PARTITION-DEPENDENT FRAMING EFFECTS IN LAB AND FIELD PREDICTION MARKETS*

ULRICH SONNEMANN

COLIN CAMERER

CRAIG R. FOX

THOMAS LANGER

Many psychology experiments show that individually judged probabilities of the same event can vary depending on the partition of the state space (a framing effect called "partition-dependence"). We show that these biases transfer to competitive prediction markets in which multiple informed traders are provided economic incentives to bet on their beliefs about events. We report results of a short controlled lab study, a longer field experiment (betting on the NBA playoffs and the FIFA World Cup), and naturally-occurring trading in macro-economic derivatives. The combined evidence suggests that partition-dependence can exist and persist in lab and field prediction markets.

This version: 2-22-2008

JEL classification: D8, G1

Keywords: prediction markets, framing effects

* Comments welcome. We thank Justin Wolfers for providing the economic derivatives market data used in Section IV, 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 for feedback. This research was supported by the DFG-grant LA1316/3-1, NSF grant (CRF), the Moore Foundation, NSF-HSD and HSFP grants (CFC). Direct correspondence to: Thomas Langer (thomas.langer@wiwi.uni-muenster.de).

I. INTRODUCTION

A number of recent psychological experiments have shown that the judged probability distribution of a continuous variable, such as the closing price of a stock index, depends on the particular intervals into which the variable's possible values are divided, a phenomenon called "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. This 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.

Our study investigates partition-dependence in experimental and field "prediction markets" for three types of naturally-occurring event domains (financial, sports, and weather outcomes). In prediction markets, people typically trade a set of all-or-nothing contingent claims on actual events. A claim pays off if and only if its associated event occurs. The price of the contingent claim is thought to reflect the market's collective probability judgment about the event's likelihood (Manski 2006; Wolfers and Zitzewitz 2005b).

Most economists are instinctively skeptical of psychology experiments that use simple abstract or hypothetical events, modest (or no) performance-based financial incentives, and little opportunity for learning. These concerns are addressed in our experiments where all choices involve prediction-market bets on actual events, with substantial payoffs linked to choices and outcomes, and trading takes place over time periods lasting from ten minutes to several weeks, which provide a substantial opportunity for learning. Taking advantage of the complementarities of lab and field methods, we report a lab experiment, a field experiment, and some analysis of naturally-occurring field data.

These results may interest both psychologists and economists. For psychologists, the magnitude and persistence of these effects in prediction markets tell us something about their psychological nature: Are they transient slips of the mind that are quickly displaced by effortful thought, and erased by competition? Or do the concrete boundaries of a presented partition persistently influence cognition? For economists, partition-dependence is a distinct type of framing effect—the way in which an event is described or "framed" influences its judged likelihood. This phenomenon violates a bedrock principle of rationality that Arrow (1982) referred to as *extensionality* and Tversky & Kahneman (1986) called *description invariance*: "The chosen element depends on the opportunity set from which the choice is to be made, independently of how that set is described" (Arrow 1982, p. 6). In fact, there are already examples of large-scale effects in economic field data that are consistent with a partition-dependent 1/*N* bias (see Section I.A.), in personal and corporate resource allocation, and race-track odds.

Some background about both partition-dependence and prediction markets is useful to present before proceeding to the details of the data and what we found.

I.A. Partition-dependence

It is now well established in the psychology literature that limited attention and awareness can lead to reliance on judgmental heuristics, which can deviate systematically from normative logical standards (Kahneman, Slovic, and Tversky 1982; Gilovich, Griffin, and Kahneman 2002).

An early example is "fault-tree bias" (which set the stage for later studies). A fault tree is a hierarchical display with branches showing possible mechanical explanations for an observed system failure (such as an airplane crash or a car failing to start). Increasing levels of detail are shown further down the tree branches. Engineers often create fault trees and assign likelihoods to the branches representing possible causes of a system failure.

Normatively, when statistically important branches are omitted from a fault tree, the subjective probability assigned to those fault branches should be reassigned to a residual "other causes" branch. However, experiments showed that the increase in "other causes" probability when large fault tree branches are omitted is too small, even when the subjects are highly knowledgeable about likely faults. For instance, when experienced auto mechanics were asked to estimate the relative frequency of six categories of reasons why a car might fail to start (battery, starting system, fuel system, ignition system, engine, mischief, all other problems) the mean proportion assigned to "all other problems" was .060. However, in another treatment where three of these categories (starting system, ignition system, mischief) were pruned from the original tree the proportion assigned to "all other problems" was only .215 rather than .441 implied by responses given to the unpruned tree (Fischhoff, Slovic, and Lichtenstein 1978).

Four psychological mechanisms have been proposed for fault tree bias of this sort: (1) enhanced psychological "availability" of explicitly mentioned categories[1], resulting in higher judged probability; (2) ambiguity or vagueness about omitted branches[2]; (3) an ecologically valid belief that the presented fault tree conveys information about likelihood, because omit-

[1] Tversky and Kahneman (1973), Fischhoff, Slovic, and Lichtenstein (1978), Van der Pligt, Eiser, and Speark (1987), Dubé-Rioux and Russo (1988), Russo and Kolzow (1994), and Ofir (2000).

[2] Hirt and Castellan (1988).

ted branches are likely to be rare;[3] and (4) a bias toward an ignorance prior of 1/*N* on each of the *N* identified events[4]. Which mechanism is driving the bias is important because different mechanisms imply different limiting conditions, moderators and "de-biasing" techniques.

Fox and Clemen (2005) distinguish among these explanations by asking participants to judge the likelihood of chance nodes of decision trees that had been partitioned in one of two different ways. In one study expert members of the Decision Analysis Society (an international association of professional decision analysts and leading scholars of decision analysis) were asked 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.) 58 of 169 contacted members participated (34%) and were randomly assigned to either a low group or a high 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+]. The judged probability that future membership will reside in the upper interval (>1000) was 10% in the low group, for whom that interval is represented by a single event. The comparable judgment was 35% in the high group, for whom the (>1000) interval is partitioned into four separate events.

This example is notable because the subjects are highly expert and responded self-selectively. The first three psychological mechanisms described above cannot explain the difference in judgments between the low and high partition groups. The categories cover all possible ranges of events (i.e., there is no "other" partition), categories are not ambiguous, and participants were told about both possible partitions so that no information was conveyed by a single partition structure. Only the remaining explanation, a natural bias toward an ignorance prior across the categories, can explain the effect. A pure ignorance prior over presented categories would yield 1/*N* judgments of 20% and 80% in the low and high groups, respectively. The actual results of 10% and 35% are partway between those 1/*N* judgments and a common subjective probability for the interval (>1000) that is partition-independent.

Other experiments have shown substantial robustness of partition-dependence to many variables. Partition-dependence was exhibited in controlled learning environment (See, Fox, and Rottenstreich 2006), using "linguistic priming"[5], in solving probability puzzles, such as a

[3] Fischhoff, Slovic, and Lichtenstein (1978), Dubé-Rioux and Russo (1988); see also Sher and McKenzie (2006).

[4] Fox and Clemen (2005).

[5] Linguistic priming means that different descriptions of the same event can influence the relative salience of alternative partitions of the state space. For example, when subjects are asked the likelihood that "tomorrow General Motors' (GM's) stock price will rise by more than any other stock on the DJIA" their judgments are

version of the Monty Hall three-door problem (Fox and Levav 2004), in valuation of insurance policies (Johnson et al. 1993), and with incentive-compatible payoffs (Fox and Rottenstreich 2003; Fox and Levav 2004; Fox and Clemen 2005).

Partition-dependence has also been shown when resources, rather than probability, are allocated to categories. Benartzi and Thaler (2001) show bias toward 1/*N* in experimental 401(k) investment decisions and in an empirical analysis of retirement savings plan data. Langer and Fox (2005) show partition-dependence more explicitly, in allocations among investments in simple gambles with incentive-compatible payoffs. Other experiments on risky choice are showing that splitting positive-outcome events into sub-events seems to increase preference for those choices (e.g., Humphrey [1996]). Bardolet, Fox, and Lovallo (2007) find in archival data that corporations allocate less capital to divisions when there are more divisions under the same corporate parent, consistent with a 1/*N* bias (see also Scharfstein and Stein [2000]). They also find experimental evidence that experienced managers are statistically biased toward 1/*N* in their hypothetical capital allocations even though they are not aware of their bias. Many studies with many years of data in different countries show a favourite-longshot bias in horse race betting odds: Unlikely winners (longshots) are generally overbet and favourites are underbet, which is consistent with a bias toward 1/*N* probability for every horse (e.g., Wolfers and Zitzewitz [2005a]).[6] Fox, Ratner, and Lieb (2005) show 1/*N* bias in experiments allocating money to beneficiaries, consumption to time periods, and choices to menus of options that are grouped by different attributes.[7]

Note that the basic phenomenon underlying partition-dependence is the tendency of concrete, salient categorization to influence attention, thought, and judgment. This effect of salience based on how possible outcomes are described is ubiquitous in human communication, because complicated ranges of outcomes are rarely categorized naturally. Instead, a discrete categorical structure is typically *chosen*, or implicitly conveyed by a choice of words.

higher than when the event is phrased "Tomorrow on the DJIA, the stock whose price will rise by the greatest amount will be General Motors (GM)." The first phrasing, by mentioning the target event at the outset, primes a partition into the target event and its complement ("GM stock will be the highest" or "GM stock will *not* be the highest") and an ignorance prior of 1/2, whereas the second phrasing, by mentioning the equivalence class at the outset suggests a partition of the state space into 30 stocks and an ignorance prior of 1/30 (Fox and Rottenstreich [2003]; see also Fox and Levav [2004]).

[6] However, Plott and Roust (2005) find that in some lab settings with parimutuel betting on abstract events, the favourite-longshot bias is a disequilibrium phenomenon which disappears when some institutional changes are made.

[7] Note that in many cases, allocating resources equally could be optimal (e.g., consumption smoothing over time when utility of consumption is time-separable and concave). But the point of these studies is that the way in which categories are unpacked or combined influences allocations, which is not optimal. For instance, allocating consumption equally among months will produce (slightly) different results than allocating consumption equally among weeks.

For example, in February 2003, a month before the onset of the Iraq war, U.S. Defense Secretary Donald Rumseld said "It is unknowable how long that conflict [the war in Iraq] will last. It could last six days, six weeks. I doubt six months" (Page 2003). Rumsfeld's wording invites consideration of a partition of possible war lengths into intervals of [0, 6 days], (6 days, 6 weeks], (6 weeks, 6 months], and (6 months+). If Rumsfeld had worded his sentence differently (e.g., "six months, six years, or six decades"), it might have established a different partition, with a different public perception of likely outcomes, with different political ramifications.

In most cases, partition-dependence is difficult to expunge because talking about continuous variables often leads to a division of possible outcomes into lumpy natural-language categories. So if partition-dependence is prominent when there are clear historical frequencies lurking behind the cognitive walls of the presented partition, as in the naturally-occurring event domains used in our experiments (financials, sports, and weather), then it might be even more prominent when "unknowable" distributions such as the length of a war are divided into discrete numerical intervals. Tversky and Koehler note that (1994, p. 565), "the need to consider unavailable possibilities"…"is perhaps the fundamental problem of probability assessment". They suggest that immunity of judgments to a particular partition is "normatively unassailable but practically unachievable", because people "cannot be expected to think of all relevant conjunctive unpackings or to generate all relevant future scenarios".

Economic theorists have recognized the importance of cognitive availability that underlies partition-dependence and begun to model it formally. Dekel, Lipman, and Rustichini (1998, p. 524) note that "an unforeseen contingency is not necessarily one the agent *could not* conceive of, just one he *doesn't* think of at the time he makes his choice." Interest in "unforeseen contingencies" is generated by potential applications like simplified employment contracts and the desire for flexibility when it may be difficult to imagine future events or judge their likelihood (Kreps 1979). Ahn and Ergin (2007) show that partition-dependence revealed by choices can be modeled by allowing subjective probability to be non-additive in a particular way.[8]

I.B. Prediction Markets

The assets (or shares) used in our experimental studies are usually referred to as "winner-take-all" contracts (or "all-or-nothing" contracts, or contingent claims) in prediction mar-

[8] Their paper contains a particularly thoughtful review of the psychological literature motivating their axiomatization.

kets studies (Wolfers and Zitzewitz 2004). Under reasonable assumptions, the prices from these assets can be directly interpreted as market generated probability estimates of the occurrence of these events (Wolfers and Zitzewitz 2005b). The first large-scale prediction markets were created in 1988 at the University of Iowa (Forsythe et al. 1992; Berg and Rietz 2005), to trade assets linked to political events.[9] Over the years, prices in the Iowa markets have been shown to forecast political outcomes more accurately than many polls about 75% of the time, in hundreds of elections. Around 2001, websites emerged where people can trade contingent claims on a wider range of event domains including political, financial and entertainment events, such as "American Idol" outcomes and when Osama Bin-Laden will be captured (Intrade: http://www.intrade.com, cf. Wolfers and Zitzewitz [2004]). Firms have also created internal markets to predict outcomes of corporate interest, such as new product sales (Chen and Plott 2002; Ho and Chen 2007).

Partition-dependence creates a challenge for prediction market design. In these markets, continuous variables such as political vote shares, movie box office grosses, new product sales, timing of event occurrences, and values of macroeconomic variables, must be necessarily partitioned into numerical intervals by the market designers. Unlike categorical markets such as the winner of the Academy Award for Best Picture or the winner of the Super Bowl, there is typically no natural partition. If the way in which partitions are constructed matters for actual prediction-market prices, this will affect the quality of the resulting market-wide estimates (as shown in Section IV). Designers should treat partition-dependence as a cognitive constraint that must be understood and anticipated in a design, much as website screen displays and menu features are chosen to satisfy design goals based on an understanding of visual and motor activity.

A well-designed prediction market will eliminate ambiguity in the definition of events, and can control for the information conveyed by the partition choice if traders know how partitions are created. However, the natural bias toward $1/N$ in the assessment of interval probability cannot necessarily be designed away. Indeed, in naturally-occurring prediction markets, only a single partition for an event is always used. So without experiments like ours that compare market prices for different partitions of the same state space, there is no way to know for sure whether there is a bias toward $1/N$.

I.C. Plan of the Paper and Preview of Results

[9] The Iowa markets were inspired by lab evidence that simple abstract experimental markets can aggregate diverse information well; see Plott and Sunder (1982), and Sunder (1995).

The next three sections report analyses of three types of data. In Section II we describe short-run experimental markets (two 10-minute trading periods) for three naturally-occurring event domains in which we can compare judgments and prices for different partitions of the same numerical interval. These data largely replicate the persistence and magnitude of partition-dependence reported in many psychology experiments (like the canonical Fox-Clemen study mentioned in Section I.B.). Section III describes a longer-run experiment conducted on the web, lasting several weeks. Subjects traded assets linked to team victories in the NBA playoffs and to FIFA World Cup soccer goal scoring. There is noticeable partition-dependence but its magnitude is smaller than in the first lab study. Section IV describes data from naturally-occurring markets for numerical values of important statistics that macroeconomists follow, called an "economic derivatives market", created by Goldman Sachs and Deutsche Bank. A structural model of these data, which assumes that observed prices mix a $1/N$ ignorance prior belief with other information, enables us to back out a de-biased distribution that predicts more accurately than the observed prices and suggests some degree of partition-dependence.

All three analyses have strengths and weaknesses that are partly compensated for by the other studies (i.e., they are scientific complements). The lab experiments are the easiest to run and replicate, and they provide an initial estimate of whether partition-dependence exists and persists in the short-run. However, lab experiments do not tell us whether the effects would persist in the longer run. The field experiments on the NBA playoffs and soccer World Cup involve a longer span of trading and self-selection of traders who know a lot about the event domains and follow them closely (if not fanatically). The field data on economic derivatives do not compare different partitions for the same variable, as we can do in the lab, but they involve higher implicit stakes and attract more sophisticated (and highly-paid) participants than we can ordinarily use in the lab.

All three studies show evidence of partition-dependence. They provide an example in which a simple observation first discovered in straightforward psychology experiments is robust to market learning opportunities, increases in the span of trading, and the sophistication of traders. These results do not imply that partition-dependence can never be eliminated under any conditions. The results simply establish that some conditions that *might* eliminate partition-dependence do not appear to do so, although in some cases (e.g., the first lab study) evidence of partition-dependence seems to decrease over time.

Instructions (translated from German) and many technical details are gathered in a set of Appendices [not intended for publication], along with some analyses that were omitted for brevity.

II. STUDY 1: A LABORATORY EXPERIMENT

Our first study is designed to see whether partition-dependence occurs and persists in short-run experimental markets. We also compare effects expressed in probability judgments (both before and after trading) with effects revealed by prediction-market trading prices.

II.A. Experimental Design

Twelve two-hour experimental trading sessions were conducted in April 2007 with 16 traders in each session, divided into two self-contained groups (markets) with 8 traders in each. Subjects were N=192 undergraduate finance students (134 male, 58 female) from the University of Muenster (in Germany).[10] The sessions spanned one week and took place in a computerized lab environment where participants were separated from each other by dividers during the trading periods. The instructions (see Appendix II) were read out aloud to ensure that all information about the experiment was common knowledge.

The essential part of the experiment consisted of several trading rounds in a set of three simple assets that are 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 that corresponded 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 I 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 that the DAX index value is in the intervals [0, 7327.99], [7328, 7496.99], or 7497 and above (denoted

[10] This experiment corrected a small flaw in an earlier pilot study. More details on the design and results of the pilot study are provided in Appendix I.

[7497+]). In partition 2 (the high partition) the events are based on 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.[11] 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).

The main treatment variable is the way in which the state space is partitioned into events. 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. As Figure I that was not shown to the subjects illustrates, to create these partitions, each state space was initially divided into four disjoint and exhaustive intervals (I_1 to I_4). 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.[12]

[**FIGURE I**: CONSTRUCTION OF ASSETS FOR THE TWO DAX PARTITIONS]

For the weather and sports event domains the interval boundaries were chosen rather arbitrarily based on historical outcomes, so there is no conclusive way to link probabilities expressed by individual judgments or inferred from market prices to objective probabilities.

[11] The weather partitions are: [-, 15.9], [16.0, 19.9], [20.0+] (low partition) and [-, 19.9], [20.0, 23.9], [24.0+] (high partition). The sports partitions are [0, 20], [21, 25], [26+] (low partition) and [0, 25], [26, 30], [31+] (high partition).

[12] Thus, to the extent that any information is conveyed by the partitions described in the instructions, it is that the experimenter thinks that the dividing point between intervals 2 and 3 is special (perhaps it divides the state space into regions of relatively equal expected likelihood). However, because there is no informational asymmetry between conditions, partition dependence cannot be rationalized on the basis of information conveyed to participants by the partitions presented.

However, 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 the four intervals represented a particular percentile of the expected PDF.[13]

For each of the three event domains two completely identical and independent trading rounds were run successively, resulting in six trading rounds per participant and experimental group, as shown in Figure II. Each trading round lasted ten minutes (with short breaks between rounds). The order in which the participants traded assets from the three event domains varied for each experimental session and was perfectly counterbalanced (i.e., for each of the six possible event domain orders there were two experimental sessions) to avoid any order effects. In each of the six trading rounds the participants were initially endowed with a combination of assets (i.e., unit portfolios spanning the set of assets) and cash, to the value of €20 in total.[14] Participants were compensated based on their final cash and asset holdings for an afterwards randomly chosen trading round, at the point when 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, although traders could use their available cash to buy unit portfolios and then sell the underlying assets. No explicit transaction costs were imposed for trading.

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 IV 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;[15] in particular they could not trade across markets with different partitions. During instruction

[13] This was done to allow—in addition to our main treatment effect—us to pursue comparison of our experimental market prices to historical frequencies. We will not discuss this issue in this paper. Note, however, that since different experimental sessions were spread out over a week, the DAX intervals were adjusted for each experimental session (based on the recent DAX index close) to preserve the percentiles (see Appendix III). This day-by-day adjustment of interval boundaries also enhanced comparability and aggregation of the data from sessions on different dates. Traders were not told about the procedure for constructing and adjusting the intervals, since doing so would instruct subjects about expected probabilities and constitute an additional treatment effect.

[14] In each market (of eight participants) groups of two traders were randomly endowed with one of the four different combinations: 16 unit portfolios + 400 cents, 12/800, 8/1200, 4/1600, all of them representing a value of €20. For each trader her initial endowment was the same over the six trading rounds.

[15] The only exception was trading the unit portfolio which was always executed immediately against the experimenter.

and a practice trading round, participants were told how to exploit arbitrage opportunities by trading unit portfolios with the experimenter for cash.

Before the first trading round for each event domain, and after the second (and final) trading round, the participants were asked to provide their individual probability judgments for the occurrence of the three events they traded. These judgments were not incentivized.[16] 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], and Ganguly, Kagel, and Moser [2000]). Like those studies, one question our method can address is whether partition-dependence 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]).

[**FIGURE II:** EXAMPLE OF THE TIME COURSE OF AN EXPERIMENTAL SESSION]

II.B. Results: Judged Probabilities

We start by considering the 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 I). Hypothesis 1 states the partition-dependence prediction for intervals I_1 and I_2 (H_0(a)) and intervals I_3 and I_4 (H_0(b))[17].

Hypothesis 1:

H_0(a): $p(I_1) + p(I_2) > p(I_1 \cup I_2)$ and

H_0(b): $p(I_3) + p(I_4) > p(I_3 \cup I_4)$

[16] Participants were also asked in advance for their self-rated competence in making such probability judgments in the domain of the specific stimulus (scale 1 [incompetent] to 7 [very competent]). At the end of the session they were further asked to provide some personal information including age, self-rated knowledge in the field of statistics and econometrics, or trading experience. As mentioned above, the sessions were constructed of equally-competent people (in general knowledge about soccer) to see if group-level competence would affect the degree of partition-dependence, but we could find no such effects.

[17] Hypothesis H_0(b) is reported solely to correspond to later market prices presentation. Note that H_0(b) is, in fact, redundant for judged probabilities, since the results (i.e., effect size) have to be the same for both hypotheses by construction.

Table I 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).

[**TABLE I:** AVERAGE PRE-TRADING INDIVIDUAL PROBABILITY JUDGMENTS]

The mean difference between summed probabilities of unpacked intervals and of the packed interval is .312, .278 and .261 for the finance, weather, and sports event domains. (All reported differences are statistically highly significant, based on a Kruskal-Wallis test ($p<.0001$)).

Note that if people were applying an ignorance prior of 1/*N* of the judged probability to each of the three events they faced, the difference between the sum of the segregated-interval judgments and the packed-interval judgments should be one third. Thus, these judgments show a very strong effect of partition-dependence, as seen in earlier psychology experiments (e.g., Fox and Clemen [2005] or See, Fox, and Rottenstreich [2006]).

II.C. Results: Market Prices

Of course, probability judgments elicited from individuals might reflect thoughtless errors, which are strongly or weakly diminished in two 10-minute trading periods. Markets are, after all, a kind of dollar-activity-weighted opinion poll that also provide substantial time for reflection and opportunities for learning from others. We now turn to probabilities inferred from market prices.

We first present some general facts on trading volume and market efficiency. Table II shows that the average number of trades per market (except for unit portfolio trades) was 42.83, an average of 14.28 trades for each of the three assets, and total shares traded were about 140 in each market. Trading was relatively continuous across the 10-minute trading period and similar across all three event domains.[18]

[**TABLE II:** TRADING VOLUME STATISTICS IN STUDY 1]

[18] The percentage of trades in each one-minute interval averaged between 8.5 and 11.8% (the latter in the second minute).

Recall that buying or selling the unit portfolio could have been used to exploit arbitrage opportunities rapidly. 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 are exploited after a few seconds.[19]

Figure III 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 unpacked assets. The gap between the paths shows the size and persistence of the 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).[20] These estimates imply that there is some slow convergence that motivates our later experiments with much longer time periods.

[**FIGURE III:** DEVELOPMENT OF PRICE DIFFERENCES OVER TIME FOR THE SPORTS ASSETS IN STUDY 1]

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 hypotheses of partition-dependence in prices are parallel to those above for judgments:

Hypothesis 2:

H_0(a): $P^*(I_1) + P^*(I_2) > P^*(I_1 \cup I_2)$ and

H_0(b): $P^*(I_3) + P^*(I_4) > P^*(I_3 \cup I_4)$

The experiment generates twelve equilibrium prices per asset for each partition. Table III shows the mean prices (divided by 100 cents to make them comparable to probabilities) for the three assets of each partition. For comparison we also report the average judgments after

[19] For ask arbitrage, out of 144 markets, 87 have no arbitrage opportunities at all. Emerging arbitrage opportunities are exploited 10.83 seconds (median) after the opportunity first appeared. For bid arbitrage, 20 of 144 markets have no arbitrage opportunities. Emerging arbitrage opportunities are exploited 12.24 seconds (median) after the opportunity first appeared. Furthermore, the chance to earn an arbitrage profit of more than 5 cents (e.g. asset prices summing to less than 95 cents or more than 105 cents) is incredibly rare; so even when arbitrage opportunities exist they are small. More details on arbitrage opportunities are provided in Appendix V.

[20] Graphs and time-trend estimations for the finance and weather event domains can be found in Appendix VI. We present the sports data in the text because the degree of convergence for the sports event domain (average slope difference .0127) is between the respective values for finance (.0099) and weather (.0148).

the trading and the partition difference (PD) from pre-trading judgments. Medians are reported in Appendix VII.

[**TABLE III:** MEAN EQUILIBRIUM PRICES (2ND TRADING ROUND) AND INDIVIDUAL JUDGMENTS (POST-TRADING AND PRE-TRADING)]

The difference between summed prices (rescaled) of unpacked assets and the packed asset, averaged across the two partitions, is .267, .149, and .229.[21] (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 ex-ante probability judgments are .312, .278, and .261 and from post-trading judged probabilities are 257, .226, and .256. Market-price partition-dependence is slightly reduced compared to probability judgments for finance and sports event domains, and cut in half for weather events. However, conservative tests using session-level data only show a statistically significant reduction for the weather event domain.[22]

Market experience also creates a slight reduction in partition-dependence between pre-trading and post-trading judgments. For a more powerful within-subject test we analyze for each subject the difference between the ex post and the ex ante probability judgments for the packed intervals. Since the general direction of the bias suggests the judgments for the packed intervals are too low, we define a subject's judgments to reflect a bias reduction due to trading experience if this difference is positive, i.e. if post-trading judgments for the packed intervals are higher than pre-trading judgments. Averaging the differences for the three event domains per subject, values are positive in 52.1% of the cases (100 out of 192 subjects), and negative in 34.4% of the cases (66 out of 192), a significant asymmetry.[23] Trading does not influence the remaining 26 subjects in any direction. Overall, the two 10-minute trading periods apparently have a modest de-biasing effect on individual judgments, though a substantial degree of partition dependence remains after 20 minutes.

[21] That is, we average the magnitude of partition-dependence in equilibrium market prices for the upper and the lower panel (e.g., (.289 + .245)/2 = .267 for finance). Note that the partition difference is not necessarily the same for the lower and higher intervals in market prices, in contrast to the judgments where this difference has to be the same for the lower and higher intervals due to the fact that, by instructions, they always summed to 1.0.

[22] 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 H_0(a) and .35 for H_0(b) for finance, .53 and .10 for sports, and .02 and .02 for weather.

[23] A sign test shows that the effect is significant ($p<.01$).

Summary: The lab Study 1 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.

Note also that the size of the effects—a difference of around .2–.3 between the sum of unpacked event probabilities and the packed event probability—is quite close to the gap of .35 in the example from Fox and Clemen (2005) presented in the introduction. The effect is even larger when stated in terms of percentage changes: the median probabilities of the packed events are around .35, so an increase of .2–.3 in probability by unpacking that event into two components inflates its perceived relative probability by more than one-half.

III. Study 2: An NBA/FIFA Field Experiment

The modest effects of trading experience on partition-dependence seen in Study 1, after twenty minutes of trading, suggest the possibility that with much longer trading spans, and perhaps with more knowledgeable traders, partition-dependence could be reduced more strongly or wiped out. Study 2, a field experiment lasting several weeks, was designed to test this hypothesis.

III.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 the CASSEL list at UCLA, Los Angeles (United States).[24]

Contracts are all-or-nothing contingent claims on intervals of 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

[24] We used two different channels of recruitment as we planned to analyze second-order effects (We expected U.S. students to feel more competent about NBA events whereas German students should feel more competent in the FIFA soccer World Cup events). Because U.S. participation in the World Cup markets was low, there is little statistical power to detect such effects so they will not be discussed further.

goals). As in Study 1, for each event domain, there are two partitions that combine sub-events differently, as shown in Figures IVa and IVb. For example, in the NBA markets the first partition packs the victory intervals [4, 7] and [8, 11] into a single interval [4, 11], and unpacks the interval [12, 16] into two components of [12, 15] and [16]. In the instructions, the participants were explicitly informed what the two different partition sets were (and that they were randomly assigned to only one partition), to control for the concern that offering one partition would convey information to subjects about likelihoods.[25] Every participant was allowed to trade assets based on four different teams—called "team markets"—using the same numerical partitions for each of the four teams.

[**FIGURE IVa:** CONSTRUCTION OF ASSET PARTITIONS (PLAYOFFS VICTORY TOTALS)]

[**FIGURE IVb:** CONSTRUCTION OF ASSET PARTITIONS (WORLD CUP GOAL TOTALS)]

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[26]. 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 likely based on the results from the three previous World Cups.

The experimental protocol was similar to Study 1, but was adapted for the Web (see Appendix VIII).[27] Participants were instructed about the composition of assets and markets

[25] Note that subjects did not see the actual Figures IVa and IVb which show the clear links between packed and unpacked events, but they were informed of the two sets of partition intervals (Appendix VIII contains the complete instructions of Study 2. Information about the partitions can be found in subsection 3.1.of the instructions.).

[26] Each fixture is a best-of-seven match, so the first team to win four games wins the round and advances. Therefore, betting on the interval [0, 3] is equivalent to betting that the team has to leave the playoffs in the first round, because a team that only wins a total of 0–3 games will be eliminated by an opponent that wins four. The interval [12, 15] is equivalent to winning three rounds but losing in the fourth (and final) round. The interval [16] is only reached by the NBA champion, who wins four games in all four rounds.

[27] Participants were randomly assigned to one four-team market for NBA playoffs games, then to another four-team market for FIFA World Cup. Groups were reshuffled for the World Cup markets so that students only faced the same traders again by coincidence. Participants from the two different recruitment channels were never assigned to the same market, though. (This happened for reasons of a second order competence analysis that we will not discuss in this paper.) In addition, the students didn't know anything about the identity of their counterparts in a market. Each experimental group initially had twenty traders, but some dropped out over time. NBA four-team markets included two teams from each of the two conferences (Eastern and Western). World Cup four-team markets used four official tournament "groups", out of the eight groups created by FIFA organizers, which were supposed to generate the most interest and trade. (The groups we used were group A: Germany, Costa Rica, Poland, Ecuador; group C: Argentina, Côte d'Ivoire, Serbia/Montenegro, Netherlands; group E:

(including the partitions of assets they could trade, and the alternative partition), how to use the trading system, and some details about the NBA playoffs and the FIFA World Cup, by e-mail. They also had Internet access to a homepage with study details, FAQs, and a practice market. The market was open continuously. As in Study 1, the trading mechanism was a multi-unit continuous double auction (CDA) with a hidden order book, so that subjects 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.[28] At the end of the experiment we randomly drew one out of the four teams for each experimental group to compensate the 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 the intervals corresponding to the assets they traded.

It is important to stress that the participants only had direct access to their own four-team NBA and World Cup markets. They could not directly observe market data (like prices or quotes) from other experimental groups trading different partitions.[29]

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.

Italy, Ghana, USA, Czech Republic; and group F: Brazil, Croatia, Australia, Japan.) These design choices should not have any obvious biases in creating or diminishing partition-dependence.

[28] In each market (of twenty participants) always four traders were randomly endowed with one of the five different combinations: 9 unit portfolios + 100 cents, 7/300, 5/500, 3/700, and 1/900, all of them representing a value of €10. For each trader the composition of her initial endowment was the same for the four "team markets" in the NBA playoffs part, but was randomized again for the FIFA World Cup markets.

[29] Of course, we cannot exclude the possibility that students were informed about these prices by friends that happen to trade exactly the same teams but the other partition of the state space. However, even in this case, arbitrage opportunities across markets could not jointly be exploited as it was not guaranteed that the same team was chosen for the incentive-compatible payment in both groups.

III.B. Results

The analysis of results is similar to the analysis from lab Study 1 in Section II. First we test for partition-dependence in the individual probability judgments elicited before the beginning of trade. Next we look for partition-dependence in the bids, asks, and trading prices in the markets. We also test whether the probabilities for the lowest-outcome event ([0, 3]), which is the same interval in both partitions, happen to differ in the differently-partitioned markets. There is no reason to expect that these probabilities will differ (because the ignorance prior probability is 1/4 for this event in both partitions), but any difference provides a measure of sampling error.

Because some participants did not submit probability judgments before the first playoffs game was played (and 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, there are N=263 judgment sets submitted by German participants before the opening game was played.[30]

Tables IVa and IVb show the median differences in judgments (×100) for the two different partitions and significance by a Kruskal-Wallis test. Not surprisingly, the differences in the commonly partitioned event [0, 3] between partitions (in the first column) are close to zero and not statistically significant.[31] The differences between the summed probabilities of the unpacked events and the probability for the corresponding packed event are positive, almost always highly significant, and are comparable in magnitude to the effects reported earlier (approximately a .20 increase in probability when the interval is unpacked).

[**TABLE IVa:** PARTITION-DEPENDENCE IN PRE-TRADING JUDGMENTS FOR NBA EVENTS]

[**TABLE IVb:** PARTITION-DEPENDENCE IN PRE-TRADING JUDGMENTS FOR FIFA EVENTS]

We now look at activity in the two most liquid experimental markets for each team (as measured by the overall number of trades).[32] In the NBA playoffs, the most liquid markets were the Dallas Mavericks (DAL) and Miami Heat (MIA) (partly because they became the two finalists, so their assets were traded for the longest span of time). For DAL, there were

[30] We do not report any results for U.S. participants since there was an extensive dropout of U.S. students for this part of the study.

[31] This is just a test for whether there are systematically different beliefs in the two markets, and also gives a measure of statistical variability that is useful for judging the size of any partition-dependence effect.

[32] For this purpose, we choose from the market clones the most active team-markets in each partition for each team and match them for further analyses.

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 Va:** PRICE CHART (DALLAS MAVERICKS, DAL)]

Figures Va and Vb show 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 is no current trading. Vertical lines indicate the beginning of a game. The (s_1), (s_2), etc. at the top of the charts indicate the number of cumulated wins after each game. For example, Figure Va shows that DAL won the first four games ((s_1) – (s_4)), lost the next game (status remains (s_4)), won the sixth game ((s_5)) 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 Vb:** PRICE CHART (MIAMI HEAT, MIA)]

Figure Vb shows similar patterns for the MIA markets. Partition-dependence is evident for the first thirty days and diminishes afterwards. Note that for both markets, the sum of prices within a partition is often above 100, indicating potential arbitrage opportunities. However, because trading was thin and bid-ask spreads were typically wide, there were few executable arbitrage opportunities.

For the World Cup, Figures VIa and VIb show price charts for Germany (59 and 57 trades in partitions 1 and 2) and Italy, the eventual World Cup champion (65 trades in both partitions).

Both markets show persistent partition-dependence of recent prices, i.e., the red line is above blue in the middle panel and vice versa in the lower panel.

[**FIGURE VIa:** PRICE CHART (GERMANY, GER)]

[**FIGURE VIb:** PRICE CHART (ITALY, ITA)]

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 the lab Study 1). Therefore partition-dependence is measured 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.

For brevity, one method is described in detail and the second is described in Appendix IX. The method described in Appendix IX calculates the time-weighted pseudo-arbitrage profits from selling the unpacked-interval assets and buying the equivalent packed-interval asset, 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. (We provide more summary details below.)

We now describe our second method, the “interpolated-price hypothetical arbitrage”. The past and future prices are used to *interpolate* a trade price continuously. 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 they cannot trade across markets, participants 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, selling trades are assumed to be 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 for the intervals I_1 and I_2 (intervals I_3 and I_4) at time t is:

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

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

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 you are *buying* the asset, we assume you 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 you are *selling* the asset, we assume you could sell it for the *lower* price of 42 during days 20 to 25.

[**FIGURE VIIa:** INTERPOLATED-PRICE HYPOTHETICAL ARBITRAGE (DALLAS MAVERICKS, DAL)]

[**FIGURE VIIb:** INTERPOLATED-PRICE HYPOTHETICAL ARBITRAGE (ITALY, ITA)]

Figure VIIa shows the interpolated-price hypothetical arbitrage profit over time for 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 in definition (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 in definition (1.2)). The panel two and four hypothetical profits from selling the unpacked-interval assets and buying the packed-interval asset (blue lines) are often positive and large in magnitude.

Figure VIIb shows the corresponding data from trades on 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 VIIa and VIIb would be above zero, but this is never the case. The fact that there is virtually no reverse effect proves partition-dependence in the expected direction (as indicated by the blue curves) to be 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 VIIa and VIIb, and divided by the total trading time (in days).[33] These statistics are provided for each team and interval in Table V.

[**TABLE V:** PER-DAY PROFITABILITY OF INTERPOLATED-PRICE HYPOTHETICAL PSEUDO-ARBITRAGE STRATEGIES]

The average per-day hypothetical profit from exploiting partition-dependence, 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 partition-dependence, across the 32 teams from both 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 IX for full details). 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 for exploiting partition-dependence are .03 and .35 for intervals I_1 and I_2, and .07 and 1.50 for

[33] Note that the relevant trading time ends either when the last auction for an asset of the relevant interval occurred or when the corresponding interval asset I_1 (or I_3) expired worthless.

intervals I_3 and I_4; the average across intervals is .92. Importantly, profits are higher from exploiting partition-dependence compared to reverse partition-dependence in 38 of 46 teams ($z=5.88$, $p<.001$ by a sign test, excluding 18 teams for which both profits are zero).

The hypothetical profits from these two measures could be treated as a lower and upper bound 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 artificially liquefies the market by essentially assuming there is always a latent trade waiting to occur at the right price, so this method provides an upper bound (though it is still conservative because it assumes trades would be executed at the worst of the most recent and next future prices).

Summary: Study 2 documents the persistence of partition-dependent pricing effects 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. These experiments address concerns about the generalizability of lab experiments due to the limited involvement of traders and short span of trade. Probability judgments before trading begins 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, gives 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 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 an artifact of random error.

IV. STUDY 3: NATURALLY-OCCURRING PREDICTION MARKETS FOR ECONOMIC DERIVATIVES

The lab and field experiments document the existence of partition-dependence when different partitions are traded (in separate markets) for the same event domain. An open question is whether these effects can be inferred from naturally-occurring prediction markets that rely on a single partition. Study 3 addresses this question.

In October 2002, Goldman Sachs and Deutsche Bank launched large-scale prediction markets for bets on the outcomes of macroeconomic indicators such as the growth in non-farm payrolls, retail sales, levels of the Institute for Supply Management's manufacturing diffusion index (a measure of business confidence), initial unemployment claims and the Euro-area harmonized CPI. These "economic derivatives" (ED) markets are designed to give professionals such as institutional traders (like hedge 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.

The basic contracts offered in these markets are similar to the assets in Studies 1 and 2. For each underlying numerical variable (i.e., the release of a specific numerical macroeconomic indicator) there exists a number of "digital (binary) options". Each digital option spans an interval of the possible outcomes of the indicator. Buying a set of options can be used to form "all-or-nothing" contracts covering disjoint intervals of possible outcomes for the released statistic value. All contracts together cover the whole state space.

The 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. Gürkaynak and Wolfers (2006), who report data covering the first 2½ years of these markets, conclude that market-generated forecasts based on prices of these prediction markets are more accurate than the "survey forecast" released by Money Market Services (MMS) on the Friday before a data release, much as the Iowa political market prices are typically more accurate than comparable political polls (Berg and Rietz 2005).

For each upcoming data release, 10–20 assets are offered in different ranges. The ranges are deliberately set to reflect the likely range of the outcomes. For example, the ranges for the non-farm payrolls auction increase range from 0 (i.e., unchanged from the previous level) to +300,000 in increments of 25,000 jobs.

The market mechanism employed is a parimutuel system, which is common in horse race betting (in contrast to the CDA mechanism used in our studies 1 and 2) (see Appendix X for a screenshot of the interface). In this mechanism the prices of the instruments are based solely on relative demand for their implied outcomes and enables market clearing without discrete matching of buy and sell orders. In parimutuel markets 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 in horse betting, the trading system periodically discloses interim prices showing what the payouts would be if no further orders were submitted.[34]

These auctions typically take place in the morning before the economic statistic is released and are sometimes preceded by another auction on the same statistic release one or two days before (e.g., non-farm payrolls auctions are held on both the morning the data are released and one day before). Thus, these markets usually have a very short-term forecast horizon and thereby offer hedging opportunities against so-called event risks. Figure VIII shows the implied probabilities from one set of digital options, for a retail trade statistic announced in April 2005.

[**FIGURE VIII:** "DIGITAL OPTION" PRICES ON RANGES OF RETAIL TRADE STATISTICS (GÜRKAYNAK AND WOLFERS 2006)]

First note that each ED market presents a single partition of possible event outcomes to participants (the digital option outcome ranges). As a result, we cannot compare prices in two differently-partitioned events on the same interval to estimate the degree of partition-dependence, as was done through experimental manipulation in Studies 1 and 2. However, we can posit a simple econometric model to estimate the degree of partition-dependence: For each event category x, assume

$$(2) \qquad f_{obs}(x) = (1 - \lambda) \cdot f_{true}(x) + \lambda \cdot f_{1/N}(x)$$

where $f_{obs}(x)$ is the observed implied probability distribution, $f_{true}(x)$ 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 was 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

[34] Since investors in the ED market are allowed to place limit orders, the parimutuel trading mechanism may result in multiple equilibriums. This problem is addressed by using a special auction-clearing tool that chooses the equilibrium prices such that the number of total trades will be maximized. (This clearing-algorithm was developed by Longitude Inc. and is called "Parimutuel Derivative Call Auction Technology"; see Baron and Lange [2003]). As in many traditional auctions, all trades (at a given strike) that occur are executed at the same price, regardless of the limit price.

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.[35] 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.[36]

From (2) it follows that M_{obs} is a linear function $(1 - \lambda) \cdot M_{true} + \lambda \cdot M_{1/N}$ and if we call the actual realization of the economic statistic X and apply a little algebra we see that the observed forecast error can be written as:

$$(3) \qquad M_{obs} - X = [M_{true} - X] - \lambda/(1 - \lambda) \cdot [M_{obs} - M_{1/N}]$$

That is, the observed forecast error has two components. The first component is the error term from a de-biased forecast based on $f_{true}(x)$ (which is expected to have expectation 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 there is partition-dependence contaminating $f_{obs}(x)$, 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 strength of the negative correlation can be used to estimate $-\lambda/(1 - \lambda)$ and the implied λ.

Table VI summarizes the results of estimating regression (3) for markets for four different statistics. There is modest support for a negative correlation between forecast errors and the forecast – $1/N$ gap, which is consistent with bias toward a $1/N$ prior. One of the event domains (initial unemployment claims) shows no bias; the other three event domains show substantial bias. However, small sample sizes for individual event domains make the effects statistically marginal. The coefficient estimated from pooling all the event domains, –.77 is more significant and 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 of λ from .39 to .56.

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

[36] Regarding the midpoints of the tail intervals, we also follow Gürkaynak and Wolfers (2006, p. 6, fn. 9): "For the tails we impute an upper- and lower-bound so that the midpoint would be equal to the mean of that bin if the PDF were normal."

[**TABLE VI:** RESULTS OF REGRESSIONS OF FORECAST ERRORS ON THE DIFFERENCE BETWEEN OBSERVED FORECAST AND 1/*N* FORECAST]

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 VI. 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 about .06.

Taken together, these calculations suggest a small degree of partition-dependence in all four ED market prices (a total sample of 153 separate markets), and a substantial degree of partition-dependence for two of the four statistics.

V. CONCLUSION

Partition-dependence is the finding that judged probabilities—as expressed by individuals directly or implied by market prices for event-contingent claims in prediction markets—vary systematically with the set of exclusive and exhaustive events into which a state space happens to be "partitioned". This phenomenon was first discovered in a cumulative series of psychology experiments beginning in the late 1970s. The basic finding in those studies is that "unpacking" a single category or interval into two or more component intervals which are logically equivalent increases the original interval's total expressed probability. The present paper tackles the question of to what extent market forces are able to eliminate (or at least mitigate) partition-dependence observed in individual judgments. Our experiments show that the bias transfers to competitive markets and that the phenomenon is robust to variations in the events that are being judged, the self-selection of participants, the length of the markets, and whether the markets are experimentally-created or are created by large market firms.

Study 1 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 2 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%. Study 3 examined 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 partition-independent probabilities and an "ignorance prior" ($1/N$ 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.

We first 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 $I_1 \cup I_2$, arbitrage would bring the sum of event I_1 and I_2 prices into line with the price of $I_1 \cup I_2$.

However, in practice there is usually no reason why two different partitions would be created and traded simultaneously. Therefore, the relevant question is whether revealed prices 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 3 suggests the answer is "Yes", that such partition-dependence could influence prices.

If one prefers a single explanation for many different findings, the combination of experimental methods used in the three studies suggests a basic behavioral propensity to bias judgments over N intervals toward $1/N$ is a component of what is observed. No other plausible explanation can explain the results of all three studies. The apparent bias cannot be due to information conveyed by the choice of partition because the subjects in Studies 1 and 2 were told about both partitions (any information conveyed should affect both markets equally). The apparent bias is not likely to be entirely due to the nature of trading institutions since Studies 1 and 2 used double auctions and the Study 3 data come from a parimutuel auctions. The apparent bias cannot be entirely due to naïveté of subjects, since there is self-selection of active traders in Study 2 (NBA and World Cup) and Study 3 (economic derivatives) markets.

More generally, these studies suggest two important themes in thinking about the implication of psychology for economics. One theme is that attention is grabbed by salient presentations of intervals, and people do not spontaneously compensate for the effect of attention-grabbing. In a similar vein, 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%. Unpacking an interval into the two components increases attention to those components and seems to increase implied probability. Because there is typically no canonical normative way to partition a continuous variable, how the variable's possible outcomes are divided into intervals can inexorably influence perceptions of the likely value of that variable.

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, but its robustness and persistence over time should be explored in further studies.

UNIVERSITY OF MUENSTER
CALIFORNIA INSTITUTE OF TECHNOLOGY
UCLA ANDERSON SCHOOL OF MANAGEMENT
UNIVERSITY OF MUENSTER

REFERENCES

Ahn, David S., and Haluk Ergin, "Framing Contingencies," Working Paper, University of California, Berkeley and Massachusetts Institute of Technology, 2007.

Arrow, Kenneth J., "Risk Perception in Psychology and Economics," *Economic Inquiry,* 20 (1982), 1-9.

Bardolet, David, Craig R. Fox, and Dan Lovallo, "Naïve Diversification and Partition Dependence in Capital Allocation Decisions," Working Paper, UCLA Anderson School, 2007.

Baron, Ken, and Jeffrey Lange, "From Horses to Hedging," *Risk*, 16 (2003), 73-77.

Benartzi, Shlomo, and Richard H. Thaler, "Naive Diversification Srategies in Defined Contribution Saving Plans," *American Economic Review*, 91 (2001), 79-98.

Berg, Joyce E., and Thomas A. Rietz, "The Iowa Electronic Market: Lessons Learned and Answers Yearned," in Paul Tetlock, and Robert Litan, eds., *Information Markets: A New Way of Making Decisions on the Public and Private Sectors* (Washington, D.C.: AEI-Brookings Joint Center, 2005).

Camerer, Colin F., "Do Biases in Probability Judgement Matter in Markets? Experimental Evidence," *American Economic Review*, 77 (1987), 981-997.

Camerer, Colin F., and Ernst Fehr, "When Does 'Economic Man' Dominate Social Behavior?" *Science*, 311 (2006), 47-52.

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

Chen, Kay-Yut, and Charles R. Plott, "Information Aggregation Mechanisms: Concept, Design and Implementation for a Sales Forecasting Problem," Social Science Working Paper No. 1131, California Institute of Technology, 2002.

Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini, "Recent Developments in Modeling Unforeseen Contingencies," *European Economic Review*, 42 (1998), 523-542.

Dubé-Rioux, Laurette, and J. Edward Russo, "An Availability Bias in Professional Judgement," *Journal of Behavioral Decision Making*, 1 (1988), 223-237.

Fehr, Ernst, and Jean-Robert Tyron, "Individual Irrationality and Aggregate Market Outcomes," *Journal of Economic Perspectives*, 19 (2005), 43-66.

Fischhoff, Baruch, Paul Slovic, and Sarah Lichtenstein, "Fault Trees: Sensitivity of Estimated Failure Probabilities to Problem Representation," *Journal of Experimental Psychology: Human Perception and Performance*, 4 (1978), 330-344.

Forsythe, Robert, Forrest Nelson, George R. Neumann, and Jack Wright, "Anatomy of an Experimental Political Stock Market," *American Economic Review*, 82 (1992), 1142-1161.

Fox, Craig R., and Robert T. Clemen, "Subjective Probability Assessment in Decision Analysis: Partition Dependence and Toward the Ignorance prior," *Management Science*, 51 (2005), 1417-1432.

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

Fox, Craig R., Rebecca K. Ratner, and Daniel Lieb, "How Subjective Grouping of Options Influences Choice and Allocation: Diversification Bias and the Phenomenon of Partition Dependence," *Journal of Experimental Psychology*, 134 (2005), 538-551.

Fox, Craig R., and Yuval Rottenstreich, "Partition Priming in Judgment under Uncertainty," *Psychological Science*, 14 (2003), 194-200.

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

Gilovich, Thomas, Dale W. Griffin, and Daniel Kahneman, *Heuristics and Biases: The Psychology of Intuitive Judgment* (Cambridge, UK: Cambridge University Press, 2002).

Gürkaynak, Refet, and Justin Wolfers, "Macroeconomic Derivatives: An Initial Analysis of Market-Based Macro Forecasts, Uncertainty, and Risk," Working Paper, National Bureau of Economic Research, 2006.

Hirt, Edward R., and N. John Castellan, "Probability and Category Redefinition in the Fault Tree Paradigm," *Journal of Experimental Psychology: Human Perception and Performance*, 14 (1988), 122-131.

Ho, Teck-Hua, and Kay-Yut Chen, "New Product Blockbusters: The Magic and Science of Prediction Markets," *California Management Review*, 50 (2007), 144-158.

Humphrey, Steven J., "Do Anchoring Effects Underlie Event-Splitting Effects? An Experimental Test," *Economic Letters*, 51 (1996), 303-308.

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

Kahneman, Daniel, Paul Slovic, and Amos Tversky, *Judgment under Uncertainty: Heuristics and Biases* (Cambridge, UK: Cambridge University Press, 1982).

Kreps, David M., "A Representation Theorem for "Preference for Flexibility"," *Econometrica*, 47 (1979), 565-577.

Langer, Thomas, and Craig R. Fox, "Biases in Allocation under Risk and Uncertainty: Partition Dependence, Unit Dependence, and Procedure Dependence," Working Paper, University of Muenster, University of California at Los Angeles, 2005.

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

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

Ofir, Chezy, "Ease of Recall vs Recalled Evidence in Judgment: Experts vs Laymen," *Organizational Behavior and Human Decision Processes*, 81 (2000), 28-42.

Page, Susan, "Confronting Iraq," *USA Today* (2003), April 1, 2003.

Plott, Charles R., and Kevin A. Roust, "The Design of Information Aggregation Mechanisms: A Two-Stage Parimutuel Market to Avoid Mirages Bubbles," Working Paper, California Institute of Technology, 2005.

Plott, Charles R., and Shyman Sunder, "Efficiency of Experimental Security Markets with Insider Information: An Application of Rational-Expectations Models," *Journal of Political Economy*, 90 (1982), 663-697.

Russo, J. Edward, and Karen Kolzow, "Where Is the Fault in Fault Trees?" *Journal of Experimental Psychology: Human Perception and Performance*, 20 (1994), 17-32.

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

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

Sher, Shlomi, and Craig R. M. McKenzie, "Information leakage from logically equivalent frames," Cognition, 101 (2006), 467-494.

Sunder, Shyam, "Experimental Asset Markets: A Survey," in John H. Kagel, and Alvin E. Roth, eds., *The Handbook of Experimental Economics* (Princeton: University Press, 1995), pp. 445-500.

Tversky, Amos, and Daniel Kahneman, "Rational Choice and the Framing of Decisions," *Journal of Business,* 59 (1986), 251-278.

——, "Availability: A Heuristic for Judging Frequency and Probability," *Cognitive Psychology*, 5 (1973), 207-232.

Tversky, Amos, and Derek J. Koehler, "Support Theory: A Nonextensional Representation of Subjective Probability," *Psychological Review*, 101 (1994), 547-567.

Van der Pligt, J., J.R. Eiser, and R. Speark, "Comparative Judgments and Preferences: The Influence of the Number of Response Alternatives," *British Journal of Social Psychology*, 26 (1987), 269-280.

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

——, "Five Open Questions about Prediction Markets," Working Paper, The Wharton School - University of Pennsylvania, Stanford GSB, 2005a.

——, "Interpreting Prediction Market Prices as Probabilities," Working Paper, The Wharton School - University of Pennsylvania, Stanford GSB, 2005b.

TABLE I

AVERAGE PRE-TRADING INDIVIDUAL PROBABILITY JUDGMENTS

		event domain		
		finance	**weather**	**sports**
treatment	**pre-trading ind. judgm.**	**mean**	**mean**	**mean**
1	$p(I_1)$	0.219	0.144	0.279
1	$p(I_2)$	0.497	0.333	0.398
	sum $p(I_1)+p(I_2)$	*0.717*	*0.477*	*0.678*
2	$p(I_1 \cup I_2)$	0.405	0.199	0.417
	PD difference	***0.312***	***0.278***	***0.261***
2	$p(I_3)$	0.397	0.349	0.378
2	$p(I_4)$	0.198	0.451	0.205
	sum $p(I_3)+p(I_4)$	*0.595*	*0.801*	*0.583*
1	$p(I_3 \cup I_4)$	0.283	0.523	0.322
	PD difference	***0.312***	***0.278***	***0.261***

TABLE II

TRADING VOLUME STATISTICS IN STUDY 1

	No. of trades per market (min – mean – max)	Average no. shares per trade (min – mean – max)	Total shares traded per market (min – mean – max)
All	21 – 42.83 – 66	1.93 – 3.39 – 5.90	78 – 143.85 – 284
Finance	21 – 43.08 – 62	1.93 – 3.37 – 5.42	81 – 144.17 – 284
Weather	26 – 41.40 – 64	2.31 – 3.49 – 5.90	85 – 143.48 – 242
Sports	25 – 44.00 – 66	2.07 – 3.31 – 4.37	78 – 143.92 – 210

TABLE III

MEAN EQUILIBRIUM PRICES (2ND TRADING ROUND) AND INDIVIDUAL JUDGMENTS (POST-TRADING AND PRE-TRADING)

treatm.	**mean prob./prices**	equil. prices 2nd round	post-trading judgment	equil. prices 2nd round	post-trading judgment	equil. prices 2nd round	post-trading judgment
1	$p(I_1)$	0.152	0.205	0.048	0.116	0.230	0.252
1	$p(I_2)$	0.561	0.494	0.256	0.307	0.490	0.432
	sum $p(I_1)+p(I_2)$	*0.713*	*0.699*	*0.303*	*0.422*	*0.720*	*0.684*
2	$p(I_1 \cup I_2)$	0.424	0.442	0.149	0.196	0.439	0.428
	PD difference	***0.289***	***0.257***	***0.154***	***0.226***	***0.281***	***0.256***
	(pre-trading)		***(0.312)***		***(.278)***	***0.281***	***(0.261)***
2	$p(I_3)$	0.404	0.382	0.354	0.352	0.416	0.403
2	$p(I_4)$	0.177	0.176	0.496	0.452	0.152	0.169
	sum $p(I_3)+p(I_4)$	*0.581*	*0.558*	*0.850*	*0.804*	*0.568*	*0.572*
1	$p(I_3 \cup I_4)$	0.336	0.301	0.707	0.578	0.391	0.316
	PD difference	***0.245***	***0.257***	***0.143***	***0.226***	***0.177***	***0.256***
	(pre-trading)		***(0.312)***		***(.278)***	***0.177***	***(0.261)***

TABLE IVa

PARTITION-DEPENDENCE IN PRE-TRADING JUDGMENTS FOR NBA EVENTS

Δ_{Median}	Whole population (N=302x4), German and U.S. subjects (pooled).			
Team	event 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***	36/37
CLE	2.0	20.0***	15.0***	36/37
DAL	0.0	24.5***	20.0***	37/40
DEN	-2.0	17.5***	18.0***	34/41
DET	2.5	25.0***	36.5***	34/41
IND	-10.0	25.0***	12.0***	34/41
LAC	5.0	5.0	10.0***	49/51
LAL	-5.0	22.5***	10.0**	37/40
MEM	15.0	25.0***	23.0***	36/37
MIA	-10.0	20.0***	15.0***	37/40
MIL	-5.0	20.0***	10.0**	49/51
NJN	-10.0*	30.0***	20.0***	49/51
PHX	0.0	30.0***	27.5***	36/37
SAC	-12.5	25.0***	2.0*	34/41
SAS	0.0	30.0***	40.0***	49/51
WAS	-10.0	10.0***	10.0	37/40

Notes. 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 team. N_1 (N_2) indicates the number of subjects in partition 1 (partition 2) that provided probability judgments for the team. Each subject (N=302) provided judgments for four different teams resulting in 1,208 judgments in total. *, **, *** indicate significance at the 10%, 5%, and 1% level, respectively (two-tailed) based on a Kruskal-Wallis test for each team.

TABLE IVb

PARTITION-DEPENDENCE IN PRE-TRADING JUDGMENTS FOR FIFA EVENTS

Δ_{Median}	Whole population (N=263x4), German subjects.			
Team	event 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***	34/30
AUS	-5.0	0.0	5.0	34/33
BRA	0.0	22.0***	20.0***	34/33
CIV	6.5	10.0*	8.5***	34/30
CRC	0.0	15.0*	9.0***	32/35
CRO	10.0**	10.0**	20.0***	34/33
CZE	8.0*	35.0***	37.5***	30/35
ECU	10.0	10.0	9.0***	32/35
GER	0.0	35.0***	37.5***	32/35
GHA	-2.5	12.5**	5.0*	30/35
ITA	4.0	25.0***	27.5***	30/35
JPN	5.0**	0.0	6.0***	34/33
NED	0.0	32.5***	30.0***	34/30
POL	5.0	30.0***	30.0***	32/35
SCG	10.0	0.5*	14.0***	34/30
USA	0.0	15.0***	7.0***	30/35

Notes. See Table IVa notes. Each subject (*N*=263) provided judgments for four different teams resulting in 1,052 judgments in total.

TABLE V

PER-DAY PROFITABILITY OF INTERPOLATED-PRICE HYPOTHETICAL PSEUDO-ARBITRAGE STRATEGIES

	low intervals		high intervals	
Team	Arbitrage PD (sell 1,2, buy 12)	Arbitrage reverse PD (buy 1,2, sell 12)	Arbitrage PD (sell 3,4, buy 34)	Arbitrage reverse PD (buy 3,4, sell 34)
NBA				
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				
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

TABLE VI

RESULTS OF REGRESSIONS OF FORECAST ERRORS ON THE DIFFERENCE BETWEEN OBSERVED FORECAST AND $1/N$ FORECAST

		regression (3) results			Implied weight λ on $1/N$	
	no. events	Coef $-\frac{\lambda}{1-\lambda}$	t-stat.	p-value (one-tailed)	λ implied by regression	Error-minimization
All statistics pooled	153	–.77	–2.60	.01	.44	.06
Initial unemployment claims	64	+.13	0.16	.44	–.15	.04
Business confidence	30	–.64	–1.88	.04	.39	.08
Non-form payrolls	33	–1.29	–1.53	.07	.56	.57
Retail sales (excluding autos)	26	–1.01	–1.32	.10	.50	.50

Note. P-value for test of regression coefficient different than zero is one-tailed.

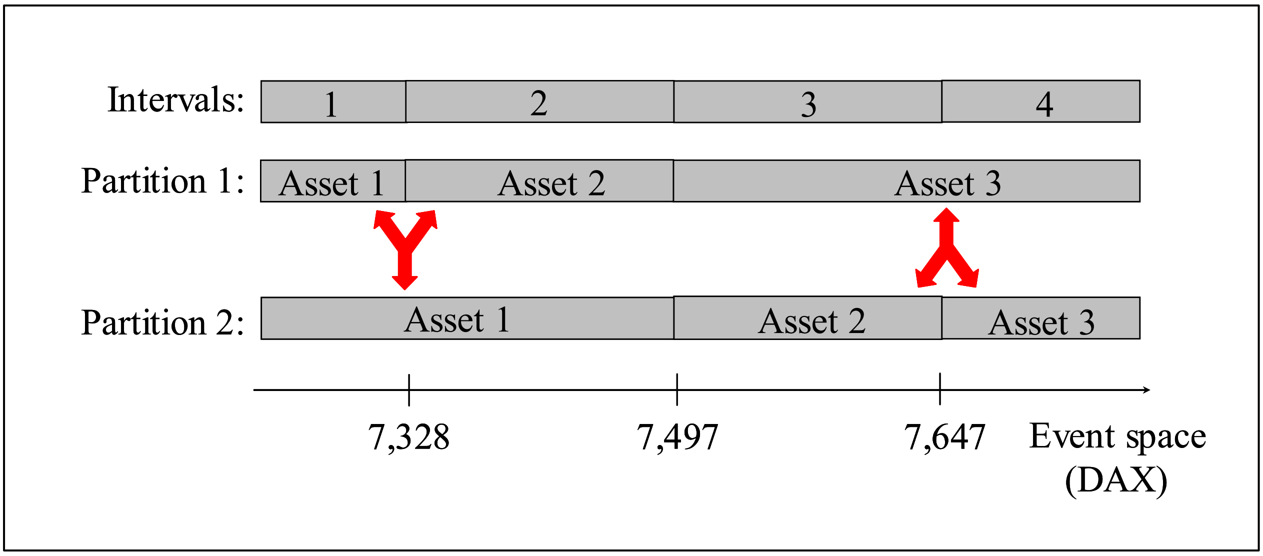


FIGURE I

Construction of Assets for the two DAX Partitions

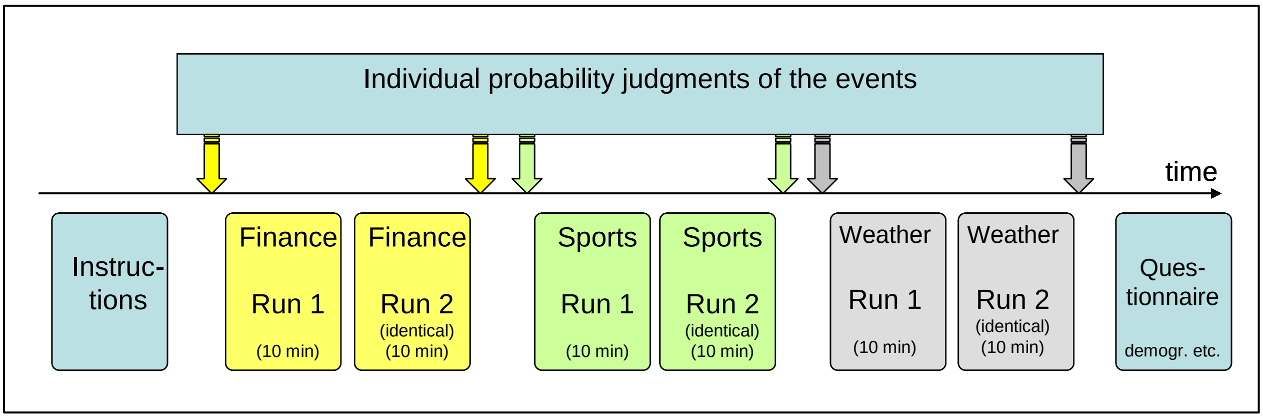


FIGURE II

Example of the Time Course of an Experimental Session

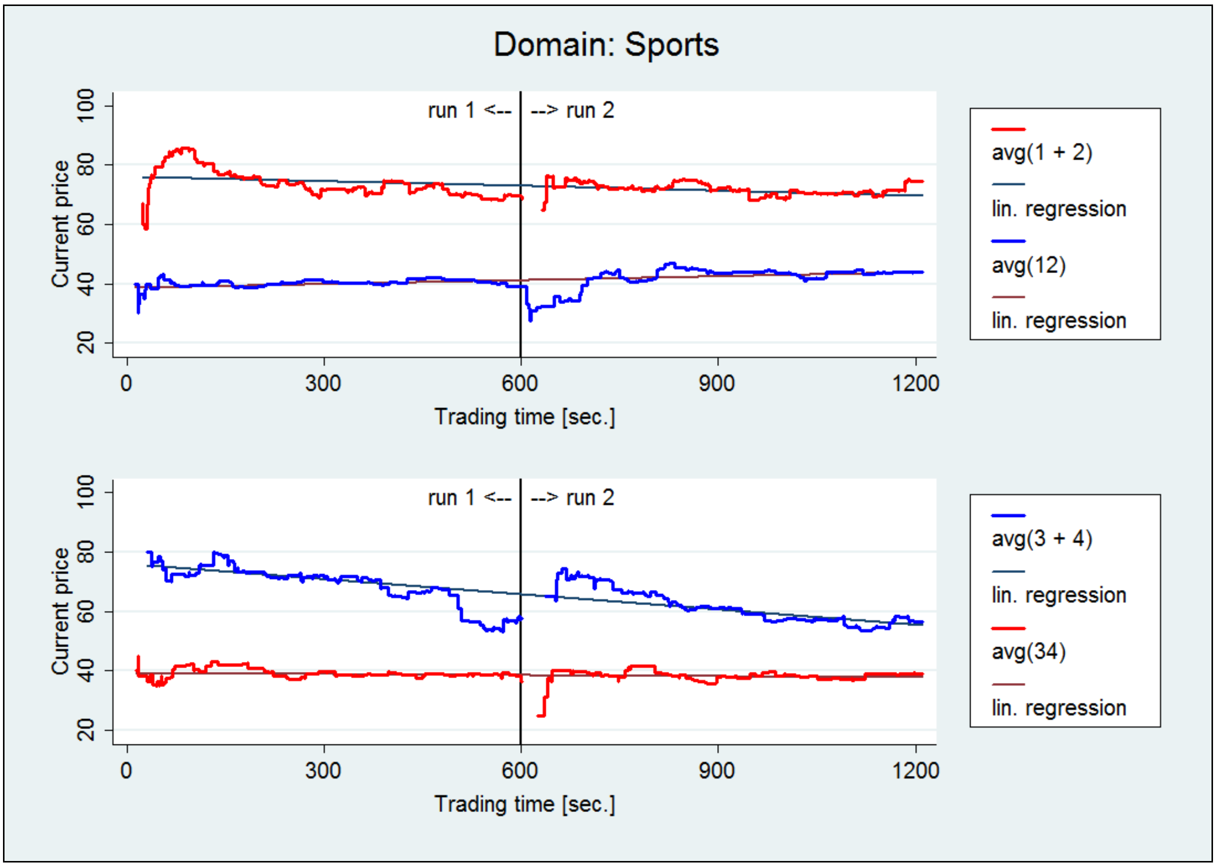


FIGURE III

Development of Price Differences over Time for the Sports Assets in Study 1

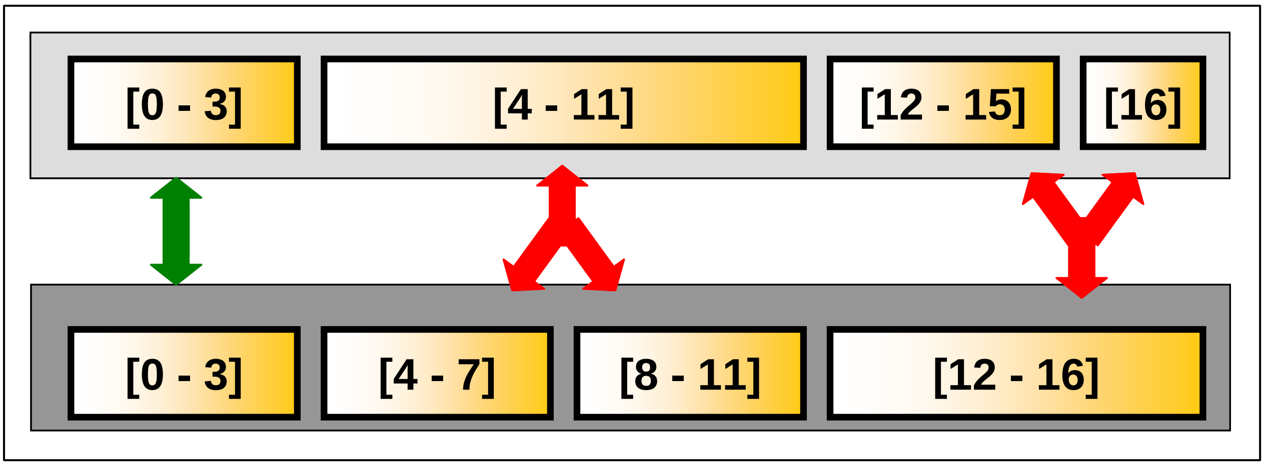


FIGURE IVa

Construction of Asset Partitions (NBA Victory Totals)

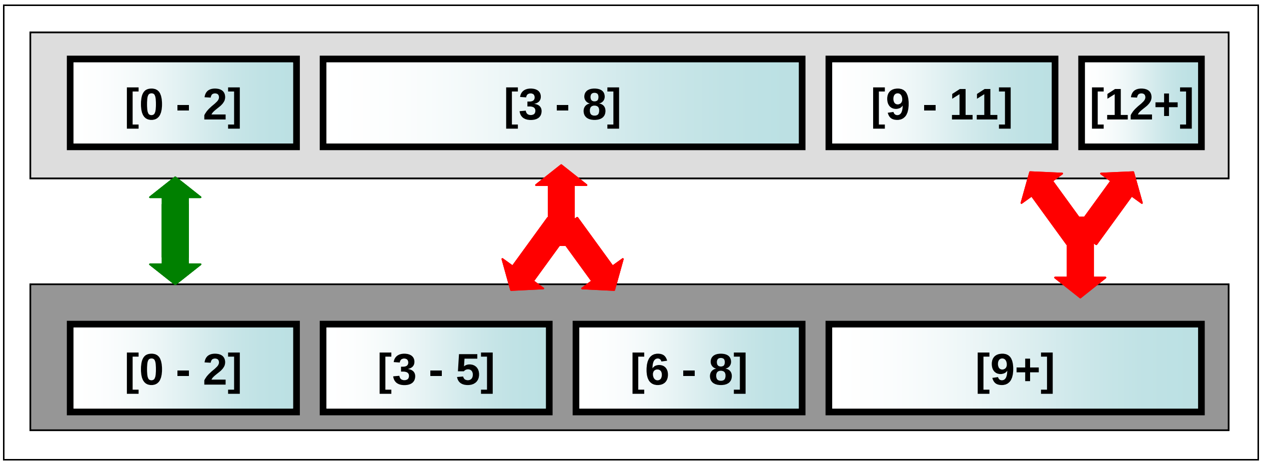


FIGURE IVb

Construction of Asset Partitions (World Cup Goal Totals)

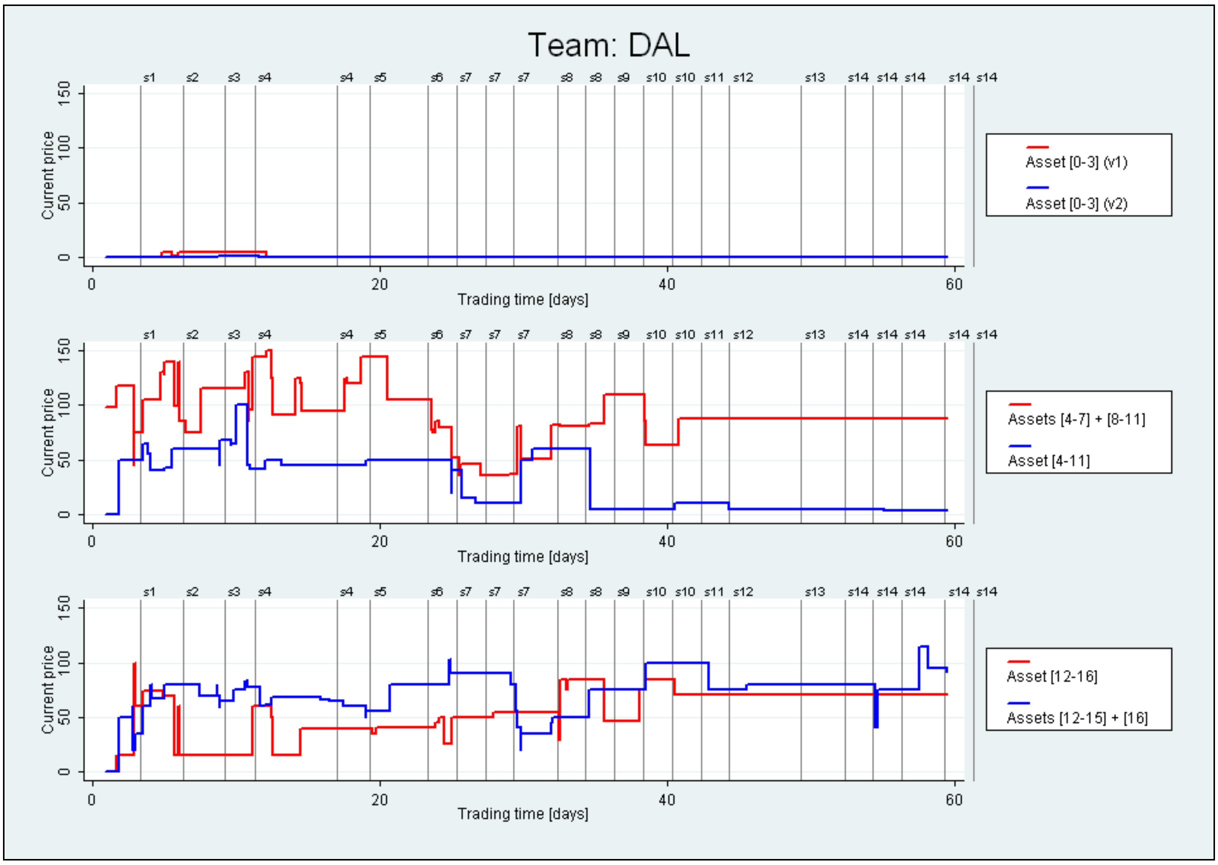


FIGURE Va

Price Chart (Dallas Mavericks, DAL)

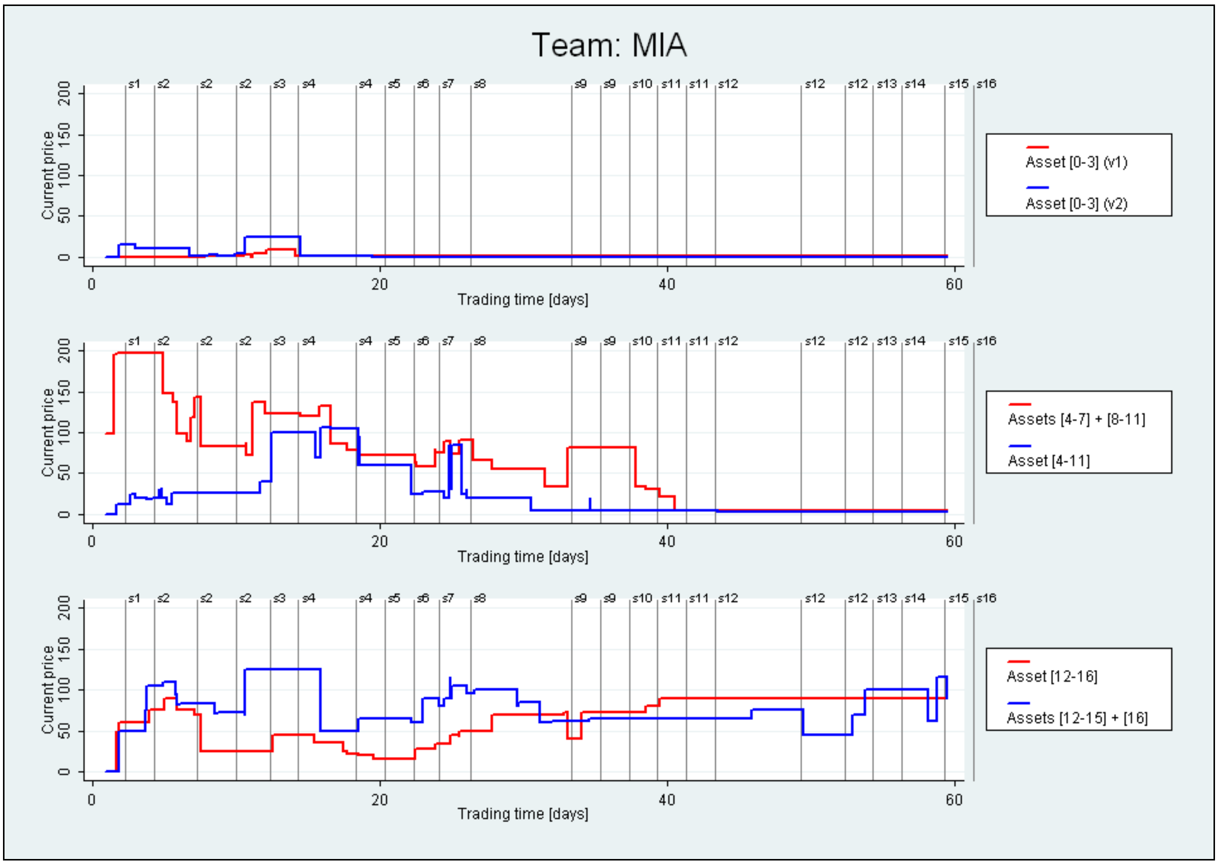


FIGURE Vb

Price Chart (Miami Heat, MIA)

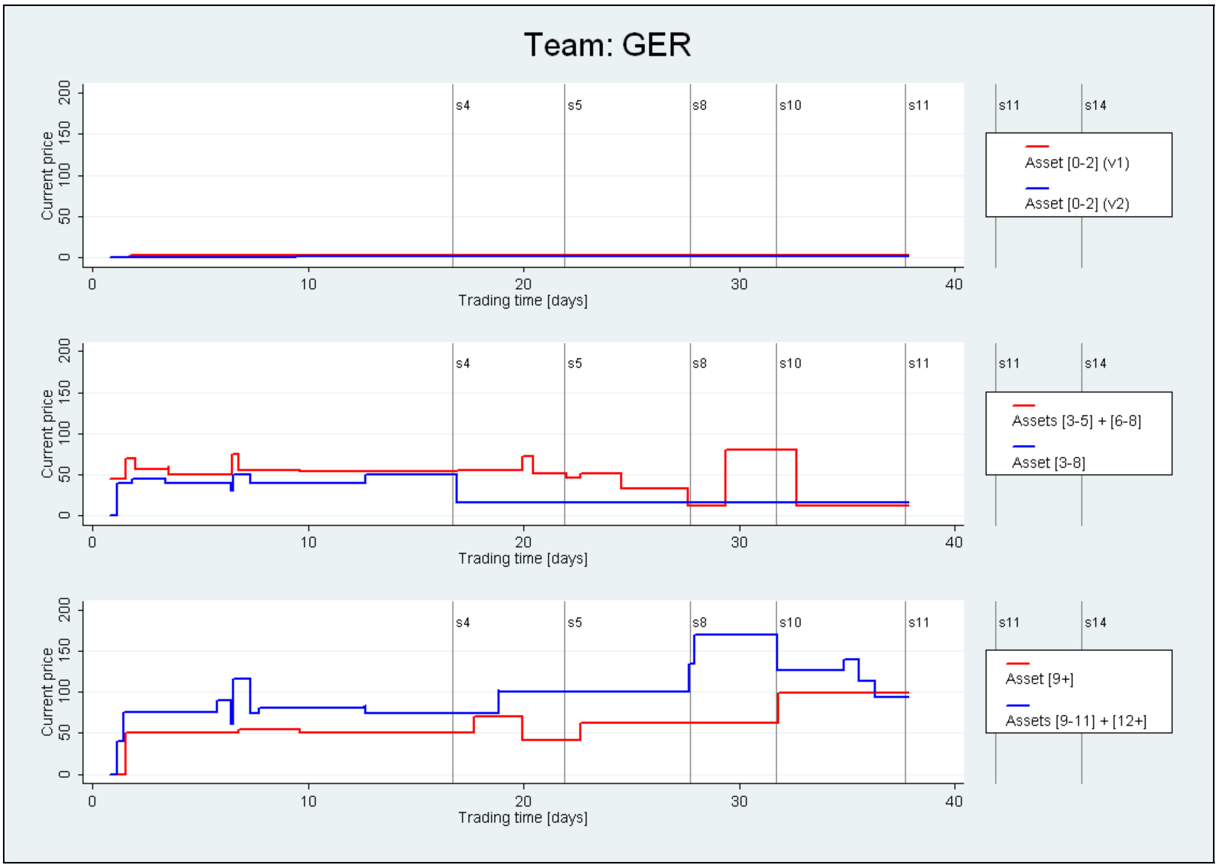


FIGURE VIa

Price Chart (Germany, GER)

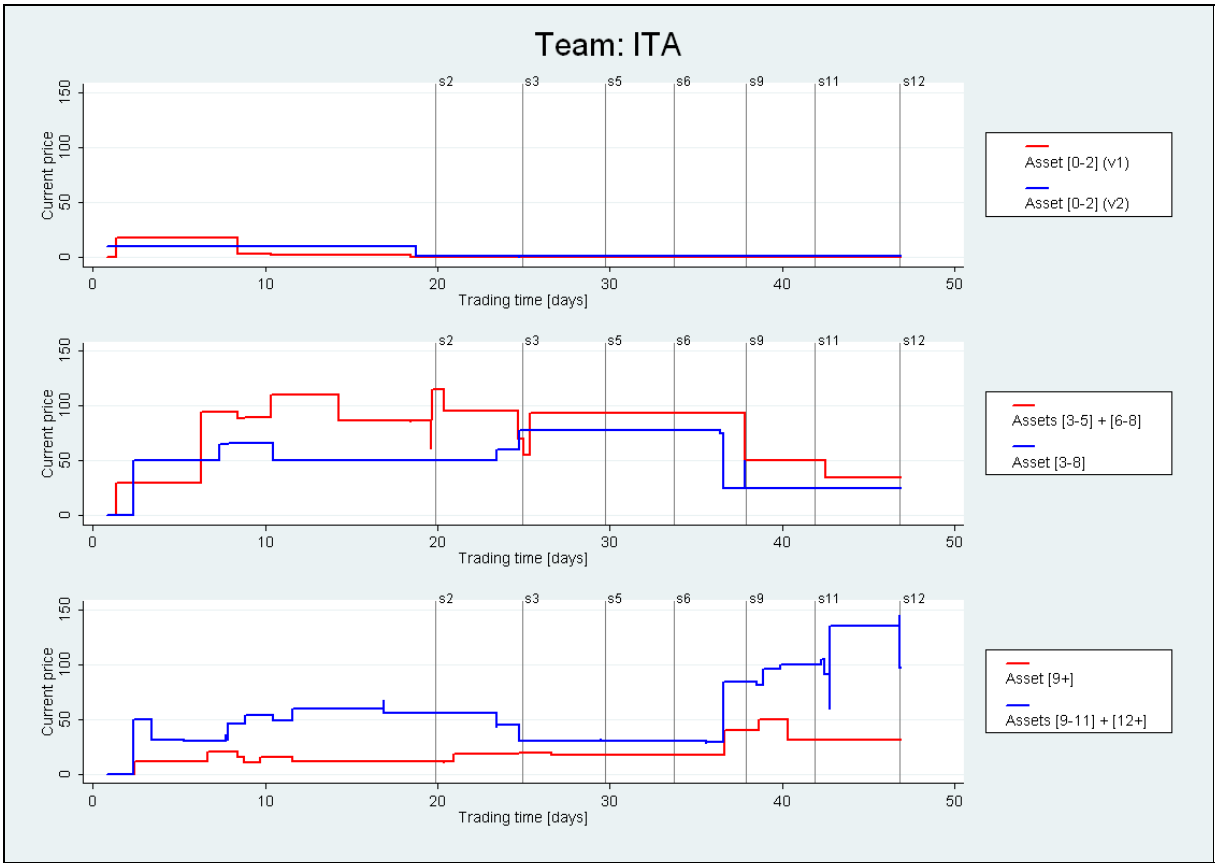


FIGURE VIb

Price Chart (Italy, ITA)

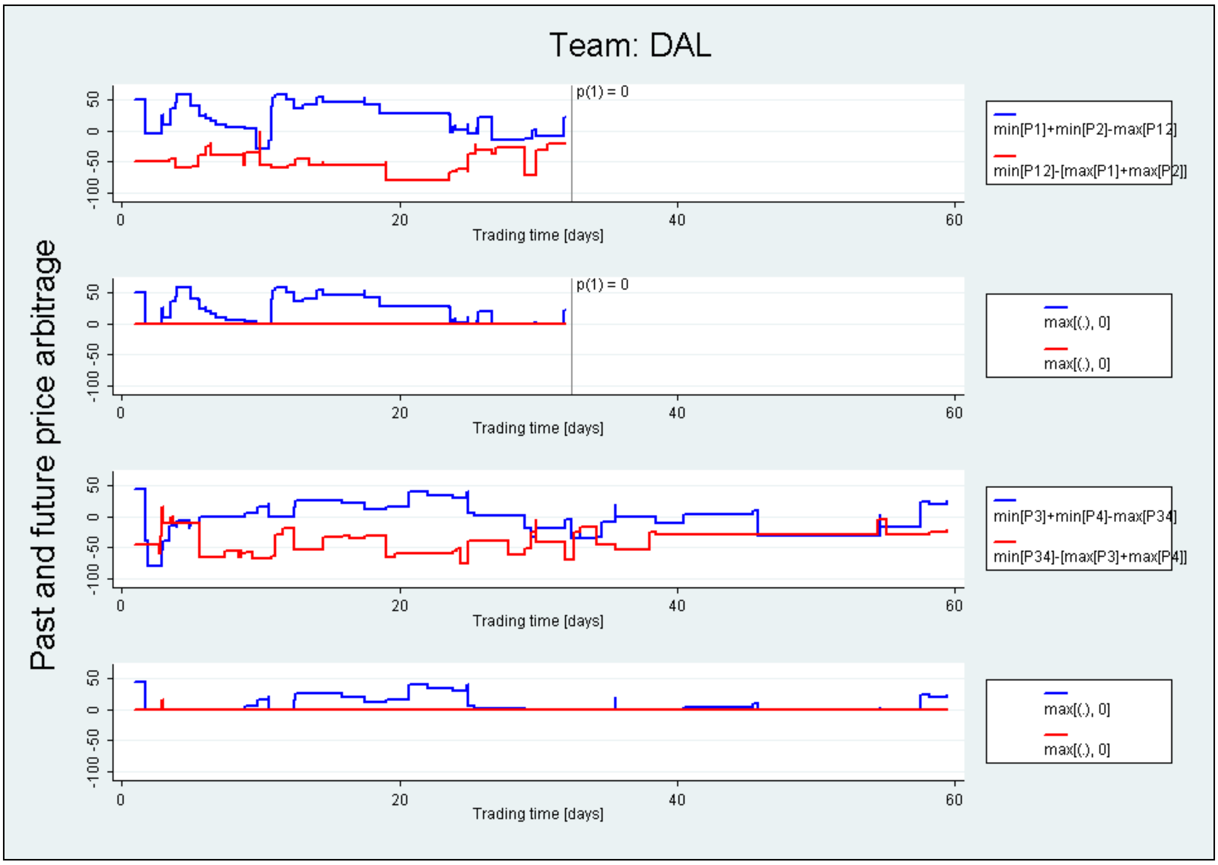


FIGURE VIIa

Interpolated-price Hypothetical Arbitrage (Dallas Mavericks, DAL)

FIGURE VIIb

Interpolated-price Hypothetical Arbitrage (Italy, ITA)

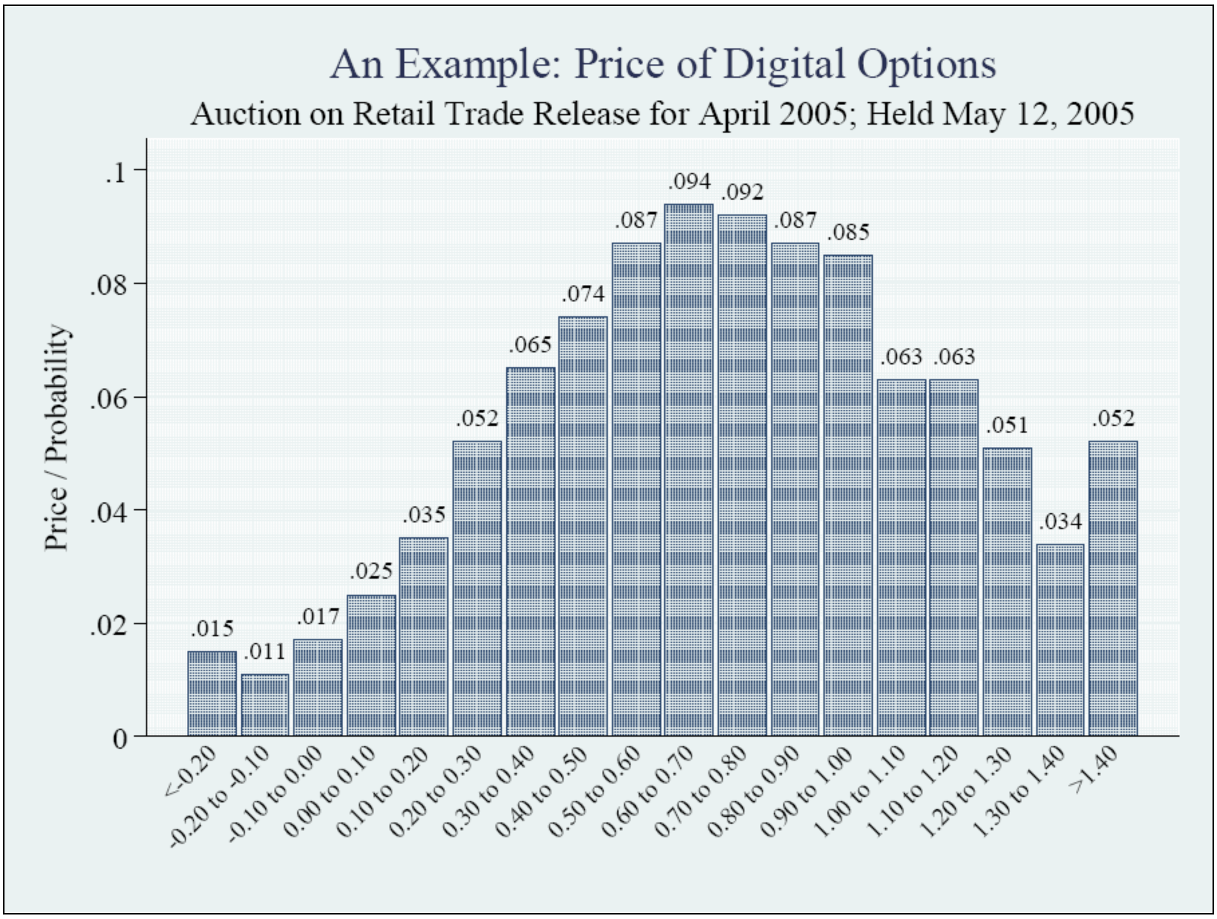


FIGURE VIII

“Digital Option” Prices on Ranges of Retail Trade Statistics (Gürkaynak and Wolfers 2006)

LIST OF APPENDICES:[37]

[37] Intended for online availability only, not intended for publication.

APPENDIX I: MAIN RESULTS FROM THE PILOT STUDY PRECEDING STUDY 1

This Table directly compares to Table III in the main text. The pilot study was almost identical to Study 1 (and results were quite similar) but had to be rerun due to a small flaw in the description of the events. Packed and unpacked events differed with respect to the inclusion of the boundary values. This is a design flaw because if subjects are highly confident that the numerical value is exactly at the boundary, then the interval prices for packed and unpacked events could rationally be different. For the DAX stock market (scaled in the hundreds) and for temperatures (rounded to the nearest 0.1°C) this boundary mismatch is probably a minor problem), but for the sports outcomes (number of goals) it creates an interpretive problem.

		event domain								
		finance			weather			sports		
treatm.	**mean** prob./prices	pre-trading ind. judgm.	equil. prices 2^{nd} round	post-trading ind. judgm.	pre-trading ind. judgm.	equil. prices 2^{nd} round	post-trading ind. judgm.	pre-trading ind. judgm.	equil. prices 2^{nd} round	post-trading ind. judgm.
1	I_1	0.244	0.245	0.247	0.199	0.159	0.180	0.287	0.270	0.268
1	I_2	0.434	0.399	0.432	0.358	0.386	0.377	0.393	0.377	0.410
	I_{1+2}	*0.678*	*0.644*	*0.679*	*0.557*	*0.545*	*0.557*	*0.680*	*0.646*	*0.678*
2	I_{12}	0.428	0.407	0.437	0.300	0.316	0.302	0.426	0.439	0.434
	PD difference	***0.250***	***0.237***	***0.242***	***0.257***	***0.229***	***0.255***	***0.254***	***0.207***	***0.244***
2	I_3	0.369	0.347	0.366	0.399	0.385	0.397	0.358	0.386	0.374
2	I_4	0.202	0.226	0.197	0.302	0.310	0.301	0.217	0.202	0.192
	I_{3+4}	*0.572*	*0.574*	*0.563*	*0.701*	*0.694*	*0.698*	*0.575*	*0.588*	*0.566*
1	I_{34}	0.322	0.344	0.321	0.443	0.438	0.443	0.320	0.331	0.321
	PD difference	***0.250***	***0.230***	***0.242***	***0.258***	***0.256***	***0.255***	***0.255***	***0.257***	***0.245***

		event domain								
		finance			weather			sports		
treatm.	**median** prob./prices	pre-trading ind. judgm.	equil. prices 2^{nd} round	post-trading ind. judgm.	pre-trading ind. judgm.	equil. prices 2^{nd} round	post-trading ind. judgm.	pre-trading ind. judgm.	equil. prices 2^{nd} round	post-trading ind. judgm.
1	I_1	0.200	0.272	0.210	0.200	0.118	0.200	0.285	0.296	0.245
1	I_2	0.400	0.394	0.400	0.350	0.396	0.400	0.355	0.355	0.400
	I_{1+2}	*0.700*	*0.647*	*0.700*	*0.575*	*0.576*	*0.600*	*0.700*	*0.658*	*0.668*
2	I_{12}	0.400	0.396	0.400	0.300	0.336	0.300	0.400	0.393	0.400
	PD difference	***0.300***	***0.251***	***0.300***	***0.275***	***0.240***	***0.300***	***0.300***	***0.265***	***0.268***
2	I_3	0.333	0.342	0.350	0.400	0.351	0.400	0.350	0.364	0.400
2	I_4	0.200	0.209	0.200	0.300	0.307	0.300	0.200	0.222	0.150
	I_{3+4}	*0.600*	*0.591*	*0.600*	*0.700*	*0.687*	*0.700*	*0.600*	*0.616*	*0.600*
1	I_{34}	0.300	0.361	0.300	0.425	0.424	0.400	0.300	0.333	0.332
	PD difference	***0.300***	***0.230***	***0.300***	***0.275***	***0.263***	***0.300***	***0.300***	***0.283***	***0.268***

APPENDIX II: INSTRUCTIONS FOR THE EXPERIMENT "STOCK MARKET IN THE LABORATORY" (TRANSLATED FROM GERMAN; ORIGINAL GERMAN INSTRUCTIONS AVAILABLE FROM THE AUTHORS)

See extra file.

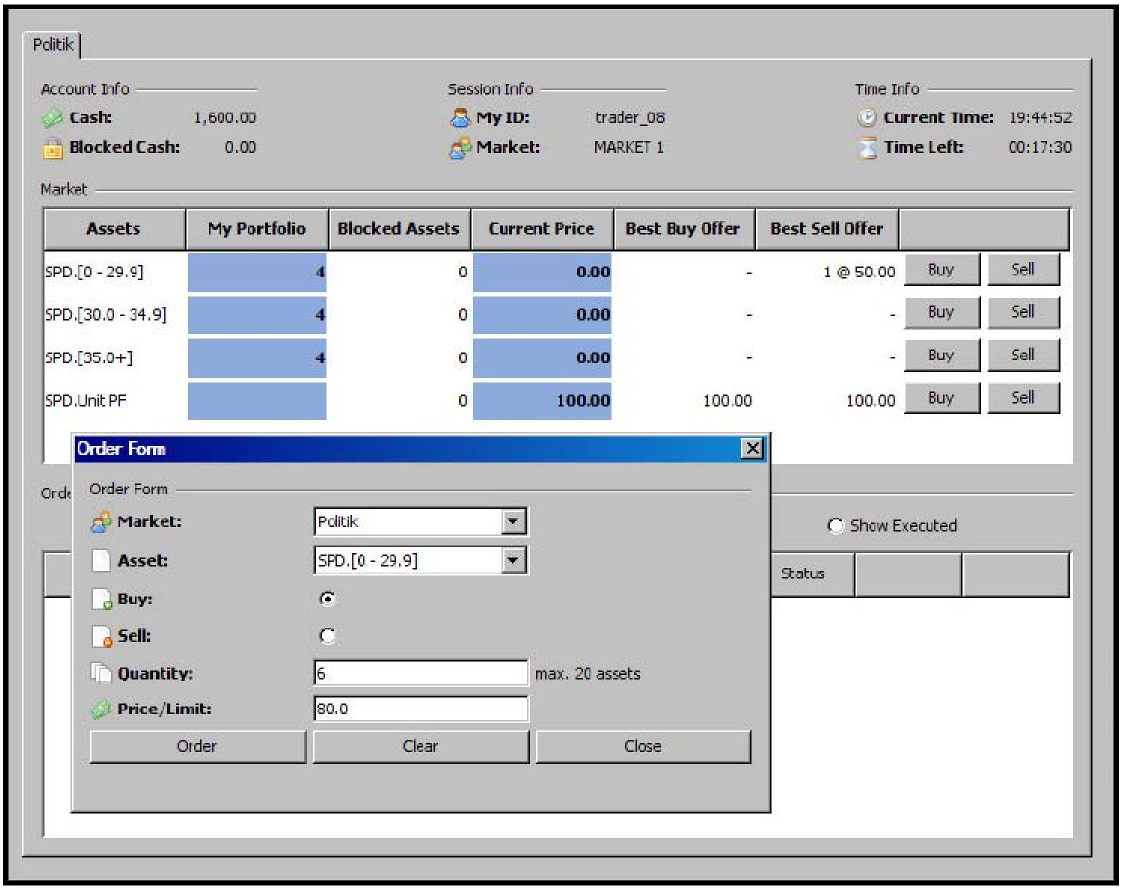
APPENDIX III: BOUNDARIES FOR DAX-PARTITIONS USED IN STUDY 1

These are all DAX boundaries used in the study

slot	exp. date	partition 1			partition 2		
		asset1	asset2	asset3	asset1	asset2	asset3
1	5/2/2007	[0 - 7327.99]	[7328 - 7496.99]	[7497+]	[0 - 7496.99]	[7497 - 7646.99]	[7647+]
2	4/24/2007	[0 - 7248.99]	[7249 - 7415.99]	[7416+]	[0 - 7415.99]	[7416 - 7563.99]	[7564+]
3	4/25/2007	[0 - 7181.99]	[7182 - 7347.99]	[7348+]	[0 - 7347.99]	[7348 - 7494.99]	[7495+]
4	4/26/2007	[0 - 7256.99]	[7257 - 7423.99]	[7424+]	[0 - 7423.99]	[7424 - 7572.99]	[7573+]
5	4/26/2007	[0 - 7256.99]	[7257 - 7423.99]	[7424+]	[0 - 7423.99]	[7424 - 7572.99]	[7573+]
6	4/27/2007	[0 - 7300.99]	[7301 - 7468.99]	[7469+]	[0 - 7468.99]	[7469 - 7618.99]	[7619+]
7	5/2/2007	[0 - 7327.99]	[7328 - 7496.99]	[7497+]	[0 - 7496.99]	[7497 - 7646.99]	[7647+]
8	4/25/2007	[0 - 7181.99]	[7182 - 7347.99]	[7348+]	[0 - 7347.99]	[7348 - 7494.99]	[7495+]
9	4/25/2007	[0 - 7181.99]	[7182 - 7347.99]	[7348+]	[0 - 7347.99]	[7348 - 7494.99]	[7495+]
10	4/26/2007	[0 - 7256.99]	[7257 - 7423.99]	[7424+]	[0 - 7423.99]	[7424 - 7572.99]	[7573+]
11	4/27/2007	[0 - 7300.99]	[7301 - 7468.99]	[7469+]	[0 - 7468.99]	[7469 - 7618.99]	[7619+]
12	4/27/2007	[0 - 7300.99]	[7301 - 7468.99]	[7469+]	[0 - 7468.99]	[7469 - 7618.99]	[7619+]

APPENDIX IV: SCREENSHOT AND FURTHER INFORMATION ON THE TRADING SOFTWARE

The trading software was exclusively developed for the study. It is based on Java Runtime Environment technology and was set up on a web-based client-server structure. The graphical user interface (GUI) was divided into three areas: information area, market area and order history (see screenshot below). Participants could submit, edit or cancel buy or sell orders via an order form. Orders were processed and executed by the system within split seconds; the trading screen was updated real-time.



APPENDIX V: MORE DETAILS ON ARBITRAGE OPPORTUNITIES IN STUDY 1

More details on arbitrage opportunities in Study 1:

No. of trading rounds/markets with not a single arbitrage opportunity

	Bid arbitrage			Ask arbitrage		
	Run 1	Run 2	Total	Run 1	Run 2	Total
Finance	3	3	6	13	10	23
Weather	2	7	9	15	17	32
Sports	3	2	5	15	17	32
Total	8	12	20	43	44	87

time period [sec.] with arb.opport. per 10-min. trading round (N=144)
(min--median--mean--max)

	Bid arbitrage			Ask arbitrage		
	Run 1	Run 2	Total	Run 1	Run 2	Total
Finance	0--47.30--84.01--296.11	0--37.36--70.46--305.72	0--37.36--77.24--305.72	0--0--22.95--106.84	0--6.19--28.15--188.90	0--2.89--25.55--188.90
Weather	0--50.90--82.71--436.26	0--46.02--75.24--375.25	0--48.93--78.98--436.26	0--0--16.46--225.88	0--0--37.85--484.85	0--0--27.16--484.85
Sports	0--72.87--85.25--210.18	0--59.68--75.20--230.82	0--62.63--80.22--230.82	0--0--20.75--268.31	0--0--31.47--395.86	0--0--26.11--395.86
Total	0--55.40--83.99--436.26	0--46.87--73.64--375.25	0--50.71--78.81--436.26	0--0--20.05--268.31	0--0--32.49--484.85	0--0--26.27--484.85

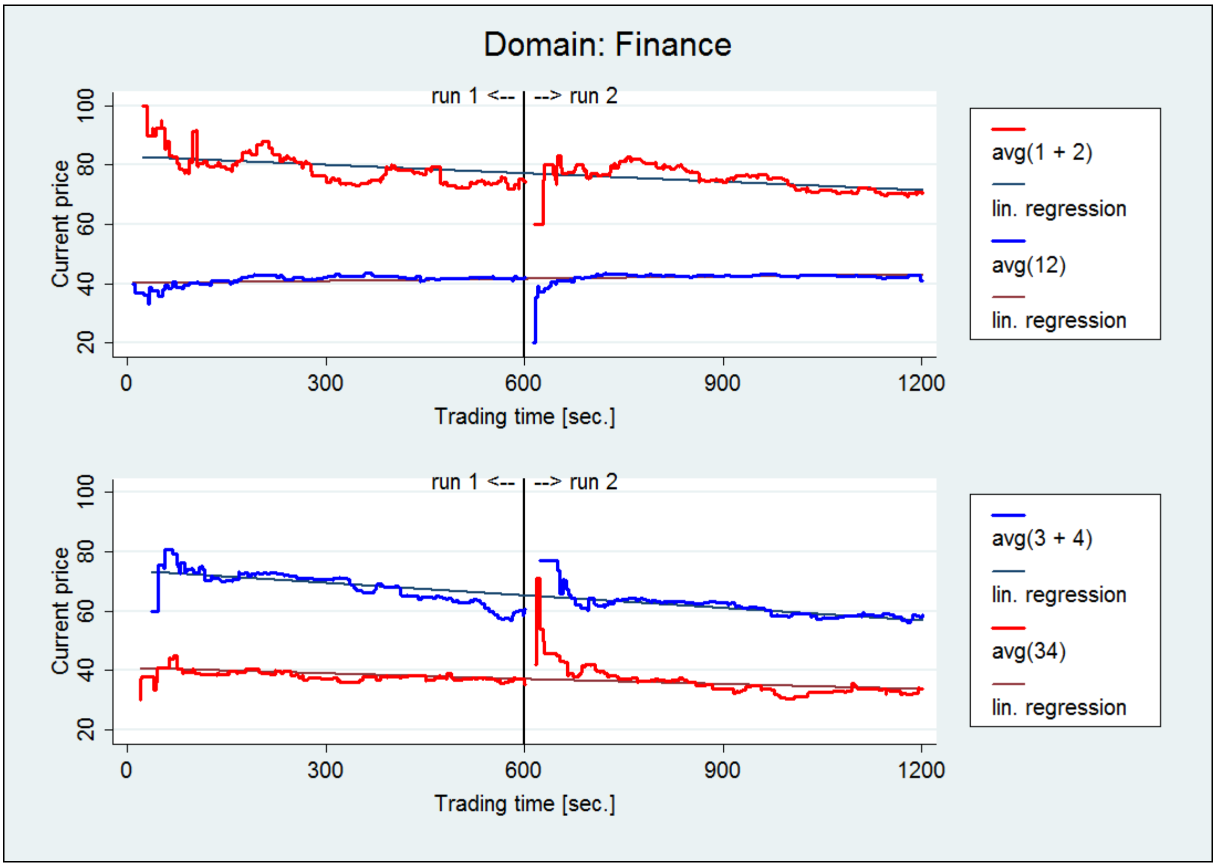
time period [sec.] until exploitation
(min--median--mean--max)

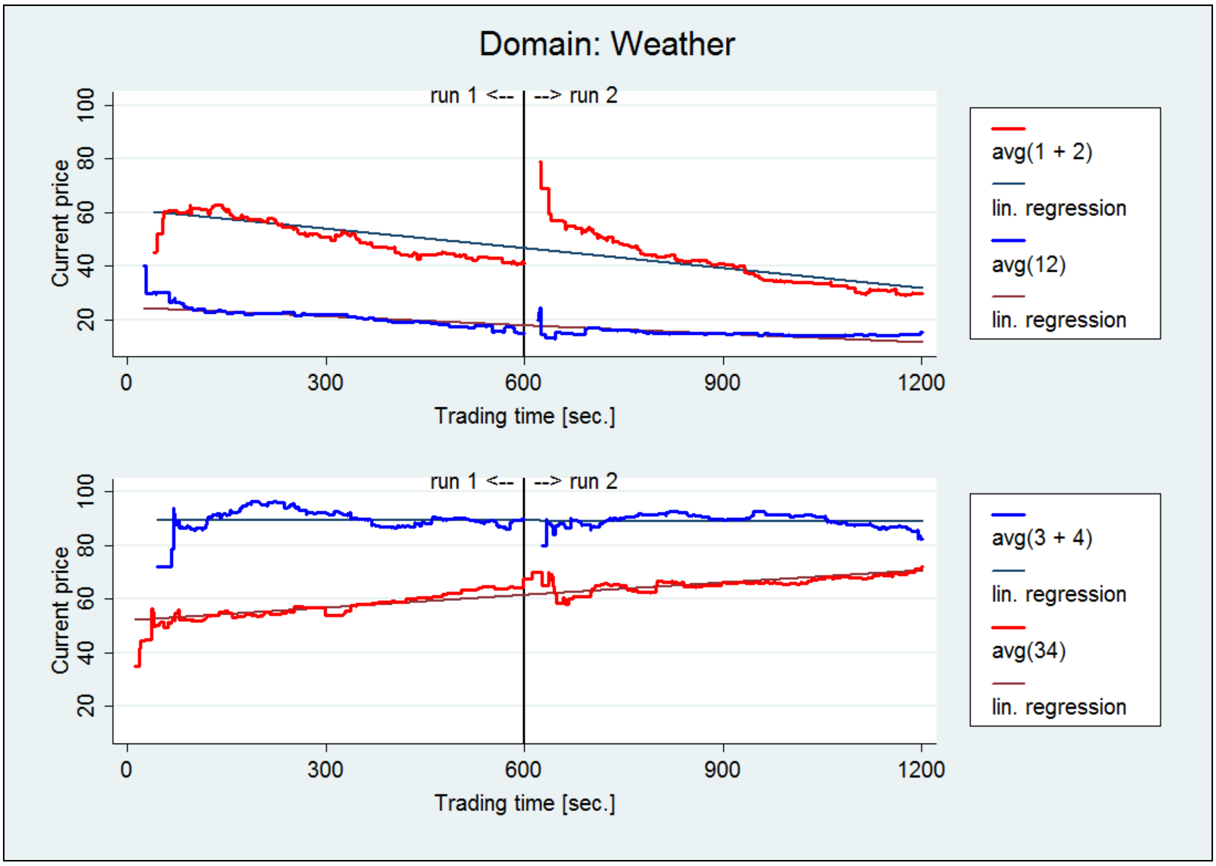
	Bid arbitrage			Ask arbitrage		
	Run 1	Run 2	Total	Run 1	Run 2	Total
Finance	1.16--15.46--26.88--141.87	1.07--8.43--21.14--208.87	1.07--12.13--23.92--208.87	1.85--10.05--20.40--101.43	1.24--9.68--21.80--185.09	1.24--9.86--21.15--185.09
Weather	1.23--9.95--28-77--369.73	1.00--12.59--24.74--188.38	1.00--11.45--26.70--369.73	1.05--6.69--21.95--180.26	1.08--20.77--47.81--455.34	1.05--9.58--35.23--455.34
Sports	1.22--16.18--29.65--133.64	1.05--12.42--19.00--126.36	1.05--12.75--23.48--133.64	1.85--10.08--33.20--151.67	1.40--15.61--29.05--334.24	1.40--13.08--30.57--334.24
Total	1.16--12.84--28.39--369.73	1.00--11.45--21.38--208.87	1.00--12.24--24.62--369.73	1.05--9.52--24.06--180.26	1.08--12.42--30.78--455.34	1.05--10.83--27.82--455.34

time-weighted amount per arbitrage occurrence
(min--median--mean--max)

	Bid arbitrage (N=461 occurrences)			Ask arbitrage (N=136 occurrences)		
	Run 1	Run 2	Total	Run 1	Run 2	Total
Finance	.49--5.60--11.85--119.00	.25--3.03--6.77--50.00	.25--5.00--9.22--119.00	.20--3.50--4.19--17.00	.38--2.00--3.24--13.77	.20--3.00--3.68--17.00
Weather	.05--5.00--6.94--33.00	.62--3.00--3.77--15.40	.05--3.91--5.31--33.00	.01--1.55--2.57--7.94	.81--3.56--8.72--37.00	.01--3.20--5.73--37.00
Sports	.62--8.38--10.54--44.99	.01--2.55--5.21--51.00	.01--4.44--7.45--51.00	.50--1.94--2.49--6.36	.50--2.65--4.42--15.63	.50--2.00--3.71--15.63
Total	.05--5.60--9.84--119.00	.01--3.00--5.29--51.00	.01--4.10--7.39--119.00	.01--2.13--3.28--17	.38--3.00--5.01--37.00	.01--3.00--4.25--37.00

APPENDIX VI: TIME SERIES OF MEAN DIFFERENCES BETWEEN PRICES FOR PACKED AND UNPACKED INTERVALS FOR FINANCE AND WEATHER EVENTS IN STUDY 1




Domain: Weather
Current price
run 1 <-- | --> run 2
Trading time [sec.]
avg(1 + 2)
lin. regression
avg(12)
lin. regression
Current price
run 1 <-- | --> run 2
Trading time [sec.]
avg(3 + 4)
lin. regression
avg(34)
lin. regression

Appendix VII: Median Equilibrium Prices (2nd Trading Round) and Individual Judgments (Pre-trading and Post-trading)

Median values to be compared with the mean values in Table III. The comparison is useful for checking whether means are influenced by a modest number of subjects.

		event domain								
		finance			weather			sports		
treatm.	**median prob./prices**	pre-trading ind. judgm.	equil. prices 2^{nd} round	post-trading ind. judgm.	pre-trading ind. judgm.	equil. prices 2^{nd} round	post-trading ind. judgm.	pre-trading ind. judgm.	equil. prices 2^{nd} round	post-trading ind. judgm.
1	I_1	0.200	0.129	0.200	0.120	0.051	0.100	0.250	0.244	0.250
1	I_2	0.500	0.563	0.500	0.320	0.244	0.300	0.400	0.469	0.400
	I_{1+2}	*0.700*	*0.714*	*0.700*	*0.500*	*0.271*	*0.400*	*0.700*	*0.716*	*0.700*
2	I_{12}	0.400	0.453	0.400	0.200	0.130	0.200	0.400	0.422	0.400
	PD difference	***0.300***	***0.261***	***0.300***	***0.300***	***0.141***	***0.200***	***0.300***	***0.294***	***0.300***
2	I_3	0.400	0.417	0.400	0.350	0.361	0.350	0.360	0.385	0.400
2	I_4	0.200	0.177	0.160	0.400	0.474	0.400	0.200	0.149	0.150
	I_{3+4}	*0.600*	*0.570*	*0.600*	*0.800*	*0.842*	*0.800*	*0.600*	*0.553*	*0.600*
1	I_{34}	0.300	0.358	0.300	0.500	0.690	0.600	0.300	0.414	0.300
	PD difference	***0.300***	***0.212***	***0.300***	***0.300***	***0.152***	***0.200***	***0.300***	***0.139***	***0.300***

APPENDIX VIII: NBA/FIFA STUDY 2 INSTRUCTIONS

See extra file.

APPENDIX IX: HYPOTHETICAL ARBITRAGE BASED ON AVAILABLE BIDS AND ASKS (STUDY 2)

The main text presents details of per-day profitability of a hypothetical arbitrage strategy based on recent and future prices. This appendix describes the analogous results for a hypothetical arbitrage using current bids and asks. That is, we ask whether there would be arbitrage opportunities if traders could actually trade in both markets, given the available bids and asks which can be used for trade. Keep in mind that participants could only trade in one market at a time, so they *could not actually execute these arbitrage trades* (which is why we refer to them as "pseudo-arbitrage"). Asking how large the opportunities are is simply a way to characterize the economic size of the partition-dependence, using all the information on bids and asks.

Consider intervals I_1 and I_2 which are traded separately (unpacked) in partition 2 and packed in partition 1. If there is partition-dependence, then the bids for assets I_1 and I_2 will be high (compared to bids for the packed asset $I_1 \cup I_2$). So one kind of pseudo-arbitrage is to take the sum of the current bids for assets I_1 and I_2 (i.e., the prices at which one could sell those assets) and to subtract the current ask for the equivalent asset $I_1 \cup I_2$ (i.e., the price at which one could buy that asset). If this difference is positive, then a trader with access to both markets could sell the two unpacked assets of intervals I_1 and I_2 for more than she could buy the packed interval asset $I_1 \cup I_2$. If there is reverse partition-dependence, then the opposite strategy would be profitable (buying the components and selling the packed asset). The size of these arbitrage strategies is represented in our notation as $B(I_1) + B(I_2) - A(I_1 \cup I_2)$ (i.e., selling unpacked and buying packed, arbitraging partition-dependence) and $B(I_1 \cup I_2) - [A(I_1) + A(I_2)]$ (selling packed and buying unpacked, arbitraging reverse partition-dependence) where $B(I_k)$ is the best (highest) bid quote for interval I_k and $A(I_k)$ is the best (lowest) ask quote for interval I_k.

Figure A.9a below shows these statistics over the life of the experiment for NBA team DAL. Look at the top panel first. The top panel shows $B(I_1) + B(I_2) - A(I_1 \cup I_2)$ (in blue) and $B(I_1 \cup I_2) - [A(I_1) + A(I_2)]$ (in red). The second panel shows the maxima of each of these spreads and zero (i.e., it only shows their values when they are positive, when pseudo-arbitrage is profitable). The blue spikes in the second panel indicate that there are pseudo-arbitrage opportunities, which are sometimes quite large in magnitude but are sporadic and usually short-lived. The red line at zero indicates that there is never a set of available bids and asks consistent with profitable arbitrage against reverse partition-dependence. The horizontal lines at the bottom of the second panel indicate the spans of time during which *any* bid or ask exists in the market for each of the assets in the arbitrage strategy. When those lines are interrupted there is no liquidity and hence no opportunity for arbitrage.[38]

The third and fourth panels show the same time series for the pseudo-arbitrage of intervals I_3 and I_4. There are frequent interruptions in the bid-ask existence series (at the bottom of the fourth panel), so pseudo-arbitrage opportunities are rare.

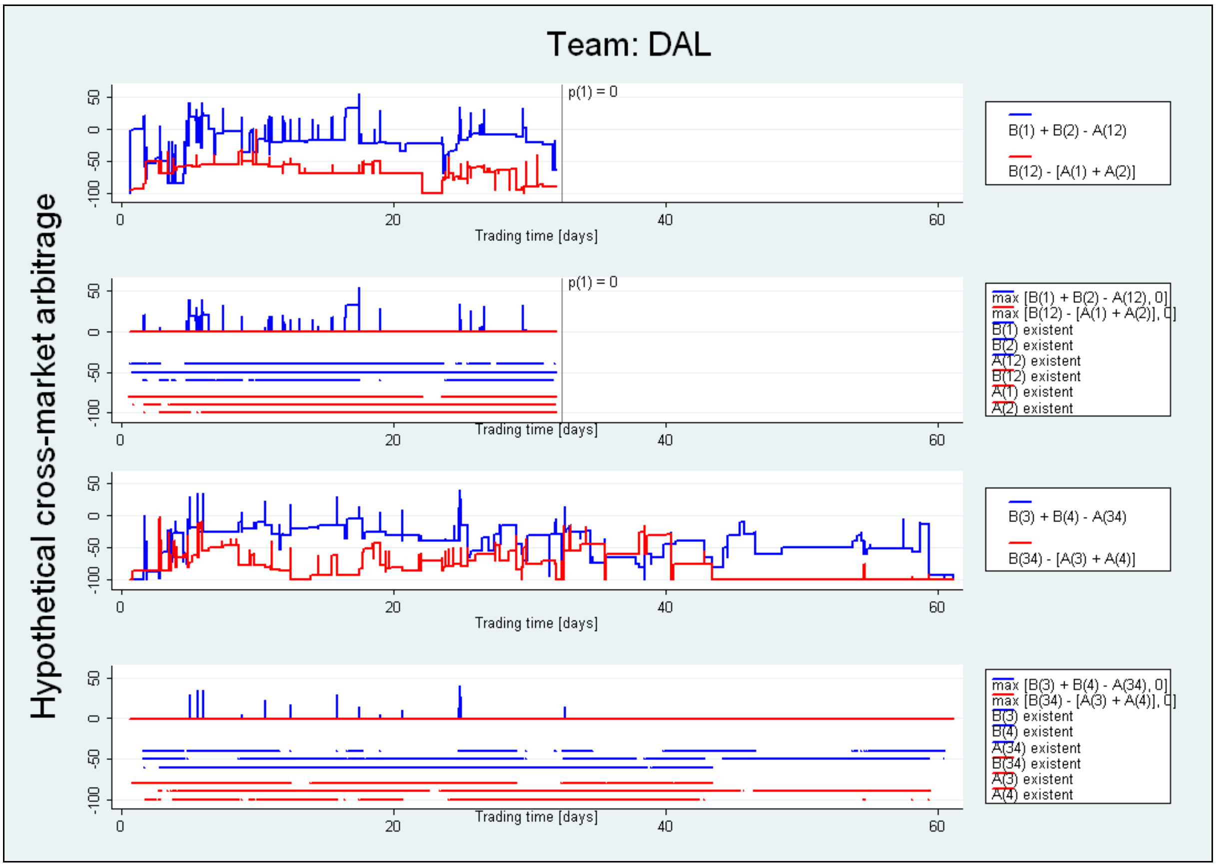


FIGURE A.9a

Cross-Market Pseudo-Arbitrage (Dallas Mavericks, DAL).

[38] Missing ask quotes were set to $+\infty$ and missing bid quotes were set to zero. The values of the positions are therefore calculated as

$\min[B(I_1) + B(I_2), 100] - \min[A(I_1 \cup I_2), 100]$ and $\min[B(I_1 \cup I_2), 100] - \min[A(I_1) + A(I_2), 100]$, and

$\min[B(I_3) + B(I_4), 100] - \min[A(I_3 \cup I_4), 100]$ and $\min[B(I_3 \cup I_4), 100] - \min[A(I_3) + A(I_4), 100]$.

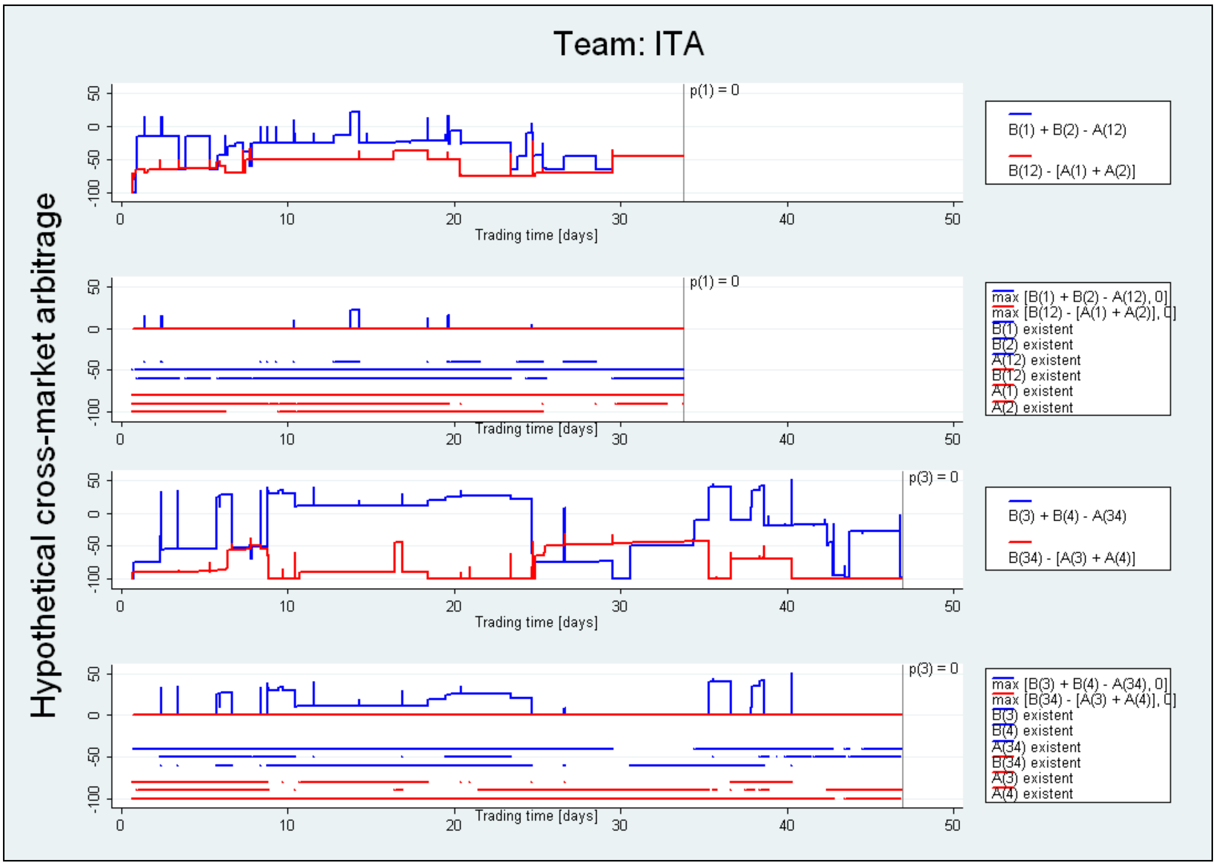


FIGURE A.9b

Cross-Market Pseudo-Arbitrage (Italy, ITA).

Figure A.9b shows the same time series for the World Cup winning team ITA. There are few pseudo-arbitrage opportunities for the low-goal intervals I_1 and I_2, but quite a bit of pseudo-arbitrage for intervals I_3 and I_4. From days 9 through 25, there is a persistent gap in the bids of unpacked events I_3 and I_4 and the ask for event $I_3 \cup I_4$. These examples illustrate the advantage of using the continuous bid and ask information. Trades are rather rare for ITA events (only about one trade per day across all assets) but bids and asks are common enough to show persistent gaps in (potential) prices.

Table A.1 reports the value of the time-weighted pseudo-arbitrage statistics for all teams. These are the area under the blue and red curves in the second and fourth panels of Figures A.9a and A.9b, divided by the total trading time.[39] The profitability of strategies exploiting partition-dependence (in columns (2) and (4)) is often very low, but is above 1.0 for 9 of 32 teams. Furthermore, pseudo-arbitrage against reverse partition-dependence is much less profitable. For 38 of the 46 team-partition comparisons, arbitraging against partition-

[39] Note that the relevant trading time ends either when the last auction for assets I_1, I_2 or $I_1 \cup I_2$ (I_3, I_4 or $I_3 \cup I_4$) occurred or when the corresponding interval asset I_1 (or I_3) expired worthless.

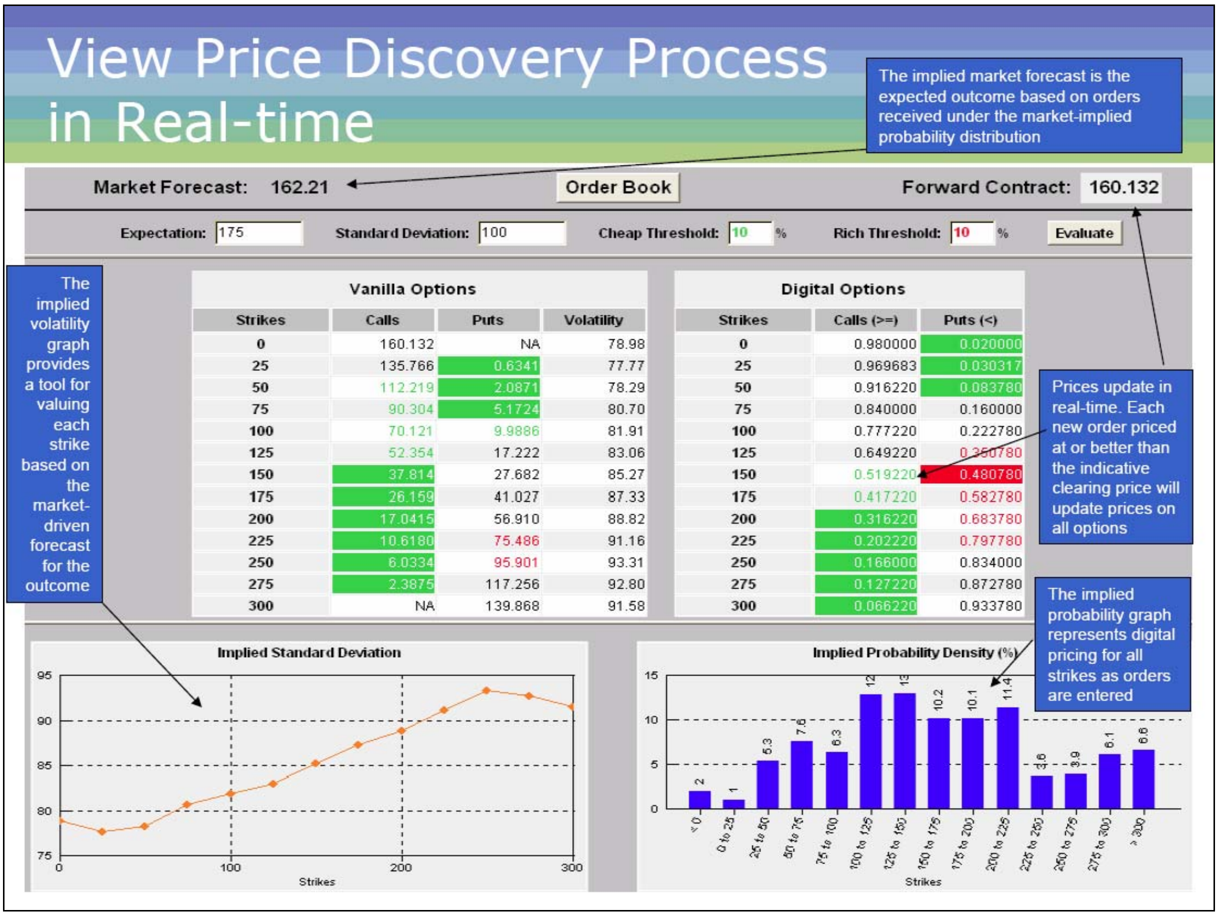
dependence is more profitable than arbitraging against reverse partition-dependence (excluding 18 team-partition cases in which both figures are zero), a fraction significantly lopsided by a conservative sign test (z=5.88, p<.001).

TABLE A.1

PER-DAY PROFITABILITY OF BID/ASK HYPOTHETICAL PSEUDO-ARBITRAGE STRATEGIES

	low intervals		high intervals	
Team	Arbitrage PD (sell 1,2, buy 12)	Arbitrage reverse PD (buy 1,2, sell 12)	Arbitrage PD (sell 3,4, buy 34)	Arbitrage reverse (buy 3,4, sell 34)
NBA				
CHI	0.00	0.04	0.06	0.00
CLE	0.01	0.00	0.11	0.00
DAL	1.95	0.00	0.03	0.00
DEN	0.00	1.08	0.00	0.00
DET	1.36	0.26	0.00	5.06
IND	0.12	0.00	0.20	0.00
LAC	0.16	0.00	0.71	0.00
LAL	0.00	0.00	5.14	0.00
MEM	0.00	0.00	2.61	0.00
MIA	0.96	0.00	0.75	0.00
MIL	0.00	0.04	0.00	0.00
NJN	0.65	0.00	7.90	0.00
PHX	0.25	0.00	1.46	0.00
SAC	0.00	0.00	0.00	0.00
SAS	0.44	0.00	0.24	0.00
WAS	0.05	0.03	0.00	0.00
FIFA				
ARG	0.05	0.00	0.15	0.00
AUS	0.55	0.01	0.48	0.00
BRA	0.00	0.11	0.07	0.00
CIV	0.00	0.00	0.00	0.00
CRC	0.00	0.06	0.00	0.00
CRO	0.03	0.00	0.00	0.00
CZE	0.64	0.00	9.93	0.00
ECU	0.00	0.00	0.00	0.00
GER	0.22	0.00	0.79	0.00
GHA	0.00	0.04	0.00	0.00
ITA	0.35	0.00	8.36	0.00
JPN	3.05	0.00	0.00	0.00
NED	0.29	0.00	0.00	0.00
POL	0.00	0.02	0.23	0.00
SCG	0.01	0.00	0.11	0.00
USA	0.00	0.00	0.00	0.00

APPENDIX X: SCREENSHOT OF TRADING INTERFACE FOR ECONOMIC DERIVATIVES MARKETS



Vanilla Options

Strikes	Calls	Puts	Volatility
0	160.132	NA	78.98
25	135.766	0.6341	77.77
50	112.219	2.0871	78.29
75	90.304	5.1724	80.70
100	70.121	9.9886	81.91
125	52.354	17.222	83.06
150	37.814	27.682	85.27
175	26.159	41.027	87.33
200	17.0415	56.910	88.82
225	10.6180	75.486	91.16
250	6.0334	95.901	93.31
275	2.3875	117.256	92.80
300	NA	139.868	91.58

Digital Options

Strikes	Calls (>=)	Puts (<)
0	0.980000	0.020000
25	0.969683	0.030317
50	0.916220	0.083780
75	0.840000	0.160000
100	0.777220	0.222780
125	0.649220	0.350780
150	0.519220	0.480780
175	0.417220	0.582780
200	0.316220	0.683780
225	0.202220	0.797780
250	0.166000	0.834000
275	0.127220	0.872780
300	0.066220	0.933780