Appendix for Estimating Coalition Majorities during Political Campaigns based on Pre-election Polls

Contents

1 Updating the Dynamic Multinomial-Dirichlet Model

The multinomial density is given by:

$$p(\boldsymbol{Y}_t|\boldsymbol{\pi}_t) = \frac{N_t!}{\prod_{k=1}^K y_{kt}!} \prod_{k=1}^K \pi_{kt}^{y_{kt}}. \tag{1}$$

the posteriori at $t-1$ is Dirichlet distributed based on information available until this time time-point D_{t-1}.

$$p(\boldsymbol{\pi}_{t-1}|D_{t-1}) = \frac{1}{C(\boldsymbol{\alpha}_{t-1})} \prod_{k=1}^K \pi_{k,t-1}^{\alpha_{k,t-1}}, \tag{2}$$

where $C(\boldsymbol{\alpha}_{t-1})$ is a known normalizing constant and $\boldsymbol{\alpha}_{t-1}$ is know from past updating steps. The power-discount model supposes that

$$p(\boldsymbol{\pi}_t|D_{t-1}) \propto p(\boldsymbol{\pi}_{t-1}|D_{t-1})^{\delta}, \tag{3}$$

allowing to construct the posterior distribution of π_t at t using Bayesian Updating:

$$p(\boldsymbol{\pi}_t|\boldsymbol{Y}_t, D_{t-1}) \propto p(\boldsymbol{Y}_t|\boldsymbol{\pi}_t)p(\boldsymbol{\pi}_t|D_{t-1}). \tag{4}$$

which after manipulation leads to

$$\begin{aligned} p(\boldsymbol{\pi}_t|\boldsymbol{Y}_t, D_{t-1}) &\propto \prod_{k=1}^K \pi_{kt}^{y_{kt}} \left(\prod_{k=1}^K \pi_{kt}^{\alpha_{k,t-1}} \right)^{\delta} \\ &\propto \prod_{k=1}^K \pi_{kt}^{y_{kt}+\delta\alpha_{k,t-1}}, \end{aligned} \tag{5}$$

which shows that the posterior distribution is Dirichlet distributed with $\boldsymbol{\alpha}_t = \boldsymbol{Y}_t + \delta\boldsymbol{\alpha}_{t-1}$.

2 Results German Federal Elections 1994-2017

Table 1 reports two ordinary least square (OLS) models that depict the evolution of the coalition options measurement over time. All models include coalition-election fixed effects (not reported in the table). With the coalition-election fixed effects, the model analyses the changes in probability for a majority of each coalition option over the last year. Model 1 and 2 studies the dynamics for the same type of coalitions as in the main text: Coalitions with final seat share below 40%, between 40%-45%, between 45% and 50%, between 50% and 55%, between 55% and 60% and above 65%. Model 1 includes linear trends for those different coalitions and echoes the findings from the main text. Coalition below 40% of final seat share, and above 60% there is no substantial effect of time. For the coalitions that at the end clearly manage to pass the cut off (55%-60%), or clearly did not (40%-45%) the measure crystallizes this over time. E.g. the effect of weeks for a coalition with seats of 55%-60% is negative (direct effect plus indirect effect $\sim$ -0.27), indicating that the expected value is higher one week before the election than 10 weeks before the election. For coalition with seat share between 40%-45% the estimate is positive suggesting the inverse. For coalitions that fall just below the cut-off (seat share 45%-50%) the linear effect is reversed (direct effect plus indirect effect $\sim$ -0.37) compared to coalitions that clearly fail to attain a majority. This implies that the closer to the election higher the chances for the majority. When considering the specific intercept for those cases, this actually mirrors the pattern described in the main text: for those cases, the chances for a coalition majority tend towards 50% in expectation. The same holds for seat share of 50%-55%. Model 2 reports that the results do not substantially change when including the second polynomial. The model fit is almost identical suggesting that the interpretation of the linear model suffices to understand the dynamics.

	Model 1	Model 2
Seat Share 40%-45% $\times$ Weeks to election	0.12	0.65
	(0.12)	(0.49)
Seat Share 45%-50% $\times$ Weeks to election	-0.37^{**}	-0.53
	(0.12)	(0.49)
Seat Share 50%-55% $\times$ Weeks to election	0.41^{***}	0.71
	(0.12)	(0.48)
Seat Share 55%-60% $\times$ Weeks to election	-0.27^{*}	-1.00^{*}
	(0.12)	(0.49)
Seat Share >60% $\times$ Weeks to election	0.00	0.00
	(0.12)	(0.49)
Seat Share 40%-45% $\times$ Weeks to election squared		-0.53
		(0.48)
Seat Share 45%-50% $\times$ Weeks to election squared		0.16
		(0.48)
Seat Share 50%-55% $\times$ Weeks to election squared		-0.30
		(0.47)
Seat Share 55%-60% $\times$ Weeks to election squared		0.73
		(0.48)
Seat Share >60% $\times$ Weeks to election squared		-0.00
		(0.48)
Seat Share 45% $-$ 50%	2.08^{***}	2.11^{***}
	(0.10)	(0.13)
Seat Share 50% $-$ 55%	1.69^{***}	1.64^{***}
	(0.10)	(0.13)
Seat Share 55% $-$ 60%	2.03^{***}	2.15^{***}
	(0.11)	(0.14)
Seat Share >60%	1.48^{***}	1.48^{***}
	(0.09)	(0.12)
Weeks to election	-0.00	-0.00
	(0.11)	(0.46)
Weeks to election squared		0.00
		(0.45)
R^2	0.90	0.90
Adj. R^2	0.90	0.90
Num. obs.	2142	2142
RMSE	0.23	0.23

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. The model includes coalition-election fixed effects that are not reported in the table.

Table 1: Evolution of coalition options. Results from OLS regression with coalition majority probabilities as the dependent variable.

3 Coalition Majorities in the 2017 Election

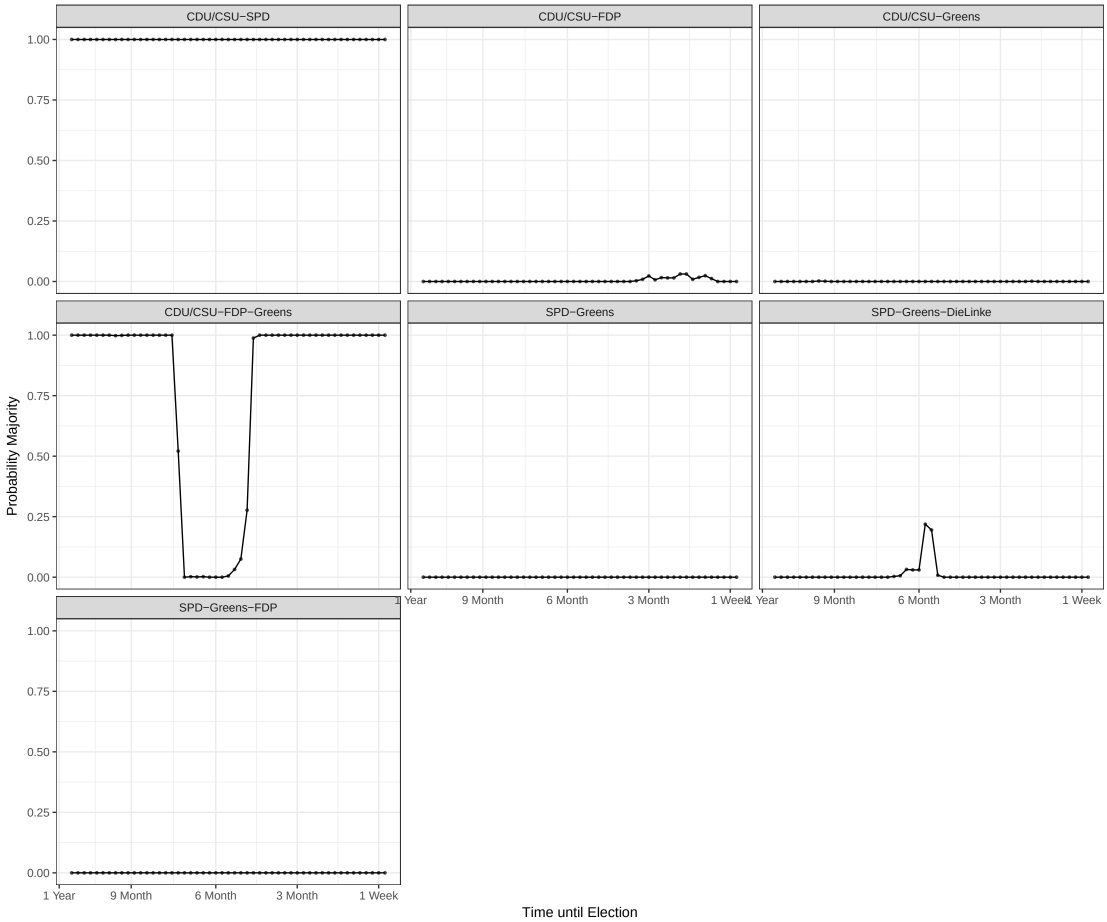

Figure 1: Estimated coalition majorities for ideological connected coalitions during the last year before German Federal Elections 2017

4 Discount Factor Estimates

	Discount-factor	S.E.
1994	0.60	0.02
1998	0.59	0.02
2002	0.74	0.02
2005	0.35	0.01
2009	0.39	0.01
2013	0.32	0.01
2017	0.37	0.01

5 Comparisons to Alternative Measurements

5.1 Directly Calculated from the Polls

An alternative measurement for the majority probabilities can be generated from the polls directly. For this purpose one can add up the shares of coalition partners and calculate the standard error for the proportion. This works for each individual poll, but also for the weekly weighted averages, as we used in the main text. For example, if we are interested in the probability that a coalition between party 1 and 2 will get a majority, we can add up their shares $c_{12,t} = y_{1,t} + y_{2,t}$ and calculate the sampling error given the number of respondents: $s_{c_{12,t}} = \sqrt{\frac{c_{12,t}*(1-c_{12,t})}{N_t}}$. Sampling theory tells us that the probability that the share of the coalition is above 50% is than given by $Pr[c_{12,t} > 0.5] = \Phi\left(\frac{c_{12,t}-0.5}{s_{c_{12,t}}}\right)$. This yields a measurement derived from the poll results directly.

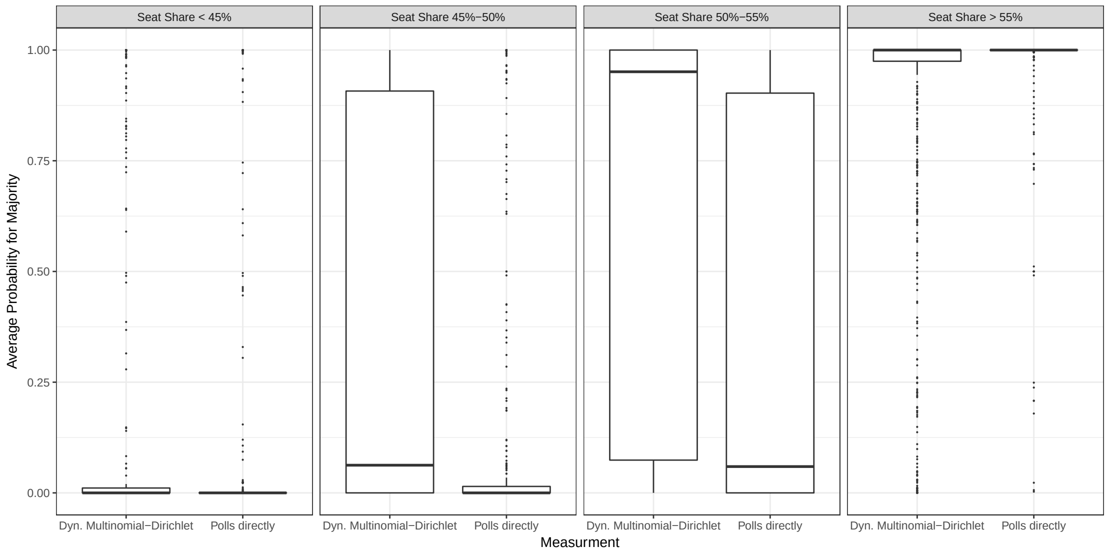

Figure 2: Comparison to alternative measurement of coalition majorities. The alternative measurement is calculated directly from standard deviation of the weekly proportions.

In the following, we will compare the alternative measurement with the measurement from the dynamic multinomial-Dirichlet model. Figure 2 shows the distribution of the coalition majority probabilities for the ideologically connected coalitions in the German Federal Elections 1994-2017, once for the alternative measurement (Polls directly) and the measurement from the main text (Dyn. Multinomial-Dirichlet). The box-plots are displayed for different final seat-margins. For clear majorities (Seat Share > 55%) and clear none-majorities (Seat Share < 45%) the measurements do equally well. Both have a set of significant outliers, but the majority of cases is assigned a small chance or a high chance respectively. With fewer outlines, the polls direct measurement is more certain about the outcome for those cases. While this is valuable for clear cases, it can be dangerous for cases where the seat margin is not as clear. This can most clearly be seen for the coalitions that just managed to secure a majority, with a seat share of 50% - 55%. Here the median from the polls directly measurement is with 0.1 far too low for coalitions that made it. For the cases that fall just below the cut-off (Seat share of 45% - 50%) the measurement is accurate, though. The dynamic Multinomial-Dirichlet is more conservative as that the distribution is more variable. With the median of the estimates tending towards the right direction, we think that this is actually a more valid estimate of the uncertainty. At 50% seat share, the probability should also be 50% of gaining a majority of seats. While the analysis in the main text highlights that the dynamic Multinomial-Dirichlet picks this up, the alternative measurement fails to do so.

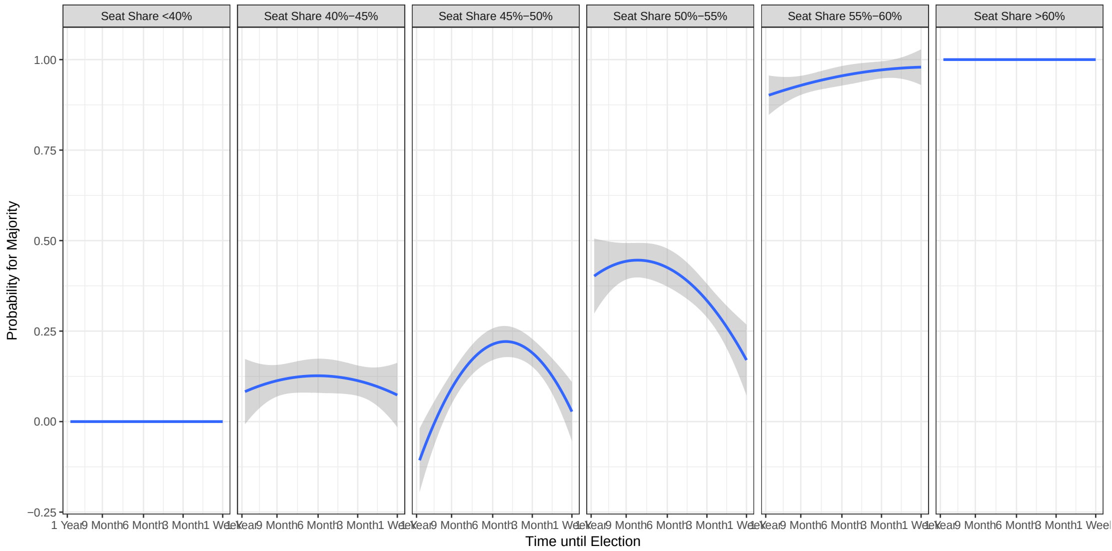

Figure 3: Evolution of coalition options within the last year for different final seat share margins

To underscore this point further, Figure 3 replicates the figure in the main text with the alternative polls directly measurement. The evolution of coalition majorities looks less distinct: While the measurement does well in estimating clear cases, for the cases in the middle the probabilities do not develop as expected. Especially, in closes cases, the pattern highlights that the poll direly measurement does not capture the uncertainty accurately. Neither do both coalitions converge to a fifty-fifty chance, nor does the just-above majority cut-off cases develop in a meaningful direction.

5.2 Dynamic Linear Model

In this section, we describe the results from a dynamic linear model and compare them to the Multinomial-Dirichlet specification in the main text. Similar to Walther (2015) and Jackman (2005), we employ an independent random walk model that can be estimated for each party-at-a-time to specify the dynamic evaluation of latent support for the parties. A detailed introduction to dynamic linear models can be found in West and Harrison (1997) and Petris et al. (2008). Here, we briefly outline the model specification for a random walk model.

We define s_{jt} as the share of respondents that intend to vote for party $k \in (1, \ldots, K)$ in poll of size N_t at time point t. We define an observation equation in which the observed vote share from the poll is related to the underlying support θ_{jt} and additional measurement error.

$$y_{jt} = \theta_{jt} + v_{jt} \quad , where \quad v_{jt} \sim N(0, V_{tj}). \tag{6}$$

The measurement error v_{jt} is assumed to be normal distributed and centered around zero with variance of V_{tj}. As in the standard applications of the model to pre-election polls (see e.g. Jackman, 2005), the standard error for the share can be used to define the measurement variance: $V_{tj} = \frac{s_{jt}(1-s_{jt})}{N_{jt}}$. Next, we assume that latent support evolves over time using an evolution equation:

$$\theta_{jt} = \theta_{jt-1} + w_{jt} \quad , where \quad w_{jt} \sim N(0, W_j). \tag{7}$$

The evolution equation specifies how the latent support for party j evolves from $t-1$ to t. The random walk specification supposes that the latent support at t is equal to the support at $t-1$ plus random deviations w_{jt}. The random deviations are assumed to be normally distributed with evolution variance W_j.

To obtain filtered latent support for a party over the time period under study from the dynamic linear model we can employ the Kalman filter (Kalman, 1960). The Kalman filter requires an estimate of the evolution variance W_j, which we obtain using Maximum Likelihood. The Maximum Likelihood estimation of W_j is similar to the estimation of the discount factor in the dynamic Multinomial-Dirichlet in the main text. For a detailed discussion of the procedure of using Maximum Likelihood to estimate unknown parameters in dynamic linear models please refer to Petris (2010, p.144ff). We implement the estimation using the dlm package in R Petris (2010).

We estimate the model for the same parties and elections than in the main text. The polls are aggregated to weeks (using a weighted sum of the polls published in a particular week) to make the results comparable to the results from the dynamic Multinomial-Dirichlet model.

The results from the model are portrayed in Figure 4. At first, the resulting latent support levels appear similar to the one obtained in the main text using the dynamic Multinomial-Dirichlet model. In particular, in the cases where we observe a lot of polls with a large number of respondents, the latent support closely follows the poll average with relatively little uncertainty around the mean latent support. The latent support

Figure 4: Latent support for the last year of the German Federal Elections 1994-2017 from Dynamic Linear Model.

adapts even stronger to the oberved poll results. The most striking deviation emerges in situations where poll results are rare, for example in the first months of the 2002 year election year. The uncertainty increases whenever there is no new information. Especially, the uncertainty about the support for the SPD and the CDU/CSU increases strongly. In the dynamic Multinomial-Dirichlet model the increase in uncertainty for this period is not as strong. Similar patterns emerge in weeks with a small number of polls in the 1998 election, but also in some weeks of the 2009 and 2013 election.

There are two reasons why the dynamic Multinomial-Dirichlet model of the main text is preferred to estimate the majority probabilities. First, the separate dynamic linear models do not hold a constraint that the latent support of all parties should sum to one. This will result in sums that are larger or smaller than the one when combining the latent supports from the party-by-party dynamic linear models. To see this, we plot the sum of the latent support from the party-by-party model over time for the different elections under study in Figure 5. For each week we sum up draws from the posterior distribution for each party. The dots indicate the mean over the different draws along the 95% range. While the means arguably level around one, we can observe be strong

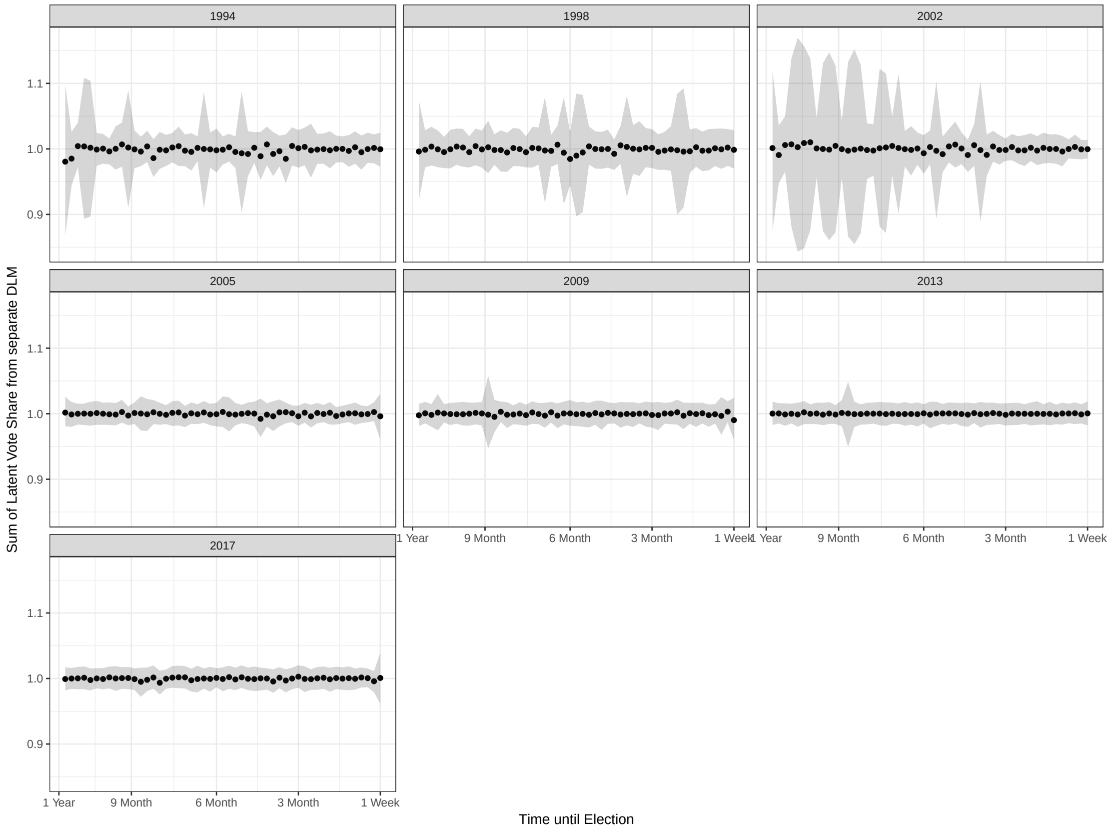

Figure 5: Comparison to alternative measurement of coalition majorities. The alternative measurement is calculated directly from standard deviation of the weekly proportions.

deviations, especially in weeks with a small number of survey respondents. The most challenging period is the first three months of the 2002 election year, where the sums range from below 0.9 to above 1.1. Whenever the uncertainty increases for the latent support, the sums more strongly deviate from one. In other cases, with informative poll results, the sums are closer to one.

Second, this can matter when using the model to analyze majority probabilities. In Figure 6 we compare the coalition majority probabilities for coalitions obtained from the dynamic Multinomial-Dirichlet model in the main text to the probabilities from the dynamic linear model. We employ the same majority generating function as in the main text, using the random draws from the posterior distribution of the separately estimated dynamic linear models for the respective parties. To validate the usefulness of the two measures, we plot the range of probabilities over four different final seat share ranges of the respective coalitions. To the left, we see that both methods give low majority probabilities for coalitions that in the end hold below 45% of seats. The distribution of majority probability for a coalition that just falls below the threshold

(45% - 50%) is almost similar between the two, with the dynamic Multinomial-Dirichlet model given a slightly wider spread. Both indicate a median value clearly below 10%, which speaks for the validity of the methods in indicating coalitions that will fail to secure a parliamentary majority. For coalitions that hold seat shares between 50% and 55% the dynamic Multinomial-Dirichlet model indicates a median majority of above 90%, while the Dynamic linear model the median majority probability from the dynamic linear model is close to 50%. Here, the dynamic Multinomial-Dirichlet model appears to be more consistent, as the distribution mirrors the inverse of the spread for coalitions just below 50%. For a coalition that clearly manages to gain a seat majority both methods again give high chances, with the dynamic linear model presenting slightly more spread.

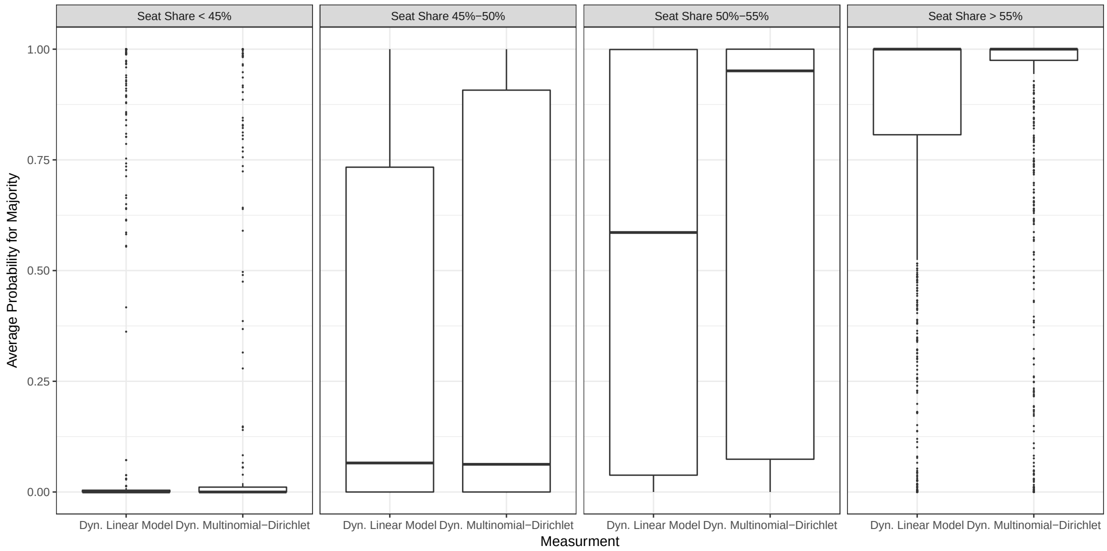

Figure 6: Comparison to alternative measurement of coalition majorities. The alternative measurement is calculated using a set of separatly estimated dynamic linear models.

References

Simon Jackman. Pooling the polls over an election campaign. *Australian Journal of Political Science*, 40(4):499–517, 2005.

R E Kalman. A New Approach to Linear Filtering and Prediction Problems 1. *Journal of Fluids Engineering*, 82:35–45, 1960.

Giovanni Petris. An R Package for Dynamic Linear Models. *Journal of Statistical Software*, 36(12):1–16, 2010.

Giovanni Petris, Sonia Petrone, and Patrizia Campagnoli. *Dynamic Linear Models with R*. 2008.

Daniel Walther. Picking the winner(s): Forecasting elections in multiparty systems. *Electoral Studies*, 40:1–13, 2015.

Mike West and P Jeff Harrison. *Bayesian Forecasting and Dynamic Models, Bayesian Forecasting and Dynamic Models.* 1997.