Intrade Prediction Market Accuracy and Efficiency: An Analysis of the 2004 and 2008 Democratic Presidential Nomination Contests

Ian Saxon

17 September 2010

Dissertation Supervisor: David Harvey

Word Count (Excluding Tables and References): 13,268

This Dissertation is presented in part fulfillment of the requirement for the completion of an MSc in the School of Economics, University of Nottingham. The work is the sole responsibility of the candidate.

Acknowledgements

I want to thank my supervisor, Associate Professor & Reader in Econometrics David Harvey, for his helpful advice and guidance. Thank you to John Delaney for making Intrade market data available and Ed Lenox for answering questions about those data sets.

Abstract

This dissertation uses a total of 300 days of political polling data and 1,981 days of Intrade.com market data to evaluate the relative accuracy of Intrade's predictions for the 2004 and 2008 Democratic presidential nomination contests, as well as the efficiency characteristics of the contract prices on the exchange. For all periods analyzed, market prices tend to be superior predictors of the ultimate winner of the two Democratic presidential nomination contests, as well as delivering stronger signals about who the winner would be when making correct predictions but similar signals about the winner when making incorrect predictions. Differences in variance between market prices and polls do not appear to explain this difference in predictive power. Pricing anomalies for closing prices are frequent, but tend to be corrected on the following trading day. Internal coherence metrics suggest that the 2008 market was likely more efficient than the 2004 market, although differences in liquidity between the two markets cannot explain the difference in efficiency.

Contents

1 Introduction

This dissertation seeks to expand on the existing literature that analyzes prediction market accuracy and efficiency. A combined total of 300 days of political polling data and 1,981 days of Intrade.com market prices are used to evaluate the relative accuracy of Intrade's predictions for the 2004 and 2008 Democratic presidential nomination contests. Then, several measures of market efficiency are identified and evaluated using the market price series for each contract listed on the exchange.

Section 1.1 gives an overview of prediction markets, giving some examples of their current and potential uses. Section 1.2 reviews previous relevant research and discusses some of the results of that research. Section 1.3 describes the data that is used in this dissertation, showing graphs for the 8 time series related to the 2004 Democratic presidential nomination contest and the 26 series related to the 2008 Democratic presidential nomination contest.

Section 2.1 outlines the methodology used to evaluate the relative accuracy of Intrade market prices as predictions of the winners of the 2004 and 2008 Democratic presidential nomination contests and Section 2.2 discusses the results of this analysis. Given that market prices appear to be better predictors than polling data, Section 2.3 calculates the likelihood that this could have occurred by chance. In other words, if market prices and polling data actually have the same underlying ability to predict winners in the political contests analyzed here, what is the probability that we would have observed the particular mix of market wins compared to polling wins reported in this dissertation? Section 2.4 considers the relative strength of predictions made by both market prices and polls in instances where those predictions are either correct or incorrect, seeking to identify a difference in information content in the two generators of predictions. Section 2.5 suggests and assesses one potential source of the observed superiority of predictions generated by market prices compared to those generated by polling data. In addition, there is a discussion of the difficulty of assessing the accuracy of market prices and polling data in the context of having just two real world event outcomes to compare against the relevant predictions.

Several aspects of expected contract pricing (taking into account the effects of fees and margin requirements) and observed anomalies are analyzed in Section 3.1. Section 3.2 shows under what circumstances and how frequently those pricing anomalies are corrected. Section 3.3 features an analysis of how the mean and variance of the market portfolio (the sum of all available contracts within a category) change over time, and whether market liquidity can explain any of that change. Tests for random walk behaviour in the series for individual contracts as well as for the market portfolio are presented in Section 3.4, with a view toward assessing the degree of unpredictability present in contract pricing as well as internal pricing coherency.

Finally, Section 4 summarizes the findings of this dissertation and draws conclusions.

1.1 Overview of Prediction Markets

Prediction markets, which are futures markets based on the outcomes of real world events like elections, movie box office receipts, and even various environmental events, have gained great currency in recent years as potentially useful guides to the future. Manski (2006) notes that it has become common to interpret market prices for prediction market contracts as probabilities that an event will happen[1]. If prediction markets are efficient in the sense that Fama (1970) famously outlined , Wolfers and Zitzewitz (2004) point out that:

> ...the market price will be the best predictor of the event, and no combination of available polls or other information can be used to improve on the market-generated forecasts.

A few examples of prediction markets include the University of Iowa's Iowa Electronic Markets (IEM), the Hollywood Stock Exchange, The Foresight Exchange, the University of British Columbia's Election Stock Market (ESM), and Intrade. Most of these markets operate using real money as the medium of exchange, while HSX uses a simulated currency called Hollywood Dollars and the Foresight Exchange uses FX-bucks.

To provide some idea of the type of things prediction markets features on their exchanges, consider a few of the contracts on offer at Intrade.com on 21 August 2010:

- Whether Roger Clemens, the former Major League Baseball player, would plead guilty or be found guilty of at least one count of perjury, making false statements, or obstruction of congress
- Whether same sex marriages would resume in California before midnight 31 Dec 2010
- Whether Robert Gibbs would depart as White House Press Secretary before the end of US president Obama's first term
- Whether Aung San Suu Kyi, the Burmese opposition politician, would be released from house arrest before midnight 31 Dec 2010
- Whether global average temperatures for 2019 will be higher than for 2009

There is a dearth of data on the identities and intentions of users of Intrade or other prediction markets, which is unfortunate for those who wish to study how prediction markets are used by individuals, organizations, and businesses. With respect to the 2004 market for the Democratic presidential nomination, the open interest[2] was low enough (it peaked at $62,166) that it is unlikely that anyone but relatively small time speculators were involved in the market. The market for the 2008 Democratic presidential nominee, however, garnered considerably more attention. With open interest peaking at $639,022, it is more likely that some of the participants were engaging in the market to hedge political risks. Intrade certainly promotes itself as both a speculative and hedging tool: the website claims to have "created an exchange for you to trade (speculate on) events that directly affect your life, like politics, entertainment, financial indicators, weather, current events and legal affairs." Presumably, events that directly affect the lives of potential bettors also offer a chance for individuals to place bets that act as a hedge against adverse outcomes that may impact them.

[1] Manksi goes on to show under what equilibrium conditions this interpretation is acceptable.

[2] Open interest is one of two measures of volatility that were included in the data, but the figures for trade volume were too error prone to use. Thus, open interest is used in this dissertation whenever liquidity is discussed.

The current legality of prediction markets is somewhat of a gray area, which is likely an impediment to their widespread use as hedging tools by individuals, organizations, and institutions. Intrade.com is hosted in Ireland which has fairly liberal gambling laws, and the website advises bettors to be aware of the gambling laws in their home countries before placing bets. In the United States, most states have legislation prohibiting large scale gambling. However, since placing bets online (whether in poker, horse racing, or political futures markets) is a relatively new phenomenon, it is still not clear whether it is permissible for US citizens to place bets on, say, Intrade's exchanges situated in Ireland. The US Commodities Futures Trading Commission has reportedly fined Ireland's Trade Exchange Network, the owner of Intrade and TradeSports for soliciting US customers (McAfee, 2007), while other legislation has sought to curb the use of non-US betting markets by making it difficult to transmit the required funds through US intermediary banks (Phillips, 2006).

Prediction markets have also come under fire as creating the potential for abuse and manipulation (Hanson, 2006). Some have warned that prediction markets for sensitive events (such as whether a political leader will be assassinated) may make those unwelcome events more likely than they would be if no betting market were in place. Thus, the legitimate hedging uses of prediction markets are not as well-recognized as similar uses on traditional commodities and stock exchanges. These reputation issues are likely to continue to limit the use of prediction markets as hedging tools on a large scale.

The challenges that prediction markets face in gaining widespread acceptance was recently made obvious when the US Senate Agriculture Committee passed legislation preventing box-office events from being included in real money futures markets (Flint, 2010). The US film industry lobbied heavily against the ability of the Hollywood Stock Exchange to begin implementing real money markets in box-office receipts, claiming that manipulation and insider trading would damage the movie-making industry.

Interestingly, Oprea et al. (2006) and Hanson and Oprea (2009) have argued that manipulation can actually aid prediction market accuracy, with a greater number of large, known manipulators increasing the incentive for more informed traders to take advantage of the increased liquidity in the market and profit from the introduction of the noise produced by the manipulative intent. Hanson et al. (2006) went on to confirm and demonstrate this theoretical result in an experimental market setting.

A long-time proponent of prediction markets, Hanson (2007) has suggested using prediction markets as a major input into public policy decisions, labelling this form of government "futarchy". Hanson envisions a process whereby voters would elect politicians who would promise to adhere closely to the recommendations made by speculators about how best to achieve numerous policy objectives and national measures of welfare.

Some businesses, including Google, Hewlett-Packard, and Siemens have made use of prediction markets internally to improve decision making and forecasting of sales (Cowgill, 2005; Wolfers, 2004). This use of prediction markets is likely to continue, and there are a number of companies (such as Consensus Point and Inkling[3]) that have sprung up to help businesses implement in-house prediction markets.

[3] See consensuspoint.com and inklingmarkets.com

1.2 Previous Research

Berg et al. (2008a) established that the Iowa Electronic Market's (IEM) election eve prices for political contracts provided good relative predictive accuracy, beating election eve polling indicators via a measure of absolute error from the final result in 9 out of 15 elections. The elections examined were held in the United States, Austria, Denmark, France, Germany, and Turkey. For the 15 elections in question, the average absolute deviation for the polls was 1.91% while the average absolute deviation produced by the market prices was 1.49%. In the process, the authors noted with interest that the IEM contract prices were more stable over time than the polling figures gathered from a variety of polling organizations, which were quite volatile.

Berg et al. (2008b) followed up their paper analyzing election eve forecasts by analyzing prediction market accuracy in the long run. They did so by comparing polling predictions against market predictions (again using data from the IEM as the relevant market prices) of ultimate vote share for candidates in the 1988, 1992, 1996, 2000, and 2004 presidential elections. Because Berg et al's analysis has some similarities to the analysis presented in this dissertation, their methodology and findings are briefly outlined here.

For each day that polling data was available, Berg et al paired the corresponding IEM market price to that figure and then computed what they termed the average absolute error of the poll predictions and the market predictions. They also calculated a p-value statistic to indicate whether a particular instance of market outperformance was likely to be due to chance or whether it was unlikely that the observed level of out-performance could have occurred by chance. These computations were conducted over a range of time periods, which provided some insight into the relative predictive prowess of market prices over the long term versus the short term. Table 1 shows the results of this analysis.

Table 1: Selected Results (Averages Over All Five Elections Considered) From Berg Et Al (2008)

	Days Prior To Election					
	Last 5	**6-31**	**32-65**	**66-100**	**101+**	**Whole Period**
Mkt. Vs. Polls	68%	73%	68%	84%	74%	74%
Mkt. Vs. Mov. Avg.	67%	70%	68%	85%	70%	71%
Mkt. Vs. Most Recent	72%	75%	72%	78%	76%	75%

Mkt. Vs. Polls: Percent of observations where market prices had a lower average absolute error than contemporaneous poll results. Mkt. Vs. Mov. Avg.: Percent of observations where market prices had a lower average absolute error than a moving average of poll results. Mkt. Vs. Most Recent: Percent of observations where market prices had a lower average absolute error than the most recent poll(s).

Considering all time periods and all elections, Berg et al found that market prices outperformed polling predictions for the ultimate vote share results 74% of the time. The relevant p-value was given as 0.000, which indicated that the result was extremely unlikely to occur by chance.

Pennock et al. (2000) examined internal coherency and accuracy in two play money markets and one real money market. The Hollywood Stock Exchange and the Foresight Exchange were the play money markets, where participant reputation is the motivation for succeeding as a trader rather than pecuniary profits. The real money market Pennock et al looked at was the IEM, which was used as a comparison for their findings on internal coherency. Analyzing contracts for the Oscar and Emmy options markets on the HSX and IEM exchanges, Pennock et al pointed out the relevance of internal pricing coherence in the sense that bundles of contracts in the same category ought to sum to some predictable value, such that the implied probability of the sum of the contracts listed never rises above one. Indeed, they found that while bundles of contracts in the same category sometimes became grossly over- or under-priced (sometimes by as much as 40%), there was a marked tendency for these contracts to revert to the coherent price.

Of particular interest was their finding that after taking into account of the reversionary tendencies of contracts that strayed from the coherent price, prices for bundles of contracts in the same category were more likely to remain over- rather than under-priced. A similar finding is discussed in this dissertation.

Much of the literature probing prediction market accuracy has focused on the relative success of markets versus poll and pundit predictions, but it may become more common to compare market predictions against other forecasting tools. As an example of this type of research, Asur and Huberman (2010) performed an analysis comparing HSX's play money markets to predictions resulting from the frequency of "tweets" on twitter.com for popular movies. They considered a total of 24 films released between December 2009 and January 2010 and found that a basic econometric model incorporating twitter data moderately outperformed a similar model using HSX data.

Tetlock (2008) looked at the TradeSports[4] exchange to determine the effect of liquidity on efficiency. Using three measures of liquidity, including bid-ask spreads, trading volume, and market depth (number of open contracts), Tetlock found that listed sporting and financial events featuring greater amounts of liquidity often did not see an improvement prediction accuracy. In fact, surprisingly, Tetlock found that greater liquidity sometimes resulted in a decrease in forecasting accuracy. He suggested that this may be a result of traders executing automated trading strategies during information events. These traders, who request that a trade be placed when the price reaches a certain level, may add noise to the process of price discovery because they are trading on old information. This would slow the market's movement toward the new, more accurate equilibrium.

Tetlock also comments on the fee structure of the TradeSports exchange, discussing how it may affect incentives to trade when arbitrage opportunities present themselves. Of particular note, the fee structure at TradeSports was such that arbitragers can profit from incorrect pricing if the mis-pricing was merely by one point, often the smallest increment possible. The fee structure of the Intrade markets will be reviewed in this dissertation, and some discussion will be devoted to pointing out the effects these fees have on trader incentives.

Servan-Schreiber et al. (2004) analyze one real money market (TradeSports.com) and one play money market (NewsFutures.com) to find out if real money markets are more accurate than play money markets. Using predictions related to the 2003-2004 NFL season, they found that both markets significantly outperformed individual human predictions but that the two markets themselves were fairly evenly matched in terms of prediction accuracy. Their research indicated

[4] This exchange is now defunct.

that 65.9% of the teams that were favourites on the TradeSports exchange eventually won their games and 66.8% of the teams predicted to win on the NewsFutures exchange eventually won their games. Servan-Schreiber et al conclude that despite a common refrain among those who study or have commented on prediction markets that real money markets are likely to be more accurate because they require traders to "put their money where their mouth is", the distinction may not actually be very important.

1.3 Data Description

Polling data for the 2004 and 2008 Democratic Presidential Primaries in the United States was collected from www.pollingreport.com (Abramowitz et al., 2004, 2008). The 2004 polling data features polls from 3 January 2003 to 27 February 2004, providing 84 polling days (days on which at least one poll was conducted) and 103 individual polls. The 2008 polling data features polls from 11 December 2006 to 24 August 2008, providing 216 polling days and 295 individual polls.

Intrade market data[5] was collected from Intrade.com for all seven market-listed candidates for the 2004 nomination (including John Kerry, Joseph Lieberman, Hillary Clinton, Dick Gephardt, John Edwards, Howard Dean, and John McCain) and all 25 market-listed candidates for the 2008 nomination (including Barack Obama, Hillary Clinton, John Kerry, Al Gore, Bill Richardson, Howard Dean, John Edwards, Wesley Clark, Joseph Lieberman, Evan Bayh, Patrick Leahy, Chris Dodd, Tom Vilsack, Joseph Biden, Harold Ford, Ed Rendell, Colin Powell, Mark Warner, Russ Feingold, Jon Corzine, Phil Bredesen, Brian Schweitzer, Mike Easley, Tom Daschle, and Rod Blagojevich). Some politicians, like John McCain in 2004, did not officially run but were included in Intrade's markets anyway, presumably because market interest existed for those contracts. Market data for the 2004 nomination covers a period from 8 November 2002 to 28 July 2004, comprising 613 days and 4,426 observations. Market data for the 2008 nomination covers a period from 4 November 2004 to 28 August 2008, comprising 1368 days and 33,446 observations. The market data from both the 2004 and 2008 Democratic Primaries comprises (nearly continuous[6]) daily closing prices for each of the candidates listed on the exchange. Unlike most stock exchanges, Intrade markets operate on weekends and holidays.

The Intrade market price series for the 2004 Democratic presidential nomination are shown in Figures 1 and 2. All series have the same starting date (8 November 2002), but the series for Joseph Lieberman, Dick Gephardt, and John McCain do not run to the date of the 2004 Democratic National Convention because those candidates either dropped out or made it clear that they would never run (as was the case with John McCain). In particular, these candidates' contracts ceased trading just prior to one of the critical milestones in the presidential primaries, the Iowa Caucuses, held on 19 January 2004. However, on the day these contracts cease to be listed in the historical data, there appear to have been holders of those contracts still participating in the market. Liquidity figures indicate that McCain contracts had open interest of $1,120, Lieberman contracts had open interest of $8,104, and Gephardt contracts had open interest of

[5] Variable names and descriptions for data collected from Intrade can be found in Table 10 in the Appendix.

[6] On visual inspection, two dates in the 2004 records had obviously erroneous data. In addition, there were were 17 unexplained gaps in the dates. In the 2008 data, there were 23 unexplained gaps in the dates.

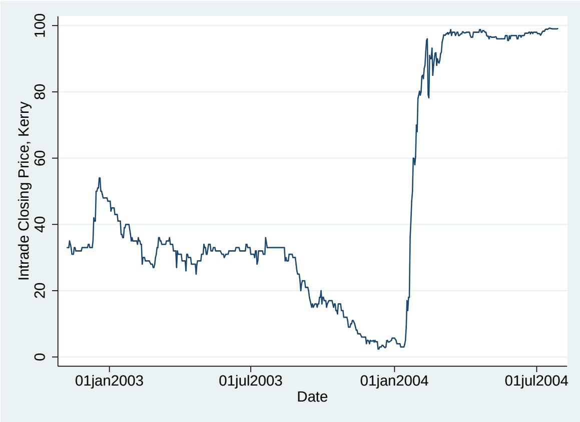
$4,801. Whether this represents an error in the data is not clear, although the prices for these three contracts (with the possible exception of Dick Gephardt's which was at 3.6[7]) were so low at the end of the time series that their disappearance is not likely to have a significant impact on any analysis that follows.

The Rest of Field contract, which represents the likelihood that some candidate not listed on the exchange would go on to win the nomination, was available from the start of the 2004 market. This contract achieved considerable liquidity and some high prices at some points, making it crucial to include it in the analyses that follow.

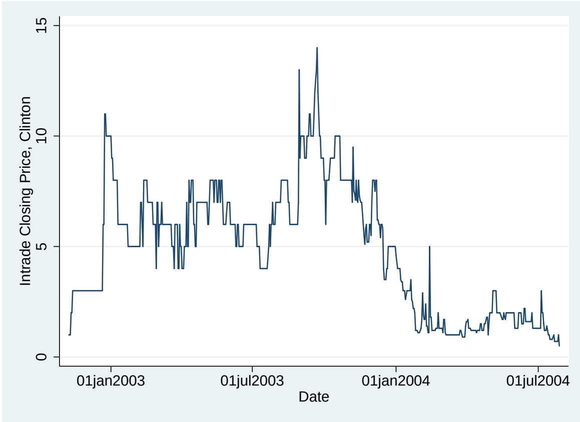
The Intrade market price series for the 2008 Democratic presidential nomination are shown in Figures 3, 4, 5, 6, and 7. In direct contrast to the market in 2004 candidates, all contracts for 2008 candidates finish on the same date (28 August 2008), but have different starting dates. Oddly given the well-traded nature of the Rest of Field contract in the 2004 market, the Rest of Field contract was not introduced in the 2008 market until 15 September 2007. It may be that the large number of candidates included on the exchange made the Rest of Field contract largely irrelevant. In any case, the inclusion of a Rest of Field contract has some interesting implications for the discussion of internal pricing coherence in section 3.1.

[7]Market prices on Intrade are quoted as values between 0 and 100, so a price of, say, 3.6 actually means that a single contract trades for $0.036 USD.

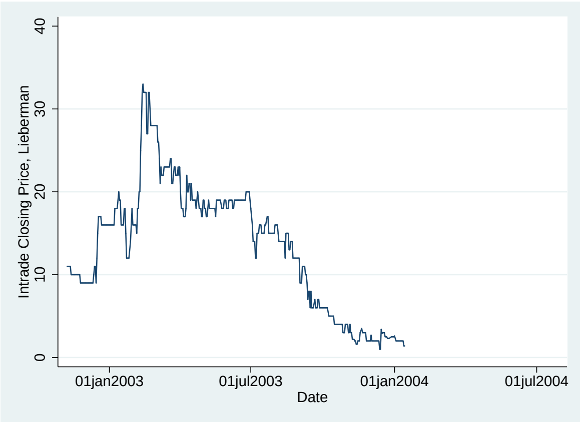
Figure 1: Intrade Market Prices, 2004 Democratic Party Presidential Nomination



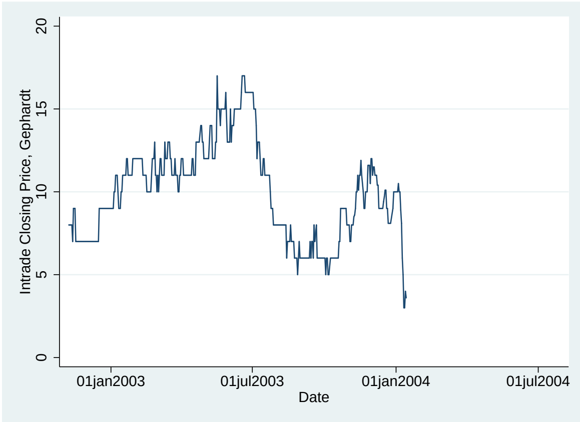
(a) John Kerry



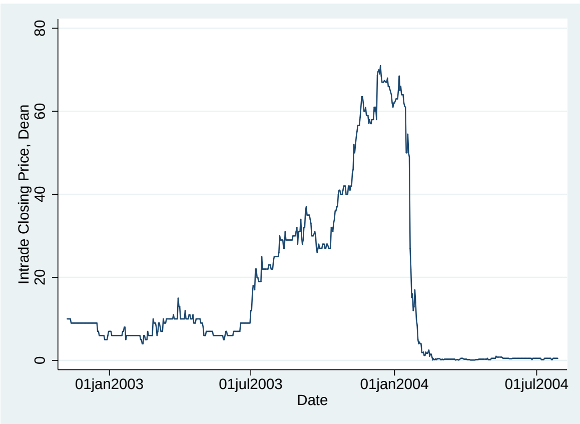
(b) Hillary Clinton



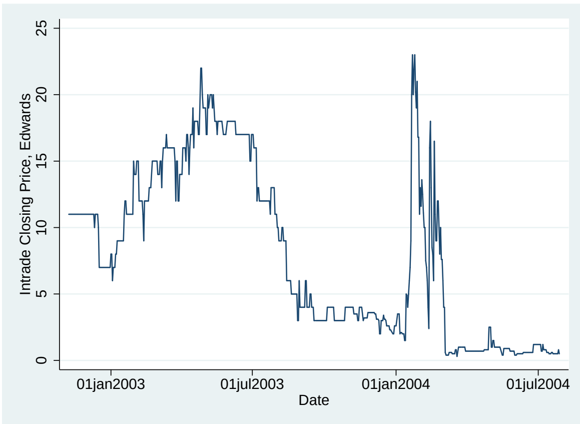
(c) Joseph Lieberman



(d) Dick Gephardt

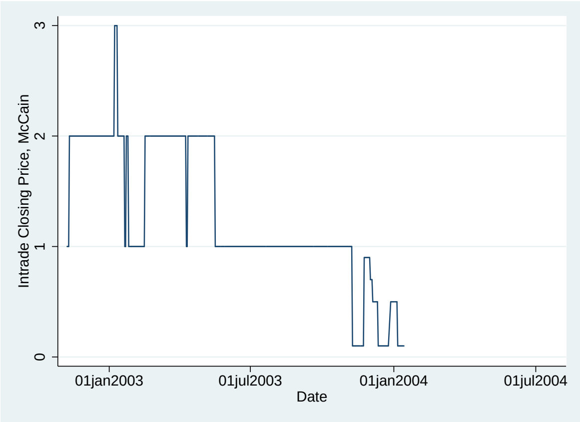


(e) Howard Dean

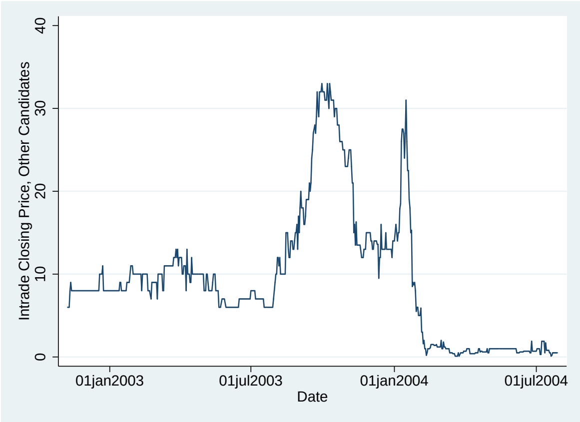


(f) John Edwards

Figure 2: Intrade Market Prices, 2004 Democratic Party Presidential Nomination

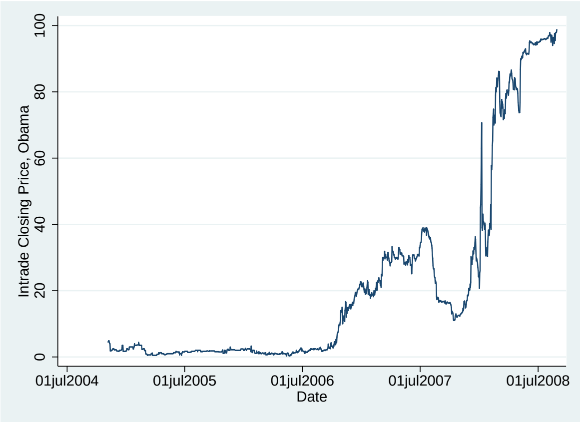


(a) John McCain

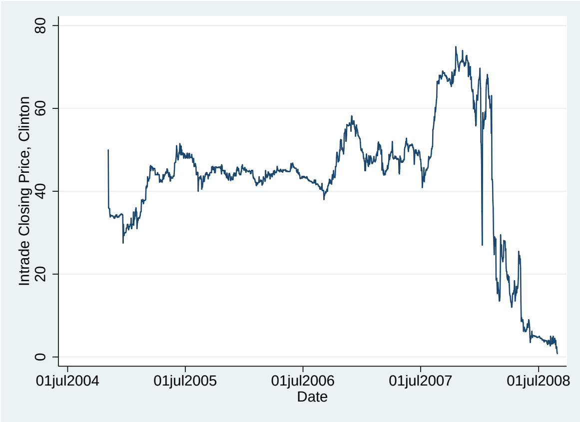


(b) Rest of Field

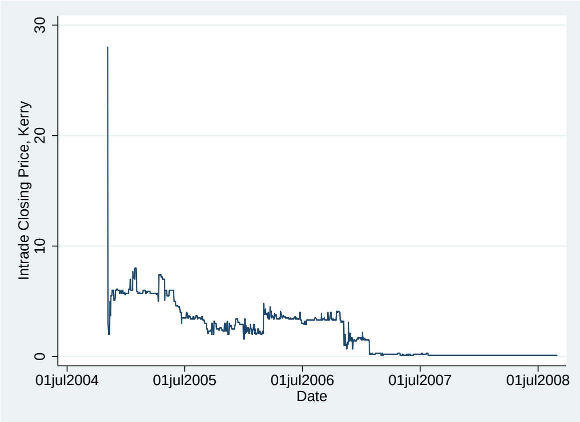
Figure 3: Intrade Market Prices, 2008 Democratic Party Presidential Nomination



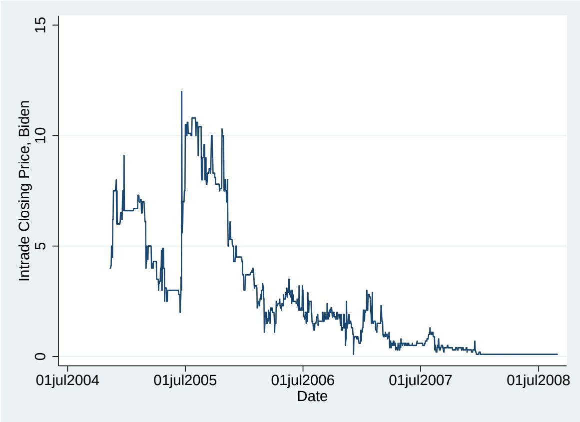
(a) Barack Obama



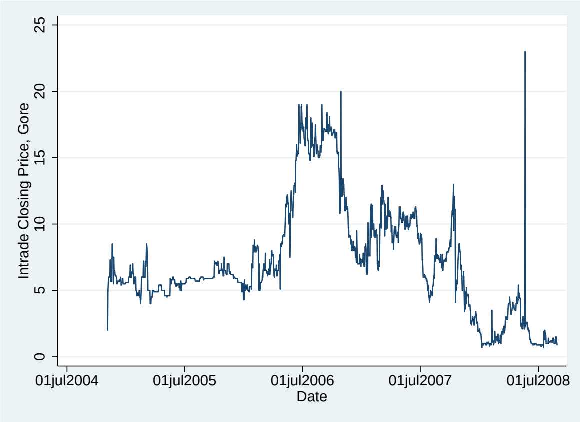
(b) Hillary Clinton



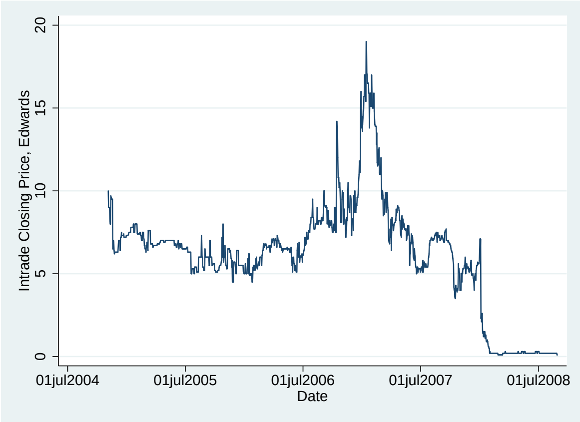
(c) John Kerry



(d) Joe Biden

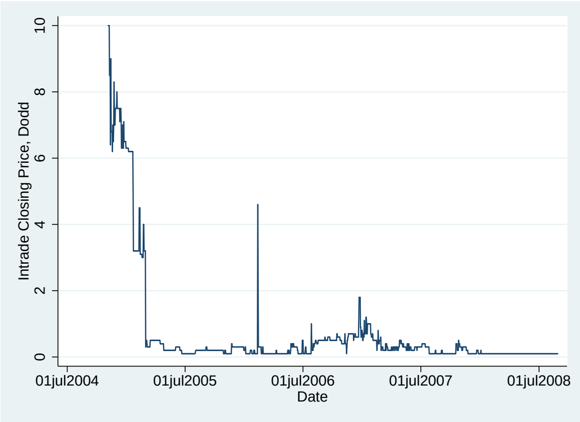


(e) Al Gore

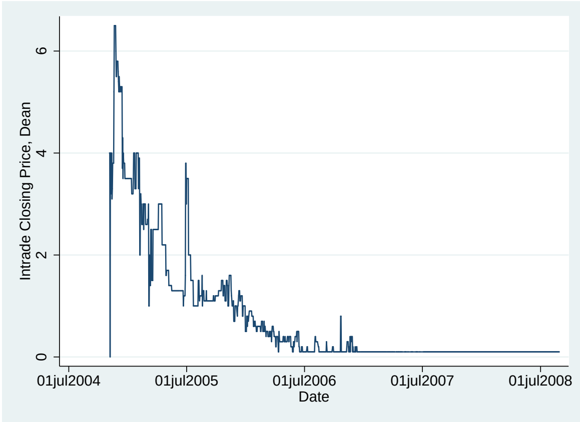


(f) John Edwards

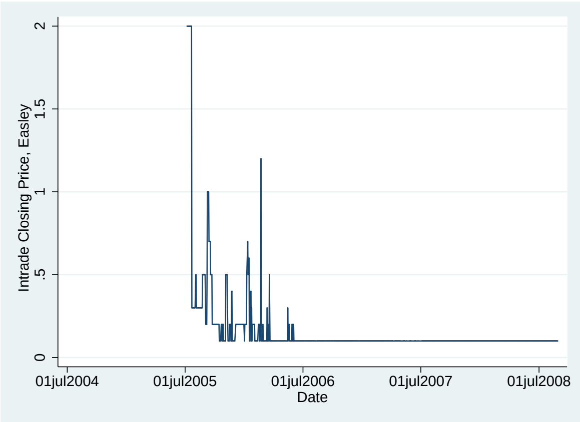
Figure 4: Intrade Market Prices, 2008 Democratic Party Presidential Nomination



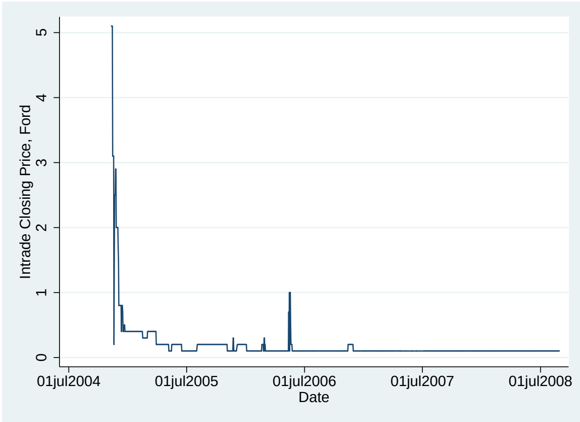
(a) Christopher Dodd



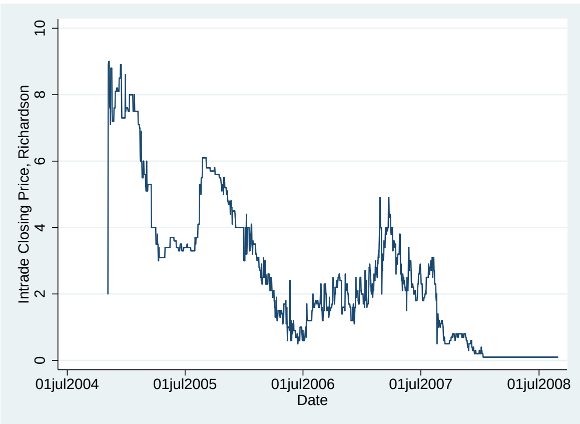
(b) Howard Dean



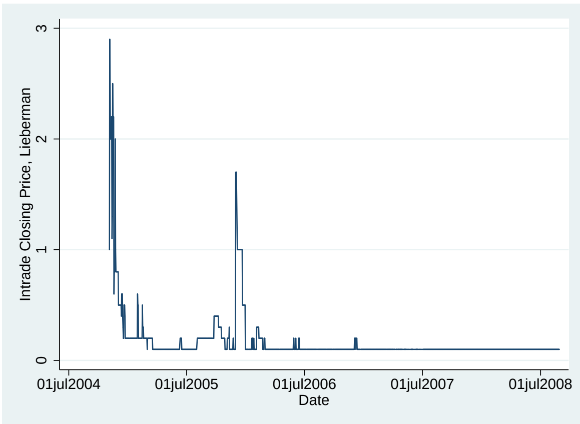
(c) Mike Easley



(d) Harold Ford

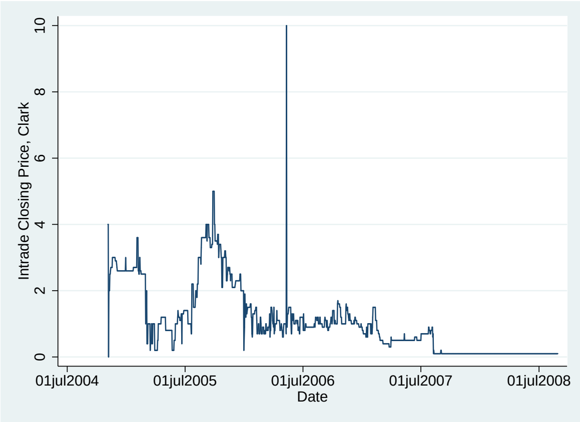


(e) Bill Richardson

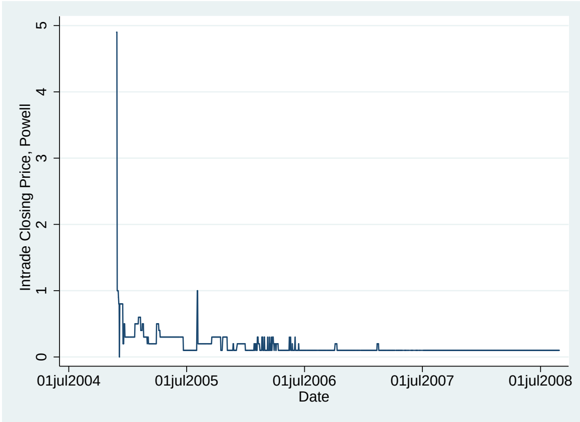


(f) Joseph Lieberman

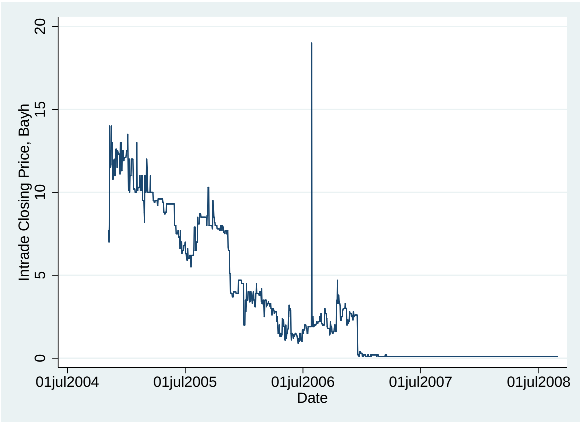
Figure 5: Intrade Market Prices, 2008 Democratic Party Presidential Nomination



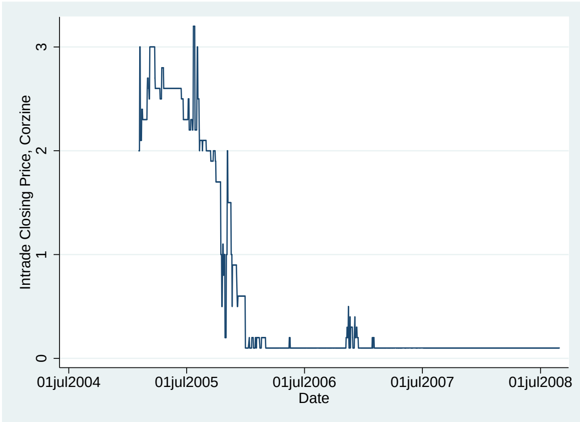
(a) Wesley Clark



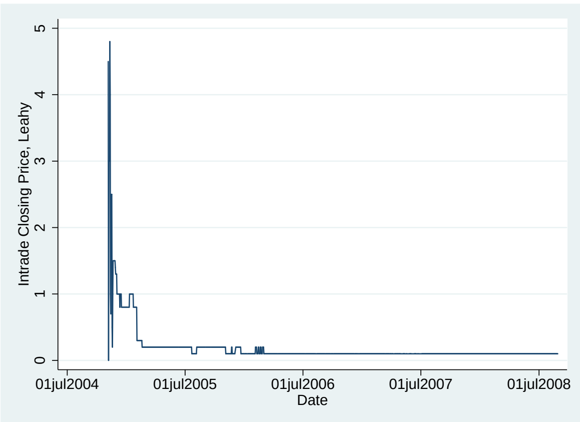
(b) Colin Powell



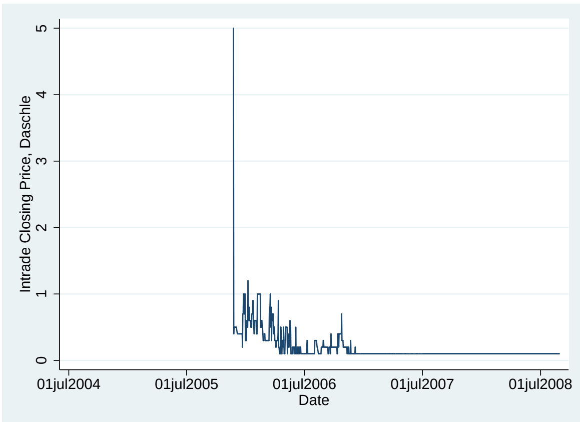
(c) Evan Bayh



(d) Jon Corzine

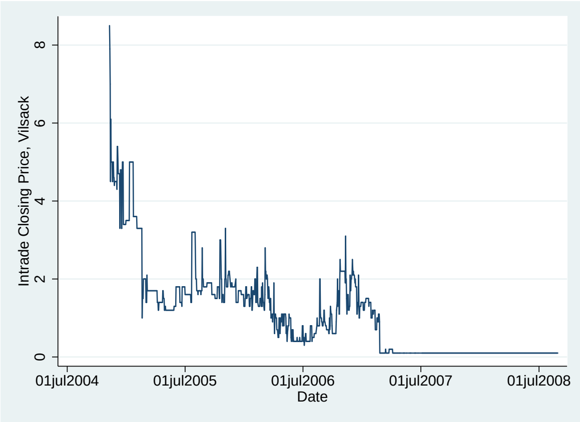


(e) Patrick Leahy

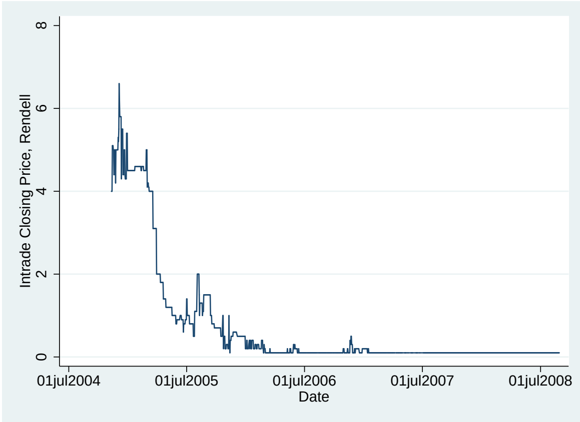


(f) Tom Daschle

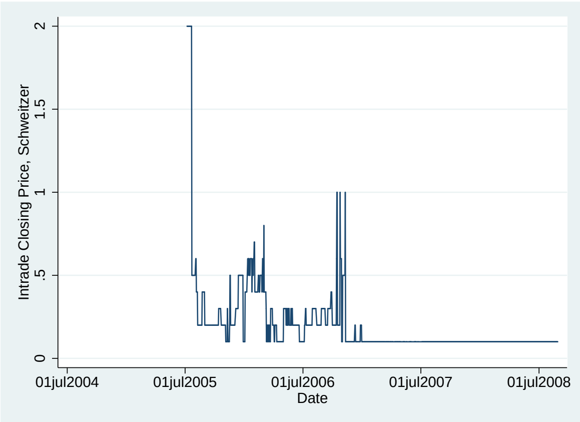
Figure 6: Intrade Market Prices, 2008 Democratic Party Presidential Nomination



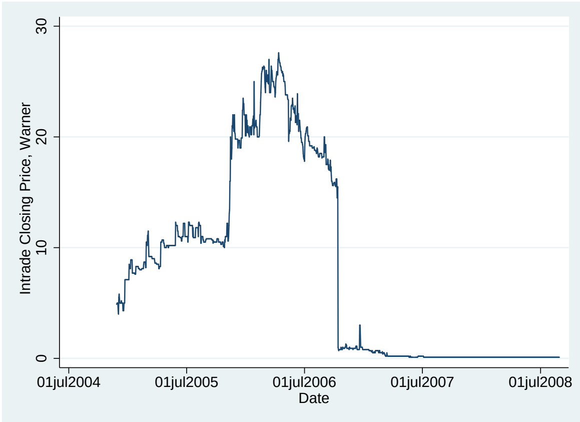
(a) Tom Vilsack



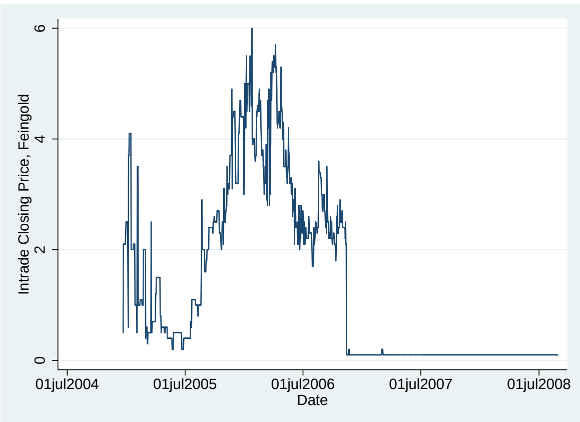
(b) Ed Rendell



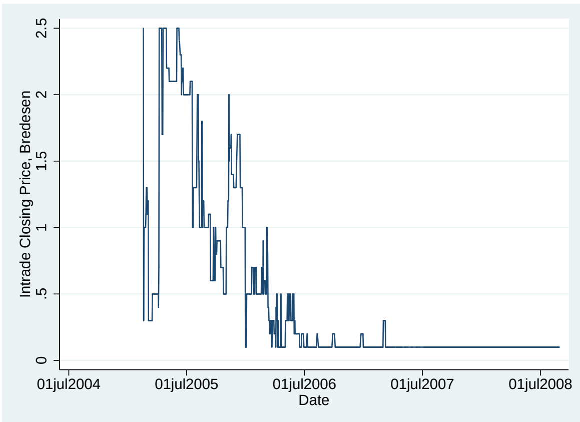
(c) Brian Schweitzer



(d) Mark Warner

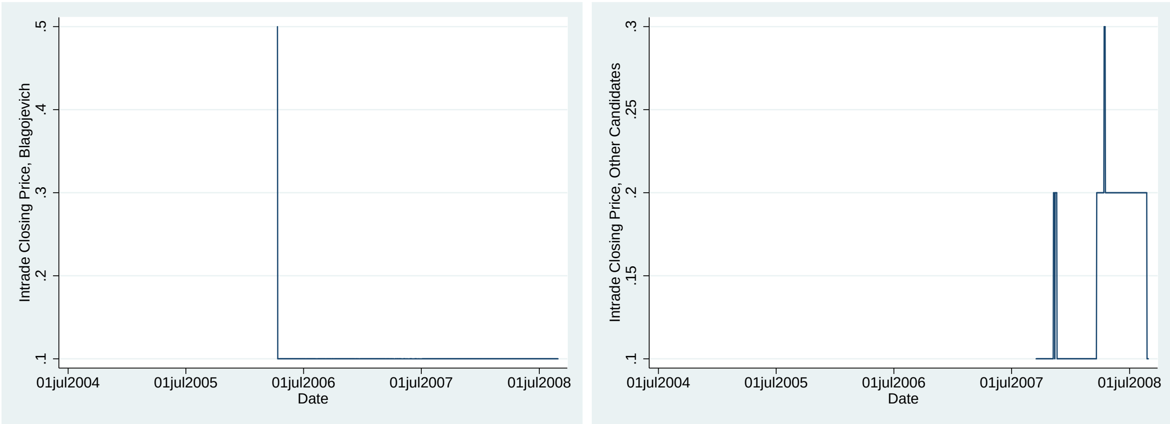


(e) Russ Feingold



(f) Phil Bredesen

Figure 7: Intrade Market Prices, 2008 Democratic Party Presidential Nomination

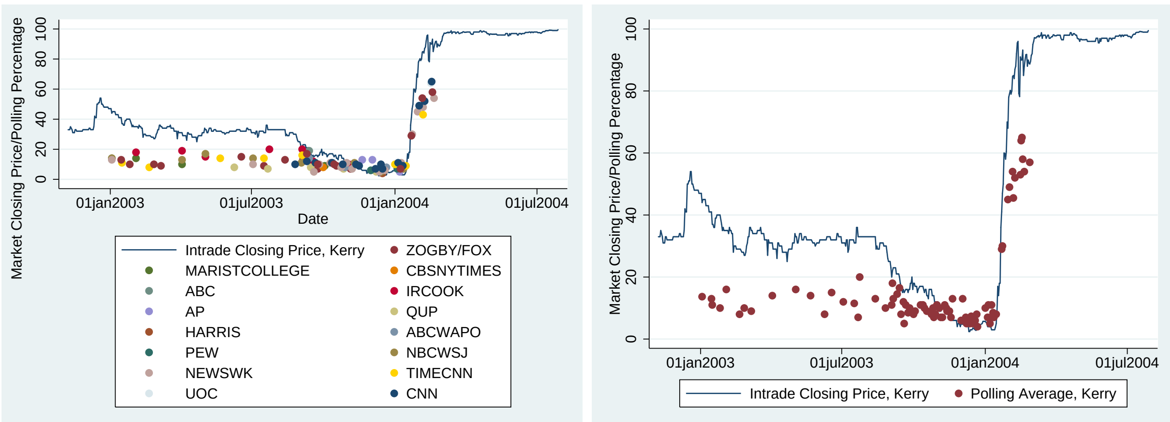


(a) Rod Blagojevich (b) Rest of Field

2 Intrade Prediction Market Accuracy Compared to Polls

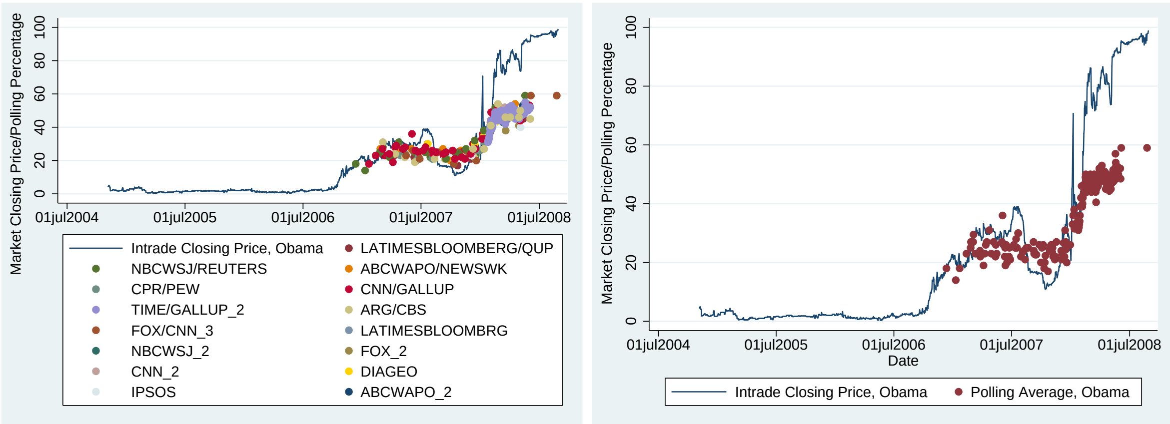
Figures 8 and 9 show the Intrade market price time series and predicted vote shares according to polling data for the two candidates who eventually won the 2004 and 2008 Democratic presidential nomination contests. This section will use polling data from 18 polling organizations that were active for the 2004 Democratic presidential nomination contest and another 18 polling organizations for the 2008 process as well as market price data taken from Intrade.com in order to analyze the relative predictive accuracy of the Intrade market compared to polls.

Figure 8: Market prices compared to polls for John Kerry for the 2004 Democratic Presidential Nomination. In Figure 8(b), polls taken on the same day were averaged.



(a) Market Prices vs. Individual Polls, Kerry (b) Market Prices vs. Average of Polls, Kerry

Figure 9: Market prices compared to polls for Barack Obama for the 2008 Democratic Presidential Nomination. In Figure 9(b), polls taken on the same day were averaged.



(a) Market Prices vs. Individual Polls, Obama

(b) Market Prices vs. Average of Polls, Obama

2.1 Methodology

Following the methodology of Berg et al, every instance of a poll and its prediction of vote share for the candidates featured within the poll was paired with market prices representing the likelihood of the relevant candidates winning the nomination. The purpose of this analysis of prediction market accuracy is to compare the success of Intrade's market data in predicting the ultimate winner of the two Democratic presidential nomination contests.

Polling data collected from the 2004 Democratic presidential primaries covers a period from 3 January 2003 to 27 February 2004. This data features 103 individual polls and 84 polling days, meaning that 19 polls were conducted on the same day (either by different polling organizations or, in some cases, the same polling organization). Polling data collected from the 2008 Democratic presidential nomination covers a period from 11 December 2006 to 24 August 208. This data features 295 individual polls and 216 polling days, meaning that 79 polls were conducted on the same day. If a poll by a single organization was conducted over several days, the last day of polling is used as the date of the poll.

Intrade market data collected from the 2004 Democratic presidential nomination covers a period from 8 November 2002 to 28 July 2004[8], a nearly continuous 613 days. Intrade market data collected from the 2008 Democratic presidential election covers a period from 11 December 2006 to 24 August 2008[9], a nearly continuous 1,368 days.

While Berg et al compared market prices corresponding to vote shares to polling data corresponding to vote shares, the data collected from Intrade for the purposes of this dissertation represents the likelihood that a particular candidate would win the nomination and so is not

[8]The Democratic National Convention for the 2004 presidential election cycle was held on July 29, 2004. John Kerry and John Edwards were officially nominated by the Democratic Party to be the candidates for President and Vice President of the United States (Toner and Seelye, 2004).

[9]The Democratic National Convention for the 2008 presidential election cycle was held on August 28, 2008 where Barack Obama was officially nominated to be the Democratic candidate for the president of the United States (Nagourney, 2008).

directly comparable to polling figures. The following procedure was conducted to make the two data sets comparable:

1. For days on which more than one poll was conducted, the forecasted vote shares from each poll were averaged. (Thus, the 103 individual polls conducted in the run-up to the 2004 Democratic Nation Convention became 84 polling days and the 295 individual polls taken prior to the 2008 Democratic National Convention became 216 polling days.)

2. For polls, the candidate with the highest forecasted vote share was labelled as the predicted winner, while other candidates were labelled as predicted losers

3. For market data, the candidate with the highest market price was labelled as the predicted winner, while the other candidates were labelled as predicted losers

Similar to Berg et. al, the number of days prior to the Democratic National Conventions over which both polling data and market data were available were split up into subperiods. This provided the opportunity to observe variability in predictive power among polls and market prices depending on the amount of time prior to the concluding events those predictions were made. Next, the number of correct predictions for both polls and Intrade's market prices was tallied across each of the subperiods, as well as for the entire period. These results are shown in Table 2.

Before discussing the main results, it is necessary to consider the differences between polls and market data. Essentially, what is being "asked" of Intrade's market participants prior to placing a bet can be quite different from what is being asked of polling participants. Intrade's contracts ask market participants to place a bet on who they think will ultimately win the Democratic presidential nomination. In contrast with participants in polls, bettors must put aside their own desires about which candidates they would *like* to win and must instead focus on the real world question of who *will* win. Prediction markets are considered to be accurate partly because of this feature, which requires participants to "put your money where your mouth is".

As a result, and in keeping with Manski (2006); Wolfers and Zitzewitz (2006), market prices can be interpreted as probabilities. If, say, a contract that pays out $1 if Hillary Clinton wins the Democratic nomination to run as president is currently selling for $0.60, observers can reasonably infer that the market sanctioned probability for the event "Hillary Clinton wins Democratic nomination" is p=0.60. If an individual disagrees with this probability, he is free to buy or sell the contract accordingly in order to profit from the perceived discrepancy. If enough market participants begin to see the probability of a given candidate winning being different from that implied by the current price, the price will change to reflect that via the buying and selling behaviour of the market participants.

Whereas it is relatively easy to interpret market prices, interpreting polling data is slightly more challenging. There are several reasons for this, each of which will be discussed:

1. Different polling organizations and, indeed, different polls conducted by the same polling organization, ask different questions. The following is an example of a question that was asked by CBS News in 2003 about the 2004 Democratic presidential nomination: "From

what you have heard or read, can you name any of the candidates running for the 2004 Democratic nomination for president?" If "Yes": "Who is the first one who comes to mind?" In contrast, consider the question asked for a poll conducted by Zogby American Poll just a few days earlier: ". . . If the Democratic primary for president were held today and the candidates were [see below], for whom would you vote?"

Sometimes the underlying question asked is nearly identical, but the phrasing is different. For example, one poll may ask participants "Who would you like to see the Democratic Party nominate as its presidential candidate in 2004?" and another might ask, "If a 2004 Democratic primary for president were held today, which ONE of the following candidates would you most likely vote for?" Without conducting a separate poll for the purposes of determining the effect of the alternate phrasing on the responses, the phrasing differences can leave us uncertain about the results.

2. Polls almost always ask participants who they *want* to win, not who they *believe* will win. Polling organizations count on the stated desires of poll participants to infer voter turnout for various candidates, but most polls do not measure the strength of those desires which may be critical in forecasting actual turnout. As an example of a possible scenario, an equal number of poll respondents may say they want candidates X and Y, respectively, to win, but those who want candidate X to win may be twice as motivated to vote and make it a reality. In such a scenario, polls may give a badly biased estimate of voter sentiment.

3. Unlike markets, polls do not give participants an incentive to answer honestly. Perhaps there are a core group of respondents who will answer honestly, the rest will answer at random, and the noisy responses will cancel each other out. Then again, perhaps there is a tendency for dishonest answers to cluster around particular candidates, skewing the results and providing an inaccurate view of the voting tendencies of the population.

4. Poll participants are a random selection of the population, whereas market participants are a self-selected sample.

5. A poll may suggest that, if voters vote the way they responded to the polls, one candidate will have a small margin of victory over another candidate. Certain details left out of a poll, like how stable that margin has been over time, might indicate that the narrow margin is more of a sure thing than it appears. Market data may reflect this by giving the candidate with a narrow margin of voter support a much higher probability of winning than his opponent, even though looking at a single poll would not suggest that such a probability is warranted. In essence, polling data may contain less information than market prices.

6. Votes for political candidates in the primary races are just one of several inputs into the ultimate selection of the Democratic presidential nominee. Besides the voting that takes place in the primaries, caucuses and superdelegates (party leaders and elected officials) have considerable influence over who eventually wins.

7. Many polls allowed respondents to choose an "Other" option if they did not wish to vote for one of the listed candidates. At times and depending on how it was calculated, the

Other category had enough support to count as either the primary or runner-up prediction. However, polls were too heterogeneous in their definition of the Other category and it was not possible to reconcile these differences. Instead, this category was ignored for the purposes of this dissertation.

Each of these features of polls means that any comparison between polling data and market data will be less than perfect. However, the reason people pay attention to polling data is because they hope it will provide some guidance as to who the ultimate winner will be. Thus, it is common practice in the media and among consumers of news to discuss polls *as if* they are predictions of who will win. From a practical perspective, then, it is worthwhile to ask whether polling data provides greater predictive accuracy than market data when both are available. With that in mind, this dissertation treats all polling data homogeneously. The relative frequency of stated preferences for particular candidates in polling data are assumed to be predictions about who will win.

2.2 Results

In every time period considered, market prices were a better predictor of the ultimate winner of both the 2004 and 2008 Democratic nomination contests. As Table 2 shows, Intrade prices correctly predicted the winner of the 2004 nomination 36% of the time polling data was available over all dates considered, while the polls themselves made the correct prediction only 13% of the time. For the 2008 nomination, market prices correctly predicted the winner 54% of the time polling data was available over all dates considered, while the polls themselves made the correct prediction only 45% of the time. It is also true that, for both the 2004 and 2008 nominations, no sub-period featured any deviation from that overall result.

For the 2004 nomination, no polls were available closer than 100 days prior to the convention on 29 July 2004. For the 2008 nomination, only one poll was available closer than 65 days prior to the convention on 25 August 2008. This is likely because the conventions in both years were mere formalities, since the state primaries (when votes are cast) and pledges by superdelegates may have, for all practical intents and purposes, already decided the eventual nominees. Polling organizations ceased conducting polls at those points, presumably because the results would no longer be considered newsworthy. However, it is interesting to note that market prices for most candidates continued to fluctuate around a fairly narrow band right up until the actual conventions, which means some uncertainties about who would ultimately win the nominations remained.

Because polls were not available in the immediate days prior to the Democratic National Conventions in 2004 and 2008, any analysis of short- versus long-term predictive capabilities needs to be shifted back to the period when the most recent polls were taken. Unfortunately, this makes the results here less comparable to those produced by Berg et al, although some interesting observations about short- versus long-term predictions can be made in any case. For the 2004 nomination, all but one sub-period featured some proportion of market prices that made correct predictions. On the other hand, only the final sub-period of time studied featured any polls that made correct predictions. For the 2008 nomination, the entire share of correct

Table 2: Predictions Made by an Average of All Polls Taken on Any Given Day Compared to Predictions Made by Intrade Markets

Days Included In Sample	**Item**	**2004 Nomination**	**2008 Nomination**
All (from beginning of polling)	# Polling Days	84	216
	# Polls Correct (%)	11 (13%)	97 (45%)
	# Market Prices Correct (%)	30 (36%)	117 (54%)
	p-value (one-sided)	0.0000	< 0.0105
501 to 622	# Polling Days	8	14
	# Polls Correct (%)	0 (0%)	0 (0%)
	# Market Prices Correct (%)	6 (75%)	0 (0%)
	p-value (one-sided)	0.0000	1
401 to 500	# Polling Days	5	24
	# Polls Correct (%)	0 (0%)	0 (0%)
	# Market Prices Correct (%)	5 (100%)	0 (0%)
	p-value (one-sided)	0.0000	1
301 to 400	# Polling Days	17	23
	# Polls Correct (%)	0 (0%)	0 (0%)
	# Market Prices Correct (%)	6 (35%)	0 (0%)
	p-value (one-sided)	0.0000	1
201 to 300	# Polling Days	36	39
	# Polls Correct (%)	0 (0%)	1 (3%)
	# Market Prices Correct (%)	0 (0%)	2 (5%)
	p-value (one-sided)	1	0.2642
101 to 200	# Polling Days	18	98
	# Polls Correct (%)	(61%)	79 (81%)
	# Market Prices Correct (%)	(72%)	97 (99%)
	p-value (one-sided)	0.2345	0.0000
66 to 100	# Polling Days	0	17
	# Polls Correct (%)	–	17 (100%)
	# Market Prices Correct (%)	–	17 (100%)
	p-value (one-sided)	–	1
32 to 65	# Polling Days	0	0
	# Polls Correct (%)	–	–
	# Market Prices Correct (%)	–	–
	p-value (one-sided)	–	–
6 to 31	# Polling Days	0	0
	# Polls Correct (%)	–	–
	# Market Prices Correct (%)	–	–
	p-value (one-sided)	–	–
Last 5	# Polling Days	0	1
	# Polls Correct (%)	–	1 (100%)
	# Market Prices Correct (%)	–	1 (100%)
	p-value (one-sided)	–	1

poll and market price predictions were generated in the last half of the total period studied. In the first 322 days of polling, no correct predictions were made by either polls or market prices.

2.3 Better Than Chance?

It is worth asking whether the general result that market prices are a more accurate predictor of the ultimate winner is better than chance. Consider the result for between 101 to 200 days prior to the election with respect to the 2004 primaries. Market prices accurately predicted the ultimate winner 72% of the time that polling data was available, which means that market prices were correct 13 out of 18 times. Over the same time period, polls correctly predicted the ultimate winner 61% of the time, which means poll values were correct 11 out of 18 times. The difference of between the 13 correct predictions from market prices and 11 correct predictions from polls does not appear so large that the observed result could not have happened by chance. To formally assess whether market prices outperform polls by an amount greater than could be expected by chance alone, the following procedure was carried out:

1. Find the most probable data generating process for correct vs. incorrect predictions made by polls.
2. Assuming that market prices are operating under the same data generating process as polls (that is, market prices are no better at predicting than polls on average), find the probability that we would discover the particular set of observations in the sample that we did.

The particular example discussed above will be analyzed and then a discussion of the results for other periods will follow. It is possible to think of polls as predicting the ultimate winner with some probability and predicting any other candidate to win with some other probability. Given that polls predicted the winner 11 out of 18 times in the example, the most probable data generating process for poll predictions will feature a distribution where the expected value is 11. Quite simply, it is assumed that polls correctly predict the winner with a probability of 11/18.

Next a table of values is generated (See Table 8 in the Appendix) that identifies the complete probability density function for this data generating process. The following variables are inputs:

ψ = total number of poll observations

ξ = probability of a correct prediction

α_i = the number of correct predictions that are possible, from 0 to 18.

β_i = Factorial(number of correct predictions), which represents the number of ways (or permutations) a given number of predictions can be observed.

δ_i = the first column inverted, or the number of ways incorrect predictions can be made.

ϵ_i = Factorial(total observations), which is the number of ways any group of 18 things can be arranged.

$\phi_i = \frac{\epsilon_i}{\delta_i \beta_i}$, which is number of combinations of correct and incorrect predictions that can be made, and it strips out all of the redundancies that exist in the permutations[10].

$\rho_i = \xi^{\alpha_i}(1-\xi)^{\psi-\alpha_i}$, which is the single permutation probability, or the probability of observing a particular pattern of correct and incorrect predictions.

$\chi_i = \rho_i \phi_i$, which is the probability of observing any particular value of α_i given the assumed data generating process. $\sum_{i=0}^{\psi} \chi_i = 1$ by definition.

$\sum_{i=n}^{\psi} \chi_i$ is the probability of observing at least n correct predictions, and comprises the one-sided p-value that indicates the likelihood that market prices actually have the same predictive prowess as polls (have the same underlying data generating process) even though it is observed that market prices predicted the ultimate winner a greater number of times than did polls.

In the example currently under consideration, $\sum_{i=13}^{\psi} \chi_i =$ gives the probability that market prices would predict the ultimate winner correctly at least 13 times out of 18 if market prices actually have the same underlying predictive power as polls. As shown in Table 2 and illustrated in Figure 10, the one-sided p-value for this case is 0.2345. This value is relatively high compared to every other one-sided p-value reported in the table. If the relevant metric is the 5% level of significance, it indicates that, in this particular instance, it is not possible to reject the hypothesis that market prices are no better than polls at predicting the ultimate winner[11]. However, all of the other results indicate that, at the 5% level of significance, it is possible to reject the hypothesis that market prices are no better than polls at predicting the ultimate winner. Thus, overall it is reasonable to assume that market prices are, in fact, better predictors of the ultimate winner than are polls.

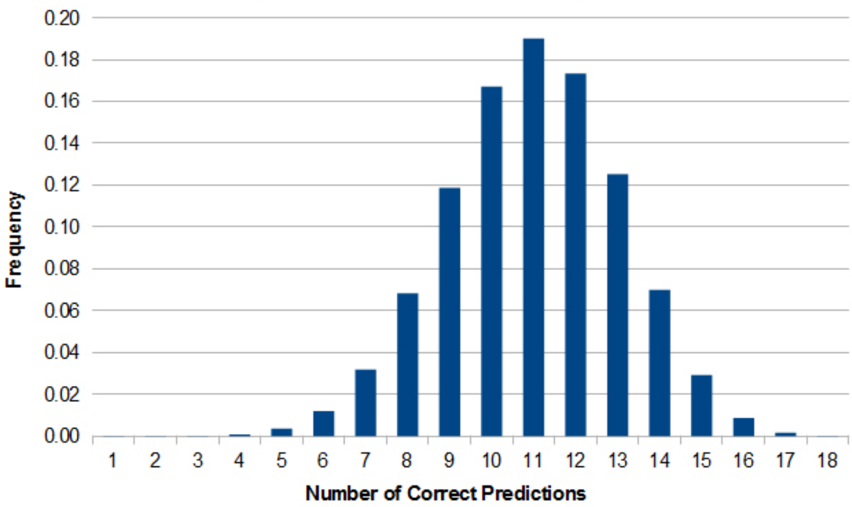
2.4 Prediction Strength

During portions of the 2004 and 2008 campaigns when both polls and market prices were making incorrect predictions about the ultimate winners, market prices and polls appeared to have somewhat similar differentials between the candidate predicted to win versus the candidate predicted to be the runner up. In the late stages of the 2004 and 2008 campaigns, when it was becoming more clear who would ultimately win the nominations, market prices seemed to feature much greater differentials between the predicted winner and the runner up. This section seeks

[10] Suppose a correct prediction is represented by the letter c and an incorrect prediction is represented by the letter i. There are three permutations where one correct prediction is made and two incorrect predictions are made (iic, ici, cii), but only one unique combination (because order doesn't matter when computing combinations).

[11] Note that the calculations involved in solving for the one-sided p-values reported in Table 2 are rather rough estimates, an unfortunate and unavoidable result of the inability of any known commercial software program to accurately store the values for large factorials. Calculating the one-sided p-value for all time periods for the 2008 primaries posed particular challenges because it was not possible to get even an estimate of any factorial beyond 170. Thus, to find the likely p-value for 216 total observations, it was necessary to calculate the p-value for 150, 160, and 170 observations, using a correct prediction percentage of 54% for each, and infer from the change in the p-value from 150 to 160 and from 160 to 170 that the p-value for 216 observations was < 0.0105.

Figure 10: Probability Distribution Function for 18 Observations if Correct Prediction Made With Probability 0.61



to quantify that differential and raises the possibility that there may be important differences in information content between the two modes of forecasting.

The following equations were created to develop a metric to measure the relative strength of predictions made by Intrade market prices and polling figures:

$$\beta_{correct} = \pi_{poll}/\kappa_{poll} \tag{1}$$

where $\beta_{correct}$ is prediction strength when a correct prediction is made by polls, π_{poll} is the poll predicted vote share for the ultimate choice for nominee (John Kerry in 2004 and Barack Obama in 2008), and κ_{poll} is the poll predicted vote share for the candidate with the second highest ranking in the polls. The calculation for prediction strength is similar when an incorrect prediction is made:

$$\beta_{incorrect} = \pi_{poll}/\psi_{poll} \tag{2}$$

where ψ is the predicted vote share for the leading candidate.

Similarly, prediction strength for market prices is given by

$$\alpha_{correct} = \pi_{mkt}/\kappa_{mkt} \tag{3}$$

and

$$\alpha_{incorrect} = \pi_{mkt}/\psi_{mkt} \tag{4}$$

where $\alpha_{correct}$ is prediction strength when a correct prediction is made by market prices, π_{mkt} is the market given probability for the ultimate choice for nominee to win the nomination, κ_{mkt} is the market given probability for the nominee with the next highest probability of winning, and ψ_{mkt} is the market given probability with the candidate with the highest market given probability of winning when that candidate was not the ultimate choice for nominee (i.e., someone other than John Kerry in 2004 and Barack Obama in 2008).

These calculations were performed for each instance where a market price or poll is available. The values reported in Table 3 are averages of those figures and are categorized by whether the prediction was ultimately correct or not.

Table 3: Average Relative Strength of Predictions Made by Market Prices Compared to Polls

Date	Prediction Outcome	Strength Mkt Prices (α)	Strength Polls (Average) (β)
2004	Correct	4.45	2.39
	Incorrect	0.3	0.31
2008	Correct	6.82	2.39
	Incorrect	0.5	0.68

An interesting relationship emerges from these results. In 2004, when market prices and polls were both giving the wrong signal about who the nominee would be (predicting something other than a Kerry win), the strength of the signal sent by the market was almost identical to the strength of the signal sent by polls. However, when market prices and polls were correctly predicting the eventual nominee, the market strength indicator was nearly twice as large as that for polls. For the 2008 data, when both market prices and polls were producing incorrect predictions (predicting something other than an Obama win), polls were 36%[12] "less wrong" than were market prices. On the other hand, when both polls and market prices were correct, the strength of the predictions coming from market prices was, on average, about 285% greater than that for polls.

Great care is required when interpreting these figures. It must be noted once again that polling data "asks" a different question of participants than is asked of market participants when they choose to place a bet. Polls could predict a narrow margin of victory for a particular candidate, but that narrow margin could nonetheless be a fine prediction of the ultimate outcome. So just looking at a single poll and the relative strength of a poll's prediction compared to the corresponding market prediction may be unfair.

On the other hand, the observed higher prediction strength for market prices compared to polls may simply reflect greater information content. In the example given above, perhaps market participants would recognize that the narrow margin of expected victory was quite stable over time and bid up the market price such that the implied probability could be much higher than the winner's vote share predicted by polls. The resulting strength of market price predictions could then be seen as a genuine improvement in predictive power. In contrast, each individual poll would simply reflect the thoughts of any given sample of polling participants at a single moment in time. The lower strength of the polls' predictions could be seen to be a result of a lack of information content in polling responses.

[12] 0.68/0.50 = 1.36, meaning that the strength of the prediction for Obama was stronger for polls than it was for markets by about 36%

2.5 Reasons for Superior Forecasting Ability of Market Prices

In Berg et al. (2008b), one of the reasons given for the result that market prices so consistently out-performed polling predictions was that polling predictions were highly volatile. In their paper, Berg et al produce a five-poll moving average with the intention of removing some of the volatility and hopefully improving the measured predictive power of polls. Indeed, they found that the moving average of polls was moderately more accurate than the individual polls, although market prices were still better predictors overall.

Thus, when considering why the market prices from Intrade tended to be better predictors of the ultimate winner of the two Democratic presidential nomination contests, variance is one factor that ought to be considered. Table 4 reports variance of market prices and polling percentages by candidate, providing separate statistics for individual polls and an average of polls.

The first interesting feature of these figures is that the variance for market prices is often higher than the variance for either individual polls or an average of the polls. The average variance of market prices for the 2004 market calculated over all candidates, at 175, is almost twice as high as the average variance of polling percentages calculated over all candidates at 93 (96 for the average of all polls). In the 2008 market, a similar story emerges. The average variance of market prices for the 2008 market calculated over all candidates, at 53, is more than twice as high as the average variance of polling percentages calculated over all candidates at 21 (23 for the average of all polls). If only those candidates that the 2008 market and polls have in common are considered, the ratio widens even further, since the average market price variance over all candidates becomes 129.

Interestingly, the average of all polls has a lower variance than the the individual polls for some candidates, but the opposite is true for other candidates. Overall, the average of the polls actually has a higher average variance over all candidates than the individual polls.

At least for the 2004 and 2008 Democratic presidential nomination contests, then, lower relative variance of market prices cannot be the cause of the superior performance of market prices as predictors of the ultimate winner.

It must be acknowledged that an unfortunate shortcoming of the analysis of the accuracy of prediction markets in this dissertation is a result of limited data. Ultimately, this dissertation has assessed the ability of Intrade's bettors and polling respondents to correctly predict that the winner of the 2004 Democratic presidential nomination contest would be John Kerry and that the winner in 2008 would be Barack Obama. Although the long run accuracy of both market prices and polls is analyzed by considering a number of distinct time periods over which data was available, that should not disguise the fact that there are just two events across which accuracy can be judged.

Furthermore, the two events analyzed in this dissertation are not identical: several states altered their election rules between 2004 and 2008, and although this is unlikely to make the comparison completely invalid, it is important to recognize that the analysis has some drawbacks when compared to, say, a controlled experiment where all factors are held constant across experiments.

Ideally, accuracy of any prediction mechanism is best computed by analyzing a large number

Table 4: Variance (σ^2) of Market Prices and Polling Percentages by Candidate

Primary	**Candidate**	**Mkt Prices**	**Polls (Individual)**	**Polls (Average)**
2004	Kerry	684.26	215.01	244.54
	Lieberman	49.66	25.22	22.92
	Clinton	8.61	99.57	120.89
	Gephardt	7.58	10.64	8.79
	Edwards	27.59	16.75	17.08
	Dean	537.8	62.35	58.29
	McCain	0.32	–	–
	Rest Of Field	82.46	220.22	199.7
2008	Obama	798.62	133.8	143.21
	Clinton	475.95	17.1	15.51
	Kerry	0.02	24.5	24.5
	Gore	5.24	4.94	4.94
	Richardson	1.47	16.6	19.77
	Edwards	9.64	15.11	17.11
	Clark	0.06	0.59	0.51
	Dodd	0.02	0.32	0.26
	Vilsack	0.25	0	0
	Biden	0.21	1.32	1.25
	Corzine	0.872	–	–
	Feingold	2.40	–	–
	Lieberman	0.06	–	–
	Bayh	15.71	–	–
	Leahy	0.12	–	–
	Ford	0.15	–	–
	Rendell	1.88	–	–
	Powell	0.05	–	–
	Warner	75.62	–	–
	Bredesen	0.42	–	–
	Schweitzer	0.06	–	–
	Easley	0.06	–	–
	Daschle	0.06	–	–
	Blagojevich	0.00	–	–
	Dean	1.44	–	–
	Rest Of Field	0.00	–	–

Notes: Missing entries for polls reflect the fact that polling organizations did not gather data on a number of candidates who did not officially announce that they wished to be considered for the Democratic nomination. Market data was available for these unofficial candidates because Intrade often adds contracts at the request of market participants even if the contract requested is in some way non-standard.

of event outcomes and comparing the outcomes to the predictions. While this is reasonably easy to do in some scenarios (such as for sporting events that occur on a regular basis), this is often impossible in the case of other real world events that we would like to study. Many events occur just once, and for those that do have a longer track record, data may not be available. The two major political parties have been holding conventions to nominate a presidential candidate for decades, although it is only for the two most recent ones at the time of writing that Intrade has featured a market in these events.

Despite these caveats, the results presented in this dissertation are in keeping with other, similar research projects (such as those found in Berg et al). Although any one paper analyzing the predictive accuracy of market prices compared to other forecasting tools may be subject to numerous shortcomings and problems, the sum of them ought to give a reasonably accurate overview of the true state of affairs.

3 Intrade Prediction Market Efficiency

This section uses three techniques to analyze the efficiency of the Intrade prediction markets. The first describes the expected features of prices of individual contracts and also the relevant market portfolios (the sum of the individual portfolios), taking note of the frequency of pricing anomalies. Some pricing anomalies can be expected to persist in an efficient market, but extreme anomalies should not persist. The second is an analysis of the observed corrections of pricing anomalies. The third discusses the case for increased pricing efficiency with respect to a particular measure of market liquidity, open interest. Finally, the fourth is an investigation into whether the prices of the individual contracts follow a random walk and whether they cointegrate in such a way that the market portfolio displays characteristics of internal market coherence.

3.1 Pricing Anomalies

If Intrade's markets were operating efficiently, it would be reasonable to expect the combination of all market prices (the "market portfolio") to never wander very far from a value of 100. If ever the market portfolio dropped to a value of, say, 90, it would be possible for market participants to purchase the market portfolio at 90 and receive a virtually guaranteed payoff of 100. Naturally, that process would continue until the market price of the various contracts increased such that the nearly riskless strategy described here ceased to be profitable. It should be noted, however, that this strategy is riskless only if one of the market listed candidates is sure to win (or a Rest of Field contract is available). It is always possible that an unforeseen and heretofore unlisted candidate enters and wins the contest after one enters into the agreement to purchase the market portfolio, meaning that the initial investment goes entirely unrecompensed.

Similarly, if the market portfolio had a value of, say, 110, market participants could short sell all listed contracts for a gain of 110 and be forced to pay out only 100 on the contract representing the one winner. It is interesting to note that if the last strategy was nearly risk free, this strategy has the potential to be truly risk free. In the unlikely event that an unforeseen outsider wins after one has purchased the market portfolio, the proceeds of the short sale are pure profit, because all short sold contracts will expire at zero in this scenario[13]. From this information alone, it is possible to infer that any pricing irregularities with respect to the complete portfolio are more likely to be found at values below 100 (at least in the 2008 market), although other factors that may counterbalance this effect will be described shortly.

The market portfolio price is calculated by summing the closing prices for all contracts, where closing prices represent the value of the last trade at market close. The problem for market participants who wish to engage in the trading strategy above is that they may be unable to trade at the last trading price, depending on the bid-ask spread at the time. For low volume contracts, the bid ask spread can be considerable. Unfortunately, the data collected from Intrade for the purposes of this dissertation did not include bid-ask details, making it impossible to determine the magnitude of the effect and the potential for preventing prices for the market portfolio from closely hewing to 100.

Another factor affecting arbitrage opportunities at Intrade is the fee structure. Market participants pay either \$0.03 USD (for price makers[14]) or \$0.05 USD (for price takers[15]) per lot (a lot is a bundle of 10 contracts, and the minimum tradable quantity) when entering into an agreement to buy or sell contracts. Additionally, contracts that expire at 100 (winning contracts) are subject to an expiry fee of \$0.10 per lot.

To provide an idea of the magnitude of the effect of transaction fees on arbitrage opportunities, consider a scenario where the market portfolio of seven candidates plus a Rest of Field[16] contract is valued at 110. One lot of each contract could be short sold, providing an immediate gain of \$11. From this, assuming the seller is a price taker, an initial trading fee of 0.05x8 = 0.40 is subtracted, leaving \$10.60. The expiry fee of \$0.10 for the one contract that ultimately winds up closing at 100 further reduces the seller's profit to \$10.50.

Given this analysis, the lowest value for the market portfolio that would present a no risk trading strategy is 105.1 (\$10.51 minus 10.50 in fees, leaving \$0.01 as the net gain). The outlook for this strategy worsens as the number of contracts that comprise the market portfolio increases. Given that the market portfolio in the 2008 market as comprised of 25 contracts[17], the lowest value that would present a profitable strategy for selling the market portfolio here is 113.6 (\$11.36 from the initial sales less \$1.25 in initial fees less \$0.10 in expiry fees leaves \$0.01 as the net gain).

From a buyer's perspective, the highest value of the market portfolio that would present a profitable opportunity to purchase the market portfolio in the case where there are seven market

[13] For this to be true, it should be the case that there is no contract that represents all other possible candidates. In the 2004 market, there was a contract for Rest of Field, which had this purpose. In the 2008 market, however, that contract was introduced only in late 2007 trading, so there was time for traders involved in short selling the market portfolio to limit their risk to zero when the market portfolio was priced highly enough.

[14] A Price Maker Order is defined as any order placed on the Exchange that is not immediately matched/filled by an existing BID/ASK on the system.

[15] According to Intrade's guidelines, A Price Taker Order is defined as any order placed on the Exchange that is immediately matched/filled by an existing BID/ASK on the system.

[16] A Rest of Field contract was available for the duration of the operation of the market for the 2004 nomination.

[17] For most of the duration of the operation of the market for the 2008 nomination, a Rest of Field contract was *not* available and when it did become available the price stayed so low as to be negligible. As discussed earlier, the period when the Rest of Field contract was not available actually decreased the risk to traders involved in selling the market portfolio, since there was some probability that no listed contract would expire at 100 (that is, no listed candidate would win the nomination).

participants and a contract that represents the rest of the field is 94.9. The lots comprising the portfolio will cost \$9.49, the initial trading fee will be \$0.40, and the expiry fee on the lot that eventually closes in the money will be \$0.10, so that the profit from selling the expired contract for \$10 will be reduced to \$9.50, leaving a \$0.01 net gain. Again, the outlook for this type of strategy worsens if we consider the 2008 market with 25 contracts. The highest value of the market portfolio that would present a profitable opportunity to purchase the market portfolio under these circumstances is 86.4[18]. The lots comprising the portfolio will cost \$8.64, the initial trading fee will be \$1.25, and the expiry fee on the lot that eventually closes in the money will be \$0.10, so that the profit from selling the expired contract for \$10 will be reduced to \$8.55, leaving a \$0.01 net gain.

Thus, it is expected that market portfolio values in the range of 95 to 105 are relatively stable for the 2004 market, while market portfolio values in the range of 86 to 114 are relatively stable for the 2008 market. One caveat to this is that a market portfolio price below or above 100 suggests that at least one contract is improperly valued, and if market participants could identify it they could profitably purchase it without incurring the same level of trading fees required of those following the riskless trading strategy. On the other hand, the discovery costs of identifying the single mispriced contract are undoubtedly much higher than the costs of identifying and taking advantage of the "dumb" strategy of buying or selling all contracts.

Finally, it has been suggested that Intrade's margin rules may discourage short selling strategies (Friedman, 2007). Intrade requires 100% margins, so a seller of, say, all seven contracts would need to keep his Intrade account stocked with at least \$70 for the duration of his bet. Of course, it is not possible that all seven contracts will expire at 100, but Intrade's margin rules do not carry exceptional provisions for those who are engaged in this type of arbitrage. In any event, money that must be kept in Intrade's accounts is money that cannot be earning interest elsewhere. Intrade pays interest on account balances of at least \$20,000 USD (Delaney, 2010), but the rates are relatively low (similar to those accruing to standard checking accounts). It is difficult to assess the magnitude of this effect because it would require an estimate of the discount rates that Intrade market participants have, but the direction of the effect is clear. A 100% margin rule for short sellers that does not take into account situations where a short seller is selling all listed contracts within a single category will tend to raise the upper bound of the expected trading range. In other words, the effect is to increase the likelihood that the market portfolio will trade at values greater than 100.

Table 5 shows the number of days (and percentage of time) that the market portfolio price fell between selected ranges. Of particular interest is the number of times the portfolio price fell into extreme ranges. Of the dates that witnessed total portfolio prices in excess of 116 for the 2008 nomination, the great majority were in 2004 and only two such events occurred after January 2005. That is, only in the early operating days of the market were extreme values at the high end somewhat common. Still, there were 80 days on which the closing price was between 111 and 116, comprising 6% of the total market operating period.

For both the 2004 and 2008 markets, it is clear that overpricing of the market portfolio is

[18]This example is calculated assuming a Rest of Field contract is not purchased. To account for that in the later days of trading in this market, simply include an additional trading fee of \$0.05 into the analysis.

more likely than underpricing. Pennock et al. (2000) noted the same thing in their analysis of Hollywood Stock Exchange prices for portfolios of same-category Oscar nominees as well in the Iowa Electronic Markets for portfolios of political candidates where there was only one possible winner. Pennock et al suggested that this was due to a preference among market participants to buy rather than sell, but did not describe the mechanisms or incentives that might have encouraged that preference.

With respect to the Intrade markets, it was found that short selling provides one advantage (the potential for pure profit if an individual who was not listed on the exchange ultimately won) that would suggest that pricing irregularities for the market portfolio might be more likely to be found at values below 100. But short selling also has a major disadvantage (the 100% margin requirements) that would suggest the opposite. Based on the evidence in Table 5, it may be reasonable to conclude that the magnitude of the disadvantage related to short selling outsizes the specified advantage. Perhaps that is the mechanism that gives rise to the preference that Pennock et al noted for market participants to prefer buying rather than selling, although it is of course possible that there is some as yet unidentified influence that trumps the possibilities discussed here.

One such influence that could fill that role is "favourite-longshot bias", the oft-noted tendency for favourites to be underpriced and for longshots to be overpriced (Snowberg and Wolfers, 2010; Weinbach and Paul, 2008; Koch and Shing, 2008). That is, bets on events with high prices have higher expected returns than bets on events with low prices. The upshot is that if longshots tend to be over-bet by more than favourites are under-bet, an overall overpricing of the entire category would result. This possibility will not be further analyzed here, but it is noted that it may be a useful avenue for further research.

Table 5: Days (and Percentage of Time) That the Market Portfolio Price Fell Between Selected Ranges

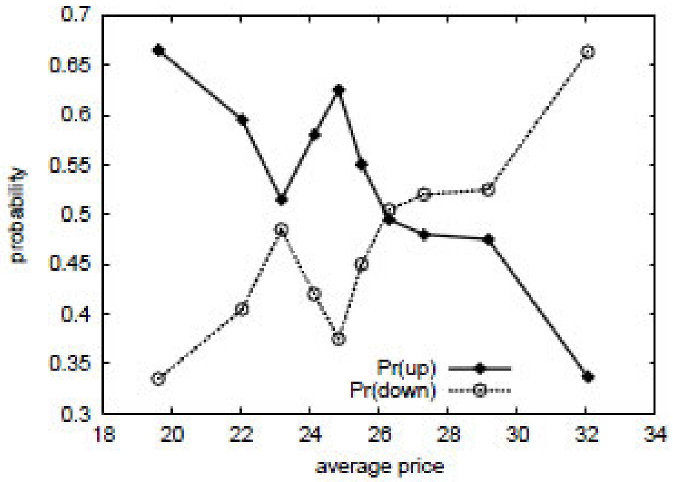
Price Range	**2004 Nomination**	**2008 Nomination**
81 to 90	34 (6%)	0 (0.0%)
91 to 94	3 (0%)	2 (0.1%)
95 to 105	433 (71%)	1028 (75%)
106 to 110	111 (18%)	238 (17%)
111 to 116	32 (5%)	82 (6%)
117 to 125	– –	18 (1%)
Total	613 days	1368 days

3.2 Corrections of Pricing Anomalies

In their analysis of the HSX and IEM Pennock et al. (2000) note that bundles of contracts in the same category ought to sum to some consistent value so that the implied probability of the sum of the events never rises above one. For example, bundles of contracts for the winner of an Oscar or Emmy award should not rise above H$25 on the HSX exchange and a complete bundle of contracts for candidates competing for the New York Senate seat in 2000 should not rise above $1 on the IEM exchange.

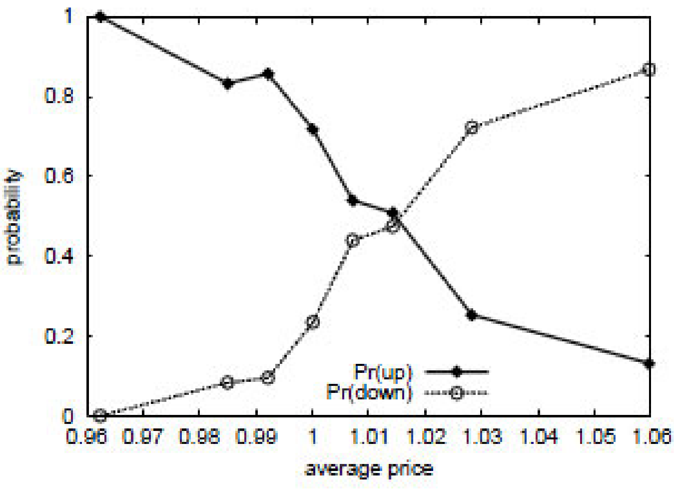
In their paper, Pennock et al found that prices do sometimes converge from these upper bounds (by as much as 40%), although they are very likely to move back towards their expected prices in subsequent trading periods. By dividing the bundles into groups, tracking their prices, and observing the fractions of the bundles that moved up or down in price in subsequent trading periods, Pennock et al composed two graphs that give the probability that a bundle will rise or fall in price given its current price. These graphs are reproduced in this dissertation as Figures 11 and 12.

Figure 11: Fraction of Bundles That Move Up (Down) in Price After Four Hours, HSX



Source: Pennock et al. (2000)

Figure 12: Fraction of Bundles That Move Up (Down) in Price After Four Hours, IEM



Source: Pennock et al. (2000)

Pennock et al point out that the most interesting feature of both of these graphs is that the "crossover point", where prices are as likely to rise in the following trading period as fall, is above H$25 in the HSX example and above $1 in the IEM example. In other words, as was discussed in Section 3.1, the authors found that there was a tendency to overprice complete bundles of candidates.

To compare Pennock et al's results to the results found in this dissertation, two graphs were produced in the fashion of Figures 11 and 12. The graphs produced using Intrade data for the 2004 and 2008 Democratic presidential nominations can be found in Figures 13 and 14. Note that, unlike Pennock et al's charts, the probabilities do not always sum to one because of instances where the price moved neither up nor down, but sideways.

Figure 13: Probability of Price Movements Based on Current Price Range, 2004 Democratic Presidential Nominee Market

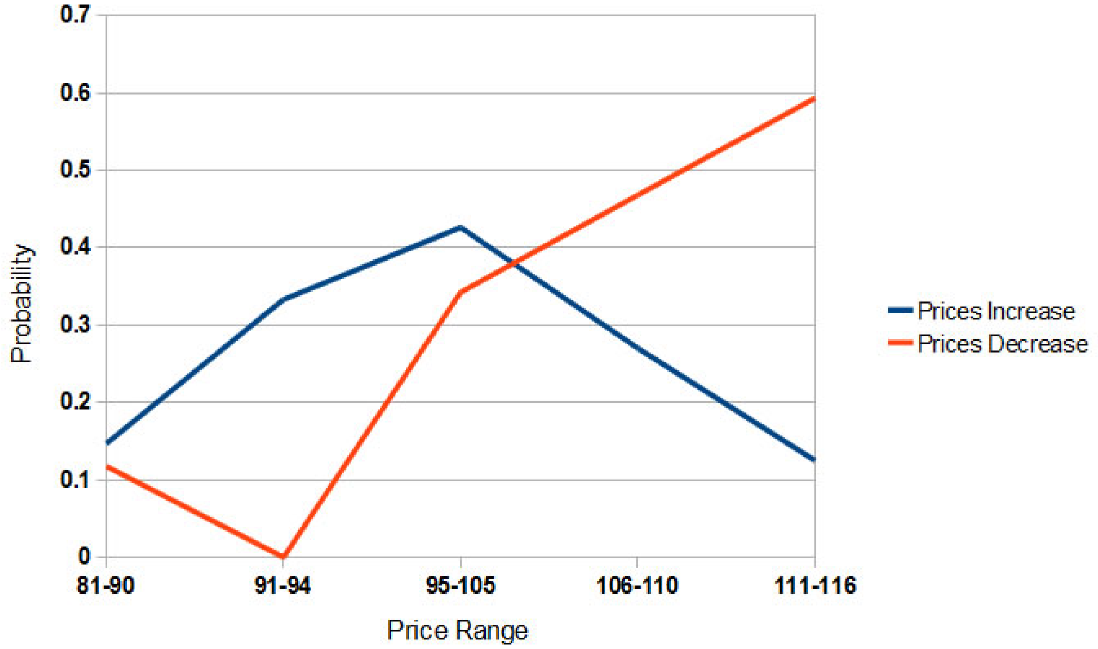
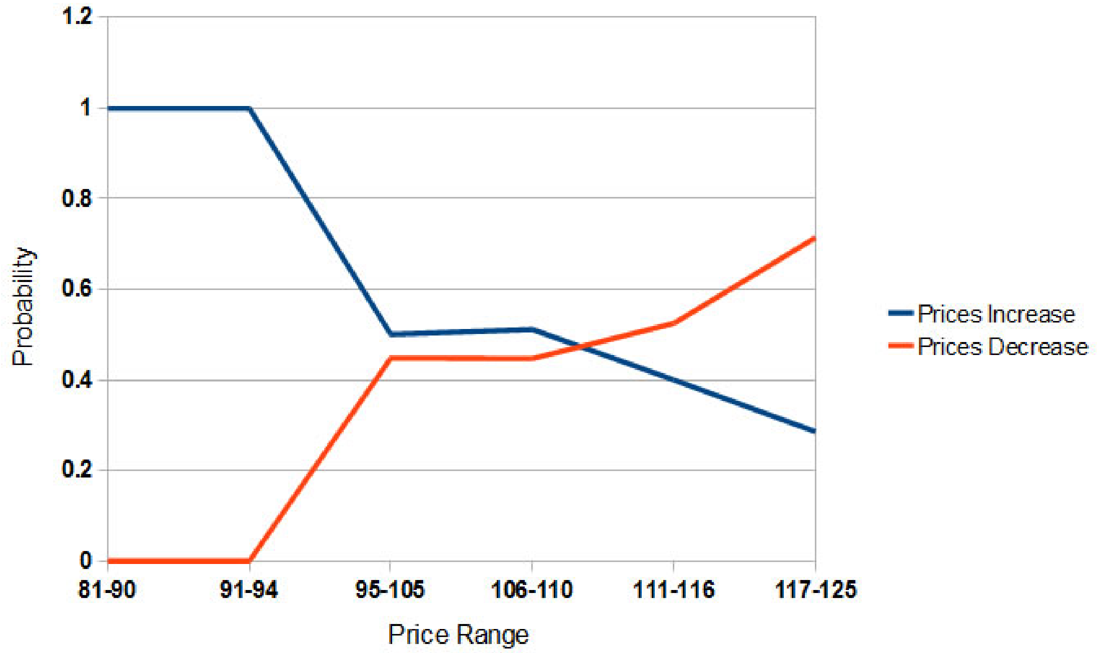


Figure 14: Probability of Price Movements Based on Current Price Range, 2008 Democratic Presidential Nominee Market



As was the case with Pennock et al's results, low prices for the market portfolios in the 2004 and 2008 markets show a tendency to rise and high prices show a tendency to fall. Furthermore, the "crossover point" for the 2004 market is at a price higher than 100; the crossover point for the 2008 market is higher still.

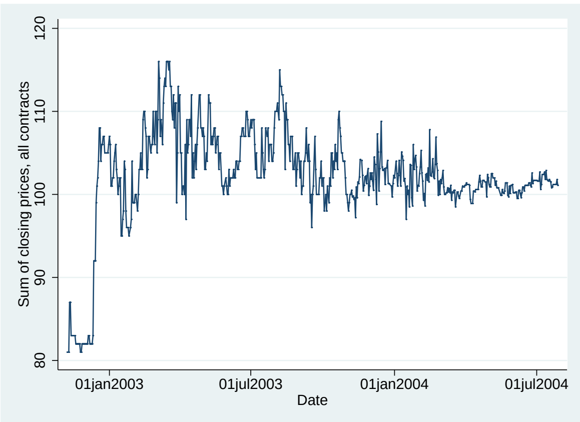
3.3 Variance of the Market Portfolio and Market Liquidity

Two characteristics of efficient markets is that the average value of the market portfolio should hew closer to 100 as the liquidity of the market increases. Another characteristic is that the variance around the mean value should decrease as the liquidity of the market increases. The liquidity measure available in the data was open interest, which measures the total value, in US dollars, of all contracts held by market participants at any given date.

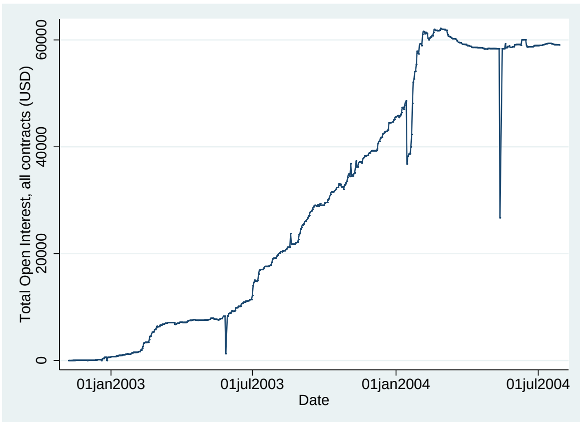
For the 2004 market, total open interest began at \$0, hit an all time high of \$62,166 on 27 February 2004, and closed at \$59,049 just prior to the Democratic convention on 28 July 2004. For the 2008 market, total open interest began at \$0, hit an all time high of \$639,022 on 27 August 2008 and closed at \$637,874 just prior to the Democratic convention on 28 August 2008.

A visual analysis of Figures 15(a), 15(b), 16(a), and 16(b) suggests that average prices tended to be less volatile and maintain a position closer to 100 as the date of the conventions approached and as liquidity in the markets increased.

Figure 15: Price of complete portfolio (all contracts) and total open interest at Intrade, 2004 Democratic Presidential Primaries

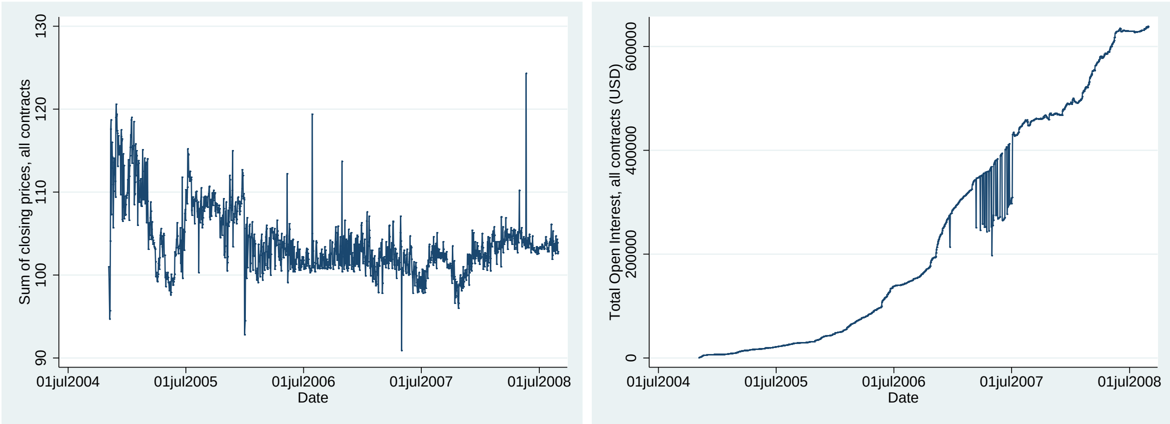


(a) Price of complete portfolio



(b) Total open interest

Figure 16: Price of complete portfolio (all contracts) and total open interest at Intrade, 2008 Democratic Presidential Primaries

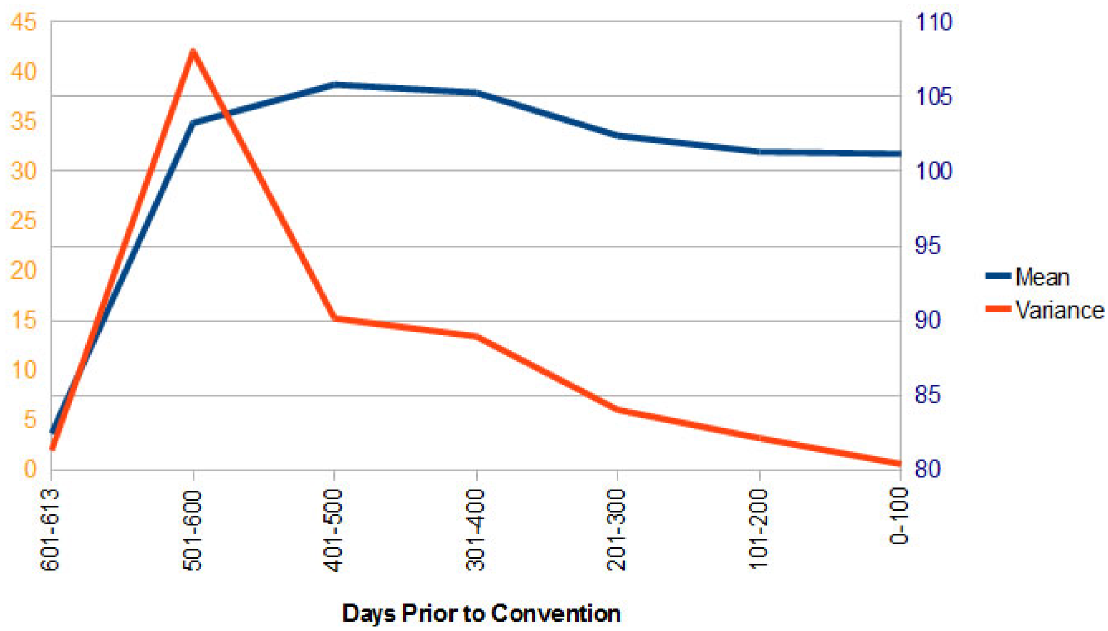


(a) Price of complete portfolio

(b) Total open interest

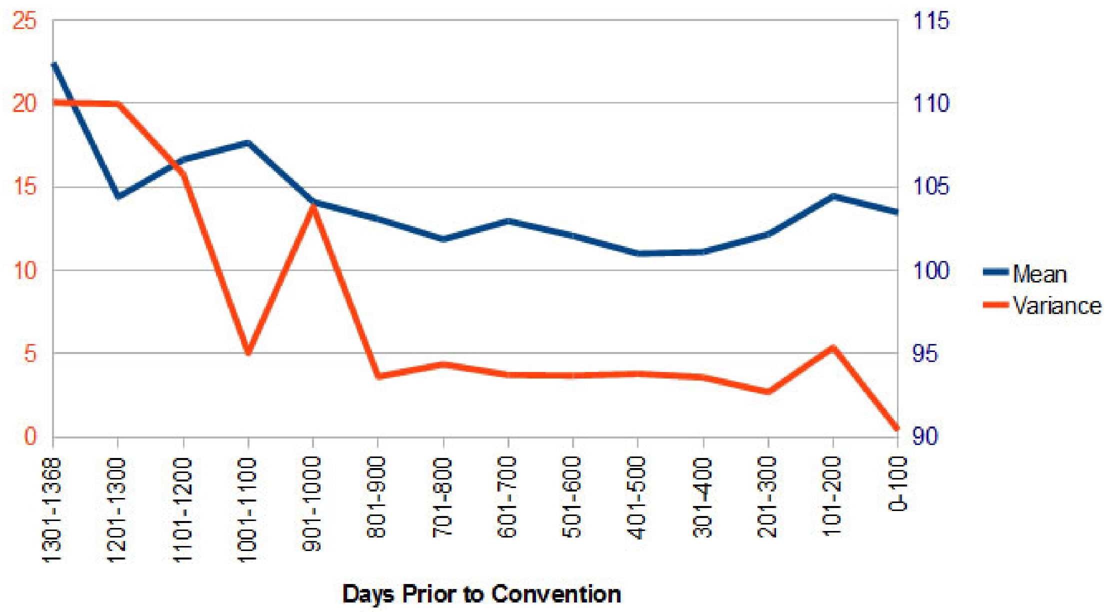
To quantify this, prices for the market portfolio were analyzed in six 100-day increments and one 13 day increment for the 2004 data and 13 100-day increments and one 68 day increment for the 2008 data. The mean and variance were then calculated for each of those periods. These figures can be found in the Appendix in Table 9 and are represented in the graphs in Figures 17 and 18.

Figure 17: Market Portfolio Mean and Variance For Select Timer Periods, 2004



Based on these graphs, it does appear that the mean became more tightly wound around values closer to 100 and the variance decreased as liquidity increased in both the 2004 and 2008 markets. In the 2008 market, there is a slight departure from that trend with respect to the mean beginning about 500 days prior to the convention and with respect to the variance

Figure 18: Market Portfolio Mean and Variance For Select Timer Periods, 2008



beginning about 300 days prior to the election. For both the mean and the variance, beginning at about 200 days prior to the convention, the characteristics of increased efficiency reassert themselves.

One confounding factor in this analysis is that observing an increase in liquidity along with the movement of two characteristics of efficient markets hardly demonstrates causality. Instead of looking at the 2004 and 2008 markets in isolation, it may be instructive to compare the two. It is worthwhile to note that the 2008 market for the Democratic presidential nomination had, at its height, 10 times greater liquidity if liquidity is measured as open interest. If liquidity is sufficient to produce improvements in how closely the mean hugs a price range near 100 and how low the variance is, the vast increase in liquidity in the 2008 market compared to the 2004 market should produce a correspondingly large improvement in those two characteristics.

In fact, a comparison of the mean and variances of the market portfolio in 2004 and 2008 does not make a clear case for liquidity's contribution to efficiency. Overall, the 2008 data for the market portfolio has a mean of 103.9 and a variance of 14.4, while the 2004 data for the market portfolio has a mean of 102.3 and a variance of 41.8. The figures for the means would suggest that the 2004 market was slightly more efficient than the 2008 market, but the figures for the variances suggest the opposite.

3.4 Testing For Unit Roots In Market Prices

If Intrade's markets are efficient, it is expected that certain characteristics of market prices will be present. One such characteristic is that prices are unpredictable. Market participants should not be able to predict future prices based on past price information and the best estimate of future prices should be the current price. In econometric parlance, prices should contain unit roots.

In order to test this, each contract's time series was subjected to an augmented Dickey-Fuller test, which tests the null hypothesis that γ in the following specification (Enders, 2010) is equal to zero:

$$\Delta y_t = a_0 + \gamma y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + \epsilon_t \tag{5}$$

Where:

y_t is the market closing price for a contract at time t
a_0 is an intercept term
$\gamma = -\left(1 - \sum_{i=1}^{p} a_i\right)$
$\beta_i = -\sum_{j=1}^{p} a_j$
Δy_{t-i} are the lags of y_t

The null and alternative hypotheses are:

H_0: $\gamma = 0$, suggesting that the process y_t contains a unit root
H_1: $\gamma \neq 0$, suggesting that the process y_t is stationary

Based on a visual inspection of each of the contract's prices over time, it was determined that Equation (5) should contain an intercept but not a trend term. An intercept is necessary because none of the time series has a zero mean. A time trend was deemed unnecessary, because there is no underlying reason to think any of the series increase or decrease at some deterministic rate. On the date of the Democratic convention, one series will inevitably go to 100 and the rest will inevitably go to zero, but the rate at which they go to those points is not constant and it's not obvious which contracts will fall into which category until the end of the time series. Furthermore, because the introduction of additional parameters for estimation reduces the degrees of freedom and therefore the power of the test to reject the null hypothesis of a unit root (Enders, 2010), including a time trend was deemed an unwarranted cost.

While too many lags will reduce the power of the test to reject the null of a unit root, too few lags may mean that the regression residuals do not behave like white noise. Thus, Akaike's information criterion (AIC) and Bayesian information criterion (BIC)[19] were used to determine the appropriate lag length for each variable tested for a unit root. In cases where the AIC and the BIC gave conflicting suggestions for lag length, unit root tests were performed using both possibilities.

In some instances, both the AIC and BIC criteria suggested using an extremely large number of lags. For example, in the case the Corzine, Easley, and Lieberman time series in 2008 and the McCain time series in 2004, at least one of the information criteria suggested lags greater than is recommended by a common rule of thumb proposed by Schwert (1989). Schwert suggested a

[19]The formula for the AIC is $2 * k - 2 * ln(n)$, while the formula for the BIC is $k * ln(n) - 2 * ln(L)$. k is the number of parameters in the model, n is the sample size, and L is the log likelihood for the model. The AIC and particularly the BIC criteria (because the latter punishes additional parameters more severely) help avoid overfitting.

maximum lag length (p_{max}) as follows:

$$p_{max} = int\left[12 \cdot \left(\frac{T}{100}\right)^{1/4}\right] \tag{6}$$

where T is the total number of observations and *int* means taking the integer portion of the content in brackets.

For the 2004 data, which contains 613 observations, the maximum lag length suggested by Scwhert's rule of thumb is 18. For the 2008 data, which contains 1368 observations, the maximum lag length suggested by Schwert's rule of thumb is 23. To avoid using an unnecessarily large number of lags, Schwert's rule of thumb was used to cap the number of lags used in the Augmented Dickey-Fuller test.

Table 6 reports the augmented Dickey-Fuller test statistic for each time series of closing market prices for the eight contracts (seven candidates plus a Rest of Field contract) on Intrade in 2004 and the 26 contracts (25 candidates plus a Rest of Field contract) with contracts on Intrade in 2008. With the exception of Edwards, Lieberman, Dodd, and Bredesen in the 2008 data, it is safe to conclude that all series examined are non-stationary. To ensure that these series are I(1) and are not integrated of some higher order, it was necessary to conduct a second set of Augmented Dickey-Fuller tests, this time in first differences. This test uses the same specification as Equation (5), except that y_t now represents the first difference of the variable to be tested rather than the level of the variable. Table 7 shows the results of this test. For all but the McCain series in the 2004 data, it is possible to reject the null hypothesis that the processes remain non-stationary. In the 2008 data, all series are shown to be stationary in first differences. Thus, with the exception of McCain in 2004, all series that were concluded to be non-stationary in the Augmented Dickey-Fuller tests in levels can be concluded to be I(1).

Except for the cases of Edwards, Lieberman, Dodd, and Bredesen for the 2008 data, the various lag lengths used did not change the results of the Dickey-Fuller tests. In the latter three of the four cases mentioned, the test using the shorter of the two lag lengths led to rejection of the null hypothesis, as would be expected by the greater power of the tests with fewer lag lengths. In the case of Edwards, it was the longer lag length that led to rejection of the null hypothesis. In the 2004 data, the choice of lag length did not have an impact on whether the null hypothesis of a unit root was rejected. In all cases in 2004, the null could not be rejected.

Another characteristic that is expected from Intrade's markets if they are efficient is internal coherence. As discussed in section 3.1, prices for the market portfolio should add to approximately 100. If so, the price series for the market portfolio should be stationary, even though the price series that contribute to it may be non-stationary. That is, it is expected that the contract prices in any given category exhibit a long-run equilibrium relationship (are cointegrated) and a very specific linear combination of those prices (the addition of all price series) is stationary. In fact, preliminary evidence in the form of an augmented Dickey-Fuller test on the market portfolio series bears this out. The test statistic of -4.137 (with two lags) in 2004 and -4.056 (six lags) and -3.948 (9 lags) in 2008 mean we can comfortably reject the null hypothesis of a unit root.

As robustness check, two additional tests were performed on the variables representing the market portfolio. The first was an informal test involving the production of correlograms at varying lag lengths for the market portfolio variable. If the correlograms showed that the autocorrelations died out quickly, then it could be concluded that the series was stationary. For the 2004 market, the correlogram showed very persistent autocorrelations, suggesting that the series for the market portfolio might not be stationary. For the 2008 market, the correlogram again showed very persistent autocorrelations, calling into doubt the previous conclusion that the series was stationary.

The second test conducted was the Augmented Dickey-Fuller GLS test, which has greater power to reject the null hypothesis of a unit root than other tests (Baum and Sperling, 2001; Cook, 2004), including the standard Dickey-Fuller test. For the 2004 market, this test was conducted for lags of the market portfolio variable of up to 18 (according to the Schwert rule of thumb). The resulting test statistic (τ_μ) of -0.705 for 2 lags was compared to the 5% critical value of -1.949 (with no trend included in the regression), which indicates that we cannot reject the null hypothesis that this series contains a unit root. This result was robust to testing of other lag lengths. For the 2008 market, this test was conducted for lags of the market portfolio variable of up to 23. The test statistic for six and nine lags was -3.901 and -3.612, respectively. Comparing those values to the 5% critical values of and -1.958 (for six lags) and -1.956 (for nine lags) suggests that we can reject the null hypothesis that the series contains a unit root. As in the case with the 2004 data, the conclusion drawn here is robust to tests at other lag lengths.

Based on the two types of Augmented Dickey-Fuller and Augmented Dickey-Fuller GLS test results, it seems reasonable to conclude that the 2008 market portfolio series is stationary, which suggests that there must be a long run equilibrium relationship between the non-stationary series that represent contract prices for individual candidates. Furthermore, the linear combination that produces this result can be identified as the simple sum of all contract values.

However, because evidence from correlograms and Augmented Dickey-Fuller GLS tests conflicted with the evidence from the standard Agumented Dickey-Fuller test, it is harder to draw a conclusion about the order of integration of the 2004 market portfolio series. In this case, it cannot be said with any confidence that the non-stationary variables that represent contract prices for the individual candidates sum up to create a stationary variable.

4 Conclusion

In keeping with similar research by Berg et al and others, this dissertation found that market prices tend to be better predictors of political event outcomes than corresponding polls. This superiority was observed over all time periods studied, with Intrade market prices correctly predicting the winner 36% of the time in the 2004 Democratic presidential nomination contest (compared to 13% for polls) and 54% of the time in the 2008 Democratic presidential nomination contest (compared to 45% for polls). Furthermore, the p-values generated from these results showed that these outcomes were unlikely to have occurred by chance, making it possible to reject that possibility at the 5% level for most time periods considered.

It was also noted that market prices seem to give off stronger signals than polls when the predictions they make turn out to be true, while the strength of the predictions is equal to or

Table 6: Augmented Dickey-Fuller tests for unit roots in levels

Variable	2004 Primary	2008 Primary
Clinton	-2.283(2)	-0.700(1)/-0.564(17)
Kerry	0.029(1)/-1.166(11)	-1.442(3)/-1.955(18)
Gore	–	-1.672(5)
Richardson	–	-2.102(4)/-1.775(15)
Dean	-1.101(3)/-1.383(16)	-2.534(3)/-2.342(7)
Edwards	-2.014(8)	-1.964(5)/-3.147(19)
Clark		-3.106(6)/-3.500(8)
Lieberman	-0.570(1)/-0.502(3)	-8.058(1)/-1.591(23)
Bayh	–	-2.324(15)/-2.031(18)
Obama	–	0.756(1)/0.514(12)
Leahy	–	-7.199(18)
Dodd	–	-5.892(1)/-2.234(19)
Vilsack	–	-2.772(5)/-1.396(16)
Biden	–	-2.597(12)/-2.254(18)
Ford	–	-10.457(9)/-12.140(11)
Rendell	–	-0.458(15)/-0.228(19)
Powell	–	-8.181(4)/-3.188(12)
Warner	–	-1.182(2)/-0.876(9)
Feingold	–	-2.563(4)/-2.466(8)
Corzine	–	-1.625(9)/-0.860(23)
Bredesen	–	-3.069(1)/-1.299(13)
Schweitzer	–	-7.682(3)/-6.523(18)
Easley	–	-1.739(23)
Daschle	–	-2.492(4)/-0.848(16)
Blagojevich	–	–
Gephardt	-1.296(1)/-1.370(2)	–
McCain	-0.963(5)/-0.214(18)	–
Rest of Field	-1.701(3)/-2.432(13)	-1.994(5)/-1.899(8)
Complete Portfolio	-4.137(2)	-4.056(6)/-3.948(9)

Notes: Lags are in brackets. Lag lengths were chosen by AIC and BIC criteria unless those criteria suggested a lag length greater than that recommended by Schwert's rule of thumb (18 for 2004 or 23 for 2008). The test statistics presented above must be assessed using nonstandard distributions, as they do not follow the standard t-distribution. The critical value for the Augmented Dickey-Fuller statistic at the 5% level of significance is -2.860. The inclusion of additional lags reduces degrees of freedom and so reduces the power of the test, although the additional lags do not change the critical values for the test statistic (Enders, 2010). It was impossible to test Blagojevich due to limited number of observations.

Table 7: Augmented Dickey-Fuller tests for unit roots in first differences (determining order of integration)

Variable	2004 Primary	2008 Primary
Clinton	-15.088(2)	-28.420(1)/-8.037(17)
Kerry	-16.228(1)/-3.925(11)	-21.313(3)/-8.604(18)
Gore	–	-18.327(5)
Richardson	–	-18.892 (4)/ -8.931(15)
Dean	-10.093(3)/-4.059(16)	-19.242(3)/-12.880(7)
Edwards	-7.969(8)	-17.632(5)/-8.229 (19)
Clark		-18.810(6)/-15.775(8)
Lieberman	-14.261(1)/-8.962(3)	-31.483(1)/-4.753(23)
Bayh	–	-10.950(15)/-15.760(18)
Obama	–	-26.954(1)/-8.823(12)
Leahy	–	-5.529 (18)
Dodd	–	-25.073(1)/-7.036(19)
Vilsack	–	-17.360(5)/-6.983(16)
Biden	–	-10.061(12)/-8.827(18)
Ford	–	-10.422(9)/-5.177(11)
Rendell	–	-5.846(15)/-3.624(19)
Powell	–	-24.937(4)/-10.501(12)
Warner	–	-20.151(2)/-15.921(9)
Feingold	–	-18.482(4)/-13.719(8)
Corzine	–	-10.731(9)/-4.935(23)
Bredesen	–	-26.361(1)/-8.091(13)
Schweitzer	–	-17.360(3)/-7.607(18)
Easley	–	-6.557(23)
Daschle	–	-17.851(4)/-8.226(16)
Blagojevich	–	–
Gephardt	-14.145(1)/-11.538(2)	–
McCain	-9.309(5)/-2.682(18)	–
Rest of Field	-10.340(3)/-5.223(13)	-9.246(5)/-5.844(8)
Complete Portfolio	-17.728(2)	-19.681(6)/-14.513(9)

Notes: These test statistics were produced by using the same number of lags used in the unit root test in levels. Experimentation with the number of lags did not materially alter any of the results.

only slightly different from the strength of the predictions made by polls when making incorrect predictions. This may be due to market prices containing greater information content than polling data.

When considering the greater overall predictive accuracy of market prices compared to polls, it was noted that, contrary to Berg et al, lower volatility of market prices was not a factor. However, this section also discussed the difficulty of making firm conclusions based on an analysis of prediction results across just two event outcomes.

Trading fees and margin requirements may have a significant impact on the observed pricing patterns in the market for Democratic presidential nominees on Intrade, driving prices away from values that would be expected of an efficient market. Similar to the results of Pennock et al, prices for the market portfolio in this dissertation were often observed to be overvalued, sometimes by as much as 25%. However, there was a marked tendency for market portfolio prices that were too high or too low to move toward the rational value.

A comparison of market liquidity over time within the 2004 and 2008 markets as well as between the two markets indicated that it would be difficult to ascribe declining volatility or more rational mean valuations of the market portfolio to greater measures of open interest.

Finally, unit root tests indicated that the majority of the individual contract prices follow a random walk process (which is expected if the market is operating efficiently), while a minority do not. An analysis of the order of integration of the market portfolios in 2004 and 2008 suggested that the 2004 market portfolio could not be conclusively said to be stationary, while the 2008 market portfolio was very likely to be stationary. Thus, if cointegration of the individual contracts that comprise the market portfolio is a prerequisite for efficiency, the market for the 2004 Democratic presidential nomination appears to fail that test while the market for the 2008 Democratic presidential nomination passes it.

Overall, the evidence on the efficiency of the 2004 and 2008 Intrade markets for the Democratic presidential nomination is mixed, although it is safe to conclude that both markets have some aspects of pricing that may not be as inefficient as they at first appear (due to the impact of fees and margin requirements) while other aspects are likely to be inefficient.

5 Appendix

Table 8: One-Sided P-Value to Determine if Market Outperformance of Polls is Better Than Chance (11 out of 18 Example)

Total Obs. (ψ)	18
DGP (ξ)	0.61
p-value (one-sided)	0.2345

(α_i)	(β_i)	(δ_i)	(ϵ_i)	(ϕ_i)	(ρ_i)	(χ_i)
0	1	6.40E+015	6.40E+015	1	0.000000044	0.000000044
1	1	355687428096000	6.40E+015	18	0.000000068	0.000001227
2	2	20922789888000	6.40E+015	153	0.000000107	0.000016307
3	6	1307674368000	6.40E+015	816	0.000000167	0.000136034
4	24	87178291200	6.40E+015	3060	0.000000261	0.000797894
5	120	6227020800	6.40E+015	8568	0.000000408	0.003494367
6	720	479001600	6.40E+015	18564	0.000000638	0.011842022
7	5040	39916800	6.40E+015	31824	0.000000998	0.031752234
8	40320	3628800	6.40E+015	43758	0.000001561	0.068287657
9	362880	362880	6.40E+015	48620	0.000002441	0.118676556
10	3628800	40320	6.40E+015	43758	0.000003818	0.167060074
11	39916800	5040	6.40E+015	31824	0.000005971	0.190035702
12	479001600	720	6.40E+015	18564	0.000009340	0.173387275
13	6227020800	120	6.40E+015	8568	0.000014609	0.125167146
14	87178291200	24	6.40E+015	3060	0.000022849	0.069919376
15	1307674368000	6	6.40E+015	816	0.000035739	0.029162953
16	20922789888000	2	6.40E+015	153	0.000055899	0.008552597
17	355687428096000	1	6.40E+015	18	0.000087432	0.001573781
18	6.40E+015	1	6.40E+015	1	0.000136753	0.000136753
Total						1.0000

Notes: The other tables in this series are far too large to be included in this document. Full tables and calculations available on request.

Table 9: Mean and Variance of Market Portfolio for Selected Periods Prior to Convention

.

2008

Days Prior To Market Close	**Mean**	**Variance**
0-100	103.48	0.42
101-200	104.45	5.39
201-300	102.17	2.71
301-400	101.11	3.6
401-500	101.01	3.81
501-600	102.07	3.69
601-700	102.97	3.74
701-800	101.87	4.37
801-900	103.09	3.64
901-1000	104.11	13.81
1001-1100	107.65	5.06
1101-1200	106.65	15.78
1201-1300	104.4	19.99
1301-1368	112.46	20.08

2004

Days Prior To Market Close	**Mean**	**Variance**
0-100	101.17	0.61
101-200	101.3	3.21
201-300	102.37	6.05
301-400	105.26	13.43
401-500	105.8	15.23
501-600	103.24	42.03
601-613	82.45	1.98

Table 10: Variable Names and Descriptions for Data Collected From Intrade

.

Variable Name	**Description**
contract_id	Unique identifier for each candidate's Intrade contract. Example: 6225
contract_symbol	Another unique identifier for each candidate's Intrade contract. Example: DEM.2004.GEPHARDT
timestamp	Market closing time and date. Example: 2002.11.08 12:00:00 GMT
auth_session_close_price	Last trading price for a given contract at market close. Example: 8
session_high	Highest trading value for a given contract for the day. This variable was too error prone to use. Example: 9
session_low	Lowest trading value for a given contract for the day. This variable was too error prone to use. Example: 2
contract_high	Highest trading value since the start of trading. This variable was too error prone to use. Example: 12
contract_low	Lowest trading value since the start of trading. This variable was too error prone to use. Example: 0
open_interest	Value of all contracts for a particular candidate held in US Dollars. Example: $4,801
execution_time	Time a given trade was executed on the exchange. Example: 2002.11.08 18:48:25 GMT
quantity	Number of lots (a bundle of ten contracts) traded on the exchange at any given execution time. Example: 20
price	Value (from 0 to 100) at which the contracts traded at execution time. Example: 43. One lot of contracts trading at 45 would cost $4.50 and would pay out either $0 or $10 upon expiry.

References

Alan Abramowitz et al. White House 2004: Democratic Nomination. http://www.pollingreport.com/wh04dem.htm, February 2004. PollingReport.com.

Alan Abramowitz et al. White House 2008: Democratic Nomination. http://www.pollingreport.com/wh08dem.htm, August 2008. PollingReport.com.

Sitaram Asur and Bernardo A. Huberman. Predicting the Future with Social Media. *CoRR*, abs/1003.5699, 2010.

Christopher Baum and Richard Sperling. DFGLS: Stata Module to Compute Dickey-Fuller/GLS Unit Root Test. Statistical Software Components, Boston College Department of Economics, 2001. URL http://econpapers.repec.org/RePEc:boc:bocode:s410001.

Joyce Berg, Robert Forsythe, Forrest Nelson, and Thomas Rietz. Results from a Dozen Years of Election Futures Markets Research. 1:742–751, February 2008a. URL http://ideas.repec.org/h/eee/expchp/5-80.html.

Joyce E. Berg, Forrest D. Nelson, and Thomas A. Rietz. Prediction Market Accuracy in the Long Run. *International Journal of Forecasting*, 24(2):285–300, 2008b. URL http://ideas.repec.org/a/eee/intfor/v24y2008i2p285-300.html.

Steven Cook. Finite-Sample Properties of the GLS-Based Dickey-Fuller Test in the Presence of Breaks in Innovation Variance. *Austrian Journal of Statistics*, 33:305314, 2004.

Bo Cowgill. Putting crowd wisdom to work. http://googleblog.blogspot.com/2005/09/putting-crowd-wisdom-to-work.html, September 2005. http://googleblog.blogspot.com/.

John (CEO) Delaney. Intrade Exchange Rules. http://www.intrade.com/jsp/intrade/help/index.jsp?page=general.html, September 2010. Intrade.com.

Walter Enders. *Applied Econometric Time Series, 3nd Edition.* Wiley, 2010.

Eugene F. Fama. Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2):383–417, 1970. ISSN 00221082. URL http://www.jstor.org/stable/2325486.

Joe Flint. Box-Office Futures Are Dealt A Blow. http://www.pollingreport.com/wh04dem.htm, April 2010. latimes.com.

Patri Friedman. InTrade Fee Structure Discourages Selling the Tails? http://www.overcomingbias.com/2007/11/intrade-fee-str.html, November 2007. overcomingbias.com.

Robin Hanson. Designing Real Terrorism Futures. *Public Choice*, 128(1):257–274, July 2006. URL http://ideas.repec.org/a/kap/pubcho/v128y2006i1p257-274.html.

Robin Hanson. Shall We Vote on Values, But Bet on Beliefs? Working Paper, 2007. George Mason University: Department of Economics.

Robin Hanson and Ryan Oprea. A Manipulator Can Aid Prediction Market Accuracy. *Economica*, 76(302):304–314, 04 2009. URL http://ideas.repec.org/a/bla/econom/v76y2009i302p304-314.html.

Robin Hanson, Ryan Oprea, and David Porter. Information Aggregation and Manipulation in An Experimental Market. *Journal of Economic Behavior & Organization*, 60(4):449–459, August 2006. URL http://ideas.repec.org/a/eee/jeborg/v60y2006i4p449-459.html.

Alexander K. Koch and Hui-Fai Shing. Bookmaker and Pari-Mutuel Betting: Is a (Reverse) Favourite-Longshot Bias Built-In? *Journal of Prediction Markets*, 2(2):29–50, September 2008. URL http://ideas.repec.org/a/buc/jpredm/v2y2008i2p29-50.html.

Charles F. Manski. Interpreting the Predictions of Prediction Markets. *Economics Letters*, 91(3):425–429, June 2006. URL http://ideas.repec.org/a/eee/ecolet/v91y2006i3p425-429.html.

Melonyce McAfee. Are “Political Futures” Illegal? http://www.slate.com/id/2164916/, April 2007. slate.com.

Adam Nagourney. Obama Wins Nomination; Biden and Bill Clinton Rally Party. http://www.nytimes.com/2008/08/28/us/politics/28DEMSDAY.html?ref=democratic_national_convention, August 2008. New York Times.

Ryan Oprea, David Porter, Chris Hibbert, Robin Hanson, and Dorina Tila. Can Manipulators Mislead Prediction Market Observers? (08-01), November 2006. URL http://ideas.repec.org/p/chu/wpaper/08-01.html.

David M. Pennock, Steve Lawrence, C. Lee Giles, and Finn Arup Nielsen. The power of play: Efficiency and Forecast Accuracy in Web Market Games. 2000.

Kate Phillips. House Backs Crackdown on Gambling on Internet . http://www.nytimes.com/2006/07/12/washington/12gamble.html?ei=5088&en=fb44926ad3840c4e&ex=1310356800..., July 2006. nytimes.com.

G William Schwert. Tests for Unit Roots: A Monte Carlo Investigation. *Journal of Business & Economic Statistics*, 7(2):147–59, April 1989. URL http://ideas.repec.org/a/bes/jnlbes/v7y1989i2p147-59.html.

Emile Servan-Schreiber, Justin Wolfers, David M. Pennock, and Brian Galebach. Prediction Markets: Does Money Matter? *Electronic Markets*, 14(3), 2004.

Eric Snowberg and Justin Wolfers. Explaining the Favorite-Longshot Bias: Is it Risk-Love or Misperceptions? (3029), 2010. URL http://ideas.repec.org/p/ces/ceswps/_3029.html.

Paul Tetlock. Liquidity and Prediction Market Efficiency. *Working Paper Series*, October 2008.

Robin Toner and Katherine Q. Seelye. Edwards Gives Strong Tributeas Democrats Nominate Kerry. http://www.nytimes.com/2004/07/29/politics/campaign/29campaign.final.html, July 2004. New York Times.

Andrew P. Weinbach and Rodney J. Paul. The Link Between Information and the Favorite-longshot Bias in Pari-mutuel Wagering Markets. *Journal of Gambling Business and Economics*, 2(1):30–44, May 2008. URL http://ideas.repec.org/a/buc/jgbeco/v2y2008i1p30-44.html.

Justin Wolfers. Guessing Games. http://bpp.wharton.upenn.edu/jwolfers/Press/Mentions/GuessingGames%28The%20Economist%29.pdf, November 2004. Economist.com.

Justin Wolfers and Eric Zitzewitz. Prediction Markets. *Journal of Economic Perspectives*, 18(2):107–126, Spring 2004. URL http://ideas.repec.org/a/aea/jecper/v18y2004i2p107-126.html.

Justin Wolfers and Eric Zitzewitz. Five Open Questions About Prediction Markets. *NBER Working Paper Series*, (12060), February 2006. URL http://www.nber.org/papers/w12060.