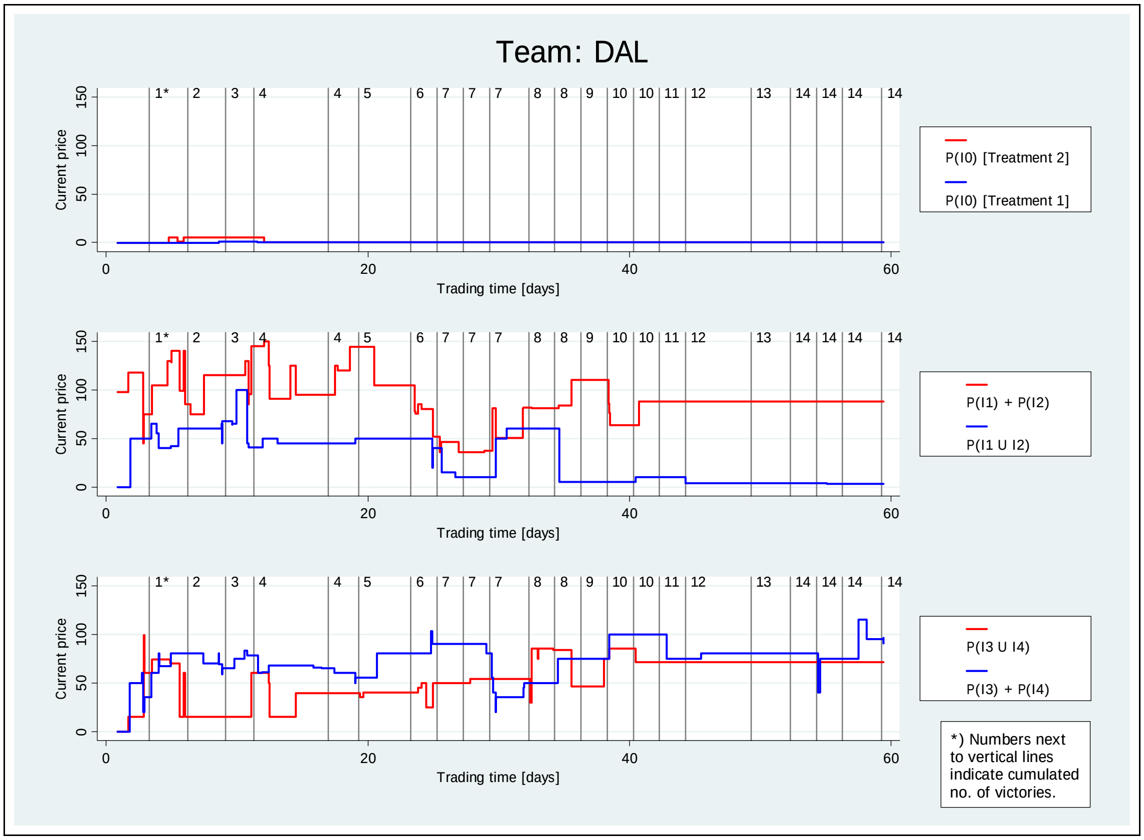

Figure 6. Price chart (Dallas Mavericks, DAL).

For the Dallas market, Figure 6 shows the most recent market price (in cents) plotted against the number of days since trading began, for assets corresponding to different partitions. Because there were only about two trades per day across all assets, there are many horizontal flat spots in the time series, which indicate the level of the last trade price when there was no current trading. Vertical lines indicate the beginning of a basketball playoff game. The numbers next to vertical lines at the top of the charts indicate the number of cumulated wins after each game. For example, Figure 6 shows that DAL won its first four games, lost its next game (status remains 4), won its sixth game and so on. The upper panels compare prices for the asset 0 interval I_0 for partition 1 (blue line) and partition 2 (red line) (these prices are low, usually zero, since DAL and MIA were expected to win many games, and prices do not differ between the two partitions for each team).

In the second panel a blue line indicates the current market price of the packed asset [4, 11] of partition 1 and a red line shows the sum of the market prices for unpacked assets [4, 7] and [8, 11] of partition 2. The fact that the red line lies above the blue line reflects partition-dependence.

The third panel shows a red line for the current market price of packed asset [12, 16] of partition 2 and the blue line represents the sum of the market prices for the unpacked assets [12, 15] and [16] of partition 1. The fact that the blue line is above the red line, for most of the time, indicates partition-dependence.

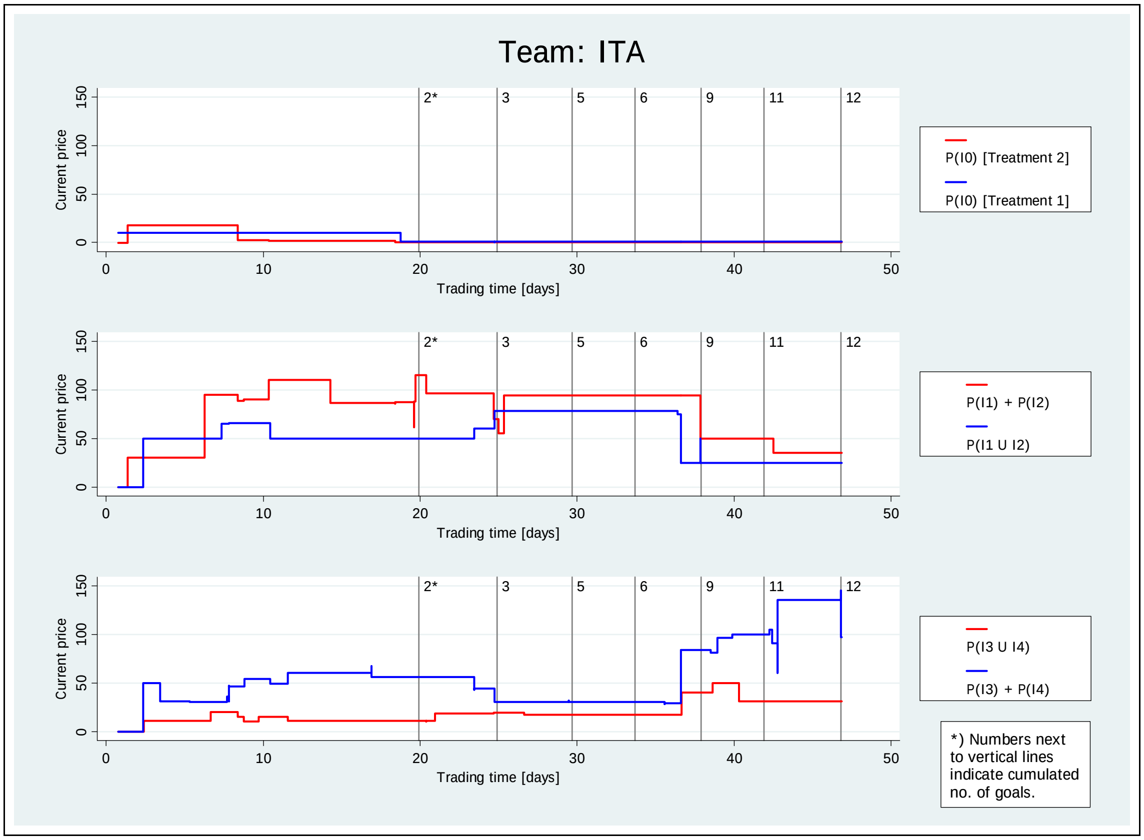

Figure 7. Price chart (Italy, ITA).

For the World Cup, Figure 7 shows the price chart for Italy, the eventual World Cup champion. This chart was chosen for presentation because the depicted events were the most actively traded, with 65 trades for both partitions (the chart for the second most actively traded team, Germany, is presented in Appendix H). Both markets show persistent partition-dependence of recent prices in the predicted direction, i.e., the red line is above blue in the middle panel and vice versa in the lower panel.

Because prices are constantly changing in response to new information over the several weeks of these tournaments, the ”equilibrium market prices” for a static event toward the end of trading cannot easily be used to determine the degree of partition-dependence revealed by prices (as in Study 3). Therefore, we measured partition-dependence in two more nuanced ways. Both methods measure the hypothetical “pseudo-arbitrage” available by comparing the summed prices for the two unpacked-interval assets (traded in one market) with the price for the equivalent packed-interval asset (traded in a different market). Note that these calculations are not true arbitrage opportunities because traders cannot actually trade in the markets with different partitions; they simply provide an economically interesting measure of the partition-dependent gap in prices between the two markets.

Our first method calculates the time-weighted pseudo-arbitrage profits from selling the unpacked-interval assets in one market and buying the equivalent packed-interval asset in the other market, at available bid and ask prices. Because trading is often quite thin, there are long stretches of time when bids and asks are not available on all assets and the measured partition-dependence is zero. For a more detailed description of this method, see Appendix I.

In our second method, “interpolated-price hypothetical arbitrage,” the most recent and subsequent prices are used to *interpolate* a continuous trade price. The arbitrage is *hypothetical* because it summarizes price differences in separate markets and assumes that trades can take place when there are no standing bids or asks. Because participants cannot trade across markets, they cannot directly act on these pseudo-arbitrage opportunities. This method is conservative because it assumes that hypothetical trades would only be executed at the worst of the observable prices. That is, even when there are no asks available, a trader seeking to buy is presumed to be able to always execute a trade, but only at the higher of the last previous trade price and the next future price. (Similarly, we assume that sell orders are executable at the lower of the last price and the next future price.) This method assumes, counterfactually, that there is a continuous flow of prices at which trades could occur (because there is latent willingness to trade that is not revealed by posted bid and asks). Note that basketball games and soccer matches are occurring during the continuous flow of trading, so using the worst of the last and next prices often means that traders are (hypothetically) betting against unfavorable public information, which adds to the conservatism of this measure.

In formal notation, the interpolated-price hypothetical arbitrage profit at time t is:

(1.1) $\min[P_{t-r}(I_1), P_{t+n}(I_1)] + \min[P_{t-r}(I_2), P_{t+n}(I_2)] - \max[P_{t-r}(I_1 \cup I_2), P_{t+n}(I_1 \cup I_2)]$

(1.2) $\min[P_{t-r}(I_3), P_{t+n}(I_3)] + \min[P_{t-r}(I_4), P_{t+n}(I_4)] - \max[P_{t-r}(I_3 \cup I_4), P_{t+n}(I_3 \cup I_4)]$

for the intervals I_1 and I_2 (intervals I_3 and I_4), where $P_s(I_j)$ is the trade price at time s for interval j, and $t–r$ and $t+n$ are the times of the most recent and next trades.

To illustrate further, suppose the trade prices of a thinly-traded asset are 42 at day 20 and 48 at day 25, and there are no trades between those dates. If a trader is *buying* the asset, we assume he could buy it at the *higher* price of 48 during days 20 to 25 even though there is no trading during those days (and even if there are no posted bids or asks). If he is *selling* the asset, we assume he could sell it for the *lower* price of 42 during days 20 to 25.

Figure 8 shows the interpolated-price hypothetical arbitrage profit over time for Dallas (DAL) in the NBA event domain. The blue line in the first panel shows the hypothetical arbitrage profits from selling at the minimum interpolated prices for unpacked intervals I_1 and I_2, and buying at the maximum interpolated price for interval $I_1 \cup I_2$ (as defined in (1.1)). The red line, by contrast, shows the hypothetical profits from the reverse arbitrage strategy, i.e., selling at the minimum interpolated price for the packed interval $I_1 \cup I_2$, and buying at the maximum interpolated prices for unpacked intervals I_1 and I_2. Because this "profit" can be positive or negative, the second panel shows the value of this hypothetical profit when it is above zero (i.e., the profit conditional on it being positive). Panels three and four show the same calculations for the assets based on unpacked intervals I_3 and I_4 and the packed interval $I_3 \cup I_4$ (as defined in (1.2)). The hypothetical profits from selling the unpacked-interval assets and buying the packed-interval asset (blue lines), depicted in panels two and four, are often positive and large. Figure 9 shows the corresponding data from trades on Italy (ITA) in the World Cup. The results are similar. Note that if there were reverse partition-dependence (the packed-interval asset price is higher) the red lines in Figures 8 and 9 would be above zero, but this is never the case. The fact that there is virtually no reverse effect shows that partition-dependence in the expected direction (as indicated by the blue curves) is systematically positive and not merely the result of random error.

To measure the daily average interpolated-price hypothetical pseudo-arbitrage profit for each team, we calculated the area under the blue and red curves in the second and fourth panels of Figures 8 and 9, and divided by the total trading time (in days). These statistics are provided for each team and interval in Table V. Note the figures are always positive because if the return to pseudo-arbitrage is negative, we assume the trade would not be made (i.e., only positive profits are averaged).

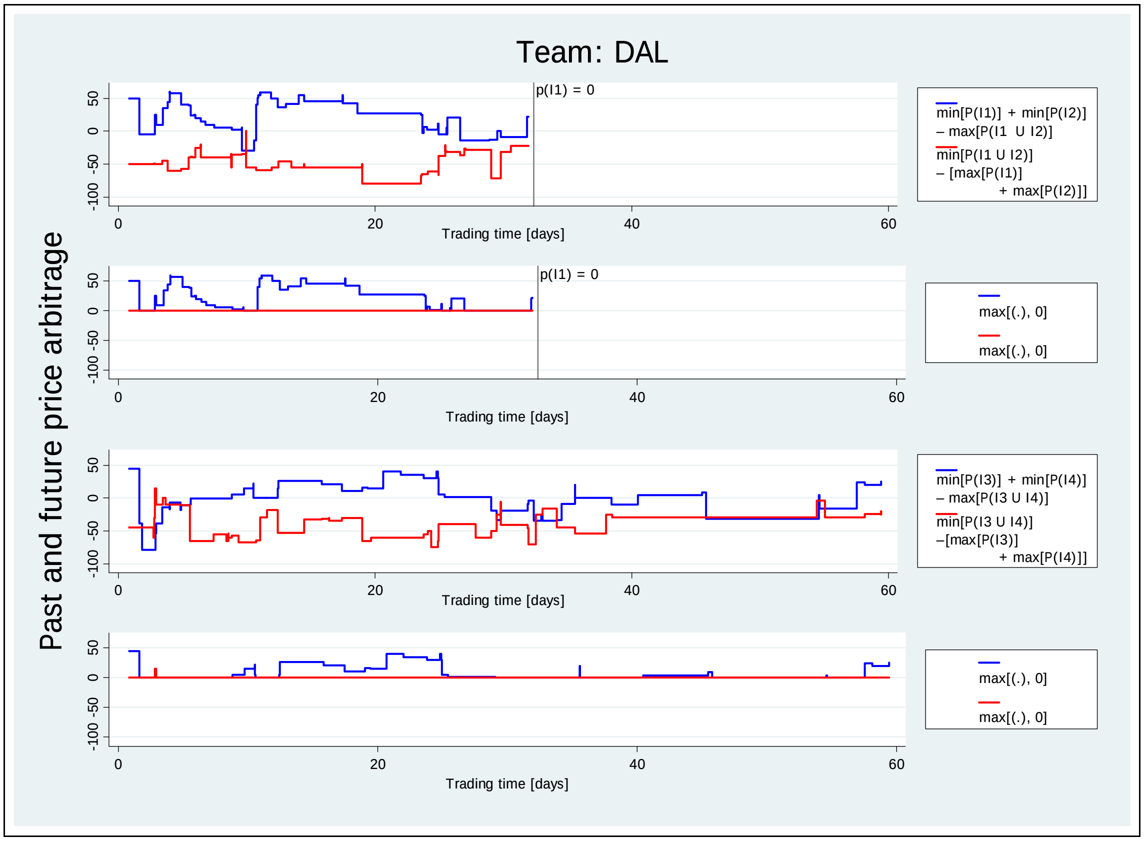

Figure 8. Interpolated-price hypothetical arbitrage (Dallas Mavericks, DAL).

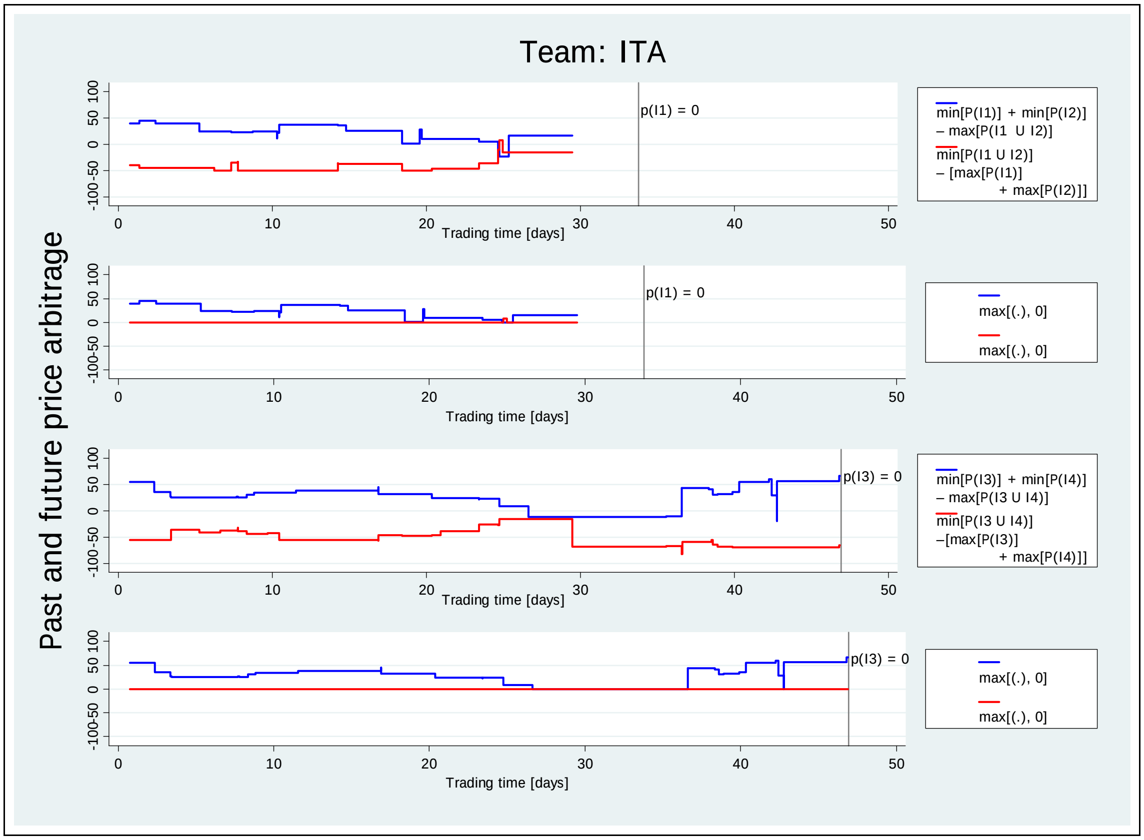

Figure 9. Interpolated-price Hypothetical Arbitrage (Italy, ITA).

The average per-day hypothetical profit from exploiting bias toward the ignorance prior by selling the unpacked-interval assets and buying the packed-interval asset, is higher than for the reverse strategy (buying unpacked and selling packed) for 21 out of 32 teams for intervals I_1 and I_2, and for 27 out of 32 for intervals I_3 and I_4 (significant by sign test at $p<0.1$ and $p<.001$ respectively). The median per-day pseudo-arbitrage profit exploiting such partition-dependence, across all teams and sports, is 5.61 for intervals I_1 and I_2 and 6.41 for intervals I_3 and I_4; the average of this median across intervals is 6.01.

As noted above, we also computed the hypothetical arbitrage profit from executing trades only when bids and asks are available on all assets (see Appendix I). These strategies cannot be executed most of the time due to thin markets. As a result, the daily average hypothetical profits are low. The median and mean daily profit averaged across teams by exploiting bias toward the ignorance prior are .03 and .35 for intervals I_1 and I_2, and .07 and 1.50 for intervals I_3 and I_4; the average across intervals is .92. Profits are higher from exploiting this form of partition-dependence compared to reverse partition-dependence in 38 of 46 teams ($z=5.88$, $p<.001$ by a sign test, excluding zero-profit teams).

The hypothetical profits from these two measures could be considered lower and upper bounds on the financial magnitude of partition-dependence. Profitability as measured using simultaneously-available bids and asks provides a lower bound because there are so many stretches of time with incomplete bids and asks. Profitability as measured by the interpolated-price method provides an upper bound because it artificially liquefies the market by assuming there is always a latent trade waiting to occur at the right price (though it is still conservative

Table V

Per-day profitability of interpolated-price pseudo-arbitrage strategies

Team	Low Intervals: Arbitrage PD (Sell I_1, I_2, Buy $I_1 \cup I_2$)	Low Intervals: Arbitrage Reverse PD (Buy I_1, I_2, Sell $I_1 \cup I_2$)	High Intervals: Arbitrage PD (Sell I_3, I_4, Buy $I_3 \cup I_4$)	High Intervals: Arbitrage Reverse (Buy I_3, I_4, Sell $I_3 \cup I_4$)
NBA Playoffs teams				
CHI	1.24	0.85	2.04	0.00
CLE	3.71	0.72	13.48	0.00
DAL	22.39	0.00	7.38	0.02
DEN	3.49	5.48	2.43	1.54
DET	16.10	2.70	0.48	7.50
IND	7.98	0.19	0.00	0.00
LAC	9.41	0.00	7.76	0.00
LAL	8.56	0.00	8.29	0.00
MEM	0.16	2.14	11.51	0.00
MIA	27.51	0.00	13.52	2.03
MIL	0.00	2.97	2.75	0.00
NJN	5.87	1.69	24.03	0.00
PHX	14.53	0.23	8.17	0.00
SAC	0.38	0.90	0.15	0.58
SAS	28.59	0.00	16.07	0.33
WAS	8.27	0.00	5.23	0.00
FIFA World Cup teams				
ARG	0.53	2.55	13.05	0.00
AUS	2.33	1.07	7.24	0.36
BRA	0.00	5.63	0.57	3.50
CIV	0.04	4.53	0.08	0.72
CRC	1.01	6.30	0.01	0.24
CRO	0.66	2.24	1.22	0.00
CZE	21.85	0.00	29.21	0.00
ECU	12.24	0.05	9.74	0.01
GER	11.87	0.04	9.35	0.20
GHA	0.00	23.55	1.98	0.00
ITA	22.84	0.06	27.66	0.00
JPN	8.11	0.51	1.68	0.00
NED	10.73	0.00	6.54	2.14
POL	0.71	0.19	0.81	0.46
SCG	0.60	0.20	0.25	0.00
USA	5.45	8.08	1.35	0.00

because it assumes trades would be executed at the worst of the most recent and next future prices).

In sum, Study 4 documents the persistence of systematic partition-dependent pricing in a field experiment in which self-selected participants trade assets whose value depend on the outcomes of events in which they take great interest—the NBA playoffs and the World Cup tournament—and for which trading lasts for several weeks. Probability judgments made before trading began exhibit partition-dependence that is similar in magnitude to previous psychological studies—e.g., the sum of unpacked intervals (e.g., [4, 7] NBA wins plus [8, 11] wins) is judged to about 20 per cent larger in absolute probability than the packed interval [4, 11].

The partition-dependence revealed by actual prices of event assets can be roughly bounded by two different methods. Using the possibility of hypothetical cross-market arbitrage at available bids and asks yields an average daily profit of about 1% (largely because there are long stretches of time where there is not a simultaneous set of bids for the unpacked assets and an ask for the packed asset). Using an interpolated-price procedure, which assumes that trades could take place continuously, but only at the worst price from the last trade and the next future trade, yields hypothetical arbitrage profits around 6%. The two measures represent likely lower and upper bounds on the practical profitability from exploiting partition-dependence, and therefore bound its likely economic magnitude in markets like these. For both measures and for a large majority of team markets, these potential profits are much larger than profits from executing the opposite strategy (buying unpacked intervals and selling the equivalent packed interval), indicating that partition-dependence is a systematic bias rather than random error.

Finally, we observe a tendency for traders to both buy and sell unit portfolios (sets of claims that span the partition) at a rate comparable to that which we observed in the laboratory study: About 77% and 44% of traders buy and sell unit portfolios at least once, respectively. Conditional on trading, the average number of purchases and sales per participant are 2.29 and 1.56, respectively, and the average volume is 5.37 and 4.63 units per trade. Thus, selling portfolios is less frequent and a little less vigorous than buying, but there is substantial amount of both sorts of trading.

VI. CONCLUSION

A series of psychology experiments beginning in the late 1970s showed that "unpacking" a single category or interval into two or more component intervals that are logically equivalent increases the total expressed probability, a phenomenon called "partition-dependence". The present paper asks to what extent market forces mitigate partition-dependence observed in individual judgments. The question is important for financial markets since some assets (such as credit default swaps) are claims based on events defined using a particular partition. It is possible that partition-dependence influences the prices of these assets.

We report evidence from four data sets, using naturally-occurring data, laboratory and field experiments. For each of these data sets, there are possible explanations for apparent partition-dependence in prices which do not explain the other data sets. Our goal is to provide a single explanation that can account for the common patterns across all four data sets.

Study 1 uses naturally-occurring data from prediction markets for macroeconomic statistics with a single partition in each market. Econometric techniques suggest that probabilities implied by prediction market prices are a convex combination of a partition-independent probability and an "ignorance prior" probability ($1/N$ for each of N intervals), with a weight λ for the prior distribution of around .50 in two of four markets and .05–.10 in the other two markets. Study 2 reports a reanalysis of subjective probabilities implied by betting odds (a kind of market price) in all American horse races from 1992-2001. These data show a typical "favorite-longshot" bias, a tendency for subjective probabilities of unlikely winners (longshots) to be higher than objective probabilities estimated from the frequencies of actual wins. Comparing races with different numbers of horses, we find that subjective probabilities are higher in races with fewer horses (i.e.,when $1/N$ is higher), consistent with a bias toward the ignorance prior.

It is often difficult to rule out alternative explanations in field settings like the derivatives and horse race markets. For example, it is possible that investors in the derivative markets infered that offered numerical intervals reflect the market designer's intent to choose intervals that are equally likely (in which case a bias toward $1/N$ is a rational inference rather than a bias). Likewise, one might try to argue that the structure of information and microstructure of parimutuel betting in the horse race markets could conceivably be part of the explanation for the differences in races with few and many horses. Therefore, it is useful to conduct experiments in which we can control for these alternative explanations.

Our laboratory and field experiments show that the partition-dependence bias is also evident in competitive markets. The size of the bias sometimes shrinks with trading experience, but it does not disappear even with substantial variations in the nature of events being evaluated, self-selection and expertise of participants, and the length of time that markets are open. Study 3 demonstrated pronounced partition-dependence under standard lab conditions for short-run (20 minute) markets. Furthermore, although market experience mitigates partition-dependence it does not eliminate the bias. Unpacking one event interval (of three) into two component intervals (out of three) increases its judged probability by about .25. Study 4 documents similar partition-dependence in longer-run markets (several weeks) on events in which our participants took great interest (NBA playoffs and FIFA soccer World Cup). Unpacking event intervals led to hypothetical arbitrage profits of 1–6%.

Note that if markets were opened with two different partitions, and traders were allowed to trade in both markets, there is little doubt that arbitrage would erase obvious differences. That is, if the price of events I_1 and I_2 were both higher than the identical packed event $I_1 \cup I_2$, then arbitrage would bring the sum of event I_1 and I_2 prices into line with the price of $I_1 \cup I_2$. If different partitions were actually traded on the same events it is unlikely that partition-dependence would be directly observed in equilibrium.

This said, in practice there is no reason why two different partitions would ever be created and traded simultaneously (unless competing exchanges did so deliberately or as part of an experiment). Therefore, the relevant question is whether revealed prices for the unique partitions that are actually traded could conceivably be a combination of highly accurate prices for each interval and a $1/N$ price for each interval when a single partition is traded. The analysis of the 153 separate economic derivatives markets reported in Study 1 suggest an affirmative answer, and the field and lab data show similar tendencies in the same direction (though to varying degrees).

The combination of methods used in the four studies provides converging evidence for our interpretation that markets are affected by a pervasive bias toward an ignorance prior probability distribution. No alternative account can explain the results of *all* four studies. We will revisit some candidate explanations now.

The apparent bias cannot be attributed entirely to information conveyed by the choice of partition because the subjects in the lab and field experiment Studies 3-4 were explicitly told about *both* partitions and randomly assigned to one. Thus, any information conveyed by the choice of partition should have affected both markets equally. The apparent bias cannot attributed entirely to the nature of trading institutions since field Studies 1 and 2 used dynamic parimutuel auctions and experimental Studies 3 and 4 used double auctions. The apparent bias cannot be attributed entirely to naïveté of subjects, since there was self-selection of active traders in field markets (Studies 1 and 2) and in the experimental market of Study 3 (NBA and World Cup). The apparent bias cannot be attributed entirely to simple overweighting of low probabilities (say, below .10) because many of the sub-event probabilities are not low. Moreover, the horse racing Study 2 controls for probability and finds a systematic bias toward $1/N$ across races with different field sizes N.

We concede that these alternative explanations may certainly contribute in different ways, in particular markets, to prices which exhibit a statistical bias toward $1/N$. However, none of these explanations can explain the results across *all* studies. The single explanation that *can* account for results across all studies is the simple hypothesis that subjective probabilities are biased toward the ignorance prior distribution implied by the set of events into which the state space happens to be partitioned, and that these judgments are also reflected by market prices in a variety of field and experimental markets.

More generally, these studies suggest two general themes in thinking about the implication of psychology for economics. One theme is that people do not always compensate for the effect of event salience. For example, Morwitz, Johnson, and Schmittlein (1993) report that simply asking people whether they will buy a car in the next year increases their tendency to buy a car—by 50%. Similarly, unpacking an interval into the two components increases attention to those components and seems to increase implied probability. Because there is typically no canonical or normative way to partition a continuous variable, the particular intervals into which a variable is divided can inexorably influence perceptions of the likely value of that variable. Gennaioli and Shleifer (2009) offer a model of judgment under selected (primed), limited memory that embodies this type of phenomenon (see also Ahn and Elgin (2009)).

The second theme is that the extent to which individual psychological processes influence market prices will depend on the processes and on the markets. As Camerer and Fehr (2006) note, in some market institutions the biases of a small number of traders will be amplified by strategic complementarity, and in other institutions biases are reduced because unbiased traders can profit by extinguishing biases created by other traders. The partition-dependence discussed here seems to exist to various extents in different experimental and field markets for different assets, with different microstructure, but its robustness and persistence over time demands attention in future studies.

REFERENCES

Ahn, David S., and Haluk Ergin, 2009, Framing Contingencies, *Econometrica*, forthcoming.

Axelrod, Boris S., Ben J. Kulick, Charles R. Plott, and Kevin A. Roust, 2009, The design of improved parimutuel-type information aggregation mechanisms: Inaccuracies and the long-shot bias as disequilibrium phenomena, Journal of Economic Behavior & Organization 69, 170-181.

Barberis, Nicholas and Ming Huang, 2008, Stocks as Lotteries: The Implications of Probability Weighting for Security Prices, American Economic Review98, 2066-2100.

Bardolet, David, Craig R. Fox, and Dan Lovallo, 2009, Naïve Diversification and Partition Dependence in Capital Allocation Decisions, UCLA Anderson School Working paper.

Beber, Alessandro, and Michael W. Brandt, 2009, Resolving macroeconomic uncertainty in stock and bond markets, Review of Finance 13, 2009, 1-45.

Benartzi, Shlomo, and Richard H. Thaler, 2001, Naive Diversification Strategies in Defined Contribution Saving Plans, American Economic Review 91, 79-98.

Berg, Joyce E., and Thomas A. Rietz, 2003, Prediction markets as decision support systems, Information Systems Frontiers 5, 79-93.

Berg, Joyce E., and Thomas A. Rietz, 2006, The Iowa Electronic Market: Lessons Learned and Answers Yearned, Information Markets: A New Way of Making Decisions, ed.

Robert W. Hahn and Paul C. Tetlock, AEI-Brookings Joint Center, Washington, D.C..

Bliss, Robert R., and Nikolaos Panigirtzoglou, 2004, Option-implied risk aversion estimates, Journal of Finance 59, 407–446.

Camerer, Colin F., 1987, Do Biases in Probability Judgment Matter in Markets? Experimental Evidence, American Economic Review 77, 981-997.

Camerer, Colin F., and Ernst Fehr, 2006, When Does ‘Economic Man’ Dominate Social Behavior? Science 311, 47-52.

Camerer, Colin F., George Loewenstein, and Martin Weber, 1989, The Curse of Knowledge in Economic Settings: An Experimental Analysis, Journal of Political Economy 97, 1232-1254.

CME, 2005, An Introduction to CME Economic Derivatives (Chicago Mercantile Exchange, Chicago et al.).

DellaVigna, Stefano, 2009, Psychology and Economics: Evidence from the Field, Journal of Economic Literature 47, 315-372.

Fehr, Ernst, and Jean-Robert Tyran, 2005, Individual Irrationality and Aggregate Outcomes, Journal of Economic Perspectives 19, 43-66.

Filippov, Andrei, 2005, Macroeconomic Derivatives: Overview and Sovereign Debt Market Applications, Master Thesis, Norges Handelshoyskole, Bergen.

Fischhoff, Baruch, Slovic, Paul, and Sarah Lichtenstein, 1978, Fault Trees: Sensitivity of

Estimated Failure Probabilities to Problem Representation. Journal of Experimental Psychology: Human Peception and Performance, 4, 330-344.

Fox, Craig R., and Robert T. Clemen, 2005, Subjective Probability Assessment in Decision Analysis: Partition Dependence and Bias toward the Ignorance Prior, Management Science 51, 1417-1432.

Fox, Craig R., and Jonathan Levav, 2004, Partition–Edit–Count: Naive Extensional Reasoning in Judgment of Conditional Probability, Journal of Experimental Psychology 133, 626-642.

Fox, Craig R., and Yuval Rottenstreich, 2003, Partition Priming in Judgment under Uncertainty, Psychological Science 14, 195-200.

Ganguly, Ananda R., John H. Kagel, and Donald V. Moser, 2000, Do Asset Market Prices Reflect Traders' Judgment Biases? Journal of Risk and Uncertainty 20, 219-245.

Gennaioli, Nicola and Andrei Shleifer, 2009, What comes to mind, Harvard Department of Economics Working paperW15084, http://isites.harvard.edu/fs/docs/icb.topic470237.files/articles%20spring%202009/whatcomestomind_february_200914.pdf (accessed October 7, 2009).

Gjerstad, Steven, 2005, Risk Aversion, Beliefs, and Prediction Market Equilibrium, Economic Science Laboratory, University of Arizona Working paper.

Gürkaynak, Refet, and Justin Wolfers, 2006, Macroeconomic Derivatives: An Initial

Analysis of Market-Based Macro Forecasts, Uncertainty, and Risk, National Bureau of Economic Research Working paper W11929.

Hsu, Ming, Ian Krajbich, Chen Zhao, and Colin Camerer, 2009, Neural Response to Reward Anticipation under Risk Is Nonlinear in Probabilities Journal of Neuroscience 29, 2231-2237.

Johnson, Eric J., John Hershey, Jacqueline Meszaros, and Howard Kunreuther, 1993, Framing, Probability Distortions, and Insurance Decisions, Journal of Risk and Uncertainty 7, 35-51.

Kahneman, Daniel, and Amos Tversky, 1979, Prospect theory: An analysis of decision under risk, Econometrica 47, 263-291.

Langer, Thomas, and Craig R. Fox, 2009, Biases in Allocation under Risk and Uncertainty: Partition Dependence, Unit Dependence, and Procedure Dependence, University of Muenster, University of California at Los Angeles Working paper.

Majumder, Saikat Ray, Daniel Diermeier, Thomas A. Rietz, and Luís A. Nunes Amaral, 2009, Price dynamics in political prediction markets, Proceedings of the National Academy of Sciences of the United States of America 106, 679-684.

Manski, Charles F., 2006, Interpreting the Predictions of Prediction Markets, Economic Letters 91, 425-429.

Morwitz, Vicki G., Eric Johnson, and David Schmittlein, 1993, Does Measuring Intent Change Behavior?, The Journal of Consumer Research 20, 46-61.

Ottaviani, Marco and Peter Norman Sorensen,2009, Surprised by the parimutuel odds? American Economic Review, in press.

Scharfstein, David S., and Jeremy C. Stein, 2000, The Dark Side of Internal Capital Markets: Divisional Rent-Seeking and Inefficient Investment, Journal of Finance 55, 2537-2564.

See, Kelly E., Craig R. Fox, and Yuval S. Rottenstreich, 2006, Between Ignorance and Truth: Partition Dependence and Learning in Judgment under Uncertainty, Journal of Experimental Psychology 32, 1385-1402.

Snowberg, Erik and Justin Wolfers,2007, Explaining the favorite-longshot bias: Is it risk-love or misperceptions?, http://www.hss.caltech.edu/~snowberg/papers/Snowberg-Wolfers%20Risk%20Love%20or%20Decision%20Weights3.pdf

Tversky, Amos, and Daniel Kahneman, 1992, Advances in prospect theory: Cumulative representation of uncertainty, Journal of Risk and Uncertainty 5, 297-323.

Wolfers, Justin, and Eric Zitzewitz, 2004, Prediction Markets, Journal of Economic Perspectives 18, 107-126.

Wolfers, Justin, and Eric Zitzewitz, 2007, Interpreting Prediction Market Prices as Probabilities, Wharton, University of Pennsylvania, and Stanford GSB Working paper.