A multilevel model for coalition governments: Uncovering dependencies within and across governments due to parties

Benjamin Rosche

benrosche.com – fbr33@cornell.edu

Abstract

Coalition research increasingly focuses on party-level explanations of coalition outcomes. This work, however, ignores the complex multilevel structure between parties and governments: Many parties become nested in multiple governments over time, and governments are often nested in coalitions of multiple parties. In this paper, I show that this crisscrossing structure (i) induces dependencies among observations that inflate both Type-I and Type-II error rates if ignored, and (ii) reflects an inherent aggregation problem in party-level explanations. I then advance a novel multilevel model to account for the dependencies among observations and explicitly model the aggregation of party effects to the government level. I demonstrate the model's ability to better integrate theoretical expectations into empirical models with an application to coalition government survival as predicted by parties' financial dependencies. The results show that the more parties' financial resources comprise of member contributions, the higher the termination hazard of governments including those parties.

Acknowledgements

I am grateful to Jeroen Weesie who was an early collaborator on this project and to Alison Schultz who continually supported me on this journey. I would also like to thank Alec McGail, David Kretschmer, Filiz Garip, Fiona Coleman, George Leckie, Jeff Gill, Lanny Martin, Markus Wagner, Thomas Meyer, Thomas Saalfeld, and Zac Greene for helpful comments. An earlier version of this paper has been presented at the 2019 APSA Annual Meeting in Washington DC.

1 INTRODUCTION

The majority of advanced industrial democracies are parliamentary systems in which coalition governments are the norm. Accordingly, the study of the formation, governance, and termination of coalition governments is a central focus of comparative scholarship. A vast body of literature has established that such outcomes depend on (i) critical events, such as economic recessions, (ii) systemic properties at the country level, such as electoral institutions, and (iii) structural attributes at the government and parliament level, such as ideological heterogeneity or minimal winningness (Strøm, Müller, and Bergman 2008).

In recent years, coalition research has directed attention to party-level explanations of coalition outcomes (Giannetti and Benoit 2009). Efforts to redress the lack of large-scale comparative data on party organization¹ have led to a rise of empirical contributions (Bäck 2008; Bäck, Debus, and Dumont 2011; Ceron 2016; Druckman 1996; Greene 2017; Martínez-Cantó and Bergmann 2019; Martínez-Gallardo 2010; Saalfeld 2009). These studies, however, ignore the complex nonstandard multilevel structure between parties and governments. As Figure 1 shows, in addition to their nesting in countries, parties often become nested in multiple governments over time, and governments are often nested in multiple parties (i.e., are coalition governments). Acknowledging this crisscrossing relationship between parties and governments is key to studying party effects on government outcomes for two reasons.

Figure 1: The multilevel structure of coalition governments in Israel

¹ New data collection efforts include the Political Party Database Project (Poguntke, Scarrow, and Webb 2017), ParlGov (Döring and Manow 2016), the Integrated Party Organization dataset (Giger and Schumacher 2015), and the Party Facts Project (Bederke, Döring, and Regel 2020).

First, from a statistical perspective, the multilevel structure implies dependencies among observations. While dependencies arising from the clustering in countries have been considered by previous research, dependencies that arise from the clustering in parties are ignored in all existing empirical analyses. These dependencies manifest in substantial correlations among outcomes of governments containing the same parties. If ignored, they bias both regression coefficients and standard errors and, thereby, affect the conclusions we draw about party effects.

Second, from a theoretical perspective, the multilevel structure reflects interdependencies of political processes in the parliamentary arena. The outcomes of coalition governments depend on all constituent parties and on the interdependencies among them. Consequently, the interplay of coalition parties in their joint effect on the government must be considered. Analyzing parties' joint effect constitutes cross-level inference as the effects of parties are aggregated to the government level. Multilevel analysis provides a suitable framework to model and test theories on how this aggregation process depends on the features and the configuration of coalition parties. While emphasized in theoretical work (Cross and Katz 2013; Lupia and Strøm 1995; Müller and Strøm 1999; Strøm 1990), the interplay of intra-party and inter-party politics is not properly modeled in empirical research on coalition government outcomes.

In this paper, I propose a novel statistical technique to represent the complex multilevel structure of coalition governments, which is based on the multiple membership multilevel model (MMMM) (Goldstein 2011). The proposed model makes explicit that many parties are members of several governments over time and that many governments are multiparty coalitions. This is accomplished by including effects of parties in all governments in which they have participated and modeling the total party effect on a government as the aggregated effect of all parties within that coalition. Dependencies within and across governments that originate at the party level are thereby taken into account.

The proposed technique further allows to consider interdependencies among coalition parties in their effect on government outcomes. This is accomplished by endogenizing the weights of the MMMM so that the importance of parties in their effect on governments may depend on their relationship to each other (e.g., balance of power). The approach relaxes the implicit assumption of conventional single-level models (SLMs) that party weights are fixed within and across governments. In doing so, the MMMM opens up opportunities to model and test theories on the interplay of agents in their collective impact on entities, which are impossible to test with SLMs.

While developed to study coalition outcomes, the proposed method has wide applicability to cases characterized by similar interdependence structures, such as data on multi-party wars, treaties, and international organizations. With this paper, I provide the R package ‘rmm’ available on CRAN to estimate the generalized MMMM for a variety of outcomes (linear, logit, conditional logit, Cox, Weibull) in a user-friendly way.

I illustrate the importance of recognizing the multilevel structure of coalition government outcomes by simulation and with an empirical application. In the simulation study, I use real variables from widely employed datasets and manipulate their effect on simulated outcomes (linear and survival) to assess the consequences of party-level clustering and interdependencies for recovering the true effects.

The results show that the SLM underestimates standard errors when the outcome is linear and underestimates regression coefficients and standard errors when the outcome is survival. The proposed MMMM, by contrast, recovers the correct regression coefficients and standard errors for both outcomes. The impact of the bias on the SLM’s false positive rate is considerable: Even though no effect was specified, the regression coefficients were significant 20-60 percent of the time depending on outcome and level of covariate. Moreover, available robust standard errors did not redress the problem. It is, therefore, crucial to model the multilevel structure in order not to primarily pick up spurious effects. This is the case for all studies with coalition government

outcomes, regardless of the level of the explanation, because estimates at all three levels (party, government, and country) are affected.

While the party-level clustering mainly affects the SLM's false positive rate, interdependencies in the aggregation process decrease the model's power – both in the linear and in the survival case. Interdependencies like the one identified in the empirical application reduce the power to detect actual effects at all three levels – in some cases to less than 50 percent. Consequently, researchers are well-advised to model interdependencies when present to uncover actual effects that otherwise are suppressed.

In the empirical application, I apply the MMMM to study the effect of parties' financial dependency on their members on the survival of coalition governments. The results indicate that (i) the more a party's financial resources comprise of contributions from its members, the higher the termination hazard of a government including this party, and that (ii) this effect is contingent upon the power balance in the coalition. These findings would have not been detected using an SLM, owing to its downward-biased regression coefficients.

2 THE MULTILEVEL STRUCTURE OF COALITION GOVERNMENT DATA

Coalition governments exhibit a complex multilevel structure that induces dependencies among observations. The first source of dependencies is the relationship between governments and countries. Governments are nested in countries as evident from Figure 1. This structure is known as *perfect hierarchy*, in which lower-level units belong to a single unit at each higher level. Hierarchical multilevel modeling is commonly used to deal with this type of data structure (Goldstein 2011). The second source of dependencies is the relationship between governments and parties. This relationship is not hierarchical: parties are members of no, one, or several governments over time and, vice versa, many governments are multiparty coalitions. As shown in Figure 1, Israel's labor party, HaAvoda, was a member of all three governments between 2003 and 2011, all of which are coalition governments. This form of *imperfect hierarchy* is known as

multiple membership multilevel structure. The MMMM models this complex data structure (Goldstein 2011).

Figure 2 illustrates such party-level dependencies for government survival in Israel. The median government duration in Israel equals 317 days. The duration of governments including the center party, by contrast, is 200 days shorter. Governments including the center party are therefore more alike than governments overall as they have (at least) one thing in common – the center party. SLMs ignore this form of dependency and assume the independence of observations. SLMs with party-level covariates relax this assumption by introducing observed explanations for this dependency. However, all unknown or unmeasured attributes of the center party that affect government survival are left unmodeled.

In linear regression, ignoring dependencies among observations will lead to downward-biased standard errors while regression coefficients will be unaffected. In nonlinear models, such as logistic and survival models, both regression coefficients and standard errors will be downward-biased. This result holds even when the dependencies are uncorrelated with the covariates in the model (Box-Steffensmeier and Jones 2004, 141–54).

In the simulation below, I demonstrate that the bias is pronounced and that it affects the conclusions we draw about party effects by inflating Type-I and II error rates. The reason for the pronounced bias is that the multiple membership structure in coalition government data is the rule, not the exception. Figure 3 shows the distribution of government participations in the Woldendorp, Keman and Budge dataset (2000). In this data, parties participate in 7.2 coalition governments on average since WWII. The distribution, however, is heavily right skewed, with some parties, such as the Italian Christian democrats, being part of over 40 governments. The bias is likely also pronounced in other fields of political science in which multiple membership structures are encountered.

Previous research has employed two strategies to overcome rather than incorporate the multilevel structure – disaggregating government outcomes to the party level or aggregating party features to the government level. Both strategies are problematic. The disaggregation strategy assigns coalition outcomes to all parties in the coalition and thereby treats the outcomes as independent realizations of each coalition party. This approach thus ignores interdependencies among coalition partners that may have led to the observed outcome. It also ignores dependencies of government outcomes containing the same parties. Moreover, the approach induces Type-I error because data are artificially multiplied. A three-way coalition, for instance, generates three observations in the dataset even though only one outcome is observed, inflating the sample size on which standard error estimates are based. Finally, the estimated effect sizes are often misinterpreted. As I show in the simulation study, regression coefficients estimated at the party level cannot be interpreted at the government level.

Aggregating party features to the government level recognizes that government outcomes depend on all constituent parties. However, by imposing a fixed aggregation function, such as the arithmetic mean, the aggregation strategy also precludes analyzing the interdependencies of parties in their effect on government outcomes. Moreover, the dependencies among observations resulting from the multiple membership clustering are likewise ignored with this approach. Finally, owing to defaults in popular statistical programs such as Stata, it is common practice to take the mean of party features *over available information* and thus to ignore nonrandom missingness at the party level. For instance, in a three-way coalition with missing information on one party, Stata's mean function will take the mean over the features of the available two parties. Research on imputation demonstrates that in the presence of nonrandom missingness and small groups, this strategy induces bias and almost any imputation strategy outperforms this form of listwise deletion (Newman and Sin 2009). Such coding choices are often obscured to the reader (and sometimes the

researcher) because the party level is not explicitly modeled. The following model proposal resolves these problems.

Figure 2: Government duration in Israel between 1950 and 2014 by party participation.

Figure 3: The distribution of government participations in the WKB dataset

3 MODELING THE MULTILEVEL STRUCTURE

For simplicity, the following model structure proposition uses the notation of a linear regression model. However, the model structure can be implemented in a large class of models, including generalized linear models and survival models. Let y_i^G be a coalition government outcome, where subscript $i = 1, \dots, I$ indexes governments and superscript G indicates that the outcome is located at the government level.

I propose to model this outcome in terms of a government-level effect θ_i^G , an aggregated party-level effect θ_i^P , and a country-level effect $\theta_{C(i)}^C$ to recognize the three levels at which explanations of coalition government outcomes originate:

$$Y_i^G = \theta_i^G + \theta_i^P + \theta_{C(i)}^C \quad (1)$$

3.1 Government and country level

The government-level effect θ_i^G in equation (1) is determined by a systematic component of observed explanatory variables at the government level $\mathbf{x}_i^G$ with regression coefficients $\boldsymbol{\beta}^G$ and a random component u_i^G :

$$\theta_i^G = \mathbf{x}_i^{G'} \boldsymbol{\beta}^G + u_i^G \quad (2)$$

The random component at this level is a disturbance term, which is assumed to be normally distributed with a mean of zero and variance $\sigma_{u^G}^2$, i.e. $u_i^G \sim N(0, \sigma_{u^G}^2)$. This part of the model represents the conventional single-level model structure.

The country-level effect $\theta_{C(i)}^C$ in equation (1) models the hierarchical relationship between governments and countries in which each government i is associated with one country k . That is, subscript $k = 1, \dots, K$ indexes countries and the indexing function $C(i)$ returns country k to which government i belongs. The country-level effect is determined by a systematic component of observed explanatory variables at the country level $\mathbf{x}^C$ with regression coefficients $\boldsymbol{\beta}^C$ and a random component u^C representing the joint impact of unobserved country-level effects:

$$\theta_k^C = \mathbf{x}_k^C \boldsymbol{\beta}^C + u_k^C \quad (3)$$

Including the same random component in all governments of the same country captures the unobserved heterogeneity at the country-level that causes coalition outcomes in the same country to be more similar after conditioning on observed variables. The random components are assumed to be normally distributed with a mean of zero and variance $\sigma_{u^C}^2$, i.e. $u_k^C \sim N(0, \sigma_{u^C}^2)$. This part of the model represents the hierarchical linear model structure (known as shared frailty structure in survival modeling).

3.2 Party level

The aggregated party-level effect θ_i^P in equation (1) models the multiple membership relationship between parties and governments. It is determined by a weighted sum of the effect of each party j on government i , p_{ij} , in the set of parties $P(i)$ constituting government i :

$$\theta_i^P = \sum_{j \in P(i)} w_{ij} p_{ij} \quad (4)$$

That is, subscript $j = 1, \dots, J$ indexes parties and the indexing function $P(i)$ returns all parties j that participate in government i . The multiple membership construct aggregates individual party effects p_{ij} by taking their weighted sum with weights $0 \leq w_{ij} \leq 1$ for all parties $j \in P(i)$.

The individual party effects p_{ij} are determined by a systematic component of observed explanatory variables at the party level $\mathbf{x}^P$ with regression coefficients $\boldsymbol{\beta}^P$ and a random component u^P , representing the joint impact of unobserved variables:

$$p_{ij} = \mathbf{x}_{ij}^P \boldsymbol{\beta}^P + u_{ij}^P \quad (5)$$

The random component accounts for the higher similarity of governments containing the same party by including a party's random component in all governments in which it has participated. If we assume that the effects of unobserved variables are constant across a party's government participations, parties each have one random component, i.e. $u_j^P \sim N(0, \sigma_{u^P}^2)$. We may also assume that the effects of unobserved variables change across a party's government participations by

including a random walk centered at the party's random effect of its previous government participation, i.e. $u_{ij}^P \sim N(u_{G(i,j),j}^P, \sigma_{uP}^2)$ ². In this case, parties have as many random components as participations in government. Either way, the random effects are assumed to follow a normally distribution with variance σ_{uP}^2 .

The MMMM aggregates party effects by an additive function. Whether an additive function captures the true link between the party and government level depends on the underlying process that cause parties to affect the government outcome. In general, more complex aggregation functions representing dynamic or emergent effects are possible (Grabisch et al. 2009). However, as I demonstrate next, using a weighted sum already allows to model complex nonlinearities.

In contrast to conventional MMMMs, which only aggregate the random components, I also include the systematic components (i.e., the observed party-level variables) into the aggregation process. The advantages of doing this are threefold. First, the aggregation function is consistent across the systematic and the random part of the model. Second, it draws attention to the fact that we always choose an aggregation function for the party-level variables. I suspect that often enough the function is chosen ad hoc when recoding the variables. By making the aggregation function a modeling choice, we are nudged to base this decision instead on theoretical grounds. Third, and most importantly, it allows to endogenize the aggregation process. Endogenizing the aggregation function not only allows to base the aggregation on theoretical grounds, but it also makes it possible to test those theories and thereby advance our theoretical understanding of the interplay of parties in their effect on governments – an issue to which I turn next.

² The indexing function $G(i, j)$ returns government i^* that party j participated *before* participating in i .

3.3 Endogenizing the aggregation process to model interdependencies

Past research studying party effects often simply averages features of coalition parties to capture the party effect on coalition governments, reducing the aggregation process to taking the arithmetic mean (Bäck 2008; Druckman 1996; Greene 2017). In equation (4), this is equivalent to assigning $w_{ij} = 1/n_i$ with n_i being the number of parties in coalition i . The arithmetic mean is a suitable aggregation function, however, only if the mean captures the underlying processes that causes the party features under scrutiny to affect the government outcome. Previous research using the arithmetic mean to aggregate party effects implicitly assumes that (i) *the weights of coalition parties are identical* and (ii) *the sum of weights is constant across governments*. In many cases these assumptions are not justified from a theoretical perspective. For instance, some coalition parties, such as the prime minister party, may be more (or less) influential than other coalition parties, which would violate (i). Moreover, the overall importance of party effects may depend on structural attributes of parties, coalitions, parliaments, or even countries, which would violate (ii). The effect of parties' internal cohesion, for instance, may be more important in ideologically dispersed coalition governments.

To relax the two assumptions, I model the weights as a nonlinear regression of unobserved w on observed explanatory variables $\mathbf{z}$, instead of assigning fixed weights to each party. I propose the following general form for the weight regression function:

$$w_{ij} = \frac{1}{n_i \exp(-\{\mathbf{z}_{ij}'\boldsymbol{\rho}\})} \quad (6)$$
$$\text{s.t. } \sum_i \sum_j w_{ij} = N$$

where $\mathbf{z}$ are explanatory variables with regression coefficients $\boldsymbol{\rho}$, n_i is the number of parties in coalition i , and N equals the total number of observations at the party level.

The weight regression coefficients estimate the impact of explanatory variables on the specific weight of a party in its effect on government outcomes. If the weight variables have no impact on

the aggregation process, i.e. $\boldsymbol{\rho} = \mathbf{0}$, the weights reduce to $w_{ij} = 1/n_i$ (mean aggregation) and assumptions (i) and (ii) hold. If $\boldsymbol{\rho} \geq \mathbf{0}$, the weights reveal a more complex interplay of parties in their effect on governments. That is, weights will deviate from $w_{ij} = 1/n_i$ and are no longer constant within and between governments. Instead, they depend on $\mathbf{z}^P$, $\mathbf{z}^G$, or $\mathbf{z}^C$, i.e., attributes of parties, governments, or countries.

Weight variables that vary *within governments* allow examining whether some parties are more important than others in their effect on the coalition outcome, i.e. $w_{ij} \geq w_{ij'}$. Weight variables that vary *between governments* (or parliaments, or countries) allow examining whether party effects are more important in some governments (or parliaments, or countries) than in others. Such cross-level interactions make it possible to model the interplay of intra- and inter-party processes.

The functional form $w_{ij} = \frac{1}{n_i \exp(-\{z_{ij}^P \boldsymbol{\rho}\})}$ constrains the weights between 0 and 1 and the sum of weights within governments to, at most, n_i . That is, $0 \leq w_{ij} \leq 1$ and $\forall i: \sum_j w_{ij} \leq n_i$. In governments in which all $w_{ij} = 0$ and, accordingly, $\sum_{j \in P(i)} w_{ij} = 0$, party effects do not matter in determining the outcome. In governments in which all $w_{ij} = 1/n_i$ and, accordingly, $\sum_{j \in P(i)} w_{ij} = 1$, the aggregated party effect is best described by the arithmetic mean of the effect of each government party. In such governments, the outcome changes by $\boldsymbol{\beta}^P$ if all coalition parties jointly increase $\mathbf{x}^P$ by one unit. Such an effect is *contextual* in that it requires all coalition parties to change for the effect to be realized. In governments in which all $w_{ij} = 1$ and, accordingly, $\sum_j w_{ij} = n_i$, the aggregated party effect is the unweighted sum of the effect of each government party. Such an effect is *autonomous* in that increasing $\mathbf{x}^P$ of one party alone by one unit impacts the outcome by $\boldsymbol{\beta}^P$. Moreover, each additional coalition party will increase the total impact of party effects on government outcomes.

To provide identification and prevent a rescaling of the party effects themselves, I constrain the sum of all weights to equal the total number of governments N , i.e. $\sum_i \sum_j w_{ij} = N$. Mean aggregation also constrains the sum of all weights to N by constraining the weights to sum to 1 in each government. In contrast to the constraint imposed by mean aggregation, this constraint allows some governments to have sums larger than 1 and some governments to have sums smaller than 1. However, it does not change the overall sum of weights, which still equals N .

Two examples from government survival

1. *Weight variable varying between parties.* Let x^P be an index describing parties' internal cohesion, z^P be a variable indicating whether a party holds the prime ministership, and the outcome be government survival. If $\rho = 0$, the effect of party cohesion on government survival does not depend on holding the prime ministership. In this case, the effect of party cohesion is best described by the effect of the average cohesion among the coalition parties. Such an effect is contextual in that it requires all coalition parties to increase their cohesion by one unit for it to impact survival by β^P . By contrast, if $\rho \geq 0$, the cohesion of parties holding the prime ministership is more (less) important for government survival than the cohesion of non-prime minister parties.
2. *Weight variable varying between governments.* Let x^P be an index describing parties' internal cohesion, z^G be the ideological heterogeneity of coalitions, and the outcome be government survival. If $\rho \geq 0$, the effect of party cohesion on survival is more (less) important in ideologically dispersed coalitions. Moreover, assume that for ideologically dispersed coalitions, we find that $\sum_j w_{ij} = n_i$. That is, the aggregated party effect is the unweighted sum of the effect of each coalition party. Such an effect is autonomous in that it only requires a one-unit-increase in cohesion of one coalition party to impact survival by β^P .

To conclude, compared to a conventional SLM, the here proposed MMMM model makes three innovations. First, party random effects u_{ij}^P are included so that the dependencies between governments that originate at the party level are taken into account. Second, the aggregation of party effects $\sum w_{ij}p_{ij}$ is explicitly incorporated into the statistical model so that the function can be theoretically grounded. Finally, the weights of parties w_{ij} in the aggregation process are endogenized so that aggregation theories, such as theories about interdependence, can be confronted with data rather than just imposed.

3.4 Model estimation and the rmm package

I employ Bayesian Markov chain Monte Carlo (MCMC) to estimate the model parameters because the likelihood function of the proposed model has no known closed algebraic form and because Bayesian estimation offers advantages. First, Bayesian estimation of multilevel models naturally accounts for uncertainty in the variance estimates while likelihood estimation treats variance estimates as fixed values. Moreover, Bayesian estimation provides variance estimates and plausible credibility intervals in cases where likelihood estimation would give zero variance estimates and no confidence intervals. The carried-out hypothesis tests are therefore more precise (Hodges 2014; Hox, Moerbeek, and van de Schoot 2017). Second, coalition government datasets cover the entire population of governments within a certain spatio-temporal frame. Frequentist statistics models sampling uncertainty even though there is none – the entire population is observed. Bayesian statistics, by contrast, is not based on sampling uncertainty but on a belief-based definition of uncertainty. Such a definition better conveys the nature of uncertainty in coalition outcomes, which is that they are realizations of a probabilistic process.

The model is fitted in JAGS (Plummer 2015; v.4.3) from within R using the rmm package, which is provided with this paper and available on CRAN. The rmm package offers a user-friendly interface to estimate MMMMs with weight functions for a variety of outcomes (linear, logit,

conditional logit, Cox, Weibull). It allows for modeling (i) structural relationships at the party, government, and country level and (ii) the aggregation of party effects to the government level. The three-level multiple membership multilevel version of a Weibull proportional hazard model, which is employed in the application below can specified in the rmm package as follows:

```
rmm(Surv(govdur, earlyterm) ~ 1 + majority + minimalwinning +
    mm(id(pid, gid), mmc(findep), mmw(w ~ 1/offset(n)^exp(-(seatshare+hetero)))) +
    hm(id=cid, type=FE),
    family="Weibull", data=dat)
```

In the `mm()` container, the multiple membership structure between parties and governments is specified in three steps. First, party and government ids must be given in `id()`. Second, party-level covariates can be included in `mmc()`. Third, the aggregation function can be specified in `mmw()`. The here shown functional form corresponds to equation 6. However, researchers can specify their own ideas in a flexible way. Mean aggregation, for instance, is specified by $w \sim 1/\text{offset}(n)$. Finally, in the `hm()` container it is specified that parties and governments are hierarchically nested in countries and that country fixed-effects shall be estimated.

3 SIMULATION STUDY

Rather than generating entirely artificial data, the simulation is based on the same datasets as the subsequent empirical application. In the simulation, I generate a linear and a survival outcome by manipulating the effect of three structural variables (x^P, x^G, x^C) and three weight variables (z^P, z^G, z^C) to those outcomes. The linear outcome, I generate by

$$Y_i^G = \beta^G x_i^G + u_i^G + \beta^C x_{C(i)}^C + u_{C(i)}^C + \sum_{j \in P(i)} \frac{1}{n_i \exp(-\{\rho^P z_{ij}^P + \rho^G z_{ij}^G + \rho^C z_{ij}^C\})} (\beta^P x_{ij}^P + u_j^P) \quad (7)$$

The survival outcome is generated using the same linear predictor and the approach described by Bender et al. (2005). Table 1 details the variables used and the manipulated parameter values.

The data generation process in equation (7) models the actual multilevel structure of coalition government data. In the simulation, I examine the consequences of ignoring this structure by

comparing the performance of the SLM vis-à-vis the MMMM in recovering the true effect of the structural variables (i.e., $\hat{\beta}^P, \hat{\beta}^G, \hat{\beta}^C$) under increasing:

- (i) similarity of observations at the party level (i.e., variance of the random effect σ_{uP}^2)
- (ii) number of multiple memberships m (i.e., # of parties' government participations over time)
- (iii) interdependencies in the aggregation process (i.e., effects of weight variables ρ).

Specifically, I examine the bias of the regression coefficients and their uncertainty estimates, as well as the ensuing consequences for power and false positive rate. To examine bias and power, I specify $\beta^P = \beta^G = \beta^C = 1$. To assess the false positive rate, I specify $\beta^P = \beta^G = \beta^C = 0$. In order to separate the consequences of clustering and interdependencies, I set the weight coefficients (ρ^P, ρ^G, ρ^C) to 0 when assessing the consequences of party-level clustering (i.e., increasing (i) + (ii)). Conversely, I set the party-level clustering (σ_{uP}^2) to 0 when examining the consequences of ignoring interdependencies in the aggregation process (i.e., increasing (iii)).

In total, I examine 33 simulation conditions (i.e., combinations of parameter values) and run each condition 10,000 times. The SLMs are estimated with maximum likelihood and the MMMMs are estimated with Bayesian MCMC. Frequentist properties of Bayesian models are regularly studied in objective Bayesian statistics (Bayarri and Berger 2004; Berger 2006). The Type-I error rate of Bayesian models, for instance, is not known a priori (as compared to the frequentist promise of 5% over the iterations of the simulation when setting the nominal α to 0.05) (Gelman and Tuerlinckx 2000). Bayesian models, however, often have good frequentist properties and it is, therefore, useful to study them.

The following results provide a concise summary of the central findings. More details on the simulation and the full results are reported in Appendix A1. Section 3.1 summarizes the consequences of ignoring party-level clustering and section 3.2 summarizes the consequences of ignoring interdependencies in the aggregation process.

Table 1: Parameters manipulated in the simulation

Level	Variable type	Variable name	Parameter values
Party N=194	Structural	x^P : Parties' financial dependency on members	$\beta^P = \{0,1\}$
	Weight	z^P : Parties' relative seat share in government	$\rho^P = \{-5,-2,0,2,5\}$
	Clustering	Similarity of obs. (variance of random effect)	$\sigma_{u^P}^2 = \{0,0.1,1,2,4\}$
		Number of multiple memberships	$m = \{1,2,4, \text{act. distr.}\}$
Gov. N=401	Structural	x^G : Majority status	$\beta^G = \{0,1\}$
	Weight	z^G : Governments' ideological heterogeneity	$\rho^G = \{-5,-2,0,2,5\}$
	Disturbance		$\sigma_{u^G}^2 = 0$
Country N=18	Structural	x^C : Investiture requirement	$\beta^C = \{0,1\}$
	Weight	z^C : Prime ministerial powers	$\rho^C = \{-5,-2,0,2,5\}$
	Clustering	Similarity of obs. (variance of random effect)	$\sigma_{u^C}^2 = 0$

3.1 Consequences of ignoring party-level clustering

Three conclusions can be drawn about the consequences of ignoring party-level clustering by estimating an SLM. First, it biases the regression coefficients of survival SLMs at all three levels towards zero. The degree of clustering I find in the subsequent empirical application ($\sigma_{u^P}^2 \approx 1$) is associated with underestimating the regression coefficients at all three levels by about 15 percent. The bias is more severe for higher degrees of clustering.

Second, it biases the standard error of both linear and survival SLMs at all three levels towards zero. Robust standard errors, unfortunately, do not redress the bias. Table 2 details the standard error bias by outcome and level of covariate when using a heteroscedasticity-consistent standard error. As visible in Table 2, the bias is greater with a linear outcome.

Table 2: Robust standard error bias in percent by outcome and covariate level for $\sigma_{u^P}^2 = 1$.

Outcome type / Covariate level	Party	Government	Country
Linear	-74%	-58%	-76%
Survival	-43%	-22%	-41%

Linear outcome. Panel A of Figure 4 shows that the false positive rate increases with parties' participations in government. For the actual distribution of parties' participations in governments (shown in Figure 3), the party- and country-level regression coefficients were significant in over 60 percent of the 10,000 runs even though no effect was specified. The government-level

regression coefficient was significant in over 40 percent of the runs. Panel B demonstrates that the false positive rate is affected as soon as there is some similarity at the party level ($\sigma_{u^p}^2 > 0$) and largely depends on the distribution of government participations. This finding demonstrates the importance of accounting for the increased similarity of observations at government level due to party-level dependencies. As parties have accumulated government participations over time, the bias of ignoring party-level clustering has increased to an unacceptable level. Panel D shows that for higher degrees of party-level clustering, the power to detect effects goes down (because the variance of the estimator increases).

Survival outcome. Panel A of Figure 5 shows that survival models also suffer from an excess of Type-I error but to a smaller extent, owing to the downward bias in the regression coefficients. Still, the false positive rate is at 18 percent for government-level variables and 38 percent for party- and country-level variables. Power is hardly affected despite the negative bias in the regression coefficient with an effect size of 1. The power does go down, however, for smaller effect sizes and higher degrees of party-level clustering.

Figure 4 and 5 show that the MMMM's false positive rate and power are unaffected by the party-level clustering. This is because the MMMM estimates both regression coefficients and posterior standard deviations at all three levels without bias.

Figure 4: Percent false positive and percent true positive (power) for a linear outcome.

Figure 5: Percent false positive and percent true positive (power) for a survival outcome

A false positive is recorded when the confidence interval (SLM) / posterior credible interval (MMMM) (CI) does not include 0 even though $\beta^P = \beta^G = \beta^C = 0$. Conversely, a true positive is recorded when the CI does not include 0 and $\beta^P = \beta^G = \beta^C = 1$. Panel A and C show percent false positive and power for $\sigma_{UP}^2 = 1$. Panel B and D show percent false positive and power for the actual distribution of parties' government participations.

3.2 Consequences of ignoring interdependencies in the aggregation process

The simulation demonstrates that imposing mean aggregation when weights are not identical among coalition partners and the sum of weights are not constant across governments biases the regression estimates. The simulation results indicate that both regression coefficients and standard errors of the SLM are affected. This is the case for both linear and survival outcomes and variables at all three levels. The direction of the bias depends on the level of the structural and weight variables. Figure 6 presents the consequences for regression coefficients and standard errors when interdependencies originate at the party level, as found in the empirical application. The results for effects of weight variables at other levels can be found in Appendix A1.

Figure 6: Bias (in %) in regression coefficients and uncertainty estimates due to interdependencies at the party level. Small absolute biases are set to 0 as the associated bias in percent is misleading.

Linear model. Panel A of Figure 6 shows that interdependencies arising at the party level do not bias the regression coefficients at this level but the regression coefficients at the government and country level. Government and country regression coefficients are sharply overestimated when the effect of the weight predictor is positive. Panel C shows that the standard errors at the government and country level are likewise affected. They are overestimated regardless of the sign of ρ^P . In practice, this will offset some of the negative standard error bias found in section 3.1.

Survival model. The bias in the regression coefficients is different for survival outcomes. Panel B indicates that interdependencies amplify the negative bias already found in section 3.1. Moreover, party-level regression coefficients are also affected. The bias in the standard errors, however, follows the same pattern as for linear outcomes. In Appendix A1, I show that the combination of underestimated regression coefficients and overestimated standard errors results in a stark loss of power to detect effects at the government and country level. Therefore, modeling interdependencies – if present – will help uncover effects that are otherwise suppressed by them. One might think that the bias could be lowered by including interactions between structural and weight variables³ in the SLM. In Appendix A1, I show that this is not the case.

Figure 6 shows that the MMMM estimates regression coefficients and posterior standard deviations again without appreciable bias.

³ An SLM including interactions between weight and structural variables and an MMMM including a weight function are not identical. The SLM models the effect of weight variables on the outcome and the MMMM models the weight variables on the weights, which are scaled differently than the outcome.

4 APPLICATION: PARTIES' FINANCIAL DEPENDENCY ON THEIR MEMBERS AND THE INTERDEPENDENCE STRUCTURE OF GOVERNMENT SURVIVAL

To illustrate the advantages of modeling the crisscrossing relationship between parties and government, I apply the MMMM to study the effect of parties' financial dependency on their members on the survival of coalition governments.

4.1 Theoretical considerations

I hypothesize that the more a party's financial resources comprise of contributions from its members, the higher the termination hazard of a government including this party. The reason is that the less financial capital a party receives from sources other than its members, the more it depends on them. Most members cannot be rewarded with office spoils but are reimbursed with promises about future public policy. Such promises confine party leaders who need to build and maintain inter-party agreement and, consequently, increase the risk of political gridlock (Müller and Strøm 1999; Strøm 1990).

Regarding the aggregation of party effects, I examine whether Gamson's law (1961) extends to government survival by testing whether the impact of parties on government survival is proportional to their (relative) legislative seat share. Allowing party weights to depend on their seat share relaxes the first of the two tacit assumptions made when mean-aggregating party effects – that all coalition parties are equally important. Furthermore, I hypothesize that party effects are more important in ideologically heterogeneous coalitions because intra- and inter-party conflicts amplify each other (Tsebelis 2002). Allowing the sum of weights to depend on the coalition's ideological heterogeneity relaxes the second tacit assumption of mean aggregation – that party effects are equally important across governments.

4.2 Data and Measurement

The analysis is based on the Woldendorp, Keman and Budge dataset (2000) to source start and end date of each government, reason for termination, government parties, and their distribution of seats in parliament, on the Comparative Manifestos Project (Volkens et al. 2016) to source political

positions of the government parties, and on Round 1a of the Political Party Database (Scarow, Poguntke, and Webb 2017) to operationalize parties' financial dependency on their rank-and-file.

Government duration is measured as the time between the investiture of a government and new elections. Due to the focus on political gridlock, I differentiate terminations by their underlying motivation (escape gridlock / seize opportunity) rather than by the employed constitutional mechanism (nonelectoral replacements / early elections). To do so, I only consider terminations following some form of conflict and more than one year before of the official end of term as events and censor other terminations.

Parties' financial dependency is measured by the share of financial contributions from party members in the total funding of parties, which the Political Party Database obtained from parties' financial reports.

Parties' relative seat share in government is measured as $S_{ij} = \left(\frac{seats_{ij}}{\sum_j seats_{ij}} - \frac{1}{N_i} \right) \left(\frac{N_i}{N_i - 1} \right)$, where $seats_{ij}$ is the seat share of party j in government i , and N_i is the number of government parties. $S_{ij} = 0$ if j has as many seats as the other parties, $S_{ij} = 1$ if j holds all seats, and $S = -1$ if j holds no seats within the coalition.

Following Warwick (1994), I compute the standard deviation of coalition parties' right-left score to measure governments' ideological heterogeneity. Benoit and Laver (2007) point to the incomparability of the left-right score across countries and time, which is why I divide the coalition's standard deviation by the left-right standard deviation of all parties in the parliament that year. The measure is therefore relative to the ideological distribution at time and place.

I control for whether the government is in the majority in parliament, a minimal winning coalition, and country fixed effects to eliminate all unobserved time-constant country confounds. The final dataset comprises 401 governments from 18 countries (Australia, Austria, Belgium, Czech Republic, Denmark, France, Germany, Hungary, Ireland, Israel, Italy, Netherlands,

Norway, Poland, Portugal, Spain, Sweden, United Kingdom) between 1944 and 2014, and a total of 194 parties. I refer to Appendix A2 for more details on theoretical considerations, data, and measurement.

4.3 Results

The results are divided into three parts. Section 4.3.1 compares the regression coefficients, uncertainty estimates, goodness of fit, and predictive accuracy of the SLM and the MMMM. Section 4.3.2 illustrates what the weight coefficients estimated by the MMMM can tell us about the interdependence structure of governments. Finally, section 4.3.3 presents the results from using the MMMM to decompose the total variance in government survival into party, government, and country components.

4.3.1 The effect of financial dependency

Table 3 presents the regression results. Model 1 is an SLM estimated on party-level data (i.e., each row constitutes a party's participation in government). All variables are significant because this approach drastically underestimates the standard error at all three levels. (In Appendix A1, I demonstrate by simulation that this strategy underestimates standard errors and that regression coefficients cannot be interpreted at the government level.)

Model 2 is an SLM estimated on government-level data, in which the share of contributions from party members in the total funding of parties has been aggregated using the arithmetic mean. The effect of parties' financial dependency is nonsignificant despite a larger effect size because the standard error is larger as well.

Model 3 is an MMMM that accounts for the similarity of observations caused by parties repeatedly appearing in the dataset, but it also uses the arithmetic mean as aggregation function. The effect of parties' financial dependency is considerably larger but remains nonsignificant because the MMMM estimates the actual uncertainty in the effect – roughly double of what the SLM estimates.

Table 3: Regression results

Level	Variable type	Variable	(1) SLM at party level	(2) SLM at gov level	(3) MMMM with $w=1/N$	(4) MMMM with $w=f(Z)$
Party	Structural	Financial dependency	0.134[†] (0.0976)	0.293 (0.2429)	0.376 (0.4655)	0.692* (0.4003)
	Weight	Relative legislative seat share	-	-	-	8.617* (5.0884)
Government	Structural	Majority	-0.279* (0.1201)	-0.385* (0.2154)	-0.418* (0.2358)	-0.388* (0.2397)
	Structural	Minimum winning	-0.897*** (0.1170)	-0.843*** (0.2083)	-0.888*** (0.2313)	-0.903*** (0.2251)
	Weight	Ideological heterogeneity	-	-	-	1.375 (4.8609)
Country	Structural	Fixed effects	yes	yes	yes	yes
N		Parties	1285		194	194
		Governments		401	401	401
		Countries			18	18
DIC			9171	2806	2851	2812
Predictive R^2			0.039	0.039	0.060	0.075

All four models are estimated with Bayesian MCMC. Details on estimation can be found in Appendix A2. The MCMC diagnostics suggest that the chains successfully converged to the posterior distributions. The dependent variable is the hazard rate of government termination, and the reported estimates are the mean and standard deviation of the posterior distribution. I refer to the standard deviation of the posterior as the standard error and to the one-sided tail probability of the posterior as p-value. Note, however, that the standard deviation of the posterior is a more direct measure of uncertainty than the standard error, and the Bayesian analog to the frequentist p-value is a test of direction rather than existence (Marsman and Wagenmakers 2016). The results can be interpreted without this reference to the frequentist paradigm. I chose this framing so that the advantages of modeling the multilevel structure are not mistaken for advantages of a Bayesian approach. Continuous variables are standardized, dividing them by two times their standard deviation, so that the regression coefficients are comparable to those of binary predictors (Gelman 2008). *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, [†] $p < 0.1$.

Model 4 is another MMMM that includes parties' relative seat share and coalitions' ideological heterogeneity as weight predictors. Instead of using a static function, the model allows the aggregation process to depend on the coalition's power balance and potential for conflict. The effect of parties' financial dependency is two times as large as the effect in the SLM. Moreover, the standard error is smaller than in the MMMM using mean aggregation. Therefore, the effect of parties' financial dependency has been suppressed by forcing the weights to be $1/N$ instead of letting them depend on the coalition's interdependence structure.

This result reinforces the conclusion from the simulation study that modeling the aggregation process is key to uncovering party effects when weights are not uniform within and across governments. The impact of parties' financial dependency is considerable. The probability of reaching the 4-year mark decreases by about 13 percentage points per 10 percent of funding by

members. This more than the difference in survival probability between minority and majority governments (10 percentage points).

The simulation study also shows that regression coefficients and standard errors of variables at other levels are affected. This is visible, for instance, for the effect of minimal winningness, which is larger in the MMMM than in the SLM.

The MMMM enjoys good model-fit and higher predictive power than the SLM. The reported predictive R^2 measures the proportion of explained variance by the linearly corrected prediction function, which quantifies the potential predictive power of survival models (Li and Wang 2019). This power is two times higher for the MMMM ($R_{M4}^2 = 0.08$) than for the SLM ($R_{M2}^2 = 0.04$). The deviance information criterion (DIC) indicates that modeling the aggregation process improves model-fit considerably. Like the AIC, the DIC is based on the trade-off between model-fit and complexity and allows comparing non-nested models. Smaller DICs indicate better model-fit and differences greater than 10 are considered significant (Spiegelhalter et al. 2002). The DIC of the MMMM including weight predictors is 39 points lower than the MMMM with mean aggregation. The DIC, however, is 6 points higher than the SLM. This is because the DIC favors models with a smaller number of parameters. Unbiased hypothesis tests, however, are worth the increase in complexity in my view.

4.3.2 The interdependence structure of government survival

The results provide strong evidence that the balance of power in the coalition influences the degree to which a coalition party's financial dependency impacts government survival. The financial dependency of a party with more seats relative to other coalitions partners is more important in the effect on government survival than the financial dependency of a party with fewer seats. Panel A of Figure 7 illustrates this relationship. Mean aggregation thus overestimates the impact of parties with fewer seats and underestimates the impact of parties with more seats, which is shown in Panel B of Figure 7. To give an example, in a 3-party coalition with one party holding

60 percent ($S = 0.4$) of the coalition's seat share and the other two each holding 20 percent ($S = -0.2$), the MMMM estimates the weights to be $w = (0.7, 0.1, 0.1)$. That is, the MMMM reveals that primarily the first party's financial dependency impacts the coalition's termination hazard. Mean aggregation, by contrast, assigns $w = (0.\bar{3}, 0.\bar{3}, 0.\bar{3})$ without consulting the data, which is the source of effect suppression in model (2) and (3). Unequal seat distributions like in this example ($S \geq |0.4|$) are very common; 59 percent of the coalitions fall into this category.

The results do not provide evidence that parties' financial dependency is more or less important in ideologically heterogeneous coalitions. Panel C of Figure 7, however, shows that the importance of parties' financial dependency nonetheless varies across governments. The sum of weights peak at $S = 0$, which suggests that parties' financial dependency matters most in coalitions with equal power distribution.

Figure 7: Party weights as a function of their relative seat share in the coalition.

4.3.3 The origin of variance in government survival

Finally, the multilevel approach provides estimates of how much variation in government survival originates at each level. As Table 4 shows, unsurprisingly, most variability in government survival is located at the government level. However, 34 percent of the variation originates from differences between governments at the party level, and 24 percent originates from differences between governments at the country level. Consequently, while it is standard in the literature to

include country-level predictors, this result vindicates recent efforts to integrate party-level explanations into explanations of government survival.

Table 4: The variance in government survival at each level

		Variance	Percent of total variance
Level	Party	0.940 (0.3726)	34.4
	Government	1.130 (0.0761)	41.3
	Country	0.666 (0.2442)	24.3

MMMM without covariates to estimate the variance at each level.
Mode (standard deviation) of the posterior distribution.

5 CONCLUSIONS

Research on coalition governments increasingly directs attention to party-level explanations of coalition outcomes (Giannetti and Benoit 2009). This research, however, ignores the complex multilevel structure between parties and governments. In addition to their nesting in countries, parties often become nested in multiple governments over time and governments are often nested in coalitions of multiple parties.

In this paper, I show that the crisscrossing relationship between parties and governments (i) implies dependencies among observations that need statistical control, and that it (ii) exposes interdependencies between parties in their effect on governments that need theorization. I then advance a new statistical method to address both issues. The proposed method represents the multilevel structure of coalition governments by modeling the total party effect on a government as the aggregated effect of all parties in the coalition and by including effects of parties in all governments in which they have participated.

The approach is based on the multiple membership multilevel model (MMMM), which I generalize in two ways. First, while a standard MMMM only aggregates the random part of the model, the proposed model aggregates both fixed and random parts to provide a consistent aggregation of party effects to the government level. Second, I propose to endogenize the weights

of an MMMM so that the weights of parties in their effect on a government are no longer fixed but can depend on interdependencies among coalition partners. With this paper, I provide the R package ‘rmm’ to estimate the generalized MMMM for a variety of outcomes (linear, logit, conditional logit, Cox, Weibull) in a user-friendly way.

I illustrate the importance of modeling the multilevel structure by simulation and with an empirical application. The simulation demonstrates that the generalized MMMM recovers the correct regression coefficients and standard errors in the presence of party-level clustering and interdependencies and enjoys good model-fit and higher predictive accuracy than the conventional single-level model (SLM). The SLM, ignoring party-level clustering and interdependencies, by contrast, is biased.

Ignoring party-level clustering leads to underestimated standard errors in linear regression and underestimated regression coefficients and standard errors in survival analysis. This behavior has considerable consequences for hypothesis testing and a precise identification of causal effects. Both false positive rate and power are affected: On datasets widely employed in the literature, government-level regression coefficients were significant over 40 percent and party- and country-level regression coefficients were significant in over 60 percent of the time even though no effect was specified in the simulation. The power to detect effects went down by 20 percentage points for higher degrees of party-level clustering because the variance of the estimator surges.

In survival analysis, the government-level regression coefficients were significant over 20 percent and party- and country-level regression coefficients were significant over 40 percent of the time even though no effect was specified. The power to detect effects went down by 25 percentage points for higher degrees of party-level clustering, owing to the downward bias in the regression coefficients and an increased variance of the estimator.

Ignoring interdependencies in the aggregation process biases both regression coefficients and standard errors at all three levels both in linear regression and survival analysis. The bias is

particularly consequential in survival analysis where it further decreases the model's power. Consequently, when interdependencies are present, modeling them is key to uncover effects that, otherwise, are suppressed.

To summarize, the bias of ignoring party-level clustering and interdependencies is substantial and affects covariates at all three levels. Worse yet, the bias continues to grow as parties continue to accumulate government participations over time and their vote shares continue to decline, leading to larger coalitions (Drummond 2006). The simulation shows that available robust standard errors do not redress the problem. It is, therefore, key to model the dependencies within and across governments regardless the level of the covariates considered.

In the empirical application, I employ the MMMM to examine the effect of a party's financial dependency on its members on government survival. The results indicate that the more a party's financial resources comprise of contributions from its members, the higher the termination hazard of a government including this party. This finding provides evidence for the idea that intra-party politics confines party leaders who need to build and maintain inter-party agreement. The results also show that the financial dependency of parties with more seats relative to other coalitions partners is more important in the effect on government survival. Thus, the effect of a party's financial dependency is contingent upon the power balance in the coalition. Owing to the downward bias in the regression coefficients, the effect of a party's financial dependency is not detected using a SLM. Finally, the MMMM allows partitioning the variation in government survival into components at each level. While, unsurprisingly, most variance is located at the government level (42%), the results reveal more variance at the party level (35%) than at the country level (23%). This finding, therefore, vindicates efforts to integrate intra-party politics into explanations of government survival.

In conclusion, this paper contributes to the literature on coalition governments by dissecting the analytical levels of explanations of coalition outcomes and the link between those levels. The

paper also proposes a novel way to model the interdependence structure of political processes, not wholly unlike the spatial autoregressive model (SARM) (Juhl 2020). The difference is that the SARM models how parties affect each other's outcomes while the MMMM models interdependences in the joint effect of parties on outcomes at a higher level. This feature makes it possible to test theoretical expectations on aggregation processes. Examples include the aggregation of player strategies in game-theoretic models (Lupia and Strøm 1995) and spatial theory (Laver and Shepsle 1996), or indices that aggregate party features to a higher level, such as power indices (Banzhaf 1964). Theoretical expectations on such processes can be tested with the MMMM by translating them into weight functions.

The paper demonstrates what a multilevel approach can add to research on coalition governments. The proposed approach, however, has the potential to advance our understanding of other political research areas characterized by similar interdependence structures, such as multiparty wars, treaties, and international organizations.

REFERENCES

- Bäck, Hanna. 2008. "Intra-Party Politics and Coalition Formation." *Party Politics* 14 (1): 71–89.
- Bäck, Hanna, Marc Debus, and Patrick Dumont. 2011. "Who gets what in coalition governments?: Predictors of portfolio allocation in parliamentary democracies." *European Journal of Political Research* 50 (4): 441–78.
- Banzhaf, John F. 1964. "Weighted voting doesn't work: A mathematical analysis." *Rutgers Law Review* 19 (2): 317–43.
- Bayarri, M. J., and J. O. Berger. 2004. "The Interplay of Bayesian and Frequentist Analysis." *Statistical Science* 19 (1).
- Bederke, Paul, Holger Döring, and Sven Regel. 2020. "Party Facts project." 2020. <https://partyfacts.herokuapp.com/>.
- Benoit, Kenneth, and Michael Laver. 2007. "Estimating party policy positions: Comparing expert surveys and hand-coded content analysis." *Electoral Studies* 26 (1): 90–107.
- Berger, James. 2006. "The case for objective Bayesian analysis." *Bayesian Analysis* 1 (3).
- Box-Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event history modeling: A guide for social scientists*. Cambridge: Cambridge University Press.
- Ceron, Andrea. 2016. "Inter-factional conflicts and government formation." *Party Politics* 22 (6): 797–808.
- Cross, William P., and Richard S. Katz. 2013. *The challenges of intra-party democracy. Comparative Politics*. Oxford: Oxford University Press.
- Döring, Holger, and Philip Manow. 2016. *ParlGov 2015 Release*.
- Druckman, J. N. 1996. "Party Factionalism and Cabinet Durability." *Party Politics* 2 (3): 397–407.

- Drummond, Andrew J. 2006. "Electoral Volatility and Party Decline in Western Democracies: 1970–1995." *Political Studies* 54 (3): 628–47.
- Gamson, William A. 1961. "A Theory of Coalition Formation." *American Sociological Review* 26 (3): 373–82.
- Gelman, Andrew, and Francis Tuerlinckx. 2000. "Type S error rates for classical and Bayesian single and multiple comparison procedures." *Computational Statistics* 15 (3): 373–90.
- Giannetti, Daniela, and Kenneth Benoit. 2009. "Intra-Party Politics and Coalition Governments." In *Intra-Party Politics and Coalition Governments*, eds. Daniela Giannetti and Kenneth Benoit. London, New York: Routledge, 3–24.
- Giger, Nathalie, and Gijs Schumacher. 2015. *Integrated Party Organization Dataset (IPOD)*.
- Goldstein, Harvey. 2011. *Multilevel statistical models*. 4th ed. Oxford: Wiley.
- Grabisch, Michel, Jean-Luc Marichal, Radko Mesiar, and Endre Pap. 2009. *Aggregation functions*. Cambridge: Cambridge University Press.
- Greene, Zachary. 2017. "Working through the issues: How issue diversity and ideological disagreement influence coalition duration." *European Political Science Review* 9 (04): 561–85.
- Hodges, James S. 2014. *Richly parameterized linear models: Additive, time series, and spatial models using random effects*. Boca Raton: CRC Press.
- Hox, Joop J., Mirjam Moerbeek, and Rens van de Schoot. 2017. *Multilevel Analysis: Techniques and Applications*. 3rd ed. *Quantitative methodology series*. New York: Routledge.
- Juhl, Sebastian. 2020. "The Sensitivity of Spatial Regression Models to Network Misspecification." *Political Analysis* 28 (1): 1–19.
- Laver, Michael, and Kenneth A. Shepsle. 1996. *Making and breaking governments: Cabinets and legislatures in parliamentary democracies*. Cambridge: Cambridge University Press.

- Li, Gang, and Xiaoyan Wang. 2019. "Prediction Accuracy Measures for a Nonlinear Model and for Right-Censored Time-to-Event Data." *Journal of the American Statistical Association* 114 (528): 1815–25.
- Lupia, Arthur, and Kaare Strøm. 1995. "Coalition Termination and the Strategic Timing of Parliamentary Elections." *American Political Science Review* 89 (3): 648–65.
- Marsman, Maarten, and Eric-Jan Wagenmakers. 2016. "Three Insights from a Bayesian Interpretation of the One-Sided P Value." *Educational and Psychological Measurement* 77 (3): 529–39.
- Martínez-Cantó, Javier, and Henning Bergmann. 2019. "Government termination in multilevel settings. How party congruence affects the survival of sub-national governments in Germany and Spain." *Journal of Elections, Public Opinion and Parties* 97 (2): 1–21.
- Martínez-Gallardo, Cecilia. 2010. "Out of the Cabinet: What Drives Defections From the Government in Presidential Systems?" *Comparative Political Studies* 45 (1): 62–90.
- Müller, Wolfgang C., and Kaare Strøm, eds. 1999. *Policy, office, or votes?: How Political Parties in Western Europe Make Hard Decisions*. Cambridge: Cambridge University Press.
- Newman, Daniel A., and Hock-Peng Sin. 2009. "How Do Missing Data Bias Estimates of Within-Group Agreement?" *Organizational Research Methods* 12 (1): 113–47.
- Plummer, Martyn. 2015. *JAGS Version 4.0.0 user manual*.
- Poguntke, Thomas, Susan Scarrow, and Paul Webb. 2017. *Political Party Database Version 1a.3*.
- Saalfeld, Thomas. 2009. "Intra-Party Conflict and Cabinet Survival in 17 West European Democracies, 1945-1999." In *Intra-Party Politics and Coalition Governments*, eds. Daniela Giannetti and Kenneth Benoit. London, New York: Routledge, 169–86.

- Scarrow, Susan, Thomas Poguntke, and Paul Webb, eds. 2017. *Organizing Representation. Political Parties, Participation, and Power*. Oxford: Oxford University Press.
- Spiegelhalter, David J., Nicola G. Best, Bradley P. Carlin, and Angelika van der Linde. 2002. "Bayesian measures of model complexity and fit." *Journal of the Royal Statistical Society: Series B* 64 (4): 583–639.
- Strøm, Kaare. 1990. "A Behavioral Theory of Competitive Political Parties." *American journal of political science*: 565–98.
- Strøm, Kaare, W. C. Müller, and T. Bergman, eds. 2008. *Cabinets and Coalition Bargaining: The Democratic Life Cycle in Western Europe. Comparative Politics*. Oxford: Oxford University Press.
- Tsebelis, George. 2002. *Veto Players: How Political Institutions Work*. Princeton: Princeton University Press.
- Volken, Andrea, Pola Lehmann, Theres Matthieß, Nicolas Merz, and Sven Regel. 2016. "The Manifesto Data Collection. Version 2016b." Berlin: WZB.
- Warwick, Paul. 1994. *Government Survival in Parliamentary Democracies*. Cambridge: Cambridge University Press.
- Woldendorp, Jaap, Hans Keman, and Ian Budge. 2000. *Party Government in 48 Democracies (1945 -1998): Composition-Duration-Personnel*. Dordrecht: Springer.