Picking the winner(s): Forecasting elections in multiparty systems

CrossMark

Daniel Walther

University of Umea, Department of Political Science, 901 87, Umea, Sweden

ARTICLE INFO

ABSTRACT

From the 1970s onwards, a wide range of forecasting techniques have been developed in the literature on electoral forecasting. However, these models have primarily been applied in two-party, presidential democracies, with the US being by far the most popular country to investigate. The question thus arises whether the same techniques that have proved successful in this context can also be applied to the more complex, multiparty democracies in northern Europe. This paper seeks to answer this question and in the process makes two main contributions. Firstly, the popular dynamic linear model (Jackman, 2005) is tried and tested in Germany and Sweden where it is shown that reasonable forecasts can be made despite the complexity of the systems and the emergence of new parties. A novelty is then introduced when cyclical changes in party support are modelled through a *seasonal component*. This extension of the dynamic linear model helps to significantly lower the error in early forecasts and is thus something that could be useful in future applications of the model.

1. Introduction

Predicting election results is a relatively recent and increasingly popular part of political science research. Competitive elections are the hallmark of modern democracy and being able to foreshadow who wins them is a tantalizing skill that has garnered significant scientific attention (Fisher et al., 2011; Lewis-Beck and Bellucci, 1982; Lock and Gelman, 2010; Gibson and Lewis-Beck, 2011; Jackman, 2005). Election forecasting stands out from many other types of political science research in a number of ways. It is highly data-driven, focused on a very concrete and delimited task, and in most studies the goal is not to explain election outcomes but to describe and predict them. In that sense, the question 'how' rather than the standard scientific question 'why' is in focus.

The question 'how' is still highly relevant from a scientific perspective. In order to answer it with reasonable accuracy you need to make the most of limited and flawed polling data, while controlling for seasonal fluctuations in public opinion, variability in measurements and bias associated with particular polling houses. In overcoming these problems we can shed more and better light on public opinion by overcoming the flaws inherent in individual polls. In addition, the circumstances facing election polling are also

faced by many scientists working with other types of data survey data. The main difference with political polling is that here we get a perfectly unbiased measurement, the national election, which makes it possible to test our estimations. This allows us to progressively develop techniques that can be used also in many other disciplines as well.

The goal of the present paper is to test whether it is possible to predict elections also in difficult parliamentary systems where a wide range of parties are competing for power, and if this can be done with reasonable lead time. For this purpose, two countries with a long tradition of multiparty competition, namely Germany and Sweden, have been selected. These cases provide a compelling tests since most of the previous studies have focused on more stable two-party systems and the methods have also been developed to fit such political environments.

Sweden, in contrast, has 8 parties represented in parliament and has seen large shifts in the electoral fortunes of the parties during the first decade of the 21st century. In the latest German elections in 2013, 6 parties received more than 4% of the vote. Such a set-up requires our models to take many more active players into consideration.

Previous studies have focused on a wide range of countries, but the U.S. has received the lion's share of scholarly attention. Other countries, though, such as the U.K.(Fisher et al., 2011), France (Foucault and Nadeau, 2012), Australia (Carlsen, 2000), and Italy (Lewis-Beck and Bellucci, 1982), have also been studied. The most popular technique used in these studies (see Bartels

E-mail address: daniel.walther@umu.se.

and Zaller (2001); Hibbs Jr (2000) for early overviews) is some type of structural model. This approach treats incumbent vote share as the dependent variable and a range of economic and political measures are used as explanatory variables that are believed to have a bearing on the outcome (e.g. Lewis-Beck (2005).

In this paper, in contrast, it is argued that pure structural models are difficult to apply in the multiparty context. To accurately forecast the results of all parliamentary parties it appears necessary to include some kind of polling data[1] in the model. An increasingly popular way of doing this is through a dynamic linear model that belongs to the general family of space state models (Jackman, 2005; Harrison and West, 1997; Pickup and Johnston, 2007). A latent trend of popular support is estimated by aggregating polling results into a time series through Kalman filtering. The election forecast is then made by extrapolating this trend into the future.

The *main contributions* of this paper are firstly the demonstration that reasonable election forecasts in multiparty systems can be made through a dynamic linear model. Structural models appear to be of limited utility though, at least if not complemented by polling data. The average error of the DLM predictions is found be around 0.69 percentage points per party in Germany and 0.78 i Sweden for the last three elections. These results are on par, or better, than what many of the previous studies have achieved in the more stable two party systems.

A second contribution is the introduction of what has been termed a *seasonal component* (Kitagawa and Gersch, 1984; Bell and Hillmer, 1984). In the economics literature a seasonal component is used to capture predictable seasonal trends (such as a boom in sales just before Christmas), and it has been argued that the development in support of political parties shows similar predictable patterns (Sanders, 1991). The stronger these reoccurring trends are, the more the standard DLM benefits from the inclusion of the seasonal component. The seasonal component is empirically tested in both Sweden and Germany. It is found to help us make reasonable forecasts at an earlier stage and it reduces the error compared to the standard DLM one month before the election with around 17% in the Swedish case and 8% in Germany. Adding the seasonal component makes a model that is already good at *nowcasting* better at the more difficult art of *forecasting*.

This paper starts with an overview of the main forecasting techniques that have been used in the past. I then discuss applications of structural models and find that we have theoretical reasons and empirical findings from the field of economic voting that suggest that such techniques hold little utility in the multiparty case. The paper then moves on to the dynamic linear model and applies it to both the German and Swedish cases. In the final section I apply the model with the addition of a seasonal component. The conclusion is that the dynamic linear works well but that more work is needed to incorporate explanatory factors into the model to further improve early forecasting.

2. The main approaches to election forecasting

Historically speaking, election forecasts are a relatively recent phenomenon. The polling company Gallup attempted forecasts of the US presidential elections in the 30s and 40s (with very modest success) and by the 50s more pollsters, not just in the US, took a stab at predicting the election (see Lewis-Beck (2005) for an overview).

In the 70s and 80s, economists and political scientists also started taking an interest in forecasting and a number of competing models emerged. The polling companies had, perhaps unsurprisingly, relied mainly on their own estimations of vote intentions that they got from their pre-election polls. The scientists, in contrast, introduced regression-based approaches. These relied on economic and political variables assumed to influence government popularity and inserted them into structural models that were used to forecast election outcomes. OLS regression was the dominant approach.

2.1. Predicting elections through structural models

Most applications of structural models in electoral forecasting have been of the type:

$$Government\ election\ result = political\ measures \\ + economic\ performance + error$$

The models have mainly differed in how they have operationalized political and economic performance. The political situation has been captured for example through measures of popularity of the president/PM, general left- or right-wing sentiment among the populace as well as current length of stay in office (Bartels and Zaller, 2001; Foucault and Nadeau, 2012; Abramowitz, 2008). The belief is that such variables can capture the general political mood in the country rather than just the current support for a specific political party. By tapping into the general mood we can learn what the political backdrop to the election will look like and this situation is then modified (exacerbated or ameliorated) through the government's economic performance.

Measures of the government's *economic track record* have also been plentiful. The most popular ones include changes in GDP per capita, inflation and unemployment, but growth in real income and the subjective beliefs among voters about the government's economic performance have also often been used (Anderson, 2000; Bartels and Zaller, 2001). The causal theory here is clear: a government that has handled state finances successfully and improved the prosperity of its citizens will be rewarded come Election Day, and a government that has failed on these measures will be punished (Healy and Lenz, 2014; Saalfeld, 2008).

Estimates for the coefficients used in the structural models are generated by applying the model to as many past elections as are available to learn what effect the variables have had in the past. Predicting the next election therefore becomes a simple matter of inserting the relevant values for the political and economic indicators as they stand in the election year and then multiplying them with the coefficients that previous elections tell us provide the best fit.

One illustrative (and very Spartan) example of such an approach is the famous bread and peace model developed by Douglas Hibbs (see Hibbs Jr (2000) for an early overview). Hibbs argues that only two measures are needed to predict the outcome of the US presidential election: namely the weighted cumulative income growth during the full term of the government and the number of US soldiers killed in foreign wars (particularly in Korea and Vietnam). Using only those two indicators he manages to account for 90% of the variation in election results. Controlling for other variables that have been suggested in the literature does nothing to improve predictive capacity, Hibbs argues (Hibbs Jr, 2000).

Hibbs' diminutive model nicely illustrates the two compelling advantages structural models have over polls. Firstly, the

---

[1] The word *data* is throughout this paper treated as a singular noun.

information you need to make your prediction is usually readily available well before the next election. Once the model has been applied to previous elections and the coefficients are at hand, making the prediction becomes a simple matter of inserting the latest data. In this view, what political parties do in the run up to the election is primarily to inform and convince people of the underlying economic and political reality. But political scientists who know where to look can find the relevant data well before the parties go into campaign mode and consequently know how the campaigns will eventually come to change public opinion.

A second advantage of structural models is that they deal with actual *causes* of the outcome of the election. It makes intuitive sense that an unpopular president from a party that has been in office for a long time and has a poor economic track record fares worse at the polls. Structural models can test exactly how robustly such factors are associated with the election outcomes and can thus test what matters the most to voters.

2.2. Election forecasts through polling-based methods

The second popular, and increasingly dominant, method for predicting elections generally disregards objective economic and political indicators and instead looks at people's subjectively stated voting intentions. Opinion polls are becoming ever more plentiful in most Western countries and now provide frequent insights into public opinion. Polls are now common not just in the run-up to an election but throughout the electoral cycle and, thanks to technological advances, now also have more respondents.

In Sweden, for example, the average number of polls was fairly stable between the 1970s and the late 90s, but then increased manifold during the first decade of the 21st century. An overview of this trend is available in Fig. 1. Apart from the notable spikes in election years, we can see a clear gradual progression in the past decade towards the current figure of close to 8 polls per month (which rose to over 10 in 2014). This gives us a lot more data to work with than has been available in the past and makes polling data more feasible as the basis for forecasting.

In Germany the number of national polls was already quite high in the first years of the 21st century. Still, since then the number of polls has continued to rise from an average of 16 per month in 2002 to just over 21 in the 2013.

Predictions using polls range from the more simplistic weighted poll averages (aggregating the polls while accounting for poll size) to some highly complex dynamic linear models (Jackman (2005) provides a good introduction). The dynamic linear model (DLM) has become the staple horse technique for polling based forecasts (Pickup and Johnston, 2007; Fisher et al., 2011; Linzer, 2013) and it is therefore worth going over the technique in some detail.

2.3. The dynamic linear model

DLMs rely on Markov chains with random walk and a Kalman filter (Harrison and West, 1997) to estimate the underlying public support for each party. Each poll that comes in is taken to be a slightly flawed measure of the real support for the party at time *t*, and the polls are then pooled into a time series that tracks party support over time. By doing so the hope is to overcome the limitations and biases of a single poll to establish a more realistic aggregate measure. So two of the main advantages of applying the DLM are firstly that the model constitutes a highly effective way to combine many polls over time into an estimate that conforms with the laws of probability. Secondly, the model allows us to perform a real-time tracking of party support, which gives us more continuous information than the one off estimate provided by structural models.

The core of a DLM is defined by the following set of equations (Petris et al., 2009):

$$P_i = \mu_i + \sigma_i^2, \sigma^2 \sim N\left(0, \sigma_i^2\right) \tag{1}$$

$$\mu_i = \mu_{i-1} + \delta_i, \delta \sim N\left(0, \delta_i^2\right) \tag{2}$$

$$\mu_0 \sim N(m_0, C_0) \tag{3}$$

Where equation (1) defines the observed time trend (i.e. the actual polling data) and (2) the assumed 'real' underlying trend. (3) is the starting value that sets the Markov chains in motion. We can see that the mean of the polling data (u_i) is the estimation from the underlying trend, which in turn is a result of the previous estimations plus a variance term (δ_i). The actual polls (P_i) are then simply the estimated latent trend with the addition of some
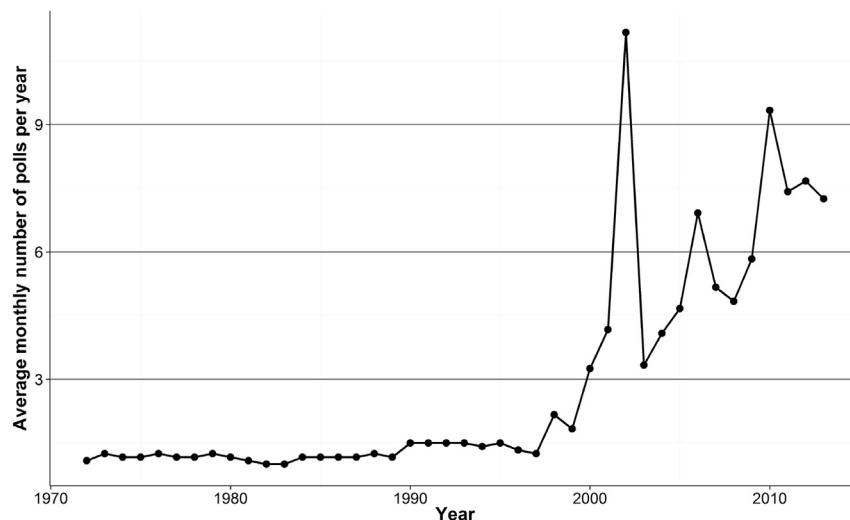

Fig.. 1 Developments in the popularity of polling in Sweden.

variance. Everything is assumed to follow a Gaussian (normal) distribution.

The process works by continuously trying to predict the next point in the time series. So a prediction is made, and once the next result is actually measured (i.e. new polling data comes in) we calculate how far off we were, update our estimation of the true latent state, and then make a new prediction for the next observation. The actual election is then simply another upcoming observation we are trying to predict.

The scholars that have applied DLMs have adapted the standard model outlined above in ways particularly suited to election forecasts. There are a couple of things one might wish to control for, such as the size of the poll, bias associated with a particular houses (Pickup and Johnston, 2007; Fisher et al., 2011). These factors modify how much confidence we should have in a new poll that comes our way (Silver, 2012).

Various techniques to achieve this have been suggested. We can control for the size of the poll by adding a prior to the variance of the sigma term in equation (1):

$$\sigma^2 = \frac{p(1-p)}{N} \tag{4}$$

Pickup and Johnston (2007) suggested a way to control for bias among polling houses by using the median polling house as an anchor (reference category) and then calculating whether certain houses differ significantly from this. You can run a series of regressions with dummy variables for the polling houses, according to a variation of equation (1):

$$P_i = \mu_i + \beta_i X_i \tag{5}$$

Where the X's are dummy variables representing each polling house. The question in the regression thus becomes: which polls differ significantly from the median estimate once we have controlled for the DLM estimate? If a particular house has a coefficient that shows that it consistently differs from the median estimate, its impact on our forecast can be weighted down in proportion to how far off it is on average. The problem is that we need to use one house as the reference point (the median in Pickup and Johnston's model) and there are clearly no guarantees that the polling industry as a whole gets it right on average. It could very well be that the extreme outlier is in fact the best poll out there.

Still, if the polling industry as a whole gets it significantly wrong on average our poll based model is unlikely to be of much use anyway. We are already assuming that aggregating the polls gives us a better estimate and that we improve our estimate by treating the polls as a continuous time-series. If those assumptions hold, this technique should be a useful way to measure the running bias associated with particular houses and will thus be applied here.

2.4. Adding a seasonal component

One problem with the DLM, even with the modifications of the basic model listed above, is that it is built exclusively on the information that has been present in the polls up until that point in time. This means that even if we know that certain changes in party support are likely to happen in the coming months, the basic model cannot take this into account.

For example, it is well known that the support of government parties tends to drop in the middle of its term in office only to be partly or completely regained when the next election is approaching (Sanders, 1991). This is the so-called political business cycle. Similarly, at least in some systems, smaller parties that are

Table 1
Comparison of DLM models with and without controls for seasonality.

Party	R^2 DLM with seasonal component	R^2 standard DLM	Country
CDU/CSU	0.872	0.872	Germany
SPD	0.440	0.429	Germany
GRUENE	0.823	0.820	Germany
FDP	0.515	0.503	Germany
LINKE	0.325	0.316	Germany
S	0.226	0.217	Sweden
V	0.542	0.540	Sweden
MP	0.223	0.209	Sweden
M	0.783	0.783	Sweden
FP	0.042	0.020	Sweden
C	0.225	0.202	Sweden
KD	0.133	0.110	Sweden

The best fit DLM model and the best fit DLM model with a seasonal component were computed for each party. The time series ran for a total of 24 months.

close to the parliamentary threshold in the polls also tend to slump in the middle of the election period only to regain support in the run-up to the election. This is because voters hesitate to let them disappear completely, especially if they are part of a larger coalition (Fredén, 2014). Also, following a scandal or another unpopular event, parties can temporarily lose support. If we were to base our forecast on the party's standing just after the event we likely underestimate their eventual result. In all of these cases, the development of party support can be seen as seasonal or cyclical since it tends to revert back from temporary outliers.

So if we do believe that there are seasonal trends in how party support develops, how can the dynamic linear model be extended to incorporate this information before we see it in the polls? One possible solution is to extend the time series of polling data to incorporate a greater number of polls and then model the reoccurring patterns through what in the economics literature is known as a 'seasonal component'.[2] This is often used to capture predictable economic events such as a slump during summer or a sales boom in the days leading up to Christmas. Mathematically, this can be modelled through (Kitagawa and Gersch, 1984; Scott, 2014):

$$S_t = -\sum_{i=1}^{L-1} S_{t-1} + \varepsilon_t \sim N(0, \sigma) \tag{6}$$

Where S is the season and L the number of periods in each season.[3]

Seasonal adjustment of time series dates back to the 1920s and a wide range of different techniques have been suggested (see Bell and Hillmer (1984) for an early overview). The techniques, naturally, have different strengths and weaknesses, and the method developed by Kitigawa and Gersch was selected for two main reasons. First, it is highly flexible and can be applied to long time series of different durations. This can be contrasted with the popular X12-Arima model (used e.g. by the US census bureau) which works best for monthly or quarterly data (Hyndman and Athanasopoulos, 2014; Jain, 2001). Second, this way to model seasonality can easily be added as a new layer in the DLM model and can thus be calculated separately alongside the other components. Unlike some other techniques, this makes it possible to set priors for some of the terms and to calculate their individual variance (Jain, 2001). This ensures that the seasonal component

---

[2] Technically, a seasonal component is of fixed duration (e.g. every six months), whereas a cyclical component has a more flexible life cycle. Here the terms are used more or less interchangeably.

[3] See the appendix for a discussion of how this is applied here.

can be successfully integrated into the overarching Bayesian set-up.

But even though we have theoretical reasons to expect seasonal swings, how can we be sure that adjusting for it will improve the model? Testing for the extent to which seasonality is present in the party time series here is less straightforward than in normal economic time series, since the trends are not necessarily as predictable as those that occur on a daily or monthly basis. Still, one possible solution, proposed by Hyndman and Athanasopoulos (2014), is to fit two models: one that controls for seasonality and one that doesn't and then compare their one-step-ahead prediction errors. Since the DLM assumes linearity, the total ability of the models to pick up the variance of the time series can be measured through R^2. The result of this exercise can be seen in Table 1.

The results indicate that the predictive accuracy improves, slightly but reliably, when controlling for seasonality. In percentages the improvement was slightly larger in the Swedish case. Moreover, the test employed here (accuracy in predicting the next poll in the time series) should be the standard DLM's strongest suits since the polls are temporally close. It is likely that controlling for seasonality will be even more useful when predicting elections that are further away and this expectation will be put to the test in Section 6.

2.5. The reasons for adopting a Bayesian approach

The main reason for adopting a model based on Kalman filtering and a general Bayesian recursive model is accuracy. The Kalman filter can be proved to be optimal when trying to model a linear time series subject to Gaussian noise, and it can easily be extended within a Bayesian framework to make it more adapted to electoral forecasting (Harrison and West, 1997).

The main benefit in our case is that our confidence in the polling estimates can be quantified. We can control for both the size of polls and the reliability of the polling houses through priors to determine exactly how much influence each poll should have on our estimations. This way to explicitly model confidence is difficult to emulate in classical frequentist time series models and it offers a highly flexible way of including other information in our model not directly available in the polls.

That being said, Bayesian statistics will be utilized in a fairly pragmatic fashion here. For example, the estimates of the variance that we get from the DLM will be used to calculate confidence intervals even though this practice is sometimes shunned in Bayesian circles (but is done e.g. by Jackman (2005)).

3. Why structural models are problematic in the multiparty case

Both structural models and DLM models been tried and tested in a wide range of studies and we know that both of them produce reasonable results. At least in stable settings with few competing parties. When trying to translate the techniques to a multiparty framework a number of interesting challenges emerge for the structural models. There are three main reasons for this, namely:

1. Lack of a clear dependent variable
2. Difficulty in assigning economic and political responsibility to individual parties
3. Difficulty in dealing with new parties

The common theme in these three factors is that multiparty systems display greater flux, more frequent emergence of new actors, and, given the greater number of active players here, problems with translating particular economic and political results to precise vote shares for all the different parties. These problems seem inherent in the set-up of multiparty systems and each will be discussed briefly in turn.

3.1. Lack of a clear dependent variable

As shown above, almost all applications of structural models in election forecasting use incumbent vote share as the dependent variable and then use various political and economic indicators as predictors. This works well when you have two parties, since if party A is in government and you can forecast its result, party B gets 100 minus the score of party A. This is useful, because it means that you only have to run a single regression on your sample since this allows you to estimate the result of both main parties.

Even in other cases, where there are two main parties but also other minor parties in the competition (e.g. the UK, France), the regression estimates can still answer the question of how the competition for the reins of governments will end. For example, in Foucault and Nadeau's attempt at election forecasting in France (Foucault and Nadeau, 2012) they used the results of the conservative party in the 2nd round as the dependent variable.

In systems with proportional representation though, where parties usually have to enter into a coalition after the election to form a government (Mitchell and Nyblade, 2008), the precise results of all the parties matter. For example, even if you could accurately forecast how the German conservatives and social democrats will do in the election, you still would not be able to predict who would eventually be in power. That depends too much on how the liberals, greens, left party and others do.

This means that we cannot simply use incumbent party vote share as our dependent variable since there is no way of knowing how the share of votes will be distributed among the various opposition parties. Instead we need to have separate models for each party and thus generate unique coefficients for each player in the party system. But adopting this approach immediately leads to new, even more damaging problems.

3.2. Difficulty assigning responsibility to individual parties

Using incumbent vote share as the dependent variable is advantageous not only for practical reasons, as discussed above, but also because of our understanding of how the causal processes work here. Incumbent parties should be affected by the economic and political situation of the country, since these parties are *responsible for it*.

However, using every party in the country as the dependent variable in a series of independent regressions, as we need to do if we want to estimate the support of many players in a multiparty system, neglects this logical link between results and responsibility. If e.g. the Social Democrats are in power and the unemployment rate increases by 1.5%, how does this impact the Christian Democrats, the Greens and the Left parties in opposition? Will all opposition parties be affected in the same way? Even if we assume that the voters want to punish the incumbent social democrats, there is little reason to expect that all of the opposition parties will be affected in a proportional and predictable manner as would be the case in a two party system.

Moreover, if there is a coalition government in power, are all parties held equally responsible by voters? This seems unlikely given that the coalition members have had different responsibilities in the governance of the country and generally are associated with different overarching policy areas in the eyes of voters.

The argument that this causal link matters receives both theoretical and empirical support from the field of economic voting (e.g, Paldam (1991); Lewis-Beck and Paldam (2000); Anderson (2000); Bengtsson (2004); Nadeau et al. (2013); Goodhart (2009); Narud and Valen (2008)). Anderson summarized the main findings of the field in 2000 by stating that three key features of the political system mediate the effect of the economy on the government's election result, namely:

- Institutional clarity
- Governing party size
- Availability of alternatives

These three factors have one feature in common that can be termed 'clarity of responsibility'. If the institutional set-up makes it clear who is in charge, if there is one dominant governing party and it is apparent what the alternative is to the incumbent ruler, then responsibility for policy outcomes is clear (Goodhart, 2009). When there is clarity of responsibility the voters know who to blame or credit for the economic conditions and then the government's track record plays an important role in predicting the election result.

This shows why economic performance plays an important role in presidential systems where one party government is the norm. In such systems it is abundantly clear who is responsible for national economic management and usually there is one main opposition party that provides a clear alternative to the current government. In multiparty systems where coalition government is the norm and there are numerous opposition parties, the relationship is likely to be far less clear-cut.

3.3. Dealing with new parties

One additional problem that makes structural models problematic in multiparty systems is that the number of parties in such systems tends to change over time. In Sweden, three parties have gained and kept parliamentary representation since the late 80s, and many other countries with proportional systems, especially in Eastern Europe, have seen even swifter changes of the political landscape (Jungerstam-Mulders, 2006). In Germany, one new party has gained a parliamentary foothold and a few others (e.g. the Pirate Party and Alternative für Deutschland) have been close in national elections and have successfully secured representation in regional parliaments.

But when new parties emerge we have no previous elections to estimate model coefficients on and thus no reasonable way to gauge how certain political and economic conditions should influence a given party. Thus, even if we could get around the problems with what to use as a dependent variable and how to model the causal chain of responsibility, the added flux in multiparty systems compared to their majoritarian equivalents means that forecasting techniques relying exclusively on some kind of structural model are bound to run into problems.

3.4. How previous studies have dealt with these issues

Five recent articles have applied some kind of structural model to the German or Swedish cases (Jérôme et al., 2013; Norpoth and Gschwend, 2010; Graefe, 2015; Kayser and Leininger, 2013; Sundell and Lewis-Beck, 2014). The question then inescapably arises how these have dealt with the theoretical problems outlined above. One popular solution, opted for by Kayser and Leininger (2013); Norpoth and Gschwend (2010); Sundell and Lewis-Beck (2014), is to focus only on the subset of parties that make up the current government. The incumbent coalition is treated as one unitary actor which means that you have one agent responsible for economic and political developments.

If the horse race for power is the key question, this could be a suitable strategy. The 'chancellor model' offered by Norpoth and Gschwend, for example, has been remarkably successful since 2002. Still, in many cases we are interested in the electoral fate not only of the government parties, but also of the parties in the opposition. The current Swedish government consists of two parties, with six in opposition. The models employed in these studies, even if successful, would only be able to predict the results of a small subset of the parties. And even then we would only get the aggregate result of the entire government, not the results of the individual coalition members.

Additionally, sometimes the estimates are simply off the mark. The prediction offered by Sundell and Lewis-Beck was for the Swedish right-wing coalition government to get 49.7% of the votes in the 2014 election. This was 10.3 percentage points away from the actual result of 39.4, which can be compared with the DLM prediction from the same time of 37.8.

Another strategy, pursued by Jérôme et al. (2013), is to combine estimates from a structural model with some polling data. Graefe (2015) goes even further and combines a weighted combination of regression estimates, polls, expert judgements and betting markets. The study by Jerome et al. is interesting in its parsimony, because it uses economic and political variables for the two larger parties in the German system, and then combines this with polling data (primarily focussing on voting intentions) for the smaller parties. They produce a prediction that is both reasonable and has good lead time, and, as I will argue below, such an integrated model has the potential to combine the strengths of both polling and structural models.

In general, then, the existing scholarship suggests that the problems associated with structural models in multiparty system are difficult to overcome. You either have to focus on a subset of parties or reinforce your model through polling data. So how much, exactly, can we do with polling data in multiparty systems?

4. The dynamic linear model in the multiparty case

Many of the problems associated with structural models can, at least seemingly, be overcome by utilizing polling data instead. The problems with the lack of a sensible dependent variable, too few data points, and how to handle new parties all disappear when we instead look at polls. Indeed, systems where structural models work well (e.g. the US, the UK), do not necessarily have any advantage when it comes to poll based techniques, other than that there are fewer parties to estimate.

Some of the key challenges of the DLM have been touched upon above and have also already been discussed in more detail elsewhere (e.g. Jackman (2005); Linzer (2013); Fisher et al. (2011)). But before we proceed to applying the DLM, we should first deal with a few challenges to the standard model that arise in particular in *the multiparty case*.

One statistical problem at the outset is that the DLM assumes a normal distribution. But technically speaking, the underlying data generating process in the multiparty case should follow the multinomial distribution, since the starting point of the data is the individual survey respondent that chooses between a discrete number of different parties. As has been noted in the statistical literature (Severini, 2005), though, the multinomial has a normal approximation. Jackman (2005) comes to the same conclusion when he argues that the binomial distribution in the Australian

party system can be approximated through the univariate normal in that case.

Moreover, in a recent attempt to apply the DLM in the multiparty system of Norway (Stoltenberg, 2013), Stoltenberg develops a new DLM assuming a multinomial distribution. He applies both the standard, Gaussian model and his own multinomial model to the Norwegian election of 2013 and gets very similar results. This suggests that we have both theoretical and practical reasons to assume that the DLM can be applied also in multiparty systems.

From a practical standpoint we need to deal with unreliability in the polls stemming from bias associated with particular polling houses according to the regression method outlined in equation (5). This was done in both the Swedish and German cases. In total two polling companies in Sweden (Skop and United Minds) and three in Germany (FGW, Forsa and INSA) were found to deviate consistently from our DLM estimate and were weighted down in accordance with their average deviance (generally between 5 and 10%).

4.1. Applying the DLM to Germany

Let us now turn to an empirical test of whether our theoretical belief in the suitability of the DLM in the multiparty case is also practically justified. We start by using the standard DLM that doesn't control for seasonality. Here the predictions were made on the day before the elections, whereas forecasts with better lead time will be made in the next section.

First, let us turn to the three German elections that took place in 2005, 2009 and 2013. There were five main parties competing in 2005 and 2009 and two more (die Piraten and AfD) in 2013. Fig. 2 gives an overview of how popular support for the five main parties has developed since 2005 according to the DLM estimates.

The developments in party support over time show that there has been fairly extensive fluctuations. CDU/CSU went from a high of close to 50% a few months before the 2005 election to just over 30% in 2007. Some of the smaller parties, especially the Green party, show even larger fluctuations relative to their size. This nicely illustrates the large variations in party support in multiparty systems

and thus why it is advantageous to estimate support as a continuous trend.

With this in mind, let us turn to an application of the DLM to the elections that took place in 2005, 2009 and 2013. In Fig. 3 we can see the point predictions for Election Day in those three elections. These predictions are a snapshot from the time trends in Fig. 2 where have zoomed in on the relevant dates but only used data up until the day before the election. The added confidence intervals show the uncertainty of the estimate.

The overall impression is that the model works well. All results (represented by the grey dots) fit within the 95% confidence intervals (although there are a few close cases, such as FDP in 2005 and CDU/CSU in 2013), and the average error is around 0.69 percentage points. The model seems to work well also for parties that have displayed large variation over time (such as CDU and the Greens) and performs equally well in 2005 and 2013 despite the slightly lower availability of polls in 2005. 2009 was the best election for the model with an average error of only 0.5.

Another promising result is that the two new parties in 2013 did not seem to be more difficult to predict. For AfD there was only polling available from April 2013 onwards, which meant that the model had only around 5 months of solid data to create the estimates. The actual result of AfD still fit comfortably within the 95% confidence interval which suggests that the DLM works reasonably well even when a new party emerges late in the process and data is scarce.

A final thing to note is that the uncertainty in the estimates stems not only from the number of polls available, but also from how much estimates from different polling houses vary and from how much support levels of the party have changed over time. For example, the confidence intervals of the Social Democrats are noticeably wider in 2013 than in 2005 since in 2013 the party was fluctuating both up and down and the polling companies were less in agreement about the direction in which the party was going.

4.2. Applying the DLM to Sweden

With these results at hand, let us now undertake a second empirical test by applying the model to Sweden. The Swedish party
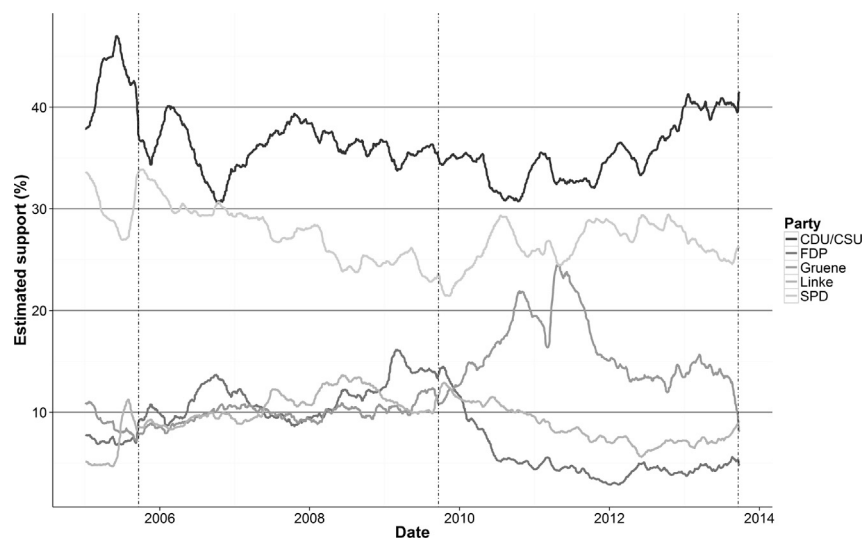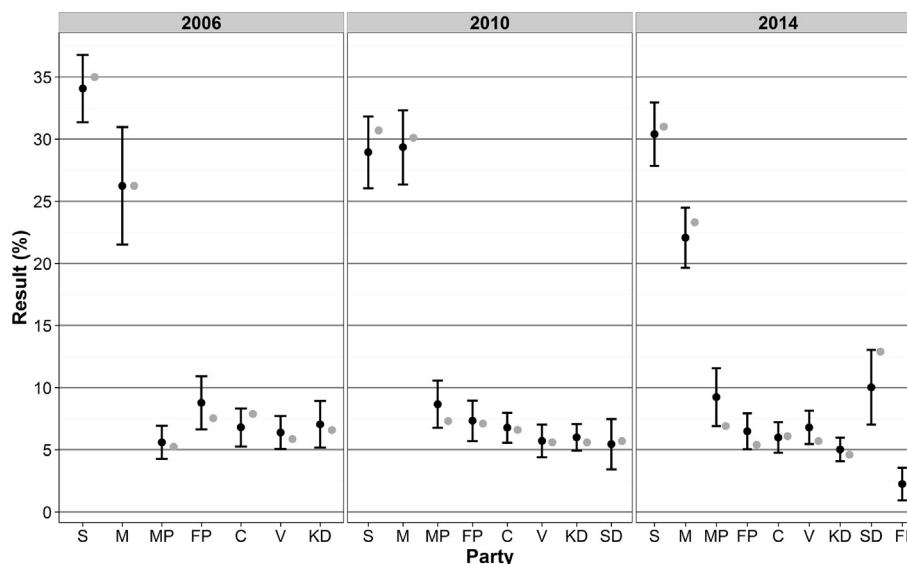

The graph shows the level of support for each party as estimated by the dynamic linear model. The dashed lines indicate the timing of the elections that took place in 2005, 2009 and 2013.

Fig. 2. Development in party support in Germany from 2005 onwards.
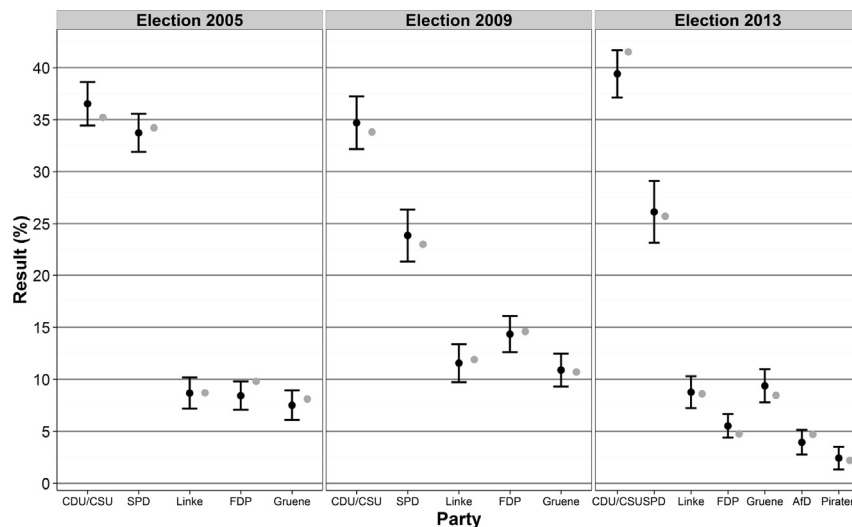
The black dots show the estimated level of support for each party on the day before the respective elections. The surrounding wings are the 95% confidence intervals and the grey dots are the actual election results. An explanation of the party acronyms is available in the appendix.

Fig. 3. Election forecasts in Germany.

The black dots show the estimated level of support for each party on the day before the respective elections. The surrounding wings are the 95% confidence intervals and the grey dots are the actual election results. An explanation of the party acronyms is available in the appendix.

Fig. 4. Election forecasts on the party level in Sweden.

system is even more complicated with 9 viable parties competing for power in the latest election in 2014.

To make matters worse, the four right-wing parties have since 2004 stood jointly in the elections as one right-wing bloc known as "the Alliance". This means that the parties have entered elections both as individual actors and as parts of larger pre-electoral coalitions. Such a set-up made it necessary for voters to also think strategically about who they favoured in the left vs. right horse race for power since two clear government alternatives were present. In order to capture this two forecasts have been made in the Swedish case — one at the party level, the other at the level of the blocs.

The party predictions in the three latest elections are available in Fig. 4. Again all the actual results fit within the 95% confidence intervals, but with a few close calls such as the Sweden Democrats (SD) and Greens (MP) in 2014 and the Centre Party (C) in 2006. The average prediction error was 0.78 and thus slightly higher than in the German case. However, the error for the two first elections was only 0.63 percentage points on average so the slightly poorer performance here is in large parts driven by the unexpected results in 2014 from the Sweden Democrats and the Greens.

Interestingly, many of the prediction errors were in the same direction for each election. The large parties were always
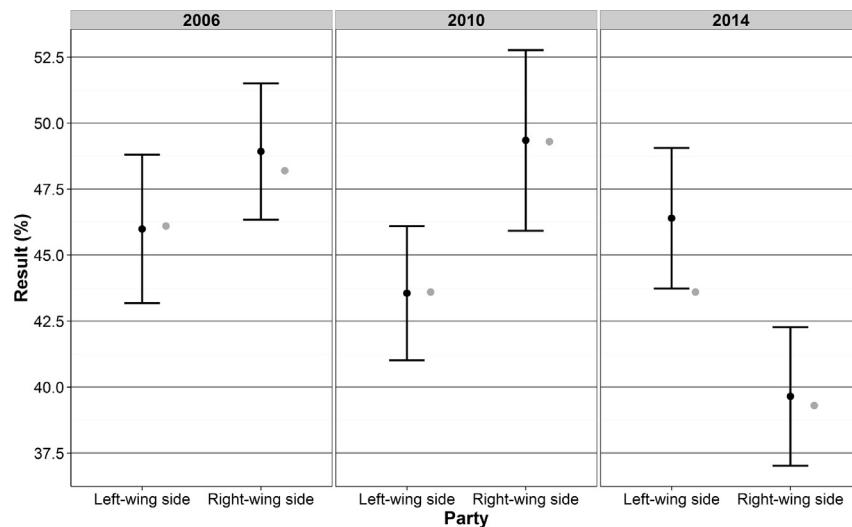
The black dots show the estimated level of support for each party on the day before the respective elections. The surrounding wings are the 95% confidence intervals and the grey dots are the actual election results.

Fig. 5. Election forecasts on the bloc level in Sweden.

underestimated, whereas most of the smaller ones tended to be overestimated. Moreover, the trend was consistent for many of the individual parties. For example, the results of the Greens and the Liberals (FP) were noticeably lower than the forecasts in all the elections while the Social Democrats performed significantly better. All in all though, the standard DLM with the set-up and weights used here seems to work reasonably well even with 9 active players to forecast.

The model generally works even better at the bloc level, as we can see in Fig. 5. Here the average error was only 0.6, even though the confidence intervals for the 2014 election left-wing side missed the actual result by a whisker. This means that the total error in the left vs. right horse race for power is far lower than the sum of the individual errors. This is an encouraging result since the issue of who will get control of the reins of the government after the election is usually a key question to answer in pre-electoral forecasts.

The greater precision at the bloc level here suggests that at least some of the uncertainty in the model comes from voters who switch between parties within the same bloc. It is likely that some voters first make a general decision about which overarching bloc to support, and then at a later stage make decisions about which specific party to give their vote to.[4]

Pre-electoral coalitions are quite common in parliamentary democracies. Golder (2006) finds that out of the 292 elections in her study, 44% had at least one pre-electoral coalition. Around 1/4 of the governments that formed were based on some kind of pre-electoral agreement. If the DLM generally works better when individual party forecast errors cancel out on the aggregate, bloc level, this could be a useful strength that could help improve predictions.

---

[4] At least in the 2014 Swedish election, there was quite a lot of movement among voters between the parties within the two blocs, but hardly any movement between the blocs: http://www.svt.se/svts/article2323667.svt/binary/SVT_ValuResultat_riksdagsval_2014_PK_0914.pdf.

5. Adding a seasonal component to capture re-occurring trends

The applications above show that the DLM is a feasible option for making forecasts in multiparty systems. The average error is relatively small and all the information needed to make the forecast is available before the election. But the forecasts were made using polling data up until the day before the election, and in real world applications we would ideally be able to say something meaningful at an earlier stage.

As argued above, one way of improving the early forecasts is to include a *seasonal component* in the model. If we know that there are predictable swings and that the parties are likely to revert to a particular level, we should be able to foreshadow this trend through seasonal adjustment before it is visible in the polls (Jain, 2001). But the extent to which this boosts accuracy is tied to the magnitude of the seasonal trends. As we saw in Table 1, controlling for seasonality appears to improve model accuracy for most parties generally, but slightly more so in the Swedish case.

In the models shown in Figs. 6 and 7, a second prediction of the German election in 2013 and Swedish election in 2014 has been made, but this time *one month before the elections took place*. For comparison, the model with the seasonal component is compared and contrasted with the prediction from the standard DLM model also used in the preceding section. The standard DLM basically employs its current estimation of the true level of the support for each party as its election forecast, and is thus the 'nowcast' as it stands one month before the election.

In Fig. 6 we can see that controlling for seasonality leads to a measurable, but inconsistent, improvement in the accuracy of the early forecast in Germany. The average error of 1.52 percentage points is about 8% lower than the error of the standard DLM. The main improvement came from the seasonal model's ability to correctly foreshadow that the Greens would revert to a lower level than polling up until that point had suggested. For most other parties the difference between the models was small and for the Left party the seasonal model nudged the prediction slightly in the
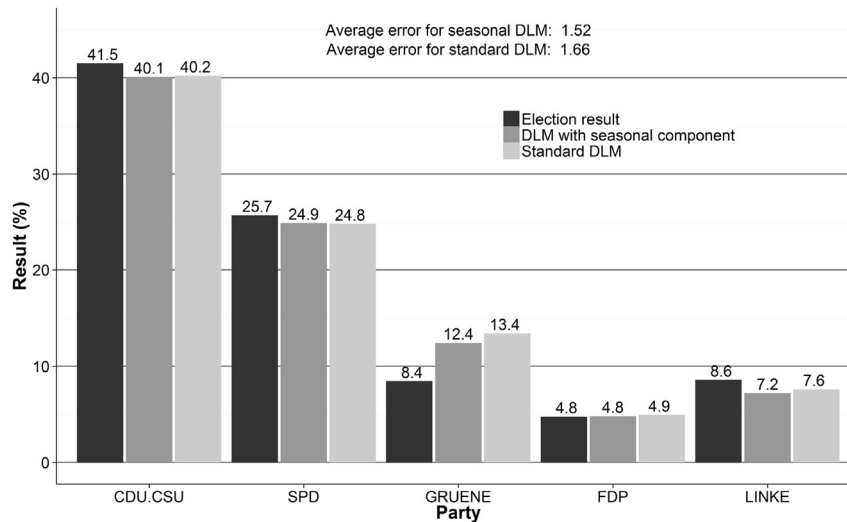
Fig. 6. Comparison of predictions with and without a seasonal component for the German election in 2013.

wrong direction compared to the standard DLM. So overall we can see a clear, but not resounding, improvement when incorporating the seasonal adjustment into the model.

So will forecasts for the Swedish parties, that on average displayed more noticeable seasonal trends, benefit more from adjusting for seasonality? The results in Fig. 7 seem to suggest that they do. The model with the seasonal component performs noticeably better. The mean error for the standard DLM is 1.28 percentage points and this is cut by around 17% by the model with the seasonal component to 1.06. Like in the German case, seasonal adjustment seems to be particularly useful in cases where the standard DLM was far off, e.g. for the Greens (MP) and the Moderates. For the Greens, for example, the standard DLM missed by 3.5 percentage points, but this is cut to 2.1 when we control for seasonality. In total then, one month before the election, the DLM with the seasonal component offered noticeable improvements and was able to predict the result of the seven parties with an error of just over 1 percentage point per party.

Taken together, the results in Germany and Sweden suggest that the DLM does benefit from seasonal adjustment of the time series. And the stronger the historical seasonal trends in the party system, the larger the benefit. Including a seasonal component thus seems to be one way of improving the early forecasting capacity of the DLM.

However, there are also many other ways to improve the standard DLM. In cases where economic and political fundamentals play a predictable role, the DLM can be combined with regression terms for forecasts with better lead time (Jérôme et al., 2013). This should be true especially for the larger parties.
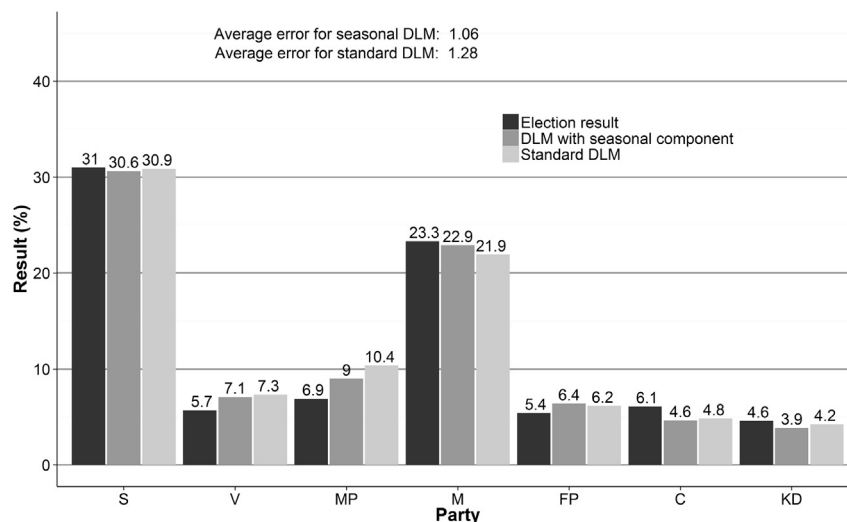

Fig. 7. Comparison of predictions with and without a seasonal component for the Swedish election in 2014.

Another way to improve the forecast could be to tie the predictions for each party more closely together and carry out joint estimations. Since the results for all parties, and the 'other' category, have to sum to 100 this can be placed as an overarching constraint so that the parties are modelled through a joint multivariate time series. Significant improvements are thus still waiting to be made to squeeze even more knowledge out of the data available to us.

6. Conclusion

The goal of this paper has been to apply methods and insights from the extensive literature on electoral forecasting to the 'difficult' multiparty cases of Germany and Sweden. Having a wide range of competing actors and significant changes in electoral fortunes between elections, these countries provide fruitful testing grounds for popular forecasting techniques.

Theoretical arguments were presented against the use of structural models in the multiparty case. Unlike systems with fewer competing parties, in multiparty systems it is often difficult to assign economic and political responsibility to individual cabinet parties (Anderson, 2000; Goodhart, 2009) and it is unclear how, if at all, the various opposition parties will be affected. The frequent emergence of completely new parties for which we have no prior data also make it difficult to accurately forecast elections using this approach.

Instead the polling based *dynamic linear model* was applied. With average errors of 0.68 percentage points in Germany and 0.78 in Sweden in the three latest elections, the application of the DLM shows that forecasts with reasonable accuracy can be made also in volatile multiparty settings. In the Swedish case, where the parties in most cases entered the elections as parts of pre-electoral coalitions, separate forecasts were made on the bloc level. Here a significant share of the prediction errors for individual parties cancelled out since some of the uncertainty in the predictions is between parties within the same ideological bloc. This approach can also be adopted in other multiparty systems where pre-electoral coalitions take place.

Finally, in order to improve the early predictions from the DLM, inspiration was drawn from the economics literature and a seasonal component was added to the model. This term can capture predictable cyclical fluctuations in party support, and incorporating seasonal adjustment into the DLM helped improve the forecast in both Sweden and Germany. A forecast one month before the 2014 election from a standard DLM and from a model with the seasonal component added shows that the component taps into seasonal patterns and cuts the error rate by 8% in Germany and 17% in Sweden.

All in all, the results above are encouraging since they demonstrate that reasonable forecasts are possible also in a multiparty system when a DLM is used. This is an interesting finding in its own right that can be applied not only to party support, but also to other types of regularly occurring polls where we want to make as much as possible of the scarce data available to us. To make the model even more practically interesting, though, in future iterations it needs to do more in the way of explaining why these particular results are likely to occur by also looking into causes, potentially by incorporating regression terms.

Acknowledgements

Appendix

1. Technical model details

The models here were estimated in R using the Bayesian Structural Time Series Package (Scott, 2014; Scott and Varian, 2014) and Markov Chain Monte Carlo simulations. 10 000 simulations were made for each party, with the first 1000 being discarded as 'burn-in' (Jackman, 2009; Kruschke, 2010). The estimated level of party support at any particular time point is taken to be the mean of the posterior distribution for the 10 000 simulations. One prior was consciously specified at the outset, namely the sigma term in equation (1). This was given an inverse gamma prior corresponding to the square root of the sample size in the poll (after having adjusted for polling house reliability). The other terms outlined in equations (2), (3) and (6) (the seasonal adjustment), all assumed to be normal, were given vague inverse gamma priors that were uninformative at the outset. Trace plots and Geweke diagnostics were calculated to check that the chains converged. These are available below.

Another technical decision is what length to assign to the seasonal component. In equation (6) we have both the season (S) and the sub-period (L) that need to be specified. S is simply the term of office which is the same for all parties. The smaller sub-periods in contrast are what we use to capture the seasonal fluctuations and these are calculated based on how quickly cyclical changes in party support are estimated to occur. Like with any time series (unless we have strong theoretical reasons to expect a certain dynamic), the precise length of the seasonal component should be empirically estimated on past data to see which level fits best with the actual development of the series. Here seasonal components were chosen that maximized the historical performance of the model up until the time when the prediction was made (i.e. the components that maximized R^2 in Table 1). The periods were found to generally last between 2 and 6 months.

### 2. Abbreviations of party names

Germany

CDU/CSU = Christian Democratic Union/Christian Social Union
SPD = Social Democrats (Sozialdemokratische Partei Deutschlands)
Gruene = Green Party (die Grünen)
FDP = Liberals (Freie Demokratische Partei)
Linke = Left Party (die Linke)
Piraten = Pirate Party
AfD = Alternative for Germany (Alternative für Deutschland)

Sweden

S = Social Democrats
V = Left party (Vänsterpartiet)
MP = Greens (Miljöpartiet)
M = Moderates
FP = Liberals (Folkpartiet)
C = Centrist party
KD = Christian Democrats
SD = Sweden Democrats
FI = Feminist Initiative

3. Traceplots and Geweke diagnostics

The traceplots are from the final draws of the MCMC chains in the 2013 and 2014 elections.
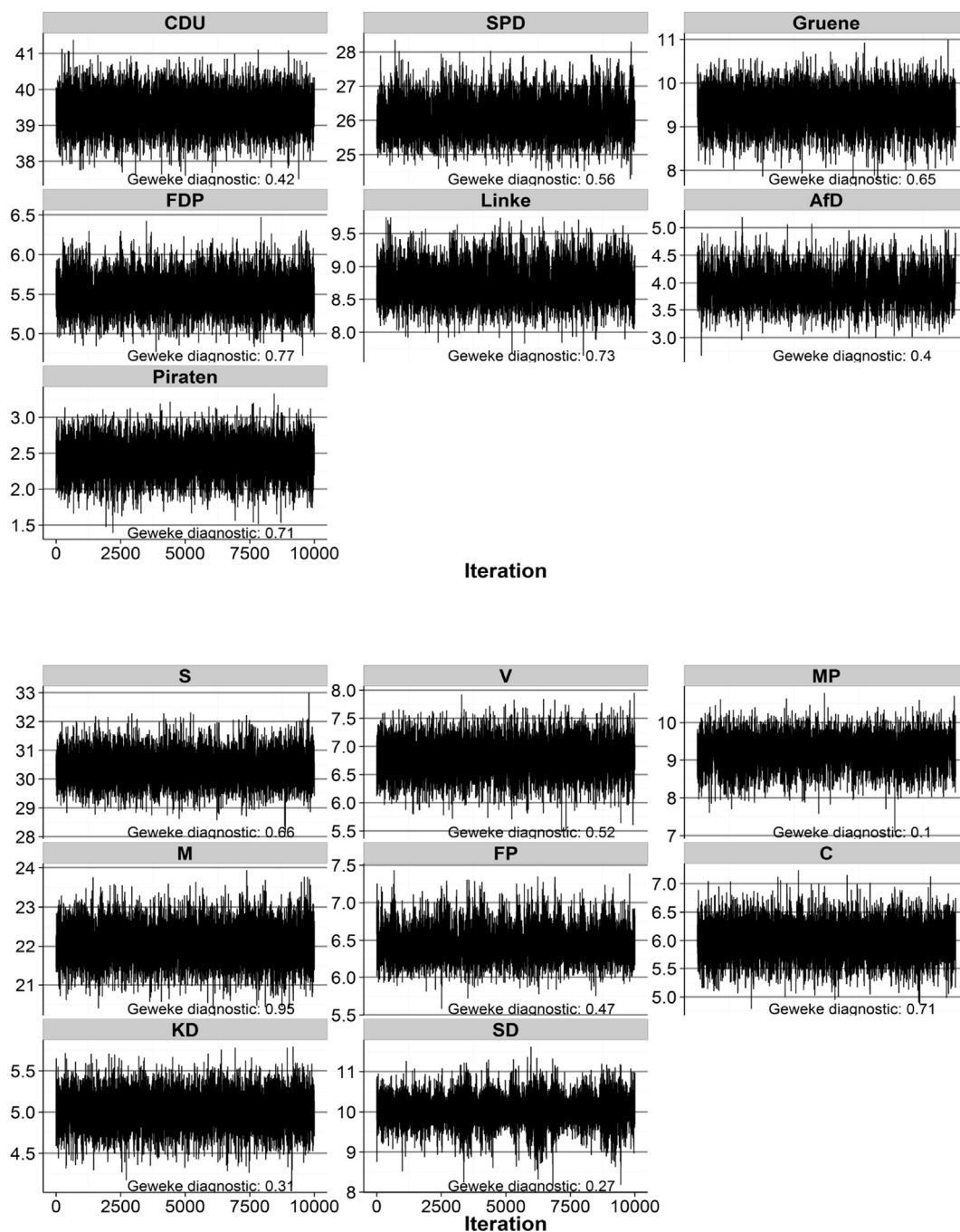
References

Abramowitz, A.I., 2008. Forecasting the 2008 presidential election with the time-–for–change model. PS: Polit. Sci. Polit. 41 (04), 691–695.

Anderson, C.J., 2000. Economic voting and political context: a comparative perspective. Elect. Stud. 19 (2), 151–170.

Bartels, L.M., Zaller, J., 2001. Presidential vote models: a recount. Polit. Sci. Polit. 34 (01), 9–20.

Bell, W.R., Hillmer, S.C., 1984. Issues involved with the seasonal adjustment of economic time series. J. Bus. Econ. Stat. 2 (4), 98–127.

Bengtsson, A., 2004. Economic voting: the effect of political context, volatility and turnout on voters assignment of responsibility. Eur. J. Polit. Res. 43 (5), 749–767.

Carlsen, F., 2000. Unemployment, inflation and government popularityare there partisan effects? Elect. Stud. 19 (2), 141–150.

Fisher, S.D., Ford, R., Jennings, W., Pickup, M., Wlezien, C., 2011. From polls to votes to seats: forecasting the 2010 british general election. Elect. Stud. 30 (2), 250–257.

Foucault, M., Nadeau, R., 2012. Forecasting the 2012 french presidential election. PS: Polit. Sci. Polit. 45 (02), 218–222.

Fredén, A., 2014. Threshold insurance voting in pr systems: a study of voters strategic behavior in the 2010 swedish general election. J. Elect. Public Opin. Parties 24 (4), 473–492.

Gibson, R., Lewis-Beck, M.S., 2011. Methodologies of election forecasting: calling the 2010 uk hung parliament. Elect. Stud. 30 (2), 247–249.

Golder, S.N., 2006. Pre-electoral coalition formation in parliamentary democracies. Br. J. Polit. Sci. 36 (02), 193–212.

Goodhart, L., 2009. Context, Clarity, and Signals: Economic Voting for Political Parties (Unpublished manuscript).

Graefe, A., 2015. German election forecasting: comparing and combining methods for 2013. German Polit. 24 (2), 195–204.

Harrison, J., West, M., 1997. Bayesian Forecasting and Dynamic Models. Springer, New York.

Healy, A., Lenz, G.S., 2014. Substituting the end for the whole: why voters respond primarily to the electionyear economy. Am. J. Polit. Sci. 58 (1), 31–47.

Hibbs Jr., D.A., 2000. Bread and peace voting in us presidential elections. Public Choice 104 (1–2), 149–180.

Hyndman, R.J., Athanasopoulos, G., 2014. Forecasting: Principles and Practice. OTexts.

Jackman, S., 2005. Pooling the polls over an election campaign. Aust. J. Polit. Sci. 40 (4), 499–517. http://dx.doi.org/10.1080/10361140500302472.

Jackman, S., 2009. Bayesian Analysis for the Social Sciences, vol. 846. John Wiley & Sons.

Jain, R.K., 2001. State space model-based method of seasonal adjustment. a. Mon. Lab. Rev. 124, 37.

Jérôme, B., Jérôme-Speziari, V., Lewis-Beck, M.S., 2013. A political-economy forecast for the 2013 german elections: who to rule with angela merkel? PS: Polit. Sci. Polit. 46 (03), 479–480.

Jungerstam-Mulders, S., 2006. Post-communist EU Member States: Parties and Party Systems. Ashgate Publishing Company.

Kayser, M.A., Leininger, A., 2013. A Benchmarking Forecast of the 2013 Bundestag Election. Unpublished manuscript. Available at: http://goo.gl/1lSIEX.

Kitagawa, G., Gersch, W., 1984. A smoothness priorsstate space modeling of time series with trend and seasonality. J. Am. Stat. Assoc. 79 (386), 378–389.

Kruschke, J., 2010. Doing Bayesian Data Analysis: a Tutorial Introduction with R. Academic Press.

Lewis-Beck, M.S., 2005. Election forecasting: principles and practice. Br. J. Polit. Int. Relat. 7 (2), 145–164.

Lewis-Beck, M.S., Bellucci, P., 1982. Economic influences on legislative elections in multiparty systems: France and italy. Polit. Behav. 4 (1), 93–107.

Lewis-Beck, M.S., Paldam, M., 2000. Economic voting: an introduction. Elect. Stud. 19 (2), 113–121.

Linzer, D.A., 2013. Dynamic bayesian forecasting of presidential elections in the states. J. Am. Stat. Assoc. 108 (501), 124–134.

Lock, K., Gelman, A., 2010. Bayesian combination of state polls and election forecasts. Polit. Anal. 18 (3), 337–348.

Mitchell, P., Nyblade, B., 2008. Government formation and cabinet type. In: Strøm, K., Müller, W., Bergman, T. (Eds.), Cabinets and Coalition Bargaining - the Democratic Life Cycle in Western Europe. Oxford University Press, Oxford.

Nadeau, R., Lewis-Beck, M.S., Blanger, E., 2013. Economics and elections revisited. Comp. Polit. Stud. 46 (5), 551–573.

Narud, H.M., Valen, H., 2008. Coalition membership and electoral performance. In: Strøm, K., Müller, W., Bergman, T. (Eds.), Cabinets and Coalition Bargaining - the Democratic Life Cycle in Western Europe. Oxford University Press, Oxford.

Norpoth, H., Gschwend, T., 2010. The chancellor model: forecasting german elections. Int. J. Forecast. 26 (1), 42–53.

Paldam, M., 1991. How robust is the vote function? a study of seventeen nations over four decades. Econ. Polit. Calc. Support 9–31.

Petris, G., Petrone, S., Campagnoli, P., 2009. Dynamic Linear Models with R. Springer, New York.

Pickup, M., Johnston, R., 2007. Campaign trial heats as electoral information: evidence from the 2004 and 2006 canadian federal elections. Elect. Stud. 26 (2), 460–476.

Saalfeld, T., 2008. Institutions, chance and choices: the dynamics of cabinet survival. In: Strøm, K., Müller, W., Bergman, T. (Eds.), Cabinets and Coalition Bargaining - the Democratic Life Cycle in Western Europe. Oxford University Press, Oxford.

Sanders, D., 1991. Government popularity and the next general election. Polit. Q. 62 (2), 235–261.

Scott, S.L., 2014. Bayesian Structural Time Series in R. http://cran.r-project.org/web/packages/bsts/bsts.pdf.

Scott, S.L., Varian, H.R., 2014. Predicting the present with bayesian structural time series. Int. J. Math. Model. Numer. Optim. 5 (1), 4–23.

Severini, T.A., 2005. Elements of Distribution Theory, vol. 17. Cambridge University Press, Cambridge.

Silver, N., 2012. The Signal and the Noise: Why So Many Predictions Fail-but Some Don't. Penguin.

Stoltenberg, E., 2013. Bayesian Forecasting of Election Results in Multiparty Systems. Unpublished manuscript. Available at: https://www.duo.uio.no/handle/10852/36975?show=full.

Sundell, A., Lewis-Beck, M.S., 2014. Forecasting the 2014 Parliamentary Election in Sweden. Available at: SSRN 2450229.