

INFORMATION FLOW BETWEEN PREDICTION MARKETS, POLLS AND MEDIA: EVIDENCE FROM THE 2008 PRESIDENTIAL PRIMARIES

URMEE KHAN AND ROBERT P. LIELI

ABSTRACT. We analyze the direction and timing of information flow between prices, polls, and media coverage of events traded on prediction markets. Using data for the 2008 Democratic Party nomination race between Barack Obama and Hillary Clinton, we construct time series variables that reflect ‘pollster’s surprise’ in each primary election, as well as an index that describes the extent of media coverage received by the candidates. Trivariate Granger causality tests are then conducted between the day-to-day percentage change in the price of the ‘Obama wins nomination’ security and the two aforementioned variables. Results suggest that prediction market prices anticipate at least some of the discrepancy between the actual outcome and the latest round of polls before the election. Nevertheless, prices also seem to be *driven* partly by election surprises, suggesting that there is an element of the pollster’s surprise that is genuine news for the market. Furthermore, in line with economic theory, prices seem to capture information earlier than its revelation in the news media.

1. INTRODUCTION

Opinion polls have a long history as a tool to gauge the state of public opinion, particularly in the context of measuring the likelihood of political outcomes. Accordingly, a rich body of literature is concerned with assessing the effectiveness of opinion polls as information aggregators and predictors of political outcomes. In the last couple of decades, with the advent and spread of digital technology, alternative tools have been developed to serve similar purposes. Asset markets designed for and exclusively dedicated to gathering information about probable outcomes of future events have come to existence. These markets, popularly

Key words and phrases. Prediction markets, opinion polls, information aggregation, Granger causality.

The paper is based on a PhD dissertation chapter of the first author. Many thanks to Péter Kondor and Max Stinchcombe for helpful comments. All errors are our responsibility.

known as prediction or information markets, have attracted attention for making accurate predictions about election outcomes [Berg et al., 2003, 2008,?], product sales [Chen and Plott, 2002], film box office and myriad other variables of interest. With some operational variations, the fundamental design of such markets is as follows. The future event of interest is formalized as a random variable whose outcome would depend on an unobserved underlying state about which information is assumed to be dispersed among potential participants in the market. A set of complete contingent assets are introduced, whose eventual payoffs are tied to the future realization of the outcome of interest. The participants are expected to trade in the contingent assets based on their private information about the underlying state. Subject to assumptions about arbitrage possibilities and risk attitudes of the traders, the resultant market prices could be considered as a probabilistic prediction of the outcome of interest, at least in theory.

The performance of such markets in predicting the respective events of interest have created an optimism about its potential as a forecasting technology (Arrow et al. [2008], Sunstein [2007], Berg and Rietz [2003]). There is a nascent industry around this technology, and a growing stream of literature trying to understand various aspects of the information aggregation process in prediction markets [Tziralis and Tatsiopoulos, 2007]. Such comprehensive understanding would be vital for future improvement in this technology. Nevertheless, there remains a number of open questions, both theoretical [Wolfers and Zitzewitz, 2004] and empirical. The set of open questions include, but are not limited to, the following: i) What kind of information is built into the price? Do prices contain information beyond that collected by traditional mechanisms such as polls and the news media? ii) How fast does new (public) information get built into the price? iii) How does the market perform in terms of forecasting relative to other mechanisms (like polls)?

In comparing polls and prediction markets, the existing empirical literature has focused mainly on the last question. In a series of papers, Berg et al. have argued, using Iowa Electronic Market (IEM) data, that prediction market prices have significantly lower forecast errors both in the short and the long run compared to contemporaneous polls [Berg

et al., 2003, 2008,?]. In response, Erikson and Wlezien [2008] point out that since polls and prediction markets ask inherently different questions, they are not directly comparable. The standard question asked in a pre-election poll is this: “if the election were held today, who would you vote for?” Therefore, polls only capture *voters’ preferences at the moment* and should not be interpreted as forecasts in themselves; rather, sophisticated forecasts should build on the information provided by the polls. Presumably, markets do just that, as traders take into account their private and public sources of information before placing their bid, providing, in effect, an answer to the question: “who do you think will win”? Nevertheless, other sophisticated forecasts may exist. Using data from the same time period as Berg et al. [2008], Erikson and Wlezien [2008] construct projections of vote shares and win probabilities based on daily polls and show that those projections can outperform prediction markets. In response, Rothschild [2009] employing a different ‘debiasing’ technique and adding state-level data, shows the contrary. Overall, it seems that the empirical evidence on the question of forecasting performance of polls relative to prediction markets is still inconclusive, and the question warrants further investigation from different angles and using different kinds of data. Our paper is an attempt to provide a separate, and complementary, approach to those taken by the previous authors to empirically investigate the way in which polls and prediction markets aggregate information and interact with each other.

More specifically, we investigate the information flow between prediction market prices and two other conventional aggregators of information, namely media and opinion polls. In contrast to the aforementioned papers (and most others), we are not directly concerned with forecasting performance; rather, we seek to establish stylized facts that go towards answering questions (i) and (ii) above. To this end, we examine the sequence of primary elections in the 2008 race for the Democratic nomination for president, and the evolution of prices in the market for the “Obama wins nomination” security. We construct time series variables that reflect the “pollster’s surprise” in each primary election, derived from the difference between Obama’s actual vote share and vote share predicted by the latest poll(s) before the primary, as well as indices that describe the extent of daily media coverage devoted

to him. We conduct Granger causality tests between the daily percentage change in the price of the “Obama wins nomination” security and these information variables. These tests provide answers to the following operationalized versions of question (i): Do prediction market prices have information that is not reflected in contemporaneous polls and media stories? Conversely, do prices react to information that appears to be news for pollsters or increases in media exposure? Further, the time horizons over which Granger causality relationships can be established provide answers to question (ii).

Based on the seminal work by Granger ([Granger \[1969\]](#), [Granger \[1980\]](#)), Granger causality has become a widely used concept aimed at studying whether a time series is informative about another across various horizons. In intuitive terms, the variable X Granger causes Y , if the history of X helps predict future values of Y even after taking the own history of Y , and possibly other information, into account. Indeed, Granger causality can be sensitive to additional ‘control’ variables included in the analysis. While our primary focus is on the relationship between polls and prediction markets, we include media coverage as an auxiliary variable for several reasons. First, media is a potential mediator between markets and polls; news obtained by the market might become known to the wider public only through media coverage. Conversely, there is a perennial question about how campaign coverage influences public opinion, or whether it does at all (see [Erikson and Wlezien \[2012\]](#) for a comprehensive study of electoral campaigns and voter intentions). As the current (2016) election cycle has demonstrated once again, qualitative political analysis presented in the media often contradicts, and rivals, more quantitative sources like polls or prediction markets. Despite massive and sustained leads in polls, a presidential contender’s win might be assessed as unlikely by the media, based on perceived ‘qualitative’ factors. Therefore, media coverage might help predict pollster’s surprise or even price changes if it indeed contains relevant information not fully internalized by polls or markets. To explore these additional questions empirically, we bring in a unique dataset, created during the 2008 election cycle by the Pew Research Center, that quantifies the extent of media coverage of the candidates on a

daily basis through a number of variables, which we collapse into a principal component (see Section and Appendix B for details).

To conduct the empirical analysis we use the Granger causality test procedure developed by Hill (2007), specifically designed for trivariate setups. The method, described in more detail in Section 3, tests for direct as well as indirect causality (conveyed by a third variable) over multiple horizons, in a logically consistent way. Our main finding regarding the relationship between price changes and the pollster's surprise is strong two-way Granger causality. Thus, on the one hand, part of the pollster's surprise is predictable by previous price movements, so market prices appear to contain information not contained in the polls. On the other hand, the pollster's surprise also contains information about future price movements even when the history of prices is taken into account. Therefore, there is an element of the pollster's surprise that is genuine news for the market as well. Our method also allows us to analyze causality over different time horizons, and results suggest that prices incorporate the new information coming out of election booths very quickly, within a day, while prices anticipate part of pollster's surprise up to a week ahead.

Regarding the relationship between prediction markets and media coverage, there is strong evidence of one-way Granger causality in the short run from price changes towards media index. Evidence for causality over longer horizon is less robust. These test results are consistent with prices capturing election-relevant information somewhat earlier than its revelation in the news media. There is little evidence of causality in the other direction. Media coverage, particularly in the election season, contains both useful information and hype (i.e., noise). Therefore, even if prediction markets are occasionally informed by media, the quantitative level of media coverage need not be a particularly strong empirical predictor of prices.

Regarding the relationship between media coverage and pollster's surprise, there is very strong evidence of direct causality from the latter to the former, all the way up to a week-long horizon, as one would expect surprise election results to generate high volume of coverage. On the other hand, there is no evidence of causality going the other way. Taken together, the results do not preclude the possibility that media influences public opinion, but they do

suggest that media coverage does not contain *additional* information that could be used to improve forecasts of election outcomes relative to polls and markets.

In summary, existing research on polls versus prediction markets focus on relative forecasting accuracy, and different approaches to ‘read’ and ‘debias’ polls drive differences in empirical findings. We take a different operational approach, and investigate information flow over time between markets, polls and media. Using Granger causality, and a measure for ‘pollster’s surprise’ that stands in for information not contained in the polls, we find that there is bi-directional information flow between the surprise index and prices. This result suggests that markets are actually more informative about future election outcomes than raw poll figures (as they anticipate part of the pollster’s surprise), but they far from being perfectly informed about the state of public opinion (as election surprises relative to polls induce short run price movements).

Admittedly, this is a small dataset pertaining to only one electoral primary race, so questions of external validity remain. But the particular race chosen made it possible to use concurrent data on prices and polls for a sufficiently long period to construct meaningful time series variables, given that the race was long and stayed contested till the very end. Also, the media index dataset we were able to use is quite unique, since that particular indexing exercise was only conducted by Pew Research Center for a limited time period overlapping with the nomination race. Our aim was to investigate and present the stylized facts, and our findings, while suggestive rather than conclusive, are insightful.

2. DATA SOURCES AND VARIABLE DEFINITIONS

For prediction market prices, we use daily data from the IEM winner-take-all market for predicting the winner of the Obama vs. Clinton nomination race in 2008 for the period January 1, 2008 to June 15, 2008 (167 days overall).¹ The raw data is a daily time series of

¹In the IEM winner-take-all market, one share of a candidate pays off 1 dollar if the candidate wins and nothing if the candidate loses. A portfolio of one unit of each candidate pays exactly 1 dollar. A trader who buys one unit of a candidate at, say 30 cents, wins either 1 dollar (a 70 cent profit) or nothing (a 30 cent loss)

prices (recorded every day at midnight, the single price being the last transaction price before midnight). To facilitate the econometric analysis, the price series is rendered stationary by taking its log difference, i.e., transforming it into day-to-percentage changes. For simplicity, we will continue to refer to the transformed data as the ‘price’.

The data on state polls regarding primary outcomes come from the Real Clear Politics website². The polls in the respective states are conducted on likely Democratic voters, essentially asking the question who they would support in the upcoming primary (since polls were conducted by different organizations, the exact version of the question might have varied). For each primary election, we consider the difference between Obama’s actual vote share and his predicted vote share as measured by the latest poll data available before the election in that state/territory (typically one or two days before). This raw difference is then multiplied by the fraction of total delegates at stake in the election (hence, a one percentage point difference in actual versus predicted vote share in a ‘big’ state counts as a larger surprise than the same observed difference in a ‘small’ state). To compute total surprise for a given day, we add the weighted surprise measures for all elections held that day. If there are no elections on the day in question, we set the surprise measure to zero. There are some smaller states and territories where elections are held but for which no poll data is available. These are ignored altogether in constructing the surprise measures (i.e., their contribution to total surprise is set to zero). Overall, the elections covered by the available poll data account for 78.4% of all delegates at stake. The poll data and the construction of the surprise index is described in more detail in Appendix A.

For constructing the media variables, we use the data library of the Pew Research Centre Project for Excellence in Journalism (PEJ)³. They maintain a continually updated database of news stories in various kinds of media outlets which are monitored at a daily basis in regular intervals during the day. Stories are classified according to topic, lead newsmaker if the contract is held until market closing following the election. If the trader buys at 30 cents and sells at, say 70, the profit is 40 cents. For further details, consult the IEM website, <http://www.biz.uiowa.edu/iem/>

²<http://www.realclearpolitics.com>

³<http://www.journalism.org>

and placement prominence. Word counts for print and web stories and duration-in-seconds for broadcast/cable stories are also counted. For details about the sampling process, see Appendix B.

For our study, we selected the stories that were categorized under the topic ‘campaign’ in the relevant date range (January 1, 2008 to June 15, 2008), had Obama as the lead newsmaker, were news items rather than opinion pieces, and had relatively high prominence in their placements. For stories matching all the criteria above, we constructed time series variables representing total number of stories and total word counts per day for print media, total number of web stories, total number of TV and radio stories, and total duration-in-seconds for TV and Radio stories. Finally, we combined as well as compressed the information contained in the various measures of media exposure by extracting their first principal component. The variable so obtained will simply be referred to as ‘media’ or ‘media exposure’ in the rest of the paper.

The final time series used in the Granger causality analysis are depicted in Figure 1.

3. METHODOLOGY

We employ the framework proposed by Hill [2007] for testing causality in trivariate processes. This hierarchical procedure is designed to test for direct as well as indirect causal routes over multiple horizons in a logically consistent way. We also report results using the simpler direct test based on Dufour and Renault [1998]. For convenience, we restate some basic definitions from the literature and provide a brief description of these methodologies. Let $W_t = (X_t, Y_t, Z_t)'$ be a trivariate covariance stationary process. Let $I_W(t)$ denote the information set consisting of the history of W_t up to time t , i.e., $I_W(t) = \{1, X_t, Y_t, Z_t, X_{t-1}, Y_{t-1}, Z_{t-1}, \dots\}$. Similarly, $I_{XZ}(t)$ denotes the smaller information set that consists of the history of X and Z only. In this trivariate setting we say that

- Y does not (linearly) cause X at horizon $h > 0$, denoted $Y \not\overset{h}{\rightarrow} X$, if the linear projection of X_{t+h} on $I_W(t)$ coincides with the linear projection of X_{t+h} on $I_{XZ}(t)$;

- Y does not (linearly) cause X up to horizon $h > 0$, denoted $Y \xrightarrow{(h)} X$ if $Y \xrightarrow{j} X$ for $j = 1, \dots, h$;
- Y does not (linearly) cause X at any horizon, denoted $Y \xrightarrow{(\infty)} X$ if $Y \xrightarrow{h} X$ for any $h > 0$.

The Hill (2007) test procedure consists of the following steps.

Step 1: Testing $Y \xrightarrow{(\infty)} X$. Either of the conditions $Y \xrightarrow{1} (X, Z)$ or $(Y, Z) \xrightarrow{1} X$ is sufficient for the complete lack of causality from Y to X . Formally, one tests $H_0^{(1.1)}$: whether $(Y_{t-j}, j \geq 0)$ has zero coefficients in the linear projection of X_{t+1} on $I_W(t)$ as well as the linear projection of Z_{t+1} on $I_W(t)$; and $H_0^{(1.2)}$: whether $(Y_{t-j}, Z_{t-j}, j \geq 0)$ has zero coefficients in the linear projection of X_{t+1} on $I_W(t)$.

Step 2: Testing $Y \xrightarrow{1} X$, $Y \xrightarrow{1} Z$ and $Z \xrightarrow{1} X$. This step amounts to testing for one-step-head direct causality and the existence of an indirect causal chain through the auxiliary variable Z . Formally, one tests $H_0^{(2.1)}$: whether $(Y_{t-j}, j \geq 0)$ has zero coefficients in the linear projection of X_{t+1} on $I_W(t)$; $H_0^{(2.2)}$: whether $(Y_{t-j}, j \geq 0)$ has zero coefficients in the linear projection of Z_{t+1} on $I_W(t)$ and $H_0^{(2.3)}$: whether $(Z_{t-j}, j \geq 0)$ has zero coefficients in the linear projection of X_{t+1} on $I_W(t)$.

Step 3: Testing $Y \xrightarrow{(h)} X$, $h \geq 2$. In this step one tests $H_0^{(3,h)}$: whether $(Y_{t-j}, j \geq 0, Z_{t-j}, j = 0, 1, \dots, h-2)$ has zero coefficients in the linear projection of X_{t+1} on $I_W(t)$.

In addition to the joint test in Step 3, we also conduct and report horizon-specific individual tests of causality based on Dufour and Renault (1998).

Step 3': Testing $Y \xrightarrow{h} X$, $h \geq 2$. This is equivalent to testing $H_0^{(3,h)'}$: whether $(Y_{t-j}, j \geq 0)$ has zero coefficients in the linear projection of X_{t+h} on $I_W(t)$.

Remarks. Step 1 is basically an ‘entry test’ for any further causality investigation; in case of non-rejection of either hypothesis, the procedure could in theory stop. In Step 2, the test for an indirect causality chain is the most interesting when there is no evidence of direct

causality over a one period horizon. This is because for non-causation to be followed by causation over a longer horizon, a causality chain must exist; see Hill (2007, Theorem 2.1). Thus, if all hypotheses in Step 2 are accepted, the procedure could again stop with the conclusion of no causation over any horizon. The tests in Step 3 will generally reject if there is causation already at horizon $h = 1$; thus, these tests are interesting mostly when Step 2 reveals no direct causality but supports a causality chain. If there is causality already at horizon $h = 1$, it is Step 3' that provides more information about other horizons.

Testing for linear causality is somewhat restrictive; if there is no linear causality, higher order causality might still be present. Nevertheless, with the limited amount of data available the implementation of more demanding nonlinear or nonparametric causality tests is rather pointless.

Implementation issues. The linear projection of W_{t+h} on $I_W(t)$ involves an infinite number of lags; in practice it is of course necessary to approximate this projection with a finite order VAR(p) process. Given our relatively short sample, we consider $p \leq 18$. Various information criteria suggest rather different choices from this set; specifically, $p = 1$ (BIC), $p = 8$ (H-Q), and $p = 18$ (AIC). However, we discard small values of p for a number of reasons. First, with daily observations potential day-of-the-week effects are better captured with 7 lags or more. Second, the pollster's surprise variable is obviously a highly nonlinear process, which means that small values of p are unlikely to provide a good approximation to its linear projection on $I_W(t)$. Third, to implement Step 3 of the Hill procedure for a given value of h , one needs to estimate at least a VAR of order $h - 1$. E.g., testing causality up to a week-long horizon requires a VAR(6). In the end, we estimate several models with lags $p = 7, 8, 10, 14$ and 18 and report results for all specifications. Thus, one can assess the sensitivity of the empirical results to the choice of p directly.

Regarding technical details of the estimation, we write the VAR as a SUR-system (see, e.g., Wooldridge 2002, Ch. 7.3) and use system OLS to jointly estimate the three VAR equations. The advantage of this formulation is that it is straightforward to construct the heteroskedasticity-robust version of the Wald statistic to test the zero restrictions stated in

Steps 1 through 3' (we use the chi-squared limit distribution with the appropriate degrees of freedom to compute p-values). Denoting the 3×1 VAR error term by ϵ_t , we thus allow the conditional variance of ϵ_t given W_{t-1}, \dots, W_{t-p} to be an arbitrary function of the conditioning variables. This is an important feature of the test given the inherently nonlinear nature of some of the variables.

A final concern is that using conventional nominal significance levels for individual tests might lead to overrejection due to specification search and the large number of tests performed. We deal with this issue rather informally; when inspecting the evidence in the next section, we require smaller than usual p-values and results reasonably robust to the choice of p to declare the existence of causality.

4. RESULTS

We investigate, in turn, the nature of Granger causality between price and surprise (Table 1, Table 2), surprise and media (Table 3 and Table 4) and price and media (Table 5 and Table 6) in the presence of the third variable. More specifically, we do this by assigning all six permutations of the variables price, pollster's surprise and media to the vector (X, Y, Z) , and then test causality going from Y to X using the procedure described in Section 3.

4.1. Price and Pollster's Surprise. There is strong evidence of causality from pollster's surprise (S) to price (P) in the presence of media (M); see Table 1). The null hypothesis of no causality at any horizon is rejected outright (Step 1), and Step 2 indicates strong direct causality at the one day horizon. There is also some evidence of an indirect causality chain conveyed by media, though this would be hard to rationalize theoretically. The evidence here is however not conclusive—the media to price link is only borderline significant at conventional levels and is sensitive to the lag length in the VAR. In accordance with Step 2, the null hypothesis of no causality at any horizon up to h is rejected for larger values of h as well (see Step 3). Nevertheless, Step 3' shows that this rejection is due solely to the causality found at the one day horizon; there is virtually no evidence of poll surprises

Granger causing prices at longer horizons. This striking result also points toward the media to price link being spurious, as the chain $S \xrightarrow{1} M \xrightarrow{1} P$ would generally imply $S \xrightarrow{h} P$ for $h > 1$ (see Hill 2007, Theorem 2.1, Remark 2). In sum, the evidence shows that there is some new information in the pollster’s surprise that is not already known to the market, but any new information gets incorporated very quickly into the price process, over the horizon of a single day.

Conversely, Table 2 shows that the null hypothesis of no causality from prices to pollster’s surprise is also rejected, though the outcome of one of the Step 1 ‘entry’ tests is somewhat sensitive to the VAR specification. Going into more detail, Step 2 shows direct causality at the one day horizon, but the indirect causal chain going through media is clearly broken at the $M \xrightarrow{1} S$ link. Causality from prices to media is expected if both variables react to the same events but markets process information faster. On the other hand, media anticipating pollster’s surprise would mean that media exposure is a better predictor of election outcomes than polls are. The data show that this is decidedly not the case. In accordance with one-step-ahead direct causality, most tests in Step 3 also reject. More interestingly, Step 3’ shows that for the three smaller VARs prices anticipate pollster’s surprise several days ahead. That is, part of the information about pollster’s surprise that the market captures is built into prices well in advance. Nevertheless, the larger models support causality only at the one day horizon.

4.2. Pollster’s Surprise and Media. There is very strong evidence of direct causality from pollster’s surprise to media starting from the one day horizon all the way up to the 8 day horizon (see Table 3). This is of course expected—a surprising election outcome should generate a lot of media coverage for several days after the fact, and as we saw, part of the pollster’s surprise does represent genuine news not captured even by the market. On the other hand, there is very little evidence of causality at any horizon in the other direction, i.e., from media exposure to pollster’s surprise (Table 4). Of particular note, we have already argued in Section 4.1 that the hint of an indirect causality chain from media to prices to surprise is likely spurious. Indeed, the existence of Granger causality from media to pollster’s

surprise would imply that media coverage contains information about election outcomes not reflected by prior polls or price movements. The lack of Granger causality does not rule out media coverage actually influencing polls; in fact, it says precisely that the full information content of media coverage gets built into polls one way or another. For example, media reports could directly influence polls over time or media could cover the same events that influence polls directly, only with a lag.

4.3. Prices and Media. Looking at the direction price to media (see Table 5), we find strong evidence of causality at horizon $h = 1$, and less robust evidence of causality over longer horizons, depending on the VAR specification. Part of the causality works through the price→pollster’s surprise→media chain, but the direct channel is also active, meaning that prices anticipate events outside election surprises that are likely to receive media coverage later on. The opposite direction, causality from media towards prices, has already been discussed in Section 4.1; nevertheless, for completeness Table 6 shows the full set of test results. Some of the Step 1 ‘entry tests’ already cast doubt on causality in this direction. The evidence provided by the Step 2 one-day-ahead test has already been discounted, and the longer horizon direct tests are almost all insignificant (Step 3’). The Step 3 tests reject only because they include a test of the surprise to price link, which is highly significant in itself. In sum, there appears to be little news in the media for prediction market traders, and if there is any at all, it gets incorporated into the price within a day.

5. CONCLUSION

Using data on the race between Barack Obama and Hillary Clinton in the 2008 Democratic primaries for presidential nomination, we investigate whether prediction market prices have information that is not reflected in contemporaneous polls and media stories and conversely, whether prices react to information that appears to be news for pollsters or is prominently featured by the media. We test for trivariate Granger causality between day-to-day percent change in prediction market prices, a constructed measure of the surprise element in primary

results (information that is not reflected in the polls), and indices constructed to capture the extent of media coverage received by a candidate.

The main qualitative finding of our exercise, based on the direction of Granger causality found in the data, is that prediction market prices seem to capture some, but not all, of the surprise element in the primary results. Also, there is some evidence that, at any given time, media reflects information that is already anticipated by polls and prediction markets, rather than the other way round.

DATA APPENDIX

A. Pollster’s surprise. Tables 7 and 8 summarize the information used in constructing the pollster’s surprise variable. For each democratic primary between January and June 2008 we record the date of the election, the number of delegates at stake, the proportion of total delegates at stake, Obama’s predicted and actual vote share, and the difference between the two. Predicted vote share is computed as the weighted average of polls that closed on the day before the election, the weights being given by the relative sample sizes when applicable. We could not find public poll results for a considerable number of primaries; nevertheless, the elections covered by the data account for 78.4% of all delegates at stake.

The final pollster’s surprise variable is constructed as follows. On days with no elections (or elections with no poll data), the variable is set to zero. On days with elections we multiply the raw surprise in each election with the fraction of delegates at stake and then sum over elections. (Raw surprise is the difference between Obama’s actual and predicted vote share, as described above.)

B. Data on media coverage. The following is a brief description of the sampling and coding methodology followed by PEJ in constructing their News Coverage Index data library. For further details, see the PEJ website <http://www.journalism.org>.

The main categories of news sources for the data are as follows: Network TV News, Newspapers, Online News Sites, Cable News and Radio News.

The major broadcast channels ABC, CBS, and NBC make up the broadcast segment. Stories are monitored through different time-slots during the day for 2 out of 3 channels on a rotation basis as follows:

- Commercial Evening News: Entire 30 minutes of 2 out of 3 programs each day (60 minutes)
- Commercial Morning News: 1st 30 minutes of 2 out of 3 programs each day (60 minutes)

- PBS NewsHour: Rotate to code the 1st 30 minutes one day, the 2nd 30 minutes the next day and then skip

This results in either 2 or 2.5 hours of programming each day. Similar method is used on a rotation basis for cable channels CNN, MSNBC and Fox News. During daytime, on a rotation basis, two out of three 30-minute daytime slots each day (60 minutes a day) are coded. During prime time, the following are included

- Two 30-minute segments for Fox News (60 minutes)
- Two 30-minute segments for CNN (60 minutes)
- Two 30-minute segments for MSNBC (60 minutes)

Newspapers are categorized into 3 tiers according to subscription levels, national prominence and regional location. Representative newspapers are chosen from each tier. Here is the list for newspapers in the sample:

- 1st Tier: The New York Times, LA Times, USA Today, Wall Street Journal
- 2nd Tier: Washington Post, Tampa Tribune, Seattle Times, Columbus Dispatch
- 3rd Tier: The Day, Rome News Tribune, Ventura News

For each of the papers selected, only articles that begin on page A1 (including jumps) are picked. This results in a newspaper sample of approximately 20 stories a day.

The websites included in the PEJ sample for our selected date range are as follows: Yahoo News, MSNBC.com, CNN.com, NYTimes.com, Google News, AOL News, Foxnews.com, USAToday.com, Washingtonpost.com, ABCNews.com, HuffingtonPost.com, and Wall Street Journal Online.

For the online news sites, the database captures each site once a day and code the top 5 stories that appear on the site at the time of capture. The time of the day that the sample captures the Web sites is rotated between 9-10 am Eastern time and 4-5 pm Eastern time. The 4-5 PM time-slot was added after April 28, 2008.

The sample of radio stories are collected as follows:

- News: 30 minutes of NPR each day, rotating between Morning Edition and All Things Considered, as broadcast on a selected member station.
- Talk: The first 30 minutes of either one or two talk programs each day. Every weekday, a total of 3 conservatives and 2 liberals were coded during the period of our sample.
- Headlines: Two headline segments each day (one from ABC Radio and one from CBS Radio), about 10 minutes total.

This results in a sample of roughly 1 or 2 hours of programming a day.

REFERENCES

- Arrow, K. J., R. Forsythe, M. Gorham, R. Hahn, R. Hanson, J. O. Ledyard, S. Levmore, R. Litan, P. Milgrom, and F. D. Nelson (2008). Economics: The promise of prediction markets. *Science* 320(5878), 877.
- Berg, J., R. Forsythe, F. Nelson, and T. Rietz (2008). Results from a dozen years of election futures markets research. *Handbook of Experimental Economics Results* 1, 742–751.
- Berg, J., F. Nelson, and T. Rietz (2003). Accuracy and forecast standard error of prediction markets. *Tippie College of Business Administration, University of Iowa*.
- Berg, J. E., F. D. Nelson, and T. A. Rietz (2008). Prediction market accuracy in the long run. *International Journal of Forecasting* 24(2), 285–300.
- Berg, J. E. and T. A. Rietz (2003). Prediction markets as decision support systems. *Information Systems Frontiers* 5(1), 79–93.
- Chen, K. Y. and C. R. Plott (2002). Information aggregation mechanisms: concept, design and field implementation. *California Institute of Technology Social Science Working Paper* 1131.
- Dufour, J.-M. and E. Renault (1998). Short run and long run causality in time series: Theory. *Econometrica* 66(5), 1099–1125.
- Erikson, R. S. and C. Wlezien (2008). Are political markets really superior to polls as election predictors? *Public Opinion Quarterly*.
- Erikson, R. S. and C. Wlezien (2012). *The timeline of presidential elections: How campaigns do (and do not) matter*. University of Chicago Press.
- Granger, C. W. J. (1969). Investigating causal relations by econometric models and Cross-Spectral methods. *Econometrica* 37, 424–438.
- Granger, C. W. J. (1980). Testing for causality:: A personal viewpoint. *J. Econ. Dyn. Control* 2, 329–352.

- Hill, J. B. (2007). Efficient tests of long-run causation in trivariate VAR processes with a rolling window study of the money–income relationship. *Journal of Appl. Econometrics* 22(4), 747–765.
- Rothschild, D. (2009). Forecasting elections comparing prediction markets, polls, and their biases. *Public Opin. Q.*.
- Sunstein, C. R. (2007). Deliberating groups vs. prediction markets (or hayek’s challenge to habermas). *Episteme: A Journal of Social Epistemology* 3(3), 192–213.
- Tziralis, G. and I. Tatsiopoulos (2007). Prediction markets: An extended literature review. *The Journal of Prediction Markets* 1(1), 75–91.
- Wolfers, J. and E. Zitzewitz (2004). Prediction markets. *Journal of Economic Perspectives* 18(2), 107–126.

DEPARTMENT OF ECONOMICS, UNIVERSITY OF CALIFORNIA, RIVERSIDE

E-mail address: `urmee.khan@ucr.edu`

DEPARTMENT OF ECONOMICS, CENTRAL EUROPEAN UNIVERSITY, BUDAPEST

E-mail address: `rlieli@gmail.com`

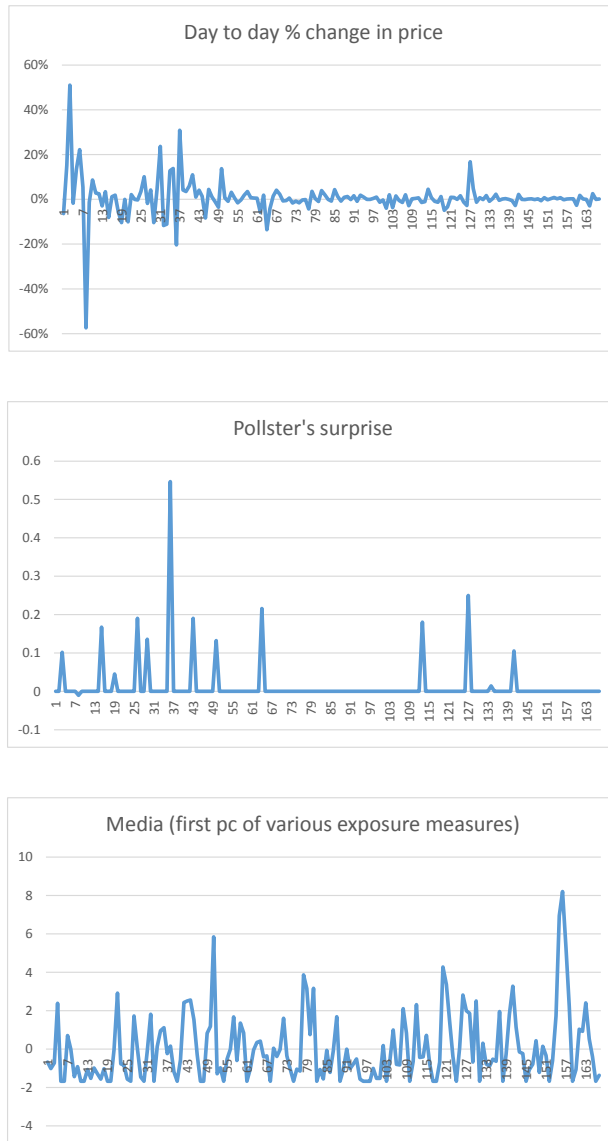


FIGURE 1. The price, surprise and media variables

TABLE 1. Testing [Surprise \Rightarrow Price] in the presence of Media; $(X, Y, Z) = (P, S, M)$

VAR order	STEP 1:		STEP 2:				
	$S \xrightarrow{1} (P, M)$	$(S, M) \xrightarrow{1} P$	$S \xrightarrow{1} P$	$S \xrightarrow{1} M$	$M \xrightarrow{1} P$		
7	0.000	0.003	0.005	0.000	0.011		
8	0.000	0.002	0.003	0.000	0.023		
10	0.000	0.001	0.017	0.000	0.044		
14	0.000	0.000	0.008	0.000	0.050		
18	0.000	0.000	0.000	0.000	0.098		
	STEP 3: $S \xrightarrow{j} P$ jointly for $1 \leq j \leq h$						
	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$	$h = 8$
7	0.009	0.015	0.007	0.007	0.008	0.011	0.003
8	0.006	0.010	0.003	0.002	0.003	0.005	0.002
10	0.025	0.025	0.008	0.007	0.009	0.010	0.001
14	0.010	0.012	0.004	0.004	0.002	0.003	0.003
18	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	STEP 3': $S \xrightarrow{h} P$ individually for $h = 2, \dots, 8$						
	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$	$h = 8$
7	0.069	0.509	0.332	0.014	0.403	0.513	0.443
8	0.500	0.365	0.155	0.021	0.552	0.574	0.335
10	0.218	0.422	0.181	0.125	0.449	0.674	0.643
14	0.651	0.642	0.158	0.004	0.441	0.557	0.334
18	0.030	0.026	0.003	0.112	0.185	0.086	0.013

Note: All p -values are from robust Wald tests conducted on the coefficients of a three-variable VAR.

TABLE 2. Testing [Price \Rightarrow Surprise] in the presence of Media; $(X, Y, Z) = (S, P, M)$

VAR order	STEP 1:		STEP 2:				
	$P \xrightarrow{1} (S, M)$	$(P, M) \xrightarrow{1} S$	$P \xrightarrow{1} S$	$P \xrightarrow{1} M$	$M \xrightarrow{1} S$		
7	0.000	0.055	0.008	0.056	0.226		
8	0.000	0.106	0.024	0.021	0.303		
10	0.000	0.000	0.000	0.005	0.345		
14	0.000	0.000	0.000	0.001	0.614		
18	0.000	0.130	0.036	0.001	0.620		
	STEP 3: $P \xrightarrow{j} S$ jointly for $1 \leq j \leq h$						
	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$	$h = 8$
7	0.009	0.015	0.015	0.024	0.035	0.040	0.055
8	0.016	0.026	0.026	0.037	0.054	0.066	0.079
10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
14	0.000	0.000	0.000	0.000	0.000	0.000	0.000
18	0.034	0.036	0.048	0.058	0.034	0.024	0.031
	STEP 3': $P \xrightarrow{h} S$ individually for $h = 2, \dots, 8$						
	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$	$h = 8$
7	0.018	0.019	0.004	0.047	0.000	0.000	0.241
8	0.012	0.000	0.012	0.001	0.000	0.000	0.366
10	0.000	0.000	0.000	0.000	0.344	0.311	0.696
14	0.116	0.072	0.082	0.334	0.201	0.342	0.855
18	0.117	0.223	0.478	0.695	0.596	0.402	0.500

Note: All p -values are from robust Wald tests conducted on the coefficients of a three-variable VAR.

TABLE 3. Testing [Surprise \Rightarrow Media] in the presence of Price; $(X, Y, Z) = (M, S, P)$

VAR order	STEP 1:		STEP 2:				
	$S \overset{1}{\nrightarrow} (M, P)$	$(S, P) \overset{1}{\nrightarrow} M$	$S \overset{1}{\nrightarrow} M$	$S \overset{1}{\nrightarrow} P$	$P \overset{1}{\nrightarrow} M$		
7	0.000	0.000	0.000	0.005	0.056		
8	0.000	0.000	0.000	0.003	0.021		
10	0.000	0.000	0.000	0.017	0.005		
14	0.000	0.000	0.000	0.008	0.001		
18	0.000	0.000	0.000	0.000	0.001		
	STEP 3: $S \overset{j}{\nrightarrow} M$ jointly for $1 \leq j \leq h$						
	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$	$h = 8$
7	0.000	0.000	0.000	0.000	0.000	0.000	0.000
8	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
14	0.000	0.000	0.000	0.000	0.000	0.000	0.000
18	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	STEP 3': $S \overset{h}{\nrightarrow} M$ individually for $h = 2, \dots, 8$						
	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$	$h = 8$
7	0.000	0.000	0.000	0.000	0.000	0.000	0.000
8	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
14	0.000	0.000	0.000	0.000	0.000	0.000	0.000
18	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note: All p -values are from robust Wald tests conducted on the coefficients of a three-variable VAR.

TABLE 4. Testing [Media \Rightarrow Surprise] in the presence of Price; $(X, Y, Z) = (S, M, P)$

VAR order	STEP 1:		STEP 2:				
	$M \xrightarrow{1} (S, P)$	$(M, P) \xrightarrow{1} S$	$M \xrightarrow{1} S$	$M \xrightarrow{1} P$	$P \xrightarrow{1} S$		
7	0.018	0.055	0.226	0.011	0.008		
8	0.102	0.106	0.303	0.023	0.024		
10	0.151	0.000	0.345	0.044	0.000		
14	0.046	0.000	0.614	0.050	0.000		
18	0.022	0.130	0.620	0.098	0.036		
	STEP 3: $M \xrightarrow{j} S$ jointly for $1 \leq j \leq h$						
	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$	$h = 8$
7	0.305	0.212	0.233	0.253	0.318	0.386	0.055
8	0.389	0.280	0.275	0.296	0.362	0.432	0.086
10	0.415	0.468	0.478	0.364	0.425	0.487	0.000
14	0.671	0.734	0.787	0.774	0.808	0.843	0.093
18	0.682	0.727	0.778	0.812	0.830	0.858	0.639
	STEP 3': $M \xrightarrow{h} S$ individually for $h = 2, \dots, 8$						
	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$	$h = 8$
7	0.268	0.213	0.441	0.552	0.310	0.260	0.437
8	0.290	0.263	0.603	0.456	0.358	0.360	0.784
10	0.432	0.370	0.478	0.533	0.690	0.579	0.819
14	0.620	0.676	0.759	0.756	0.719	0.715	0.925
18	0.521	0.577	0.869	0.832	0.847	0.852	0.969

Note: All p -values are from robust Wald tests conducted on the coefficients of a three-variable VAR.

TABLE 5. Testing [Price \Rightarrow Media] in the presence of Surprise; $(X, Y, Z) = (M, P, S)$

VAR order	STEP 1:		STEP 2:				
	$P \xrightarrow{1} (M, S)$	$(S, P) \xrightarrow{1} M$	$P \xrightarrow{1} M$	$P \xrightarrow{1} S$	$S \xrightarrow{1} M$		
7	0.000	0.000	0.056	0.008	0.000		
8	0.000	0.000	0.021	0.024	0.000		
10	0.000	0.000	0.005	0.000	0.000		
14	0.000	0.000	0.001	0.000	0.000		
18	0.000	0.000	0.001	0.036	0.000		
	STEP 3: $P \xrightarrow{j} M$ jointly for $1 \leq j \leq h$						
	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$	$h = 8$
7	0.087	0.118	0.158	0.006	0.006	0.001	0.000
8	0.033	0.045	0.065	0.003	0.005	0.000	0.000
10	0.008	0.010	0.014	0.001	0.001	0.000	0.000
14	0.001	0.002	0.003	0.000	0.001	0.000	0.000
18	0.002	0.003	0.004	0.000	0.000	0.000	0.000
	STEP 3': $P \xrightarrow{h} M$ individually for $h = 2, \dots, 8$						
	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$	$h = 8$
7	0.589	0.244	0.457	0.540	0.935	0.602	0.031
8	0.501	0.283	0.089	0.408	0.764	0.188	0.000
10	0.055	0.057	0.239	0.156	0.002	0.004	0.000
14	0.001	0.003	0.012	0.007	0.002	0.078	0.004
18	0.001	0.004	0.001	0.003	0.000	0.000	0.000

Note: All p -values are from robust Wald tests conducted on the coefficients of a three-variable VAR.

TABLE 6. Testing [Media \Rightarrow Price] in the presence of Surprise; $(X, Y, Z) = (P, M, S)$

VAR order	STEP 1:		STEP 2:				
	$M \xrightarrow{1} (P, S)$	$(M, S) \xrightarrow{1} P$	$M \xrightarrow{1} P$	$M \xrightarrow{1} S$	$S \xrightarrow{1} P$		
7	0.018	0.003	0.011	0.226	0.005		
8	0.102	0.002	0.023	0.303	0.003		
10	0.151	0.001	0.044	0.345	0.017		
14	0.046	0.000	0.050	0.614	0.008		
18	0.022	0.000	0.098	0.620	0.000		
	STEP 3: $M \xrightarrow{j} P$ jointly for $1 \leq j \leq h$						
	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$	$h = 8$
7	0.000	0.001	0.001	0.003	0.001	0.003	0.003
8	0.004	0.006	0.008	0.013	0.002	0.003	0.002
10	0.002	0.000	0.001	0.001	0.000	0.000	0.000
14	0.000	0.000	0.000	0.000	0.000	0.000	0.000
18	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	STEP 3': $M \xrightarrow{h} P$ individually for $h = 2, \dots, 8$						
	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$	$h = 8$
7	0.132	0.167	0.348	0.085	0.091	0.113	0.089
8	0.408	0.123	0.052	0.157	0.164	0.039	0.161
10	0.034	0.017	0.142	0.071	0.079	0.111	0.250
14	0.205	0.143	0.223	0.399	0.326	0.398	0.139
18	0.413	0.130	0.067	0.029	0.027	0.014	0.017

Note: All p -values are from robust Wald tests conducted on the coefficients of a three-variable VAR.

TABLE 7. Data used in constructing pollster's surprise for Obama

State	Primary date	Delegates at stake	% Delegates at stake	Predicted vote share	Actual vote share	Surprise (raw)
Iowa	3-Jan	45	1.32	29	37.5	7.7
New Hampshire	8-Jan	22	0.65	38	36.4	-1.6
Michigan	15-Jan	64	1.88	31	39.9	8.9
Nevada	19-Jan	25	0.73	39	45.2	6.2
South Carolina	26-Jan	45	1.32	41	55.4	14.4
Florida	29-Jan	92.5	2.71	28	33.0	5.0
California	5-Feb	370	10.85	45.5	42.3	-3.2
New York	5-Feb	232	6.80	39	39.9	0.9
Illinois	5-Feb	153	4.49	66.0	64.5	-1.5
New Jersey	5-Feb	107	3.14	42.0	44.0	2.0
Massachusetts	5-Feb	93	2.73	41.7	40.8	-0.9
Georgia	5-Feb	87	2.5	49.0	66.4	17.4
Minnesota	5-Feb	72	2.11			
Missouri	5-Feb	72	2.11	45.0	49.2	4.2
Tennessee	5-Feb	68	1.99	35	40.5	5.5
Colorado	5-Feb	55	1.61			
Arizona	5-Feb	56	1.64	41	41.9	0.9
Alabama	5-Feb	52	1.53	45	55.8	10.8
Connecticut	5-Feb	48	1.41	48	50.7	2.7
Arkansas	5-Feb	35	1.03			
Oklahoma	5-Feb	38	1.11			
Kansas	5-Feb	32	0.94			
New Mexico	5-Feb	26	0.76			
Utah	5-Feb	23	0.67			
Delaware	5-Feb	15	0.44			
North Dakota	5-Feb	13	0.38			
Idaho	5-Feb	18	0.53			
Alaska	5-Feb	13	0.38			
American Samoa	5-Feb	3	0.09			

Notes: Predicted vote share for Obama is measured by the weighted average of polls that closed on the day before the election (the weights are given by the sample sizes). The final pollster's surprise is constructed as follows. On days with no elections, the variable is set to zero. On days with elections we multiply the raw surprise in each election with the fraction of delegates at stake and then sum over elections. Elections for which data was available account for 78.3% of all delegates.

TABLE 8. Data used in constructing pollster's surprise for Obama (continued)

State	Primary date	Delegates at stake	% Delegates at stake	Predicted vote share	Actual vote share	Surprise (raw)
Washington	9-Feb	78	2.29			
Louisiana	9-Feb	56	1.64			
Nebraska	9-Feb	24	0.70			
Virgin Islands	9-Feb	3	0.09			
Maine	10-Feb	24	0.70			
Virginia	12-Feb	83	2.43	60	63.6	3.6
Maryland	12-Feb	70	2.05	55	60.0	5.0
District of Columbia	12-Feb	15	0.44			
Democrats Abroad	12-Feb	7	0.21			
Wisconsin	19-Feb	74	2.17	52	58.1	6.1
Hawaii	19-Feb	20	0.59			
Texas	4-Mar	193	5.66	44	47.4	3.4
Ohio	4-Mar	141	4.14	43.5	44.1	0.6
Rhode Island	4-Mar	21	0.62			
Vermont	4-Mar	15	0.44			
Wyoming	8-Mar	12	0.35			
Mississippi	11-Mar	33	0.97			
Pennsylvania	22-Apr	158	4.63	41.5	45.4	3.9
Guam	3-May	4	0.12			
North Carolina	6-May	115	3.37	48.8	56.2	7.4
Indiana	6-May	72	2.11	45	49.3	4.3
West Virginia	13-May	28	0.82	24	25.7	1.7
Oregon	20-May	52	1.53	53.1	58.2	5.1
Kentucky	20-May	51	1.50	28.1	29.9	1.8
Puerto Rico	1-Jun	55	1.61			
Montana	3-Jun	16	0.47			
South Dakota	3-Jun	15	0.44			

Notes: Predicted vote share for Obama is measured by the weighted average of polls that closed on the day before the election (the weights are given by the sample sizes). The final pollster's surprise is constructed as follows. On days with no elections, the variable is set to zero. On days with elections we multiply the raw surprise in each election with the fraction of delegates at stake and then sum over elections. Elections for which data was available account for 78.3% of all delegates.