When are prediction market prices most informative?

March 23, 2018

Abstract

Prediction markets are a popular platform for eliciting incentivised crowd predictions. In this paper, we examine variation in the information contained in prediction market prices by studying Intrade prices on U.S. elections around the release of opinion polls. We find that poll releases stimulate an immediate uptick in trading activity. However, much of this activity involves relatively inexperienced traders and, as a result, price efficiency declines in the immediate aftermath of a poll release. It is not until more experienced traders enter the market in the following hours that price efficiency recovers. More generally, this suggests that information releases do not necessarily improve prediction market forecasts, but may instead attract noise traders who temporarily reduce price efficiency.

JEL Classification: C53, G14, D17, D82, D83.
Keywords: Prediction markets, opinion polls, price efficiency, information efficiency.

1 Introduction

Prediction markets allow for trading on the outcome of future events. They are a popular method for forecasting, as individuals are well-incentivised to acquire information and produce accurate forecasts. In addition, the prices produced by these markets aggregate dispersed information and therefore harness the 'wisdom of the crowd' (Galton, 1907; Surowiecki, 2004).

But when are prediction market prices most informative? To answer this question we study a rich dataset from the now-defunct prediction market Intrade. We analyse every transaction on U.S. elections (national and state-level) between 2008 and 2012. Crucially, the data include anonymised trader identification. We marry this data with information on poll releases from Gallup, the oldest polling firm in the U.S. This allows us to examine trader activity in the hours surrounding poll releases — who traded and how much — and also allows us to examine the effect of this activity on the accuracy, or efficiency, of prices.

We find that poll releases stimulate increased trading volume. However, it is the inexperienced traders (with less prior trading activity) who initially respond to the poll release. Furthermore, these traders seem to respond more to the incidence of a poll rather than to its content. As a result, we observe a significant decline in price efficiency in the hour after a poll release. This decline is interrupted by the arrival of more experienced traders in the following hours, who correct an element of this mispricing (Hanson and Oprea, 2009). This suggests that information releases may actually temporarily harm prediction market price accuracy, as these releases attract the attention of less experienced noise traders (De Long et al., 1990). In short, if an individual is basing their forecast on prediction market prices, it may be worth turning away after a significant information event.

This paper contributes to the literature on the accuracy of prediction market prices. Authors have compared prediction markets to opinion polls (Leigh and Wolfers, 2006; Chen et al., 2005; Sjöberg, 2009; Vaughan Williams and Reade, 2016a; Wang et al., 2015), have combined them (Graefe et al., 2014; Rothschild, 2015), have considered the predictive power of social media content (Huberty, 2015; D'Amuri and Marcucci, 2017; Peeters, 2018), and also married up prediction market forecasts with social media content (Brown et al., 2017; Vaughan Williams and Reade, 2016b). There are perhaps three papers most closely related to ours. Croxson and Reade (2014) studied the accuracy of betting (prediction) market prices in the immediate aftermath of soccer goals. They found that prices respond almost instantaneously, and indeed accurately, to information. Another closely related paper is Page and Clemen (2013) who studied prediction market accuracy over time. They found that prediction market accuracy improved as the event in question approached. Notwithstanding this general increase in prediction market accuracy over time, Page (2012) shows that in-play betting prices toward the end of matches overestimate the likelihood of low probability outcomes (e.g. the losing team winning the match). We identify more noise around information events than Croxson and Reade (2014), and also show more bumps in the road towards efficiency relative to Page and Clemen (2013).

The rest of the paper is structured as follows. Section 2 introduces our dataset, Section 3 details our methodology, and Section 4 presents our results. Section 5 concludes.

2 Data

We use data from the now-defunct Intrade prediction market between 2008 and 2012. The dataset begins after the 2008 US presidential election and runs through to the end of the 2012 presidential election. The overwhelming majority of contracts are associated with US politics over the four-year period.

The data includes a timestamp (in Co-ordinated Universal Time, UTC) for each trade, a unique identifier for both the buyer and seller, the number of contracts traded and the price they were traded at. A description of the contract being traded is also provided. In addition, there is information on which trader was the *aggressor*, which is the trader taking available liquidity.

The dataset includes 952,141 recorded trades of 25,192,197 contracts, implying that on average, traders traded 26.5 contracts per trade. There are 8,236 individual traders who buy and sell contracts over the time period. Half of those traders trade 18 or fewer contracts, and 807 trade just once. At the other end of the spectrum, 226 traders trade over a thousand times, 23 over 10,000 times and just 3 traders more than 100,000 times over the four year period.

This dataset has also been analysed by Rothschild and Sethi (2016), who build up an ecosystem of traders over the two week period immediately before the 2012 presidential election. Amongst other things, they attempt to distinguish different types of trader behaviour based on aggression. Do traders take existing liquidity, or instead provide it? Our data allows us to consider this, because although we cannot know *when* the passive (non-aggressive) trader placed their trade, we know that they purposely placed an order at a price at which no existing contracts were available, hence providing liquidity.

We take opinion polls for the 2008–2012 cycle from Pollster. We focus on Gallup polls because of their frequency, and also because of the historical stature of the Gallup polling company. From May 2012 onwards, Gallup sampled for six or seven days, before releasing the subsequent poll the day after polling closed, at 1pm Eastern Standard Time. A press release, and a Tweet on Twitter accompanied each release. Over the cycle, Gallup released 198 polls, and between April 15 and October 28, a week before the election, it released a poll daily. Gallup polled likely and registered voters over that period,

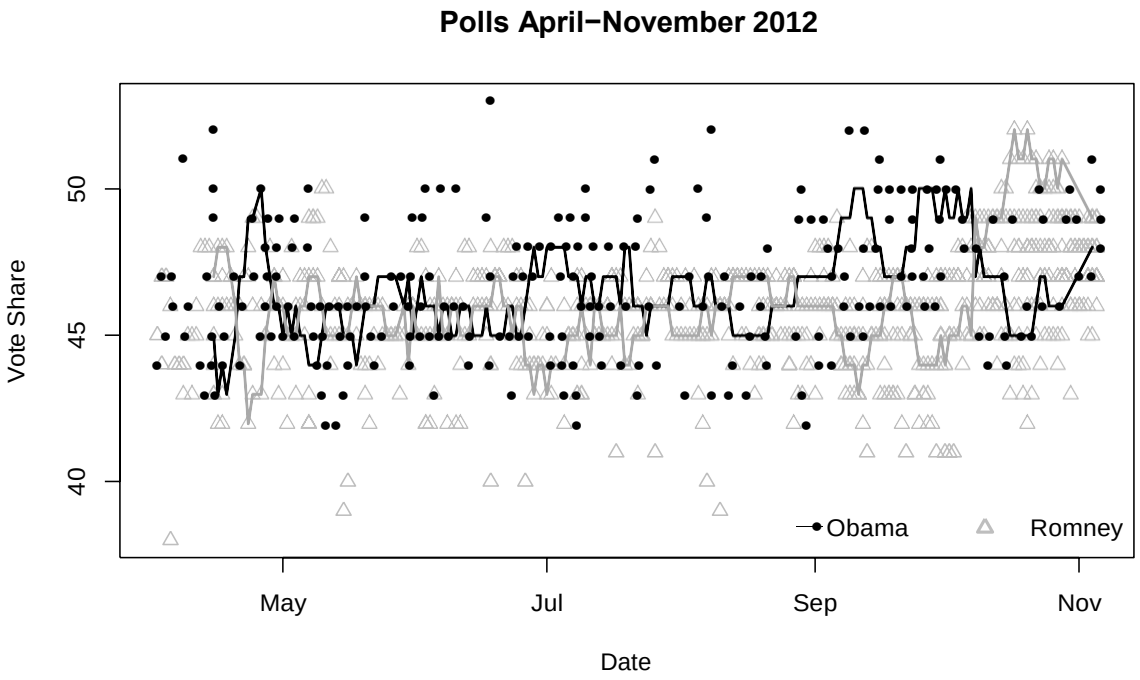

Figure 1: Polled vote shares from all released polls in the seven months up to the 2012 US Presidential election. The solid lines link all Gallup polls during the same period.

with samples sizes generally between two and three thousand.

We focus on Gallup also as, for this election, they were notably out of line with other polls in the final days before the election itself, as Figure 1 shows. The black circles are polled vote shares for Obama, the grey triangles are vote shares for Romney. The lines link up all Gallup polls released over the same period, which runs from April through to the election in early-November of 2012. From mid-October Romney polls ahead of Obama consistently, although the gap closes to only one point by the final poll.

Trades on Intrade are recorded at Coordinated Universal Time (UTC), which until November 4 was four hours ahead of Eastern Standard Time (EST), and hence as Gallup polls were released at 1pm Eastern Standard Time, they were released at 5pm UTC.

3 Methodology

Our methodology is to run a number of linear regressions in order to determine the significance or otherwise of salient information around the time of Gallup poll releases. Considering prediction market movement on the day of poll releases in general is relatively uninformative since on most days in 2012 a poll was released. However, Gallup always released polls at 1pm Eastern Time.[1] Our strategy is to create dummy, or indicator variables for each trade in our dataset that are 1 if that trade:

[1]Daylight saving time ended on November 4 2012, two days before the election, and the day before the release of the final Gallup poll. We take note of this in our analysis.

- Occurred between 11am and 12 noon on the day that a Gallup poll was released at 1pm (all times EST).
- Occurred between 12 noon and 1pm on the day that a Gallup poll was released at 1pm.
- Occurred between 1pm and 2pm on the day that a Gallup poll was released at 1pm.
- Occurred between 2pm and 3pm on the day that a Gallup poll was released at 1pm.

We then include these dummy variables in a range of regression models, and comment on their significance (or otherwise). We consider simple regressions of the form:

$$Y_i = \beta_0 + X_i\beta + ttallup_i\gamma + u_i,\ u \sim N(0, \sigma^2 \Sigma), \tag{1}$$

where X_i is a matrix of fixed effects and control variables, while $ttallup_i$ is a matrix of Gallup-poll specific indicator variables, and Y_i is some variable of market activity of interest (quantity traded, trader experience, aggression index). We anticipate $\gamma \neq 0$ in order to identify some impact of Gallup poll releases on market activity.

When considering the impact on the market price, we employ a Mincer and Zarnowitz (1969) (MZ) regression test. This is a regression of an outcome variable, z_i, here whether or not an Intrade contract paid out, on a forecast variable, $\hat{z}_i$, here the market price from Intrade:

$$z_i = \beta_0 + \beta_1 \hat{z_i} + u_i,\ u \sim N(0, \sigma^2 \Sigma). \tag{2}$$

The principle of the MZ test is to determine whether the market price alone is a sufficient statistic for considering the outcome of the event being traded. This implies a null hypothesis of $\beta_0 = 1 - \beta_1 = 0$ such that the forecast error is mean zero with constant variance: $z_i - \hat{z}_i \sim N(0, \sigma^2 \Sigma)$ The further implication is that no additional variables ought to be significant in (2), as only $\hat{z}_i$ should be necessary to predict z_i. The MZ test is often employed in the context of market prices such as that of Intrade, in order to determine whether or not the price is an efficient forecast of outcomes. We thus add dummy variables for the incidence of a Gallup poll, and interact the dummy variables with the price variable to maximise the potential ways in which a poll can influence market pricing:

$$z_i = \beta_0 + \beta_1 \hat{z_i} + \gamma_0 ttallup_i + \gamma_1 \hat{z_i} \times ttallup_i + u_i, u \sim N(0, \sigma^2 \Sigma), \tag{3}$$

and consider whether $\gamma_0 = \gamma_1 = 0$ as a test of the impact of a Gallup poll on market activity. In addition, in the context of betting markets, a coefficient of $\beta_1 > 1$ implies the existence of the favourite-longshot bias, one particular form of market inefficiency where more likely events (favourites) occur relatively more frequently than their predictions imply that they should.

4 Results

4.1 Gallup relative to the polling average

4.2 Influence of polling outcomes on prediction market activity

We firstly consider whether there was any impact of a Gallup poll on the total amount of trading on Intrade. We aggregate by hour of the day, and consider the total quantity of contracts traded each hour.

The first column of Table 1 displays the result of regressing total quantity traded per hour on dummies for the proximity of a Gallup poll release. The results suggest that for the four hour window around a Gallup poll there is a significant increase in trading activity, with the peak occurring in the hour after the poll was released at 1pm.

This is a rather basic regression, however, and does not reflect that the total quantity traded varied by proximity to the election, by day of the week, and hour of the day. Including indicator variables for all of these seasonal patterns, in addition to a variable calculating the length of time until each contract expires, will provide a more accurate idea of whether trading was greater around a Gallup poll release. The first column in Table 2 plots the same regression coefficients as in Table 1, but from a regression with these extra variables added.[2] A similar pattern results, with the hour after the poll release having the greatest increase in trading, of around a thousand contracts. The subsequent hour, between 2 and 3pm, is about half the size at 554 contracts, but is insignificant with a p-value of 0.11.

Hence trading activity increases with the existence of a Gallup poll release.

4.3 Influence of polling outcomes on prediction market pricing

To consider the impact of polling outcomes from Gallup on the efficiency of pricing on Intrade, we run a Mincer-Zarnowitz (MZ) regression test. The null hypothesis of an efficient market is an intercept coefficient of zero, and a slope coefficient of one, and furthermore that any additional variables added be insignificant.

The second column of Table 1 reports the estimation of (3), where $ttallup_i$ contains dummies for the four hour window around a 1pm release of a Gallup poll. An F-test for the restriction that $\beta_0 = 1 - \beta_1 = \gamma_0 = \gamma_1 = 0$ is emphatically rejected. The size of the t-statistics makes clear the deviation from efficiency. The finding of a favourite-longshot bias, by itself, however, is in line with previous findings. **?**, for example, find that although smaller on prediction markets, a favourite-longshot bias is nonetheless found for horse racing betting. Reade (2014) confirms this finding for betting on English football.

The finding of interest, though, is that the inefficiency (and favourite-longshot bias) is increased around the release of a Gallup poll. In the hour immediately after the poll is released, the inefficiency is strongest as the slope coefficient increases from 1.24 to 1.37. The following hour the slope coefficient falls back down to 1.3. The changes in the β_0 coefficient are also consistent with this temporal pattern in inefficiency.

Again, however, it may be that there are other systematic influences on the efficiency of prices that should be controlled for in our analysis. As with trading quantity, we thus add control variables for the hour of the day, for the day of the week, for the month and for the year, along with a variable calculating the length of time until the contract expires. The same coefficients reported in Table 1 are reported in Table 2 from a regression including all of these control variables. We interact these extra variables with the price variable, also. A considerable number of these coefficients were highly significant, evidence against the efficiency of markets. In principle, traders could make use of such time-related characteristics in market pricing to develop a profitable trading strategy. Of most interest though are the coefficients on the Gallup Poll dummy variables. They remain significant, and in the same direction as before: pointing towards inefficiency around the time of a Gallup poll release.[3]

[2] In the interests of space we do not report these coefficients.

[3] Note that the β_1 coefficient is -0.58, which is very different to that in Table 1. This is because of the influence of the added variables. For example, the indicator variable for the year 2012 is around 1.5, and for October and November is 0.5, implying that indeed markets remain inefficient in the direction of favourite longshot bias, as in Table 1.

Table 1: Regressions of total quantity traded (column 1), and price (column 2), on indicator variables for the hours surrounding the release of a Gallup opinion poll.

	Dependent variable:	
	Trade Quantity	Outcome
	(1)	(2)
Constant	1,236.279***	−0.047***
	(33.032)	(0.0003)
11am – 12 noon on Gallup release day	2,817.171***	−0.038***
	(315.312)	(0.002)
12 noon – 1pm on Gallup release day	2,655.504***	−0.044***
	(315.312)	(0.002)
1pm – 2pm on Gallup release day	2,865.459***	−0.055***
	(315.312)	(0.002)
2pm – 3pm on Gallup release day	2,535.125***	−0.035***
	(315.312)	(0.002)
price		1.235***
		(0.001)
price × 11am – 12 noon on Gallup release day		0.099***
		(0.005)
price × 12 noon – 1pm on Gallup release day		0.104***
		(0.005)
price × 1pm – 2pm on Gallup release day		0.136***
		(0.005)
price × 2pm – 3pm on Gallup release day		0.074***
		(0.005)
Observations	18,636	952,141
R^2	0.015	0.707
Adjusted R^2	0.015	0.707
Residual Std. Error	4,412.421 (df = 18631)	1.183 (df = 952131)
F Statistic	72.128*** (df = 4; 18631)	255,867.300*** (df = 9; 952131)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 2: Linear regression of the number of trades in an hour on a number of explanatory variables, including the incidence of a Gallup poll at the start of that hour.

	Dependent variable:	
	Hourly Trade Quantity	Outcome
	(1)	(2)
Constant	−9,095.539***	0.231***
	(913.084)	(0.036)
Gallup Poll (hour)	1,076.221***	−0.026***
	(340.007)	(0.003)
Time until contract expiry	−3.987***	0.0002***
	(0.495)	(0.00001)
Time until contract expiry × price		−0.001***
		(0.00001)
price		−0.583***
		(0.094)
price × Gallup Poll (hour)		0.045***
		(0.007)
Observations	18,588	952,023
R^2	0.124	0.729
Adjusted R^2	0.122	0.729
Residual Std. Error	4,161.927 (df = 18541)	1.139 (df = 951929)
F Statistic	57.031*** (df = 46; 18541)	27,555.950*** (df = 93; 951929)
Note:		*p<0.1; **p<0.05; ***p<0.01

4.4 Influence of polling outcomes on prediction market participants

How did a prominent pollster releasing polling information affect the market? We look at two aspects. We consider how experienced traders were when trading, as we can calculate this for each trader in each trade. Specifically, we count how many trades had a trader previously been involved in when engaging in a recorded trade, and add our dummy variables for the proximity of a Gallup poll release, as well as fixed effects for the hour of the day, day of the week, month of the year and year. This can inform us as to whether the release of information coincided with periods where systematically more experienced traders were trading, or coincided with such traders trading less.[4]

We also consider an index of aggression, calculated in real time in the same manner as Rothschild and Sethi (2016). That is, the proportion of all trades before the current trade being recorded that were aggressive (taking liquidity) or passive (providing liquidity). Did the release of polling information from Gallup occur at times with more aggressive, or more passive traders? We run the same kind of linear regression as for experience.

Table 3 records these regression outputs, showing just the coefficients relating to Gallup poll releases, and the constant term.[5] Considering column (1) for trader experience, trader experience increases over the four hours surrounding a poll release. Traders are relatively inexperienced (over a thousand fewer trades per trader, on average) between 11am and 1pm, then insignificantly less experienced in the hour that the poll was released, and the following hour significantly more experienced (around 1600 trades per trader, on average). Hence it appears that more experienced traders arrived considerably after the poll had been released, and avoided the market before that.

This seems somewhat at odds with the idea expressed in Hanson and Oprea (2009), namely that an event that influences market pricing away from some previous level (in their example, a market manipulator) will be exploited by traders quickly in order to restore an efficient price level. If it is only with some delay that experienced traders enter the market and take liquidity, it rather suggests that such traders are taking time for distorted market positions to emerge in light of an information event before trading.

Considering column (2) for the index of aggression, this is significantly higher before the poll is released (between 11am and 1pm), by about 0.7 percentage points, but during the hour of the release is less significantly higher (p-value 0.074), and the next hour after the release (2pm to 3pm) this is 1 percentage point higher on average. Hence only with a delay does the average trader become more aggressive, after a Gallup poll release.

Hence we can characterise the market before the poll release as less experienced, but a little more aggressive in terms of taking liquidity, the market in the immediate hour after the poll release as less experienced, and the hour following that as more experienced traders who are more aggressively taking liquidity. This helps to explain the finding that markets are less efficient in the hour that a Gallup poll is released: traders are less experienced, and less aggressive. Only once more experienced and aggressive traders return over an hour after the poll does the market inefficiency fall.

4.5 Do Salient Polls Matter More?

Third, we look at the impact of a poll that was out of line with general polling.

[4] As the data is anonymised, we are only able to determine a trader's experience using Intrade since 2008.

[5] Note that the constant term is interpretable as the mean value of the dependent variable *conditional* on all the explanatory variables, hence can be difficult to interpret.

Table 3: Linear regressions of number of trades and aggression index on fixed effects for hour of day, day of week, month of year and year, as well as dummies for Gallup poll releases.

	Dependent variable:	
	Trader experience	Aggressor index
	(1)	(2)
Constant	−28,238.140***	0.419***
	(1,487.204)	(0.010)
11am – 12 noon on Gallup release day	−1,153.619***	0.007***
	(242.524)	(0.002)
12 noon – 1pm on Gallup release day	−955.320**	0.007**
	(432.907)	(0.003)
1pm – 2pm on Gallup release day	−603.155	0.005*
	(422.176)	(0.003)
2pm – 3pm on Gallup release day	1,602.567***	0.010***
	(263.188)	(0.002)
Observations	1,904,282	1,904,282
R^2	0.096	0.006
Adjusted R^2	0.096	0.006
Residual Std. Error (df = 1904233)	40,429.300	0.281
F Statistic (df = 48; 1904233)	4,197.874***	241.098***
Note:		*p<0.1; **p<0.05; ***p<0.01

We include the distance each Gallup poll was from the average of the other polls at the time, and we do so using indicator saturation. Indicator saturation allows us to calculate a trend line for opinion polls that reflects secular movements in polls, whilst controlling for outliers. We then subtract each Gallup poll from this trend line in order to measure how salient, or out of line, any released Gallup poll was.

The dependent variable we use for this process is the polled difference between Obama and Romney. This way, we capture information about the polled vote shares for both candidates, which is usually the focus of any opinion poll release. Indicator saturation inserts an indicator variable for each observation in a sample (in batches to avoid multicollinearity), and retains only the significant indicator variables (Hendry et al., 2008).[6] It can be employed to model the mean of a series subject to outliers and structural shifts, as the polling average in an election campaign might. While outlier indicator variables might be used, it is more efficient to include step indicator variables for each observation in a sample, adding them in batches in order to avoid perfect multicollinearity, and selecting over all of the indicators in order to determine which are retained (Hendry et al., 2013).[7] It is a method for detecting both outliers and structural breaks (see, for example, Hendry et al., 2008; Johansen and Nielsen, 2009; Hendry and Pretis, 2011).[8]

Indicator saturation involves running regression of the form:

$$\Delta poll_{it} = \beta_0 + D\beta + u_{it}, \qquad u_{it} \sim \mathsf{N}\left[0, \sigma^2\right], \tag{4}$$

where D is a matrix of step indicator variables, and $\Delta poll_{it}$ is the difference between the Obama and Romney vote share in poll i at time t. When adding in batches of indicators, D takes a different form, and in the final output of the process, D represents the retained step indicator variables.

Note that $\Delta poll_{it}$ is not a conventional time series, but neither is it a panel dataset given that only few pollsters report on a regular basis. However, we purport a polling average to exist, and indeed many such averages are calculated by various media outlets in and around election times.

Figure 2 plots as a solid line the average polling difference between Obama and Romney found using indicator saturation through all the polls released by all companies. Only the segment of the line between April and November 2012 is plotted as this corresponds to the period when Gallup released a poll most days. Each Gallup poll is plotted as a cross on the diagram. The vertical axis is the difference between the Obama poll share and the Romney poll share. The Gallup polls before the election were thus considerably away from the general range of polling outcomes at the time. Equally, however, in July Gallup produced a number of polls more generous to Obama than the general consensus at the time. As such, Gallup often deviated substantially from the consensus amongst pollsters.

All of the retained step dummies are reported in the rows of Table 4, and the naming convention is not particularly illuminating, as it is merely the observation number. Nonetheless, we note that 11 such step shift indicators are retained. Two adjacent step indicators, such as for observations 900 and 901 (sis900, and sis901 respectively), model an impulse indicator, or an outlier. The flexibility of modelling outliers with two step indicators is that the subsequent mean level of polling can differ, and indeed this is the case, as the outlier is a full 17 polling points out from the average before that poll, but subsequent

[6] Indicator variables are often also referred to as dummy variables, or deterministic variables. They are terms in a regression model that take particular values by design, rather than being random variables like standard regressors. The most common indicator variable is an outlier indicator, which takes the value one for one observation (the outlier), and zero for all others.

[7] By step indicator we mean a variable which is equal to zero for all observations less than some point t_0, and equal to one thereafter. Such indicator variables are commonly used to model structural breaks in time series data, but equally two step indicators adjacent to each other (say a second with $t_1 = t_0 + 1$) is the same as an outlier indicator.

[8] We use the R package Preti et al. (2017) to carry out our analysis.

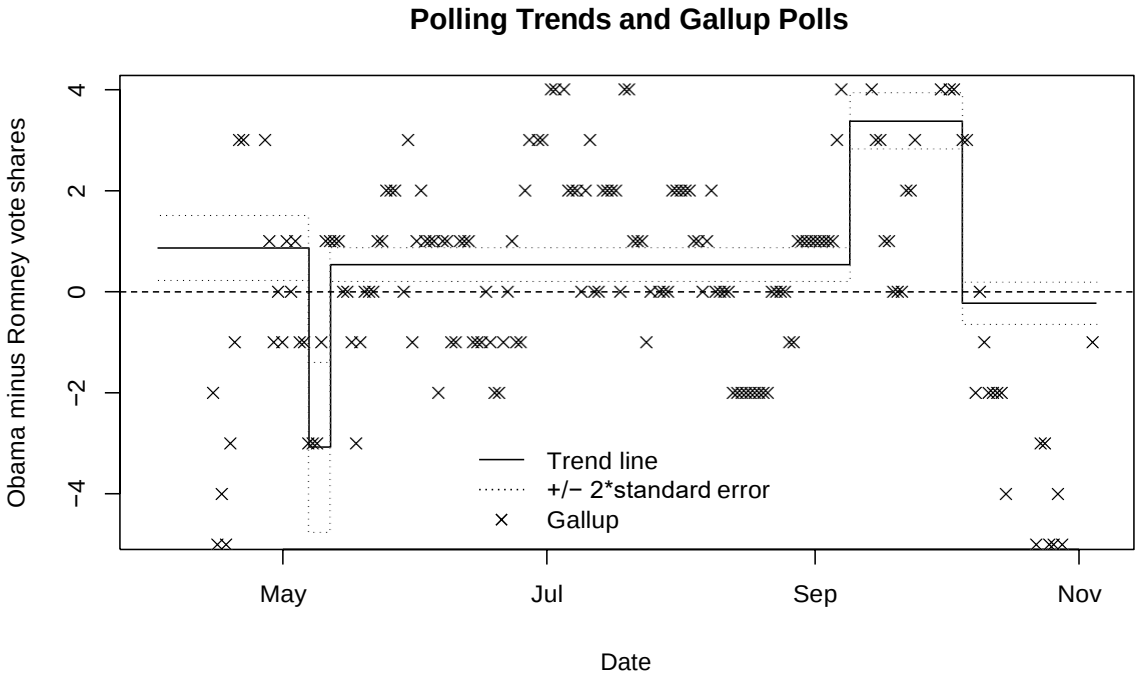

Figure 2: Trend line from polled vote shares from all released polls in the seven months up to the 2012 US Presidential election plotted as the solid line, with Gallup polls plotted as crosses.

polls are still two polling points below those that came before, as can be inferred by the fact that the subsequent dummy (note that the data are ordered such that a lower observation number is a more recent poll) is 19 polling points in size.

Table 4: Regression output from using indicator saturation to determine the average polling difference between the candidates in the 2012 US presidential election. Dependent variable is the polling difference between Obama and Romney, and the sample is all opinion polls reported on Pollster.

	coef	std.error	t-stat	p-value
Constant	-0.23	0.21	-1.09	0.28
sis212	3.61	0.35	10.38	0.00
sis331	-2.84	0.32	-8.78	0.00
sis663	-3.61	0.86	-4.22	0.00
sis676	3.94	0.90	4.38	0.00
sis765	2.34	0.65	3.62	0.00
sis794	-6.96	1.62	-4.31	0.00
sis798	8.80	1.58	5.56	0.00
sis842	-2.05	0.61	-3.38	0.00
sis900	-19.00	3.05	-6.22	0.00
sis901	17.26	3.06	5.65	0.00
sis954	2.29	0.54	4.24	0.00
sis1032	3.66	0.88	4.17	0.00

	Chi-sq	df	p-value
Ljung-Box AR(1)	5.99	1.00	0.01
Ljung-Box ARCH(1)	1.80	1.00	0.18
Jarque-Bera	109.71	2.00	0.00

Figure 3 plots the polling trend line against time for the *entire sample period* (Figure 2 considered only the months up to the election), along with *all* polls released over the time period. The majority of polls occurred during 2012, and this is reflected by the weight of observations plotted towards the right hand end of the scale. Most of the variation in the trend line also occurs in 2012, also. Outliers detected can be spotted via the poll, with three in 2011, and one in 2012, with a number of step shifts throughout, and a gradual convergence of the poll difference between the candidates towards zero as the election day became closer.

With the trend measure created, the next step is to determine which measure of Gallup polls relative to the trend to make use of in determining whether more salient Gallup polls had more impact. We experimented with both the residual, and the absolute value of the residual, in order to determine which, if either, generated larger coefficients. In the case of quantity traded, the absolute value would appear more appropriate since it might be expected that surprise polling releases in either direction would generate significant market activity.

In Table 5 we repeat the regressions for trading quantity, and market price, including the absolute value of the residuals for Gallup polls, and find that the impact on trading quantity is positive but insignificant (after controlling for other seasonal patterns), whereas the impact on market price is significant (again after seasonal controls inserted).

In Table 6 we repeat the regressions for trader experience and trader aggression, finding that in both cases there is an impact of the absolute size of the residual for a Gallup poll on experience and aggression. That is, the larger is a residual, the more experienced, on average, is a trader trading, and the more aggressive, on average, is a trader trading.

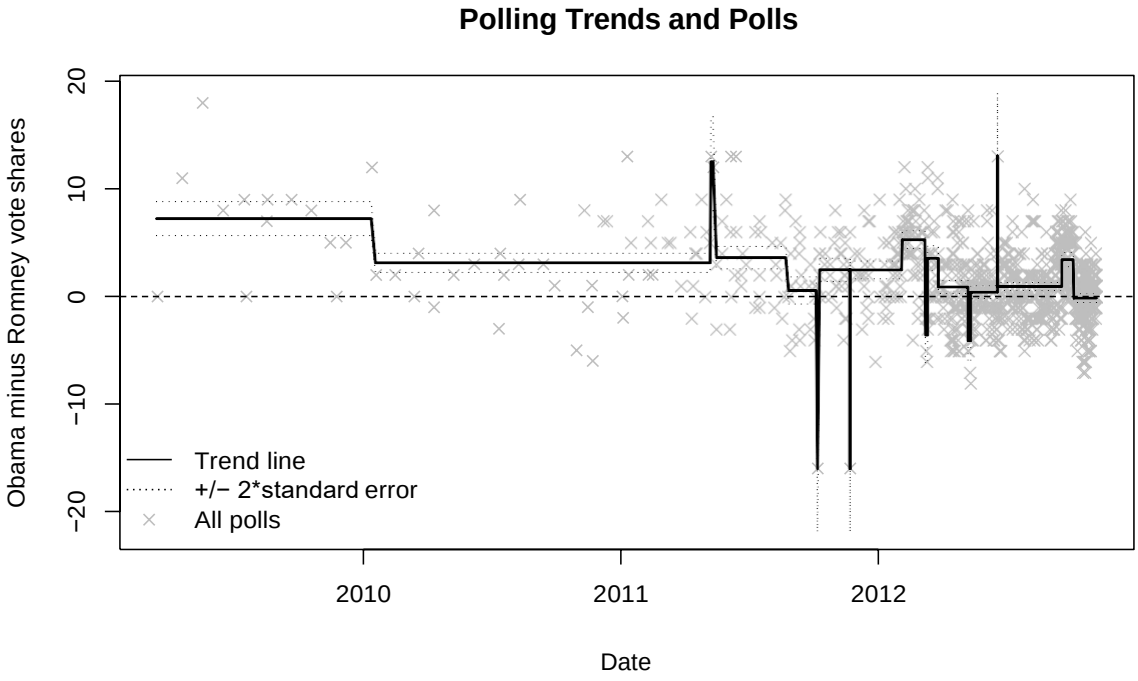

Figure 3: Plot of the polling trend line as selected via indicator saturation, along with each polling difference, plotted against time.

In general, the effect when controlling for the salience of a Gallup poll is of a less clear magnitude. This suggests that simply the existence of a Gallup poll generates market activity, rather than a Gallup poll that is particularly outstanding compared to other opinion polls. This is consistent also with the pattern of aggressiveness and experience; less experienced (noise) traders react to any poll, rather than necessarily its content, in particular one that is an outlier relative to the polling average.

5 Conclusions

Prediction markets often vie with opinion polls as forecasting tools for elections. In this paper we study the accuracy of prediction market prices around the release of opinion polls. We find an immediate decline in price accuracy, as inexperienced traders respond noisily to the incidence of a poll rather than to its content. This decline is temporary, as more experienced traders enter the market in the following hours. More generally, it would appear that information releases do not necessarily improve prediction market accuracy in the very short-term.

Table 5: Regressions including residuals of Gallup polls.

	Dependent variable:	
	Trade.Quantity	outcome2
	(1)	(2)
Constant	−8,994.465***	0.231***
	(912.054)	(0.036)
abs(residuals)	148.970	−0.004***
	(127.689)	(0.001)
price		−0.584***
		(0.094)
price:abs(residuals)		0.014***
		(0.002)
Observations	18,636	952,141
R^2	0.123	0.729
Adjusted R^2	0.121	0.729
Residual Std. Error	4,167.577 (df = 18589)	1.139 (df = 952047)
F Statistic	56.932*** (df = 46; 18589)	27,558.120*** (df = 93; 952047)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 6: Regressions including residuals of Gallup poll

	Dependent variable:	
	trader.trade.no	aggressor.index
	(1)	(2)
Constant	−28,209.650***	0.418***
	(1,487.174)	(0.010)
abs(residuals)	263.086***	0.003***
	(70.310)	(0.0005)
Observations	1,904,282	1,904,282
R^2	0.096	0.006
Adjusted R^2	0.096	0.006
Residual Std. Error (df = 1904236)	40,430.210	0.281
F Statistic (df = 45; 1904236)	4,475.553***	256.175***
Note:		*p<0.1; **p<0.05; ***p<0.01

References

A. Brown, D. Rambaccussing, J.J. Reade, and G. Rossi. Using Social Media to Identify Market Inefficiencies: Evidence from Twitter and Betfair. *Economic Inquiry*, 2017. URL http://onlinelibrary.wiley.com/doi/10.1111/ecin.12506/abstract.

Y. Chen, C. Chu, Mullen T., and D. M. Pennock. Information markets vs. opinion polls: An empirical comparison. *Proc. of the 6th ACM Conference on Electronic Commerce (EC)*, pages 58–67, 2005.

K. Croxson and J.J. Reade. Information and Efficiency: Goal Arrival in Soccer Betting. *Economic Journal*, 124:62–91, March 2014.

J. B. De Long, A. Shleifer, L. H. Summers, and R. J. Waldmann. Positive Feedback Investment Strategies and Destabilizing Rational Speculation. *Journal of Finance*, 45:379–395, 1990.

F. D'Amuri and J. Marcucci. The predictive power of Google searches in forecasting US unemployment. *International Journal of Forecasting*, 33(4):801–816, 2017.

F. Galton. Vox Populi. *Nature*, 1949:450–451, 1907.

A. Graefe, J.S. Armstrong, R.J. Jones Jr, and A.G. Cuzán. Combining forecasts: An application to elections. *International Journal of Forecasting*, 30(1):43–54, 2014.

R. Hanson and R. Oprea. A Manipulator can Aid Prediction Market Accuracy. *Economica*, 76(302): 304–314, 2009.

D.F. Hendry and F. Pretis. Anthropogenic Influences on Atmospheric CO2. Economics Series Working Papers 584, University of Oxford, Department of Economics, December 2011.

D.F. Hendry, S. Johansen, and C. Santos. Automatic Selection of Indicators in a Fully Saturated Regression. *Computational Statistics*, 23(2):337–339, April 2008.

D.F. Hendry, J.A. Doornik, and F. Pretis. Step-indicator Saturation. Economics Series Working Papers 658, University of Oxford, Department of Economics, June 2013.

M. Huberty. Can we vote with our tweet? On the perennial difficulty of election forecasting with social media. *International Journal of Forecasting*, 31(3):992–1007, 2015.

S.J. Johansen and B. Nielsen. An Analysis of the Indicator Saturation Estimator As a Robust Regression Estimator. In J.L. Castle and N. Shephard, editors, *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, Oxford and New York, 2009. Oxford University Press.

A. Leigh and J. Wolfers. Competing Approaches to Forecasting Elections: Economic Models, Opinion Polling and Prediction Markets. *The Economic Record*, 82(258):325–340, 2006.

J.A. Mincer and V. Zarnowitz. The evaluation of economic forecasts. In *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, pages 1–46. NBER, 1969.

L. Page. 'It ain't over till it's over.' Yogi Berra bias on prediction markets. *Applied Economics*, 44(1): 81–92, 2012.

L. Page and R.T. Clemen. Do prediction markets produce well-calibrated probability forecasts? *Economic Journal*, 123(568):491–513, 2013.

T. Peeters. Testing the Wisdom of Crowds in the field: Transfermarkt valuations and international soccer results. *International Journal of Forecasting*, 34(1):17–29, 2018.

F. Pretis, J.J. Reade, and G. Sucarrat. *gets: General-to-Specific (GETS) Modelling and Indicator Saturation Methods*, 2017. URL `http://CRAN.R-project.org/package=gets`. R package version 0.12.

J. J. Reade. Information and predictability: Bookmakers, prediction markets and tipsters as forecasters. *The Journal of Prediction Markets*, 8(1):43–76, 2014.

D. Rothschild. Combining forecasts for elections: Accurate, relevant, and timely. *International Journal of Forecasting*, 31(3):952–964, 2015.

D.M. Rothschild and R. Sethi. Trading Strategies and Market Microstructure: Evidence from a Prediction Market. *The Journal of Prediction Markets*, 10(1), 2016.

L. Sjöberg. Are All Crowds Equally Wise? A Comparison of Political Election Forecasts by Experts and the Public. *Journal of Forecasting*, 28(1):1–18, 2009.

J. Surowiecki. *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*. Little, Brown, 2004. ISBN 0-316-86173-1.

L. Vaughan Williams and J.J. Reade. Forecasting Elections. *Journal of Forecasting*, 35(4):308–328, 2016a.

L. Vaughan Williams and J.J. Reade. Prediction Markets, Social Media and Information Efficiency. *Kyklos*, 69(3):518–556, August 2016b.

W. Wang, D. Rothschild, S. Goel, and A. Gelman. Forecasting elections with non-representative polls. *International Journal of Forecasting*, 31(3):980–991, 2015.