

Nielsen, Morten Ørregaard; Shibaev, Sergei S.

Working Paper

Forecasting daily political opinion polls using the fractionally cointegrated VAR model

Queen's Economics Department Working Paper, No. 1340

Provided in Cooperation with:

Queen's University, Department of Economics (QED)

Suggested Citation: Nielsen, Morten Ørregaard; Shibaev, Sergei S. (2015) : Forecasting daily political opinion polls using the fractionally cointegrated VAR model, Queen's Economics Department Working Paper, No. 1340, Queen's University, Department of Economics, Kingston (Ontario)

This Version is available at:

<https://hdl.handle.net/10419/122030>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Queen's Economics Department Working Paper No. 1340

Forecasting daily political opinion polls using the fractionally cointegrated VAR model

Morten Åyrregaard Nielsen
Queen's University and CREATES

Sergei S. Shibaev
Queen's University

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

6-2015

Forecasting daily political opinion polls using the fractionally cointegrated VAR model*

Morten Ørregaard Nielsen[†]
Queen's University and CREATES

Sergei S. Shibaev
Queen's University

June 14, 2015

Abstract

We examine forecasting performance of the recent fractionally cointegrated vector autoregressive (FCVAR) model. The model is applied to daily polling data of political support in the United Kingdom for 2010–2015. We compare with popular competing models and at various forecast horizons. Our findings show that the precision of forecasts generated by the FCVAR model is better than all multivariate and univariate models in the portfolio, and the four variants of the FCVAR model considered are generally ranked as the top four models in terms of forecast accuracy. Furthermore, the FCVAR model significantly outperforms the standard cointegrated VAR (CVAR) model at all forecast horizons and the relative forecast improvement is highest at longer forecast horizons, where the root mean squared forecast error of the FCVAR model is up to 20% lower than that of the CVAR benchmark model. In an empirical application to the prediction of vote shares in the 2015 UK general election, forecasts generated by the FCVAR model leading into the election appear to provide a more informative assessment of the current state of public opinion on electoral support than that suggested by the hung government prediction of the opinion poll. Specifically, the FCVAR model projects the correct direction for the realized vote shares in the election for both the Conservative and Labour parties.

Keywords: forecasting, fractional cointegration, opinion poll data, vector autoregressive model.

1 Introduction

In this paper we investigate the forecasting performance of the recently developed fractionally cointegrated vector autoregressive (FCVAR) model of [Johansen \(2008\)](#) and [Johansen](#)

*We are grateful to Martin Burda, Jianjian Jin, Søren Johansen, Michal Popiel, seminar participants at Queen's University, and conference participants at the DWAE 2015 and the CEA 2015 for helpful comments and suggestions. Nielsen thanks the Canada Research Chairs program, the Social Sciences and Humanities Research Council of Canada (SSHRC), and the Center for Research in Econometric Analysis of Time Series (CREATES, funded by the Danish National Research Foundation, DNRF78) for financial support.

[†]Corresponding author. Address: Department of Economics, Dunning Hall, Queen's University, 94 University Avenue, Kingston, Ontario, K7L 3N6, Canada; E-mail: mon@econ.queensu.ca

and Nielsen (2012) relative to a portfolio of competing models at various forecast horizons. The FCVAR model generalizes the concept of cointegration, and in particular generalizes Johansen's (1995) cointegrated VAR (CVAR) model to fractionally integrated time series, and hence allows estimating long-run equilibrium relationships between fractional time series.

The FCVAR model is a very recently developed statistical model, and it is therefore of particular interest to examine the gains this model can deliver for the purposes of forecasting. The choice of data set for applying the model should reflect a current and relevant issue for forecasting. A prominent example is the desire to predict political support and election vote share outcomes. This paper addresses this task by applying the FCVAR model to a novel data set which is comprised of polling results of political support in the United Kingdom for the period 2010–2015 at the business-daily observation frequency. The fractional integration behavior of political opinion polling data has been well established in the literature, albeit for time series at lower frequencies (monthly, quarterly), e.g. Box-Steffensmeier and Smith (1996), Byers et al. (1997), Dolado et al. (2002, 2003), and Jones et al. (2014). For a more general reference on fractional integration methods in political time series data, see Box-Steffensmeier and Tomlinson (2000). It therefore appears natural to apply a fractional time series model such as the FCVAR to model and forecast political opinion polls.

The industry standard for measuring the current state of political support is through opinion polling. The demand for polling and survey methodology is largely driven by the clients desire to form an accurate understanding of the current state of opinion on a particular question. The poll evidence then serves as an input into the decision making process. When polls are conducted at regular intervals, it seems natural to use a statistical model to extract the full potential of the information contained in these time series of poll results by using them to forecast public opinion beyond the most recent poll date. However, long time series of poll data are scarce, and, to the best of our knowledge, all previous studies that have analyzed time series of political opinion polls have used data observed at the monthly frequency or lower, see, e.g., above references. Authors of these studies have noted that an ideal data set would have all observations contained within a single government regime spanning only one political cycle, while providing a large enough sample to conduct meaningful statistical analysis. The data set used in this paper fully satisfies both desired properties: it spans the entire UK political cycle following the 2010 UK general election, it is conducted at a high observation frequency (business-daily), and it is very recent and on-going, and hence very relevant also for forecasting poll standings which can be viewed as the predicted vote shares for each political party in an election.

The long time series provided in our data set facilitates forecast accuracy evaluation using several forecast evaluation procedures. In particular, it allows the formation of a large number of training sets from which each statistical model can produce forecasts. We apply two standard procedures to assess forecast accuracy: the rolling window and recursive forecasting schemes. The main distinction between the two schemes is how they select the training sets used for estimating the models. The rolling window scheme uses a fixed training set length (commonly referred to as a window) that moves across the data set, and the recursive scheme uses an expanding training set length with a fixed start date. The portfolio of models we consider consists of eight statistical models, four of which are variants of the FCVAR model. These are then evaluated on their forecasting ability relative to a group of four popular competing models. Among the latter, the CVAR model serves as the

main multivariate benchmark model. The simple ARFIMA(0, d , 0) and the more general ARFIMA(p , d , q) models serve as the fractional univariate benchmarks, where the former was found by, e.g., [Byers et al. \(1997\)](#) and [Dolado et al. \(2002\)](#), to fit (monthly) UK polling data well. Finally, we include the ARMA(p , q) model as the classical univariate benchmark. Forecast accuracy is assessed at seven out-of-sample forecast horizons: 1, 5, 10, 15, 20, 25, and 50 steps ahead.

The forecasting analysis in this paper shows that the FCVAR model delivers valuable gains in predicting political support. Both forecasting schemes agree on this finding. The accuracy of forecasts generated by the FCVAR model is better than all multivariate and univariate models in the portfolio, and overall the four variants of the FCVAR model are ranked as the four top performing models. Not only do they perform better relative to the other models, but the forecasting performance of all FCVAR variants is within very close range of each other. When compared to the main benchmark model, the FCVAR model substantially outperforms the CVAR model in 56 of 56 cases, and the relative forecast improvement is highest at the 15–50 steps ahead forecast horizons, where the root mean squared forecast error (RMSFE) of the FCVAR model is up to 20% lower than that of the CVAR benchmark model.

As an empirical application, we consider the prediction of the vote shares for the 2015 UK general election from a range of forecast horizons. In this application, we find that the FCVAR model has advantages for predicting vote shares and appropriately complements the industry standard of basing predictions on the latest opinion poll standings.

The remainder of the paper is structured as follows. Section 2 introduces the concept of fractional integration, the classic arguments for fractional integration in polling data, and describes our data set. In Section 3 we describe the FCVAR methodology and in Section 4 we give an example of the estimation of the model to our data set. Section 5 presents the main forecasting analysis including the empirical application to the 2015 UK general election. Finally, Section 6 provides some concluding remarks.

2 Fractional integration, polling data, and summary statistics

In important early contributions, [Box-Steffensmeier and Smith \(1996\)](#) and [Byers et al. \(1997\)](#) show that political popularity, as measured by public opinion polls, can be modeled as fractional time series processes. The fractional (or fractionally integrated or just integrated) time series models are based on the fractional difference operator,

$$\Delta^d X_t = \sum_{n=0}^{\infty} \pi_n(-d) X_{t-n}, \quad (1)$$

where the fractional coefficients $\pi_n(u)$ are defined in terms of the binomial expansion $(1 - z)^{-u} = \sum_{n=0}^{\infty} \pi_n(u) z^n$, i.e.,

$$\pi_n(u) = \frac{u(u+1) \cdots (u+n-1)}{n!}. \quad (2)$$

For details and for many intermediate results regarding this expansion and the fractional coefficients, see, e.g., [Johansen and Nielsen \(2015, Appendix A\)](#). Efficient calculation of

fractional differences, which we apply in our analysis, is discussed in [Jensen and Nielsen \(2014\)](#).

With the definition of the fractional difference operator in (1), a time series X_t is said to be fractional of order d , denoted $X_t \in I(d)$, if $\Delta^d X_t$ is fractional of order zero, i.e. if $\Delta^d X_t \in I(0)$. The latter property can be defined in the frequency domain as having spectral density that is finite and non-zero near the origin or in terms of the linear representation coefficients if the sum of these is non-zero and finite, see, e.g., ([Johansen and Nielsen, 2012](#), p. 2672). An example of an $I(0)$ process is the stationary and invertible ARMA model.

The standard reasoning for political opinion poll series being fractional relies on [Robinson’s \(1978\)](#) and [Granger’s \(1980\)](#) aggregation argument, and can briefly be described as follows. Suppose individual level voting or polling behavior is governed by the (possibly binary) autoregressive process

$$x_{i,t} = \delta_{i,1} + \delta_{i,2}x_{i,t-1} + u_{i,t}, \tag{3}$$

where $i = 1, \dots, N$ denotes individuals and $t = 1, 2, \dots$ as usual denotes time. The important point here is that the autoregressive coefficients $\delta_{i,2}$ differ across individuals. Some individuals have coefficients $\delta_{i,2} \approx 0$ and are referred to as “floating” voters, whereas others have coefficients $\delta_{i,2} \approx 1$ and are referred to as “committed” voters.¹ If it is assumed that the distribution of $\delta_{i,2}$ across individuals in the population follows a $\text{Beta}(u, v)$ distribution, then the aggregate vote share or polling share $X_t = N^{-1} \sum_{i=1}^N x_{i,t}$ is fractionally integrated of order $d = 1 - v$ when N is large, i.e., $X_t \in I(1 - v)$. For more details, see [Box-Steffensmeier and Smith \(1996\)](#) or [Byers et al. \(1997\)](#).

The above theoretical argument in favor of modeling opinion poll data as fractional time series has been supported in empirical work by a large number of authors. For example, [Box-Steffensmeier and Smith \(1996\)](#) estimate fractional models for US data, [Byers et al. \(1997\)](#) and [Dolado et al. \(2002\)](#) analyze UK data, [Dolado et al. \(2003\)](#) analyze Spanish data, and [Jones et al. \(2014\)](#) analyze Canadian data. All find strong evidence in support of fractional integration with estimates of d around 0.6 – 0.8. In addition, [Byers et al. \(2007\)](#) analyze an updated version of the sample in [Byers et al. \(1997\)](#) and show that the change to phone interviews had no effect on estimates of d and did not appear to constitute a structural break.

The aggregate polling data set we analyze is from the on-going YouGov daily poll of voting intention for political parties in the United Kingdom. Each business day survey participants are asked the question: “If there were a general election tomorrow, which party would you vote for? Conservative, Labour, Liberal Democrat, Scottish Nationalist/Plaid Cymru, some other party, would not vote, don’t know?”. This poll, and hence the data series, is business-daily and began on May 14th, 2010 (so shortly after the 55th UK general election of 2010 held on May 6th). With the next general election held on May 7th, 2015, this on-going survey provides a long series of polling data all within the tenure of a single government regime. The results presented in this paper use May 6th, 2015, as the end date, which was the last day the poll was conducted before the election, for a total of 1227 business-

¹“Floating” voters are defined as those who do not have a strong alliance to one party and may be more easily swayed by current events, media, etc., and “committed” voters, on the other hand, are those who consistently vote for a particular party and are less inclined to change their voting preference.

Table 1: Summary statistics (data in percentage)

Series	Mean	SD	Min	Max	Skew	Kurt	Start date	End date	Obs
Conservative (CP)	34.67	3.29	27	44	0.76	2.89	2010/5/14	2015/05/06	1227
Labour (LP)	39.39	3.32	30	45	-0.44	2.33	2010/5/14	2015/05/06	1227
Liberals (LD)	9.41	1.87	5	21	1.42	7.75	2010/5/14	2015/05/06	1227

Notes: The table presents summary statistics for the polling data (expressed in percentages). The statistics presented are the sample mean, standard deviation, minimum, maximum, skewness, kurtosis, start and end dates, and number of observations.

Table 2: Summary statistics (logit transformed data)

Series	Mean	SD	Min	Max	Skew	Kurt	ELW(m)		
							$m = T^{0.6}$	$m = T^{0.7}$	$m = T^{0.8}$
Conservative (CP)	-0.63	0.14	-0.99	-0.24	0.67	2.81	0.79	0.70	0.65
Labour (LP)	-0.43	0.14	-0.84	-0.20	-0.50	2.41	0.88	0.71	0.64
Liberals (LD)	-2.28	0.20	-2.94	-1.32	0.50	4.59	0.85	0.66	0.64

Notes: The table presents summary statistics for the logit transform of the polling data. The start and end dates are the same as in Table 1, as are the statistics presented, with the addition of ELW(m), which denotes the exact local Whittle estimator of Shimotsu and Phillips (2005) with bandwidth parameter m , whose asymptotic standard error is $(4m)^{-1/2}$.

daily observations.² Previous empirical studies of political support have analyzed monthly and quarterly data spanning several decades and election cycles. Thus, this daily frequency data set is particularly attractive for estimating models within a single election cycle.

Our analysis focuses on the three major political parties in the United Kingdom: the Conservative Party (CP) and the Liberal Democrats (LD), which together constitute the British Government over the sample period (coalition formed on May 12th, 2010), and the Labour Party (LP)—the official opposition.³ In Table 1 we present some summary statistics for the polling data, where these are given in percentage vote shares.

The analysis proceeds after converting the polling data to log-odds (for more details and background, see Byers et al. (1997)). This is done to map variables on the unit interval into the real line, in order to use error terms with unbounded support in our models. The log-odds or logit transformation for a variable $Y_t \in (0, 1)$ is

$$y_t = \log \left(\frac{Y_t}{1 - Y_t} \right),$$

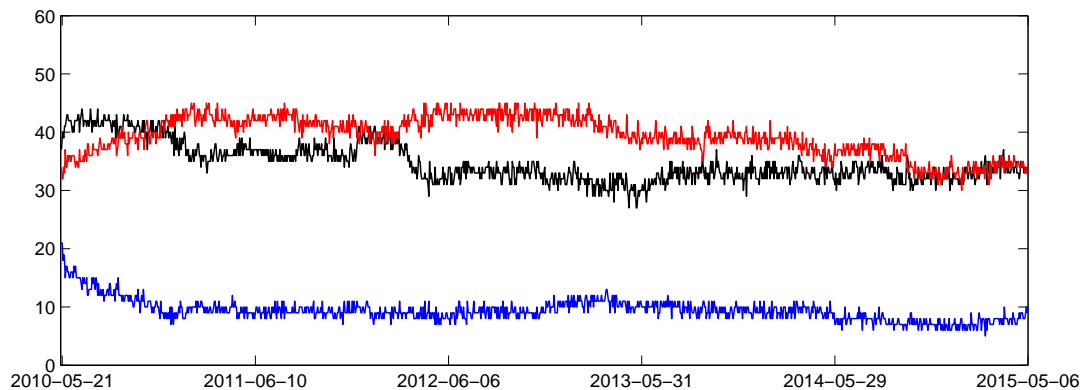
where Y_t is the original series and y_t is the logit transformed series with support $(-\infty, \infty)$.

²Starting April 7th, 2015, i.e. for the last month before the 2015 election, YouGov changed their polling frequency to daily, including non-business days. We ignore this minor change, as well as the break in polling over the Christmas holiday, and in our analysis we treat all observations as equi-distant as is standard.

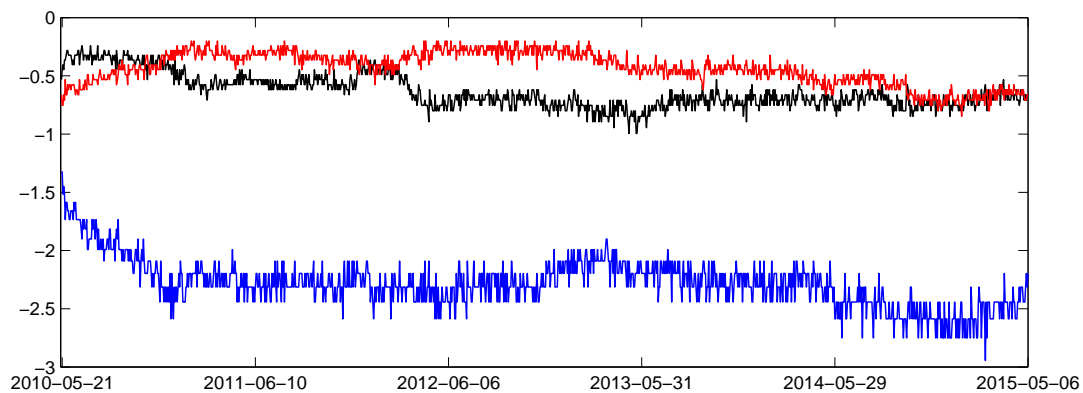
³On April 16th, 2012, YouGov removed the United Kingdom Independence Party (UKIP) from the residual category and started recording it as a separate time series. However, as is the case for the Scottish Nationalist/Plaid Cymru, the support for UKIP is quite low, and hence we do not model either of these series explicitly.

Figure 1: Time series plots of data 2010-05-14 – 2015-05-06

(a) Percentage



(b) Logit transform



Note: Black line is Conservative (CP), red line is Labour (LP), and blue line is Liberal Democrats (LD).

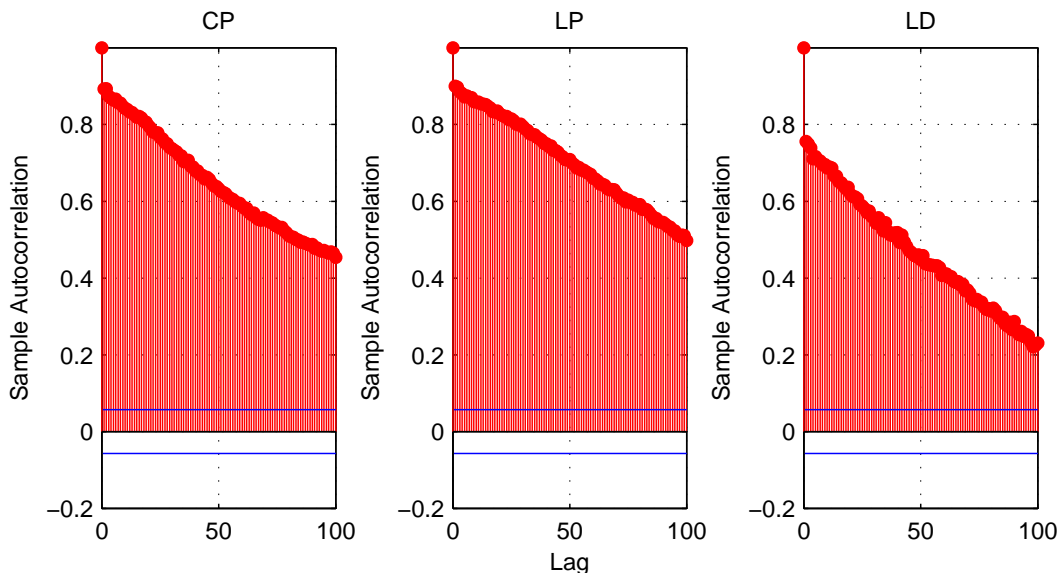
Table 2 presents summary statistics for the logit transformed data. The original data and the logit transform of the data are shown in Figures 1(a) and 1(b), respectively.

As mentioned earlier, the fractional integration characteristic of political opinion polling data has been well established in the literature for monthly and quarterly data. To add to this body of literature, we computed the sample autocorrelation functions and estimated spectral density functions of each of the three series, and we display these in Figures 2(a) and 2(b), respectively. For a fractionally integrated time series, we would expect that the sample autocorrelation functions decay very slowly (hyperbolically, as opposed to geometrically) and that the estimated spectral density functions have mass concentrated near the origin. See, e.g., Baillie (1996) for a review covering these properties. Indeed, both these features appear clearly in Figure 2.

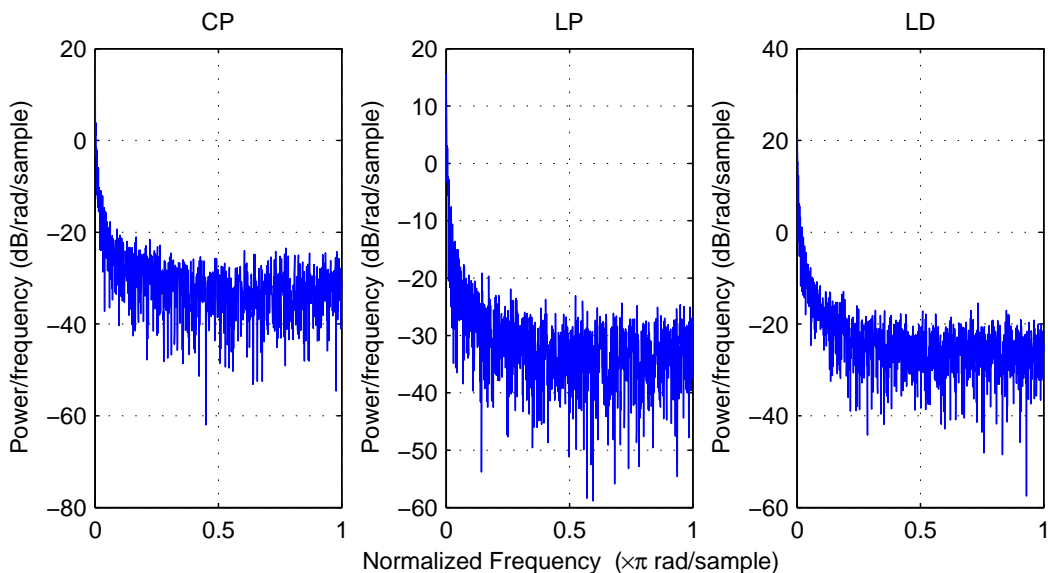
Finally, for each univariate series we have computed semiparametric estimates of the fractional integration parameter, d , that do not rely on the specification of a parametric model or lag structure. Specifically, we computed the exact local Whittle (ELW) estimates

Figure 2: Serial dependence structure

(a) Sample autocorrelation functions



(b) Estimated spectral density functions



of Shimotsu and Phillips (2005), which are displayed in the last three columns in Table 2 for three different choices of bandwidth parameter, $m = T^{0.6}$, $m = T^{0.7}$, and $m = T^{0.8}$ (when this is not an integer, the integer part of the result is used). In each case, the asymptotic standard error of the estimate is $(4m)^{-1/2}$, so for example when $m = T^{0.6}$ the asymptotic standard error of the estimate is 0.059. The results from the ELW estimates suggest that each series is fractionally integrated with a fractional integration parameter that is statistically significantly less than one. More generally, the estimates are in line with estimates from the literature, e.g., Box-Steffensmeier and Smith (1996), Byers et al. (1997), Dolado et al. (2002, 2003), and Jones et al. (2014), where estimates around 0.6 – 0.8 are commonly found

for polling data at the monthly frequency. The evidence presented here thus shows also fractional integration characteristics for the political opinion polls at the daily frequency, as suggested by the theoretical arguments discussed earlier.

3 Econometric methodology: FCVAR model

Our analysis applies the FCVAR model of [Johansen \(2008\)](#) and [Johansen and Nielsen \(2012\)](#). This model is a generalization of [Johansen's \(1995\)](#) CVAR model to allow for fractionally integrated and fractionally cointegrated time series.

3.1 Variants of the FCVAR model and interpretations

For a time series Y_t of dimension p , the well-known CVAR model with a so-called “restricted constant” term is given in error correction form as

$$\Delta Y_t = \alpha(\beta' Y_{t-1} + \rho') + \sum_{i=1}^k \Gamma_i \Delta Y_{t-i} + \varepsilon_t = \alpha L(\beta' Y_t + \rho') + \sum_{i=1}^k \Gamma_i L^i \Delta Y_t + \varepsilon_t, \quad (4)$$

where, as usual, ε_t is a p -dimensional independent and identically distributed error term with mean zero and covariance matrix Ω . The simplest way to derive the FCVAR model from the CVAR model is to replace the difference and lag operators, Δ and $L = 1 - \Delta$, in (4) by their fractional counterparts, Δ^b and $L_b = 1 - \Delta^b$, respectively, and apply the resulting model to $Y_t = \Delta^{d-b} X_t$. We then obtain⁴

$$\Delta^d X_t = \alpha \Delta^{d-b} L_b (\beta' X_t + \rho') + \sum_{i=1}^k \Gamma_i \Delta^d L_b^i X_t + \varepsilon_t, \quad (5)$$

where Δ^d is the fractional difference operator, and $L_b = 1 - \Delta^b$ is the fractional lag operator.⁵

Model (5) nests [Johansen's \(1995\)](#) CVAR model in (4) as the special case $d = b = 1$. Some of the parameters are well-known from the CVAR model and these have the usual interpretations also in the FCVAR model. The most important of these are the long-run parameters α and β , which are $p \times r$ matrices with $0 \leq r \leq p$, and ρ , which is an r -vector. The rank r is termed the cointegration, or sometimes cofractional, rank. The columns of β constitute the r cointegration (cofractional) vectors, such that $\beta' X_t$ are the cointegrating combinations of the variables in the system, i.e. the long-run equilibrium relations. The parameters in α are the adjustment or loading coefficients which represent the speed of adjustment towards equilibrium for each of the variables. The restricted constant term ρ is interpreted as the mean level of the long-run equilibria $\beta' X_t$ when these are stationary. The short-run dynamics of the variables are governed by the parameters $(\Gamma_1, \dots, \Gamma_k)$ in the autoregressive augmentation.

The FCVAR model has two additional parameters compared with the CVAR model, namely the fractional parameters d and b . Here, d denotes the fractional integration order

⁴In principle, the restricted constant term should be included as $\rho' \pi_t(1)$, where $\pi_t(1)$ denotes the coefficient in (1). This is mathematically convenient, but makes no difference in terms of the practical implementation because the infinite summation in (1) needs to be truncated in practice.

⁵Both the fractional difference and fractional lag operators are defined in terms of their binomial expansion in the lag operator, L , as in (1). Note that the expansion of L_b has no term in L^0 and thus only lagged disequilibrium errors appear in (5).

of the observable time series, while the parameter b determines the degree of fractional cointegration, i.e. the reduction in fractional integration order of $\beta'X_t$ compared to X_t itself. Both fractional parameters are estimated jointly with the other parameters, see Section 3.2. The FCVAR model (5) thus has the same main structure as the standard CVAR model (4), in that it allows for modeling of both cointegration and adjustment towards equilibrium, but is more general since it accommodates fractional integration and cointegration.

We note that the fractional difference as defined in (1) is an infinite series, but any observed sample will include only a finite number of observations. This makes calculation of the fractional differences as defined in (1) impossible. In practice, therefore, the summation in (1) would need to be truncated at $n = t - 1$, and the bias introduced by application of such a truncation is analyzed by [Johansen and Nielsen \(2015\)](#) using higher-order expansions in a simpler model. They show, albeit in a simpler model, that this bias can be avoided by including a level parameter, μ , that shifts each of the series by a constant. We follow this suggestion and also consider the unobserved components formulation

$$X_t = \mu + X_t^0, \quad \Delta^d X_t^0 = \alpha \Delta^{d-b} L_b \beta' X_t^0 + \sum_{i=1}^k \Gamma_i \Delta^d L_b^i X_t^0 + \varepsilon_t, \quad (6)$$

from which we easily derive the model

$$\Delta^d (X_t - \mu) = \alpha \beta' \Delta^{d-b} L_b (X_t - \mu) + \sum_{i=1}^k \Gamma_i \Delta^d L_b^i (X_t - \mu) + \varepsilon_t. \quad (7)$$

The formulation (7) includes the restricted constant, which may be obtained as $\rho' = -\beta' \mu$. More generally, the level parameter μ in (7) is meant to accommodate a non-zero starting point for the first observation on the process, i.e., for X_1 .

Our forecasting analysis applies the versions of the FCVAR model given in (5) and (7) and we provide comparisons with the CVAR model in (4) as our main benchmark model. Following the work of [Jones et al. \(2014\)](#) we also consider the sub-models obtained by setting $d = b$ in (5) and (7), which results in disequilibrium errors that are $I(0)$. Thus, the four variants of the FCVAR model that we consider are

1. FCVAR $_{d,b,\rho}$: (5) with two fractional parameters (d, b) and a restricted constant (ρ),
2. FCVAR $_{d,b,\mu}$: (7) with two fractional parameters (d, b) and a level parameter (μ),
3. FCVAR $_{d=b,\rho}$: (5) with one fractional parameter (d) and a restricted constant (ρ),
4. FCVAR $_{d=b,\mu}$: (7) with one fractional parameter (d) and a level parameter (μ).

In each model the fractional parameters are estimated as described in the next subsection, possibly with the restriction $d = b$ imposed as appropriate.

3.2 Maximum likelihood estimation

It is assumed that a sample of length $T + N$ is available on X_t , where N denotes the number of observations used for conditioning, for details see [Johansen and Nielsen \(2012, 2015\)](#). The models (5) and (7) are estimated by conditional maximum likelihood, conditional on N initial values. For the standard CVAR model, the arguments are well-known and choosing $N = k + 1$ leads to reduced rank regression estimation. For the FCVAR model, we proceed

similarly by maximizing the conditional likelihood function

$$\log L_T(\lambda) = -\frac{Tp}{2}(\log(2\pi) + 1) - \frac{T}{2} \log \det \left\{ T^{-1} \sum_{t=N+1}^{T+N} \varepsilon_t(\lambda) \varepsilon_t(\lambda)' \right\}, \quad (8)$$

where the residuals are defined as

$$\varepsilon_t(\lambda) = \Delta^d X_t - \alpha \Delta^{d-b} L_b(\beta' X_t + \rho') - \sum_{i=1}^k \Gamma_i \Delta^d L_b^i X_t, \quad \lambda = (d, b, \alpha, \beta, \Gamma_i, \rho), \quad (9)$$

for model (5), and hence also for sub-models of model (5), such as that obtained under the restriction $d = b$. For model (7) the residuals are

$$\varepsilon_t(\lambda) = \Delta^d (X_t - \mu) - \alpha \beta' \Delta^{d-b} L_b(X_t - \mu) - \sum_{i=1}^k \Gamma_i \Delta^d L_b^i (X_t - \mu), \quad \lambda = (d, b, \alpha, \beta, \Gamma_i, \mu), \quad (10)$$

and similarly for sub-models of model (7).

It is shown in [Johansen and Nielsen \(2012\)](#) how, for fixed (d, b) , the estimation of model (5) reduces to regression and reduced rank regression as in [Johansen \(1995\)](#). In this way the parameters $(\alpha, \beta, \Gamma_i, \rho)$ can be concentrated out of the likelihood function, and numerical optimization is only needed to optimize the profile likelihood function over the two fractional parameters, d and b . In model (7) we can similarly concentrate the parameters $(\alpha, \beta, \Gamma_i)$ out of the likelihood function resulting in numerical optimization over (d, b, μ) , thus making the estimation of model (7) somewhat more involved numerically than that of model (5).

The asymptotic analysis of the FCVAR model is provided in [Johansen and Nielsen \(2012\)](#). For model (5), [Johansen and Nielsen \(2012\)](#) show that asymptotic theory is standard when $b < 0.5$, and for the case $b > 0.5$ asymptotic theory is non-standard and involves fractional Brownian motion of type II. Specifically, when $b > 0.5$, [Johansen and Nielsen \(2012\)](#) show that under i.i.d. errors with suitable moment conditions, the conditional maximum likelihood parameter estimates $(\hat{d}, \hat{b}, \hat{\alpha}, \hat{\Gamma}_1, \dots, \hat{\Gamma}_k)$ are asymptotically Gaussian, while $(\hat{\beta}, \hat{\rho})$ are locally asymptotically mixed normal. These results allow asymptotically standard (chi-squared) inference on all parameters of the model, including the cointegrating relations and orders of fractionality, using quasi-likelihood ratio tests. As in the CVAR model, see [Johansen \(1995\)](#), the same results hold for the same parameters in the model (7), whereas the asymptotic distribution theory for the remaining parameter, μ , is currently unknown.

Likelihood ratio (trace-type) tests for cointegration rank can be calculated as well, and hypotheses on the cointegration rank can be tested in the same way as in the CVAR model. In the FCVAR model, the asymptotic distribution of the tests for cointegration rank depends on the unknown (true value of the) scalar parameter b , which complicates empirical analysis compared to the CVAR model. However, the distribution can be simulated on a case-by-case basis, or the computer programs by [MacKinnon and Nielsen \(2014\)](#) can be used to obtain either critical values of P values for the rank test. The calculation of maximum likelihood estimators and test statistics is discussed in detail in [Johansen and Nielsen \(2012\)](#) and [Nielsen and Popiel \(2014\)](#), with the latter providing Matlab computer programs that we apply in our empirical analysis.

3.3 Forecasting from the FCVAR model

We now discuss how to forecast the (logit transformed) polling data, that is X_t , from the FCVAR model (since the CVAR model is a special case obtained as $d = b = 1$, forecasts from that model are derived in the same way). Because the model is autoregressive, the best linear predictor takes a simple form and is relatively straightforward to calculate. We first note that

$$\Delta^d(X_{t+1} - \mu) = X_{t+1} - \mu - (X_{t+1} - \mu) + \Delta^d(X_{t+1} - \mu) = X_{t+1} - \mu - L_d(X_{t+1} - \mu)$$

and then rearrange (7) as

$$X_{t+1} = \mu + L_d(X_{t+1} - \mu) + \alpha\beta'\Delta^{d-b}L_b(X_{t+1} - \mu) + \sum_{i=1}^k \Gamma_i \Delta^d L_b^i(X_{t+1} - \mu) + \varepsilon_{t+1}. \quad (11)$$

Since $L_b = 1 - \Delta^b$ is a lag operator, so that $L_b^i X_{t+1}$ is known at time t for $i \geq 1$, this equation can be used as the basis to calculate forecasts from the model.

We let conditional expectation given the information set at time t be denoted $E_t(\cdot)$, and the best linear predictor forecast of any variable Z_{t+1} given information available at time t be denoted $\hat{Z}_{t+1|t} = E_t(Z_{t+1})$. Clearly, we then have that the forecast of the innovation for period $t + 1$ at time t is $\hat{\varepsilon}_{t+1|t} = E_t(\varepsilon_{t+1}) = 0$, and $\hat{X}_{t+1|t}$ is then easily found from (11). Inserting also coefficient estimates based on data available up to time t , denoted⁶ $(\hat{d}, \hat{b}, \hat{\mu}, \hat{\alpha}, \hat{\beta}, \hat{\Gamma}_1, \dots, \hat{\Gamma}_k)$, we have that

$$\hat{X}_{t+1|t} = \hat{\mu} + L_{\hat{d}}(X_{t+1} - \hat{\mu}) + \hat{\alpha}\hat{\beta}'\Delta^{\hat{d}-\hat{b}}L_{\hat{b}}(X_{t+1} - \hat{\mu}) + \sum_{i=1}^k \hat{\Gamma}_i \Delta^{\hat{d}} L_{\hat{b}}^i(X_{t+1} - \hat{\mu}). \quad (12)$$

This defines the one-step ahead forecast of X_{t+1} given information at time t .

Multi-period ahead forecasts can be generated recursively. That is, to calculate the h -step ahead forecast, we first generalize (12) as

$$\hat{X}_{t+j|t} = \hat{\mu} + L_{\hat{d}}(\hat{X}_{t+j|t} - \hat{\mu}) + \hat{\alpha}\hat{\beta}'\Delta^{\hat{d}-\hat{b}}L_{\hat{b}}(\hat{X}_{t+j|t} - \hat{\mu}) + \sum_{i=1}^k \hat{\Gamma}_i \Delta^{\hat{d}} L_{\hat{b}}^i(\hat{X}_{t+j|t} - \hat{\mu}), \quad (13)$$

where $\hat{X}_{s|t} = X_s$ for $s \leq t$. Then forecasts are calculated recursively from (13) for $j = 1, 2, \dots, h$ to generate h -step ahead forecasts, $\hat{X}_{t+h|t}$.

Clearly, one-step ahead and h -step ahead forecasts for the model (5) with a restricted constant term instead of the level parameter can be calculated entirely analogously. We will apply the forecasts (12) and (13) for both models (5) and (7) in our analysis below for several forecast horizons, h .

4 Estimation results

Before we move on to the forecasting analysis, we briefly discuss some estimation results which are presented in Tables 3–6. Each table shows FCVAR estimation results for one

⁶To emphasize that these estimates are based on data available at time t , they could be denoted by a subscript t . However, to avoid cluttering the notation we omit this subscript and let it be understood in the sequel.

Table 3: FCVAR $_{d,b,\rho}$ estimation results

Model:

$$\Delta^{\hat{d}} \begin{bmatrix} \text{CP}_t \\ \text{LP}_t \\ \text{LD}_t \end{bmatrix} = \hat{\alpha} \left(\hat{\beta}' \Delta^{\hat{d}-\hat{b}} L_{\hat{b}} \begin{bmatrix} \text{CP}_t \\ \text{LP}_t \\ \text{LD}_t \end{bmatrix} + \hat{\rho}' \right) + \hat{\Gamma} \Delta^{\hat{d}} L_{\hat{b}} \begin{bmatrix} \text{CP}_t \\ \text{LP}_t \\ \text{LD}_t \end{bmatrix} + \hat{\varepsilon}_t$$

Parameters:

$$\hat{d} = 0.774, \quad \hat{b} = 0.094, \quad \hat{\alpha} = \begin{bmatrix} 0.532 & -0.810 \\ 0.405 & 0.034 \\ 5.517 & 10.748 \end{bmatrix}, \quad \hat{\beta} = \begin{bmatrix} 1.000 & 0.000 \\ 0.000 & 1.000 \\ 2.247 & -1.962 \end{bmatrix},$$

$$\hat{\rho} = [0.019 \quad -1.536], \quad \hat{\Gamma} = \begin{bmatrix} -6.011 & 0.970 & -2.485 \\ -0.250 & -5.686 & -0.684 \\ -5.875 & -10.224 & 2.080 \end{bmatrix},$$

$$Q_{\hat{\varepsilon}} = 50.111, \quad \log(\mathcal{L}) = 4898.335$$

(0.625)

Notes: The table shows FCVAR estimation results for model (5) with two fractional parameters, (d, b) , and a restricted constant term, ρ . The multivariate Portmanteau Q-test for serial correlation up to order six in the residuals is reported as $Q_{\hat{\varepsilon}}$ and the maximized log-likelihood value is reported as $\log(\mathcal{L})$. Standard errors are in parentheses below \hat{d} and \hat{b} and P values are in parentheses below $Q_{\hat{\varepsilon}}$. The sample size is $T + N = 1227$ and the first $N = 20$ observations are used as initial values.

of the four variants considered. Specifically, Table 3 shows results for model (5) with two fractional parameters, (d, b) , and a restricted constant term, ρ , Table 4 shows results for model (7) with two fractional parameters, (d, b) , and the level parameter, μ , while Tables 5 and 6 show the corresponding results for the models with only one fractional parameter, $d = b$. In addition, the multivariate Portmanteau Q-test for serial correlation up to order six in the residuals is reported as $Q_{\hat{\varepsilon}}$ and the maximized log-likelihood value is reported as $\log(\mathcal{L})$. Standard errors are in parentheses below \hat{d} and \hat{b} and P values are in parentheses below $Q_{\hat{\varepsilon}}$.

Since it is infeasible to present results for all the many models and training sets on which our forecasting analysis is based, the results in Tables 3–6 are based simply on the full sample of size $T + N = 1227$. In the models with a restricted constant term in Tables 3 and 5, the first $N = 20$ observations are used as initial values in the estimation, while in the models with a level parameter in Tables 4 and 6, we follow Johansen and Nielsen (2015) and set $N = 0$.

For each model we initially chose the lag order by the Bayesian information criterion, and then conducted cointegration rank tests. The model was then estimated and the residuals tested for white noise (multivariate Portmanteau Q-test reported), which was not rejected for any of the four models with a lag order of $k = 1$ and cointegration rank $r = 2$. Hence we conclude that the models appear correctly specified.

Looking at the estimated parameter values in the four tables, we first note that \hat{d} is very strongly significantly different from one and that \hat{b} is strongly significantly different from both zero and one in all models. This suggests very clearly that the FCVAR model is more

Table 4: FCVAR_{d,b,μ} estimation results

Model:

$$\Delta^{\hat{d}} \left(\begin{bmatrix} \text{CP}_t \\ \text{LP}_t \\ \text{LD}_t \end{bmatrix} - \hat{\mu} \right) = \hat{\alpha} \hat{\beta}' \Delta^{\hat{d}-\hat{b}} L_{\hat{b}} \left(\begin{bmatrix} \text{CP}_t \\ \text{LP}_t \\ \text{LD}_t \end{bmatrix} - \hat{\mu} \right) + \hat{\Gamma} \Delta^{\hat{d}} L_{\hat{b}} \left(\begin{bmatrix} \text{CP}_t \\ \text{LP}_t \\ \text{LD}_t \end{bmatrix} - \hat{\mu} \right) + \hat{\varepsilon}_t$$

Parameters:

$$\hat{d} = 0.624, \quad \hat{b} = 0.273, \quad \hat{\alpha} = \begin{bmatrix} 0.009 & -0.156 \\ 0.448 & 0.647 \\ 0.064 & -0.354 \end{bmatrix}, \quad \hat{\beta} = \begin{bmatrix} 1.000 & 0.000 \\ 0.000 & 1.000 \\ 0.947 & -0.748 \end{bmatrix},$$

$$\hat{\mu} = \begin{bmatrix} -0.396 \\ -0.660 \\ -1.597 \end{bmatrix}, \quad \hat{\Gamma} = \begin{bmatrix} -1.278 & 0.283 & -0.033 \\ -0.367 & -2.105 & 0.074 \\ -0.308 & 0.478 & -1.924 \end{bmatrix},$$

$$Q_{\hat{\varepsilon}} = 57.148, \quad \log(\mathcal{L}) = 4965.089$$

(0.026) (0.084) (0.359)

Notes: The table shows FCVAR estimation results for model (7) with two fractional parameters, (d, b) , and the level parameter, μ . The multivariate Portmanteau Q-test for serial correlation up to order six in the residuals is reported as $Q_{\hat{\varepsilon}}$ and the maximized log-likelihood value is reported as $\log(\mathcal{L})$. Standard errors are in parentheses below \hat{d} and \hat{b} and P values are in parentheses below $Q_{\hat{\varepsilon}}$. The sample size is $T + N = 1227$ and $N = 0$ observations are used as initial values.

Table 5: FCVAR_{d=b,ρ} estimation results

Model:

$$\Delta^{\hat{d}} \begin{bmatrix} \text{CP}_t \\ \text{LP}_t \\ \text{LD}_t \end{bmatrix} = \hat{\alpha} \left(\hat{\beta}' L_{\hat{d}} \begin{bmatrix} \text{CP}_t \\ \text{LP}_t \\ \text{LD}_t \end{bmatrix} + \hat{\rho}' \right) + \hat{\Gamma} \Delta^{\hat{d}} L_{\hat{d}} \begin{bmatrix} \text{CP}_t \\ \text{LP}_t \\ \text{LD}_t \end{bmatrix} + \hat{\varepsilon}_t$$

Parameters:

$$\hat{d} = 0.627, \quad \hat{\alpha} = \begin{bmatrix} -0.024 & -0.024 \\ -0.013 & 0.005 \\ -0.100 & 0.132 \end{bmatrix}, \quad \hat{\beta} = \begin{bmatrix} 1.000 & 0.000 \\ 0.000 & 1.000 \\ 0.370 & -0.922 \end{bmatrix},$$

$$\hat{\rho} = [1.955 \quad -1.652], \quad \hat{\Gamma} = \begin{bmatrix} -0.552 & 0.054 & 0.029 \\ 0.049 & -0.587 & 0.031 \\ 0.049 & -0.031 & -0.554 \end{bmatrix},$$

$$Q_{\hat{\varepsilon}} = 76.692, \quad \log(\mathcal{L}) = 4875.513$$

(0.013) (0.023)

Notes: The table shows FCVAR estimation results for model (5) with one fractional parameter, $d = b$, and a restricted constant term, ρ . The multivariate Portmanteau Q-test for serial correlation up to order six in the residuals is reported as $Q_{\hat{\varepsilon}}$ and the maximized log-likelihood value is reported as $\log(\mathcal{L})$. Standard errors are in parentheses below \hat{d} and \hat{b} and P values are in parentheses below $Q_{\hat{\varepsilon}}$. The sample size is $T + N = 1227$ and the first $N = 20$ observations are used as initial values.

Table 6: FCVAR $_{d=b,\mu}$ estimation results

Model:

$$\Delta^{\hat{d}} \left(\begin{bmatrix} \text{CP}_t \\ \text{LP}_t \\ \text{LD}_t \end{bmatrix} - \hat{\mu} \right) = \hat{\alpha} \hat{\beta}' L_{\hat{d}} \left(\begin{bmatrix} \text{CP}_t \\ \text{LP}_t \\ \text{LD}_t \end{bmatrix} - \hat{\mu} \right) + \hat{\Gamma} \Delta^{\hat{d}} L_{\hat{d}} \left(\begin{bmatrix} \text{CP}_t \\ \text{LP}_t \\ \text{LD}_t \end{bmatrix} - \hat{\mu} \right) + \hat{\varepsilon}_t$$

Parameters:

$$\begin{aligned} \hat{d} &= 0.572, & \hat{\alpha} &= \begin{bmatrix} -0.022 & -0.018 \\ 0.079 & 0.055 \\ -0.058 & -0.048 \end{bmatrix}, & \hat{\beta} &= \begin{bmatrix} 1.000 & -0.000 \\ 0.000 & 1.000 \\ 2.403 & -3.679 \end{bmatrix}, \\ & (0.015) & & & & \\ \hat{\mu} &= \begin{bmatrix} -0.411 \\ -0.653 \\ -1.574 \end{bmatrix}, & \hat{\Gamma} &= \begin{bmatrix} -0.497 & 0.075 & 0.027 \\ -0.032 & -0.652 & 0.022 \\ -0.061 & 0.119 & -0.694 \end{bmatrix}, \\ & & Q_{\hat{\varepsilon}} &= 60.163, & \log(\mathcal{L}) &= 4962.473 \\ & & & (0.263) & & \end{aligned}$$

Notes: The table shows FCVAR estimation results for model (7) with one fractional parameter, $d = b$, and the level parameter, μ . The multivariate Portmanteau Q-test for serial correlation up to order six in the residuals is reported as $Q_{\hat{\varepsilon}}$ and the maximized log-likelihood value is reported as $\log(\mathcal{L})$. Standard errors are in parentheses below \hat{d} and $\hat{\beta}$ and P values are in parentheses below $Q_{\hat{\varepsilon}}$. The sample size is $T + N = 1227$ and $N = 0$ observations are used as initial values.

appropriate for this data than the non-fractional CVAR model since the latter has $d = b = 1$ imposed. Comparing across the models, it appears that the estimates of d are fairly close, ranging from 0.57 to 0.77, whereas the estimates of b are quite different in the models with $d \neq b$ in Tables 3 and 4.

Further comparison across models leads to consideration of the likelihood ratio test statistic for the null hypothesis that $d = b$. That is, for the models with a restricted constant term (and $N = 20$), we can test the null of the model in Table 5 against the alternative of the model in Table 3, and for the models with a level parameter (and $N = 0$), we can test the null of the model in Table 6 against the alternative of the model in Table 4. In the first case, the likelihood ratio test statistic is 35.002, and in the second case, the likelihood ratio test statistic is 0.84. In both cases, this is asymptotically chi-squared distributed with one degree of freedom, so the conclusions of these tests differ and consequently we proceed with the consideration of all four models.

5 Forecast comparison

This section presents and discusses the forecasting procedure and results. The four variants of the FCVAR model presented above are evaluated on their forecasting ability relative to a group of popular competing models. The CVAR model (4) serves as the multivariate benchmark model. The main univariate benchmark is the ARFIMA(p, d, q) model,

$$A(L)\Delta^d(X_t - \mu) = B(L)\varepsilon_t, \quad (14)$$

where $A(L)$ and $B(L)$ are the autoregressive and moving average polynomials, satisfying standard regularity conditions. A special case of (14) is the simple ARFIMA(0, d , 0) model,

which is also included because it was found by, e.g., [Byers et al. \(1997\)](#) and [Dolado et al. \(2002\)](#), to fit (monthly) UK polling data well. Finally, another special case of (14) is the standard ARMA(p, q) model, which we include as the classical univariate benchmark. Estimation of the univariate models is by minimization of the conditional sum-of-squares, see [Box and Jenkins \(1970\)](#) for ARMA models and [Hualde and Robinson \(2011\)](#) and [Nielsen \(2015\)](#) for ARFIMA models, while forecasting is done using the best linear predictor which appears standard for these models.

5.1 Forecasting methodology

The forecasting procedure applies the standard rolling window and recursive forecasting schemes to examine forecasting accuracy. The main distinction between the two schemes is how they select the training sets used for estimating each model to produce forecasts. The rolling window scheme uses a fixed training set length (usually referred to as a window) that moves across the data set. The recursive scheme uses an expanding training set length with a fixed start date at the beginning of the data set. In order to assess the forecasting capability of each model, it is necessary to generate predictions from a sufficiently large number of training sets used to estimate each model, and it is preferable that each training set is long enough for reliable estimation and forecasting.

For the rolling window scheme, each statistical model uses training sets with a fixed length of $T + N = 600$ observations, approximately half the length of the data set. For the recursive scheme, the first training set includes $T + N = 600$ observations and each subsequent training set includes one more observation, until the last training set which has $T + N = 1227 - h$ observations, where h is the forecast horizon. This implies that for both forecasting schemes, the total number of training sets is $1227 - 600 - h + 1 = 628 - h$, and the first training set is the same for both procedures. For the FCVAR models with a restricted constant term we use the first $N = 20$ observations of each training set as initial values, and for the FCVAR models with the level parameter we follow [Johansen and Nielsen \(2015\)](#) and use $N = 0$ initial values. The CVAR model applies estimation conditional on $N = k + 1$ initial values, such that maximum likelihood estimation reduces to reduced rank regression.

The forecasting programs use all applicable model specification criteria and tests consistently across both the multivariate and univariate model types, and all rejection rules for statistical hypothesis testing are conducted at the five percent level of significance. For all models, the model specification is based on the very first training set, and the same model specification is then applied to all training sets. That is, we maintain the same lag orders and cointegration ranks for all training sets, but all the parameters of the models are re-estimated for each training set before forecasts are calculated.

The multivariate specifications involve first selecting the lag order, k . Lag order selection is initially based on the Bayesian information criterion (BIC). Given the lag order, cointegration rank tests are performed, which determine the number of cointegrating relations, r , for each model. The multivariate model is then estimated using these values of k and r . In the next step, the program performs a multivariate Q-test for white noise up to order six on the residuals. If the white noise test rejects, then another lag is added and the rank test, estimation, and white noise test are repeated in sequence until the program fails to reject white noise for the residuals. The univariate specification differs from the multivariate only in that two lag orders, p and q , need to be selected conditional on the univariate white noise

test failing to reject.

Following the completion of the specification algorithm, the forecasting program estimates the model for all training sets and uses the estimated model parameters to generate h -step ahead forecasts for each time series. All multivariate and univariate models considered generate forecasts recursively, see Section 3.3. The ARFIMA(0, d , 0) model is included in the portfolio due to its popularity for political opinion poll data. Its fixed lag orders make it the only model in the portfolio with lag orders not determined by a decision rule in the forecasting program.

Forecasts are generated for seven out-of-sample horizons, h : 1, 5, 10, 15, 20, 25, and 50 steps ahead. As mentioned above, the number of training sets is different for each forecast horizon, and we denote this number M_h . The models are ranked based on the multivariate (system) root mean squared forecast error,

$$\text{RMSFE}_{\text{sys}} = \sqrt{\frac{1}{pM_h} \sum_{i=1}^p \sum_{j=1}^{M_h} \left(\hat{X}_{i,T_j+h|T_j} - X_{i,T_j+h} \right)^2}, \quad (15)$$

as well as the univariate root mean squared forecast errors for each series,

$$\text{RMSFE}_i = \sqrt{\frac{1}{M_h} \sum_{j=1}^{M_h} \left(\hat{X}_{i,T_j+h|T_j} - X_{i,T_j+h} \right)^2}, \quad (16)$$

where, in both cases, h denotes the forecast horizon, $p = 3$ is the number of series, i.e., the dimension of the multivariate system, $M_h = 628 - h$ is the number of training sets, and T_j is the terminal date of training set j . The individual RMSFE_i ($i = \text{CP}, \text{LP}, \text{LD}$) measures the typical magnitude of forecast errors for each individual time series, while the $\text{RMSFE}_{\text{sys}}$ measures the typical magnitude of all forecast errors produced by each model.

5.2 Forecasting results

This section discusses the forecast performance results, and in particular analyses in detail the relative performance of the FCVAR model to the multivariate benchmark CVAR model, and concludes with several figures of forecasts generated by all models in the portfolio.

Tables 7 and 8 report the RMSFE_i ($i = \text{CP}, \text{LP}, \text{LD}$) and $\text{RMSFE}_{\text{sys}}$ values for the rolling and recursive schemes, respectively. Numbers in parentheses are the corresponding rankings at each individual forecast horizon. The models are ranked based on the $\text{RMSFE}_{\text{sys}}$ because for each model it provides a single measurement of forecast accuracy for all three time series.

We begin the assessment of the forecast accuracy with a discussion of the one-step ahead forecasts. This seems natural prior to examining performance at other subjectively selected horizons that may be of interest in any given application. In the context of political opinion polls, one can easily imagine the relevance of forecasting poll standings or vote shares at many different horizons.

According to the recursive scheme, all four variants of the FCVAR model outperform all competing models at the one-step ahead forecasting horizon. According to the rolling window scheme, three of the four FCVAR variants outperform all competing models, and the $\text{FCVAR}_{d=b,\rho}$ model loses to the ARFIMA(p, d, q) model by a very small margin. Both

forecasting schemes rank a variant of the FCVAR model with two fractional parameters as the top performing model. The rolling window scheme ranks the $\text{FCVAR}_{d,b,\mu}$ and $\text{FCVAR}_{d,b,\rho}$ models (jointly) first and the $\text{FCVAR}_{d=b,\mu}$ model second, while the recursive scheme ranks the $\text{FCVAR}_{d,b,\rho}$ model first and the $\text{FCVAR}_{d=b,\rho}$ model second. An important observation is that the third and fourth ranked FCVAR specifications are very close in performance to the top performing variant, showing that the reliability of one-step ahead forecasts generated by the FCVAR model is robust to the number of fractional parameters and type of deterministic terms used in the specification, at least for this data set.

The longer forecast horizons considered, 5, 10, 15, 20, 25 and 50 steps ahead, deliver results that are in agreement with the findings for the one-step ahead horizon. The model rankings across all forecast horizons determine that the accuracy of both short, medium and long term forecasts generated by the FCVAR model is better than the other models in the portfolio. This can be seen from the fact that for 7 of 7 cases (1 to 50 steps) in both the rolling and the recursive schemes, the top two performing models are always variants of the FCVAR model and three of the top four models are always variants of the FCVAR model. Furthermore, with the exception of three forecast horizons (1, 5 and 50 steps) in the rolling window scheme and one forecast horizon (50 steps) in the recursive scheme, all four variants of the FCVAR model are ranked as the top four models. The exception occurs when only one variant of the FCVAR model underperforms by a small margin relative to the ARFIMA model. Overall, this evidence provides strong support for the application of the FCVAR model for forecasting next day (one-step), next week (5-steps), and all the way up to ten weeks ahead (50-steps) poll standings.

The results for both forecasting schemes suggest that the FCVAR model with two fractional parameters produces the smallest average forecast errors. A variant of the FCVAR model with two fractional parameters always outperforms both sub-models, the $\text{FCVAR}_{d=b,\rho}$ and $\text{FCVAR}_{d=b,\mu}$. The recursive scheme determines the $\text{FCVAR}_{d,b,\rho}$ model as the absolute favorite at all forecast horizons, while the rolling window scheme ranks the $\text{FCVAR}_{d,b,\mu}$ and $\text{FCVAR}_{d,b,\rho}$ models as the top two performing models interchangeably, with the $\text{FCVAR}_{d,b,\mu}$ variant ranked first over the majority of forecast horizons.

Furthermore, the model rankings strongly suggest that the forecasting accuracy of the FCVAR is better than the multivariate benchmark CVAR model. To assess the degree of relative performance, Tables 9 and 10 report the RMSE percentage change,

$$100 \left(\frac{\text{RMSFE}_{\text{sys}}(\text{FCVAR})}{\text{RMSFE}_{\text{sys}}(\text{CVAR})} - 1 \right), \quad (17)$$

between the two models for the rolling and recursive schemes, respectively. Negative values favor the FCVAR model and positive values favor the CVAR model.

Table 9 and 10 show strong forecast accuracy improvement for all variants of the FCVAR model relative to the CVAR model. At the one-step ahead horizon the FCVAR model delivers 4 – 6% and 6 – 8% improvement for the rolling and recursive schemes, respectively. In 56 of 56 cases, all four variants of the FCVAR model perform better than the CVAR model. For the FCVAR models with a restricted constant, the improvements in performance increase substantially as the forecasting horizon increases. The recursive scheme shows 15%, 17%, 19% and 20% improvement for the 15, 20, 25 and 50 step ahead horizons, attained by the FCVAR model with two fractional parameters and a restricted constant. The rolling

Table 7: Root mean squared forecast errors – rolling window forecast scheme

Model	Series	1 step	5 step	10 step	15 step	20 step	25 step	50 step
FCVAR _{d,b,ρ}	CP	0.0562	0.0613	0.0637	0.0656	0.0678	0.0704	0.0781
	LP	0.0499	0.0524	0.0569	0.0609	0.0647	0.0690	0.0902
	LD	0.1134	0.1161	0.1233	0.1298	0.1358	0.1415	0.1650
	System	0.0785 (1)	0.0816 (2)	0.0866 (2)	0.0910 (2)	0.0953 (2)	0.0996 (1)	0.1176 (2)
FCVAR _{d,b,μ}	CP	0.0561	0.0608	0.0632	0.0656	0.0677	0.0702	0.0781
	LP	0.0505	0.0549	0.0607	0.0669	0.0723	0.0781	0.1030
	LD	0.1130	0.1145	0.1201	0.1249	0.1295	0.1350	0.1543
	System	0.0785 (1)	0.0813 (1)	0.0858 (1)	0.0901 (1)	0.0941 (1)	0.0987 (2)	0.1162 (1)
FCVAR _{d=b,ρ}	CP	0.0569	0.0629	0.0664	0.0684	0.0708	0.0729	0.0782
	LP	0.0506	0.0525	0.0569	0.0604	0.0645	0.0687	0.0908
	LD	0.1174	0.1241	0.1329	0.1400	0.1445	0.1499	0.1663
	System	0.0808 (4)	0.0859 (5)	0.0918 (4)	0.0965 (4)	0.1001 (3)	0.1041 (3)	0.1183 (3)
FCVAR _{d=b,μ}	CP	0.0566	0.0634	0.0682	0.0731	0.0783	0.0828	0.1039
	LP	0.0502	0.0540	0.0597	0.0653	0.0705	0.0828	0.1056
	LD	0.1131	0.1169	0.1242	0.1326	0.1379	0.1450	0.1799
	System	0.0786 (2)	0.0829 (3)	0.0888 (3)	0.0952 (3)	0.1002 (4)	0.1076 (4)	0.1345 (6)
CVAR _ρ	CP	0.0605	0.0656	0.0712	0.0748	0.0800	0.0835	0.0953
	LP	0.0547	0.0577	0.0667	0.0744	0.0805	0.0865	0.1114
	LD	0.1200	0.1260	0.1386	0.1489	0.1559	0.1619	0.1828
	System	0.0838 (5)	0.0885 (7)	0.0979 (7)	0.1054 (7)	0.1113 (7)	0.1164 (7)	0.1353 (7)
ARFIMA(0, d, 0)	CP	0.0578	0.0623	0.0664	0.0700	0.0730	0.0758	0.0858
	LP	0.0531	0.0638	0.0746	0.0834	0.0909	0.0977	0.1248
	LD	0.1158	0.1244	0.1335	0.1410	0.1466	0.1521	0.1628
	System	0.0808 (4)	0.0884 (6)	0.0963 (6)	0.1029 (6)	0.1081 (6)	0.1132 (6)	0.1284 (4)
ARFIMA(p, d, q)	CP	0.0578	0.0623	0.0664	0.0700	0.0730	0.0758	0.0860
	LP	0.0531	0.0638	0.0746	0.0834	0.0909	0.0977	0.1252
	LD	0.1133	0.1184	0.1259	0.1328	0.1382	0.1438	0.1637
	System	0.0796 (3)	0.0856 (4)	0.0928 (5)	0.0991 (5)	0.1044 (5)	0.1095 (5)	0.1289 (5)
ARMA(p, q)	CP	0.0664	0.0795	0.0945	0.1051	0.1105	0.1138	0.1186
	LP	0.0641	0.1076	0.1351	0.1491	0.1567	0.1617	0.1759
	LD	0.1373	0.1833	0.1934	0.1966	0.1980	0.1992	0.2054
	System	0.0955 (6)	0.1310 (8)	0.1467 (8)	0.1548 (8)	0.1591 (8)	0.1620 (8)	0.1705 (8)

Notes: The overall performance of each model is measured by the root mean square forecast error of the entire multivariate system. Numbers in parentheses are the corresponding rankings at each individual forecast horizon. The number 1 rank is assigned to the best performing model. The number 8 rank is assigned to the worst performing model. Results are based on h -step ahead forecasts produced using $628-h$ training sets of length 600.

Table 8: Root mean squared forecast errors – recursive window forecast scheme

Model	Series	1 step	5 step	10 step	15 step	20 step	25 step	50 step
FCVAR _{d,b,ρ}	CP	0.0561	0.0616	0.0642	0.0667	0.0697	0.0727	0.0839
	LP	0.0505	0.0548	0.0611	0.0671	0.0726	0.0784	0.1062
	LD	0.1123	0.1144	0.1197	0.1251	0.1292	0.1342	0.1520
	System	0.0781 (1)	0.0814 (1)	0.0860 (1)	0.0906 (1)	0.0946 (1)	0.0991 (1)	0.1175 (1)
FCVAR _{d,b,μ}	CP	0.0563	0.0619	0.0653	0.0688	0.0728	0.0765	0.0905
	LP	0.0499	0.0524	0.0564	0.0602	0.0640	0.0686	0.0939
	LD	0.1150	0.1198	0.1288	0.1381	0.1466	0.1567	0.2024
	System	0.0793 (3)	0.0835 (3)	0.0895 (4)	0.0956 (4)	0.1015 (4)	0.1082 (4)	0.1390 (6)
FCVAR _{d=b,ρ}	CP	0.0563	0.0617	0.0646	0.0672	0.0701	0.0732	0.0846
	LP	0.0507	0.0550	0.0612	0.0673	0.0729	0.0786	0.1058
	LD	0.1144	0.1193	0.1267	0.1338	0.1387	0.1439	0.1614
	System	0.0792 (2)	0.0838 (4)	0.0894 (3)	0.0948 (3)	0.0991 (2)	0.1037 (2)	0.1217 (2)
FCVAR _{d=b,μ}	CP	0.0563	0.0624	0.0662	0.0697	0.0731	0.0766	0.0887
	LP	0.0507	0.0555	0.0622	0.0685	0.0747	0.0808	0.1113
	LD	0.1147	0.1171	0.1236	0.1309	0.1367	0.1441	0.1741
	System	0.0794 (4)	0.0830 (2)	0.0886 (2)	0.0943 (2)	0.0993 (3)	0.1051 (3)	0.1298 (3)
CVAR _ρ	CP	0.0615	0.0684	0.0712	0.0830	0.0909	0.0972	0.1160
	LP	0.0551	0.0595	0.0667	0.0828	0.0926	0.1017	0.1388
	LD	0.1208	0.1222	0.1386	0.1427	0.1500	0.1572	0.1789
	System	0.0845 (7)	0.0878 (7)	0.0979 (7)	0.1066 (7)	0.1145 (7)	0.1218 (7)	0.1469 (7)
ARFIMA(0, d, 0)	CP	0.0586	0.0628	0.0668	0.0706	0.0739	0.0770	0.0881
	LP	0.0515	0.0566	0.0634	0.0692	0.0744	0.0792	0.0991
	LD	0.1181	0.1242	0.1359	0.1468	0.1560	0.1651	0.2007
	System	0.0817 (6)	0.0842 (5)	0.0948 (6)	0.1022 (6)	0.1085 (6)	0.1147 (6)	0.1389 (5)
ARFIMA(p, d, q)	CP	0.0585	0.0627	0.0668	0.0706	0.0739	0.0770	0.0881
	LP	0.0517	0.0566	0.0634	0.0692	0.0744	0.0792	0.0991
	LD	0.1149	0.1190	0.1280	0.1374	0.1451	0.1537	0.1877
	System	0.0802 (5)	0.0843 (6)	0.0910 (5)	0.0977 (5)	0.1034 (5)	0.1093 (5)	0.1327 (4)
ARMA(p, q)	CP	0.0693	0.0836	0.1041	0.1205	0.1309	0.1379	0.1489
	LP	0.0623	0.1035	0.1386	0.1564	0.1657	0.1709	0.1813
	LD	0.1460	0.1483	0.1524	0.1615	0.1656	0.1759	0.2265
	System	0.1000 (8)	0.1150 (8)	0.1333 (8)	0.1473 (8)	0.1549 (8)	0.1624 (8)	0.1883 (8)

Notes: The overall performance of each model is measured by the root mean square forecast error of the entire multivariate system. Numbers in parentheses are the corresponding rankings at each individual forecast horizon. The number 1 rank is assigned to the best performing model. The number 8 rank is assigned to the worst performing model. Results are based on h -step ahead forecasts produced using 628- h training sets of length= 600, 601, ..., 1227 - h .

Table 9: Rolling scheme: $\text{RMSFE}_{\text{sys}}$ percentage change FCVAR vs. CVAR

Model	1 step	5 step	10 step	15 step	20 step	25 step	50 step
$\text{FCVAR}_{d,b,\rho}$	-6.24	-7.80	-11.50	-13.60	-14.43	-14.48	-13.10
$\text{FCVAR}_{d,b,\mu}$	-6.34	-8.18	-12.29	-14.44	-15.45	-15.19	-14.10
$\text{FCVAR}_{d=b,\rho}$	-3.56	-3.01	-6.14	-8.42	-10.10	-10.60	-12.52
$\text{FCVAR}_{d=b,\mu}$	-6.22	-6.39	-9.29	-9.64	-10.00	-7.58	-0.55

Notes: Negative values favor the FCVAR model. Results are based on h -step ahead forecasts produced using $628-h$ training sets of length 600.

Table 10: Recursive scheme: $\text{RMSFE}_{\text{sys}}$ percentage change FCVAR vs. CVAR

Model	1 step	5 step	10 step	15 step	20 step	25 step	50 step
$\text{FCVAR}_{d,b,\rho}$	-7.53	-7.32	-12.13	-15.07	-17.43	-18.66	-20.00
$\text{FCVAR}_{d,b,\mu}$	-6.08	-4.92	-8.53	-10.32	-11.39	-11.17	-5.36
$\text{FCVAR}_{d=b,\rho}$	-6.23	-4.61	-8.65	-11.11	-13.45	-14.88	-17.18
$\text{FCVAR}_{d=b,\mu}$	-6.06	-5.47	-9.50	-11.55	-13.24	-13.68	-11.61

Notes: Negative values favor the FCVAR model. Results are based on h -step ahead forecasts produced using $628-h$ training sets of length= 600, 601, \dots , $1227 - h$.

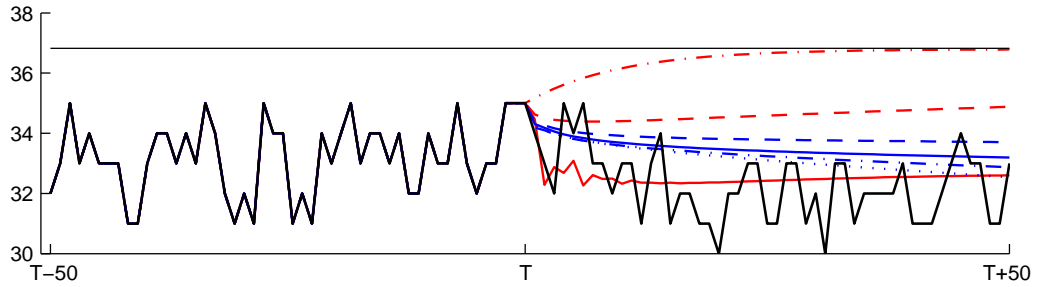
scheme shows up to 15% improvement for the FCVAR models with a restricted constant, and up to 14% improvement for the FCVAR models with a level parameter. In 30 of 56 cases, the FCVAR model delivers more than a 10% reduction in the $\text{RMSFE}_{\text{sys}}$ relative to the CVAR model. Overall the fractional models outperform the non-fractional models, and fractional cointegration substantially improves forecast accuracy, with the gains becoming more pronounced at longer forecast horizons.

To conclude the forecast comparison, examples of forecasts generated by all models in the portfolio are presented in Figure 3. The presented forecasts are generated using the first training set, which is common to both the rolling and recursive forecasting schemes. The figure shows the last 50 observations in the training set, followed by the out-of-sample observations and forecasts up to the longest horizon considered, 50, for the CP, LP and LD series in separate panels. All series and forecasts in the figure are transformed back to percentages to provide an illustration of how all models forecast political support as measured by daily opinion polls.

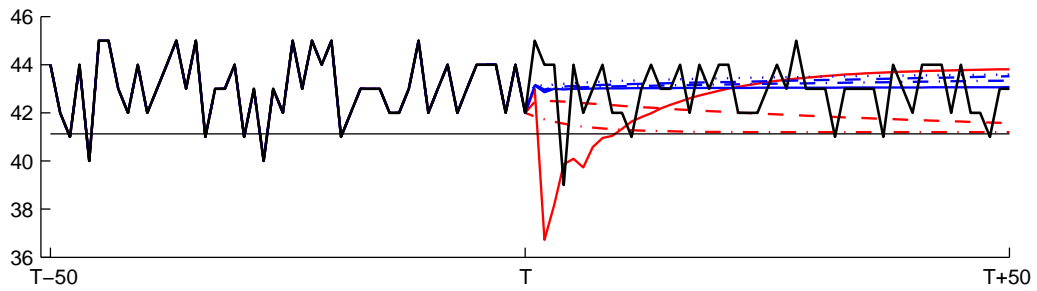
The forecasts for this training set are in agreement with the evidence presented in the model ranking exercise. An interesting observation, which is present in results from other training sets as well, is that even though the CVAR forecasts exhibit more dynamics in the short run (which is a result of a higher lag order selected compared to the FCVAR), this does not translate into more accurate predictions for short horizons. Furthermore, as the short run dynamics die out for both multivariate models, it is evident that fractional cointegration generates more precise predictions. All variants of the FCVAR produce similar predictions, and these are reasonably close to the realized data series in all three panels, while the other models in the portfolio only perform well in some cases.

Figure 3: Forecasts for first training set

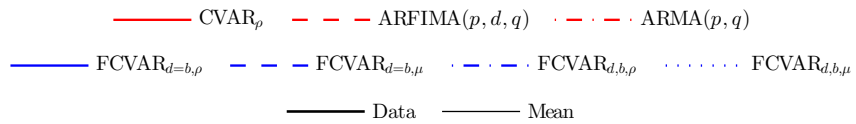
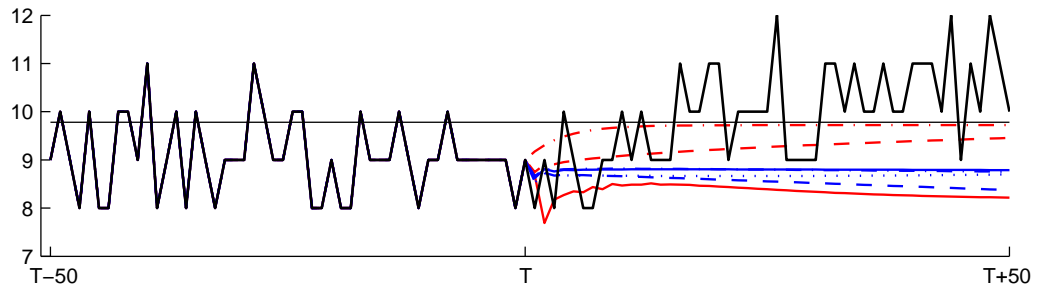
(a) Conservative (CP)



(b) Labour (LP)



(c) Liberal Democrats (LD)



Notes: Training set corresponds to the first window that is shared by the rolling window and recursive window forecasting schemes. The ARFIMA(0, d, 0) model is not included because the ARFIMA(p, d, q) model specifies both lag orders to zero.

5.3 Predicting vote shares of the 2015 UK general election

In this section we present an application to the prediction of the 2015 UK general election, which was deemed the most unpredictable election in decades in the media. Opinion poll agencies predicted a hung government, but ended up significantly underestimating the Conservative Party vote share; the party that won the election with a majority representation in Parliament. Note, however, that vote shares cannot be mapped into election outcomes in the context of the UK election process. This can be seen from the fact that the realized vote share for the Conservative Party was 36.9%, but the party won 331 out of 650 constituencies in the country; an outcome that the predicted vote share of the previous-day YouGov opinion poll, 34%, does not exclude.

We also note that the three political parties represented in our data set, are the three major political parties in the UK that have historically had the most representation in government by a strong margin over other parties running in the election. However, as the 2015 general election has shown, major losses incurred by the three major parties can be due to constituencies lost to other political parties not in the top three. In this election the UK Independence Party (UKIP), Scottish National Party and Green Party have received 12.6%, 4.7% and 3.8% vote shares, respectively, and these three parties are not accounted for in the data for the entire duration of the survey.

A key strength of our data set for the purpose of vote share prediction is that the observed time series are contained within one political cycle. This should allow application of a relatively simple statistical model, and the previous analysis has shown strong support for the FCVAR model for this task. We next present predictions of all models in the portfolio leading into the 56th UK general election held on May 7th, 2015. Specifically, Table 11 shows the opinion poll and the election vote share predictions of each model one month, one week and one day preceding the election day, as well as the final election vote share outcome. The predicted values represent forecasts of the poll standings which can be viewed as predicted vote shares for each political party. Relative to the poll predictions on May 6th, 2015, which underestimated the vote share of the Conservative Party and overestimated the vote shares for Labour and the Liberal Democrats, all models in the portfolio predicted a fall in the vote share for the Liberal Democrats, in agreement with the realized vote share. The forecasts generated on the day before the election show that, for both the Conservative Party and the Labour Party, the FCVAR model variants, CVAR model and ARFIMA(0, d , 0) model predicted vote shares very similar to the poll predictions, while the ARFIMA(p , d , q) and ARMA(p , q) models predicted lower vote shares for both parties.

The advantage of using the FCVAR model for predicting vote shares becomes evident when examining the evolution of the models predictions as the observed data approaches the election day on May 7th, 2015. Figure 4(a) shows how the analysis presented in this paper complements the industry standard of using the latest opinion poll as an indicator of future political support, by allowing the analyst to monitor the dynamics in the opinion poll and the model predictions. The evolution of the predictions for all FCVAR model variants show a clear upward trend in the support for the Conservative Party leading into the election, and a downward trend for the Labour Party for one week leading into the election. These two trends project the correct direction for the realized vote shares in the election. For the Liberal Democrats, the strong upward trend in the predicted vote share is suspect to exhibit

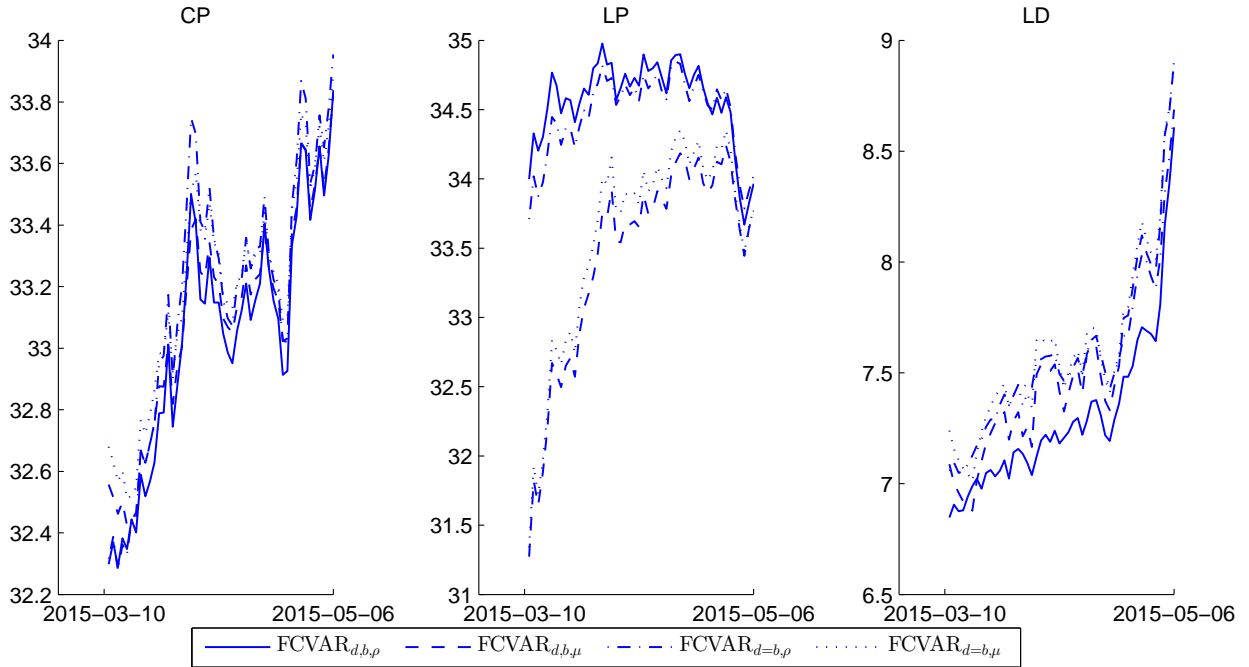
Table 11: Vote share prediction and results for the 2015 UK general election

		Conservative	Labour	Liberal Democrat
Panel A: 1 month before election				
Poll		33.00	35.00	8.00
Forecasts	FCVAR _{d,b,ρ}	33.41	34.59	7.56
	FCVAR _{d,b,μ}	33.38	33.76	7.66
	FCVAR _{d=b,ρ}	33.15	34.65	7.19
	FCVAR _{d=b,μ}	33.22	33.67	7.51
	CVAR _ρ	32.68	32.91	7.23
	ARFIMA(0, d, 0)	34.15	35.07	8.13
	ARFIMA(p, d, q)	34.01	35.18	8.33
	ARMA(p, q)	34.54	39.30	9.26
Panel B: 1 week before election				
Poll		34.00	35.00	8.00
Forecasts	FCVAR _{d,b,ρ}	33.81	34.65	8.00
	FCVAR _{d,b,μ}	33.72	34.35	8.10
	FCVAR _{d=b,ρ}	33.64	34.58	7.69
	FCVAR _{d=b,μ}	33.64	34.24	8.04
	CVAR _ρ	33.57	33.79	7.70
	ARFIMA(0, d, 0)	33.96	34.91	8.27
	ARFIMA(p, d, q)	33.96	34.91	8.32
	ARMA(p, q)	34.32	37.26	9.05
Panel C: 1 day before election				
Poll		34.00	34.00	10.00
Forecasts	FCVAR _{d,b,ρ}	33.95	34.02	8.68
	FCVAR _{d,b,μ}	33.89	33.77	8.90
	FCVAR _{d=b,ρ}	33.98	34.05	8.53
	FCVAR _{d=b,μ}	33.83	33.77	8.90
	CVAR _ρ	33.81	33.66	8.65
	ARFIMA(0, d, 0)	33.89	34.11	9.23
	ARFIMA(p, d, q)	31.77	33.05	7.18
	ARMA(p, q)	33.37	33.10	7.01
2015 election result		36.90	30.40	7.90

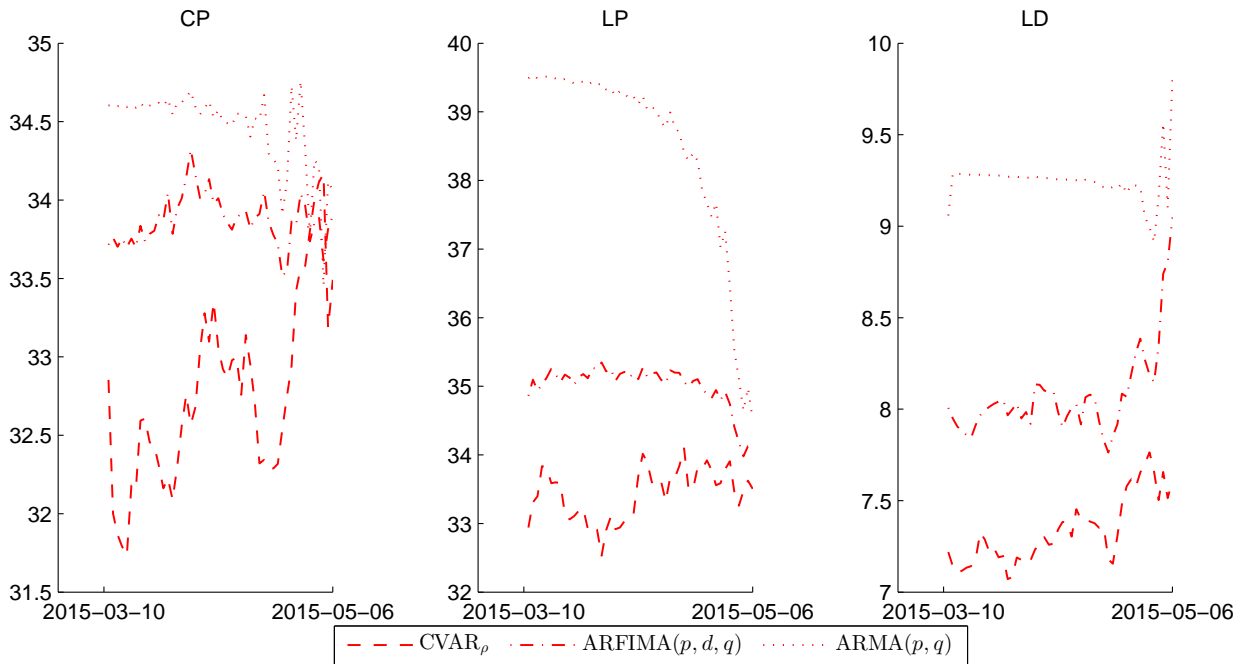
Notes: The table shows the opinion poll and the election vote share predictions of each model one month (Panel A), one week (Panel B), and one day (Panel C) preceding the election day. The last row shows the election vote share outcomes.

Figure 4: Vote share prediction over 50 polling days leading into the election

(a) FCVAR model



(b) Competing models



Notes: Each subfigure shows the evolution of election vote share predictions by all FCVAR model variants for the May 7th, 2015 general election. The predictions are formed starting with the data available 50 polling days prior to the election and continue to May 6th, 2015, the day before the election.

the same tendency as political polls, in that they tend to over-represent support for smaller political parties when compared to the election vote share outcomes and the representation in government. Overall, these dynamics leading into the election provide a more informative assessment of the current state of public opinion on electoral support than that suggested by the hung government prediction of the opinion poll.

The evolution of the vote share predictions generated by the competing models, see Figure 4(b), do not exhibit the same trends as the FCVAR models predictions, which were indicative of the realized vote shares in the election for the Conservative and Labour parties. The CVAR model captures an increase in support for the Conservative Party, but the increase is not as strong as for the FCVAR models. The CVAR model also incorrectly predicts a fall in the Conservative Party vote share for the last week leading up to the election day. The CVAR models predictions for the Labour Party do not suggest any upward or downward momentum in the vote share, and while the ARMA(p, q) model correctly finds a downward trend in the vote share, it substantially overestimates support for the Labour party throughout.

In general, the results from this empirical application show how modeling time series of political opinion polls using the FCVAR model, which is strongly favored by the model forecast comparisons, can complement the industry standard of basing predictions solely on the most recent opinion poll (e.g., the poll standings on the day preceding the election day) and provide a more informative assessment of the current state of public opinion.

6 Concluding remarks

This paper has examined the forecasting ability of the fractionally cointegrated vector autoregressive (FCVAR) model of Johansen (2008) and Johansen and Nielsen (2012) relative to a portfolio of competing models at several forecast horizons. The model was applied in the context of predicting political support in the form of opinion polls; a very relevant topic in the context of forecasting. The analysis used a novel data set of daily polling of political support in the United Kingdom over the period 2010–2015. The analysis has shown how observed time series of public opinion polls can improve forecasting public opinion beyond the most recent poll date. This complements the industry standard for measuring the current state of political support through opinion polling, and contributes to decision making processes that rely on poll evidence as inputs.

Specifically, the forecasting analysis has shown that the FCVAR model delivers valuable gains in predicting political support. The accuracy of both short, medium, and long term forecasts generated by the FCVAR model is better than all multivariate and univariate models in the portfolio. Indeed, overall, the four variants of the FCVAR model are the top performing models. Not only do they perform better relative to the other models, but the forecasting performance of all FCVAR model variants are within close range of each other. When compared to the main benchmark CVAR model, the FCVAR model significantly outperforms the benchmark at all forecast horizons and the gains are more pronounced at the longer forecast horizons, where the root mean squared forecast error is up to 20% lower than the benchmark. Overall the fractional models outperform the non-fractional models, and the evidence provides strong support for the application of the FCVAR model. Fractional cointegration substantially improves forecast accuracy, and the gains become more pronounced at longer forecast horizons.

In an empirical application to the prediction of vote shares for the 2015 UK general election, the results show how modeling time series of political opinion polls using the FCVAR model, which is strongly favored by the model forecast comparison analysis, can complement the industry standard of basing predictions solely on the most recent opinion poll. The forecasts generated by the FCVAR model leading into the election appear to provide a more informative assessment of the current state of public opinion on electoral support than the hung government prediction of the opinion polls. Specifically, the FCVAR model projects the correct direction for the realized vote shares in the election for the Conservative and Labour parties.

References

- Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics* 73, 5–59.
- Box, G. E. P. and G. M. Jenkins (1970). *Time Series Analysis, Forecasting, and Control*. San Francisco: Holden-Day.
- Box-Steffensmeier, J. M. and R. M. Smith (1996). The dynamics of aggregate partisanship. *American Political Science Review* 90, 567–580.
- Box-Steffensmeier, J. M. and A. R. Tomlinson (2000). Fractional integration methods in political science. *Electoral Studies* 19, 63–76.
- Byers, D., J. E. H. Davidson, and D. Peel (1997). Modelling political popularity: an analysis of long-range dependence in opinion poll series. *Journal of the Royal Statistical Society Series A* 160, 471–490.
- Byers, D., J. E. H. Davidson, and D. Peel (2007). The long memory model of political support: some further results. *Applied Economics* 39, 2547–2552.
- Dolado, J. J., J. Gonzalo, and L. Mayoral (2002). A fractional Dickey-Fuller test for unit roots. *Econometrica* 70, 1963–2006.
- Dolado, J. J., J. Gonzalo, and L. Mayoral (2003). Long-range dependence in Spanish political opinion poll series. *Journal of Applied Econometrics* 18, 137–155.
- Granger, C. W. (1980). Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics* 14, 227–238.
- Hualde, J. and P. M. Robinson (2011). Gaussian pseudo-maximum likelihood estimation of fractional time series models. *Annals of Statistics* 39, 3152–3181.
- Jensen, A. N. and M. Ø. Nielsen (2014). A fast fractional difference algorithm. *Journal of Time Series Analysis* 35, 428–436.
- Johansen, S. (1995). *Likelihood-based inference in cointegrated vector autoregressive models*. New York: Oxford University Press.

- Johansen, S. (2008). A representation theory for a class of vector autoregressive models for fractional processes. *Econometric Theory* 24, 651–676.
- Johansen, S. and M. Ø. Nielsen (2012). Likelihood inference for a fractionally cointegrated vector autoregressive model. *Econometrica* 80, 2667–2732.
- Johansen, S. and M. Ø. Nielsen (2015). The role of initial values in conditional sum-of-squares estimation of nonstationary fractional time series models. *Econometric Theory*, forthcoming.
- Jones, M. E., M. Ø. Nielsen, and M. K. Popiel (2014). A fractionally cointegrated VAR analysis of economic voting and political support. *Canadian Journal of Economics* 47.
- MacKinnon, J. G. and M. Ø. Nielsen (2014). Numerical distribution functions of fractional unit root and cointegration tests. *Journal of Applied Econometrics* 29, 161–171.
- Nielsen, M. Ø. (2015). Asymptotics for the conditional-sum-of-squares estimator in multivariate fractional time series models. *Journal of Time Series Analysis* 36, 154–188.
- Nielsen, M. Ø. and M. K. Popiel (2014). A Matlab program and user’s guide for the fractionally cointegrated VAR model. QED working paper 1330, Queen’s University.
- Robinson, P. M. (1978). Statistical inference for a random coefficient autoregressive model. *Scandinavian Journal of Statistics* 5, 163–168.
- Shimotsu, K. and P. C. B. Phillips (2005). Exact local Whittle estimation of fractional integration. *Annals of Statistics* 33, 1890–1933.