Online Appendix to:

"A General Approach to Measuring Electoral Competitiveness for Parties and Governments"

Axel Cronert*and Pär Nyman†

The replication materials for this paper can be found at Code Ocean `https://doi.org/10.24433/CO.7984668.v1` and at the *Political Analysis* Dataverse: `https://doi.org/10.7910/DVN/YMQFYB`.

*Department of Government, Uppsala University. Email: axel.cronert@statsvet.uu.se.
†Department of Government, Uppsala University. Email: par.nyman@statsvet.uu.se.

A1 Application on Swedish local governments

A1.1 Modeling government formation

Having surveyed much of the literature on government formation, we have identified 32 factors that have been claimed to be important in government formation processes. While we have tried to include as many factors as possible, we have omitted factors that either are irrelevant to the Swedish local context, do not vary between municipalities, or are impossible for us to measure.[1] The 32 factors are all operationalized as a variable x_{jiy} of the potential government coalition j in the government formation opportunity occurring in municipality i following the election in year y. Below, we divide the factors into four different groups. The number of each factor corresponds to that reported in Table A1 below.

The factors in the first and largest group are related to the size and ideological proximity of the parties in the potential government. Among the earliest hypotheses is that potential governments are less likely to form if they control only (1) a *minority* of the seats (see Martin and Stevenson 2010). An early refinement to this hypothesis suggests that those majority cabinets are more likely to form that are (2) *minimal-winning coalitions*, meaning a coalition from which no partner is unnecessary to the majority status of the coalition (Morgenstern and Von Neumann 1953). In such cabinets, the benefits associated with being in power are shared by as few partners as possible. As there are often more than one minimal-winning coalition, several further refinements have been made to this theory. Among them is that minimal-winning coalitions are more likely to form if they are (3) *connected*, in the sense that they contain only ideologically adjacent parties (Axelrod 1970), if they have (4) the *narrowest ideological range* (De Swaan 1973), if they contain (5) the *fewest parties* among available minimal-winning coalitions, or if they are (6) the *minimum-winning coalition* meaning that they control the smallest share of seats among available minimal-winning coalitions (Laver and Schofield 1990).

As regards cabinet size, Glasgow and Golder (2015) include the two variables (7) *cabinet seat share* and (8) *cabinet seat share squared* to capture an idea strongly related to the minimal-winning coalition theory, namely that both small minority cabinets and large surplus majority cabinets are less likely to form than cabinets whose seat share is slightly above 50 percent. As far as (9) *number of parties* are concerned, the standard account suggests that, ceteris paribus, cabinets are more likely to form the fewer parties they include (Glasgow and Golder 2015).

Later policy-centered theories, in the vein of Axelrod and De Swaan, have suggested that irrespective of their size, ideologically divided cabinets should be less attractive to potential coalitions partners than more compact ones. Following Martin and Stevenson (2001), we account for this by including a measure of (10) the *ideological range* between the two most distant parties in the potential government

[1] Examples include the formally appointed formateur, of which there are none in Swedish municipalities (Bäck 2003), the existence of an investiture vote, for which there is no variation in the sample (Martin and Stevenson 2001), or the circumstances that ended the tenure of the previous cabinet, for which there is little data at the local level (Martin and Stevenson 2010).

along the left–right continuum.[2] A related theory by Laver and Schofield (1990) holds that potential minority cabinets should be more likely to form the larger the ideological divisions within the majority opposition they would face (Martin and Stevenson 2001). We capture this by including a measure of the (11) *ideological range between the most distant parties in the opposition* and (12) by *interacting* that measure with the minority cabinet variable mentioned above.

A more recent theory about the relevance of the cabinet's ideological composition is provided by Glasgow and Golder (2015). They suggest that as the ideological distance of the partners in a potential coalition to the median left–right position in the parliament increases, the probability that these coalitions will form decreases. Following Glasgow and Golder (2015), we therefore include a measure of (13) the *ideological distance from the median* computed as the weighted mean ideological distance between the partners in the coalition and the median, with the weights for each party based on its seat share.

Several theories suggest that potential governments are more likely to form if they contain particular parties that have a strong bargaining position. Among those most frequently occurring in the literature are (14) the party that controls the *median seat* on the left–right dimension, and (15) the *largest party*. For obvious reasons, the potential government with the strongest bargaining position of them all is one that consists of (16) a *single majority party*. This is not a rare phenomenon in Swedish municipalities; out of the approximately 1,700 governments in place between 1998 and 2018, 6 percent were single party majority governments (SKL 2018).

The literature has also considered that certain parties may make potential governments less attractive to join, because doing so will incur high electoral costs to prospective partners. This is particularly the case with parties that promote 'anti-system' political views (Martin and Stevenson 2010). In the context and period studied here, this phenomenon should be properly captured by a variable indicating whether or not the potential government includes the radical right party (17) the *Sweden Democrats (SD)*. In a local setting such as ours, it also makes sense to include a variable on whether or not the potential government includes (18) a *local party*, since these are often elected on an anti-establishment agenda.

A second set of factors relate to the incumbent and to its recent electoral performance. First, incumbency theory holds that for a number of reasons a potential government is more likely to form if it is constituted by the same set of parties that formed (19) the *incumbent government* (see Martin and Stevenson 2001). Considering the greater discretion that comes with leading a government,

[2]This variable scores a 0 for single-party cabinets. We base the measure on national-level party positions from the Chapel Hill expert survey (Polk 2017). For 2008 and 2012, survey data on local politicians from the KOLFU surveys (Gilljam et al. 2015; Karlsson and Gilljam 2018) make it possible to retrieve party positions at the local level as well. In line with results from Bäck (2003) on earlier data, these party positions are strongly correlated with the national-level party positions ($r = 0.83$). Therefore, we choose to rely on the national-level data that have the advantage of being available for a longer time period. We ascribe the scores for 1999 and 2014 to the election periods 1994–1998 and 2015–2018, respectively. The empirical distribution of these scores ranges from 1.43 (the Left Party in 2010) to 8.7 (the Sweden Democrats in 2010).

the same logic would suggest that a potential government is more likely to form if it includes (20) the *party of the chief executive* (Martin and Stevenson 2010); in our case, the Mayor. In a recent contribution to the incumbency theory, Glasgow and Golder (2015) distinguishes between the incumbent coalition and the incumbent parties. Their analysis specifically links the incumbency advantage to the coalition; in case it does not re-form as a whole, other potential coalitions that include (21) *one or more but not all incumbent parties* are less likely to form.

Martin and Stevenson (2010) have also hypothesized that the recent electoral performance of the incumbent affects its prospects of returning into office because parties should be more willing to join or re-form a coalition that has performed well, even after considering the ways that the election may have changed the seat distribution. As a measure of (22) *electoral performance*, we calculate the average seat change experienced by each potential government between the most recent election and the election prior to that. Because this effect is expected to matter particularly for the incumbent government, we include (23) an *interaction* between the electoral performance measure and the incumbent government indicator.

A third set of factors have to do with parties' pre-electoral relationships. Martin and Stevenson (2010) have shown that parties who have governed together in the past are more likely to do so again, presumably because cooperation promotes mutual trust and understanding. Accordingly, we construct a measure of the (24) *familiarity* of the partners in each potential coalition, following the approach developed by Martin and Stevenson (2010) in which past governing partnership is discounted relative to more current partnership. A related but distinct factor is discussed by Bäck and Lindvall (2015), who create a measure of a coalition's (25) *commitment potential*, based on the historical patterns of cooperation among the involved parties. In short, coalitions that include parties who have recently been in government with a different set of coalition partners or who have recently governed alone are conceived as having lower commitment potential. Whereas Bäck and Lindvall (2015) do not use their measure to study government formation, we include it here because we might expect that potential coalitions with a low commitment potential are associated with higher costs and are less likely to form.

Martin and Stevenson (2010) furthermore hypothesize that if parties make pre-electoral commitments to form certain coalitions, these coalitions are more likely to form. Lacking data on actual statements on the local level, we resort to including dummy variables for the two long-standing political blocs within Swedish politics, namely (26) the *right-wing bloc* consisting of the Center Party, the Christian Democrats, the Conservative Party, and the Liberal Party, and (29) the *left-wing bloc* consisting of the Social Democrats and the Left Party.[3] Inspired by Skjæveland et al. (2007), we also include two additional variants of each of these bloc dummies. The (27, 30) *bloc-plus* variables score 1 if the coalition contains all parties from the bloc, including cases where it contains all parties from the bloc and one or more other parties. The (28, 31) *bloc-minus* variables score 1 if the

[3]In recent years, the Green Party has moved closer to the left-wing bloc, especially so at the national level. However, at the municipality level the Green Party is best perceived as bloc independent (Folke 2014).

coalition contains some but not all parties from the bloc, as well as one or more other party. The expectation is that breaking a pre-electoral coalition is a less appealing option in a government formation process than extending a pre-electoral bloc with a party from the outside. In addition, we include a variable indicating a (32) *bloc-transcending coalition* that includes parties from both blocs.

The first column of A1 reports the output from a conditional logit model that includes these 32 factors, run on the realized governing coalition outcomes[4] of the 1,719 government formation opportunities in the Swedish municipalities between 1998 and 2018. For the sake of brevity, we refrain from interpreting individual coefficients here. However, it is clear that the model provides a reasonably good fit to the data, as indicated by the Pseudo R^2 parameter of 0.58.

A1.1.1 A reduced model of government formation

Many of the aforementioned variables are computationally demanding and require data that might be difficult for researchers to get a hold of. Therefore, as a robustness check reported in the second column of Table A1, we have run a reduced version of our model that includes no more than 10 key variables that should be comparably easy to compile. A comparison shows that the predictive power of the model decreases, as the Pseudo R^2 drops from 0.58 to 0.49. As reported in Section 4 in the paper and in Section A1.3 below, our approach performs better the more data is used in this step, but also the reduced model represents a significant improvement over previous measures in terms of predicting re-election.

[4] As described in the paper, the governing coalition in Swedish municipalities refers to the majority coalition (or party) which appoints the chairs and vice chairs of the committees in the local assembly. It should be noted that each local council is also required to appoint a formal executive committee (*kommunstyrelse*). Appointments to the executive committee are mostly made proportionally, which means that (almost) all parties tend to be represented. However, among scholars as well as the municipalities themselves, the parties that hold committee chairs and vice chairs are commonly regarded as the governing parties (e.g., Bäck 2003; SKL 2018).

Table A1: The government formation models

	(1) Full		(2) Reduced	
1. Minority cabinet	−1.933	(0.276)	−0.645	(0.165)
2. Minimal-winning coalition	1.101	(0.120)	1.039	(0.100)
3. MWC: Connected	−0.173	(0.143)		
4. MWC: Narrowest ideological range	−0.293	(0.142)		
5. MWC: Fewest parties	0.196	(0.130)		
6. MWC: Minimum-winning coalition	−0.170	(0.112)		
7. Seatshare	0.557	(0.053)	0.473	(0.044)
8. Seatshare squared (/100)	−0.463	(0.046)	−0.414	(0.040)
9. Number of parties	−0.377	(0.100)	−0.424	(0.048)
10. Ideological range	−0.340	(0.039)	−0.555	(0.019)
11. Opposition ideological range	−0.218	(0.031)	−0.303	(0.020)
12. Opp. ideological range × Minority cabinet	0.222	(0.046)		
13. Ideological distance to median	0.094	(0.078)		
14. Median party	0.400	(0.099)	0.460	(0.080)
15. Largest party	0.655	(0.179)		
16. Single-party majority	1.122	(0.360)		
17. Anti-system party (SD)	−5.027	(0.390)		
18. Local party	−0.609	(0.115)		
19. Incumbent government	1.510	(0.117)	2.223	(0.096)
20. Party of incumbent Mayor	−0.489	(0.125)		
21. One or some incumbent parties	−0.510	(0.161)		
22. Electoral performance	0.043	(0.016)		
23. Incumbent gov't × Electoral performance	0.063	(0.029)		
24. Familiarity	1.927	(0.372)		
25. Commitment potential	0.464	(0.173)		
26. Right-wing bloc	2.474	(0.183)	2.373	(0.115)
27. Right-wing bloc (plus)	2.501	(0.213)		
28. Right-wing bloc (minus)	0.277	(0.161)		
29. Left-wing bloc	0.522	(0.181)		
30. Left-wing bloc (plus)	−0.122	(0.252)		
31. Left-wing bloc (minus)	−0.031	(0.126)		
32. Bloc-transcending coalition	−1.255	(0.128)		
Observations	409,113		409,113	
Government formation opportunities	1,719		1,719	
Pseudo R^2	0.584		0.489	

Standard errors in parentheses (clustered by government formation opportunity).

A1.2 Electoral competitiveness measures

This section gives an overview of the 18 measures of electoral competitiveness to which our Swedish measures are compared in Section 4.2 of the paper. For exact definitions and scales, we refer the reader to the respective referred study.

To begin, we include four measures of electoral competition that are derived from the vote share (or the seat share) of the partisan majority, the incumbent, the ruling coalition, and the largest party, respectively, developed by Clingermayer and Wood (1995), Canes-Wrone and Park (2012), Boyne (1998) and Vanhanen (2000).[5]

In addition, we include three measures that capture the closeness between the two major parties in the parliament: The *two-party margin*, the *two-party ratio*, and the *raw vote margin* (in percentage points) separating the two major parties. These measures are described in detail by Fauvelle-Aymar and François (2006). A related measure by Aidt et al. (2011) of the win-margin of the *mayor's party* over the largest opposition party is also included. Applicable specifically to the Swedish case, Högström (2017) proposes a measure of electoral closeness defined as the difference between the two major *blocs* in Swedish politics (see Footnote 3). Here, we include both versions of Högström's closeness measure, one based on *ex-ante* vote shares and one based on *ex-post* vote shares.

Next, we include three more advanced indexes based on vote shares, designed more specifically for multi-party systems: Kirchgässner and co-authors' (1992) *entropy index*, which is a measure of the instability in the election; Endersby and co-authors' (2002) *competition index*, and Capron and Kruseman's (1988, p. 33) *fractionalization index*, which "measures the probability that any two voters randomly chosen from the electorate have voted for different parties". For definitions, see Fauvelle-Aymar and François (2006).

We also include a *volatility* measure proposed by Boyne (1998), applied, in our case, on the changes in vote shares of the Mayor's party over three most recent elections. In a similar vein, following Hübscher and Sattler (2017), we also include a measure of the *replacement risk* of the largest incumbent party, which is a function of its closeness to the second party in the parliament as well as changes in vote shares among all parties over (up to) the five past elections. In addition, we include a measure of the plurality party's likelihood of losing plurality status, created following the approach outlined by Kayser and Lindstädt (2015).

Finally, we include two measures of *electoral pressure* and *political protection*, developed by Immergut and Abou-Chadi (2014). The two measures are derived from a factor analysis performed on six variables related to electoral competition: voter volatility, the disproportionality of the electoral system (LSQ index), the effective number of parties, the fraction of electoral winners in government, the size of government majority, and the size of government majority relative to the number of governing parties.[6]

[5] Although the measure developed by Canes-Wrone and Park (2012) originally used vote shares from polling data, we have resorted to using vote shares from the previous election.

[6] Immergut and Abou-Chadi (2014) apply a varimax rotation to the results of the factor analysis, and then create the two variables. We skip rotation, as the factors generated on our sample appeared unsuitable for rotation and as the unrotated versions perform better in the comparisons.

In Figure 4 in the paper, the 18 measures described above are compared to the measures constructed using our approach, with regards to their capability to predict re-election of the incumbent. As an alternative way to explore this issue, we follow Kayser and Lindstädt (2015, appendix) by considering visually how the distributions of the tested measures vary between two sub-samples of the data, one with municipality-elections after which re-election took place and one with municipality-elections after which it did not. If a measure is to have predictive capability, we should expect its scores to be unevenly distributed across the two sub-samples. Our main outcome indicator – the weighted re-election of the incumbent – is not dichotomous, thus we resort here to the indicator on re-election (or ouster) of the largest incumbent party.

The box plots reported in Figure A1 show how the scores of the 26 compared measures are distributed across the two sub-samples. Plotted in the top row, the four main versions of our re-election probability turn out to be the ones for which the distribution of the predictor varies most clearly across the two realized outcomes. Consider, for instance, that for elections after which the largest incumbent party remained in office, the re-election probability as measured in May of the year of the election was around or above 66 percent in 75 percent of the cases, whereas little more than 25 percent of the probabilities were that high for elections after which the largest incumbent party was ousted. Also, consider that the difference in the median predicted probability across the two outcome is more than 20 percent of the full range of values. An inspection of the plots for the 18 existing measures reveals that none of them comes close to being that much differently distributed across the two outcomes. Indeed, for several measures the two distributions are largely indistinguishable.

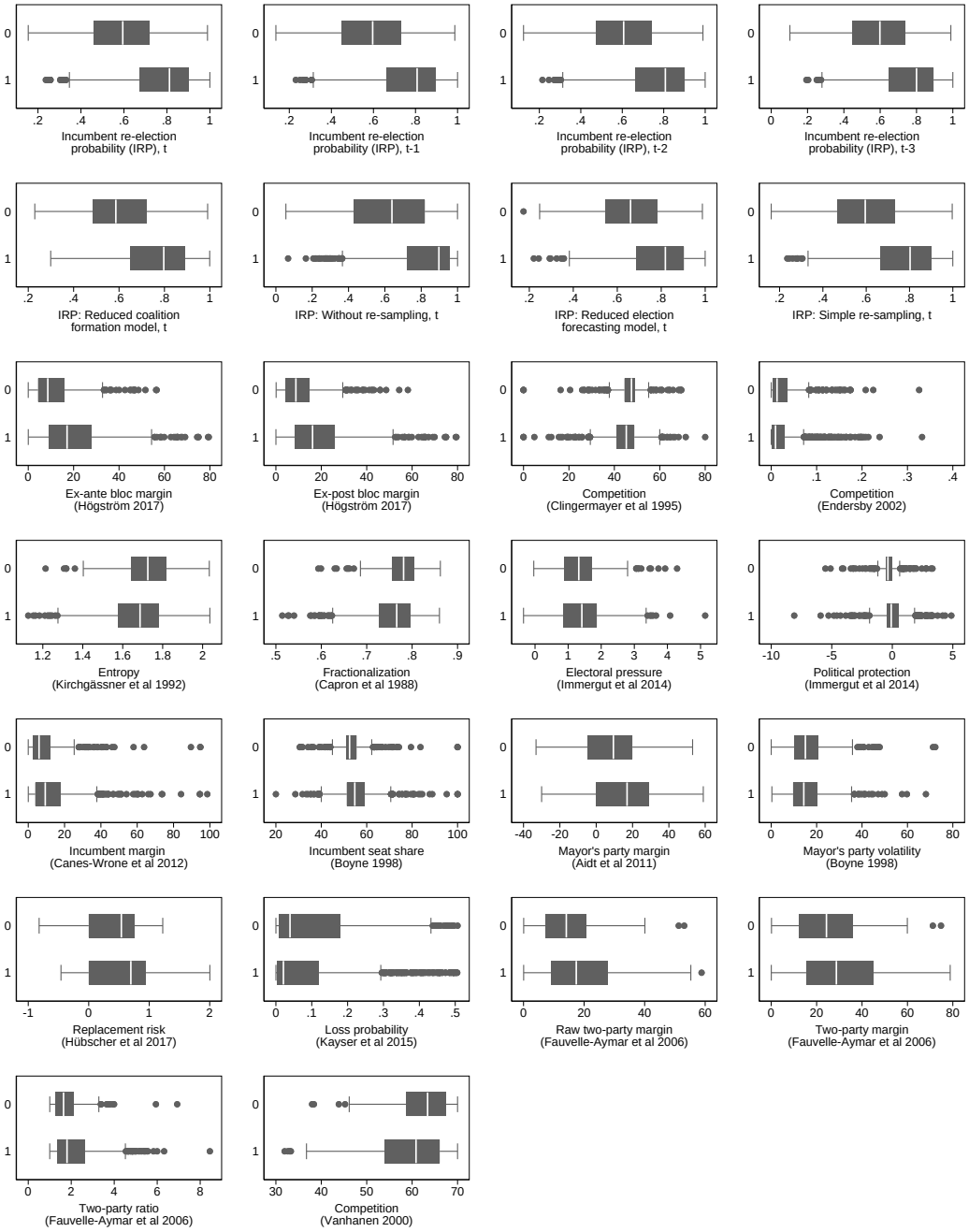

Figure A1: Box plots for re-election of the largest incumbent party

A1.3 Evaluation of simplified measures

Admittedly, to code each and every step of the approach presented in this paper is a time-consuming task. To make our approach more appealing to researchers whose main interest is not in the estimation of re-election prospects, we will now present and evaluate a number of possible simplifications to our approach.

We have investigated three different simplifications of how we generate the simulated election outcomes in Step I. First, we have used a reduced model for the election forecast, with previous election results being the only predictor. Compared to the large set of variables in Equation 2 in the paper, this is a huge simplification. Second, we have simplified the simulation procedure, by simply drawing party-specific residuals without the blocs and clusters described in Section 3.2 in the paper. This simplification means that we do not model the inter-party correlation or the relation between party size and residual variance in our re-sampling. Third, we have skipped the re-sampling altogether, effectively assuming that the election forecast will be realized, which leads us to underestimate the pre-electoral uncertainty.

As was mentioned in Section A1.1 above, we have tried two different sets of independent variables in the coalition formation model used in Step II. Our main model includes 32 independent variables, of which some are computationally demanding or require data that may be difficult to get hold of. We have therefore run a coalition formation model that only includes ten variables, chosen because they are central in the literature, have high explanatory power, and are easy to calculate.

The rest of this section evaluates these four simplifications. As in the paper, we will begin by assessing the accuracy of the uncertainty estimations, and then evaluate the predictive capability. Because all our measures are estimated probabilities, we are also able to evaluate the simplifications using Brier scores, which are reported separately in Section A3 of this appendix.

A1.3.1 Accuracy of uncertainty estimation

In the paper, we argue that it is important to model the electoral uncertainty, or else the probability estimates will be biased away from full uncertainty ($p = 0.5$). In Figure A2 we test this claim by showing binned scatter plots, with the share of elected observations on the y-axis and the estimated office probability on the x-axis, for the four different simplified versions of our measure described above. For the probabilities to be correctly estimated, they should line up along the 45° line, where the share of elected observations equal the estimated election probability.

First, the measure in the upper-left plot (I) uses the reduced election forecasting model in Step I, relying only on previous election results. This reduces the accuracy of our uncertainty estimations, as evidenced by the larger spread of the bins away from the 45° line compared to the main measure reported in the center plot of Figure 3 in the paper. Moving on to the upper-right plot (II), we find that the simplification of the procedure for re-sampling the residuals in Step I results in a minor deterioration of accuracy.

In the lower-left plot (III), we report the measures for which there is no

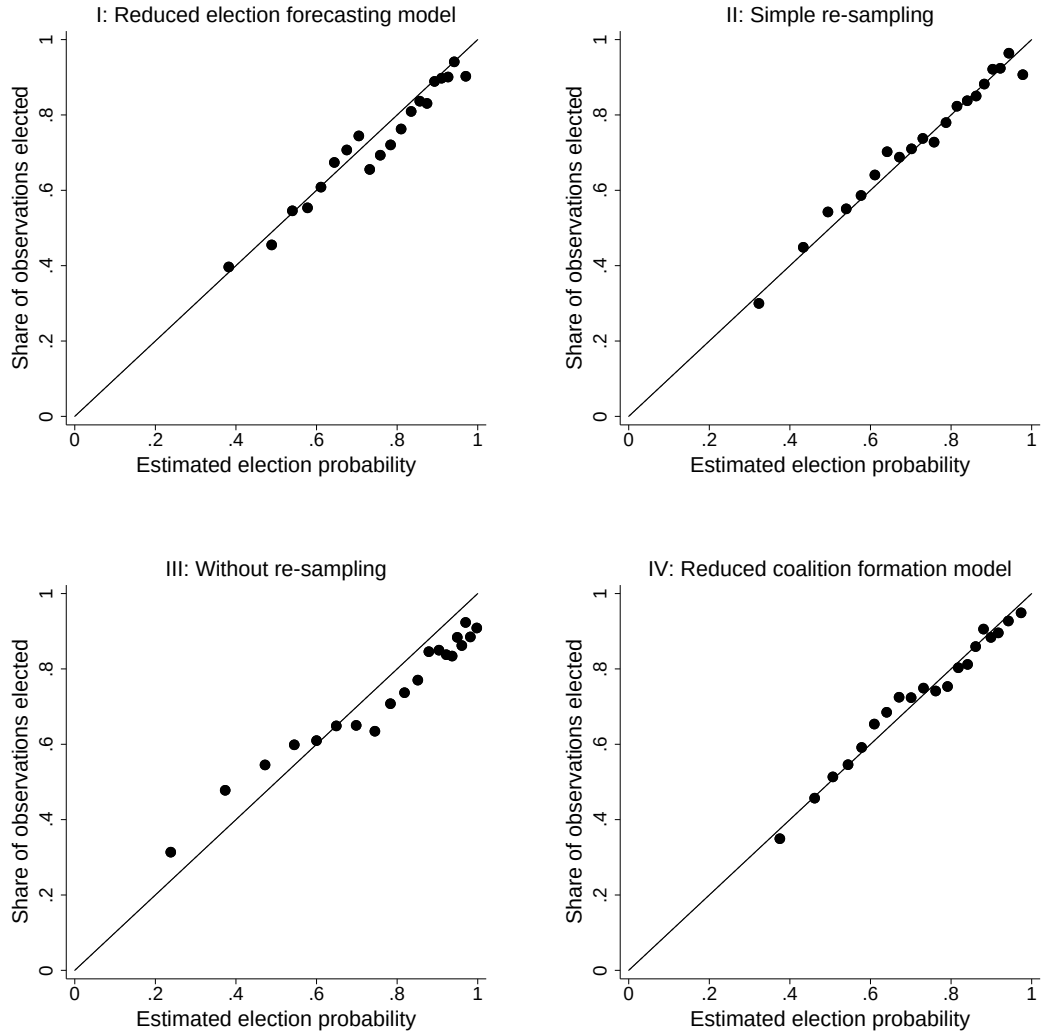

Figure A2: Comparison between predicted probabilities and average outcomes

re-sampling in Step I. Here, the plot reveals a clear pattern of underestimated uncertainty, as the share of elected observations is clearly higher than estimated for low probabilities, and vice versa for high probabilities. This result confirms that the simulation is necessary in order to produce correctly calibrated probability estimates. This finding is intuitive: When we use the predicted election outcomes as if forecasts were perfect, we fail to model one of two sources of uncertainty: pre-electoral uncertainty. Thus, predicted election losses, or wins, will be taken for granted, when in fact they should not.

Finally, the lower-right plot (IV) reports the measure that relies on the reduced coalition formation model in Step II. As illustrated by the close resemblance between this plot and the center plot of Figure 3 in the paper, the estimated probabilities are well calibrated also when using the reduced government formation model. This should not come as a surprise. While the main measure will be

estimated with greater precision, thanks to a larger number of relevant variables in the coalition formation model, there is no reason why fewer variables would provide more biased estimates.

A1.3.2 Capability to predict re-election into office

The second part of the evaluation investigates how the four simplifications of our measure affect its capability to predict re-election of the incumbent. Here we repeat the approach used in Section 4.2 in the paper by running, for each of the four simplified measures as well as the main measure, one OLS regression of the seat share weighted re-election of the incumbent parties (see Footnote 13 in the paper), one of the re-election of the largest incumbent party into the governing coalition and one of the re-election of the Mayor's party into the Mayor's office, with the respective measure in question as a single predictor. To conserve space, we only construct our simplified measures for one point in time, namely the month of May of the election year (t).

Figure A3 reports the results. For each of the five variants, three bars with Adjusted R^2 scores are presented. Given that the bars for the three outcome measures follow similar patterns, we may concentrate our discussion on the seat share weighted re-election of the incumbent. Here, the incumbent re-election probability (IRP) measure produced using our preferred election model and the full coalition formation model account for 23.1 percent of the variation in re-election. As expected, when applying the reduced coalition formation model in Step II the Adjusted R^2 score of the measure drops considerably; by around 3.8 percentage points or 16 percent in the case at hand. Yet, this loss in predictive power is after all not too dramatic, and in some applications it may be a price worth paying.

The figure furthermore shows the added value of carefully modeling the pre-electoral uncertainty, as for each of the simplified variations of Step I we notice a considerable reduction in predictive capacity. For the worst-performing variation of the IRP measure, where the election model is based solely on previous election results, around one third of the predictive capacity is lost and the Adjusted R^2 score drops to slightly more than 0.15.

In sum, the implication of the results reported in Figure A2 and Figure A3 is two-fold. On the one hand, it demonstrates that efforts to improve the quality of the models – in both steps of our approach – have the potential to pay off in terms of the accuracy of the uncertainty estimations and/or the predictive capability of the produced measures. On the other hand, it shows that even a simple election forecasting model in Step I, based solely on the previous election results, or a fairly limited government formation model in Step II, produce measures of the re-election probability of the incumbent that substantially outperform all previous measures of electoral competitiveness in terms of their capacity to predict re-election, while also conforming to Kayser and Lindstädt's (2015) six criteria for a useful measure of electoral competitiveness. This, in our view, demonstrates the added value of considering both pre-electoral and post-electoral uncertainty when devising measures of electoral competitiveness for parties and governments.

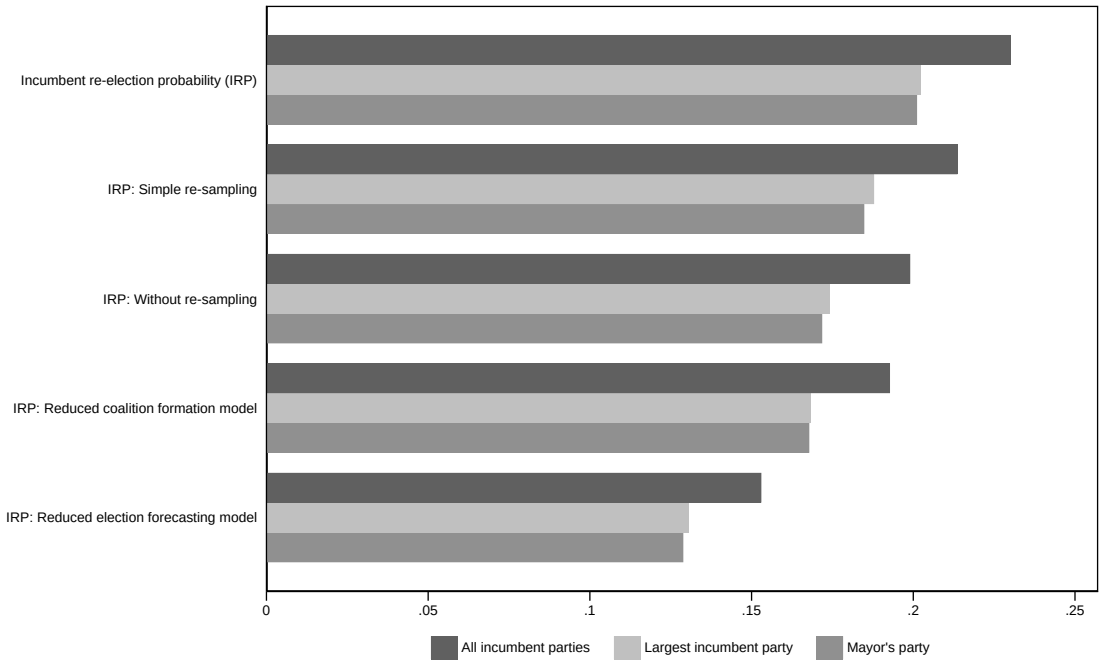

Figure A3: Capability to predict re-election into office at t (Adjusted R^2)

A2 Cross-national application

A2.1 Data considerations

For this application, we use the development version of the ParlGov database (Döring and Manow 2019). The database contains many different files, but we primarily use the files view_cabinet, view_election and election_result, as well as – but mostly for convenience – the files election and party. We also use the file party_changes to identify party splits and mergers.

Our point of departure is the universe of elections in ParlGov. We exclude elections that meet one of the following three criteria: i) it happened before 1945, or it was one of the country's first two elections after the Second World War, in which case we would expect very high electoral volatility, ii) the country has a presidential or hybrid system, in which case the result of the parliamentary election does not directly affect the cabinet composition, or iii) a party that received at least 25 percent of the votes in the previous election has experienced a split or merger during the election period, so as to avoid having to impute vote shares in the way described below for one of the most important parties. This leaves us with 420 national elections from 34 countries.[7]

It is not uncommon with multiple cabinet changes during an election period. In

[7]The countries represented in our sample are Australia, Austria, Belgium, Bulgaria, Canada, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, New Zealand, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and the United Kingdom. The observed election years range between 1950 and 2019.

those cases, the incumbent government (whose re-election probability we estimate) is defined as the last cabinet before an election observed in the ParlGov database, and which parties that are elected into office is based on the parties that took part in the first cabinet succeeding the election, excluding caretaking cabinets.

We define the parties in the same way as they appear in ParlGov's cabinet file – which is sometimes different from how they appear in the election file. For example, in Germany, CDU and CSU are presented as one party in the cabinet file, but two parties in the election files. We therefore treat them as one party, with a vote share that corresponds to the total vote share of the two parties. Conversely, in cases where multiple parties form a pre-electoral alliance, they often have a common vote share registered in the election file, although they appear as separate parties in the cabinet file. In those cases, we do our best to divide this vote share between the parties, based on how many parliamentary seats they received or how many votes they received in the previous election. Whenever a party has split during an election period, we define both new parties' previous vote share as half of their joint previous vote share. In instances of party mergers, we define the new party's previous vote share as the sum of the vote shares of the merging parties. Since this semi-automatic procedure for linking parties to alliances, mergers and splits is not flawless, and since the cabinet file sometimes excludes parties that have won a substantial vote share, the sum of the vote shares for parties included in our analysis is sometimes higher than or substantially lower than 100 percent. It is possible that some measures of electoral competitiveness would perform better if vote shares always summed to unity.

To model the emergence of new parties, we include one dummy party in each election with a previous vote share that equals zero and a new vote share that equals the total vote share for all parties that, in the previous election, received zero votes (or for which there is no party-specific information in the database).

Our only addition to the ParlGov data is vote intention poll data included in the database compiled by Jennings and Wlezien (2018). For the sake of this simple demonstration we create only one variable, by computing – for each party and election observed in both databases – the average result in all polls collected during the 365 days preceding the election.[8] Polling data are available for almost half of the elections in our sample, and the matched parties correspond to more than 80 percent of the total vote share in these elections. For parties and elections where no poll data are available, we simply replace the missing values by the previous vote share reported in the ParlGov database, to achieve a measure with no missingness.

A2.2 Modeling considerations

Whereas the cross-national application follows the same four-step procedure as the Swedish one, a number of operations have been carried out differently to accommodate the differences in institutional context and data availability.

[8] We have manually linked party names across the two databases. Polls are considered missing in instances where parties or pre-electoral alliances do not match.

A2.2.1 Step I: Forecasting election results

There are three important parts to the first step: the election forecast that predicts parties' next vote shares, the re-sampling of vote share residuals, and the translation of predicted vote shares into predicted seat shares.

In the cross-national application, the election forecast is a regression model where we regress each party's vote share on a cubic spline of the previous vote share, an indicator for being a new (or previously unsuccessful) party,[9] an indicator for incumbent government membership, an indicator for the party's family identity, and an interaction between the incumbency indicator and the previous vote share. To account for the larger electoral volatility and the less institutionalized party system in post-communist countries (Döring and Hellström 2013; Savage 2016), we also include two interactions between the indicator for new democracies[10] and the indicators for previous vote share and for being a new party. As in the Swedish application, we include election fixed effects. An advantage of this specification is that it relies on a single data source and that it does not require any polling data. However, in an attempt to improve forecast accuracy, we also run a second version of this model where, whenever possible, we replace each party's previous election result with the party's average result from all vote intention polls collected during the 365 days preceding the election.

As in the Swedish application, we model the uncertainty surrounding the electoral forecast by drawing new party-election specific residuals in each simulation. However, because the same set of parties never appear in more than a handful of elections, we cannot draw blocs of residuals as we did for Sweden. Instead we simply draw individual residuals for each party. The residuals are re-sampled within five clusters. The first cluster consists of new parties in old democracies, the second includes new parties in new democracies, and the three remaining clusters are determined by the expected absolute size of the residual.[11]

The third part of Step I, in which we allocate parliamentary seats based on how many votes each party received, was not given much attention in the Swedish application, where we were able to apply the same modified Sainte-Laguë method to all elections. However, in our cross-national application we are dealing with several different electoral systems. Even if we knew the details about how the seat distribution is decided, it would be impossible to replicate the actual seat distribution in some countries, because we lack information about how the votes were distributed at the constituency level. For instance, in countries with first-past-the-post systems, the largest party would receive all the votes if we assumed that each party received the same share of the votes in every constituency. What we do instead is to give each party the same seat share as in the previous election, plus their change in vote share multiplied by an estimated seats-to-votes elasticity.[12]

[9] This is defined as not having a previous vote share registered in the ParlGov database.

[10] We define new democracies as Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia, and Slovenia (Döring and Hellström 2013; Savage 2016). The remaining 23 countries are classified as old democracies.

[11] Expectations are derived from a regression model similar to that used in the election forecast.

[12] To estimate this elasticity we have, for one country at a time, regressed historical seat shares on a cubic spline of the corresponding vote shares, and calculated the marginal effect of a changed

A2.2.2 Step II: Modeling government formation

In this application, the conditional logit model includes 14 of the 32 potential government variables used in the Swedish application (and described in Section A1.1): *minority cabinet*, *minimal-winning coalition*, *seat share*, *seat share squared*, *number of parties*, *ideological range*, *opposition ideological range*, an *interaction* between opposition ideological range and minority cabinet, *median party*, *largest party*, *single-party majority*, *incumbent government*, *party of the incumbent prime minister*, and *one or some incumbent parties*.

In addition, to compensate for our lack of institutional knowledge regarding pre-electoral alliances, the existence of cordon sanitaires around particular parties, etc., we add two new coalition variables, which measure the *average share of their time in parliament that parties have spent in government*, and *whether or not the combination of parties have governed together before*.

Furthermore, to account for previously documented differences between old and new democracies – for instance that incumbent parties and established parties do not have the same bargaining advantage in new democracies (Döring and Hellström 2013; Savage 2016) – we include two interaction terms between an indicator for new democracies and the two aforementioned new variables. In total, we thus end up with 18 potential government variables in the model.

Because we have a much smaller number of elections in the cross-national application, the coefficients in the coalition formation model are estimated with much lower precision. Unless we account for the uncertainty about the correct parameters, our estimations will overestimate high re-election probabilities, and underestimate the probabilities in the lower part of the distribution.[13] To deal with this problem, we estimate a separate coalition formation model for each simulation, where the estimation is run on a bootstrapped sample of government formation opportunities.[14] For each simulation, it is then a unique vector of coefficients ($\boldsymbol{\beta}$) that, in Step III, is being applied to the simulated vote shares in order to predict each potential government's probability of entering office.

vote share at all relevant points of the distribution.

[13] In theory, this could be an issue also in the Swedish application, but as we had so many more elections in that case it turned out not to be a problem in practice (see Figure 3 in the paper).

[14] If we would store all the regression coefficients obtained in this procedure, we would see, for each variable, a distribution of coefficients, with a mean close to the coefficient estimated on the main sample and a standard deviation close to the estimated standard error of the coefficient.

A3 Brier scores and their decomposition

In this section, we provide an additional evaluation of the Swedish and the cross-national applications of our approach, by means of computing and comparing Brier scores for the various versions of our measures.

It is difficult to evaluate the accuracy of probability forecasts, because the true probability of a positive outcome can never be observed. Instead, forecasters or forecasting models are typically evaluated using a 'scoring rule', where repeated probability measures are compared with the observed outcomes. The most frequently used scoring rule is the Brier score (Brier 1950), which measures the average squared distance between the estimated probability of a positive outcome (in our application the office probability $p(O)_i$) and the actual outcome (a binary indicator measured as 0 or 1):

$$BS = \frac{1}{N}\sum_{i=1}^{N}(p(O)_i - O_i)^2 \tag{1}$$

The Brier score takes on a value between 0 (all negative outcomes are assigned a probability of 0 and all positive outcomes a probability of 1) and 1 (all positive outcomes are assigned a probability of 0 and all negative outcomes a probability of 1). The lower the Brier score is, the better the predictions are considered to be.

A useful feature of Brier scores is that they can be decomposed into three additive terms (Murphy 1973), which captures different attributes of the prediction: reliability (REL), resolution (RES) and uncertainty (UNC). These terms relate to the Brier score as follows, which means that to get a low Brier score, forecasts should have low scores on reliability and uncertainty, but high scores on resolution:

$$BS = REL - RES + UNC \tag{2}$$

Reliability, commonly also referred to as 'calibration', concerns whether the estimated probabilities are actual probabilities, such that for any set of predictions, the average probability estimate equals the expected share of positive outcomes. In other words, reliability is the same thing as we evaluate in Section 4.1 in the paper and Section A1.3.1 of this appendix. To compute the reliability component used in the decomposition above, one first sorts the predictions into K number of groups, based on the estimated probability, and then calculates the average squared deviation between the share of positive outcomes and the estimated probability in each group:

$$REL = \sum_{k=1}^{K}\frac{n_k}{N}(p(O)_k - \bar{O}_k)^2 \tag{3}$$

The resolution component measures how much variation there is in the estimated probabilities. In other words, this measure is similar to the regression sum of squares used when calculating the R^2 for a regression model. The resolution is calculated as the average difference between the conditional event frequencies, given the groupings, and the global share of outcomes:

$$RES = \sum_{k=1}^{K} \frac{n_k}{N}(\bar{O}_k - \bar{O})^2 \tag{4}$$

Lastly, the uncertainty component measures the global uncertainty of the outcome. It equals zero if an outcome always or never occurs, and can reach a maximum of 0.25 for outcomes that occur half the time.

$$UNC = \bar{O}(1 - \bar{O}) \tag{5}$$

For the Swedish application, we have calculated Brier scores for our preferred approach as described in the paper, but also for the four other variants where we i) use a reduced set of variables in the election forecast model, ii) skip the blocs and clusters in the re-sampling procedure, iii) skip the simulation step altogether, and iv) use a reduced set of variables in the coalition formation model.For the cross-national application, we have calculated Brier scores for the two different election forecasting models: one with and one without polling data. The scores have been calculated for two actors where the outcomes are binary: election of any specific party and the re-election of the largest party in the incumbent coalition.

The results are presented in Table A2. Let us begin with the results for the Swedish application. As indicated by the lower Brier score, our main model performs better than the four alternatives. The worst performer is the variant without re-sampling, which receives the largest Brier score for both outcomes. Skipping the simulation step leads to a worse reliability (higher REL score), in line with expectations and the visual evidence in Figure A2. However, doing so also turns out to reduce the resolution of the prediction model. A much better result is achieved if the re-sampling procedure is instead replaced with the simplified version without blocks and clusters. In fact, this variant performs almost as good as our main model. However, it should be kept in mind that the two re-sampling methods yielded larger differences when the outcome was the weighted share of re-elected government parties (Figure A3).

The two variants with fewer variables in the regression models perform better than the variant without re-sampling, but not as well as the other two. When we analyze all parties, it is primarily the reduction in the coalition formation model which affects the Brier score, while the reduced election forecasting model performs fairly well. On the other hand, the tables are turned when we restrict the sample to the largest incumbent party. This difference implies that the pre-election uncertainty is relatively more important for large incumbent parties, while it becomes more important to accurately model the coalition prospects when we are interested in smaller parties who are more dependent on ruling with others.

Simply comparing the two alternative outcomes, we see large differences with regards to uncertainty and resolution, that jointly result in a much smaller difference in the Brier score. The difference in uncertainty reflects the fact that by guessing the most common outcome, we would be right in 75 percent of the cases with respect to the largest incumbent party (they are elected after 75 percent of the elections), but in only 65 percent of the cases with respect to all parties (they have a baseline office probability of 35 percent). However, because our prediction model

Table A2: Brier scores and their decomposition

	BS	REL	RES	UNC
Sweden: All parties				
Main model	0.144	0.000	0.084	0.229
Reduced election forecasting model	0.150	0.000	0.078	0.229
Simple re-sampling	0.146	0.000	0.083	0.229
Without re-sampling	0.156	0.004	0.076	0.229
Reduced coalition formation model	0.156	0.000	0.072	0.229
Sweden: Largest incumbent party				
Main model	0.158	0.001	0.033	0.192
Reduced election forecasting model	0.168	0.003	0.026	0.192
Simple re-sampling	0.159	0.001	0.032	0.192
Without re-sampling	0.170	0.008	0.029	0.192
Reduced coalition formation model	0.163	0.001	0.029	0.192
Cross-national: All parties				
Without polls	0.152	0.002	0.065	0.215
With polls	0.149	0.003	0.068	0.215
Cross-national: Largest incumbent party				
Without polls	0.209	0.009	0.024	0.223
With polls	0.200	0.010	0.033	0.223

Columns represent Brier scores (BS), reliability (REL), resolution (RES) and uncertainty (UNC). Lower scores on BS and REL, and higher scores on RES, indicate a better forecasting model. UNC is unaffected by the predictions. All scores are calculated based on groups of 10 observations.

does a very good job of resolving the differences between large incumbent parties and other parties, the predictions for the latter end up slightly more accurate.

The main result for the more challenging cross-national setting is that our approach can be applied successfully also there, although it does not perform as well as it did in the Swedish application. Looking at the RES score, we see that our measure can resolve a similar amount of variation in the two applications, even though the resolution is slightly lower for the cross-national application. However, as we could see from the scatter plots in the paper, the probabilities generated in the cross-national application are not as well-calibrated as the ones for Sweden, with REL scores similar to the Swedish model without re-sampling. This is particularly true for the re-election probability of the largest incumbent party. The model which includes polling data performs somewhat better than the model without. This difference would be much more pronounced if we would limit this evaluation to the elections for which polling data were available.

References

Aidt, Toke S, Francisco José Veiga, and Linda Gonçalves Veiga (2011). 'Election results and opportunistic policies: A new test of the rational political business cycle model'. *Public Choice* 148.1-2, 21–44.

Axelrod, Robert (1970). *Conflict of interest*. Chicago: Markham.

Bäck, Hanna (2003). 'Explaining and predicting coalition outcomes: Conclusions from studying data on local coalitions'. *European Journal of Political Research* 42.4, 441–472.

Bäck, Hanna and Johannes Lindvall (2015). 'Commitment problems in coalitions: A new look at the fiscal policies of multiparty governments'. *Political Science Research and Methods* 3.1, 53–72.

Boyne, George A (1998). 'Party competition and local spending decisions'. *British Journal of Political Science* 28.1, 185–222.

Brier, Glenn W (1950). 'Verification of forecasts expressed in terms of probability'. *Monthly Weather Review* 78.1, 1–3.

Canes-Wrone, Brandice and Jee-Kwang Park (2012). 'Electoral business cycles in OECD countries'. *American Political Science Review* 106.1, 103–122.

Capron, Henri and Jean-Louis Kruseman (1988). 'Is political rivalry an incentive to vote?' *Public Choice* 56.1, 31–43.

Clingermayer, James C and B Dan Wood (1995). 'Disentangling patterns of state debt financing'. *American Political Science Review* 89.1, 108–120.

De Swaan, Abram (1973). *Coalition theories and government formation*. Amsterdam: Elsevier.

Döring, Holger and Johan Hellström (2013). 'Who gets into government? Coalition formation in European democracies'. *West European Politics* 36.4, 683–703.

Döring, Holger and Philip Manow (2019). *Parliaments and governments database (ParlGov): Information on parties, elections and cabinets in modern democracies*. Development version.

Endersby, James W, Steven E Galatas, and Chapman B Rackaway (2002). 'Closeness counts in Canada: Voter participation in the 1993 and 1997 federal elections'. *The Journal of Politics* 64.2, 610–631.

Fauvelle-Aymar, Christine and Abel François (2006). 'The impact of closeness on turnout: An empirical relation based on a study of a two-round ballot'. *Public Choice* 127.3-4, 461–483.

Folke, Olle (2014). 'Shades of brown and green: party effects in proportional election systems'. *Journal of the European Economic Association* 12.5, 1361–1395.

Gilljam, Mikael, David Karlsson, and Anders Sundell (2015). *Kommun- och landstingsfullmäktigeundersökningen (KOLFU) 2008. Version 1.0*. Svensk nationell datatjänst.

Glasgow, Garrett and Sona N Golder (2015). 'A new approach to the study of parties entering government'. *British Journal of Political Science* 45.4, 739–754.

Högström, John (2017). 'Does closeness matter for voter turnout in proportional systems? An examination at the local level in Sweden'. *European Political Science* 7.4, 571–586.

Hübscher, Evelyne and Thomas Sattler (2017). 'Fiscal consolidation under electoral risk'. *European Journal of Political Research* 56.1, 151–168.

Immergut, Ellen M and Tarik Abou-Chadi (2014). 'How electoral vulnerability affects pension politics: Introducing a concept, measure and empirical application'. *European Journal of Political Research* 53.2, 269–287.

Jennings, Will and Christopher Wlezien (2018). 'Election polling errors across time and space'. *Nature Human Behaviour* 2.4, 276.

Karlsson, David and Mikael Gilljam (2018). *Kommun- och landstingsfullmäktigeundersökningen (KOLFU) 2012. Version 1.0.* Database.

Kayser, Mark Andreas and René Lindstädt (2015). 'A cross-national measure of electoral competitiveness'. *Political Analysis* 23.2, 242–253.

Kirchgässner, Gebhard and Jörg Schimmelpfennig (1992). 'Closeness counts if it matters for electoral victory: Some empirical results for the United Kingdom and the Federal Republic of Germany'. *Public Choice* 73.3, 283–299.

Laver, Michael and Norman Schofield (1990). *Multiparty Government: The Politics of Coalition in Europe*. Oxford: Oxford University Press.

Martin, Lanny W and Randolph T Stevenson (2001). 'Government formation in parliamentary democracies'. *American Journal of Political Science*, 33–50.

Martin, Lanny W and Randolph T Stevenson (2010). 'The conditional impact of incumbency on government formation'. *American Political Science Review* 104.3, 503–518.

Morgenstern, Oskar and John Von Neumann (1953). *Theory of games and economic behavior*. Princeton University Press.

Murphy, Allan H (1973). 'A new vector partition of the probability score'. *Journal of applied Meteorology* 12.4, 595–600.

Polk, Jonathan (2017). 'Explaining the salience of anti-elitism and reducing political corruption for political parties in Europe with the 2014 Chapel Hill Expert Survey data'. *Research & Politics* January–March, 1–9.

Savage, Lee (2016). 'Party system institutionalization and government formation in new democracies'. *World Politics* 68.3, 499–537.

Skjæveland, Asbjørn, Søren Serritzlew, and Jens Blom-hansen (2007). 'Theories of coalition formation: An empirical test using data from Danish local government'. *European Journal of Political Research* 46.5, 721–745.

SKL (2018). *Maktfördelning i kommuner, landsting och regioner 1994–2014*. Database. Swedish Association of Local Authorities and Regions (SKL).

Vanhanen, Tatu (2000). 'A new dataset for measuring democracy, 1810–1998'. *Journal of Peace Research* 37.2, 251–265.