UNIVERSITY OF CAMBRIDGE
Faculty of Economics

Institute for
New Economic Thinking

Cambridge-INET Institute

THE BEHAVIOUR OF BETTING AND CURRENCY MARKETS
ON THE NIGHT OF THE EU REFERENDUM

Tom Auld

(University of Cambridge)

Oliver Linton

(University of Cambridge)

We study the behaviour of the Betfair betting market and the sterling/dollar exchange rate (futures price) during 24 June 2016, the night of the EU referendum. We investigate how the two markets responded to the announcement of the voting results. We employ a Bayesian updating methodology to update prior opinion about the likelihood of the final outcome of the vote. We then relate the voting model to the real time evolution of the market determined prices. We find that although both markets appear to be inefficient in absorbing the new information contained in vote outcomes, the betting market is apparently less inefficient than the FX market. The different rates of convergence to fundamental value between the two markets leads to highly profitable arbitrage opportunities.

The behaviour of betting and currency markets on the night of the EU referendum[*]

Tom Auld[†], Oliver Linton

31 October, 2017

Faculty of Economics, University Of Cambridge, The Austin Robinson Building, Sidgwick Ave, Cambridge, CB3 9DD, UK

Abstract

We study the behaviour of the Betfair betting market and the sterling/dollar exchange rate (futures price) during 24 June 2016, the night of the EU referendum. We investigate how the two markets responded to the announcement of the voting results. We employ a Bayesian updating methodology to update prior opinion about the likelihood of the final outcome of the vote. We then relate the voting model to the real time evolution of the market determined prices as results are announced. We find that although both markets appear to be inefficient in absorbing the new information contained in vote outcomes, the betting market is apparently less inefficient than the FX market. The different rates of convergence to fundamental value between the two markets leads to highly profitable arbitrage opportunities.

Keywords: EU Referendum, Prediction Markets, Machine Learning, Efficient Markets Hypothesis, Pairs Trading, Cointegration, Bayesian Methods, Exchange Rates

1 Introduction

Were currency and prediction markets efficient overnight on 24 June 2016 as the results of the United Kingdom European Union membership referendum were announced?

This question is important as the EU referendum was one of the great political shocks of 2016. The results of the vote itself provide for a unique period in market history for which both financial and prediction market efficiency can be studied. The night is a special event for a number of reasons. Firstly, referendums are rare events with no similar votes in history for market participants to base expectations[1]. There was also a strong prior belief that the UK would vote to remain in the European Union. This provided fertile ground for inefficiencies and behavioural biases to arise. Secondly, the EU referendum results was the only information to affect the market during the hours of the night. There were 382 different voting areas and results were announced and widely distributed at different times. This represented a drip feeding of information to the market for a period of a few hours. Thirdly, there are 2 markets to study, a prediction market in the Betfair betting market and the pound dollar currency market.

The Efficient Market Hypothesis (EMH) holds that financial markets immediately reflect all available information in prices. If this is true, investors cannot receive above market returns except by chance. The weakest form of the hypothesis relates only to historical price information. Opinion is split on whether this form holds. Stronger forms of the EMH relating to both fundamental (semi-strong) and private information (strong) also exist. Most studies conclude that the stronger forms of the EMH do not hold. For the night of the EU referendum, one existing working paper (Wu et al. (2017)) concludes that the pound market was slow to reflect the information contained in the vote results and hence the EMH in the semi-strong form did not hold. Regarding prediction markets, there is a consensus in the literature that prediction markets provide better estimates of future events than experts, and that the predictions of such markets are useful in a variety of situations.

This paper makes a number of contributions: This is the first high frequency study in the literature comparing a prediction market with a financial market. We agree with Wu et al. (2017) that the EMH in semi-strong form

---

[1]There have only been two other UK wide referendums. The first, the European Communities membership referendum held in 1975, would be of little use for inferring voting patterns today. The other, on an unrelated subject, was the Alternative Vote Referendum in 2011 and had a turnout of only 42.2%, opposed to a typical figure of $60-70\%$ for general elections

did not hold in the currency market during the night of the referendum, but we demonstrate this was also the case for the prediction market. Further, we show that the Betfair market was more efficient than the sterling futures market, which provides some support for the view that prediction markets yield useful predictions. Small sample inference is required to predict Brexit early on in the night of the vote and we improve upon earlier prediction methods by using a rigorous Bayesian approach that is valid for small samples. Finally, we demonstrate that the different rates at which the vote was reflected in the two markets led to arbitrage opportunities, implying a failure of the EMH in the weak form.

The remaining sections of this paper are organized as follows. The next section reviews the literature. Section 3 discusses the data we use. In Section 4 we present the electoral model updating methodology, which employs Bayesian Machine Learning. Section 5 presents a theoretical model linking the behaviour of the two asset prices (Betfair contracts and Sterling futures) under standard economic assumptions. In Section 6 we present our empirical results. Section 7 concludes.

2 Literature Review

This is a study of market efficiency, and its manifestation in a prediction and a financial market on the night of the EU referendum. We summarize the debate concerning the Efficient Market Hypothesis (EMH) and refer to two studies of referendums recently held in the UK. However a comprehensive review of the field is beyond the scope of this paper. The reader is instead referred to Horn et al. (2014) for a review of prediction markets and Graefe (2016) for a specific review of political markets. There is a growing consensus in the literature that political prediction markets are more accurate at forecasting elections than are polls or experts.

The EMH (Fama (1965)) states that the prices of financial assets immediately discount all available information and therefore investors cannot make above average returns, except by chance. There are various forms of the hypothesis. In the Weak form, financial prices instantaneously discount all *market* information; in the Semi-Strong form, prices instantaneously discount all *publicly available* information; in the Strong form, prices instantaneously discount all information *both public and private*, including privileged information available to insiders. Many authors (eg Malkiel (2003)) argue that the EMH does not imply that pricing is perfect or that mispricings never occur, just that mispricings are random and it is not possible to systematically profit from them in advance.

There are various behavioural explanations that attempt to explain why the EMH may not hold. For a recent comprehensive review see Huang et al. (2016). One such theory is that of investor inattention, which is potentially relevant to us as results were announced outside of major market times. See Hirshleifer et al. (2013), Hirshleifer et al. (2009) and Hou et al. (2009) for examples, as well as DellaVigna & Pollet (2009) where the authors claim to show that earnings announcements on a Friday take longer for the market to react to. Other behavioural explanations include anchoring and systematic overconfidence. Another idea presented in Caballero & Simsek (2016) postulates that a study of any anomalies of the EMH require an analysis of the presence or absence of any arbitrage process that may exist to bring prices rapidly back to the "correct" value. There have been many opinions and studies published on the EMH and no consensus exists as to its validity. For a recent review, see GabrielaTitan (2015).

There are many studies of betting markets and their ability to predict future events including elections (Wolfers & Zitzewitz (2004)). A number of studies have evaluated the efficiency of sports betting markets, mostly finding some inefficiency (Goddard & Asimakopoulos (2004); Vlastakis et al. (2009); Badarinathi & Kochman (1996)). There are relatively few studies of referendums because they are unique events and require event-specific approaches. During the preparation of this paper a related study was published (Wu et al. (2017)) that investigated the real time response of the exchange rate to the announced vote outcomes. Their conclusion was that the "Brexit result could have been predicted with high confidence under realistic conditions". Examining social and psychological factors as well as Betfair data prior to the vote, the authors conclude that the mispricing "indicates both generic inefficiency and a specific inertia / durable bias in the market similar to herding during bubbles". The paper also examines trading behaviour in the pound around the announcement of specific results. We agree with the conclusion of a mispricing in the pound market. However, there are some shortcomings to the methodology. This concerns the use of OLS estimators with small number of observations. These considerations are explored formally in B. We improve upon this work by using a Bayesian electoral model valid for small sample inference, by using real-time Betfair price data from the night of the vote and introducing a theoretical model to relate the prices of the two markets. One other study that uses political Betfair data is Wall et al. (2017). This study does not examine the efficiency of the betting market but instead relies on the largely accepted premise that prediction markets can provide meaningful forecasts of outcomes. Betfair data is used to control for polling shocks and isolate campaign event effects in the 2014 Scottish Independence campaign in the months leading up to the referendum.

3   Data

We first describe the data used:

- EU Referendum Results: sourced from the Electoral Commission (EC) website (Electoral Commission (2016)).

- Timing Results: we use the earliest confirmed time for each voting area from the three sources below:

  - Press Association: the time, to the nearest second, that each result was received and processed at the Association was used. This was a small number of seconds after it was publicly announced at the count.

  - Electoral Commission: returning officers for each area both inform the EC both just prior to announcement and immediately afterwards. The EC has made both times for each area available to us to the nearest second. This data set provides a window within which each area's result must have been made public.

  - Bloomberg: 67 of the 382 results were published in real time on the Bloomberg terminal and the timestamps to the nearest second have been retrieved. Although only a small subset of the total, these results can be assumed to contain the most market-sensitive, and hence most informative, information.

- Priors for each voting area: Before the vote the psephologist, C. Hanretty of the University of East Anglia, published a blog titled "The EU referendum: what to expect on the night" (Hanretty (2016)). Expectations and 90% confidence intervals (CIs) for vote share based upon publicly available data could be downloaded for all but 4 of the 382 voting areas[2]. The work was reported in the media, including on the popular financial platform Bloomberg. It is reasonable to believe that market participants were aware of this information and had ready access to it. The priors were based on a panel data analysis of the British Electoral Study (BES) from 2015 and demographic results at the local authority level. Firstly the priors are calculated directly from the BES, and secondly, a uniform swing is applied to each area to bring the results in line with polling information available on the date of Hanretty's publication (7 June). Various assumptions were used to generate the data set. Hanretty characterises them as follows:

---

[2]Northern Ireland, the Isles of Scilly, Anglesey and Gibraltar were excluded due to the lack of availability of local authority demographic data.

- the geographic patterns of the Leave and Remain votes have remained constant since May of [2015] (though the levels may have changed)
- that undecideds will break in roughly even numbers between the two camps
- Leave- and Remain-voting areas will vote at roughly equal rates.

- We make particular use of YouGov's poll on the day (YouGov (2016)). We also use other polls where there are gaps in Hanretty's priors.

- Historical General Election Data: we make extensive use of constituency level data for preceding general elections from the EC (Electoral Commission (2017)) and the website Electoral Calculus (Baxter (2017)). We also use historical polling information and measures of their accuracy for general elections. This was obtained from Wikipedia (Wikipedia (2017b)).

- GBPUSD Futures Price: we decided to use the GBPUSD future price traded on the Chicago Mercantile Exchange rather than the spot price. There are multiple exchanges where the spot trades and aggregation could be prohibitively difficult. It is well known that spot and futures prices for foreign exchange are extremely well correlated and are effectively contemporaneous on time scales of under one second when both markets are open. The data was downloaded from Bloomberg and timestamps of trades were reported to an accuracy of one second. Note that the futures contract was closed between 10 pm and 11 pm on 23 June which was before the announcement of results.

- Betfair Data: The betting website Betfair listed two contracts. These were traded on Betfair's exchange platform which acts as a limit order book. The first paid out £1 in the event of Brexit, the other paid £1 in the event of Remain. The sum of the prices of the contracts did not deviate sufficiently from £1 to enable a profitable arbitrage. Betfair supplied all trades with timestamps of one second granularity in both contracts between 10 pm on 23 June to 5 am on 24 June. We convert all prices in the Remain contract to a synthetic price in the Brexit one by subtracting from £1. The resulting Betfair data set, along with the GBPUSD future price, is shown in Figure 1. We conduct analysis on the combined set of trades, which number 182,534. £51,016,907 in total was matched during this 7 hour window. This compares with 88,246 trades in the GBP future during this time with a total notional traded
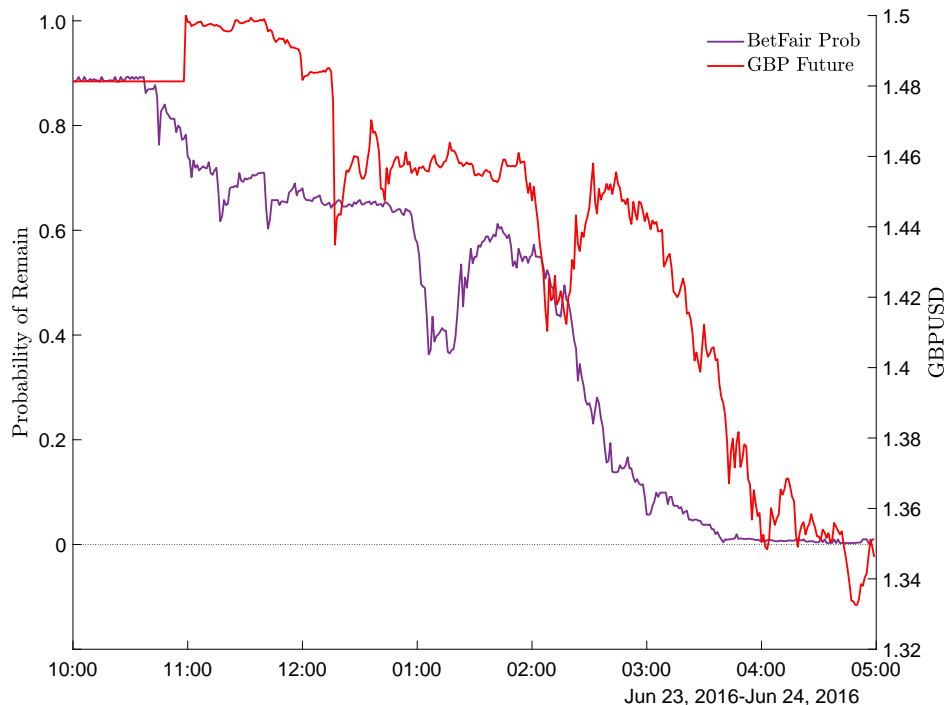
6

Figure 1: The Pound and Betfair Markets Overnight for the Period of the Announcement of Results.

of around \$5.5Bn. Although the futures market is considerably larger in notional traded terms than the betting market, the Betfair contracts moved by around 90% of their price whereas the pound moved around 10%.

4   Electoral Model

In this section we present a Bayesian model for calculating an implied probability of Brexit throughout the night. A summary of the model is presented in Appendix A.

4.1   Setup

There are n constituencies with fixed sizes which we label $s_1, s_2, \ldots, s_n \in \mathbb{N}$. Suppose $p_i, q_i \in [0, 1]$ are the proportion of voters in favour of leaving the

EU and the turnout percentage in constituency i, ordered by time. Then the proportion of the national vote is:

$$p_N = \frac{\sum_{i=1}^n p_i q_i s_i}{\sum_{i=1}^n q_i s_i} \tag{4.1}$$

and the event of leaving the EU occurs when $p_N > \frac{1}{2}$.

4.2 Gaussian Copula Prior

The variables in the model are the $2n$ values of vote share and turnout of the vector $r = (p_1, \ldots, p_n, q_1, \ldots, q_n)^\top$, ordered chronologically. We use a Gaussian copula prior. Given marginal CDFs for the unknown variables $F_1(r_1), \ldots, F_{2n}(r_{2n})$ the joint CDF is:

$$\Pr(r_1 < t_1, \ldots r_{2n} < t_{2n}) = \Phi_{\Sigma_0}(\Phi^{-1}(F_1(t_1)), \ldots, \Phi^{-1}(F_{2n}(t_{2n})))$$

where Φ_Σ is the CDF of a multivariate normal distribution with zero mean and correlation matrix Σ. We model the dependence between the variables as via two-factors being vote share and turnout, with correlations between factors allowed so that:

$$\Sigma_0 = \begin{pmatrix} \Sigma_p & \Sigma_{pq} \\ \Sigma_{pq}^T & \Sigma_q \end{pmatrix},$$

where

$$\begin{aligned}
\Sigma_p &= (1 - \rho_p)\, I_n + (\rho_p)\, i_n i_n^\top \\
\Sigma_q &= (1 - \rho_q)\, I_n + (\rho_q)\, i_n i_n^\top \\
\Sigma_{pq} &= \rho_{pq} \times \left[(1 - (\rho_q \rho_p))\, I_n + (\rho_q \rho_p) i_n i_n^\top\right].
\end{aligned}$$

The correlation matrices Σ_p, Σ_q, and Σ_{pq} are all of size $n \times n$ and represent the dependence between area voteshares $(p_1, \ldots, p_n)$, turnouts $(q_1, \ldots, q_n)^\top$ and correlation between voteshare and turnouts, respectively. The relevant correlations are the same between areas and are ρ_p, ρ_q and $\rho_{pq} \in (-1, 1)$.

4.3 Prior Probability of Brexit

Given Σ_0 and marginal distributions $F_1(r_1), \ldots, F_{2n}(r_{2n})$ an expression can be found for the prior probability of Brexit (see appendix A.2). There is no analytical form for the probability integral involved so a sampling method is required for evaluation.

4.4 Marginal Prior Distributions

Hanretty provided expectations and 90% CIs for the marginals. These CIs are implied by responses of the BES coupled with Local Authority demographic data. They do not take into account the uncertainty of the national vote. There may be an argument that the distributions that Hanretty supplies as part of his analysis of panel and local authority data are asymptotically normal. No such argument can be made once the uncertainty of the national vote share is taken into account. We interpret the prior as an expression of a degree of belief about the possible values that vote shares could take. As such, we are not constrained by the normal distribution.

4.4.1 Marginal Calibration

To calculate the expectations of the marginals, suppose that μ_H is the vector of expectations provided by Hanretty. Given an expected level for the national average vote share μ_N, then expectations for each area i, denoted μ_{p_i} can be formed by applying a fixed uniform shift, α_N to μ_{H_i}, where μ_{H_i} is the expectation provided by Hanretty.[3]

We calibrate the marginal distributions by targeting a subjective level of the prior variance of the national vote share σ_p^2 equal to a generous estimate of what we think it could be, say $\hat{\sigma}_p^2$. Either way, we add a constant variance σ_N^2 to each marginal variance and adjust σ_N^2 to achieve the result. We convert the CIs as if the marginal distribution were normal:

$$(\sigma_H^2)_{p_i} = ((90\% \text{ Confidence Interval})_i/(2 \times 1.645))^2,$$

$(\sigma_H^2)_i$ is now the unadjusted variance implied for the i'th voting area. The national vote share is uncertain and treating this variation as independent of the idiosyncratic variances implied by Hanretty's values leads us to add to each area variance a constant variance, say σ_N^2.[4]

$$\sigma_{p_i}^2 = (\sigma_H^2)_i + \sigma_N^2$$

---

[3]Note, Hanretty himself forms μ by applying a uniform shift to the priors he calculates from the BIS and census data to agree with polling data at the time of his publication

[4]We are not implying that area vote shares are independent of the national vote share, just that the idiosyncratic variation implied by Hanretty's study of survey respondents and local authority data is independent of the variation of the national vote share number. Our assumptions are not even that strong; as we apply a uniform shift to variances we are simply implying that the *difference* in the variations of the marginal distributions of individual areas is the same as the difference implied by Hanretty's study.

The prior mean for p_N is below 50% so the part of the national vote share distribution that lies above 50%, which is the model probability of Brexit, will be greater given higher variance σ_N^2.

For turnout we base area level expectations, say ν, on historical general elections. However we choose the variances for each area to be equal and labelled by σ_ν^2.

We consider the marginals distributions of the following types: Normal, Logit-Normal, Beta and Logit Student with Location and Scale[5]. In each case the hyperparameters of the marginal distributions are set so that the mean and variance are equal to μ_{p_i} and $\sigma_{p_i}^2$ respectively, for each area i.

4.5 Update

Calculation of the conditional distribution is most easily done by a re-ordering of the variables,

$$\tilde{r} = (p_1, q_1, \ldots, p_n, q_n)^\intercal.$$

Then

$$\Phi^{-1}(\tilde{F}_1(\tilde{r}_1)), \ldots, \Phi^{-1}(\tilde{F}_{2n}(\tilde{r}_{2n})) \sim N(0, \tilde{\Sigma}_0),$$

where for generic x, $\tilde{x}$ indicates a similar re-ordering of the rows and columns of x. To calculate the conditional distribution of the remaining variables after m results have been announced, partition $\tilde{\Sigma}_0$ into four block matrices as follows:

$$\tilde{\Sigma}_0 = \left(\begin{array}{cc} \tilde{\Sigma}_{m,m} & \tilde{\Sigma}_{m,\backslash m} \\ \tilde{\Sigma}_{\backslash m,m} & \tilde{\Sigma}_{\backslash m,\backslash m} \end{array} \right),$$

where: $\tilde{\Sigma}_{m,m}$, $\tilde{\Sigma}_{m,\backslash m}$, $\tilde{\Sigma}_{\backslash m,m}$ and $\tilde{\Sigma}_{\backslash m,\backslash m}$, are $2m \times 2m$, $2m \times 2(n-m)$, $2(n-m) \times 2m$ and $2(n-m) \times 2(n-m)$ matrices respectively. The multivariate conditional Gaussian copula is also a Gaussian copula. Given the observations $p_1, q_1, \ldots, p_m, q_m$, write the data as:

$$\tilde{x}_m = \left(\Phi^{-1}(\tilde{F}_1(\tilde{r}_1)), \ldots, \Phi^{-1}(\tilde{F}_{2m}(\tilde{r}_{2m})) \right).$$

Then

$$\Phi^{-1}(\tilde{F}_{2m+1}(\tilde{r}_{2m+1})), \ldots, \Phi^{-1}(\tilde{F}_{2n}(\tilde{r}_{2n})) \,|\, \tilde{r}_m \sim N(\tilde{\Pi}_{\backslash m}, \tilde{\Sigma}_{\backslash m})$$

---

[5]For the logit-student distribution the degrees of freedom parameters ν was set to 3. This has fatter tails and $logit(r_i)$ will have infinite fourth moment.

where:

$$\tilde{\Pi}_{\backslash m} = \tilde{\Sigma}_{\backslash m,m}\tilde{\Sigma}_{m,m}^{-1}\tilde{x}_m$$
$$\tilde{\Sigma}_{\backslash m} = \tilde{\Sigma}_{\backslash m,\backslash m} - \tilde{\Sigma}_{\backslash m,m}\tilde{\Sigma}_{m,m}^{-1}\tilde{\Sigma}_{m,\backslash m}.$$

As $p_1, \ldots, p_m, q_1, \ldots, q_m$ is now known, the model probability of Brexit can be computed via:

$$\Pr(\text{BREXIT})_m =$$
$$\Pr\left(\sum_{i>m}\left[p_i - \frac{1}{2}\right]q_i s_i > \sum_{i\leq m}\left[\frac{1}{2} - p_i\right]q_i s_i \;\middle|\; p_1, q_1, \ldots, p_m, q_m\right). \quad (4.2)$$

4.6   Model Properties

An advantage of the model is that it provides for closed form updates to the posterior distributions of the parameters, as the Gaussian copula has a conditional distribution. This avoids the need for a Markov Chain Monte Carlo sampling technique to calculate the integral in equation (4.2). This would be particularly arduous given the large number (764) of variables involved. An alternative copula, with well understand and closed form conditional distributions, is the Student's t copula, Ding (2016). The conditional distribution is also a Student's t copula and will have fatter joint tails than the Gaussian.

The model describes the dependence structure of the unknowns in a parsimonious way with a relatively few parameters. This leads to limitations. A better description of the correlations of the variables could be found by using more factors, such as the demographic ones used by Hanretty to calculate the published marginal statistics. Heterogeneous correlation coefficients would likely follow. However, we postulate that our model will capture the main swing to the leave vote. We will perform robustness tests using some limited forms of heterogeneous parameters to test this hypothesis.

We now comment on the expected qualitative impact of the model as parameters change. For purely independent vote share results ($\rho_p, \rho_q = 0$), convergence will solely be due to the results as they come in and the distributions of the yet to be announced results will not be affected. For higher values of ρ_p, convergence will be faster. It is expected that the value of ρ_p and the variance σ_N^2 will have the greatest effect on the speed at which predictions change. The effect of ρ_q and ρ_{pq} on the model probability are effectively second order. Given that we will be setting ρ_{pq} as negative and turnout was above expectations, there will be some small second order effects

from changes in the other parameters. Lowering ν or σ_v^2 will slow the speed of convergence whereas lower ρ_q and lower $|\rho_{pq}|$ will speed convergence, but these effects should be very small.

4.7 Parameter Choices

There will be a degree of subjectivity involved in the setting of the hyperparameters of the prior, particularly as the referendum was a one off event with little historical precedent. We use both general election and polling data to inform our choice of parameter values.

4.7.1 Turnout

National Turnout There were reports of high turnout on the day of the vote itself (Gutteridge (2016)). We will use national turnout for general elections as a guide, but note that the Scottish independence referendum had an unprecedentedly high turnout of 85%. The general election turnout figures since 1945 are shown in Table 1. The average is 66.9% (6.7%) and for the last three elections the average is 64.2%. We use 67.6% which is the three-election average weighted upwards by half the six-election standard deviation. A reasonable range of expectations would be $65 - 70\%$.

Area Turnout Voting regions for the EU referendum were not the same as the constituencies used for general elections. However, the EC categorizes both the 381 voting areas in the referendum (excluding Gibraltar) and the (most recently 650) general election constituencies by 12 region codes. This enables us to make a more granular estimate of turnout per area v_i than simply assuming a uniform expectation. We use average turnout for each region for the 2010 and 2015 general elections as outlined in Table 2. Similar to the means of the expected vote share per region, $(\nu_1, \ldots, \nu_n)$ can be uniformly shifted to achieve the required expected national turnout.

Turnout Variance Instead of setting turnout variance by region we will simply use the same level for every area and use the standard deviation figure for the last six general elections; 6.7%.

4.7.2 Turnout Correlation by Area

Given elections in time periods $t = 1, \ldots, T$ and turnouts q_{it}, if we have predictions in advance for q_{it}, $\bar{q}_{it}$, then we can model the prediction errors

Election Year	England	Wales	Scotland	N. Ireland	UK
2015	65.8%	65.7%	71.1%	58.1%	66.1%
2010	65.5%	64.7%	63.8%	57.6%	65.1%
2005	61.3%	62.6%	60.8%	62.9%	61.4%
2001	59.2%	61.6%	58.2%	68%	59.4%
1997	71.4%	73.5%	71.3%	67.1%	71.4%
1992	78%	79.7%	75.5%	69.8%	77.7%
Standard Deviation UK, 1992 - 2015					**6.7%**
1987	75.4%	78.9%	75.1%	67%	75.3%
1983	72.5%	76.1%	72.7%	72.9%	72.7%
1979	75.9%	79.4%	76.8%	67.7%	76%
1974 Feb	79%	80%	79%	69.9%	78.8%
1974 Oct	72.6%	76.6%	74.8%	67.7%	72.8%
1970	71.4%	77.4%	74.1%	76.6%	72%
1966	75.9%	79%	76%	66.1%	75.8%
1964	77%	80.1%	77.6%	71.7%	77.1%
1959	78.9%	82.6%	78.1%	65.9%	78.7%
1955	76.9%	79.6%	75.1%	74.1%	76.8%
1951	82.7%	84.4%	81.2%	79.9%	82.6%
1950	84.4%	84.8%	80.9%	77.4%	83.9%
1945	73.4%	75.7%	69%	67.4%	72.8%

Table 1: Historical UK General Election Turnout.

Region	2015 Turnout	2010 Turnout	Average Turnout
East	67.5%	67.6%	67.6%
East Midlands	66.5%	66.8%	66.6%
London	65.4%	64.5%	64.9%
North East	61.8%	61.1%	61.4%
North West	64.3%	62.3%	63.3%
Northern Ireland	58.1%	57.6%	57.8%
Scotland	71.0%	63.8%	67.4%
South East	68.6%	68.2%	68.4%
South West	69.5%	69.0%	69.2%
Wales	65.7%	64.8%	65.2%
West Midlands	64.1%	64.7%	64.4%
Yorkshire and The Humber	63.3%	62.9%	63.1%

Table 2: Turnout per EC Region in 2010 and 2015.

$\triangle q_{it} = q_{it} - \bar{q}_{it}$ as being due to a national error η_t and individual error terms ϵ_{it} i.e,

$$\triangle q_{it} = \eta_t + \epsilon_{it}, \quad \eta_t \sim N(0, \sigma_\eta^2), \quad \epsilon_{it} \sim N(0, \sigma_\epsilon^2), \quad cov(\epsilon_t, \eta_{it}) = 0$$

Then for $i \neq j$, ρ_q is given by

$$\rho_q = corr(\triangle q_{it}, \triangle q_{jt}) = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\epsilon^2}, \qquad i \neq j.$$

A regression of 2015 constituency turnout on 2010 turnout yields a coefficient of determination of 0.734 which provides evidence for simply using the turnout of the last election as the prediction. We do so. σ_η^2 is simply the variance of the national turnout $(6.7\%)^2$. Calculation of σ_ϵ^2 requires looking at errors at the constituency level for each separate election. As there was the fifth constituency boundary review in 2008 we can form no easy prediction for area turnout for the 2010 election because constituencies changed. We simply use the 2015 election to estimate σ_ϵ^2 with predictions provided by the 2010 election. This results in a estimate of 3.0% for σ_ϵ and one for ρ_q of 0.835.[6] In the absence of any other estimate or information pertinent to likely voting habits, this is what we use.

4.8 Vote Share

4.8.1 Area Vote Share

An expected national vote share μ_N is required. Polls with samples in the week preceding the referendum are shown in Table 3, along with seven polls of polls. For general elections, exit polls measure how people declare they have voted on the day itself at a selection of particular, secret, polling stations. They are much more accurate than any pre-election polling (Curtice et al. (2011)), due to the fact that there is no measurement error of respondents. There was no exit poll for the referendum as it was a one-off election. There was, however, a poll on the day conducted by YouGov which was published shortly after voting closed at 10 pm YouGov (2016). This poll measured how people voted versus how those same individuals reported their voting intention the preceding day. The result was a demographically weighted result of 48.38% which was broadly in line with recent polls. As we consider this to be the most accurate poll we set $\mu_N = 48.38\%$.

---

[6]Estimates based on sample moments are consistent due to the Law of Large Numbers.

Date(s)	Remain	Leave	Undecided	Remain Lead	Organisation
22-Jun	55%	45%	--	10%	Populus
20–22 Jun	51%	49%	--	2%	YouGov
20–22 Jun	49%	46%	1%	3%	Ipsos MORI
20–22 Jun	44%	45%	9%	1%	Opinium
17–22 Jun	54%	46%	--	8%	ComRes
17–22 Jun	48%	42%	11%	6%	ComRes
16–22 Jun	41%	43%	16%	2%	TNS
20-Jun	45%	44%	11%	1%	Survation/IG Group
18–19 Jun	42%	44%	13%	2%	YouGov
16–19 Jun	53%	46%	2%	7%	ORB/Telegraph
17–18 Jun	45%	42%	13%	3%	Survation
			Polls of Polls		
23-Jun	52%	48%	--	4%	What UK Thinks: EU
23-Jun	50.6%	49.4%	--	1.2%	Elections Etc.
23-Jun	45.8%	45.3%	9%	0.5%	HuffPost Pollster
22-Jun	46%	44%	10%	2%	Number Cruncher Politics
23-Jun	48%	46%	6%	2%	Financial Times
22-Jun	51%	49%	--	2%	The Telegraph
23-Jun	44%	44%	9%	0%	The Economist
				2.0%	**Average Poll of Polls**

Table 3: Opinion Polling Prior to the EU Referendum. Source Wikipedia (2017a).

4.8.2 Variance of Area Vote Share

Variances are chosen by shifting those implied by Hanretty by a constant amount σ_N^2 so that a generous estimate of the variance of the national vote share σ_p^2 results. Table 4 shows that the average error in opinion polls from the prior week in the last six general elections was 2.66%. General election polling is a well-researched field with plenty of historical precedent and would provide far too confident a figure. As our aim is to produce a prediction based on a conservative prior, we set $\sigma_p = 5\%$.

4.8.3 Correlation between Voting Areas

We analyse general election data in a similar manner to section 4.7.2 to estimate ρ_p. We examine the variation at the constituency versus the national level of the Conservative party voteshare to inform our choice of ρ_p. Pre-

Election	Average Poll (Prior Week)	Result	Error
2015	34%	37.8%	-3.80%
2010	35%	36.9%	-1.58%
2005	31%	33.2%	-2.20%
2001	31%	31.7%	-0.92%
1997	30%	30.7%	-0.27%
1992	37%	41.9%	-4.46%
	$\hat{\sigma}_\epsilon$		**2.66%**

Table 4: Opinion Polls and Vote Share for the Conservatives for recent general elections.

dictions for the constituency level vote are based on applying the implied swing from opinion polls from the week prior to each election, to the level of the last election. See Table 4 for these polling results and for the results of the last six elections. Complications arise due to Westminster constituency boundary reviews in 1995, 2005 and 2008. These reviews change the number of constituencies and their composition of voters. They occur periodically in order to remove variations in the number of electors in each area, and have tended to favour the Conservatives (Rallings et al. (2008)). This is a well understood problem and the website Electoral Calculus (Baxter (2017)) publishes implied election results for elections preceding a review to enable ready comparison; we use these implied figures.

The implied standard error of σ_η using the data in Table 4 is 2.66%[7]. Relying on the last six elections, the constituency level error calculation yields an estimate of 4.18% for σ_ϵ, implying a correlation $\rho_p = 0.288$. However, our constituency level errors are probably estimated at too high a level as better predictions for constituency level results exist although we do not have ready access to them. For this reason, the value of σ_ϵ^2 is likely estimated too high and ρ_p too low. Consequently, this value of ρ_p will be treated as a lower bound.

The correlation ρ_p is likely to be the parameter with the largest effect on how quickly the model prediction will converge to the true result. The implied correlation coefficient, as estimated, appears to be stable. Using only the last three elections results in an estimate of 0.324. The largest estimated value of σ_ϵ (5.69%) in any single election for the last six was in 1997, which was (unsurprisingly) also the largest error in the national vote share. If we combine this with $\sigma_\eta = 2.66\%$, $\rho_p = 0.18$ results. This is an artificially low

---

[7]As the model implicitly assumes a mean of zero this is the square root of the average of the squares of the error, not the sample variance.
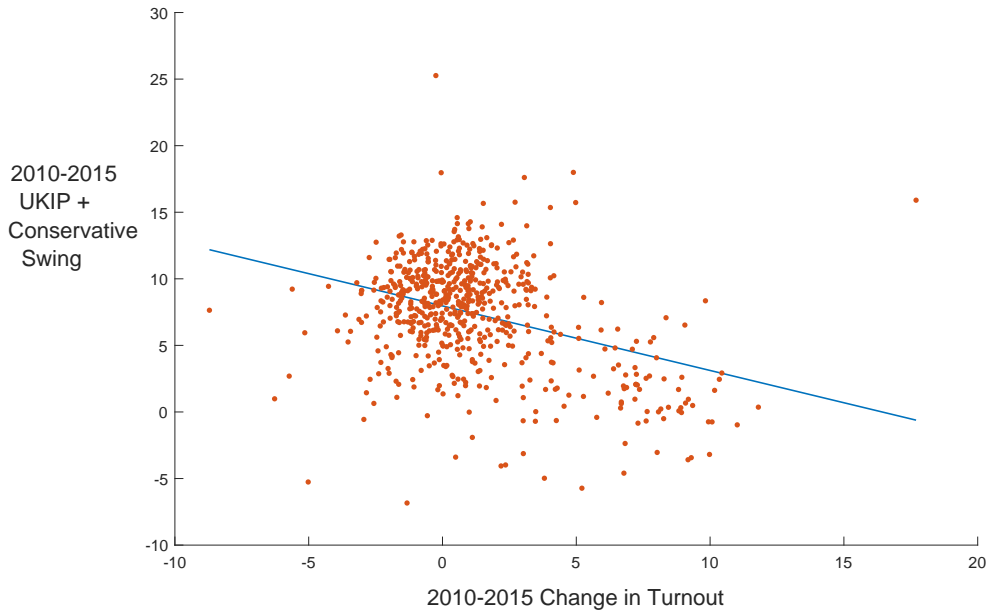
Figure 2: Change in UKIP + Conservative Vote Share versus Change in Constituency Turnout at the 2015 General Election.

estimate and parameter values below this level are highly unlikely.

4.9 Correlation between Vote and Turnout

There were conflicting reports concerning the probable impact of turnout, even within the same newspaper on the day of the result (Gutteridge (2016), Foster (2016)).

Predicting how unexpected turnout affects results requires a successful prediction of whether the difference in turnout is attributed to Leave or Remain supporters. This is a difficult problem. It was well understood in advance that younger voters, who are less likely to vote, would favour Remain, and that Brexit supporters were reported in surveys being more than twice as likely to vote as Remain ones Twyman (2016). YouGov were widely quoted as suggesting the relationship between turnout and Brexit vote share would be negative (see Foster (2016) for an example).

We take a quantitative approach based on general elections. Due to boundary changes we are restricted to studying only the 2015 general election as no implied turnouts for elections preceding a boundary review is available. We proxy support for Brexit at the 2015 general election by using

17

Parameter	Description	Value	Range
ν	National Turnout	67.6	$65-70\%$
σ_ν	Turnout Error	$(6.7\%)^2$	–
ρ_q	Area Turnout Correlation	0.835	–
μ_N	Expected National Vote Share	48.38	–
σ_p	National Vote Error	5%	–
ρ_p	Area Vote Correlation	≥ 0.288	≥ 0.18
ρ_{pq}	Area Vote and Turnout Correlation	-0.361	–

Table 5: Plausible parameter values.

combined vote shares of UKIP and the Conservatives (the parties with supporters most likely unsympathetic to the EU). We regress the swing of the combined UKIP and Conservatives vote share against the change in turnout at the constituency level for the 2010-2015 elections. Figure 2 shows the regression, which results in a statistically significant correlation coefficient of -0.361. This is indeed negative, agreeing with YouGov, and we use this value. However, we note that there may have been beliefs that this parameter may have been of different sign due to conflicting reports in the media[8].

4.10 A Note on Model Correlation and Measured Correlation

The model correlation parameters ρ_p, ρ_q and ρ_{pq} are not strictly the correlations of the random variables $r = (r_1, \ldots, r_{2n})$. They are the correlations of Φ^{-1} of the quantiles being $(\Phi^{-1}(F_1(r_1)), \ldots, \Phi^{-1}(F_{2n}(r_{2n})))$. When the marginal is normal, $\Phi^{-1}(F_i(r_i))$ will be a linear function of r_i and the correlations will be identical. When $\Phi^{-1}(F_i(\bullet))$ is non-linear, a simulation or other method would be strictly required to convert them. However, we omit this step as it is not significant in practice.

4.11 Missing Priors

Of the 382 voting areas of the Referendum, Hanretty failed to publish priors for four areas. These are listed in Table 6. The 4 areas are:

1. Gibraltar: This makes up a tiny 0.05% of the electorate, was the first

---

[8]The actual correlation observed on the night between voteshare surprised and turnout surprises at the area level was around-0.1. Although turnout was higher than expected, as was the vote for leaving the EU, at the area level, turnout was even higher on average, for area's with lower support for Brexit

area to announce, and had overwhelming support for Remain (Reyes (2016b,a)). As the population is so distinct from that of the rest of the UK, the result is not informative. We therefore take it as given and do not include it in the model.

2. The Isles of Scilly and Isle of Anglesey make up only 0.11% of the electorate and are simply ignored.

3. Northern Ireland consists of about 1.26m voters in a total electorate of roughly 46.5m. We use opinion polls for the mean and a standard deviation equal to that of the average of the other areas. We use a poll published on June 20 (Shapiro (2016)) that showed Remain 11% ahead, or 9% higher than the rest of the UK at that time. We therefore set the mean equal to $\mu_N - 9\%$.

Area	Declaration (Actual / Expected)	Electorate (%)
Gibraltar	23:36:33 / 00:01	24,119 (0.05%)
Isles of Scilly	00:49:42 / 00:01	1,799 (0.004%)
Isle of Anglesey	02:18:00 / 02:30	51,445 (0.11%)
Northern Ireland	04:37:00 / 04:00	1,260,955 (2.71%)

Table 6: The 4 Voting Areas with Missing Priors. Source Hanretty (2016).

5 Theoretical Framework

In this section we present a framework to analyse the efficiency of the betting and pound markets overnight on the Brexit vote.

Notation

p_N Vote share for leaving the EU

$\Pr(\text{Brexit})$ Generic probability of Brexit

$\mathbb{P}_{\text{Brexit}}$ True probability of Brexit

$P_{Betfair}$ Probability of Brexit implied by Betfair market

P_{GBP} Probability of Brexit implied by the GBP market

T Time at which decision of referendum announced

GBP_t Price of GBP at time $t < T$

$BETP_t$ Price of Betfair contract paying out £1 in the event of Brexit at time $t < T$

$u(\cdot)$ Bernoulli utility function

ϵ_t Stationary martingale process

5.1 Assumptions

We investigate the implications of various economic assumptions which we list in order of strength:

NS   No persistent shocks to GBP_t beyond those that affect Pr(Brexit)

EMHW   The Efficient Market Hypothesis holds in the weak form

EMHSS   The Efficient Market Hypothesis holds in the semi-strong form

CMI   Conditional Mean Independence of GBP_t from national vote share p_N given Brexit

RN   Risk Neutrality

Not all these assumptions are regarded as holding exactly but they are presented as an approximation. We discuss each one in detail below.

NS   Simply put, this assumption states that the only fundamental determinants of the GBP price on the night were the results of the vote, and that those results only affect the price through their effect on the probability of Brexit, Pr(Brexit). First we discuss the validity that the only determinant of GBP price during the hours under study were the referendum results. There are certainly other determinants of the GBP price over longer periods of time, for example information regarding the nature of future trading relationships that may became apparent after a negative vote. However on the night itself, there were no major economic releases, or other significant news events. In advance, the Econoday Economic Calendar (Econoday (2016)) listed the final market moving news releases on the 23 June as the US New Home Sales Report at 10:00 am Eastern Time (ET) and the first one for 24 June (beyond the referendum) as Durable Goods Orders at 10 am ET. They predicted that the following would be the market focus

for the 24th: "In a rare and potentially powerful wildcard, the markets will react to the Brexit outcome". This demonstrates that in advance there were beliefs that the main determinant of prices would be the outcome of the Referendum. Indeed, the authors of Wu et al. (2017) describe the circumstances as "a natural experiment" with "near perfect conditions" to study a situation. Next we discuss the assumption that only the likelihood of Brexit affects the GBP price. Before the vote there were several predictions that the pound would sell off significantly in the event of Brexit but rally a little otherwise (see Wu et al. (2017)). Thus, an assumption that GBP decreases monotonically with $Pr(\text{Brexit})$ is reasonable. This does not lead directly to the assumption that *only* $Pr(\text{Brexit})$ affects the GBP price. For instance, it could be believed that the vote share p_N has an affect on GBP over above the decision to leave the EU through the "hardness" of such a Brexit. This would still be consistent with the NS assumption as unless the support of the distribution of p_N were entirely above 50% ($P(\text{Brexit}) = 1$) any move in p_N would result in a, ceteris paribus, change of $Pr(\text{Brexit})$. This is as, as the expectation of p_N moves further from the 50% cutoff, the probability mass below 50% decreases and $P(\text{Brexit})$ decreases, at least for the distributions studied in this paper. Also, we postulate that only the binary result of the vote, and not the any particular form of trading relationship following a vote, was on the minds of investors on the night of the vote. Despite there being intense scrutiny of the negotiations between the EU and the UK on the terms of withdrawal in recent months, the term "hard" Brexit only first appeared several months after the referendum after the Conservative party conference in October 2016.

EMHW   When the EMH holds there can be an interpretation of a market probability for a prediction market, as all information has been aggregated into the price. The market price of the Betfair contract can be interpreted as $u^{-1}(P_{Betfair} \times u(\pounds 1))$. Correspondingly, $P_{Betfair,t} = u(BETP_t)/u(\pounds 1)$ where $BETP$ is the price. An implication of weak form market efficiency is that different markets will have identical beliefs. For the purposes of this study this means that although the two markets may have beliefs about the outcome of the referendum which do not agree with fundamental information contained in vote results, they must have the same beliefs and so $P_{Betfair} = P_{GBP}$.

EMHSS   EMH holding in the semi-strong form will imply that prices immediately discount all publicly available information including any results of voting areas already announced in the referendum. As such $P_{betfair} = P_{GBP} = \mathbb{P}_{\text{Brexit}}$, the true probability. For our purposes this will imply that market probabilities equal the probability of the electoral probability model. Note that EMHSS $\Rightarrow$ EMHW.

CMI   Let
$$D = \left\{ \begin{array}{ll} 1 & BREXIT \text{ at time } T \\ 0 & REMAIN \text{ at time } T \end{array} \right.$$

The CMI assumption can be written mathematically as:

$$\mathbb{E}_t(GBP_T|D, p_N) = \mathbb{E}_t(GBP_T|D),$$

where p_N is the national share of those voting for Brexit, while T is the time when the decision is announced and t is any time with $t < T$. Equivalently, the sterling rate is affected by p_N only through its affect on whether Brexit occurs. The GBPUSD price would be expected to be the same if the vote for Brexit were either 50.01% or 99.99%. This is a strong assumption, particularly when there is significant probability mass around outcomes very close $p_N = 50\%$.

RN   Again, this is a strong assumption which is not believed to hold in practice, but it is a useful approximation that is likely to be roughly valid. However, it is noted that any deviations from RN may have larger effects than they otherwise might have due to the increased risk of holding the pound during the period under study[9].

5.2   Implications

The implications of the assumptions are explored below and summarised in table 7.

5.2.1   Semi-Strong Market Efficiency

Under EHMSS:

---

[9]Higher perceptions of risk were evident from, for instance, higher implied volatility from options pricing as well as increased margin requirements from brokers for sterling related products.

Assumptions		Implication
EMHSS	$\Rightarrow$	$u(BETP_t) = \mathbb{P}_{\text{Brexit, t}} \times u(\pounds 1)$
EHMSS+RN	$\Rightarrow$	$BETP_t = \mathbb{P}_{\text{Brexit, t}}$
NS+EMHW	$\Rightarrow$	$GBP_t = g(1 - BETP_t) + \epsilon_t, \quad g(\cdot) \text{ increasing}$
NS+EMHW+CMI+RN	$\Rightarrow$	$GBP_t = GBP_L + (1 - BETP_t) \times (\triangle GBP) + \epsilon_t$

Table 7: Assumptions and their implications in the theoretical framework.

$$\mathbb{P}_{\text{Brexit, t}} = P_{Betfair,t} = \frac{u(BETP_t)}{u(\pounds 1)},$$

where t subscripts indicate evaluation at time t. The further assumption of RN implies:

$$\mathbb{P}_{\text{Brexit, t}} = BETP_t$$

5.2.2 Weak Market Efficient

NS implies

$$GBP_t = f(P_{GBP,t}) + \epsilon_t,$$

where $f(\cdot)$ is monotonically decreasing. Adding the EMHW assumption which implies $P_{Betfair} = P_{GBP}$ yields

$$
\begin{aligned}
GBP_t &= f(P_{Betfair,t}) + \epsilon_t \\
&= f(u(BETP_t)/u(\pounds 1)) + \epsilon_t \\
&= g(1 - BETP_t) + \epsilon_t,
\end{aligned}
\tag{5.1}
$$

where since $u(\cdot)$ is increasing and $f(\cdot)$ decreasing, $g(x) = f(u(1 - x))$ is increasing. Thus only NS and EMHW are required to imply that the pound price moves contemporaneously with some non-linear but monotonic function of the Betfair price.

5.2.3 Cointegration under Stronger Assumptions

Adding the stronger assumptions of CMI and RN will imply that for the form of $g(\cdot)$ in equation 5.1 is linear and the markets will be cointegrated:

$$
\begin{aligned}
GBP_t &= \mathbb{E}_t\left(GBP_T\right) + \epsilon_t \\
&= \mathbb{E}_t\left(GBP_T | D, p_N\right) + \epsilon_t \\
&= \mathbb{E}_t\left(GBP_T | D\right) + \epsilon_t \\
&= \Pr\left(D = 1 | \mathcal{F}_t\right)\mathbb{E}_t\left(GBP_T | D = 1\right) + \epsilon_t \\
&\quad + \Pr\left(D = 0 | \mathcal{F}_t\right)\mathbb{E}_t\left(GBP_T | D = 0\right) + \epsilon_t \\
&= \Pr\left(D = 1 | \mathcal{F}_t\right)\left(\mathbb{E}_t\left(GBP_T | D = 1\right) - \mathbb{E}_t\left(GBP_T | D = 0\right)\right) + \epsilon_t \\
&\quad + \mathbb{E}_t\left(GBP_T | D = 0\right) + \epsilon_t
\end{aligned}
$$

$$
\begin{aligned}
BETP_t &= \mathbb{E}_t\left(BETP_T\right) \\
&= \Pr\left(D = 1 | \mathcal{F}_t\right) \times \pounds 1 \\
&= \Pr\left(D = 1 | \mathcal{F}_t\right)
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
GBP_t &= BETP_t \times \left(\mathbb{E}_t\left(GBP_T | D = 1\right) - \mathbb{E}_t\left(GBP_T | D = 0\right)\right) \\
&\quad + \mathbb{E}_t(GBP_T | D = 0) + \epsilon_t
\end{aligned}
$$

Write $GBP_H = \mathbb{E}_t\left(GBP_T | D = 0\right)$, $GBP_L = \mathbb{E}_t\left(GBP_T | D = 1\right)$ and $\triangle GBP = GBP_H - GBP_L$ and assume that these are constants, i.e., independent of t. Prices at time t will thus satisfy:

$$
\begin{aligned}
GBP_t &= GBP_L + (1 - BETP_t) \times (GBP_H - GBP_L) + \epsilon_t \\
&= GBP_L + (1 - BETP_t) \times (\triangle GBP) + \epsilon_t, \qquad (5.2)
\end{aligned}
$$

where ϵ_t is a martingale difference (if ϵ_t were predictable EHMW would be violated). Thus the two markets will be cointegrated with cointegrating vector $(\triangle GBP, -1)$. Taking $\triangle GBP$ as given, an arbitrage opportunity may exist if the error term deviates from zero. This would be facilitated by taking a position in the Betfair contract and taking an opposing position in the pound, in the ratio of the cointegrating vector.

6 Results

Electoral Probability Model results were generated in Matlab, sampling from the relevant multivariate distributions to evaluate model probabilities. Where

calibration was required, we found that a simple gradient descent algorithm was adequate.

6.1 Semi-Strong Market Efficiency

Figure 3 shows the results for calibrating the model prior to the generous standard deviation figure of 5% using the parameters from Table 5. The kurtosis of the distributions is lower for the normal model, due to the logit mapping squeezing the distribution. The logit t-distribution has higher kurtosis, as expected. Both logit distributions are highly significantly different to the normal distribution, as shown by the Jarque-Bera (JB) test pValue. All marginals had a very generous prior 90% confidence interval of around 16.5% width. Figure 4 shows the evolution of the model forecast as the night progressed. The results of all marginals are almost identical except for the higher kurtosis logit-t marginal, which had a surprisingly quicker rate of convergence. The logit-t model predicted Brexit at 1:23:34 am on the 12th result with 95% accuracy, and on the 18th result at 1:45am with 99% accuracy. The other models took until 1:43:46am (15th result) and 2:03am (33th result) to get to those certainties respectively. This compares with the BBC projecting Brexit at 4:39:32am[10]. The Betfair market took until 04:21:00 am to imply 99% probability, and the pound appeared to take even longer to react.

In section 5 it was shown that the assumptions of semi-strong market efficiency (EMHSS) and risk neutrality (RN) leads to $\mathbb{P}_{\text{Brexit}, t} = BETP_t$, where $\mathbb{P}_{\text{Brexit}, t}$ is the true probability of Brexit. This is rejected by the data. It is theoretically possible that this rejection is due to a failure of RN and not EMHSS. If EMHSS holds, but not RN then $\mathbb{P}_{\text{Brexit}, t} = \frac{u(BETP_t)}{u(\pounds 1)}$. To investigate this possibility, we plot the forecast probability versus the Betfair contract price in Figure 5. Also shown is a polynomial regression constrained to be monotonic and unity at a 100% probability of Brexit as well as the path taken by the observed data. This would represent the shape of the utility function under EMHSS holding, but RN not. The regression does not provide a convincing fit. There are many points for which there is a low price for the Betfair contract when the forecast is implying a very high probability of leaving the EU. The poor fit is explained as the paths of the contract price and the likelihood of Brexit were not moving contemporaneously, as evidenced by the path taken by the data not adhering closely to the regression. Failure of EMHSS is the only explanation for such behaviour, not any non-linearity of utility arising due to risk aversion.

We now discuss the semi-strong efficiency of the pound market. We can

---

[10]According to our Bloomberg scrape.

Forecast Model Prior
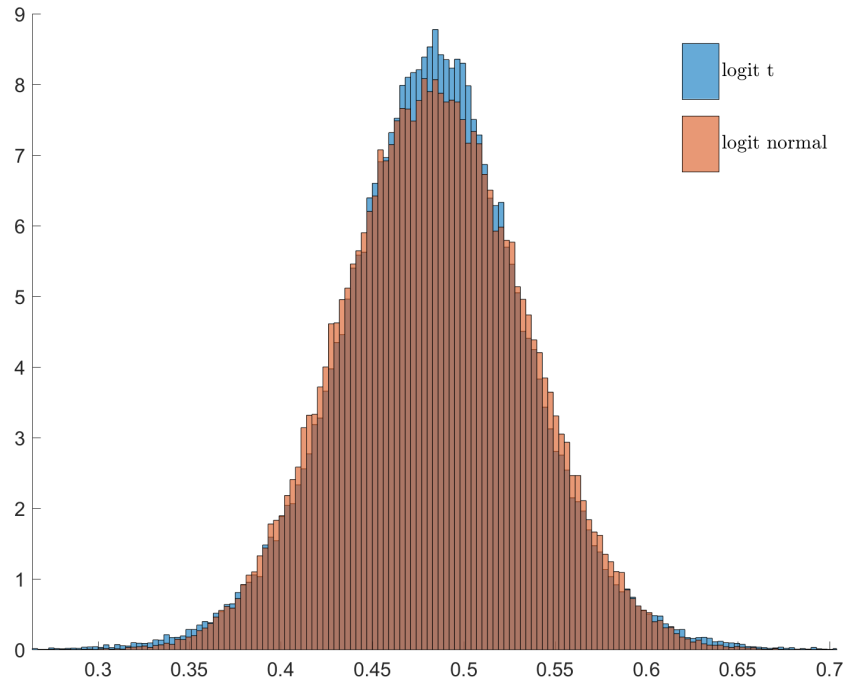
Marginal	Kurtosis	90% CI	99% CI	JB p
Normal	3.0	(40.1%, 56.7%)	(35.4%, 61.4%)	> 0.5
Logit Normal	2.9	(40.2%, 56.6%)	(35.9%, 60.1%)	< 0.001
Beta	3.0	(40%, 56.6%)	(35.5%, 61.3%)	0.29
Logit t	3.6	(40%, 56.6%)	(34.4%, 62.4%)	> 0.5

95% result	99% result
15 (1:43:46)	33 (2:03:00)
15 (1:43:46)	33 (2:03:00)
15 (1:43:46)	33 (2:03:00)
12 (1:23:34)	18 (1:45:00)

Figure 3: Prior Distributions for Forecast Model.

Forecast Model Path
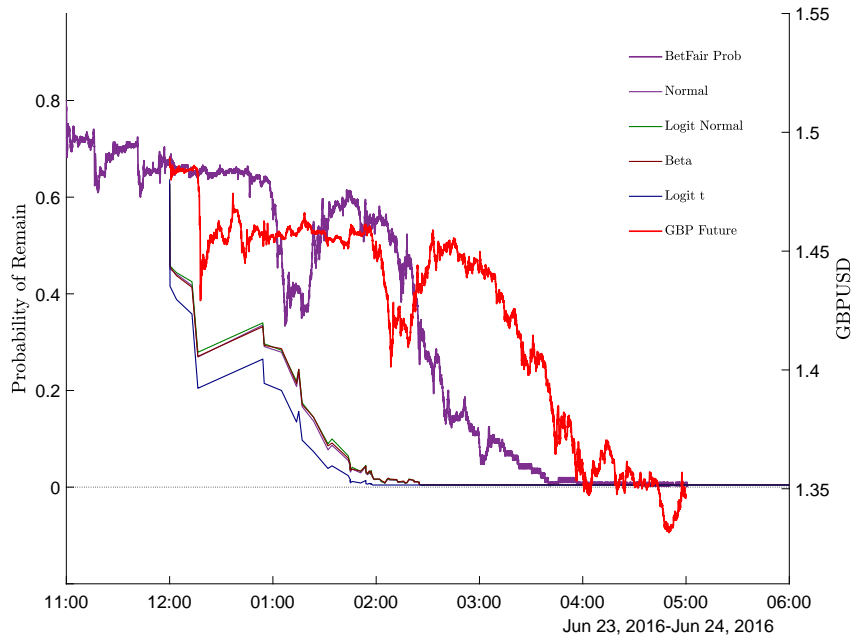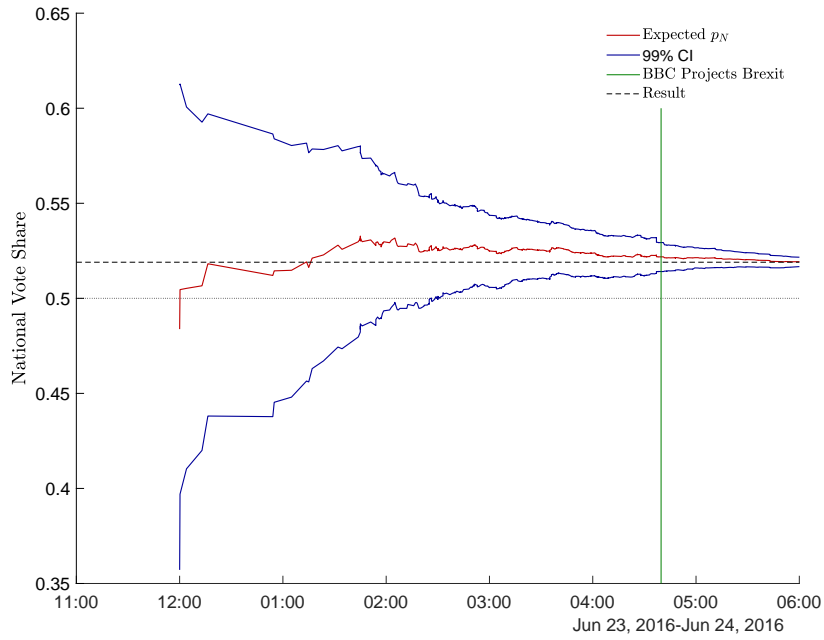
Figure 4: National Vote Distribution Evolution (Logit Normal) and Model Probability Paths for Forecast Model.
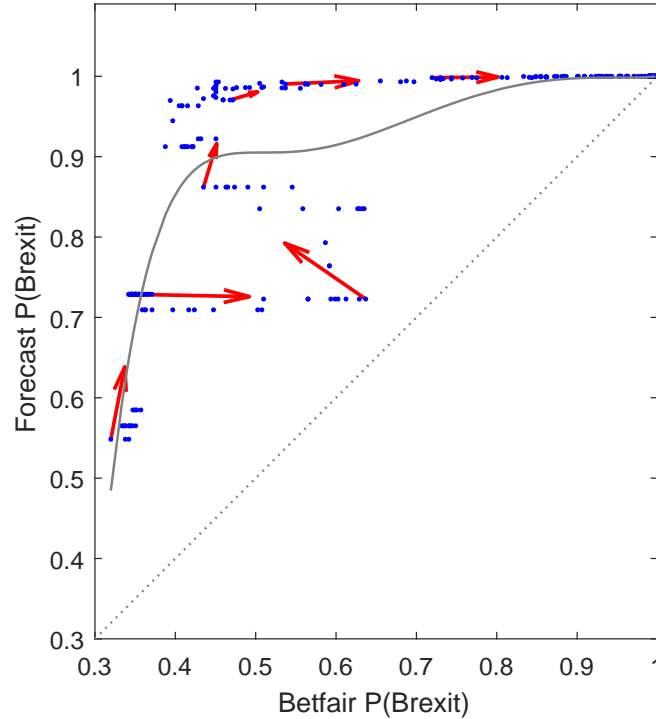
Figure 5: GBP versus Betfair Remain Contract with Monotonic Non-Linear Regression.

evaluate an implied probability of Brexit from this market under the additional CMI and RN assumptions. Figure 6 graphs the implied probabilities of the forecast model and both markets. The forecast model leads the Betfair price, which in turn leads the pound probability. The average horizontal distance on this plot between the relevant lines is 113 minutes for forecast-Betfair and 185 for forecast-pound. It appears that the fundamental information (forecast) led the betting market by nearly 2 hours which in turn led the pound by over an additional hour. Relaxation of RN and CMI in the GBP market will lead to non-linearities arising between the price and probability of Brexit relationships. Again, it is theoretically possible that the relaxation of these assumptions, and not failure of EHMSS in the GBP market, caused the apparent lag in the pound market. However, the next section will discuss this further. A plausible explanation for the behaviour observed is that the pound market discounted the information contained in the vote more slowly than the betting market, whereas non-linearities in any
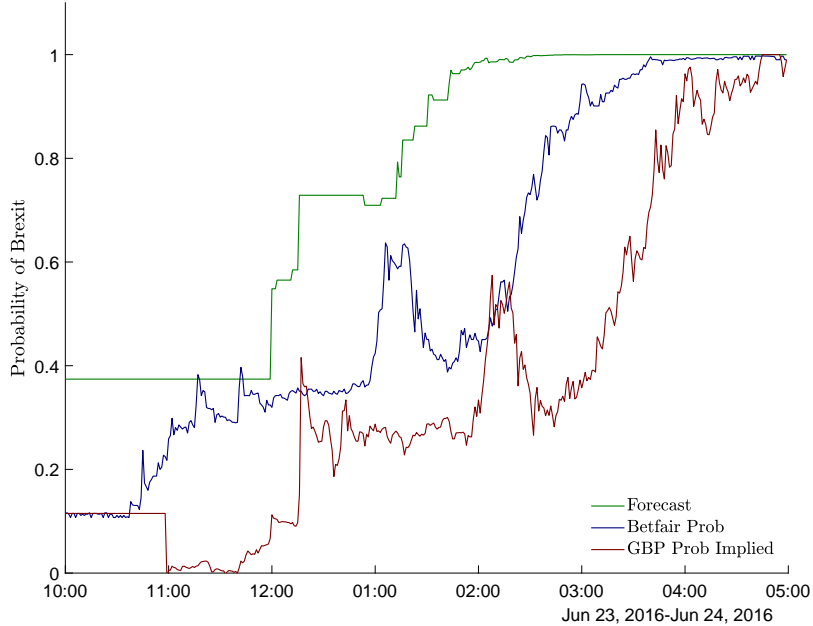
Figure 6: The Forecast Model Probability, the Betfair Probability and the GBP Implied Probability under RN and CMI Assumptions.

contemporaneous relationship caused by deviations from CMI and RN are not.

6.2 Weak Market Efficiency

The evolution of the two markets on the night is shown in Figure 7. To the eye, they do not appear to be moving together. To form a robust conclusion we rely on the theoretical framework developed in section 5. Recall that an implication of EMHW is that both markets discount the same likelihood. Using the notation of the previous section EMHW $\Rightarrow$ $P_{Betfair} = P_{GBP}$ and we explore this possibility.

Under the various assumptions of no shocks to the GBP price beyond the referendum (NS), weak market efficiency (EMHW), conditional mean independence (CMI) and risk neutral (RN) the two markets will be cointegrated. We calculate the cointegrating ratio by evaluating both markets at the beginning and end of the 7 hour period from 10 pm on 23 June to 5 am on 24 June. Following the notation of equation 5.2, this gives implied values of $GBP_L = 1.345$, $GBP_H = 1.499$ and $\triangle GBP = 0.154$. This is similar to the
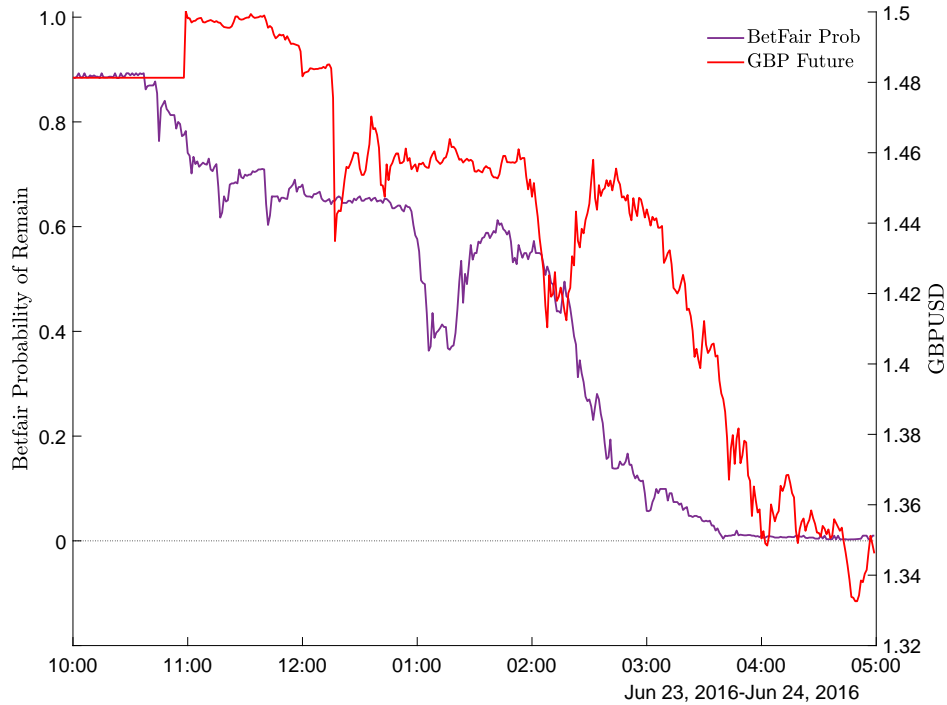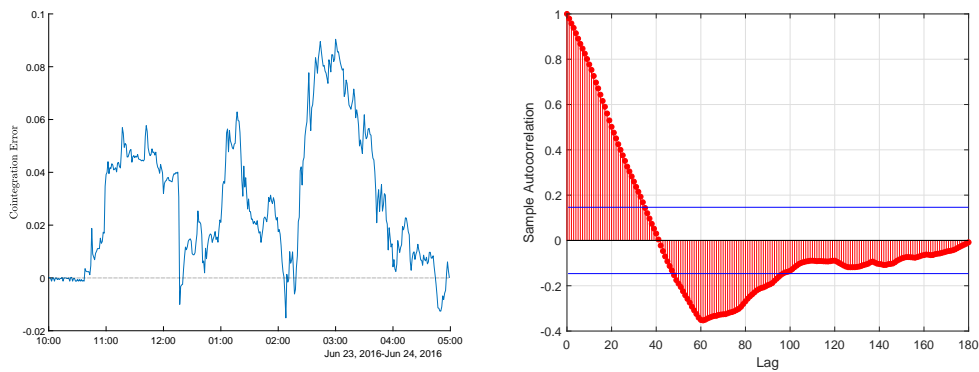
Figure 7: Last Prices by Minute.

Note: The error is the return, before transaction costs, of buying GB-
PUSD and selling the Betfair contract for Remain, in the cointegrating
ratio. There appear to be profitable arbitrage opportunities that require
no successful forecast of GBP

Figure 8: The Theoretical Cointegration Error with Sample Auto Correlation
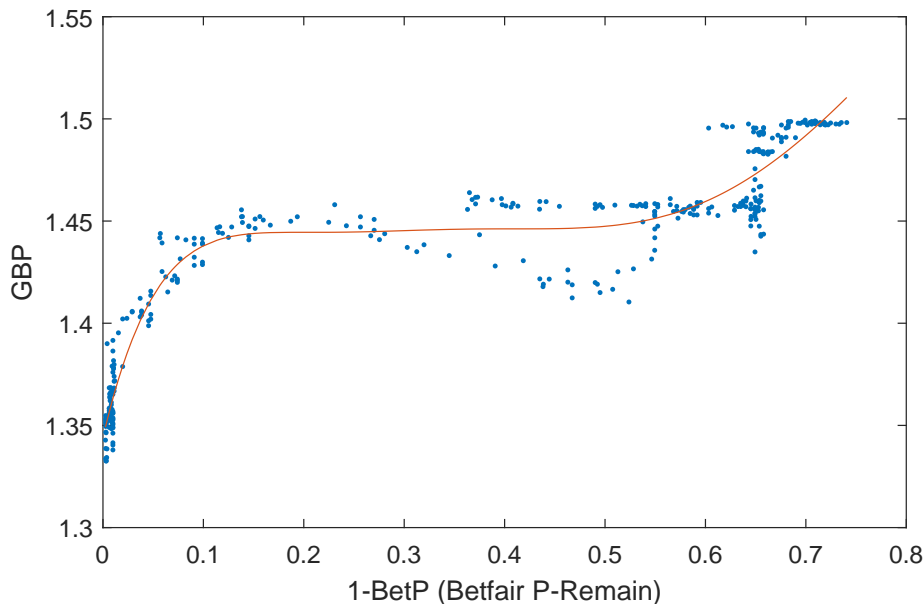Function.

Figure 9: GBP versus Betfair Remain Contract with Monotonic Non-Linear Regression.

predictions made of the pound conditional on the outcome of the referendum made by various commentators in advance (see Wu et al. (2017)).

The cointegration error $\epsilon_t = GBP_t - GBP_L - (1 - BetP) \times \triangle GBP$ is plotted in Figure 8 along with the sample Auto-Correlation Function (ACF). Testing the Null of stationarity of the error with the KPSS test results in rejection with a pValue of < 0.01[11]. This is a formal rejection of the assumptions that led to cointegration, but not necessarily of EMHW. Relaxing the CMI, RN assumptions made in section 5 but still relying on NS, implies that the markets move contemporaneously, according to some possibly non-linear but monotonic function of price $g(\cdot)$ plus a martingale difference ϵ_t. Suppose the rejection of cointegration were due to the resulting non-linearity due to deviations from CMI and RN, what would the $g(\cdot)$ look like?

Figure 9 attempts to answer this question by showing a plot of the pound

---

[11]Testing a null hypothesis of a unit root via an augmented Dickey Fuller test using the saturate and reject method of lag selection results in non-rejection of the null with a pValue of 0.168. This provides further evidence of non-stationarity.

price against that of the Betfair contract for remain for each minute of the night, along with a polynomial monotonic regression which represents $g(\cdot)$. This shows very little sensitivity of the GBP price to the price of the Betfair contracts for prices between about 15p and 50p. This is unlikely. GBP is most sensitive when the likelihood of Brexit is deemed high. This corresponds to the period of the night when the betting market was pricing a high probability of leaving the EU, but that the GBP had yet to react fully. A better explanation is that the pound market had not discounted the result of the vote whereas the betting market had. Further, figure 10 shows how the error of this regression varies with time as well as the autocorrelation function. The assumptions of NS and EMHW imply that this error should be a stationary martingale difference. Stationarity is rejected by the KPSS test with a pValue of 0.01[12] and the ACF has highly significant lags which would in themselves be a rejection of EMHW.

Discounting non-linearities in the price relationship as the cause of rejection of cointegration leads us to consider violations of NS as the only defence against a conclusion that $P_{Betfair} \neq P_{GBP}$. What would need to be assumed for the EMHW to actually hold? One would have to believe that there was a shock to the pound that caused a change in price of 9 cents, or around 7%, which was independent of the referendum results and which was subsequently reversed around 2 hours later, at around the time that Brexit became apparent. This is not plausible. Finally, we consider the case that pound investors could have marked down the price of sterling after the vote to leave the EU was known, but that the exact voteshare, and hence likely "hardness" of Brexit, was not. As can be seen from figure 4 the forecast model's expectation of the voteshare p_N actually peaked prior to 2am above 53%. It then decreased to the actual result of 51.9% during the rest of the night. To the extent that the voteshare had any effect at all beyond being either side of 50% it would suggest a softening of Brexit towards the final hours of the night. This would support GBP not exacerbate a sell-off. The only convincing explanation for the rejection of cointegration is that the referendum shocks were first felt in the betting market and later in the currency market and we conclude that EMHW was indeed violated.

6.3 Arbitrage

It appears that Betfair market led the pound market which suggests that a portfolio of selling the pound and buying the Betfair contract would be

---

[12]Evidence of non-stationarity is not as strong as in the cointegration case as the ADF test does reject a null of a unit root with a pValue of 0.047, whereas the Variance Ratio test barely avoids rejecting a unit root with a pValue of 0.071.
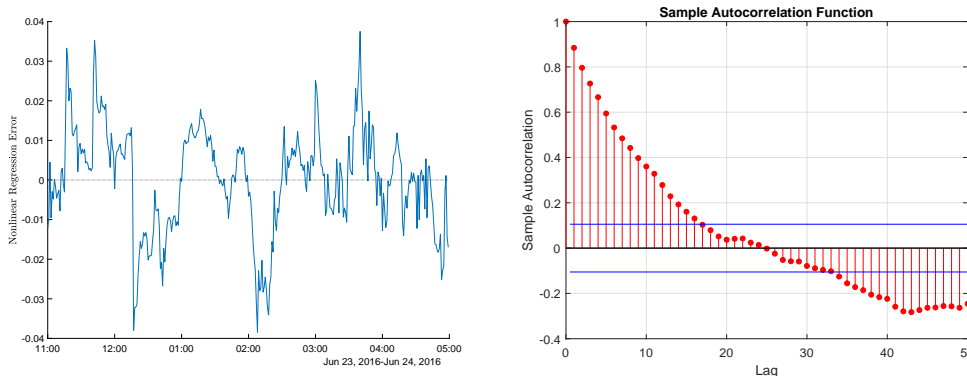
Figure 10: Non-Linear Regression Error with Sample Auto Correlation Function.

profitable. Note that the cointegration error shown in figure 8 is the return of the portfolio generated by buying £1 and selling $\triangle GBP$ Betfair Brexit contracts. Selling in this ratio at around 4 am that would have made roughly 9 cents of profit per £1 of the pound sold, or an unleveraged return of up to 7% in about two hours[13]. Supporters of the EMH could object to the conclusion that the EMH fails by simply selling the pound uncovered due to the increased risk an investor would be being exposed to (the pound, after all, was exceptionally volatile during the period). However, the arbitrage strategy of covering the short with a position in the spread bet largely eliminates risk. The arbitrage is not reliant on the forecast model predictions and further illustrates the failure of the EMH in its weak form.

6.4   Robustness Checks

We now consider the sensitivity of the electoral model to changes in the model parameter changes. Table 8 shows the 95% Brexit prediction times for various parameter changes. No changes are observed when making all parameters except σ_p and ρ_p more conservative by roughly 5% (row 1). Increasing only σ_p by 5% or reducing ρ_p by 5% (rows 2 and 3) reduces the speed of convergence a little with the 95% prediction coming on the 16th as opposed to 15th result. Lowering ρ_p further to the very conservative lower bound of 0.18 slows the result significantly but still predicts Brexit with 99%

---

[13]In terms of transaction costs, selling the pound would cost about 2-3 hundredths of a cent at that time, whereas the Betfair cost is $3-5\%$ levied on any bets that pay out. This would slightly change the ratio of the portfolio but not significantly affect profits or these conclusions.

33

probability about an hour and a half before the Betfair market. Making the correlation between vote share and turnout at the area level ρ_{pq} both zero or of opposite sign increases the speed of convergence a little, with the 95% likelihood now coming on the 14th result. This suggests, although there is some subjectivity in the hyperparameter choices, the conclusions of violations of markets efficiency are both robust to sensible changes as well as completely changing the sign of ρ_{pq}. Further we observe that changes to σ_p and ρ_p only meaningfully affect the model to a large degree, which is as expected.

Heterogeneous correlations

As noted, the electoral model is limited by having homogeneous correlation coefficients. To test this limitation we ran the model with a modified vote-share correlation matrix Σ_p. Correlations were set so that they were higher for areas with similar expected vote share, according to the following formula:

$$(\Sigma_p)_{ij} = A - B\sqrt{|\mu_i - \mu_j|},$$

where A and B are constants chosen so that the average correlation is the usual value of $\rho_p = 0.288$ and the minimum value was 0.1. This form, although crude, is justified as areas with closer expectations are more likely demographically similar and so have a greater degree of dependence. Running this model for normal marginals resulted in 95% and 99% likelihoods of Brexit at 1:15 am (10th result) and 1:17 am (11th result), respectively. This was considerably faster than the homogeneous case, suggesting that departures from homogeneity in the model would quicken convergence and lead to a strengthening of our conclusions.

7 Conclusion

This paper examined the efficiency of the Betfair and pound/dollar markets as the results of the United Kingdom European Union membership referendum were announced. This event provided a unique opportunity to study the interaction between a flow of information, a prediction market and a financial market, where there was a sole, public determinant of prices. Other work has identified the pound market as being inefficient during the period under investigation but we were able to answer questions about the efficiency of the prediction market as well as the relative speed at which the two markets digested the information flow.

Result	σ_P	ρ_p	ρ_q	ρ_{pq}	ν	σ_ν	95% result[a]
1 [b]	5%	0.288	0.877	−0.379	64.2%	6.4%	15 (1:43:46)
2 [c]	5%	0.274	0.835	−0.361	67.6%	6.7%	15 (1:43:46)
3 [d]	5.3%	0.288	0.835	−0.361	67.6%	6.7%	16 (1:44:57)*
4 [e]	5%	0.18	0.835	−0.361	67.6%	6.7%	34 (2:05:00)*
5 [f]	5.3%	0.288	0.835	0	67.6%	6.7%	14 (1:34:21)*
6 [g]	5.3%	0.288	0.835	0.361	67.6%	6.7%	14 (1:34:21)*

[a]*Indicates a change from initial values

[b]Result 1 changes all parameters except ρ_p and σ_p.

[c]Result 2 changes ρ_p only.

[d]Result 3 changes σ_p only.

[e]Result 4 sets ρ_p at the lower limit of our plausible range, 0.18.

[f]Result 5 sets $\rho_{pq} = 0$.

[g]Result 6 changes the sign of ρ_{pq}.

Table 8: How the Times to Predict Brexit Vary with More Conservative Parameter Values (Normal Marginal).

We have presented a rigorous Bayesian real-time model of the probability of Brexit for the period under consideration. This is based on a copula that is not constrained to normal distributions. The Bayesian method improves upon earlier estimation methods as it does not rely on any asymptotic properties of estimators for small samples We also demonstrate robustness of results to changes in the prior. The conclusions of the model are as follows:

1. Not only was the currency market informationally inefficient so too was the betting market, and both markets violated semi-strong EMH on the night of the vote.

2. The betting market, although inefficient, was more efficient than the currency market. The betting market took less than 2 hours to reflect the information contained in the vote whereas the currency market took over 3 hours.

3. There was a close to risk-free arbitrage opportunity in the two markets. The arbitrage result suggests that a violation of EMH in the weak form has occurred. The conclusion that there is a failure of the weak form of the hypothesis is not reliant on any flow of fundamental information or the electoral probability model.

Our results suggest that market participants suffered a behavioural bias as the results unfolded. It appears that traders and gamblers simply could not

believe that the UK was voting to leave the EU. Further, it appears that on this occasion the betting market, although slow, adjusted much more quickly than the financial markets. Any future possible UK referendum on this subject will present an opportunity to study whether this inefficiency persists or whether efficient behaviour is exhibited, possibly due to the publication of this and other studies. If an inefficiency were to persist it would be interesting to observe whether the betting markets again lead the financial markets.

A   Model summary

A.1   Variables

$p_i, q_i, s_i :$ Voting area vote percentage, turnout and size, order by time of announcement

$p_n, \mu_N :$ National vote share and prior mean for Brexit

$\sigma_P^2, \hat{\sigma}_P^2 :$ National vote share prior variance and estimate, for Brexit

$\mu_{p_i} :$ Expectation of p_i

$\mu_H :$ Vector of expectations provided by Hanretty study

$\sigma_{p_i}^2 :$ Marginal variance of p_i

$\upsilon_i :$ Expectation of q_i

$\sigma_\upsilon^2 :$ Marginal variance of q_i, independent of i

$r :$ Vector of variables $= (p_1, \ldots, p_n, q_1, \ldots, q_n)^\intercal$

$\widetilde{r} :$ Re-ordered vector by time of announcement $= (p_1, q_1, \ldots, p_n, q_n)^\intercal$

$F_i, \widetilde{F}_i :$ Marginal CDFs of i'th components of r and $\tilde{r}$

$\rho_\theta, \rho_\phi :$ Inter-area prior vote share and turnout correlation

$\rho_{\theta\phi} :$ Intra-area prior vote and turnout correlation

$\Sigma_0 :$ Covariance matrix of copula prior

$\widetilde{\Sigma}_0 :$ Covariance matrix of copula prior with sequentially reordered rows and columns

$\widetilde{\Pi} :$ Mean of prior after the announcement of m results

$\widetilde{\Sigma} :$ Covariance matrix of prior after the announcement of m results

A.2 Prior Probability of Brexit

$$P(\text{BREXIT})_0 = P\left(\frac{\sum_i p_i q_i s_i}{\sum_i q_i s_i} > \frac{1}{2}\right)$$

$$\Phi^{-1}(F_1(r_1)), \ldots, \Phi^{-1}(F_{2n}(r_{2n})) \sim N(0, \Sigma_0)$$

$$\Sigma_0 = \begin{pmatrix} \Sigma_p & \Sigma_{pq} \\ \Sigma_{pq}^T & \Sigma_q \end{pmatrix}$$

$$\Sigma_p = (1 - \rho_p)\, I_n + (\rho_p)\, i_n i_n'$$

$$\Sigma_q = (1 - \rho_q)\, I_n + (\rho_q)\, i_n i_n'$$

$$\Sigma_{pq} = \rho_{pq} \times \left[(1 - (\rho_q \rho_p))\, I_n + (\rho_q \rho_p) i_n i_n'\right]$$

A.3 Prior Marginal Calibration

$$\mu_p = \mu_H + \alpha_N \times i$$

$$\sigma_{p_i}^2 = (\sigma_H^2)_i + \sigma_N^2$$

$$\sigma_N^2, \alpha_N : \quad E(p_N) = \mu_N, \ \sigma_P = \hat{\sigma}_P^2$$

A.4 Update

$$\tilde{\Sigma}_0 = \begin{pmatrix} \tilde{\Sigma}_{m,m} & \tilde{\Sigma}_{m,\not{m}} \\ \tilde{\Sigma}_{\not{m},m} & \tilde{\Sigma}_{\not{m},\not{m}} \end{pmatrix}$$

$$\tilde{x}_m = (F_{p_1}(p_1), F_{q_1}(q_1), \ldots, F_{p_m}(p_m), F_{p_m}(q_m))'$$

$$\tilde{\Pi}_{\not{m}} = \tilde{\Sigma}_{\not{m},m}\tilde{\Sigma}_{m,m}^{-1}\tilde{x}_m$$

$$\tilde{\Sigma}_{\not{m}} = \tilde{\Sigma}_{\not{m},\not{m}} - \tilde{\Sigma}_{\not{m},m}\tilde{\Sigma}_{m,m}^{-1}\tilde{\Sigma}_{m,\not{m}}$$

$$P(\text{BREXIT})_m = P\left(\frac{\sum_{i > m} p_i q_i s_i}{\sum_i q_i s_i} > \frac{1}{2} - \frac{\sum_{i \leq m} p_i q_i s_i}{\sum_i q_i s_i}\right)$$

$$\Phi^{-1}(\tilde{F}_{2m+1}(\tilde{r}_{2m+1})), \ldots, \Phi^{-1}(\tilde{F}_{2n}(\tilde{r}_{2n}))|\tilde{x}_m \sim N(\tilde{\Pi}_{\not{m}}, \tilde{\Sigma}_{\not{m}})$$

B Review of Probability Model in Wu et al. (2017)

The model under consideration in Wu et al. (2017) performs (in the one factor case) the following Weighted Least Squares regression following the announcement of k results

$$p_i = \alpha\mu_i + \beta + \epsilon_i, \qquad \epsilon_i \sim N(0, \sigma_\epsilon^2) \quad i = 1, \ldots, k.$$

The national vote share and thus the probability of Brexit is then simulated by generating M realisations and evaluating the relevant sum. A correct application of this method would involve sampling unknowns $(\alpha, \beta, \sigma_\epsilon^2)$ from the joint distribution

$$N\left(\left(\begin{array}{c} \hat{\alpha} \\ \hat{\beta} \end{array}\right), \left(\begin{array}{cc} \hat{\sigma}_\alpha^2 & \rho_{\alpha\beta}\hat{\sigma}_\alpha\hat{\sigma}_\beta \\ \rho_{\alpha\beta}\hat{\sigma}_\alpha\hat{\sigma}_\beta & \hat{\sigma}_\beta^2 \end{array}\right)\right), \ \chi^2(k-2),$$

where the slope and the intercept from linear regression are mutually correlated. Then the correct covariance and variance of unannounced results would be:

$$
\begin{aligned}
\mathrm{cov}(p_i, p_j) &= E_{(\alpha,\beta,\sigma_\epsilon^2)}\left[\mathrm{cov}(p_i, p_j | \alpha, \beta, \sigma_\epsilon^2)\right] \\
&\quad + \mathrm{cov}_{(\alpha,\beta,\sigma_\epsilon^2)}\left[E(p_i|\alpha,\beta,\sigma_\epsilon^2), E(p_j|\alpha,\beta,\sigma_\epsilon^2)\right] \\
&= \mathrm{cov}_{(\alpha,\beta,\sigma_\epsilon^2)}\left[\alpha\mu_i + \beta, \alpha\mu_j + \beta\right] \\
&= \mu_i\mu_j\hat{\sigma}_\alpha^2 + \hat{\sigma}_\beta^2 + [(\mu_i + \mu_j)\rho_{\alpha\beta}\hat{\sigma}_\alpha\hat{\sigma}_\beta] \\
\mathrm{var}(p_i) &= \hat{\sigma}_\epsilon^2 + \mu_i\mu_j\hat{\sigma}_\alpha^2 + \hat{\sigma}_\beta^2 + [(\mu_i + \mu_j)\rho_{\alpha\beta}\hat{\sigma}_\alpha\hat{\sigma}_\beta].
\end{aligned}
\tag{B.1}
$$

However, an assumption that $\hat{\alpha}$ and $\hat{\beta}$ are uncorrelated appears to be used when in fact they are close to being perfectly anti-correlated with[14]

$$\rho_{\alpha\beta} = \frac{-\sum_1^k \mu_i}{\sqrt{n\sum_1^k \mu_i^2}} = -\frac{\bar{\mu}}{\sqrt{\bar{\mu^2}}} \approx -1.$$

The calculation in Wu et al. (2017) would thus be calculating the variance structure of the unknown referendum results as:

$$
\begin{aligned}
\mathrm{cov}(p_i, p_j) &= \mu_i\mu_j\hat{\sigma}_\alpha^2 + \hat{\sigma}_\beta^2 \\
\mathrm{var}(p_i) &= \sigma_\epsilon^2 + \mu_i\mu_j\hat{\sigma}_\alpha^2 + \hat{\sigma}_\beta^2,
\end{aligned}
$$

---

[14]A simulation of their results yielded $\rho_{\alpha\beta} \in (-0.97, -1)$.

which is different to the values in equation B.1.

Another issue with the method is that the assumption of a normal distribution for $(\hat{\alpha}, \hat{\beta})$ with correctly specified errors is highly questionable, particularly in small samples. $(\hat{\alpha}, \hat{\beta})$ only follows such a distribution in finite samples if the errors are normal, and if not, it would be biased but consistent. The normal distribution is an asymptotic result for non-normal errors, and even then a correct evaluation assumes no heteroskedasticity; otherwise error estimates are likely to be too low and implied probabilities of Brexit to be too confident. This could be overcome using robust errors, but only for large data sets. Using robust errors in small samples can produce severely biased estimators.

The model in Wu et al. (2017) and that presented in this paper use different approaches to estimate the covariance structure of the conditional distributions used to form predictions. That of Wu et al. (2017) requires no prior (beyond expectations) and attempts to infer the covariance structure from an OLS regression of results announced so far. Our model, by contrast, starts with a prior for the covariance structure and updates that prior as results come in. Both methods will produce the same covariance and results in larger samples but will be different for small samples. The different approach is illustrative of the differences between a Frequentist and Bayesian approach to inference. However, we suggest that the Frequentist approach presented in Wu et al. (2017) is not appropriate for the small numbers of results available at the times of predictions (<20 results). More sophisticated corrections for small sampling estimation would be desirable. We believe our Bayesian approach is a more suitable way to proceed in the case of this application.

Acknowledgments

References

Badarinathi, R., & Kochman, L. (1996). Football betting and the efficient market hypothesis. *The American Economist*, *40*, 52–55.

Baxter, M. (2017). Electoral Calculus; Historical Data and Plots. URL: www.electoralcalculus.co.uk Retrieved July 2017.

Caballero, R. J., & Simsek, A. (2016). *A Model of Fickle Capital Flows and Retrenchment*. Working Paper 22751 National Bureau of Economic Research.

Curtice, J., Fisher, S. D., & Kuha, J. (2011). Confounding the commentators: how the 2010 exit poll got it (more or less) right. *Journal of Elections, Public Opinion and Parties*, *21*, 211–235.

DellaVigna, S., & Pollet, J. M. (2009). Investor inattention and Friday earnings announcements. *The Journal of Finance*, *64*, 709–749.

Ding, P. (2016). On the conditional distribution of the multivariate t distribution. *The American Statistician*, *70*, 293–295.

Econoday (2016). 2016 Economic Calendar. URL: http://www.econoday.com.

Electoral Commission (2016). EU Referendum Results. URL: www.electoralcommission.org.uk retrieved July 2016.

Electoral Commission (2017). Previous General Elections. URL: www.electoralcommission.org.uk retrieved July 2017.

Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, *38*, 34–105.

Foster, A. (2016). Brexit: How important will turnout be to the EU referendum result? The Express. Retrieved July 2017.

GabrielaTitan, A. (2015). The efficient market hypothesis: Review of specialized literature and empirical research. *Procedia Economics and Finance*, *32*, 442–499.

Goddard, J., & Asimakopoulos, I. (2004). Forecasting football results and the efficiency of fixed-odds betting. *Journal of Forecasting*, *23*, 51–66.

Graefe, A. (2016). Political markets. *Forthcoming (subject to changes) in the SAGE Handbook of Electoral Behavior*, .

Gutteridge, N. (2016). Working class revolution? Reports of huge EU referendum turnout which "would favour leave". The Express. Retrieved August 2016.

Hanretty, C. (2016). The EU referendum: what to expect on the night. Blog. URL: https://medium.com/@chrishanretty/the-eu-referendum-what-to-expect-on-the-night-521792dd3eef retreived July 2016.

Hirshleifer, D., Hsu, P.-H., & Li, D. (2013). Innovative efficiency and stock returns. *Journal of Financial Economics*, *107*, 632–654.

Hirshleifer, D., Lim, S. S., & Teoh, S. H. (2009). Driven to distraction: Extraneous events and underreaction to earnings news. *The Journal of Finance*, *64*, 2289–2325.

Horn, C. F., Ivens, B. S., Ohneberg, M., & Brem, A. (2014). Prediction markets–a literature review 2014. *The Journal of Prediction Markets*, *8*, 89–126.

Hou, K., Xiong, W., & Peng, L. (2009). A tale of two anomalies: The implications of investor attention for price and earnings momentum, .

Huang, J. Y., Shieh, J. C., & Kao, Y.-C. (2016). Starting points for a new researcher in behavioral finance. *International Journal of Managerial Finance*, *12*, 92–103.

Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *The Journal of Economic Perspectives*, *17*, 59–82.

Rallings, C., Johnston, R., & Thrasher, M. (2008). Changing the boundaries but keeping the disproportionality: The electoral impact of the fifth periodical reviews by the Parliamentary Boundary Commissions for England and Wales. *The Political Quarterly*, *79*, 80–90.

Reyes, B. (2016a). Gibraltar will vote to remain in EU - poll. Gibraltar Chronicle.

Reyes, B. (2016b). Support for Remain strengthens in Gibraltar - poll. Gibraltor Chronicle.

Shapiro, Y. (2016). Eu Referendum: Growing support in Northern Ireland for Brexit but balance of power lies with the undecided. Belfast Telegraph.

Twyman, J. (2016). Analysis: the impact of turnout on the EU referendum. YouGov. Retrieved August 2017.

Vlastakis, N., Dotsis, G., & Markellos, R. N. (2009). How efficient is the European football betting market? Evidence from arbitrage and trading strategies. *Journal of Forecasting*, *28*, 426–444.

Wall, M., Costello, R., & Lindsay, S. (2017). The miracle of the markets: Identifying key campaign events in the Scottish independence referendum using betting odds. *Electoral Studies*, *46*, 39–47.

Wikipedia (2017a). Opinion Polling for the United Kingdom European Union membership referendum. URL: `www.wikipedia.org` retrieved July 2017.

Wikipedia (2017b). Opinion polling for the United Kingdom General Election, 2015, 2010, 2005, 2001, 1997, 1992. URL: `www.wikipedia.org` retrieved July 2017.

Wolfers, J., & Zitzewitz, E. (2004). Prediction markets. *Journal of Economic Perspectives*, *18*, 107–126.

Wu, K., Wheatley, S., & Sornette, D. (2017). The British Pound on Brexit Night: A Natural Experiment of Market Efficiency and Real-Time Predictability. *Swiss Finance Institute Paper no. 17-12*, .

YouGov (2016). YouGov EU Referendum On the Day Poll.