National Electoral Thresholds and Disproportionality

Tasos Kalandrakis[*]   Miguel R. Rueda[†]

July 25, 2019

Abstract

We develop a Maximum a Posteriori Expectation-Maximization (MAP-EM) algorithm to recover national electoral thresholds of representation and of electoral disproportionality from observed seats/votes data. We apply the procedure to 118 electoral systems used in 417 elections to the lower house across 36 European countries since WWII. We find that over half of these systems exhibit a statistically significant positive national threshold of representation. Furthermore, the two modal electoral system configurations involve positive thresholds with allocation for parties exceeding thresholds that does not statistically differ from perfectly proportional allocation (38.14% of all systems); and disproportional seat allocations with (statistically) negligible thresholds of representation (31.36% of all systems). We also develop procedures to evaluate model fit and to test for changes in electoral institutions.

---

[*]Department of Political Science and Department of Economics, University of Rochester. E-mail: kalandrakis@rochester.edu.

[†]Department of Political Science, Emory University.

1 Introduction

How easy is it for political parties to earn representation in the national legislature? Does the process of translating votes into seats favor larger parties? Though related, these are *distinct* questions that are central to the study of electoral systems and their consequences. An electoral system may impose a high barrier to representation in the form of electoral thresholds and at the same time exact proportional allocation for parties eligible for seats. Conversely, a system may allow even the smallest parties to earn a seat but grant marked seat advantages to larger parties.[1] These alternative combinations of electoral system provisions engender sharply different incentives for electoral competition, the former allowing many moderately sized parties to compete as equals, the latter promoting top-heavy competition between two dominant parties flanked by a possibly large number of fringe alternatives. We therefore believe that these two quantities (threshold of representation and disproportionality of allocation for parties above threshold) deserve separate measurement and evaluation of their (possibly independent) consequences. We take up the first task of empirical measurement in this paper.

We make two main contributions: First, we develop a method to obtain conceptually distinct measures of empirical thresholds of representation and of electoral disproportionality favoring larger parties from electoral returns data. We use real (instead of simulated) electoral returns to estimate these quantities, so that the measures reflect the marginal impact of vote changes on seat allocations under actual, competitive electoral forces. Our approach builds on a classic votes-to-seats curve model (e.g. Taagepera 1973; Tufte 1973; Schrodt 1981; Grofman 1983; King and Browning 1987; King 1990). We preserve the disproportionality (or responsiveness) parameter shared by these previous studies and enrich it by allowing for

---

[1]In fact, our empirical study finds that these are the modal categories of electoral systems in modern European democracies.

a stochastic national threshold of representation. To efficiently estimate this enriched model we develop a Maximum a Posteriori (MAP) estimator using an Expectation-Maximization (EM) algorithm (Dempster, Laird and Rubin 1977).[2] The resulting estimated parameters are statistics in the classic sense of providing efficient empirical summaries of the votes-to-seats relationship, and come with a gauge of confidence in these estimates in the form of standard errors.

We implement the MAP-EM procedure in 118 electoral systems used in 417 elections to the lower house across 36 European countries since WWII. We find statistically significant national thresholds of representation in over half of the electoral systems in our sample. We also find that the systems that offer more advantages to small parties (those with statistically negligible thresholds of representation and seat allocations statistically indistinguishable from proportional representation), account for only 14.41% of all systems. The two modal electoral system configurations involve either systems with positive thresholds and proportional allocation for parties exceeding thresholds (38.14% of all systems), or disproportional seat allocations with no thresholds (31.36% of all systems).[3]

As a second contribution, we develop a battery of inference procedures that allow us to evaluate both model fit and alternative hypotheses regarding the function of the electoral

---

[2]This type of MAP-EM estimator has been previously introduced by Levitan and Herman (1987) for statistical image reconstruction and by Fraley and Raftery (2007) for the optimal classification of finite mixture models.

[3]Here and whenever we refer to estimation findings in what follows, proportional seat allocation means that we cannot reject the hypothesis that seat shares equal vote shares, while disproportional systems are those for which we can reject the hypothesis of proportional allocation for parties above thresholds in favor a disproportional allocations favoring larger parties. Similarly we find significant thresholds when we can reject the hypothesis that the threshold is equal to zero in favor of a positive alternative.

system. Our tests of fit are suitable for both model evaluation and model comparison. For example, we reject the hypothesis that the data are generated from a model that assumes no threshold in favor of our model in 51 systems. Importantly, we develop a test procedure to address a long-standing problem in the literature, namely whether observed or unobserved changes in the electoral system have led to statistically significant changes in the pattern of votes-to-seats allocation? Our test allows analysts to objectively resolve whether the electoral system and the votes-to-seats relationship it determines have changed or not? To illustrate the applicability of the test, we assess three alternative definitions of electoral systems in their ability to capture changes in the votes-seats allocation over time. In the first definition an electoral system persists if there are no changes in the seat allocation formula and if the number of districts, seats, and legal threshold remain within 5% of the system's average. The second definition identifies a new electoral system only when the allocation formula changes. In the third, there is a new electoral system whenever any change relevant to seat allocation occurs in the electoral law. Our test provides conditional support for the 5% cutoff definition over the alternatives.

Our emphasis on separate and distinct measurement of national thresholds of representation is consistent with increasing interest, theoretical, normative, and empirical, on that aspect of national electoral systems. On the theoretical side, electoral thresholds figure prominently in analytical accounts that explore the determinants of the number of viable candidates in elections.[4] On the normative side, while it is tempting to dismiss the significance of national thresholds on the grounds that (by their very nature) they only affect the viability of representation for relatively small parties, it is important to appreciate that these "small" parties may represent minorities, political or ethnic, struggling for their integration in the political process. Because electoral thresholds often reflect explicit intent to bar repre-

---

[4]For example, the upper bound of $M + 1$ viable candidates at the district level of Cox (1997) is premised on the threshold of representation implicit in a district allocating M seats.

sentation to such political minorities, they are of interest to a growing community of scholars that specialize on the ethical and positive issues of minority representation, including the extensive attention to thresholds devoted by the Venice Commission, the advisory body of the Council of Europe on constitutional matters European Commission for Democracy through law (Venice Commission) (2018).[5] It is thus essential on normative grounds to be able to, at a minimum, determine the degree to which such minorities are barred from having a voice in parliament.

The importance of estimating national thresholds of representation is also highlighted by the fact that the choice to institute them, directly or implicitly, is rarely politically innocuous. When they are not blatantly intended at barring representation to politicized ethnic minorities, high thresholds are often defended citing the need to avoid the proliferation of many frivolous parties. Such justification was used by supporters of the relatively high German electoral thresholds instituted after the second world war (the proliferation of parties implicitly debited as one of the flaws of the pre-war system). In this, as in all cases, thresholds are instituted by winning majorities, broad or marginal, depending on the rules for setting the electoral system. It is hard to imagine that savvy political actors involved in these decisions do not take into account which of the larger parties are likely to benefit the most from any votes diverted from smaller and, due to thresholds, no-longer viable parties. Even a small such advantage may prove crucial for a larger party competing for government.[6]

---

[5]The unusually high 10% threshold provided by the Turkish electoral system is often cited as an obvious case aimed at preventing representation to the Kurdish minority (European Commission for Democracy through law (Venice Commission) (2018), page 11).

[6]For example, if the main existing small party affected by a threshold has a conservative ideological leaning, then it is reasonable to expect that large governing parties of the right are more willing to institute an electoral threshold, anticipating to be the main beneficiaries of any votes diverted from the smaller right-wing party that can no longer earn representation.

A challenge faced by students of electoral systems is that the actual level of thresholds of representation is often hardly discernible from the provisions of the electoral law. Even when explicitly codified in the main body of the electoral law, legal thresholds often apply at the subnational level, or are qualified by additional provisions for the allocation of seats, so that the threshold parties effectively face at the national level cannot be directly determined from the provisions of the electoral system alone. In the example of Germany, a high national legal threshold of 5% was in place for most of the post-war years but, at the same time, the electoral law allowed parties below that threshold to earn seats by winning a plurality of the vote in one of the single member districts. Therefore, despite the fact that the actual expected threshold in Germany is arguably lower than 5% by virtue of this additional provision of the law, the numerical value analysts may assign to the expected level of the threshold is both not obvious *and* ought to come with a gauge of uncertainty we can place in it. Indeed, our point estimates of the German threshold over time reflect a level lower than 5%, but with a sizable confidence interval around them. In the minority of cases when the electoral law determines the national electoral threshold explicit and unambiguously (for example, this is so for the Greek electoral systems in the period from 2003 to 2009), our estimator successfully recovers the actual threshold.

Besides providing a rigorous and comparable method to evaluate the national level of electoral thresholds, our study also offers additional advantages when it comes to measuring disproportionality. Of course, thresholds contribute to the overall pattern of disproportionality, but our second disproportionality measure is different and distinct because it is specific to parties that exceed these thresholds. We argue that this separation between the two measures provides both more accurate and more politically relevant resolution on the pattern of votes-seats allocation. Disproportionality is of primary interest to positive scholars of electoral systems as an indication of how likely it is that the largest among the larger parties captures a majority in parliament. For that purpose, we claim that it is often more informa-

tive to know whether the process of seat allocation favors larger parties among the parties *viable* for seats, rather than having an overall gauge of disproportionality. Returning to the German example, seats-votes data will show higher levels of disproportionality whenever parties close to the national threshold fail to obtain representation. But high levels of overall disproportionality need not translate to a high probability of a parliamentary majority for the largest party in each election because seats are allocated nearly perfectly proportionaly for parties above the %5 threshold. As we already pointed out in the introductory paragraph, the two quantities (thresholds and disproportionality for parties above thresholds) can be manipulated independently by authors of electoral laws and, in fact, vary independently in the data.

This discussion is independent of the related question of what kind of criterion of disproportionality a given measure captures? We eschew the latter question, though it still applies whether we choose to measure disproportionality for all parties (as the extant literature does) or only for parties above the electoral threshold, as we propose. It is by now well understood that there are many such alternative measures Gallagher (1991); Cox and Shugart (1991) (at least as many as the various proportionality formulas) and that differences between most measures matter less as the number of allocated seats increases. Our measure has the conceptual backing of Theil's Theil (1969) minimum entropy criterion and captures well the political character of disproportionality Cox and Shugart (1991).

Before we proceed with the analysis, we take the opportunity to relate our work to different strands of the literature, besides that already cited. Various models of votes-to-seats allocation related to ours have been used by numerous authors in order to theoretically and empirically characterize the pattern of allocation in two and multi-party contexts (e.g. Theil 1969; Taagepera 1973; Tufte 1973; Schrodt 1981; Grofman 1983; Jackman 1994; Calvo 2009; Linzer 2012; Calvo and Rodden 2015), to explore possible partisan bias or party-specific swing ratios (King and Browning 1987; King 1990; Linzer 2012), to understand the

6

historical process of adoption of Proportional Representation (PR) (Calvo 2009), or the effects of geographical dispersion of electoral support on party systems (Calvo and Rodden 2015). We differ from these authors in that we introduce and jointly estimate national electoral threshold parameters.

Of course, we are not the first to attempt to quantify electoral thresholds and a number of authors have proposed ways to convert observed electoral provisions into a national threshold of representation (e.g. Taagepera 1989; Taagepera and Shugart 1989; Gallagher 1992; Taagepera 1998*a,b*; Lijphart 1994; Taagepera 2002; Ruiz-Rufino 2007; Taagepera and Shugart 2017). Naturally, thresholds are determined by a confluence of formal provisions along with unobserved or harder to quantify features of the electoral system such as districting practices or the specific configuration of party forces that render representation viable in each instance. Among the advantages of our estimation approach is that it allows us to quantify the combined effect of these possible determinants of barriers to representation; and that it comes with a gauge of the confidence we can place on the resulting measurement in the form of a standard error. Furthermore, we are able to jointly estimate thresholds and disproportionality above thresholds as distinct summary statistics of the electoral system: disproportionality reflects any bias in the allocation of seats in favor of (or against) larger parties *conditional* on exceeding the national threshold, while thresholds indicate the expected vote share a party must reach in order to receive any seats. As we already emphasized, this conceptual separation is both substantively relevant on *a priori* grounds, and empirically fits the data well in a substantial fraction of electoral systems.

2   Model & Estimation

In this section we specify a statistical model of the allocation of seats as a function of vote shares that incorporates a threshold of representation and outline our estimation

strategy. Consider N parties that compete across T elections. Index parties by $i = 1, \ldots, N$ and elections by $t = 1, \ldots, T$. Let the vote share of party i in election t be denoted by $v_{t,i} \geq 0$, and denote the number of seats allocated to that party in election t by $s_{t,i} \geq 0$. Let the vector of realized seat allocations in election t be denoted by s_t and the corresponding vector of vote shares by v_t. We take the standard perspective that s_t follows a distribution given v_t and assume that s_t are conditionally independent across elections.[7]

To accommodate seat allocation processes that provide for a (possibly unobserved to the analyst) threshold of representation, we assume that in each election t a latent variable θ_t^* is drawn from a distribution $f(\theta_t^* \mid \theta, \sigma)$, independently across elections, and a national electoral threshold θ_t is realized as a function of θ_t^*. Specifically, the threshold θ_t is zero if the latent variable θ_t^* is negative and is equal to θ_t^* otherwise, that is, $\theta_t = \max\{0, \theta_t^*\}$. Turning to the distribution of seats in election t given the realized threshold, θ_t, we start with the (provisional) assumption that parties whose vote share falls below the realized threshold receive no seats, that is, $s_{t,i} = 0$ when the vote share of party i satisfies $v_{t,i} < \theta_t$. For parties that exceed the threshold in election t, we assume (as in, for example, King (1990)) that the allocation of seats follows a multinomial distribution

$$
(1) \qquad p(s_t \mid v_t, \beta, \theta_t) = Multinomial\left[q_1(v_t, \beta, \theta_t), \ldots, q_N(v_t, \beta, \theta_t); \sum_{i=1}^{N} s_{t,i}\right],
$$

where the expected seat share of party i is given by

$$
(2) \qquad q_i(v_t, \beta, \theta_t) = \begin{cases} \dfrac{v_{t,i}^{\beta}}{\displaystyle\sum_{j:v_{t,j} \geq \theta_t} v_{t,j}^{\beta}} & if \quad v_{t,i} \geq \theta_t \\[2em] \dfrac{\mathbb{I}(v_{t,i} = \max_j v_{t,j})}{\sum_k \mathbb{I}(v_{t,k} = \max_j v_{t,j})} & if \quad v_{t,i} < \theta_t. \end{cases}
$$

[7]Observe that typically a given *national* vote share may result in more than one national seat allocations because of *district*-level variation in parties' electoral strengths.

Here, β is a *disproportionality* parameter, while $\mathbb{I}(\cdot)$ is an indicator function taking the value one if the expression in parentheses is true, and zero, otherwise. The second line of (2) covers the case the realized threshold exceeds the vote share of the plurality party (admittedly a negligible event in our estimation and certainly in the data), specifying that seats are allocated with equal probability among parties tied in the plurality position.

We complete the statistical model of seat allocation by specifying a parametric normal distribution for the latent threshold variable, that is, $\theta_t^* \sim f(\theta_t \mid \theta, \sigma) := N(\theta, \sigma^2)$. This parametric form allows us to compute in closed form the expected national threshold $\bar{\theta}(\theta, \sigma)$ and its standard deviation $\bar{\sigma}(\theta, \sigma)$ as functions of the parameters θ, σ (see online Appendix D). Naturally, the expected threshold $\bar{\theta}$ is of primary importance for our purposes, though the nuisance parameter is also of potential relevance —and certainly necessary given the stochastic perspective we take on the data. In turn, parameter β serves as a natural (dis)proportionality parameter for the seat allocation among parties that exceed the threshold, implying proportional representation (in expectation) when $\beta = 1$, disproportional allocations favoring larger parties when $\beta > 1$, and disproportional allocations favoring smaller ones when $\beta < 1$.[8] Theil (1969) provides a justification for this parameter choice, but it has a long tradition in empirical models of seats-votes relationships, especially in two-party systems (e.g., Kendall and Stuart 1950; Taagepera 1973; Tufte 1973; Schrodt 1981).

With the model thus specified, we now derive a likelihood function in order to motivate our estimator and provide insight as to the nature of information present in the data that allows us to recover the threshold parameters. We exploit a number of simplifications permitted by the structure of our estimation problem. First, we infer from the data that the

---

[8]By virtue of being directional, our measure captures the "political character of disproportionality" (Cox and Shugart 1991).

threshold in period t did not exceed

$$(3) \qquad \bar{v}_t = \min_{i \in \{1,\ldots,N\}} \{v_{t,i} \mid s_{t,i} > 0\},$$

that is, the minimum vote share among parties that received seats. A second observation that we exploit is the fact that if the realized threshold θ_t falls in any interval $(\ell, u] \subset [0, \bar{v}_t]$ and there does not exist a party i with vote share $v_{t,i}$ in (ℓ, u), then

$$(4) \qquad p(s_t \mid v_t, \beta, \theta_t) = p(s_t \mid v_t, \beta, u).$$

This is simply a consequence of the fact that the exact realization of the threshold does not matter as long as the set of parties with vote share at or above the threshold remains identical. Accordingly, suppose there are n_t distinct parties with vote shares $v_{t,i_1}, \ldots, v_{t,i_{n_t}}$ that are less than or equal to the upper bound of the realized threshold $\bar{v}_t$. We index these by $z_t = 1, \ldots, n_t$ in increasing order and define $u_{t,z_t} := v_{t,i_{z_t}}$, $z_t = 1, \ldots, n_t$, so that

$$u_{t,1} = v_{t,i_1} < u_{t,2} = v_{t,i_2} < \ldots < u_{t,n_t-1} = v_{t,i_{n_t-1}} < u_{t,n_t} = v_{t,i_{n_t}} = \bar{v}_t.$$

Similarly, we can define for each $z_t = 1, \ldots, n_t$ a corresponding lower bound

$$\ell_{t,z_t} := \begin{cases} u_{t,z_t-1} & \text{if} \quad z_t > 1 \\ -\infty & \text{if} \quad z_t = 1. \end{cases}$$

We may now view z_t as the realization of a random variable Z_t with support $\{1, \ldots, n_t\}$, with the interpretation that z_t *denotes the interval* $(\ell_{t,z_t}, u_{t,z_t}]$ *within which the latent threshold is*

realized. Clearly, the probability the latent threshold $\theta_t^* \in (\ell_{t,z_t}, u_{t,z_t}]$ is

$$(5) \qquad p(z_t \mid v_t, \theta, \sigma) = P(Z_t = z_t \mid v_t, \theta, \sigma) = \int_{\ell_{t,z_t}}^{u_{t,z_t}} f(\theta_t^* \mid \theta, \sigma) d\theta_t^*.$$

To illustrate using an example, in Figure 1 we use the seat allocation in the Portuguese elections of 1979, where each row in the table corresponds to the seats and percentage of votes received by a party. Note that in this election the smallest party that gained seats received 2.24% of the votes and so $\overline{v}_t = 2.24\%$. Moreover, the number of parties whose vote shares are less than or equal to that upper bound for the threshold is eight ($n_t = 8$). The threshold must then be in one of the intervals defined by the vote shares of these eight parties and the random variable Z_t can take value 1, 2, 3, ... , 8, each corresponding to one of these intervals. The first of these intervals is $(-\infty, 0.06\%]$. The second is $(0.06\%, 0.22\%]$, all the way up to $(1.24\%, 2.24\%]$. The probability calculation in (5) is represented by the shaded area in Figure 1 for the case $z_t = 7$.

| s_t	votes (%)
121	43.71
⋮	⋮
7	2.42
1	2.24
0	1.24
0	0.91
0	0.74
0	0.63
0	0.40
0	0.22
0	0.06
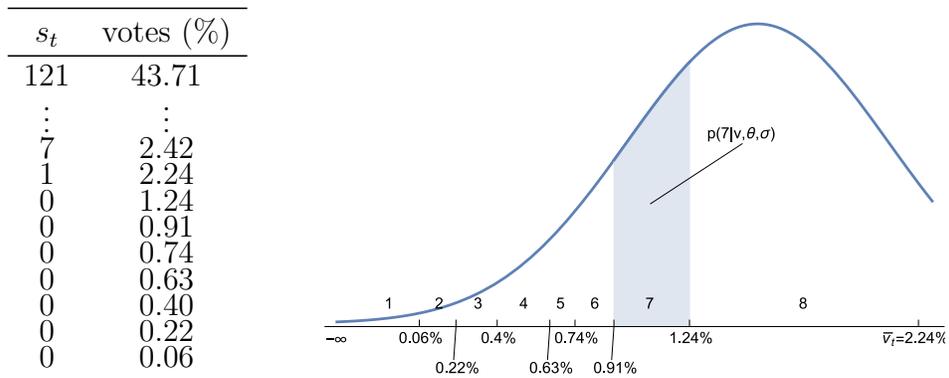
Figure 1: Threshold location (Portugal 1979)

Using (1), (4), and (5), we can now write a log-likelihood as

$$(6) \qquad L(\theta, \sigma, \beta \mid X) = \sum_{t=1}^{T} \log \left(\sum_{z_t=1}^{n_t} p(z_t \mid v_t, \theta, \sigma) p(s_t \mid v_t, \beta, u_{t,z_t}) \right).$$

11

It now becomes evident that, even though the *a priori* probability (5) may suggest otherwise, the estimates of threshold and proportionality parameters θ, σ and β, respectively, interact heavily with each other. Values of θ, σ that place high probability on low realizations of threshold interval Z_t tend to suggest more disproportionality (higher β) because more parties exceeding the threshold receive no seats. Conversely, the value of the proportionality parameter β provides sharper (compared to the logical bound $\bar{v}_t$) information about the likely realization of the random variable Z_t and the threshold. Returning to the 1979 Portuguese elections example of Figure 1, if we were to also observe a very proportional pattern of seat allocation for parties at or above $\bar{v}_t = 2.42\%$ (we actually do not), then we would update an appreciably higher probability that $Z_t = 8$ and the threshold $\theta_t \in (1.24, 2.42]$, for we would otherwise expect a party with 1.24% of the vote to also receive seats.

The log-likelihood in (6) is a first step in obtaining an estimator, with obvious candidates including a classic Maximum Likelihood estimator (ML) or a Bayesian estimator of the MCMC variety. The summations inside the log function suggest that the former may not be the most appropriate plan of attack, and an EM estimator (Dempster, Laird and Rubin (1977)) is better suited for this problem. In addition, the coarse information we sometimes encounter on electoral thresholds (with even a single election per system or otherwise badly conditioned data) suggests a regularization might be of advantage.[9] Therefore, our approach is a happy medium between the frequentist and Bayesian approaches, in the form of a Maximum a Posteriori Expectation Maximization estimator (MAP-EM) (Fraley and Raftery 2007). We regularize the likelihood with respect to the expected threshold parameter θ and the nuisance parameter σ, using the conjugate inverse-gamma-normal prior,

---

[9]In our experience with the data in this study, regularization has the main advantage that it dramatically expedites convergence. In a small fraction of cases it also produces standard errors where the non-regularized likelihood would not using the analytic approach we currently pursue.

as do Fraley and Raftery (2007). Specifically, we assume that $\sigma^2 \sim InverseGamma(\frac{\nu}{2}, \frac{s^2}{2})$ and that conditional on σ, $\theta \sim N(0, \frac{\sigma^2}{\kappa})$. In effect, this amounts to a prior whose log is (excluding the normalizing constant)[10]

$$(7) \qquad p(\theta, \sigma, \beta) = -(\nu + 3) \log(\sigma) - \frac{s^2 + \kappa\theta^2}{2\sigma^2}.$$

It is important to note that we see the prior as just one form of likelihood regularization and we do not commit to a Bayesian perspective in using the MAP-EM estimator. There is to our knowledge only an asymptotic justification of the use of the posterior mode as a Bayesian estimator. This justification is tenuous at best (e.g., Bassett and Deride (2019)), and under the conditions under which it applies, the MAP-EM estimator is also a frequentist estimator with the same asymptotic properties as the ML estimator. Moreover, upon committing on reporting the posterior mode "estimate" the distinction between Bayesian or Frequentist only becomes relevant when we perform inference. Given that we have the analytical machinery to perform all estimation computations without sampling the posterior, we opt for the expediency of frequentist tests (see also Footnote 20).[11] Nevertheless, we explicitly discuss how Bayesian versions of the tests we develop here can be performed with small modifications in Appendix G.

With this regularization in place, we may proceed to specify the EM steps. As is standard, we start by augmenting the data in order to write a (log)likelihood conditional on the data augmented by the (unobserved) component $Z = \{z_t\}_{t=1}^T$, and the unobserved latent

---

[10]As a consequence, we assume a uniform improper prior on the bias parameter β.

[11]Note also that our prior is improper or vague and it only smooths the likelihood coordinates corresponding to the threshold parameters. We introduce it uniformly across applications (though a committed Bayesian is welcome to tailor the prior to her knowledge of specific electoral systems).

variables that determine the election thresholds $\Theta^* = \{\theta_t^*\}_{t=1}^T$. Given z_t, the latent variable θ_t^* is now distributed according to the normal distribution truncated in $(\ell_{t,z_t}, u_{t,z_t}]$:

$$\frac{\mathbb{I}_{(\ell_{t,z_t}, u_{t,z_t}]}(\theta_t^*) f(\theta_t^* \mid \theta, \sigma)}{p(z_t \mid v_t, \theta, \sigma)}.$$

With a bit of algebra, we can now write a log-likelihood for the complete data as

$$(8) \qquad L(\theta, \sigma, \beta \mid X, Z, \Theta^*) \;=\; \sum_{t=1}^T \log(f(\theta_t^* \mid \theta, \sigma)) + \log(p(s_t \mid v_t, \beta, u_{t,z_t})).$$

This constitutes a considerable simplification over (6), as we have now avoided taking logs of any summation terms.

The MAP-EM estimator amounts to an iterative procedure that starts by setting some initial guess for the parameter values $(\theta_0, \sigma_0, \beta_0)$ and at the $m+1$-th iteration computing:

1. Expectation (E-step):

$$Q(\theta, \sigma, \beta; \theta_m, \sigma_m, \beta_m) = \mathbb{E}_{Z,\Theta^*}\left[L(\theta, \sigma, \beta \mid X, Z, \Theta^*) \mid X, \theta_m, \sigma_m, \beta_m\right].$$

2. Maximization (M-step):

$$(\theta_{m+1}, \sigma_{m+1}, \beta_{m+1}) = \arg\max_{(\theta,\sigma,\beta)} \{Q(\theta, \sigma, \beta; \theta_m, \sigma_m, \beta_m) + p(\theta, \sigma, \beta)\}.$$

Details on the computation of the Expectation step appear in online Appendix A. Using the first order conditions, we execute the Maximization step by setting

$$(9) \qquad \theta_{m+1} \;=\; \frac{1}{T+\kappa} \sum_{t=1}^T \sum_{k=n_t}^N \frac{p(s_t \mid v_t, \beta_m, u_{t,z_t})}{h(X_t \mid \theta_m, \sigma_m, \beta_m)} \mathbb{E}_{z_t}[\theta_t^* \mid \theta_m, \sigma_m], \text{ and}$$

14

$$(10) \quad \sigma_{m+1} = \sqrt{\frac{s^2 + \sum_{t=1}^{T} \sum_{k=n_t}^{N} \frac{p(s_t|v_t, \beta_m, u_{t,z_t})}{h(X_t|\theta_m, \sigma_m, \beta_m)} \mathbb{E}_{z_t}[\theta^{*2}_t \mid \theta_m, \sigma_m] - (T+\kappa)\theta^2_{m+1}}{T + \nu + 3}}.$$

It is not possible to solve analytically for β_{m+1}, but due to the separation of the likelihood into terms involving bias β and threshold parameters θ, σ, we can obtain β_{m+1} numerically by solving a single non-linear equation. We efficiently implement these computations, including the derivation of standard errors, at speeds comparable to conventional ML-estimators by taking full advantage of analytical derivations. More details and derivations for this estimator and the computation of standard errors appear in online Appendices A and B.

3 Data and Results

We implement our estimator on electoral returns data from elections to the lower house in 36 European democracies for all elections held after the Second World War, or since democratization, until 2010 (inclusive). We have compiled our electoral returns data from various printed sources including Mackie and Rose (1991); Caramani (2000); Nohlen and Stöver (2010) and online resources such as Alvarez-Rivera (2011), Carr (2011), the Inter-Parliamentary Union (2011), and the Norwegian Social Science Data Services (2011). We relied on official webpages of national parliaments, Election Commissions, each country's Ministry of Interior, etc., to iron out any inconsistencies in these sources and to obtain electoral results for smaller parties, which are usually lumped as "others" in electoral statistics.[12] In addition to electoral returns, using these and additional sources (e.g., Carstairs 1980; Lijphart 1994; Renwick 2010), we also coded the electoral institutions in effect for each election at a fine level of detail. Specifically, for each election, we recorded the number of seats allocated, the different tiers of allocation, the number of districts in each tier, the allo-

---

[12]A complete list of all Internet sources is included in the online Appendix J.

cation formula, the nature of allocation in upper tiers as a function of lower tier allocation, the presence of bonus seat provisions for the (national) plurality or majority party, as well as details for various types of electoral thresholds in effect at each tier of allocation. Details of our institutional coding appear in the online Appendix I.

These institutional data are necessary in order to identify distinct electoral systems in use across time within countries. The main question we have to answer is: What observed changes in electoral institutions are important enough to demarcate a new electoral system that merits estimation of a separate set of electoral system parameters? For the purposes of this analysis we identify changes in electoral systems within countries if either a change in the allocation formula (across tiers) occurs or when another recorded institutional provision (e.g., number of allocated seats, number of districts) changes by more than 5% of that system's average. The 5% cutoff is admittedly arbitrary. One of the advantages of our approach is that we can bring standard statistical inference principles to bear on the question of what constitutes a significant change in electoral institutions. We develop a test procedure for that purpose in section 4.1 and find support (though not universal) for our cutoff rule.

Figure 2 presents the estimated expected national electoral threshold on the left, proportionality estimates on the right, and their standard errors. The figure excludes systems with very large 95% confidence intervals to increase visibility, but a full set of estimates is included in the online Appendix H.[13] The threshold estimates vary significantly with values ranging from 6.81% for the second Moldavian system (MOL2) to virtually zero in systems like GBR1, ITA5–ITA7, and HUN2, among others. The median threshold estimate is 1.58%, which is close to the third system in Cyprus (CYP3 with 1.59%) and the third Portuguese system (PRT3 with 1.56%). There are a number of cases where the presence of thresholds can be rejected statistically, but the point estimate is large (e.g., the German systems DEU4 and DEU8 or the Greek system GRC1). In these cases, a small number of observations does

---

[13]Systems not included in Figure 2: HRV3, LUX2, MKD2, MLT1, MLT2, MLT4, SWE1.

not allow precise estimation.

A sign that our threshold estimates perform quite well is the notable success of the estimator at recovering the level of national legal thresholds when these are unequivocally set by the electoral law. Examples include the second Croatian system (HRV2) and the last two Greek systems GRC5 and GRC6, all with national legal thresholds of 3%. In other cases, the electoral system prescribes a national electoral threshold but qualifying provisions allow small parties to gain seats below the nominal threshold. For example, in most of the German systems, a national 5% threshold applies unless parties win a seat in one of the single member districts of the majoritarian partition. In those cases, the statistical estimates we recover generally indicate an average of the legal threshold attenuated according to the probability that alternative qualifying conditions for representation are met.

Next we turn to the estimates of the electoral proportionality. We see that a majority of systems have estimated proportionality parameters that are greater than one. Overall, there are only 18 systems for which the estimated proportionality parameter is below one, and in many of these cases the difference is in the third decimal point. Among these 18, it is only for the notoriously aberrant second French system (FRA2, with proportionality 0.78) that we can reject the hypothesis that allocation above the threshold is perfectly proportional (in expectation). Conversely, in 56 out of 118 electoral systems, we find the proportionality parameter to be significantly larger than one, that is, we find statistically significant evidence of disproportional allocations favoring larger parties.

The highest proportionality parameters are found in the third Croatian system (HRV3 with 7.83), the first Maltese system (MLT1 with 2.84), and the first Hungarian system (HUN1 with 2.33). Only HRV3 is a plurality system, while MLT1 is a Single Transferable Vote (STV) system with low district magnitude, and HUN1 is an elaborate fusion of majoritarian and PR provisions. Both HRV3 and MLT1 have a proportionality parameter that is consistent

Figure 2: Threshold and Disproportionality Estimates
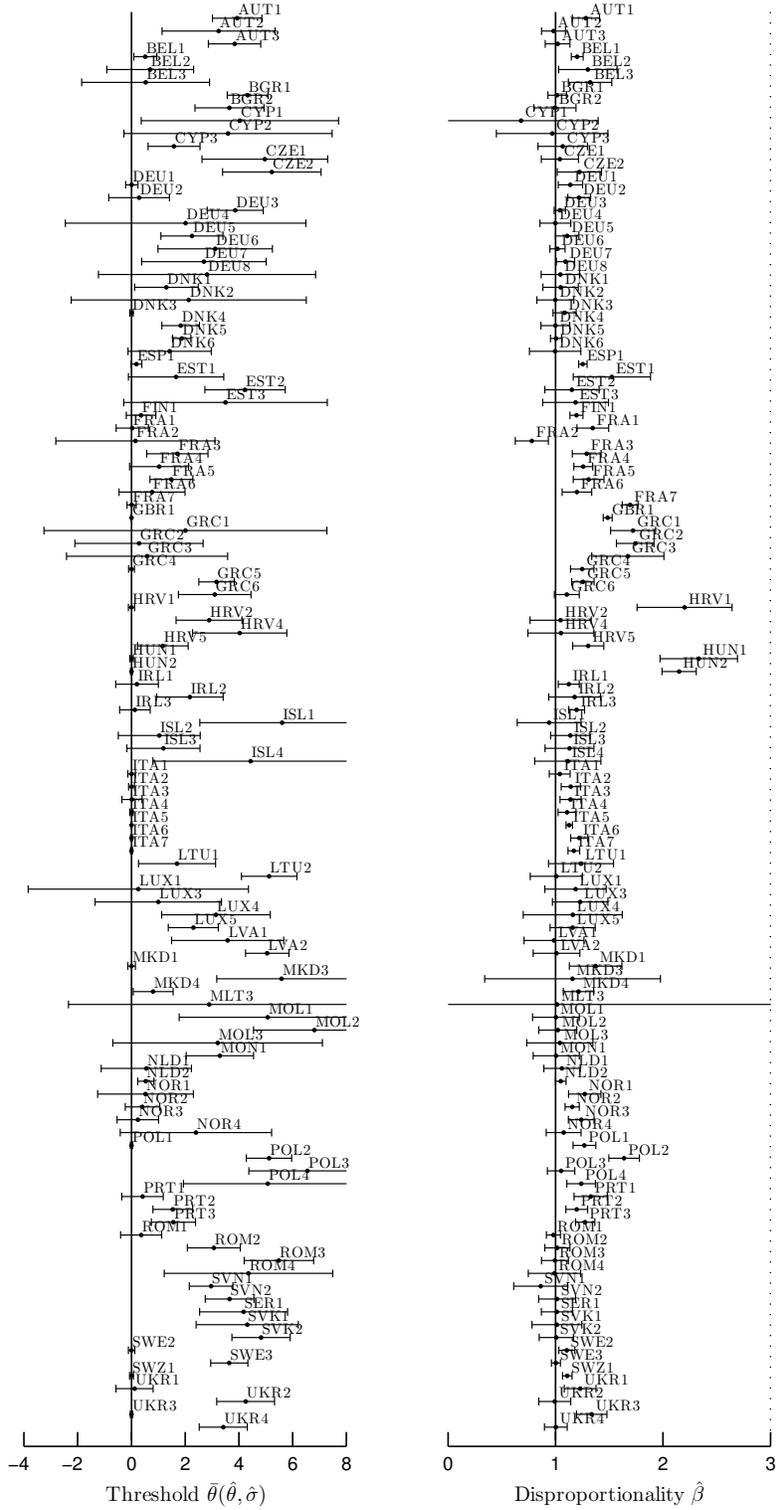
Estimates of thresholds $\bar{\theta}(\hat{\theta}, \hat{\sigma})$ and disproportionality $\hat{\beta}$ along with 95% confidence intervals. Based on point estimates and standard errors reported in the online Appendix H.

Table 1: Thresholds and Disproportionality (Counts)

	Non-proportional	Proportional	Total
Threshold	19	45	64
	[16.10]	[38.14]	[54.24]
No-Threshold	37	17	54
	[31.36]	[14.41]	[45.76]
Total	56	62	118
	[47.46]	[52.54]	[100]

Threshold denotes the count of systems for which the expected threshold is larger than zero at the 5% level of significance (one-tailed test) and No Threshold all other systems. Non-Proportional denotes the count of systems for which the disproportionality estimate is larger than one at the 5% level of significance (one-tailed test) and Proportional denotes all other systems. Percentages in brackets.

with the "Cube Law,"[14] that is, these are the two cases when we fail to reject the null of β being equal to three (3).[15] The median proportionality parameter across all systems is 1.13 and the next largest point estimates following the above three are all statistically different from 3. Even among majoritarian systems, the Cube Law finds little empirical support.[16]

Table 1 presents an aggregated summary of the results. We find 19 systems for which there are thresholds statistically larger than zero that also have disproportionality parameters larger than one. These are the systems that provide the most advantages for larger parties. Among them, the ones with larger disproportionality parameters are the second Polish system, the first Estonian, and the fifth French system (POL2, EST1, and FRA5).

---

[14]For the literature on the cube law and empirical refutations see Kendall and Stuart (1950); Schrodt (1981); Taagepera (1986).

[15]Other cases for which that null is not rejected have point estimates closer to one, but large standard errors.

[16]The systems or partitions in which all districts use majoritarian or plurality formulas are: FRA3-FRA5, FRA7, GBR1, HRV1, HRV3, ITA6, MKD1, MKD2, UKR1, and UKR3.

POL2 also has one of the highest thresholds (5.12%) and only FRA5 uses a majoritarian allocation formula in that group. On the other extreme, there are 17 systems that offer the most opportunities to small parties with perfect proportionality and no thresholds. The third Danish and first Italian systems —DNK3 and ITA1, both PR systems with some form of compensatory upper tier—fall into that category. But the most populated categories are in the off-diagonal entries of Table 1, which include systems that provide for (statistically significant) thresholds but proportional allocation for parties exceeding the threshold (45 out of 118 systems) and systems that have no statistically significant threshold but dispropor-tionality favoring larger parties (37 out of 118 systems).

4   Inference

We perform all inferences using the conventional approach that relies on model predic-tions at the point estimates and not the entire posterior. As discussed in the first paragraph after equation (7), a properly Bayesian execution of these tests is also possible and the reader can consult Appendix G for details. We start by discussing methods to evaluate the fit of the model. First, we devise a test of the hypothesis that the data are generated according to the estimated model. The test is based on a Pearson chi-square type of statistic as a weighted sum of squared deviations between actual and model predicted seat allocations,

$$
(11) \qquad P := \sum_{t=1}^{T} \sum_{i=1}^{N} \frac{(s_{i,t} - \overline{s}_{i,t}(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta}))^2}{\overline{s}_{i,t}(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta})},
$$

where $\overline{s}_{i,t}(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta}) := S_t \sum_{z_t=1}^{N} q_{i,t}(v_t, \widehat{\beta}, u_{t,z_t}) p(z_t \mid v_t, \widehat{\theta}, \widehat{\sigma})$ denotes party i's expected seats according to the estimated model. The intuition of the test is standard: If the observed allocation is close to the model's prediction, then the value of the statistic will be small, whereas large values of the statistic would cast doubt on the ability of the model to account

for the variation in the data. We compute empirical p-values of the test via Monte Carlo simulations.[17] We find that the test rejects the null of the observed seat allocation being consistent with the model's prediction for only 5 systems out of 118 (FRA2-4, FRA6, and MKD2) at the 5% significance level.[18] These results suggest that the model performs well accounting for variation in the data for the bulk of the estimated electoral systems.

We also assess the performance of the model relative to one that restricts the national threshold to be zero, a common assumption in the literature heretofore. We call the model with no threshold the *restricted model*.[19] As an initial comparison, we perform a goodness of fit test for the restricted model that is analogous to the one described above. This time, the test rejects the null that the observed seats data are generated from the restricted model in 18 systems, a larger fraction relative to the case of the model allowing thresholds. A more systematic comparison is given by a statistical test of differences of fit between the two models. The test statistic is simply $P^o - P$, where P and P^o take the form used in the above model fit tests and seat allocations under the null arise from the restricted model. Under that null hypothesis, large values of the statistic provide evidence in favor of the alternative model with thresholds. Consistent with findings so far, the empirical p-value of the test is below the 5% significance level in 51 systems.[20] As expected, we earlier found significant positive thresholds in 45 of those 51 systems .

Figure 3 presents the proportionality estimates of the full and restricted models for

---

[17]Full details on this and subsequent test procedures appear in the online Appendix G.

[18]See p-values of test in the online Appendix H.

[19]Online Appendix C describes this model and its estimation.

[20]The online Appendix H presents the p-values of the test for all systems. Note that in the restricted model the MAP-EM estimator is equivalent to the ML estimator and therefore, under the null that data are generated from the restricted model, both the test of model fit and the model comparison test exactly match the ML frequentist methodology in the extant literature studying the restricted model.

Figure 3: MAP-EM Proportionality vs. Restricted Model Proportionality
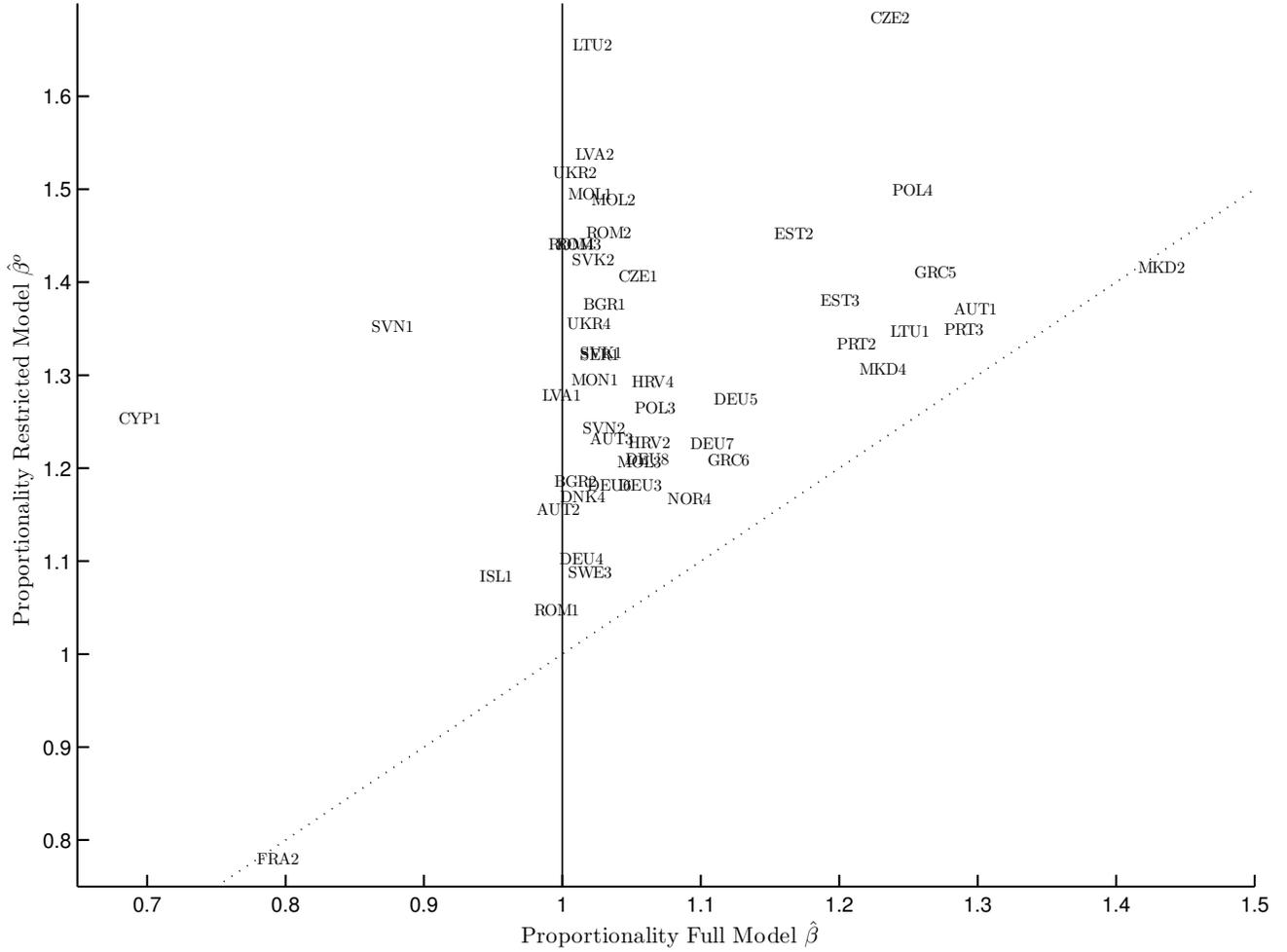
Illustrates overestimation of disproportionality in the restricted model that ignores thresholds, compared to estimates obtained from the full model with thresholds. Based on results reported in the online Appendix H.

those systems for which the test indicates significant differences in fit across the models.[21] If a system falls on the dotted line, this indicates that the unrestricted model's disproportionality estimate is the same as that of the restricted model. As expected, ignoring the threshold brings an overestimation of disproportionality, indicated by the fact that most systems in the figure are above the dotted line. In particular, there is a group of systems that are close to being perfectly proportional for parties above the threshold for which the restricted model would overstate the advantages given to large parties—those near the vertical line at one (e.g., ROM1, SWE3, AUT3, DNK4). Overall, we see that the full model does a good job accounting for the variation in the data. Even though estimating the model with thresholds presents additional technical challenges compared to the simpler restricted model used in the literature, these tests suggest that accounting for the threshold is warranted both statistically, and certainly politically, for a large fraction of systems.

4.1 Electoral System Change

We have conducted our analysis assuming an electoral system persists if i) there are no changes in the allocation formula and ii) the numerical legal thresholds, number of districts, or number of seats remain within 5% of that system's average. While practical, this definition is arbitrary and highlights a more general problem in the literature, namely, the lack of a systematic way to identify whether formal changes in the electoral law—or, even changes not codified in the electoral law (for example, new districting practices)—translate to substantial changes in the resulting pattern of seat allocation. We build on the inferential approach developed in the previous subsection to propose a statistical test to detect changes

---

[21]POL2, which has an estimated MAP-EM proportionality of 1.64 and 2.06 with the restricted model, is excluded to increase visibility of other systems in the figure. The cluster of systems near (1.01,1.32) are SVK1 and SER1. Those near (1,1.44) are ROM3 and ROM4.

Table 2: Statistically Significant Electoral System Changes

Panel A: Coarse vs. 5% cutoff definition			*Panel B*: 5% cutoff vs. finest definition		
System	Years	p-values	System	Years	p-values
CYP1-2	1981-1991	0.275	BEL1	1946-1991	0.25
DEU4-5	1987-1990	0.268	DEU3	1957-1983	0.094
DEU6-7	1994-2005	0.306	FRA1	1945-1946	0.283
EST1-2	1992-1999	0.024	FRA4	1967-1973	0.055
FRA3-5,7	1958-1981,1988-2007	0	FRA7	1988-2007	0.058
HRV1,3[a]	1992,1995	0.006	GBR1	1945-2010	0
HRV2,4[b]	1992,1995	0.171	IRL1	1948-1965	0.245
HUN1-2	1990-2010	0.542	IRL2	1969-1977	0.321
IRL1-3	1948-2007	0.032	IRL3	1981-2007	0.19
ISL1-2	1946-1983	0.025	LUX1	1945,1954-1959	0.237
ITA4-5	1958-1992	0.373	LUX3	1964-1979	0.407
LTU1-2	1992-2008	0	MLT1	1966-1971	0.29
LUX1-6	1945-2009	0.053	MOL1	1994-2010	0.245
LVA1-2	1993-2010	0.003	NOR1	1945-1949	0.395
MLT1-2	1966-1981	0.003	NOR2	1953-1985	0.055
MOL1-3	1994-2010	0.428	PRT1	1975-1976	0.018
NLD1-2	1946-2010	0.595	ROM2	1992-1996	0.389
NOR3-4	1989-2009	0.106	ROM3	2000-2004	0.319
PRT1-3	1975-2009	0	SVK2	1998-2010	0.424
ROM2-4	1992-2008	0.04	SWE2	1952-1968	0.368
UKR2,4[c]	1998,2002,2006-2007	0.175	SWE3	1970-2010	0.521
			SWZ1	1947-2007	0.063

Reports tests of change in electoral systems comparing two pairs of alternative *ad hoc* definitions. Panel A presents results of the comparison between our default 5% cutoff definition and a coarser definition. Panel B presents results of the comparison between our default 5% cutoff definition and a finer definition of electoral systems. In each case, the null hypothesis is that the coarser definition is correct (no change within the corresponding period) and small *p*-values indicate support for the finer definition. 'System' reports systems involved in the test according to our default definition. 'Years' reports the years covered by the test.
a SMD partition of mixed system.
b PR partition of mixed system.
c years 1998 and 2002 PR partition.

in electoral systems. This test is meant to provide an objective way to evaluate whether hypothesized changes in electoral systems are statistically significant or not.

In a nutshell, the test compares the fit of the estimated model under a coarser definition with that of the model when two or more sets of parameters are estimated using a finer definition which partitions the elections included in estimating the coarser system parameters. The test statistic is of a similar pedigree as previous tests in this section computing sums of squared deviations between actual and predicted seats for the coarse and finer definitions for all T elections covered by the coarser definition and taking the difference of the two. Under the null hypothesis of no change, the resulting test statistic should be small and we reject the null when the mass of the distribution of the statistic under the null that exceeds the computed value is less than the chosen significance level.

To illustrate this test, we perform two sets of comparisons between our default definition of electoral system and possible alternatives. One alternative is *coarser* and identifies a new electoral system only when the allocation formula changes (that is, ignoring any changes in seats, number of districts, or legal thresholds). The second alternative is *finer* and identifies a new electoral system whenever *any* change occurs in recorded institutions or when a change in electoral law not codified in our institutional variables is reported in our sources. When comparing the coarser definition against our default 5% cutoff, we reject the null of no change for 10 of 21 of the tests at the 5% significance level, that is, we find evidence against the coarser definition in 10 out of 21 of the cases when the two definitions classify elections into different systems.[22] When comparing our default definition with the finer definition, we reject the null in 2 out of 22 tests at the 5% level of significance. These tests are reported in Table 2. Though these comparisons are not meant to be definitive in this context, they suggest that if one has to rely on a uniform rule applied across countries, our 5% cutoff

---

[22]The null is rejected in only 35% of the tests comparing the coarsest with the finest definitions.

rule appears like a reasonable compromise. But our test procedure provides a data-driven alternative that does not rely on the application of such a uniform rule.

5   Conclusions

We have developed a new statistical model of the translation of votes into seats and obtained estimates of national electoral thresholds of representation and of disproportionality of seat allocation for parties exceeding thresholds. These measures quantify conceptually distinct and politically relevant dimensions of the electoral system, they are comparable across systems and time, provide a natural and intelligible summary of the electoral system, and come with a gauge of uncertainty in the confidence we can place in them in the form of standard errors. We also developed a battery of inference procedures tailored to this model that allow us to evaluate model fit, compare models, and evaluate significant changes in electoral provisions over time.

The introduction of threshold parameters is empirically necessary in order to fit the data when we suspect large national electoral thresholds to be present. For nearly half the electoral systems in our study, we can reject the hypothesis of zero national threshold of representation and in the bulk of these cases, the model that ignores thresholds results in a poorer fit, according to our rigorous fit-comparison test. We also identified that the modal electoral institutional arrangements in European democracies after World War II is to either adopt significant thresholds of representation along with perfectly proportional allocation for parties that exceed these thresholds or to effectively impose negligible thresholds of representation but provide substantial seat bonuses to larger parties.

References

Alvarez-Rivera, Manuel. 2011. "Election Resources on the Internet." http://electionresources.org/.

Bassett, Robert and Julio Deride. 2019. "Maximum a posteriori estimators as a limit of Bayes estimators." *Mathematical Programming, ser. B* 174:129–144.

Calvo, Ernesto. 2009. "The Competitive Road to Proportional Representation: Partisan Biases and Electoral Regime Change under Increasing Party Competition." *World Politics* 61(2):254–295.

Calvo, Ernesto and Jonathan Rodden. 2015. "The Achilles Heel of Plurality Systems: Geography and Representation in Multiparty Democracies." *American Journal of Political Science* 59(4):789–805.

Caramani, Daniele. 2000. *Elections in Western Europe since 1815: Electoral Results by Constituencies.* Cambridge: Cambridge University Press.

Carr, Adam. 2011. "Psephos Adam Carr's Election Archive." http://psephos.adam-carr.net/.

Carstairs, Andrew McLaren. 1980. *A Short History of Electoral Systems in Western Europe.* London: Allen & Unwin.

Cox, Gary. 1997. *Making Votes Count.* Cambridge: Cambridge University Press.

Cox, Gary and Mathew Soberg Shugart. 1991. "Comment on Gallagher's 'Proportionality, Disproportionality, and Electoral Systems'." *Electoral Studies* 10(4):348–352.

Dempster, A. P., N. M. Laird and D. B. Rubin. 1977. "Maximum Likelihood from Incomplete Data via the EM Algorithm." *Journal of the Royal Statistical Society. Series B (Methodological)* 39(1):1–38.

European Commission for Democracy through law (Venice Commission). 2018. Compilation of Venice Commission Opinions and Reports Concerning Thresholds that bar Parties from Access to Parliament. 004 Council of Europe https://www.venice.coe.int/webforms/documents/?pdf=CDL-PI(2018)004-e: .

Fraley, Chris and Adrian Raftery. 2007. "Bayesian Regularization for Normal Mixture Estimation and Model-Based Clustering." *Journal of Classification* 24:155–181.

Gallagher, Michael. 1991. "Proportionality, Disproportionality, and Electoral Systems." *Electoral Studies* 10(1):33–51.

Gallagher, Michael. 1992. "Comparing Proportional Representation Electoral Systems: Quotas, Thresholds, Paradoxes and Majorities." *British Journal of Political Science* 22(4):469–496.

Grofman, Bernard. 1983. "Measures of Bias and Proportionality in Seats-Votes Relationships." *Political Methodology* 9:295–327.

Inter-Parliamentary Union. 2011. "PARLINE Database on National Parliaments." http://www.ipu.org/parline-e/parlinesearch.asp.

Jackman, Simon. 1994. "Measuring Electoral Bias: Australia 1949-93." *British Journal of Political Science* 24(3):319–357.

Kendall, M. G. and A. Stuart. 1950. "The Law of the Cubic Proportion in Election Results." *The British Journal of Sociology* 1(3):183–196.

King, Gary. 1990. "Electoral Responsiveness and Partisan Bias in Multiparty Democracies." *Legislative Studies Quarterly* 15(2):159–181.

King, Gary and Robert X. Browning. 1987. "Democratic Representation and Partisan Bias in Congressional Elections." *American Political Science Review* 81(4):1251–1273.

Levitan, Emanuel and Gabor T. Herman. 1987. "A Maximum A Posteriori Probability Expectation Maximization Algorithm for Image Reconstruction in Emission Tomography." *IEEE Transactions on Medical Imaging* MI-6(3):185–192.

Lijphart, Arend. 1994. *Electoral Systems and Party Systems: A Study of Twenty-Seven Democracies, 1945-1990.* Oxford: Oxford University Press.

Linzer, Drew A. 2012. "The Relationship between Seats and Votes in Multiparty Systems." *Political Analysis* 20(3):400.

Mackie, Thomas and Richard Rose. 1991. *The International Almanac of Electoral History.* London: Macmillan.

Nohlen, Dieter and Philip Stöver. 2010. *Elections in Europe: A Data Handbook.* Berlin: Nomos Verlagsgesellschaft.

Norwegian Social Science Data Services. 2011. "European Election Database." http://www.nsd.uib.no/european_election_database/.

Renwick, Alan. 2010. *The Politics of Electoral Reform.* Cambridge: Cambridge University Press.

Ruiz-Rufino, Rubén. 2007. "Aggregated Threshold Functions or How to Measure the Performance of an Electoral System." *Electoral Studies* 26:492–502.

Schrodt, Philip A. 1981. "A Statistical Study of the Cube Law in Five Electoral Systems." *Political Methodology* 7(2):31–53.

Taagepera, Rein. 1973. "Seats and Votes: A Generalization of the Cube Law of Elections." *Social Science Research* 2:257–75.

Taagepera, Rein. 1986. "Reformulating the Cube Law for Proportional Representation Elections." *American Political Science Review* 80(2):489–504.

Taagepera, Rein. 1989. "Empirical Threshold of Representation." *Electoral Studies* 8(2):105–116.

Taagepera, Rein. 1998*a*. "Effective Magnitude and Effective Threshold." *Electoral Studies* 17(4):393–404.

Taagepera, Rein. 1998*b*. "Nationwide Inclusion and Exclusion Thresholds of Representation." *Electoral Studies* 17(4):405–417.

Taagepera, Rein. 2002. "Nationwide Threshold of Representation." *Electoral Studies* 21(3):383–401.

Taagepera, Rein and Mathew Soberg Shugart. 1989. *Seats and Votes: The Effects and Determinants of Electoral Systems.* New Haven: Yale University Press.

Taagepera, Rein and Mathew Soberg Shugart. 2017. *Votes from Seats: Logical Models of Electoral Systems.* Cambridge: Cambridge University Press University Press.

Theil, Henri. 1969. "The Desired Political Entropy." *American Political Science Review* 63(2):521–525.

Tufte, Edward. 1973. "The Relationship Between Seats and Votes in Two-Party Systems." *American Political Science Review* 67(2):540–554.

National Electoral Thresholds and Disproportionality
Online Appendix (not intended for publication)

Tasos Kalandrakis[*]   Miguel R. Rueda[†]

July 25, 2019

[*]Department of Political Science and Department of Economics, University of Rochester. E-mail: kalandrakis@rochester.edu.
[†]Department of Political Science, Emory University.

A   MAP-EM Estimator

With complete data including the unobserved $Z = \{z_t\}_{t=1}^T$ and $\Theta^* = \{\theta_t^*\}_{t=1}^T$, the data-augmented log-likelihood is given by (8) of the main text. The $m + 1$-th MAP-EM estimator steps consist of:

1. Expectation (E-step):

$$Q(\theta, \sigma, \beta; \theta_m, \sigma_m, \beta_m) = \mathbb{E}_{Z, \Theta^*} \left[L(\theta, \sigma, \beta \mid X, Z, \Theta^*) \mid X, \theta_m, \sigma_m, \beta_m \right].$$

2. Maximization (M-step):

$$(\theta_{m+1}, \sigma_{m+1}, \beta_{m+1}) = \arg \max_{(\theta, \sigma, \beta)} \{ Q(\theta, \sigma, \beta; \theta_m, \sigma_m, \beta_m) + p(\theta, \sigma, \beta) \}.$$

Recall that the prior $p(\theta, \sigma, \beta)$ is given by (7) in the main text. To execute the Expectation step, note that the probability of z_t conditional on $X, \theta_m, \sigma_m, \beta_m$ takes the form

$$\mathbb{P}(z_t \mid X, \theta_m, \sigma_m, \beta_m) = \frac{p(z_t \mid v_t, \theta_m, \sigma_m) p(s_t \mid v_t, \beta_m, u_{t,z_t})}{\sum_{z_t'=1}^{n_t} p(z_t' \mid v_t, \theta_m, \sigma_m) p(s_t \mid v_t, \beta_m, u_{t,z_t'})},$$

by an application of Bayes' rule. Furthermore,

$$\mathbb{P}(\theta_t^* \mid X, Z, \theta_m, \sigma_m, \beta_m) = \frac{\mathbb{I}_{(\ell_{t,z_t}, u_{t,z_t}]}(\theta_t^*) f(\theta_t^* \mid \theta, \sigma)}{p(z_t \mid v_t, \theta, \sigma)}.$$

We can then compute $Q(\theta, \sigma, \beta; \theta_m, \sigma_m, \beta_m)$ as follows:

$$Q(\theta, \sigma, \beta; \theta_m, \sigma_m, \beta_m) = \sum_{t=1}^{T} \sum_{z_t=1}^{n_t} \frac{p(s_t \mid v_t, \beta_m, u_{t,z_t})}{h(X_t \mid \theta_m, \sigma_m, \beta_m)} \int_{\ell_{t,z_t}}^{u_{t,z_t}} \log(p(s_t \mid v_t, \beta, u_{t,z_t})) f(\theta_t^* \mid \theta_m, \sigma_m) d\theta_t^*$$

$$+ \sum_{t=1}^{T} \sum_{z_t=1}^{n_t} \frac{p(s_t \mid v_t, \beta_m, u_{t,z_t})}{h(X_t \mid \theta_m, \sigma_m, \beta_m)} \int_{\ell_{t,z_t}}^{u_{t,z_t}} \log(f(\theta_t^* \mid \theta, \sigma)) f(\theta_t^* \mid \theta_m, \sigma_m) d\theta_t^*$$

$$= \sum_{t=1}^{T} \sum_{z_t=1}^{n_t} \frac{p(s_t \mid v_t, \beta_m, u_{t,z_t})}{h(X_t \mid \theta_m, \sigma_m, \beta_m)} \mathbb{E}_{z_t}[1 \mid \theta_m, \sigma_m] \sum_{i=z_t}^{N} s_{t,i} \log \left(\frac{v_{t,i}^{\beta}}{\sum_{j=z_t}^{N} v_{t,j}^{\beta}} \right)$$

$$+ \sum_{t=1}^{T} \sum_{z_t=1}^{n_t} \frac{p(s_t \mid v_t, \beta_m, u_{t,z_t})}{h(X_t \mid \theta_m, \sigma_m, \beta_m)} \frac{\theta}{\sigma^2} \mathbb{E}_{z_t}[\theta_t^* \mid \theta_m, \sigma_m]$$

$$- \sum_{t=1}^{T} \sum_{z_t=1}^{n_t} \frac{p(s_t \mid v_t, \beta_m, u_{t,z_t})}{h(X_t \mid \theta_m, \sigma_m, \beta_m)} \frac{1}{2\sigma^2} \mathbb{E}_{z_t}[\theta_t^{*2} \mid \theta_m, \sigma_m]$$

$$- T \left(\log(\sqrt{2\pi}\sigma) + \frac{\theta^2}{2\sigma^2} \right),$$

where $h(X_t \mid \theta_m, \sigma_m, \beta_m) = \sum_{z_t'=1}^{n_t} p(z_t' \mid v_t, \theta_m, \sigma_m) p(s_t \mid v_t, \beta_m, u_{t,z_t'})$, $X_t = (s_t, v_t)$, and with the expectation terms taking the form

(1)
$$\mathbb{E}_{z_t}[\theta_t^{*p} \mid \theta_m, \sigma_m] := \int_{\ell_{t,z_t}}^{u_{t,z_t}} \theta_t^{*p} f(\theta_t^* \mid \theta_m, \sigma_m) d\theta_t^*, p = 0, 1, 2.$$

These are available in closed form (see online Appendix E).

To execute the Maximization step, we compute the first order conditions for a maximum, taking first partial derivatives:

$$\frac{\partial Q(\theta, \sigma, \beta; \theta_m, \sigma_m, \beta_m) + p(\theta, \sigma, \beta)}{\partial \theta} = \sum_{t=1}^{T} \sum_{z_t=1}^{n_t} \frac{p(s_t \mid v_t, \beta_m, u_{t,z_t})}{h(X_t \mid \theta_m, \sigma_m, \beta_m)} \frac{1}{\sigma^2} \mathbb{E}_{z_t}[\theta_t^* \mid \theta_m, \sigma_m] - \frac{\theta(\kappa + T)}{\sigma^2},$$

$$
\frac{\partial Q(\theta,\sigma,\beta;\theta_m,\sigma_m,\beta_m) + p(\theta,\sigma,\beta)}{\partial \sigma} = -\sum_{t=1}^{T}\sum_{z_t=1}^{n_t} \frac{p(s_t \mid v_t, \beta_m, u_{t,z_t})}{h(X_t \mid \theta_m, \sigma_m, \beta_m)} \frac{2\theta}{\sigma^3} \mathbb{E}_{z_t}[\theta_t^* \mid \theta_m, \sigma_m]
$$
$$
+ \sum_{t=1}^{T}\sum_{z_t=1}^{n_t} \frac{p(s_t \mid v_t, \beta_m, u_{t,z_t})}{h(X_t \mid \theta_m, \sigma_m, \beta_m)} \frac{1}{\sigma^3} \mathbb{E}_{z_t}[\theta_t^{*2} \mid \theta_m, \sigma_m]
$$
$$
- \frac{T+\nu+3}{\sigma} + \frac{\theta^2(T+\kappa)+s^2}{\sigma^3},
$$

and

(2)
$$
\frac{\partial Q(\theta,\sigma,\beta;\theta_m,\sigma_m,\beta_m) + p(\theta,\sigma,\beta)}{\partial \beta} = \sum_{t=1}^{T}\sum_{z_t=1}^{n_t} \frac{p(s_t \mid v_t, \beta_m, u_{t,z_t})}{h(X_t \mid \theta_m, \sigma_m, \beta_m)} \mathbb{E}_{z_t}[1 \mid \theta_m, \sigma_m] \times
$$
$$
\sum_{i=z_t}^{N} s_{t,i} \left(\log(v_{t,i}) - \frac{\sum_{j=z_t}^{N} v_{t,j}^{\beta} \log(v_{t,j})}{\sum_{j=z_t}^{N} v_{t,j}^{\beta}} \right).
$$

Therefore, the updated iterates θ_{m+1} and σ_{m+1} are obtained as in equations (9) and (10) of the main text by solving the corresponding first order conditions. It is not possible to solve analytically for β_{m+1}, and we obtain it numerically by solving a single non-linear equation setting (2) to zero. We use Newton's method for that purpose and the cross partial second derivative used is given by:

$$
\frac{\partial^2 Q(\theta,\sigma,\beta;\theta_m,\sigma_m,\beta_m)}{\partial \beta \partial \beta} = -\sum_{t=1}^{T}\sum_{z_t=1}^{n_t} \frac{p(s_t \mid v_t, \beta_m, u_{t,z_t})}{h(X_t \mid \theta_m, \sigma_m, \beta_m)} \mathbb{E}_{z_t}[1 \mid \theta_m, \sigma_m] \sum_{i=z_t}^{N} s_{t,i} \times
$$
$$
\left[\frac{\sum_{j=z_t}^{N} v_{t,j}^{\beta} \log(v_{t,j})^2}{\sum_{j=z_t}^{N} v_{t,j}^{\beta}} - \left(\frac{\sum_{j=z_t}^{N} v_{t,j}^{\beta} \log(v_{t,j})}{\sum_{j=z_t}^{N} v_{t,j}^{\beta}} \right)^2 \right].
$$

Finally, before we move to estimation we must specify values for the prior parameters ν, s^2, and κ, and a termination criterion for the EM iterations. With regard to the priors, we opt

3

for a vague prior for σ by setting $\nu = s^2 = 2$. This amounts to an $InverseGamma(1,1)$ distribution and a prior mode at $\frac{1}{2}$ for σ^2. This choice bounds our estimates of the nuisance parameter away from zero, but does not unduly restrict that estimate. We also set the prior parameter $\kappa = \frac{1}{100}$, which is the value chosen by Fraley and Raftery (2007) in their implementation. We monitor convergence by measuring the distance between successive iterates and terminate the algorithm when this distance is less than 10^{-9}. To safeguard against isolating a local maximizer, we initiate the algorithm from different starting values.

B Standard Errors

Standard errors are calculated using the Hessian of the posterior evaluated at the point estimates, which is denoted by $H(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta})$. We compute this matrix using a result from Dempster, Laird and Rubin (1977, p.10) (see also Jamshidian and Jennrich (2000)) for the EM context and applied in the MAP-EM case

$$H(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta}) = (\ddot{Q}(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta}; \widehat{\theta}, \widehat{\sigma}, \widehat{\beta}) + \ddot{p}(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta})) \left(I_3 - \dot{M}(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta}) \right).$$

Here, $\ddot{Q}(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta}; \widehat{\theta}, \widehat{\sigma}, \widehat{\beta}) + \ddot{p}(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta})$ is a 3×3 matrix of second derivatives of $Q + p$ with respect to the components of its first three arguments. Also,

$$M(\theta_m, \sigma_m, \beta_m) := \arg \max_{\theta, \sigma, \beta} \{Q(\theta, \sigma, \beta; \theta_m, \sigma_m, \beta_m) + p(\theta, \sigma, \beta)\},$$

and $\dot{M}(\theta_m, \sigma_m, \beta_m)$ is the Jacobian with respect to θ_m, σ_m, and β_m. The first coordinate of $M(\theta_m, \sigma_m, \beta_m)$ is given by (9) and the second by (10) in the main text. Since the third coordinate is not available in closed form and is obtained by solving (2) numerically, its derivatives with respect to $x \in \{\theta_m, \sigma_m, \beta_m\}$ are computed using the Implicit Function theorem as follows:

4

$$-\frac{\frac{\partial^2 Q(\theta,\sigma,\beta;\theta_m,\sigma_m,\beta_m)+p(\theta,\sigma,\beta)}{\partial x \partial \beta}}{\frac{\partial^2 Q(\theta,\sigma,\beta;\theta_m,\sigma_m,\beta_m)+p(\theta,\sigma,\beta)}{\partial \beta \partial \beta}}.$$

Thus, all the necessary partial derivatives needed to calculate the Jacobian of $M(\theta_m, \sigma_m, \beta_m)$ (and the Hessian $H(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta})$) are obtained analytically (see online Appendix F for detailed derivations).

C   Estimation of Restricted Model

Without parameters θ, σ and since we have assumed a flat improper prior for β, a MAP-EM estimator of the no-threshold model is equivalent to an EM-estimator which in turn is equivalent to a classic ML estimator. The log-likelihood of the restricted model is

(3)
$$L(\beta \mid X) = \sum_{t=1}^{T} \log\left(p(s_t \mid v_t, \beta, 0)\right).$$

The first order condition for the maximization of the log-likelihood is

$$\sum_{t=1}^{T}\sum_{i=1}^{N} s_{t,i} \left(\log(v_{t,i}) - \frac{\sum_{j=1}^{N} v_{t,j}^{\beta} \log(v_{t,j})}{\sum_{j=1}^{N} v_{t,j}^{\beta}} \right) = 0.$$

We thus obtain the proportionality parameter using Newton's algorithm exploitng the fact that

$$\frac{\partial^2 L(\beta \mid X)}{\partial \beta \partial \beta} = -\sum_{t=1}^{T}\sum_{i=1}^{N} s_{t,i} \left[\frac{\sum_{j=1}^{N} v_{t,j}^{\beta} \log(v_{t,j})^2}{\sum_{j=1}^{N} v_{t,j}^{\beta}} - \left(\frac{\sum_{j=1}^{N} v_{t,j}^{\beta} \log(v_{t,j})}{\sum_{j=1}^{N} v_{t,j}^{\beta}} \right)^2 \right].$$

D   Expected Threshold and its Variance

An analytical expression for the mean of the threshold is:

$$\bar{\theta}(\theta, \sigma) := \mathbb{E}[\theta_t | \theta, \sigma] \quad = \quad P\left(\theta_t^* \geq 0 \mid \theta, \sigma\right) \mathbb{E}[\theta_t^* | \theta_t^* \geq 0, \theta, \sigma]$$

$$= \quad \left(1 - \Phi\left(-\tfrac{\theta}{\sigma}\right)\right)\theta + \sigma\phi\left(-\tfrac{\theta}{\sigma}\right).$$

To derive an expression for its standard deviation, we first compute the expectation of the squared threshold

$$\mathbb{E}[\theta_t^2 | \theta, \sigma] \quad = \quad P\left(\theta_t^* \geq 0 \mid \theta, \sigma\right) \mathbb{E}[\theta_t^{*2} | \theta_t^* \geq 0, \theta, \sigma]$$

$$= \quad \left(1 - \Phi\left(-\tfrac{\theta}{\sigma}\right)\right)(\sigma^2 + \theta^2) + \theta\sigma\phi(-\tfrac{\theta}{\sigma}),$$

to obtain

$$\bar{\sigma}(\theta, \sigma) \quad = \quad \sqrt{\left(1 - \Phi\left(-\tfrac{\theta}{\sigma}\right)\right)(\sigma^2 + \theta^2) + \theta\sigma\phi(-\tfrac{\theta}{\sigma}) - \left(\left(1 - \Phi\left(-\tfrac{\theta}{\sigma}\right)\right)\theta + \sigma\phi\left(-\tfrac{\theta}{\sigma}\right)\right)^2}$$

$$= \quad \sqrt{\left(1 - \Phi\left(-\tfrac{\theta}{\sigma}\right)\right)(\sigma^2 + \theta^2) + \theta\sigma\phi(-\tfrac{\theta}{\sigma})\left(2\Phi(-\tfrac{\theta}{\sigma}) - 1\right) - \left(1 - \Phi\left(-\tfrac{\theta}{\sigma}\right)\right)^2\theta^2 - \sigma^2\phi(-\tfrac{\theta}{\sigma})^2}.$$

Both $\bar{\theta}(\theta, \sigma)$ and $\bar{\sigma}(\theta, \sigma)$ are continuously differentiable functions of (θ, σ, β) which are asymptotically normally distributed with variance $-\mathbb{E}[H(\theta, \sigma, \beta)]^{-1}$. We may thus apply the Delta method to obtain estimates of the asymptotic variances:

$$(4) \qquad\qquad \widehat{\bar{\theta}(\theta, \sigma)} \overset{a}{\sim} N\left(\bar{\theta}(\theta, \sigma), J_{\bar{\theta}(\theta, \sigma)} V_{\theta, \sigma} J'_{\bar{\theta}(\theta, \sigma)}\right),$$

where $J_{\bar{\theta}(\theta, \sigma)} = \left(\frac{\partial \bar{\theta}(\theta, \sigma)}{\partial \theta}, \frac{\partial \bar{\theta}(\theta, \sigma)}{\partial \sigma}\right)$ and $V_{\theta, \sigma}$ denotes the matrix formed by the first two columns and rows of $-\mathbb{E}[H(\theta, \sigma, \beta)]^{-1}$.

Similarly,

$$\widehat{\bar{\sigma}(\theta, \sigma)} \overset{a}{\sim} N\left(\bar{\sigma}(\theta, \sigma), J_{\bar{\sigma}(\theta,\sigma)} V_{\theta,\sigma} J'_{\bar{\sigma}(\theta,\sigma)}\right),$$

where $J_{\bar{\sigma}(\theta,\sigma)} = \left(\frac{\partial \bar{\sigma}(\theta,\sigma)}{\partial \theta}, \frac{\partial \bar{\sigma}(\theta,\sigma)}{\partial \sigma}\right).$

The partial derivatives involved in the above Jacobians are:

$$\frac{\partial \bar{\theta}(\theta, \sigma)}{\partial \theta} = 1 - \Phi(-\theta/\sigma),$$

$$\frac{\partial \bar{\theta}(\theta, \sigma)}{\partial \sigma} = \phi(-\theta/\sigma),$$

$$\frac{\partial \bar{\sigma}(\theta, \sigma)}{\partial \theta} = \frac{1}{2\bar{\sigma}(\theta, \sigma)} 2\Phi(-\theta/\sigma)\left(\theta(1 - \Phi(-\theta/\sigma)) + \sigma\phi(-\theta/\sigma)\right) + 2\theta\phi(-\theta/\sigma)^2(\theta - 1), \text{ and}$$

$$\frac{\partial \bar{\sigma}(\theta, \sigma)}{\partial \sigma} = \frac{1}{2\bar{\sigma}(\theta, \sigma)} - 2\theta\phi(-\theta/\sigma) + 2\sigma\left(1 - \Phi(-\theta/\sigma)\right) + 2\theta\Phi(-\theta/\sigma)\phi(-\theta/\sigma) - 2\sigma\phi(-\theta/\sigma)^2.$$

E   First and Second Moments of Truncated Normal

To write closed form expressions for (1), let

$$\Phi_{z_t}(\theta_m, \sigma_m) := \int_{\ell_{t,z_t}}^{u_{t,z_t}} f(\theta_t \mid \theta_m, \sigma_m)d\theta_t.$$

With this notation we now write

$$\mathbb{E}[\theta_t \mid z_t, \theta_m, \sigma_m] = \Phi_{z_t}(\theta_m, \sigma_m)\theta_m + \left(\phi\left(\frac{\ell_{t,z_t} - \theta_m}{\sigma_m}\right) - \phi\left(\frac{u_{t,z_t} - \theta_m}{\sigma_m}\right)\right)\sigma_m,$$

and

$$
\begin{aligned}
\mathbb{E}[\theta_t^2 \mid z_t, \theta_m, \sigma_m] &= \sigma_m^2 \left(\frac{\ell_{t,z_t} - \theta_m}{\sigma_m} \phi \left(\frac{\ell_{t,z_t} - \theta_m}{\sigma_m} \right) - \frac{u_{t,z_t} - \theta_m}{\sigma_m} \phi \left(\frac{u_{t,z_t} - \theta_m}{\sigma_m} \right) + \Phi_{z_t}(\theta_m, \sigma_m) \right) \\
&\quad + 2\mathbb{E}[\theta_t \mid z_t, \theta_m, \sigma_m]\theta_m - \Phi_{z_t}(\theta_m, \sigma_m)\theta_m^2,
\end{aligned}
$$

where $\phi(.)$ is the standard normal probability density function.

F   Derivative Calculations for Standard Errors

All the necessary partial derivatives needed to calculate the Jacobian of $M(\theta_m, \sigma_m, \beta_m)$ (and the Hessian $H(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta})$) are obtained analytically and are given by the following expressions:

$$
\ddot{Q}(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta}; \widehat{\theta}, \widehat{\sigma}, \widehat{\beta}) = \begin{pmatrix} -\frac{T+\kappa}{\widehat{\sigma}^2} & 0 & 0 \\ 0 & -\frac{2(T+\nu+3)}{\widehat{\sigma}^2} & 0 \\ 0 & 0 & \frac{\partial^2 Q(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta}; \widehat{\theta}, \widehat{\sigma}, \widehat{\beta})}{\partial \beta \partial \beta} \end{pmatrix}.
$$

As for the derivatives of the Jacobian, we have:

$$
\begin{aligned}
\frac{\partial \theta_{m+1}}{\partial \theta_m} &= \frac{1}{T+\kappa} \sum_{t=1}^{T} \sum_{z_t=1}^{n_t} \frac{p(s_t \mid v_t, \beta_m, u_{t,z_t})}{h(X_t \mid \theta_m, \sigma_m, \beta_m)} \left[\frac{\mathbb{E}_{z_t}[\theta_t^* \mid \theta_m, \sigma_m]}{\sigma_m^2} \left(\sum_{z_t'=1}^{n_t} \frac{p(s_t \mid v_t, \beta_m, u_{t,z_t'})}{h(X_t \mid \theta_m, \sigma_m, \beta_m)} \left(\mathbb{E}_{z_t'}[1 \mid \theta_m, \sigma_m]\theta_m \right. \right. \right. \\
&\quad \left. \left. \left. - \mathbb{E}_{z_t'}[\theta_t^* \mid \theta_m, \sigma_m] \right) \right) + \frac{\partial \mathbb{E}_{z_t}[\theta_t^* \mid \theta_m, \sigma_m]}{\partial \theta_m} \right].
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \theta_{m+1}}{\partial \sigma_m} &= \frac{1}{T+\kappa} \sum_{t=1}^{T} \sum_{z_t=1}^{n_t} \frac{p(s_t \mid v_t, \beta_m, u_{t,z_t})}{h(X_t \mid \theta_m, \sigma_m, \beta_m)} \left[\frac{\mathbb{E}_{z_t}[\theta_t^* \mid \theta_m, \sigma_m]}{\sigma_m} \left(\sum_{z_t'=1}^{n_t} \frac{p(s_t \mid v_t, \beta_m, u_{t,z_t'})}{h(X_t \mid \theta_m, \sigma_m, \beta_m)} \left(\mathbb{E}_{z_t'}[1 \mid \theta_m, \sigma_m] \right. \right. \right. \\
&\quad \left. \left. \left. - \frac{\mathbb{E}_{z_t'}[\theta_t^{*\,2} \mid \theta_m, \sigma_m] - 2\mathbb{E}_{z_t'}[\theta_t^* \mid \theta_m, \sigma_m]\theta_m + \mathbb{E}_{z_t'}[1 \mid \theta_m, \sigma_m]\theta_m^2}{\sigma_m^2} \right) \right) + \frac{\partial \mathbb{E}_{z_t}[\theta_t^* \mid \theta_m, \sigma_m]}{\partial \sigma_m} \right].
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \theta_{m+1}}{\partial \beta_m} &= \frac{1}{T+\kappa} \sum_{t=1}^{T} \sum_{z_t=1}^{n_t} \frac{p(s_t \mid v_t, \beta_m, u_{t,z_t})}{h(X_t \mid \theta_m, \sigma_m, \beta_m)} \left[\mathbb{E}_{z_t}[\theta_t^* \mid \theta_m, \sigma_m] \left(\frac{\partial \log(p(s_t \mid v_t, \beta_m, u_{t,z_t}))}{\partial \beta_m} \right. \right. \\
&\quad \left. \left. - \sum_{z_t'=1}^{n_t} \frac{p(s_t \mid v_t, \beta_m, u_{t,z_t'})}{h(X_t \mid \theta_m, \sigma_m, \beta_m)} \frac{\partial \log(p(s_t \mid v_t, \beta_m, u_{t,z_t'}))}{\partial \beta_m} \mathbb{E}_{z_t'}[1 \mid \theta_m, \sigma_m] \right) \right].
\end{aligned}
$$

$$\frac{\partial \sigma_{m+1}}{\partial \theta_m} = \frac{1}{2\sigma_{m+1}(T+\nu+3)}\left[\sum_{t=1}^{T}\sum_{z_t=1}^{n_t}\frac{1}{T}\frac{p(s_t \mid v_t,\beta_m,u_{t,z_t})}{h(X_t \mid \theta_m,\sigma_m,\beta_m)}\left[\frac{\partial \mathbb{E}_{z_t}[\theta_t^{*2} \mid \theta_m,\sigma_m]}{\partial(\theta_m)} - \frac{\mathbb{E}_{z_t}[\theta_t^{*2} \mid \theta_m,\sigma_m]}{\sigma_m^2}\right.\right.$$

$$\left.\left.\left(\sum_{z_t'=1}^{n_t}\frac{p(s_t \mid v_t,\beta_m,u_{t,z_t'})}{h(X_t \mid \theta_m,\sigma_m,\beta_m)}\left(\mathbb{E}_{z_t'}[\theta_t^{*} \mid \theta_m,\sigma_m] - \mathbb{E}_{z_t'}(1 \mid \theta_m,\sigma_m)\theta_m\right)\right)\right] - 2(T+\kappa)\theta_{m+1}\frac{\partial \theta_{m+1}}{\partial \theta_m}\right],$$

$$\frac{\partial \sigma_{m+1}}{\partial \sigma_m} = \frac{1}{2\sigma_{m+1}(T+\nu+3)}\left[\sum_{t=1}^{T}\sum_{z_t=1}^{n_t}\frac{1}{T}\frac{p(s_t \mid v_t,\beta_m,u_{t,z_t})}{h(X_t \mid \theta_m,\sigma_m,\beta_m)}\left[\frac{\partial \mathbb{E}_{z_t}[\theta_t^{*2} \mid \theta_m,\sigma_m]}{\partial \sigma_m} - \frac{\mathbb{E}_{z_t}[\theta_t^{*2} \mid \theta_m,\sigma_m]}{\sigma_m}\sum_{z_t'=1}^{n_t}\frac{p(s_t \mid v_t,\beta_m,u_{t,z_t'})}{h(X_t \mid \theta_m,\sigma_m,\beta_m)}\right.\right.$$

$$\left.\left.\left(\frac{\mathbb{E}_{z_t'}[\theta_t^{2} \mid \theta_m,\sigma_m] - 2\mathbb{E}_{z_t'}[\theta_t^{*} \mid \theta_m,\sigma_m](\theta_m) + \mathbb{E}_{z_t'}(1 \mid \theta_m,\sigma_m)(\theta_m)^2}{\sigma_m^2} - \mathbb{E}_{z_t'}(1 \mid \theta_m,\sigma_m)\right)\right]\right.$$

$$\left.-2(T+\kappa)\theta_{m+1}\frac{\partial \theta_{m+1}}{\partial \sigma_m}\right],$$

$$\frac{\partial \sigma_{m+1}}{\partial \beta_m} = \frac{1}{2\sigma_{m+1}(T+\nu+3)}\left[\sum_{t=1}^{T}\sum_{z_t=1}^{n_t}\frac{1}{T}\frac{p(s_t \mid v_t,\beta_m,u_{t,z_t})}{h(X_t \mid \theta_m,\sigma_m,\beta_m)}\left[\mathbb{E}_{z_t}[\theta_t^{*2} \mid \theta_m,\sigma_m]\left(\frac{\partial \log(g(s_t \mid p_m,v_{tk}))}{\partial \beta_m} - \sum_{z_t'=1}^{n_t}\frac{p(s_t \mid v_t,\beta_m,u_{t,z_t'})}{h(X_t \mid \theta_m,\sigma_m,\beta_m)}\right.\right.\right.$$

$$\left.\left.\left.\frac{\partial \log(p(s_t \mid v_t,\beta_m,u_{t,z_t'}))}{\partial \beta_m}\mathbb{E}_{z_t'}(1 \mid \theta_m,\sigma_m)\right)\right] - 2(T+\kappa)\theta_{m+1}\frac{\partial \theta_{m+1}}{\partial \beta_m}\right],$$

where

$$\frac{\partial \mathbb{E}_{z_t}[\theta_t^* \mid \theta_m, \sigma_m]}{\partial \theta_m} = \left[\phi\left(\frac{l_{t,z_t} - \theta_m}{\sigma_m}\right) - \phi\left(\frac{u_{t,z_t} - \theta_m}{\sigma_m}\right)\right]\frac{\theta_m}{\sigma_m} + \mathbb{E}_{z_t}[1 \mid \theta_m, \sigma_m] -$$
$$\left[\phi'\left(\frac{l_{t,z_t} - \theta_m}{\sigma_m}\right) - \phi'\left(\frac{u_{t,z_t} - \theta_m}{\sigma_m}\right)\right],$$

$$\frac{\partial \mathbb{E}_{z_t}[\theta_t^* \mid \theta_m, \sigma_m]}{\partial \sigma_m} = \left[\frac{l_{t,z_t} - \theta_m}{\sigma_m^2}\phi\left(\frac{l_{t,z_t} - \theta_m}{\sigma_m}\right) - \frac{u_{t,z_t} - \theta_m}{\sigma_m^2}\phi\left(\frac{u_{t,z_t} - \theta_m}{\sigma_m}\right)\right]\theta_m +$$
$$\left[\left(\frac{l_{t,z_t} - \theta_m}{\sigma_m}\right)^2 \phi\left(\frac{l_{t,z_t} - \theta_m}{\sigma_m}\right) - \left(\frac{u_{t,z_t} - \theta_m}{\sigma_m}\right)^2 \phi\left(\frac{u_{t,z_t} - \theta_m}{\sigma_m}\right)\right] +$$
$$\left[\phi\left(\frac{l_{t,z_t} - \theta_m}{\sigma_m}\right) - \phi\left(\frac{u_{t,z_t} - \theta_m}{\sigma_m}\right)\right],$$

$$\frac{\partial \mathbb{E}_{z_t}[\theta_t^{*2} \mid \theta_m, \sigma_m]}{\partial \theta_m} = 3\left[\mathbb{E}_{z_t}[\theta_t^* \mid \theta_m, \sigma_m]\left(1 - \frac{\theta_m^2}{\sigma_m^2}\right) - \mathbb{E}_{z_t}[1 \mid \theta_m, \sigma_m]\theta_m\right] + \frac{\theta_m}{\sigma_m^2}$$
$$\left(2\mathbb{E}_{z_t}[\theta_t^{*2} \mid \theta_m, \sigma_m] + \mathbb{E}_{z_t}[1 \mid \theta_m, \sigma_m]\theta_m^2\right) - \sigma_m\left[\frac{l_{t,z_t} - \theta_m}{\sigma_m}\phi\left(\frac{l_{t,z_t} - \theta_m}{\sigma_m}\right)\right.$$
$$\left.\left(1 - \frac{l_{t,z_t} - \theta_m}{\sigma_m}\right) - \frac{u_{t,z_t} - \theta_m}{\sigma_m}\phi\left(\frac{u_{t,z_t} - \theta_m}{\sigma_m}\right)\left(1 - \frac{u_{t,z_t} - \theta_m}{\sigma_m}\right)\right],$$

$$\frac{\partial \mathbb{E}_{z_t}[\theta_t^{*2} \mid \theta_m, \sigma_m]}{\partial \sigma_m} = \left[\frac{l_{t,z_t} - \theta_m}{\sigma_m}\phi\left(\frac{l_{t,z_t} - \theta_m}{\sigma_m}\right) - \frac{u_{t,z_t} - \theta_m}{\sigma_m}\phi\left(\frac{u_{t,z_t} - \theta_m}{\sigma_m}\right)\right]\left(2\sigma_m + \frac{\theta_m^2}{\sigma_m}\right) +$$
$$2\mathbb{E}_{z_t}[1 \mid \theta_m, \sigma_m]\left(\sigma_m - \frac{\theta_m^2}{\sigma_m}\right) + 2\theta_m\left[\left(\frac{l_{t,z_t} - \theta_m}{\sigma_m}\right)^2 \phi\left(\frac{l_{t,z_t} - \theta_m}{\sigma_m}\right)\right.$$
$$\left.- \left(\frac{u_{t,z_t} - \theta_m}{\sigma_m}\right)^2 \phi\left(\frac{u_{t,z_t} - \theta_m}{\sigma_m}\right)\right] + 2\mathbb{E}_{z_t}[\theta_t^* \mid \theta_m, \sigma_m]\frac{\theta_m}{\sigma_m}$$
$$+\sigma_m\left[\left(\frac{l_{t,z_t} - \theta_m}{\sigma_m}\right)^3 \phi\left(\frac{l_{t,z_t} - \theta_m}{\sigma_m}\right) - \left(\frac{u_{t,z_t} - \theta_m}{\sigma_m}\right)^3 \phi\left(\frac{u_{t,z_t} - \theta_m}{\sigma_m}\right)\right],$$

$$\frac{\partial^2 Q(\gamma;\gamma_m)}{\partial\theta_m\partial\beta} = \sum_{t=1}^{T}\sum_{z_t=1}^{n_t}\frac{p(s_t\mid v_t,\beta_m,u_{t,z_t})}{h(X_t\mid\theta_m,\sigma_m,\beta_m)}\left[\frac{\mathbb{E}_{z_t}[\theta_t^*\mid\theta_m,\sigma_m]-\mathbb{E}_{z_t}[1\mid\theta_m,\sigma_m]\theta_m}{\sigma_m^2}-\mathbb{E}_{z_t}[1\mid\theta_m,\sigma_m]\right.$$

$$\sum_{z_t=1}^{n_t}\frac{p(s_t\mid v_t,\beta_m,u_{t,z_t})}{h(X_t\mid\theta_m,\sigma_m,\beta_m)}\frac{\mathbb{E}_{z_t}[\theta_t^*\mid\theta_m,\sigma_m]-\mathbb{E}_{z_t}[1\mid\theta_m,\sigma_m]\theta_m}{\sigma_m^2}\right]$$

$$\times\left[\sum_{i=z_t}^{N}s_{ti}\left(\log(v_{t,i})-\frac{\sum_{j=z_t}^{N}v_{t,j}^{\beta}\log(v_{t,j})}{\sum_{j=z_t}^{N}v_{t,j}^{\beta}}\right)\right]$$

$$\frac{\partial^2 Q(\gamma;\gamma_m)}{\partial\sigma_m\partial\beta} = \sum_{t=1}^{T}\sum_{z_t=1}^{n_t}\frac{p(s_t\mid v_t,\beta_m,u_{t,z_t})}{h(X_t\mid\theta_m,\sigma_m,\beta_m)}\left[\left(\frac{\mathbb{E}_{z_t}[\theta_t^{*2}\mid\theta_m,\sigma_m]-2\mathbb{E}_{z_t}[\theta_t^*\mid\theta_m,\sigma_m]\theta_m+\mathbb{E}_{z_t}[1\mid\theta_m,\sigma_m]\theta_m^2}{\sigma_m^2}\right.\right.$$

$$\mathbb{E}_{z_t}[1\mid\theta_m,\sigma_m]\bigg)\frac{1}{\sigma_m}-\mathbb{E}_{z_t}[1\mid\theta_m,\sigma_m]\sum_{z_t=1}^{n_t}\frac{p(s_t\mid v_t,\beta_m,u_{t,z_t})}{h(X_t\mid\theta_m,\sigma_m,\beta_m)}\frac{1}{\sigma_m}$$

$$\left(\frac{\mathbb{E}_{z_t}[\theta_t^{*2}\mid\theta_m,\sigma_m]-2\mathbb{E}_{z_t}[\theta_t^*\mid\theta_m,\sigma_m]\theta_m+\mathbb{E}_{z_t}[1\mid\theta_m,\sigma_m]\theta_m^2}{\sigma_m^2}-\mathbb{E}_{z_t}[1\mid\theta_m,\sigma_m]\right)\bigg]$$

$$\times\left[\sum_{i=z_t}^{N}s_{ti}\left(\log(v_{t,i})-\frac{\sum_{j=z_t}^{N}v_{t,j}^{\beta}\log(v_{t,j})}{\sum_{j=z_t}^{N}v_{t,j}^{\beta}}\right)\right],$$

$$\frac{\partial^2 Q(\gamma;\gamma_m)}{\partial\beta_m\partial\beta} = \sum_{t=1}^{T}\sum_{z_t=1}^{n_t}\frac{p(s_t\mid v_t,\beta_m,u_{t,z_t})}{h(X_t\mid\theta_m,\sigma_m,\beta_m)}\left[\frac{\partial\log(p(s_t\mid v_t,\beta_m,u_{t,z_t}))}{\partial\beta_m}-\sum_{z_t=1}^{n_t}\frac{p(s_t\mid v_t,\beta_m,u_{t,z_t})}{h(X_t\mid\theta_m,\sigma_m,\beta_m)}\right.$$

$$\frac{\partial\log(p(s_t\mid v_t,\beta_m,u_{t,z_t}))}{\partial\beta_m}\mathbb{E}_{z_t}[1\mid\theta_m,\sigma_m]\bigg]\mathbb{E}_{z_t}[1\mid\theta_m,\sigma_m]\times$$

$$\left[\sum_{i=z_t}^{N}s_{ti}\left(\log(v_{t,i})-\frac{\sum_{j=z_t}^{N}v_{t,j}^{\beta}\log(v_{t,j})}{\sum_{j=z_t}^{N}v_{t,j}^{\beta}}\right)\right].$$

G   Auxiliary Tests

All tests are based on Pearson's goodness-of-fit test statistic. We use simulation under the null hypothesis to compute exact p-values instead of relying on asymptotic arguments. Exact procedures are described below for each test.

G.1 Fit

We assess model fit using the test statistic

(6)
$$P((s_t)_{t=1}^T, \widehat{\theta}, \widehat{\sigma}, \widehat{\beta}) = \sum_{t=1}^T \sum_{i=1}^N \frac{(s_{t,i} - \overline{s}_{t,i}(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta}))^2}{\overline{s}_{t,i}(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta})}.$$

The null hypothesis is that the observed seat allocations are drawn from the multinomial distribution with means as specified by the model parameters, specifically, $\overline{s}_{t,i}(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta})$ is the expected number of seats of party i in election t under the null:

$$\overline{s}_{t,i}(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta}) = S_t \sum_{z_t=1}^N q_{t,i}(v_t, \widehat{\beta}, u_{t,z_t}) p(z_t \mid v_t, \widehat{\theta}, \widehat{\sigma}),$$

where $p(N \mid v_t, \widehat{\theta}, \widehat{\sigma}) = 1 - \Phi\left(\frac{u_{t,N-1} - \widehat{\theta}}{\widehat{\sigma}}\right)$. We simulate the distribution of the test statistic (and calculate the p-value) by calculating (6) using simulated seat allocations generated for the observed vote shares under the null. Specifically, to obtain one of $10,000$ realizations from the distribution of this statistic:

1. For each t, we draw a latent threshold $\widetilde{\theta}_t^*$ from the normal distribution with mean $\widehat{\theta}$ and variance $\widehat{\sigma}^2$, which determines a non-negative threshold, $\widetilde{\theta}_t$ ($\widetilde{\theta}_t = \widetilde{\theta}_t^*$ if $\widetilde{\theta}_t^* \geq 0$ and $\widetilde{\theta}_t = 0$ otherwise).

2. For each t, we draw a vector $\widetilde{s}_t$ allocating S_t total seats from the Multinomial distribution with probabilities $q_{t,i}(v_t, \widehat{\beta}, \widetilde{\theta}_t)$ as defined in (2).

3. We compute the statistic in (6) using the simulated seats $(\widetilde{s}_t)_{t=1}^T$ instead of the actual seats $(s_t)_{t=1}^T$, $P((\widetilde{s}_t)_{t=1}^T, \widehat{\theta}, \widehat{\sigma}, \widehat{\beta})$.

We repeat this process 10,000 times and compute the p-value as the fraction of the Pearson's statistics calculated with these simulated seats that is greater than or equal to P.

Similarly, the fit test of the restricted model is based on the statistic

$$(7) \qquad P^o((s_t)_{t=1}^T, \widehat{\beta}^o) = \sum_{t=1}^{T} \sum_{i=1}^{N} \frac{(s_{i,t} - \overline{s}_{i,t}^o(\widehat{\beta}^o))^2}{\overline{s}_{i,t}^o(\widehat{\beta}^o)}.$$

Party i's expected seats according to the restricted model are given by

$$\overline{s}_{i,t}^o(\widehat{\beta}^o) = S_t q_{i,t}(v_t, \widehat{\beta}^o, 0),$$

where $\widehat{\beta}^o$ denotes the estimated disproportionality parameter from the model with no thresh-old. The only difference in the computation of the p-value is that we now skip step one and set the threshold to zero instead. Step 2 is now executed with multinomial probabilities $q_{i,t}(v_t, \widehat{\beta}^o, 0)$ and the remaining steps are analogous.

G.2 Comparison with the Restricted Model

To compare the proposed model with the model with no threshold we use the following statistic:

$$(8) \qquad D = P^o - P,$$

where P^o and P are computed as in (7) and (6), respectively. The null hypothesis is that data are generated from the model with no threshold and large values of the statistic provide evidence against that model and in favor of the alternative model with thresholds. We compute the p-value by drawing seat allocations under the null (from the Multinomial distribution with probabilities $q_{t,i}(v_t, \widehat{\beta}^o, 0)$, as in the case of the fit statistic P^o), evaluating (8) using the simulated seats instead of the actual seats, and computing the fraction of simulated statistics that exceed D to obtain the p-value.

G.3 Electoral System Change

Consider a set of elections $t = 1, \ldots, T$. Under the null hypothesis one electoral system governs all T elections with estimated parameters denoted by $(\widehat{\theta}^c, \widehat{\sigma}^c, \widehat{\beta}^c)$. The alternative hypothesis is the these T elections are partitioned into two or more subsets $T_f \subset \{1, \ldots, T\}$, and a different electoral system with estimated parameters $(\widehat{\theta}^f, \widehat{\sigma}^f, \widehat{\beta}^f)$ governs all elections in each subset T_f. The comparison test computes the following statistic for elections that are part of the coarser system

$$
(9) \qquad D^{systems} = \sum_t P_t^c - \sum_f \sum_{t \in T_f} P_t^f,
$$

where

$$
P_t^k = \sum_{i=1}^N \frac{(s_{t,i} - \bar{s}_{t,i}(\widehat{\theta}^k, \widehat{\sigma}^k, \widehat{\beta}^k))^2}{\bar{s}_{t,i}(\widehat{\theta}^k, \widehat{\sigma}^k, \widehat{\beta}^k)}, k \in \{c, f\}.
$$

We simulate the distribution of this statistic under the null. We obtain one realization from that distribution as follows:

1. For each t, we draw a latent threshold $\widetilde{\theta}_t^*$ from the normal distribution with mean $\widehat{\theta}^c$ and variance $(\widehat{\sigma}^c)^2$, which determines a non-negative threshold, $\widetilde{\theta}_t$ ($\widetilde{\theta}_t = \widetilde{\theta}_t^*$ if $\widetilde{\theta}_t^* \geq 0$ or $\widetilde{\theta}_t = 0$ otherwise).

2. For each t, we draw a vector $\tilde{s}_t$ allocating S_t total seats from the Multinomial distribution with probabilities $q_{t,i}(v_t, \widehat{\beta}^c, \widetilde{\theta}_t)$ as defined in (2).

3. We compute the statistic in (9) using the simulated seats $(\tilde{s}_t)_{t=1}^T$ instead of the actual seats $(s_t)_{t=1}^T$.

We repeat this process 10,000 times and compute the p-value as the fraction of these 10,000 Pearson's statistics that is greater than or equal to $D^{systems}$.

14

G.4 Bayesian Versions of the tests

We can perform similar tests using information from the full posterior. The tests of fit can be performed using the same test statistics but using the posterior predictive distribution Gelman et al. (2004). Specifically, using (6), prior to the first step (step 1) we would obtain a sample of size one of the model parameters from the posterior distribution, say θ', σ', β'. We would then execute Steps 1-3 to obtain a statistic, drawing threshold realizations $\tilde{\theta}_t$, simulated seats $\tilde{s}_t$, and computing the statistic $P((\tilde{s}_t)_{t=1}^T, \theta', \sigma', \beta')$. This approach would yield a p-value as the fraction of times this statistic exceeds the statistic computed with the actual seats $P((s_t)_{t=1}^T, \theta', \sigma', \beta')$. We can perform the model comparison tests either using a version of the Bayesian Information Criterion (BIC)[1] or the Deviance Information Criterion Spiegelhalter et al. (2002) which has the advantage that it more naturally compensates for model complexity. The latter test is easy to compute once a sample of parameters from the posterior distribution is available.

---

[1]We perform a version of that test using the output from our computations with similar findings

H   Full Results

Table 1: MAP-EM Estimates and Comparison with Restricted Model

Country	System	T	Dates	Full $\bar{\theta}(\hat{\theta},\hat{\sigma})$	Full $\bar{\sigma}(\hat{\theta},\hat{\sigma})$	Full $\hat{\beta}$	Restricted $\hat{\beta}^o$	Tests Fit	Tests Fit^o	Tests Difference
Austria	AUT1	8	1945-1970	3.938†	0.585	1.283*	1.371*	1	0.805	0.018$^+$
				(0.471)	(0.154)	(0.064)	(0.061)			
	AUT2	6	1971-1990	3.247†	0.628	0.981	1.156*	0.995	0.506	0$^+$
				(1.074)	(0.182)	(0.057)	(0.05)			
	AUT3	6	1994-2008	3.841†	0.603	1.02	1.231*	1	0.589	0$^+$
				(0.498)	(0.17)	(0.059)	(0.05)			
Belgium	BEL1	16	1946-1991	0.518†	0.443	1.202*	1.22*	0.995	0.95	0.203
				(0.22)	(0.103)	(0.028)	(0.027)			
	BEL2	2	1995-1999	0.696	0.539	1.3*	1.326*	1	0.998	0.244
				(0.825)	(0.191)	(0.139)	(0.132)			
	BEL3	3	2003-2010	0.53	0.501	1.322*	1.326*	0.991	0.938	0.287
				(1.214)	(0.327)	(0.103)	(0.101)			
Bulgaria	BGR1	5	1991-2005	4.332†	0.564	1.015	1.377*	1	0.353	0$^+$
				(0.389)	(0.147)	(0.044)	(0.034)			
	BGR2^b	1	2009	3.652†	0.605	0.994	1.185*	0.985	0.051	0$^+$
				(0.657)	(0.175)	(0.1)	(0.09)			
Cyprus	CYP1	1	1981	4.039†	0.65	0.68	1.254	0.419	0.232	0.016$^+$
				(1.875)	(0.211)	(0.366)	(0.281)			
	CYP2	2	1985-1991	3.6†	0.646	0.969	1.173	1	0.952	0.073
				(1.979)	(0.209)	(0.265)	(0.223)			
	CYP3	3	1996-2006	1.586†	0.584	1.067	1.158	1	1	0.123
				(0.493)	(0.164)	(0.117)	(0.105)			
Czech Republic	CZE1	2	1996-1998	4.966†	0.652	1.041	1.406*	1	0.15	0$^+$
				(1.194)	(0.198)	(0.088)	(0.071)			
	CZE2	3	2002-2010	5.231†	0.647	1.223*	1.685*	0.999	0.329	0$^+$
				(0.935)	(0.193)	(0.105)	(0.083)			
Germany	DEU1	1	1949	0.012	0.073	1.139*	1.139*	0.986	0.865	0.374
				(0.116)	(1.104)	(0.058)	(0.058)			
	DEU2	1	1953	0.287	0.385	1.22*	1.225*	0.796	0.007	0.1
				(0.578)	(0.322)	(0.054)	(0.053)			
	DEU3	8	1957-1983	3.866†	0.744	1.04	1.182*	0.22	0.138	0$^+$
				(0.534)	(0.218)	(0.027)	(0.024)			

16

	Code	N	Years							
	DEU4	1	1987	2.019	0.634	0.998	1.103*	1	0.795	0.001+
				(2.286)	(0.203)	(0.073)	(0.062)			
	DEU5	1	1990	2.255†	0.586	1.109*	1.274*	0.48	0.065	0+
				(0.59)	(0.169)	(0.055)	(0.049)			
	DEU6	3	1994-1998	3.118†	0.627	1.018	1.181*	1	0.284	0+
				(1.088)	(0.186)	(0.036)	(0.03)			
	DEU7	2	2002-2005	2.696†	0.627	1.092*	1.226*	0.265	0.26	0+
				(1.183)	(0.188)	(0.044)	(0.037)			
	DEU8	1	2009	2.812	0.639	1.046	1.21*	0.931	0.1	0+
				(2.063)	(0.206)	(0.091)	(0.078)			
Denmark	DNK1	2	1945-1947	1.303†	0.586	1.047	1.083	0.992	0.978	0.178
				(0.604)	(0.156)	(0.084)	(0.081)			
	DNK2	2	1950, 4/1953	2.134	0.635	0.999	1.052	1	0.999	0.079
				(2.234)	(0.203)	(0.088)	(0.081)			
	DNK3	3	9/1953, 1957-1960	0.003	0.036	1.085	1.085	0.771	0.664	0.718
				(0.034)	(0.493)	(0.055)	(0.055)			
	DNK4	3	1964-1968	1.835†	0.536	0.998	1.169*	1	0.765	0.003+
				(0.357)	(0.136)	(0.068)	(0.059)			
	DNK5	14	1971-2005	1.874†	0.544	1.006	1.102*	0.958	0.958	0.052
				(0.172)	(0.104)	(0.027)	(0.025)			
	DNK6	1	2007	1.426†	0.598	0.997	1.051	1	0.993	0.115
				(0.791)	(0.168)	(0.122)	(0.115)			
Spain	ESP1	10	1977-2008	0.187†	0.269	1.255*	1.273*	0.973	0.628	0.054
				(0.102)	(0.078)	(0.02)	(0.018)			
Estonia	EST1	1	1992	1.664†	0.608	1.524*	1.585*	0.992	0.775	0.091
				(0.905)	(0.176)	(0.183)	(0.17)			
	EST2	2	1995-1999	4.231†	0.621	1.153	1.453*	1	0.822	0.001+
				(0.763)	(0.179)	(0.129)	(0.107)			
	EST3	2	2003-2007	3.503†	0.644	1.186	1.381*	1	0.984	0.037+
				(1.935)	(0.209)	(0.156)	(0.132)			
Finland	FIN1	18	1945-2007	0.359	0.451	1.195*	1.207*	0.991	0.946	0.476
				(0.282)	(0.157)	(0.031)	(0.029)			
France	FRA1	3	1945-1946	0.031	0.123	1.347*	1.347*	1	0.773	0.788
				(0.307)	(1.798)	(0.075)	(0.075)			
	FRA2	2	1951-1956	0.15	0.286	0.779	0.779	0.028	0.001	0.002+
				(1.512)	(2.075)	(0.078)	(0.078)			
	FRA3	2	1958-1962	1.713†	0.606	1.29*	1.359*	0.014	0	0.063
				(0.582)	(0.181)	(0.068)	(0.064)			
	FRA4	3	1967-1973	1.03†	0.563	1.258*	1.268*	0	0	0.324

Country	Code	N	Years							
				(0.56)	(0.138)	(0.045)	(0.044)			
	FRA5	2	1978-1981	1.492†	0.558	1.307*	1.426*	0.148	0	0.554
				(0.407)	(0.146)	(0.073)	(0.066)			
	FRA6	1	1986	0.77	0.54	1.199*	1.222*	0.033	0	0.947
				(0.627)	(0.16)	(0.07)	(0.067)			
	FRA7	5	1988-2007	0.008	0.059	1.694*	1.694*	0.881	0	1
				(0.086)	(0.945)	(0.038)	(0.038)			
United Kingdom	GBR1	18	1945-2010	0	0.01	1.486*	1.486*	1	0	0.976
				(0.003)	(0.084)	(0.021)	(0.021)			
Greece	GRC1	1	1974	2.013	0.635	1.722*	1.732*	1	0.04	0.217
				(2.688)	(0.202)	(0.107)	(0.105)			
	GRC2	2	1977-1981	0.285	0.39	1.742*	1.745*	1	0.195	0.806
				(1.218)	(0.702)	(0.089)	(0.088)			
	GRC3	1	1985	0.586	0.519	1.673*	1.683*	1	0.973	0.441
				(1.531)	(0.38)	(0.171)	(0.167)			
	GRC4	3	1989-1990	0.006	0.05	1.248*	1.249*	1	0.744	0.496
				(0.055)	(0.674)	(0.055)	(0.055)			
	GRC5	4	1993-2004	3.171†	0.556	1.254*	1.411*	0.952	0.362	0.003$^+$
				(0.336)	(0.145)	(0.053)	(0.049)			
	GRC6	2	2007-2009	3.102†	0.61	1.105*	1.209*	0.697	0.309	0.002$^+$
				(0.69)	(0.177)	(0.059)	(0.054)			
Croatia	HRV1^a	1	1992	0.005	0.048	2.201*	2.201*	0.167	0	0.998
				(0.059)	(0.74)	(0.225)	(0.225)			
	HRV2^b	1	1992	2.893†	0.595	1.048	1.228*	0.998	0.902	0.021$^+$
				(0.627)	(0.172)	(0.145)	(0.127)			
	HRV3^a	1	1995	0.252	0.369	7.829**	7.829**	1	1	0.578
				(3.163)	(2.164)	(3.109)	(3.109)			
	HRV4^b	1	1995	4.032†	0.628	1.05	1.294*	1	0.524	0.002$^+$
				(0.895)	(0.185)	(0.156)	(0.133)			
	HRV5	3	2000-2007	1.168†	0.652	1.305*	1.386*	0.966	0.703	0.131
				(0.481)	(0.179)	(0.074)	(0.066)			
Hungary	HUN1	1	1990	0.003	0.035	2.332*	2.332*	0.867	0.001	0.998
				(0.035)	(0.511)	(0.184)	(0.184)			
	HUN2	5	1994-2010	0.002	0.026	2.149*	2.149*	0.968	0.002	0.995
				(0.019)	(0.327)	(0.081)	(0.081)			
Ireland	IRL1	6	1948-1965	0.209	0.331	1.124*	1.128*	0.99	0.929	0.382
				(0.403)	(0.336)	(0.05)	(0.049)			
	IRL2	3	1969-1977	2.175†	0.596	1.177	1.288*	0.975	0.808	0.083
				(0.632)	(0.168)	(0.123)	(0.107)			

	IRL3	9	1981-2007	0.13	0.263	1.196*	1.2*	1	0.922	0.499
				(0.292)	(0.463)	(0.038)	(0.036)			
Iceland	ISL1	5	1946-1956,	5.614†	0.669	0.94	1.084	0.282	0.144	0.017$^+$
			6/1959	(1.564)	(0.212)	(0.151)	(0.136)			
	ISL2	8	10/1959,	1.036	0.591	1.138	1.18*	0.979	0.986	0.28
			1963-1983	(0.78)	(0.193)	(0.093)	(0.084)			
	ISL3	4	1987-1999	1.193†	0.619	1.13	1.203*	1	1	0.224
				(0.694)	(0.187)	(0.116)	(0.1)			
	ISL4	3	2003-2009	4.442†	0.655	1.114	1.258*	1	0.993	0.08
				(1.859)	(0.212)	(0.157)	(0.141)			
Italy	ITA1	1	1946	0.007	0.053	1.04	1.04	0.982	0.978	0.366
				(0.071)	(0.84)	(0.049)	(0.049)			
	ITA2	1	1948	0.005	0.046	1.142*	1.142*	0.999	0.908	0.429
				(0.056)	(0.711)	(0.046)	(0.046)			
	ITA3	1	1953	0.021	0.101	1.14*	1.14*	0.998	0.944	0.648
				(0.189)	(1.37)	(0.05)	(0.05)			
	ITA4	1	1958	0.003	0.033	1.107*	1.107*	0.999	0.985	0.711
				(0.032)	(0.479)	(0.043)	(0.043)			
	ITA5	8	1963-1992	0.001	0.016	1.127*	1.127*	1	0.926	0.492
				(0.007)	(0.161)	(0.015)	(0.015)			
	ITA6^a	2	1996-2001	0.002	0.024	1.221*	1.221*	0.94	0.167	0.488
				(0.017)	(0.297)	(0.041)	(0.041)			
	ITA7	2	2006-2008	0.002	0.025	1.17*	1.17*	0.934	0.148	0.167
				(0.018)	(0.319)	(0.028)	(0.028)			
Lithuania	LTU1^b	1	1992	1.702†	0.597	1.237	1.347*	0.827	0.591	0.041$^+$
				(0.733)	(0.171)	(0.154)	(0.134)			
	LTU2^b	4	1996-2008	5.13†	0.597	1.008	1.656*	1	0.433	0$^+$
				(0.527)	(0.165)	(0.124)	(0.09)			
Luxembourg	LUX1	3	1945,	0.258	0.373	1.186	1.187	1	1	0.453
			1954-1959	(2.093)	(1.384)	(0.146)	(0.145)			
	LUX2	2	1948-1951	0.252	0.369	0.617	0.617	0.178	0.134	0.905
				(3.162)	(2.164)	(0.235)	(0.235)			
	LUX3	4	1964-1979	1	0.59	1.23*	1.243*	1	1	0.404
				(1.2)	(0.183)	(0.131)	(0.128)			
	LUX4	1	1984	3.148†	0.626	1.161	1.272	0.998	0.954	0.163
				(1.03)	(0.186)	(0.235)	(0.223)			
	LUX5	5	1989-2009	2.307†	0.578	1.16	1.24*	0.521	0.506	0.46
				(0.475)	(0.155)	(0.107)	(0.099)			
Latvia	LVA1	1	1993	3.584†	0.631	0.986	1.279*	1	0.869	0$^+$

				(1.065)	(0.188)	(0.142)	(0.11)			
	LVA2	5	1995-2010	5.056†	0.573	1.01	1.538*	1	0.647	0$^+$
				(0.412)	(0.148)	(0.111)	(0.084)			
F.Y.R.O. Macedonia	MKD1	1	1994	0.007	0.054	1.373*	1.373*	0.096	0	0.973
				(0.073)	(0.856)	(0.125)	(0.125)			
	MKD2^a	1	1998	0.252	0.369	1.416*	1.416*	0.042	0	0.048$^+$
				(3.163)	(2.164)	(0.213)	(0.213)			
	MKD3^b	1	1998	5.594†	0.662	1.158	1.419	0.999	0.924	0.123
				(1.234)	(0.201)	(0.417)	(0.38)			
	MKD4	3	2002-2008	0.81†	0.56	1.215*	1.307*	0.998	0.978	0.039$^+$
				(0.38)	(0.139)	(0.072)	(0.06)			
Malta	MLT1	2	1966-1971	0.257	0.372	2.84*	2.84*	0.999	0.895	0.424
				(3.247)	(2.166)	(1.357)	(1.356)			
	MLT2	2	1976-1981	0.252	0.369	0.439*	0.439*	0.721	0.711	0.558
				(3.163)	(2.164)	(3.462)	(3.462)			
	MLT3	2	1987-1992	2.898	0.64	1.015*	1.854*	1	0.901	0.342
				(2.674)	(0.206)	(2.767)	(1.02)			
	MLT4	4	1996-2008	2.274	0.637	1.652*	2.099*	0.99	0.966	0.404
				(3.498)	(0.204)	(1.81)	(0.937)			
Moldova	MOL1	3	1994-1998, 2010	5.076†	0.661	1.004	1.496*	1	0.447	0$^+$
				(1.683)	(0.216)	(0.111)	(0.083)			
	MOL2	3	2001-2005, 4/2009	6.814†	0.679	1.021	1.489*	1	0.2	0$^+$
				(1.158)	(0.204)	(0.09)	(0.07)			
	MOL3	1	7/2009	3.211	0.642	1.04	1.207	1	0.774	0.015$^+$
				(1.992)	(0.207)	(0.157)	(0.133)			
Montenegro	MON1	2	2006-2009	3.295†	0.615	1.007	1.295*	1	0.764	0$^+$
				(0.64)	(0.185)	(0.109)	(0.086)			
Netherlands	NLD1	3	1946-1952	0.557	0.506	1.06	1.067	1	1	0.355
				(0.858)	(0.23)	(0.086)	(0.085)			
	NLD2	17	1956-2010	0.535†	0.402	1.049*	1.07*	1	1	0.126
				(0.153)	(0.066)	(0.025)	(0.024)			
Norway	NOR1	2	1945-1949	0.526	0.495	1.272*	1.278*	0.956	0.499	0.459
				(0.909)	(0.256)	(0.076)	(0.075)			
	NOR2	9	1953-1985	0.412	0.49	1.155*	1.17*	0.946	0.661	0.474
				(0.327)	(0.202)	(0.034)	(0.031)			
	NOR3	4	1989-2001	0.241	0.367	1.238*	1.253*	0.959	0.749	0.513
				(0.395)	(0.308)	(0.061)	(0.055)			
	NOR4	2	2005-2009	2.401†	0.63	1.076	1.168*	0.975	0.905	0.007$^+$

				(1.438)	(0.192)	(0.083)	(0.074)			
Poland	POL1	1	1991	0.002	0.028	1.269*	1.269*	0.989	0.217	0.701
				(0.023)	(0.377)	(0.054)	(0.054)			
	POL2	2	1993-1997	5.124†	0.578	1.638*	2.056*	1	0.03	0$^+$
				(0.432)	(0.155)	(0.072)	(0.064)			
	POL3	1	2001	6.556†	0.673	1.052	1.265*	0.995	0.023	0$^+$
				(1.113)	(0.202)	(0.065)	(0.057)			
	POL4	2	2005-2007	5.074†	0.661	1.238*	1.499*	1	0.139	0$^+$
				(1.599)	(0.209)	(0.068)	(0.057)			
Portugal	PRT1	2	1975-1976	0.413	0.482	1.327*	1.379*	0.071	0.01	0.989
				(0.394)	(0.184)	(0.079)	(0.071)			
	PRT2	5	1979-1987	1.538†	0.565	1.198*	1.334*	0.719	0.537	0.02$^+$
				(0.377)	(0.153)	(0.052)	(0.044)			
	PRT3	6	1991-2009	1.565†	0.55	1.276*	1.349*	1	0.978	0.027$^+$
				(0.419)	(0.144)	(0.046)	(0.042)			
Romania	ROM1	1	1990	0.364	0.412	0.98	1.047	1	1	0$^+$
				(0.391)	(0.184)	(0.033)	(0.03)			
	ROM2	2	1992-1996	3.069†	0.576	1.017	1.453*	0.877	0.4	0$^+$
				(0.503)	(0.161)	(0.06)	(0.045)			
	ROM3	2	2000-2004	5.491†	0.635	0.996	1.441*	0.999	0.49	0$^+$
				(0.66)	(0.182)	(0.064)	(0.047)			
	ROM4	1	2008	4.36†	0.651	0.99	1.441*	0.997	0.451	0$^+$
				(1.6)	(0.205)	(0.124)	(0.095)			
Slovenia	SVN1	2	1992-1996	2.965†	0.555	0.862	1.352*	0.999	0.666	0$^+$
				(0.414)	(0.149)	(0.128)	(0.096)			
	SVN2	3	2000-2008	3.654†	0.585	1.015	1.243*	1	0.928	0$^+$
				(0.463)	(0.16)	(0.088)	(0.073)			
Serbia	SER1	2	2007-2008	4.176†	0.627	1.013	1.322*	1	0.311	0$^+$
				(0.836)	(0.18)	(0.074)	(0.061)			
Slovakia	SVK1	1	1994	4.312†	0.635	1.013	1.324*	1	0.22	0$^+$
				(0.972)	(0.188)	(0.118)	(0.093)			
	SVK2	4	1998-2010	4.829†	0.599	1.007	1.424*	1	0.268	0$^+$
				(0.552)	(0.168)	(0.081)	(0.064)			
Sweden	SWE1	1	1948	0.252	0.369	1.16	1.16	0.795	0.567	0.51
				(3.163)	(2.164)	(0.105)	(0.105)			
	SWE2	6	1952-1968	0.006	0.049	1.105*	1.105*	0.91	0.703	0.955
				(0.061)	(0.754)	(0.038)	(0.038)			
	SWE3	13	1970-2010	3.644†	0.552	1.004	1.088*	1	0.397	0$^+$

				(0.353)	(0.133)	(0.021)	(0.02)			
Switzerland	SWZ1	16	1947-2007	0.004	0.04	1.109*	1.109*	1	1	0.442
				(0.035)	(0.494)	(0.023)	(0.023)			
Ukraine	UKR1	1	1994	0.117	0.249	1.23*	1.239*	0.911	0.461	0.374
				(0.352)	(0.631)	(0.075)	(0.071)			
	UKR2[b]	2	1998-2002	4.257†	0.593	0.993	1.518*	0.981	0.097	0$^+$
				(0.549)	(0.171)	(0.076)	(0.056)			
	UKR3[a]	2	1998-2002	0.002	0.028	1.337*	1.337*	0.991	0.176	0.666
				(0.022)	(0.364)	(0.072)	(0.072)			
	UKR4	2	2006-2007	3.422†	0.572	1.003	1.356*	0.963	0	0$^+$
				(0.456)	(0.155)	(0.054)	(0.046)			

Estimates of expected national electoral threshold, $\bar{\theta}(\hat{\theta}, \hat{\sigma})$, threshold standard deviation, $\bar{\sigma}(\hat{\theta}, \hat{\sigma})$, and disproportionality, $\hat{\beta}$. Estimates of disproportionality for models with no threshold, $\hat{\beta}^o$. Standard errors in parentheses.

a SMD partition of mixed system.

b PR partition of mixed system.

† Larger than zero at the 5% level of significance (one-tailed test).

* Not statistically different than three (3) at the 5% level of significance (two-tailed test).

* Larger than one at the 5% level of significance (one-tailed test).

$+$ Comparison test favors model with threshold at 5% level of significance.

I Electoral Results and Institutions Data Codebook

The data include information on votes and seats of all parties or coalitions of parties contesting elections, as well as the institutional characteristics that govern the allocation of seats for a given election. Each country data file contains up to five worksheets: *Parties*, *Coalitions*, *Votes*, *Seats*, and *Institutions*.

I.1 Electoral Data

These data are recorded in worksheets: *Parties*, *Coalitions*, *Votes*, *Seats*. The worksheet *Parties* includes the following variables:

PARTY or COALITION NUMBER: Number that identifies a party or coalition in the data set. Includes parties that did not contest elections on their own but in an electoral coalition including other parties. Coalitions are listed at the bottom of the worksheet.

ENGLISH NAME: Name of the party in English. If not available from original sources, name corresponds to the literal translation of the native name.

NATIVE NAME: Native name.

The *Coalitions* worksheets detail the composition of electoral coalitions (if present).

The *Votes* and *Seats* worksheets contain votes and seats (respectively) of parties or coalitions (corresponding to rows) per election (corresponding to columns).

I.2 Primary Institutional Variables

The *Institutions* worksheet records electoral system information per election (corresponding to columns as in the *Seats* and *Votes* worksheets). Each row corresponds to a variable, starting with election date information and continuing with institutional variables (either primary or derivative) which we detail below.

SYSTEM: Electoral systems per country are enumerated starting with 1 for the first system encountered in the data. A system persists if

- There are no changes in the allocation formula.[2]

---

[2]There is one exception in the data: A nominal change in allocation formula for the highest

- No other primary institutional variable (number of districts, number of seats, or numerical thresholds) changes by more than 5% of the system's average (that is if $I_1, \ldots, I_t$, and I_{t+1} reflect consecutive values, a change occurs at $t+1$ whenever $\frac{\max\{\frac{\sum_{t'=1}^{t+1} I_{t'}}{t+1}, I_{t+1}\}}{\min\{\frac{\sum_{t'=1}^{t+1} I_{t'}}{t+1}, I_{t+1}\}} > 1.05$).

COARSE SYSTEM: Alternative (coarser) definition of electoral system enumerated starting with 1 for first system encountered in the data. A system persists if

- There are no changes in the allocation formula, even if changes occur in other coded variables (numerical thresholds, number of districts, or number of seats).

FINEST SYSTEM: Alternative (finer) definition of electoral system enumerated starting with 1 for first system encountered in the data. A system persists if

- There are no changes in the allocation formula.

- There are no changes in numerical thresholds, number of districts, or number of seats.

TOTAL SEATS: Total number of seats to be allocated, absent provisions for seats to be added (to ensure proportionality or legislative majority) or that remain unallocated (due to failure of turnout or threshold provisions).

MAJORITY/PLURALITY BONUS: Binary variable. Takes the value 1 if national tier allocation reserves extra seats for the party with a majority or plurality of votes.

---

tier in the third Swedish system from Modified Saint Lague to Saint Lague is ignored as the divisor sequence that is actually in use at this tier is identical between the two versions of these two allocation rules.

I.2.1 First tier variables

NSEATS: Maximum number of seats out of TOTAL SEATS potentially allocated in the first tier.

NDISTRICTS: Number of districts in first tier.

NTHRES: National vote share that must be exceeded to take part in first tier allocation in any district.

DTHRES: District vote share that must be exceeded to take part in first tier district allocation.

ALLOCATION FORMULA: Formula used for first tier allocation. Possible values are

- One of **Plurality, Majority runoff, Majority plurality**: Possible Majoritarian allocation formulas.

- **STV**: Single transferable vote method.

- One of **Hare, Droop, Hagenbach-Bischoff, Imperiali, $\frac{1}{n+3}$ quota**: Respective quota rule with remainders reserved for allocation in an upper tier.

- One of **Hare, Droop, Hagenbach-Bischoff, Imperiali, $\frac{1}{n+3}$ LR**: Respective quota rule with largest remainders method applied to allocate seats that cannot be allocated on the basis of the quota.

- One of **D'Hondt, Saint Lague, Modified Saint Lague, Modified D'Hondt**: Highest divisors methods with different divisor sequences.

- Hybrids: Combinations of the above formulas, e.g., **Hare LR/Majority**.

I.2.2 Tier k=2, 3, 4

Up to 3 additional tiers of allocation are coded. Higher tiers are assigned a value $k = 2, 3, 4$, with higher values reserved for coarser geographic partitions of the country (e.g., in a country with three increasingly coarser tiers of allocation, the first tier may involve allocation at the level of the district, tier $k = 2$ at the level of the state or region, and tier $k = 3$ at the level of the country). If allocation occurs at the same level of aggregation but with two different modes or rules, these tiers are assigned consecutive levels in arbitrary fashion. For example, the first Greek electoral system (GRC1) involves district allocation (tier 1), regional allocation of remainders from tier 1 (tier 2), national allocation of remainders from tier 2 (tier 3), and national allocation of reserved seats in an "at-large" national district (tier 4), so that tiers 3 and 4 apply to the exact same geographic unit.

For each of the possible tiers $k = 2, 3, 4$, the following variables are coded:

TIERk: Binary variable; takes the value 1 if there exists a k-th tier, 0 otherwise.

NDISTRICTS: Number of districts in tier k.

ALLOCATION FORMULA: Formula used for tier allocation. Takes the same possible values as in the first tier formula.

FIXED SEATS: Binary variable. Takes the value 1 if a predetermined number of seats are reserved for allocation in this tier, 0 otherwise.

NFSEATS: Predetermined number of seats to be allocated in this tier (if FIXED SEATS = 1).

NATIONAL ACCESS THRESHOLD: National share of vote that is sufficient to take part in this tier's allocation.

TIER 1 DISTRICT ACCESS CONDITION: Lowest tier district vote share condition that is sufficient for access to this tier's allocation.

OTHER ACCESS CONDITION: Other condition that is sufficient for access to this tier's allocation.

J Electoral Results Sources

We compile electoral returns data using information from online official sources or databases of electoral results. We also relied on printed sources, notably Nohlen and Stöver (2010) and Mackie and Rose (1991), for older elections. If there were discrepancies between official country sources and the data from these books, we use the information from official sources. The following is the list of online sources.

J.1 Several Countries

Election Resources on the Internet by Manuel Alvarez Rivera. http://electionresources.org/

European Election Database. Norgewian Social Science Data Services. http://www.nsd.uib.no/european_election_database/

PARLINE Database on National Parliaments. http://www.ipu.org/parline-e/parlinesearch.asp

Psephos Adam Carr's Election Archive. http://psephos.adam-carr.net/

J.2 Country-specific

Belgium. Federal Portal

> 2007 Election Results. http://polling2007.belgium.be/en/

> 2010 Election Results. http://polling2010.belgium.be/en/

Bulgaria. Electoral Commission of Bulgaria. http://rezultati.cik2009.bg/results/proportional/rik_00.html

Czech Republic. Czech Statistical Office. http://www.volby.cz/pls/ps2010/ps2?xjazyk=EN

Croatia. State Electoral Commission of the Republic of Croatia. http://www.izbori.hr/izbori/ip.nsf/wpds/A51BF17BD1E3B7D5C125742000368C5D?open&1

Cyprus. Parliament of Cyprus. http://http://www.parliament.cy/

France. Ministry of Interior. http://www.interieur.gouv.fr/Elections/Les-resultats/Legislatives/

Germany. The Federal Returning Officer. http://www.bundeswahlleiter.de/en/bundestagswahlen/

Italy. Ministry of Interior. http://elezionistorico.interno.it/index.php

Macedonia. State Election Commission http://http://217.16.84.11/Default.aspx

Moldova. E-Democracy. http://www.e-democracy.md/elections/parliamentary/

Portugal. National Electoral Commission. http://eleicoes.cne.pt/index.html

Romania. Romanian Electoral Commission. http://www.becparlamentare2008.ro/

Spain. Ministry of Interior. http://www.infoelectoral.interior.es/min/

K   Data Issues – Country Notes

In cases when countries have adopted mixed electoral systems with two distinct partitions of the electorate into districts, two separate ballots cast in each election, one for each partition, and a separate (not necessarily independent) allocation of seats within each partition, we treat the mixed system as one electoral system using the PR vote as an input in the German and Hungarian cases in which allocations across the two partitions are not independent.[3] In all other cases, we estimate separate electoral system parameters for each partition, provided that within that partition the allocation of seats takes the ballot in that partition as the sole input, independent of the vote outcome in the other partition.

When parties or candidates lumped in the 'others' category earn seats, we amend the definition of the upper bound on the realized threshold for election t in (3) to the more conservative

$$
(10) \qquad\qquad \bar{v}_t := \min_{i \in \{0,1,\dots,N\}} \{ v_{t,i} \mid s_{t,i} > 0 \},
$$

where $v_{t,0} := 1 - \sum_{i=1}^{n} v_{t,i}$ and $s_{t,0} := S_t - \sum_{i=1}^{n} s_{t,i}$, the vote share and seat number, respectively, of parties or candidates not separately reported in electoral returns. A number of electoral systems make special provisions to ensure the representation of minorities, or special overseas districts (e.g., Finland, Italy, Croatia, Romania, Slovenia). In these cases, we faced the choice whether or not to include these seats and corresponding votes (if separately reported) as part of the estimation. Because these seats are typically allocated using

---

[3]A third case of mixed system with a similar dependency structure between the two partitions is the Italian 'scorporo' system. We only estimate the majoritarian partition of that system (ITA5) primarily due to data difficulties identifying parties competing across the two partitions.

special provisions, their inclusion in the analysis can lead to erroneous conclusions about the nature of competition induced by the electoral system. This is especially the case when very small minorities are guaranteed representation, thus forcing an inference that the electoral threshold in effect for all parties is small. We have made efforts to exclude electoral returns data from special or singular districts reserved for minority or overseas seats for which special electoral provisions apply. However, ambiguities regarding these exclusions remain in some cases, primarily due to unavailability of disaggregated results by district or inconsistencies between different sources.

The following list states exclusions regarding special districts, the treatment of mixed electoral systems, and any residual ambiguities.

Bulgaria   Data include only proportional representation seats in 2009.

Croatia   Data exclude minorities and diaspora results. Independent components of mixed systems in 1992 and 1995 are treated as separate electoral systems.

Germany   Data exclude representatives from Berlin before unification. Data include proportional representation votes and total seats (i.e., including SMD seats).

Denmark   Data exclude representatives from Faroe islands and Greenland.

Finland   Data include district of Aland results.

France   Results for Metropolitan France up to 1988. Results including overseas territories starting with the 1993 election.

Hungary   Data include proportional representation votes and total seats (i.e., including SMD seats).

Italy   Data includes Valle D'Aosta results. For the years 1996 and 2001 in which a mixed electoral system is used, the data include only the single member district results. Overseas deputies are excluded.

Lithuania   Single member district component of mixed systems is excluded in all elections.

Macedonia   Overseas representatives are excluded. The mixed system in place in 1998 is treated as two separate electoral systems.

Montenegro   Ethnic minorities' representatives are excluded.

Poland   Minorities' representatives are excluded from seat allocation.

Portugal   Representatives from Europe and the rest of the World are included.

Romania   Overseas results are included for the fourth Romanian system. Minorities' seats are excluded in systems ROM2-ROM4.

Serbia   Minority seats are excluded.

Slovenia   Minority results are excluded.

Ukraine   Mixed system components are treated as separate systems.

References

Dempster, A. P., N. M. Laird and D. B. Rubin. 1977. "Maximum Likelihood from Incomplete Data via the EM Algorithm." *Journal of the Royal Statistical Society. Series B (Methodological)* 39(1):1–38.

Fraley, Chris and Adrian Raftery. 2007. "Bayesian Regularization for Normal Mixture Estimation and Model-Based Clustering." *Journal of Classification* 24:155–181.

Gelman, Andrew, John B. Carlin, Hal. S. Stern and Donald B. Rubin. 2004. *Bayesian Data Analysis.* 2 ed. Boca Raton, Florida: Chapman & Hall/CRC.

Jamshidian, Mortaza and Robert I. Jennrich. 2000. "Standard Errors for EM Estimation." *Journal of Royal Statistical Society* 62(2):257–260.

Mackie, Thomas and Richard Rose. 1991. *The International Almanac of Electoral History.* London: Macmillan.

Nohlen, Dieter and Philip Stöver. 2010. *Elections in Europe: A Data Handbook.* Berlin: Nomos Verlagsgesellschaft.

Spiegelhalter, David J., Nicola G. Best, Bradley P. Carlin and Angelika van der Linde. 2002. "Bayesian measures of model complexity and fit (with discussion)." *Journal of the Royal Statistical Society, Series B (Statistical Methodology)* 64(4):583–639.