

# Information Incorporation in Online In-Game Sports Betting Markets

Sandip Debnath,<sup>1,4</sup> David M. Pennock,<sup>2,4</sup> C. Lee Giles,<sup>1,3,4</sup> and Steve Lawrence<sup>4</sup>

<sup>1</sup>Computer Science and Engineering

<sup>2</sup>Overture Services, Inc.

<sup>4</sup>NEC Laboratories America, Inc.

<sup>3</sup>Information Sciences and Technology

74 N. Pasadena Ave., 3rd floor

4 Independence Way

Pennsylvania State University

Pasadena, CA 91101 USA

Princeton, NJ 08540 USA

University Park, PA 16801 USA

david.pennock@overture.com

lawrence@necmail.com

debnath@cse.psu.edu, giles@ist.psu.edu

## ABSTRACT

We analyze data from 52 online *in-game* sports betting markets (where betting is allowed continuously throughout a game), including 34 markets based on soccer (European football) games from the 2002 World Cup, and 18 basketball games from the 2002 USA National Basketball Association (NBA) championship. We show that prices on average approach the correct outcome over time, and the price dynamics in the markets are closely coupled with game events, agreeing with efficient market assumptions. We also examine qualitative distinctions between the two types of games.

## Categories and Subject Descriptors

J.4 [Computer Applications]: Social and Behavioral Sciences—*Economics*; G.3 [Mathematics of Computing]: Probability and Statistics

## General Terms

Economics, Measurement

## Keywords

Information incorporation, economic efficiency, efficient markets hypothesis, in-game sports betting, market reaction, soccer, logarithmic score, entropy.

## 1. RELATED WORK

A market can serve as a tool for aggregating buyers' knowledge about items of uncertain value. One of the stronger forms of the so-called *efficient markets hypothesis* states that information is incorporated into market prices virtually instantaneously. Gambling markets epitomize trading contracts of uncertain value. Analyses of horse racing markets [4], NBA point spread markets [1], and a market in the Euro 2000 soccer tournament [3], to name a few, are largely consistent with strong efficiency assumptions.

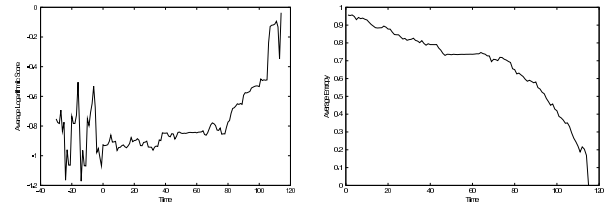


Figure 1: (a) ALS and (b) AE of 34 soccer markets.

## 2. DATA AND METRICS USED

We analyze information incorporation in markets on the World Sports Exchange where traders can bet on sporting events continuously throughout a game. The markets are double auctions in contracts that pay off \$100 if and only if the favored team wins the game by more than  $l$  points, where  $l$  is a nonnegative opening line or spread.

To measure the accuracy of implied market forecasts over time, we use the *logarithmic score*, a standard measure of the accuracy of probabilistic forecasts [5]. We took the market price to be the midpoint between the bid and ask prices, and normalized prices between 0 and 1. Let the price of the ultimate winner at time  $t$  be  $p_{win}(t)$ . Then the Average Logarithmic Score (ALS) for the market at time  $t$  is  $1/N \cdot \sum_{i=1}^N \log p_{win}(t)$  where  $N$  is the number of markets. Similarly, the Average Entropy (AE) at time  $t$  is

$$\frac{1}{N} \cdot \sum_{i=1}^N -p(t) \log p(t) - (1 - p(t)) \log(1 - p(t)).$$

Note that ALS can only be computed after the game ends, as it depends on the identity of the winning team, while AE does not depend on who wins.

## 3. SOCCER MARKETS

We analyze markets corresponding to 34 soccer games played during the 2002 World Cup. Every 10 seconds during every game, we recorded prices from the Sports Exchange as well as score changes and game clock information from CBS Sportsline.<sup>1</sup> ALS and AE for all 34 games appear in Figure 1. ALS increases roughly monotonically, indicating that prices approach the correct outcome over time. Increases in ALS reflect incorporation of evidence pertinent to

Copyright is held by the author/owner.

EC'03, June 9–12, 2003, San Diego, California, USA.

ACM 1-58113-679-X/03/0006.

<sup>1</sup><http://www.sportsline.com/>

the outcome of the game [2]. The graph seems to display two qualitatively different regions: a roughly linear increase (indicating a constant influx of information) until just after halftime, and a superlinear increase approaching the end of the game. The final few minutes (near the 104th minute on the graph, which is the 89th minute of the game after subtracting the 15 minute halftime) show the largest price movements, as any uncertainty about the outcome rapidly dissolves. AE is roughly monotonically decreasing during game play (remaining roughly constant during halftime), reflecting decreasing uncertainty. AE drops to 0 (certainty) as the game ends. The superlinear decay indicates a more rapid resolution of uncertainty near the end of the game.

#### 4. MARKET REACTION

In previous work [2], we examined the reaction of political stock markets to newsworthy events. For in-game sports betting markets, the clear dominant factors are the game score and the time remaining. Second-order factors include penalties, ball possession, perceived momentum, and fan enthusiasm, though some are hard to quantify.

**Speed of reaction.** We evaluate how promptly the market reacts to score changes in the games. We denote the time of a goal or score as  $\tau_s$  and the time of the corresponding price swing as  $\tau_p$ . Communication delays (including network and updating delays) from Sportsline and the Sports Exchange are represented as  $\delta_s$  and  $\delta_p$ , respectively. We assume that any negative time difference between  $\tau_p + \delta_p$  and  $\tau_s + \delta_s$  is due to delays rather than market foresight. We factor out delays by using the most negative difference as a threshold. This results in a conservative (over-) estimate of delay. We define the *measured delay* as

$$\Delta_m = (\tau_p + \delta_p) - (\tau_s + \delta_s).$$

Now for all the goals scored we recalculate the *thresholded delay* as  $\Delta = \Delta_m + \Theta$  where the threshold is defined as

$$\Theta = |\min(\{\Delta_m^i, \forall i\})|,$$

where the superscript  $i$  ranges over all goals scored in a game. We found that the average thresholded delay for the 74 goals scored in the 34 games is 31.633 seconds. Note that this difference reflects a conservative estimate, due to the thresholding procedure; any delay may be the result of website update delays or network delays.

**Amount of reaction.** The average percentage change in logarithmic score for the 74 goals is 111.34%. Figure 3(a) shows the extremely spiky distribution of percent change in logarithmic score. The current score and time remaining greatly influence the amount of reaction. For example, market reactions to late goals in blowout games are small or negligible.

#### 5. BASKETBALL MARKETS

We analyze markets from 18 games during the 2002 NBA Championship. Figure 2 plots ALS and AE for these games. The time axis includes the 48 minutes of actual game play, as well as the time during halftime, clock stoppages, etc. The slow initial decrease in AE indicates significant uncertainty throughout much of the game (AE remains greater than 0.8 for 55.5% of the game and greater than 0.7 for 77% of the game). Price changes in the markets are highly correlated with score changes. The right side of Figure 3(b) displays

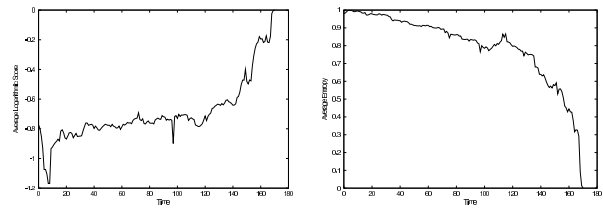


Figure 2: (a) ALS and (b) AE of 18 basketball markets.

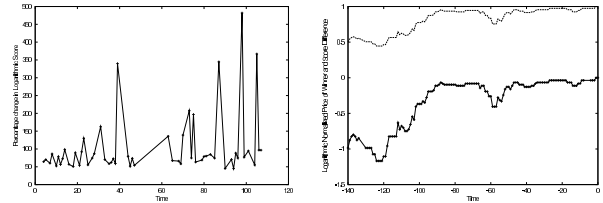


Figure 3: (a) Soccer reaction curve; (b) Correlation between logarithmic score and score difference for San Antonio vs. LA Lakers, May 7, 2002.

the correlation in one of the 18 games. The top curve is the normalized score difference between the two teams and the bottom curve is the logarithm of the price of the winner. The correlation between these two is 0.93. The average of all the correlations for all games is 0.61

#### 6. COMPARING THE GAMES

Basketball games are more uncertain for a larger proportion of the game, and so in one sense more exciting. On the other hand a comeback in soccer, being unlikely, is more dramatic than in basketball. In both cases, price is highly correlated with scoring. In soccer, where scoring is infrequent, price changes are infrequent but dramatic. In basketball, where scoring is frequent, price changes are frequent but less drastic. Meaningful goals late in a game affect price much more than early goals.

#### 7. REFERENCES

- [1] J. M. Gandar, W. H. Dare, C. R. Brown, and R. A. Zuber. Informed traders and price variations in the betting market for professional basketball games. *J. Finance*, LIII(1):385–401, 1998.
- [2] D. M. Pennock, S. Debnath, E. Glover, and C. L. Giles. Modeling information incorporation in markets with application to detecting and explaining events. In *Proc. of the 18th Conf. on Uncertainty in Artificial Intelligence*, pages 405–411, 2002.
- [3] C. Schmidt and A. Werwatz. How accurate do markets predict the outcome of an event? The euro 2000 soccer championships experiment. *Max planck Institute for Research into Economic Systems working paper*, 2002.
- [4] R. H. Thaler and W. T. Ziemba. Anomalies: Parimutuel betting markets: Racetracks and lotteries. *J. Economic Perspectives*, 2(2):161–174, 1988.
- [5] R. L. Winkler and A. H. Murphy. Good probability assessors. *J. Applied Meteorology*, 7:751–758, October 1968.