A Dynamic Duverger's Law

Jean Guillaume Forand* and Vikram Maheshri†

September 8, 2015

Abstract

Electoral systems promote strategic voting and affect party systems. Duverger (1951) proposed that plurality rule leads to bi-partyism and proportional representation leads to multi-partyism. We show that in a dynamic setting, these static effects also lead to a higher option value for existing minor parties under plurality rule, so their incentive to exit the party system is mitigated by their future benefits from continued participation. The predictions of our model are consistent with multiple cross-sectional predictions on the comparative number of parties under plurality rule and proportional representation. In particular, there could be more parties under plurality rule than under proportional representation at any point in time. However, our model makes a unique time-series prediction: the number of parties under plurality rule should be less variable than under proportional representation. We provide extensive empirical evidence in support of these results.

1 Introduction

Duverger's 'law' that plurality rule leads to two-party competition (Duverger (1951)) and its complementary 'hypothesis' that plurality rule with a runoff and proportional representation

*Assistant Professor, Department of Economics, University of Waterloo, Hagey Hall of Humanities, Waterloo, Ontario, Canada N2L 3G1. Office phone: 519-888-4567 x. 33635. Email: jgforand@uwaterloo.ca.

†Assistant Professor, Department of Economics, University of Houston, McElhinney Hall, Houston, Texas 77204-5019. Office phone: 713-743-3833. Email: vmaheshri@uh.edu. We thank Catherine Bobtcheff, Daniel Diermeier, seminar audiences at Carleton, UQAM and Ryerson, conference participants at CPEG 2013, Elections and Electoral Institutions IAST 2014, MPSA 2014, and Scott Legree for excellent research assistance.

favor multi-partism (see Benoit (2006) and Riker (1982)) have stimulated a large body of game-theoretic (e.g., Cox (1997), Feddersen (1992) and Palfrey (1989)) and empirical (e.g., Cox (1997), Lijphart (1994) and Taagepera and Shugart (1989)) research. Duverger's arguments highlight the role of strategic voting (the psychological effect) that is generated by the electoral formula that translates votes into seats (the mechanical effect). The combination of these effects provides strong incentives for voters to coordinate to winnow down the set of viable alternatives. Formal models detailing this phenomenon have been static, considering voters' and politicians' incentives only in a single election. The corresponding empirical work has focused on cross-sectional analysis of either the number of parties competing in national elections or of the number of competitive candidates at the district level.

Empirically, however, most party systems are not stable over time: there is substantial longitudinal variation in active parties in a country irrespective of its electoral system. This observation leads Chhibber and Kollman (1998) to argue that when "accounting for changes in the number of national parties over time within individual countries, however, explanations based solely on electoral systems [...] are strained. These features rarely change much within countries, and certainly not as often as party systems undergo change in some countries." As important features of political environments evolve over time, changes in the number of parties over time should be expected: an issue that existing parties have difficulty capturing can become salient, giving a new party an opportunity for entry, or an existing party can be discredited by scandal, which can lead to the disbanding of this party or its replacement by a new alternative. Nevertheless, this remark leaves open the possibility that different electoral systems endogenously induce systematically different party system dynamics.

In this paper, we present a novel empirical finding that relates the entrance and exit of parties at the national level to a country's electoral system: in a panel of 44 democracies since 1945, we find that countries with less proportional electoral systems tend to experience less entry of new parties and less exit of existing parties.[1] This result highlights the relative flexibility of party systems under more proportional electoral rules: opportunities for party formation are more easily grasped, and party system realignments through party mergers and alliances occur more often. Our finding that more new parties enter under more proportional electoral systems reflects the existence of higher barriers to party entry under plurality rule, which is consistent with existing empirical results and hence not necessarily

[1] Our data come from the Constituency-Level Elections (CLE) Dataset (Brancati (2013)).

surprising. However, our results on the exit of existing parties are more subtle and show that party systems under plurality rule are more persistent: currently active parties are more likely to compete in future elections. This implies that a party under plurality facing unfavourable circumstances in current elections is less likely than a comparable party under a more proportional system to respond by disbanding or entering into an alliance/merger with another ideologically compatible party.

We argue that our finding that parties exit less often under plurality cannot be explained by the static effects underlying Duverger's law and hypothesis, but that it can be rationalized through the dynamic incentives of party decision-makers that are generated by these standard effects. In particular, existing models of plurality elections show that a party with a small expected support is strategically abandoned by its supporters, so that its vote share is substantially lower than it would be if voters expressed themselves sincerely.[2] When models allow for costly participation by parties, these expected losers fail to compete in elections (Feddersen et al. (1990), Osborne and Slivinski (1996)), which suggests that existing small parties should be more likely to disband under plurality rule. Furthermore, if voters use past elections to help coordinate in current elections, the higher incentives for coordination under plurality rule generate barriers to entry by new parties. Our key observation, which is new to the literature, is that if forward-looking party leaders and supporters value the possibility of a party sharing their aims reemerging in the future if they disband an existing party, then these future barriers to entry under plurality rule will generate current barriers to exit. Hence, an existing party generates an option value for those that support it that is lower under more proportional systems in which new participants are more easily admitted into party systems. While this argument is theoretically straightforward, it does imply rather subtle reasoning from party leaders and activists, which makes the fact that it finds empirical support all the more striking.

To build on the intuition from above, we develop a simple dynamic model of party competition in Section 2. In the model, parties function as vehicles to promote the preferred policies of long-lived and ideologically motivated interest groups. Parties are formed, maintained, and possibly disbanded by their interest groups. Supporting a party is costly as

[2] In Myerson and Weber (1993) and Palfrey (1989), an expected loser gets no votes, while in Myatt (2007), since voters face aggregate uncertainty, an expected loser is hurt by voters' coordination but still receives some votes

it requires the resources necessary to run a serious campaign: recruiting good candidates, mobilizing party volunteers and raising advertising funds. In view of the critique of Chhibber and Kollman (1998), the key dynamic ingredient of the model is a stochastic political environment: for any number of reasons, the support garnered among the voters by the various policies preferred by the interest groups evolves over time. It follows that interest groups' incentives to support parties to represent them also evolve, so that interest groups whose policy goals are currently out of favor with voters may disband an existing party in the hopes of forming a new party in the future when voters become more receptive.

To focus on party formation and maintenance decisions, we simplify our treatment of elections by modelling them as probabilistic. First, the political environment determines the current support for possible policies. Second, support for policies is transferred to active parties who champion policies that receive the support for nearby policies that are not represented by a party. Finally, party supports are mapped into probabilities of winning the election: different electoral systems correspond to different contest success functions. Under proportional representation, a party's probability of winning is derived from its support in an unbiased way. Plurality rule differs from proportional representation through two coordination costs that are imposed on the probabilities that parties win when voters have the most incentive to coordinate (that is, when more than two parties contest an election). Under plurality rule, a party with a small expected support in the current election suffers a *minority penalty* to its probability of winning the election. And given expected support, a newly-formed party under plurality rule suffers an *entry penalty* to its probability of winning.

Minority penalties under plurality rule are motivated by the static mechanical and psychological effects. While there is some debate on whether these effects can be separately identified (see Benoit (2002)), the importance of their combined effect has been extensively documented, both at the country level (see Blais and Carty (1991), Lijphart (1994), Neto and Cox (1997), Ordeshook and Shvetsova (1994) and Taagepera and Shugart (1989)) and at the electoral district level (see Cox (1997) and Fujiwara (2011)). As noted by Cox (1997), in an equilibrium in which the voters of a district with magnitude M coordinate onto at most one non-winning alternative, the ratio of votes for candidates with ranks $M + 2$ and $M + 1$ should be zero. Interestingly, Cox (1997) finds that the proportion of districts with electoral outcomes approaching this 'Duverger' outcome shrinks as the district magnitude M increases, suggesting that the incentives promoting, and/or the effectiveness of, strategic voting is reduced under more proportional electoral systems. The evidence supporting the comparative importance of strategic voting under plurality rule also supports our assumption of entry penalties: if past voting behavior facilitates coordination, then new parties are comparatively penalized under plurality rule. Indeed, a recent paper by Anagol and Fujiwara (2015) documents a related 'runner-up effect' for individual candidates under plurality rule: second-place finishers are more likely to run in, and win, subsequent elections than third-place finishers, which they attribute to past electoral results resolving coordination problems for voters in current elections.[3]

Notably, our model does not make unambiguous cross-sectional predictions about the relationship between the number of competing parties and the disproportionality of electoral systems. Under proportional representation, we study an equilibrium in which interest groups respond closely to changes in their current political circumstances by disbanding the parties they support in unfavorable political environments and forming new parties as soon as the environment becomes more favourable. Under plurality rule, we derive two equilibria. In the first, the option value of being represented by a party is high enough that interest groups maintain an existing party through hard times, so that, in all elections, there are at least as many active parties as under the equilibrium under proportional representation. In the second, minority penalties dominate option values and push interest groups to disband existing parties in unfavourable environments. This leads to a 'Duverger' equilibrium with two parties competing in all elections, although their identities vary with the political environment. However, both equilibria under plurality rule feature less longitudinal variation in the number of active parties than the unique equilibrium under proportional representation. We provide robust empirical support for this finding in Section 3.

In the terminology of Shugart (2005), ours is a 'macro level' study in that we focus on parties' entry and exit decisions in elections to the national parliament. This aggregation is necessary, and our hypothesis cannot be evaluated at the electoral district-level: a serious party either participates in elections in a large number of districts or risks failing to be considered as a legitimate national party. In fact, Fujiwara (2011) demonstrates this when he finds that the electoral system (plurality versus plurality with a runoff) has no impact on the identities of the parties competing for the mayoralty of Brazilian cities. He attributes this to the fact that serious candidates are affiliated to a major national party, and all serious

[3]See references therein for related finding in experimental settings.

national parties field candidates in most mayoral elections. It has long been noted that the results of Duverger (1951) are naturally established at the district level, and that his arguments establishing the 'linkage' of electoral systems' effects on the number of parties at the district level with the number of parties on the national stage are incomplete (see Cox (1997)). While a growing number of empirical studies address this linkage problem (see Chhibber and Kollman (1998), Chhibber and Murali (2006), Cox (1997)), theoretical investigations of Duverger's results have mostly focused on a single electoral district. In an important exception, Morelli (2004) shows that Duverger's predictions can be reversed in a multi-district setting if there is enough heterogeneity across districts. While stylized, our model provides a novel possibility result: even abstracting from the linkage problem and considering a single district, the cross-sectional predictions of Duverger can be reversed solely due to the dynamic incentives of parties' supporters. Another key element of this paper is that we recover a clear comparative time series prediction.

While Duverger (1951) couched his arguments in dynamic terms, intertemporal approaches to the study of comparative political systems are rare. Cox (1997) highlights the importance of the dynamic incentives of parties and politicians for understanding the limits to Duverger's predictions, but he does not propose a particular model. Fey (1997) studies a dynamic process involving opinion polls to show that non-Duverger equilibria of the standard static model are unstable. Anagol and Fujiwara (2015) introduce a static model of plurality rule elections in which a public signal about parties' popularity proxies for past electoral histories. We are not aware of any other theoretical paper embedding the study of the number of parties in a dynamic framework. Some recent empirical studies have focused on the dynamics of the number of parties. Chhibber and Kollman (1998) show that in the United States and India, the number of parties decreased in periods in which the central government assumed a larger role. This result, which compares countries with plurality elections, is focused on providing conditions which support the linkage from district to the national level. Reed (2001) provides evidence that at the district level elections became increasingly bipartisan in Italy following a change of voting rule in 1993. However, Gaines (1999) finds little evidence of a trend towards local two-partism in a longitudinal analysis of Canadian elections (see also Diwakar (2007) for the case of India). The findings of Mershon and Shvetsova (2013), who establish that sitting legislators switch parties less often in single-member districts, can be interpreted, as we do for our results, as evidence that more proportional electoral systems are more adaptable to changing political circumstances. However, the mechanism underlying their results is quite different: while they focus on voters that value the predictability of politicians that maintain their party affiliation, along with the greater accountability of individual politicians when candidates are not selected through party lists, our results are driven by the differences in the aggregate electoral outcomes of third parties under more or less proportional electoral systems.

2 The Dynamics of Party Entry and Exit: Model

2.1 Setup

Elections are held in each period $t = 1, 2, ...$, after which the winning party selects a policy $x^t \in \{x_{-1}, x_0, x_1\}$, where $x_{-1} < x_0 < x_1$. A party j can be of one of three types in $\{-1, 0, 1\}$ (e.g., left, middle or right). Parties are formed and maintained by policy-motivated interest groups. Specifically, there are two long-lived interest groups of type -1 and 1, and in each period they simultaneously decide whether or not to support a party of their type to represent them. We make two simplifying assumptions that allow us to focus on the incentives of these two non-centrist interest groups to form, maintain and disband parties. First, we assume that parties cannot commit to implement any policy other than their preferred policy: if in power, party j implements policy x_j. Second, we assume that a party of type 0 is present in all elections.[4] This simple two-player environment still allows for rich dynamics for party entry and exit as well as for party structures that can feature one, two or three parties in any given election. The electoral rule, which we detail below, is either plurality rule or proportional representation.

At the beginning of each period, a *preference state* $s^t \in \{s_{-1}, s_0, s_1\}$ is randomly drawn. Preference states capture variability in the political environment, which is, by definition, absent from static models. We assume that preference states are independently and identically distributed across periods: let $Pr(s^t = s_0) = q$ and $Pr(s^t = s_1) = Pr(s^t = s_{-1}) = \frac{1-q}{2}$ for $q \in (0, 1)$.[5] Preference states have a straightforward interpretation: in state s_j, the party

[4]An alternative would be to assume that, say, a party of type 1 is present in all elections, and that two left of centre interest groups, of types -1 and 0, decide whether or not to support parties in each election. Such a model is almost equivalent to our specification and would yield closely related results.

[5]We could allow for persistence in electoral states, although this would add computational complexity

representing interest group j is favored by voters. Specifically, define $\overline{p}$, p and $\underline{p}$ such that $1 \geq \overline{p} > p > \underline{p} \geq 0$ and $\overline{p} + p + \underline{p} = 1$. Then for the two non-centrist policies, we can define the *policy support of* x_j in period t as

$$p_j^t = \begin{cases} \overline{p} & \text{if } s^t = s_j, \\ \underline{p} & \text{if } s^t = s_{-j}, \\ \frac{\overline{p}+\underline{p}}{2} & \text{if } s^t = s_0. \end{cases}$$

Note that this implies that when the voters have non-centrist preferences (i.e., $s^t \in \{s_{-1}, s_1\}$), the policy support of the centrist policy x_0 is p. Also, note that, for any preference state s^t, $p_{-1}^t + p_0^t + p_1^t = 1$.

While p_j^t is a measure of the popularity of policy x_j in the election at time t, this policy may not be championed by a party if the interest group of type j does not support a party. Conversely, a party championing policy x_j may have an expected support in excess of the support of its policy x_j since it may draw support from voters whose preferred policy is not championed by a party at t. A *party structure* ϕ^t lists the non-centrist parties supported by their interest groups in the current election: formally, $\phi^t \in 2^{\{-1,1\}}$. If a party supported by a non-centrist interest group is active under ϕ^t, then we define its *party support*, P_j^t, as equal to p_j^t, the support for policy x_j. If instead this interest group fails to support a party at t, the centrist party 0 collects the support of policy x_j. Specifically, we define the support of party 0 under ϕ^t as

$$P_0^t = p_0^t + p_{-1}^t \mathbb{I}_{-1 \notin \phi^t} + p_1^t \mathbb{I}_{1 \notin \phi^t},$$

where $\mathbb{I}$ is the indicator function

The legislative power of interest groups depends on the support garnered by their preferred policies among the voters and on whether or not they are represented in elections by a party, but it is also mediated by the electoral system. A main challenge we face is finding a model that captures both plurality rule and proportional representation yet remains tractable when embedded in a dynamic model of party entry and exit. The most thorough approach would model voter's choices explicitly and have policy outcomes be determined by legislative bargaining after elections (see Austen-Smith and Banks (1988), Austen-Smith

without affecting our central conclusions. Likewise, the simplifying assumption that non-centrist preference states s_1 and s_{-1} occur with equal probability allows us to exploit symmetry, but it is not essential.

(2000), Baron and Diermeier (2001) and Indridason (2011)), but these models are complex even when restricted to the study of a single election. Conceptually, we can represent plurality and proportional electoral systems as different mappings from the distribution of voter support for parties into the distribution of seats in the legislature and corresponding policy outcomes (see Faravelli and Sanchez-Pages (2012) and Herrera et al. (2012)). On average, legislative policy outcomes under proportional representation should be more representative of voters' views as expressed by vote shares, whereas policy outcomes under plurality rule are more heavily tilted towards the views of plurality voters. We model this mapping in a reduced form with a probabilistic voting approach that maps the party supports of active parties into these parties' probabilities of winning the election and implementing their ideal policies, which we interpret as obtaining decisive power in the legislature.[6] We recognize that our approach presents an incomplete view of legislative policy-making, but our goal is to construct a minimal dynamic model of elections that is consistent with observed patterns in party entry and exit.

Under proportional representation, we assume that the *probability of winning* of any active party j is its support P_j^t. Under plurality rule, we assume that the stronger incentives for strategic voting give rise to coordination costs for those elections in which all three parties compete. Specifically, we assume that these costs are borne by both small existing parties and new parties of all sizes. First, a non-centrist party that is active at t when the preference state is s_{-j} and party $-j$ is also active bears a *minority penalty* of $\alpha \geq 0$ to its probability of winning. As discussed earlier, this cost is generated by both the mechanical effect of the electoral formula and the psychological effect of strategic voting as highlighted by Duverger (1951) and the extensive theoretical and empirical literatures that followed. Second, in any preference state at t, if a non-centrist interest group j forms a new party and party $-j$ is active in both the election at $t-1$ and t, then party j bears an *entry penalty* of $\beta \geq 0$ to its probability of winning. This dynamic effect increases incentives for strategic voting under plurality and is consistent with the recent empirical findings of Anagol and Fujiwara (2015). Our key postulate is that the coordinating effect of a party's past electoral activity, which acts as a barrier to entry, is weaker under proportional representation. Finally, note

[6] See Crutzen and Sahuguet (2009), Hamlin and Hjortlund (2000), Ortuno-Ortin (1997) and Myerson (1993) for related reduced-form treatments of post-election legislative arrangements under proportional representation.

that because both the minority and entry penalties suffered by party j are motivated by the coordination problems that voters face under plurality when choosing between more than two parties, both penalties are conditioned on all three parties being active at t.

In our model, the coordination costs α and β alone distinguish plurality rule from proportional representation. Specifically, under plurality rule, fix time t and suppose that the party structure in the current election is such that $\phi^t = \{-1, 1\}$. Then non-centrist party j wins with probability

$$P_j^t + \alpha \left[\mathbb{I}_{s^t = s_j} - \mathbb{I}_{s^t = s_{-j}} \right] + \beta \left[\mathbb{I}_{j \in \phi^{t-1}} \mathbb{I}_{-j \notin \phi^{t-1}} - \mathbb{I}_{j \notin \phi^{t-1}} \mathbb{I}_{-j \in \phi^{t-1}} \right].$$

Meanwhile, if $\phi^t = \{j\}$, then party j wins with probability P_j^t. To ensure that active parties have non-negative winning probability in all states, we assume that $\alpha + \beta \leq \underline{p}$. Note that our formulation assumes that any coordination costs imposed on party j benefit only party $-j$. This implies that party 0 wins with probability P_0^t in any preference state under plurality rule.[7]

We do not rule out the possibility that there could be many reasons that an electorally unsuccessful party is maintained: to put pressure on more important parties in the legislature, or to keep afloat a party organization that brings benefits unrelated to electoral outcomes (employment for party workers, bribes for legislators, state subsidies for electoral participation). That a party's benefits from contesting an election are exactly its winning probability is a convenient normalization. However, we are implicitly making the assumption that a party's ability either to obtain or to profit from these non-electoral benefits are increasing in its success in both electoral systems. Hence, insofar as electoral systems affect probabilities of winning, we also assume that they influence the scale of parties' non-electoral benefits.

Supporting a party is costly for an interest group, although forming a new party is costlier than simply maintaining an existing party. Specifically, if $j \in \phi^{t-1}$, then the party maintenance cost to interest group j in the electoral cycle at t is $\underline{c}$. If instead $j \notin \phi^{t-1}$, then no party represented interest group j in the previous election, and the party formation cost at t to interest group j is $\overline{c} > \underline{c}$. Along with the variability of preference states, this wedge

[7] That the centrist party never benefits from the coordination costs imposed on minor parties is assumed for convenience and is not important for our results.

between $\overline{c}$ and $\underline{c}$, which indexes the opportunity cost of disbanding a party, generates an option value to existing parties for interest groups under both electoral systems. This option value is derived from the costs to party activities and does not depend on the electoral rule faced by the party. However, under plurality rule the opportunity cost of disbanding a party also includes future entry penalties, which generates a comparatively higher option value to an existing party. This increment in option value over proportional representation is driven by the advantages that established parties have over new entrants under plurality rule.

Interest groups are risk-neutral and have single-peaked preferences over feasible policies. A non-centrist interest group of type j has ideal policy x_j. Given any non-centrist interest group, let $\overline{u}$ be its stage payoff to its preferred policy, u be its stage payoff to its second-ranked policy, and $\underline{u}$ be its stage payoff to its third-ranked policy with $\overline{u} > u > \underline{u}$. Interest groups discount future payoffs by a common factor of δ and support parties to maximize their expected discounted sum of payoffs that consists of the expected difference between its benefits from the policy implemented by the winning party and party formation costs (where the expectation is over electoral outcomes).

2.2 Strategies and Equilibrium

We focus on Markov perfect equilibria in pure strategies in which interest groups condition their party formation and maintenance decisions at time t on the payoff-relevant *state* (s^t, ϕ^{t-1}) which encompasses the current preference state and the previous party structure. For a non-centrist interest group j, a strategy is given by $\sigma_j : \{s_{-1}, s_0, s_1\} \times 2^{\{-1,1\}} \to \{0, 1\}$, where $\sigma_j(s, \phi) = 1$ indicates that the interest group supports a party in preference state s given party structure ϕ inherited from past periods. Let $V_j(s, \phi; \sigma)$ denote the expected discounted sum of payoffs to interest group j under profile $\sigma \equiv (\sigma_{-1}, \sigma_1)$ conditional on state (s, ϕ). Profile σ^* is a *Markov perfect equilibrium* if, for all states (s, ϕ) and all profiles (σ_{-1}, σ_1),

$$V_{-1}(s, \phi; \sigma^*) \geq V_{-1}(s, \phi; (\sigma_{-1}, \sigma_1^*)) \text{ and}$$
$$V_1(s, \phi; \sigma^*) \geq V_1(s, \phi; (\sigma_{-1}^*, \sigma_1)).$$

Hereafter, the term equilibrium refers to Markov perfect equilibrium. Restricting attention to strategies in which interest groups condition only on payoff-relevant elements of histories of play limits the possibilities for intertemporal coordination between interest groups and hence refines our equilibrium predictions. It also ensures that equilibrium behavior in our model is relatively simple.

2.3 Results

The comparative equilibrium dynamics of party systems under both electoral systems depend on the values of the model's parameters: party formation and maintenance costs $(\overline{c}, \underline{c})$, coordination costs (α, β), and policy payoffs $(\overline{u}, u, \underline{u})$. For example, if $\overline{c} > \overline{u}$, then under both electoral systems no non-centrist party ever forms in any equilibrium. Conversely, if $\underline{c} = 0$ and $\underline{p} > 0$, then no existing non-centrist party is ever disbanded in any equilibrium in both electoral systems. Characterizing the full set of equilibria for all parameters is difficult: although our game is simple, its dynamic structure generates multiple equilibria and cumbersome equilibrium conditions. Our approach is to focus instead on a region of the parameter space which gives rise to equilibria with natural properties. We detail our assumptions and discuss our equilibrium selection below, but for now we note that we restrict attention to parameter values such that in the static stage game with preference state s_{-j}, interest groups of type j prefer to disband their party when anticipating that a non-centrist party j will contest the election: therefore, any equilibrium party maintenance by current minority interest groups is due solely to dynamic incentives.

We first present our results for proportional representation. Our aim is to show that lower coordination costs under proportional representation allow interest groups to better tailor their party formation and maintenance decisions to the current preference state. They can do so by supporting parties when voters' preferences favor their policy positions and disbanding parties when they do not. To this end, we introduce a strategy profile in which non-centrist interest groups support parties if and only if the current electoral state does not favor the interest group on the other side of the political spectrum. Specifically, define profile σ^{PR} such that for any non-centrist interest group j and party structure ϕ,

$$\sigma_j^{PR}(s, \phi) = \begin{cases} 1 & \text{if } s \in \{s_j, s_0\} \\ 0 & \text{if } s = s_{-j}. \end{cases}$$

Notice that under σ_j^{PR}, the party formation and maintenance decisions of interest group j

are independent of the party structure and, in particular, of whether or not interest group j was represented by a party in past elections. In the following result, we identify conditions under which the strategy profile σ^{PR} is an equilibrium under proportional representation. Furthermore, we show that under these same conditions no other equilibrium exists.[8]

Proposition 1. *Suppose that*

$$\overline{c} < \frac{1-\overline{p}}{2}[\overline{u} - u], \tag{1}$$

and that

$$\underline{c} > \underline{p}[\overline{u} - u] + \delta\frac{1+q}{2}[\overline{c} - \underline{c}]. \tag{2}$$

Then σ^{PR} is the unique Markov perfect equilibrium under proportional representation.

Condition (1) ensures that a non-centrist interest group j always supports a party in s_j and s_0, so the only remaining question is whether or not the interest group will support a party in s_{-j}. Note that under condition (2), $\underline{p}[\overline{u} - u] - \underline{c} < 0$, so in the stage game with preference state s_{-j}, interest group j prefers disbanding an existing party to maintaining it. However, maintaining an existing party in s_{-j} has an associated option value realized in s_j and s_0, which is derived from the cost savings for supporting a party in those states. Condition (2) ensures that under proportional representation, the immediate cost savings from disbanding an existing party dominates the option value of supporting it through an unfavorable election. Conditions (1) and (2) uniquely pin down the optimal party formation and maintenance decisions of both non-centrist interest groups so that no other equilibrium can exist. Also, note that while the equilibrium σ^{PR} is symmetric in strategies, we impose no ex ante symmetry restriction on equilibria.

We now turn to our results under plurality rule. Our aim is to show that in those regions of the parameter space identified in Proposition 1, the coordination costs imposed on parties under plurality rule lead interest groups' party formation and maintenance decisions to display more persistence than under proportional representation. Accordingly, we focus attention on strategy profiles in which interest groups support *existing* parties if and only if the preference state does not favor the interest group on the other side of the political spectrum. Contrary to the case of profile σ^{PR} under proportional representation, entry penalties induce interest groups to form *new* parties only when the preference state favors

[8]All proofs are in Appendix A.

them. Specifically, we restrict attention to profiles σ^{PL} with the property that for all non-centrist interest groups j,

$$\sigma_j^{PL}(s,\phi) = \begin{cases} 1 & \text{if } s = s_j, \text{ or if } s = s_0 \text{ and } \phi \neq \{-j\}. \\ 0 & \text{if } s \in \{s_0, s_{-j}\} \text{ and } \phi = \{-j\}. \end{cases} \tag{3}$$

The key question is whether interest group j supports an *existing* party when the preference state favors its opponent. On the one hand, minority penalties increase the cost of maintaining a party in unfavorable electoral circumstances. On the other hand, entry penalties increase the option value of a party that is maintained even through a string of lost elections. We consider two alternatives. Profile $\overline{\sigma}^{PL}$ denotes the strategy profile respecting (3) with maximal participation:

$$\overline{\sigma}_j^{PL}(s,\phi) = 1 \;\text{ if } s = s_{-j} \text{ and } j \in \phi,$$

while profile $\underline{\sigma}^{PL}$ denotes the strategy profile respecting (3) with minimal participation:

$$\underline{\sigma}_j^{PL}(s,\phi) = 0 \;\text{ if } s = s_{-j} \text{ and } j \in \phi.$$

In the following result, we identify conditions under which $\overline{\sigma}^{PL}$ and $\underline{\sigma}^{PL}$ are equilibria under plurality rule.[9] These conditions will depend on the entry penalty β being bounded above and below. These upper and lower bounds, denoted $\overline{\beta}$ and $\underline{\beta}$ respectively, are functions of all the parameters of the problem except the minority penalty α, and they are derived in Appendix A.

Proposition 2. *Suppose that* (1) *and* (2) *hold and that* $\beta \in (\underline{\beta}, \overline{\beta})$*. Then there exist* $\underline{\alpha}, \overline{\alpha} \in [0, \underline{p} - \beta]$ *such that* $\underline{\sigma}$ *is a Markov perfect equilibrium whenever* $\alpha > \underline{\alpha}$ *and* $\overline{\sigma}$ *is a Markov perfect equilibrium whenever* $\alpha < \overline{\alpha}$*. Furthermore,* $\underline{\alpha} \geq \overline{\alpha}$*.*

Our dynamic model provides no robust cross-sectional predictions on the number of parties under different electoral systems. In any given election under proportional representation, there could be either two or three parties competing (under σ^{PR}). Under plurality, our model allows for the standard Duverger prediction of a two-party system (under $\underline{\sigma}^{PL}$),

[9] Interest group j's actions are not yet specified if the preference state is s_{-j} and no interest groups supported parties in the previous elections (i.e., $\phi^t = \emptyset$). These histories only occur off the equilibrium path, and the details are in Appendix A.

although the identities of the parties change over time as voters' preferences evolve, but it also allows for a non-Duverger equilibrium in which three parties are always present (under $\overline{\sigma}^{PL}$). This last result may be surprising in itself: under plurality rule, parties face additional costs to participating in elections relative to proportional representation, yet in equilibrium they may contest more elections. Interest groups supporting minor parties are worse off under plurality rule and would find it optimal to disband them in a static setting. However, in a dynamic setting, forward-looking interest groups internalize the high opportunity cost under plurality of losing their vehicle for influencing policy.

Our results do support a dynamic prediction: there is greater variation in the number of active parties in equilibrium σ^{PR} under proportional representation than under either of the equilibria $\underline{\sigma}^{PL}$ and $\overline{\sigma}^{PL}$ that we identify under plurality. Intuitively, under $\overline{\sigma}^{PL}$, the option value of an existing party dominates the static costs due to minority penalties, so that parties are never disbanded in hard times and there is no variation in the number of parties, whereas under σ^{PR} there is frequent party destruction. Under $\underline{\sigma}^{PL}$, parties are disbanded when the preference state favors their opponents, yet there is less variability in the number of parties than under σ^{PR} since entry penalties lead to hesitation by interest groups to induce three-party competition, and hence to less party formation than under σ^{PR}. Specifically, under σ^{PR}, the expected number of changes to the party system in state s_0 is $1-q$, since a party enters whenever a transition to s_0 occurs from an extreme state. Meanwhile, under $\underline{\sigma}^{PL}$, the expected number of changes in the number of parties in state s_0 is 0 since no entry occurs in this state. Furthermore, under σ^{PR}, the expected number of changes to the party system in state s_j for $j \in \{-1, 1\}$ is $1 \cdot q + 2 \cdot \frac{1-q}{2} = 1$ since a single exit occurs when transitioning from s_0, and both an entry and an exit occur when transitioning from s_{-j}. Meanwhile, under $\underline{\sigma}^{PL}$, the expected number of changes to the number of parties in state s_j is $[0 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2}] \cdot q + 2 \cdot \frac{1-q}{2} = 1$ since when a transition occurs from s_0 to s_j, there is either (i) no change to the party system, if party j was active in the previous election, which occurs with probability $\frac{1}{2}$ (the probability that the last extreme state to be realized was s_j as opposed to s_{-j}), or (ii) both the entry of party j and the exit of party $-j$, if party $-j$ was active in the previous election.

Our model also supports novel predictions about the comparative persistence of party systems. Specifically, although preference states are drawn independently across periods, party structures under plurality rule are history-dependent whereas party structures under proportional representation are not. Under σ^{PR}, the probability that a party representing interest group j contests any election is $\frac{1+q}{2}$ (the probability that the preference state is either s_j or s_0), which does not depend on the realization of past preference states or of party structures. Under both equilibria under plurality, the probability that a party representing interest group j contests an election at time t depends on whether or not this party contested an election at time $t-1$. Under equilibrium $\overline{\sigma}^{PL}$, party structures are fully persistent as no party ever exits. Specifically, if $j \in \phi^{t-1}$, then party j contests an election at time t with probability 1. On the other hand, if $j \notin \phi^{t-1}$, then it contests the election with probability $\frac{1-q}{2}$, the probability that the preference state transitions to s_j. In the equilibrium $\underline{\sigma}^{PR}$, if $j \in \phi^{t-1}$, then party j contests the election at time t with probability $\frac{1+q}{2}$, the probability that the preference state is either s_j or s_0, whereas if $j \notin \phi^{t-1}$, then it contests the election with probability $\frac{1-q}{2}$.

While beyond the scope of this paper, the observation that parties are more persistent under plurality rule than under proportional representation could have implications for the normative comparison of electoral systems, which is typically organized around the trade-off between representation and accountability (Powell (2000)). The advantage of proportional representation in ensuring that the diverse opinions in the electorate are included in the legislative process is augmented by dynamic considerations: emerging constituencies are more likely to be represented by a new party than under plurality rule. However, the advantage of plurality rule in clearly attributing responsibility to office-holders is buttressed by parties' longevity: the more frequent realignment and relabelling of parties under proportional representation could further hamper voters' ability to hold politicians accountable.

To understand the conditions under which $\underline{\sigma}^{PL}$, or alternatively $\overline{\sigma}^{PL}$, are equilibria, consider interest group j in state $(s_{-j}, \{j\})$. Under $\underline{\sigma}^{PL}$, interest group j disbands its current party and waits until the preference state returns to s_j before forming a new party to represent it. However, since in that case interest group $-j$ will disband the party it forms in state $(s_{-j}, \{j\})$, interest group j faces no entry penalty when it forms a new party. Hence, $\underline{\sigma}^{PL}$ provides incentives for interest group j to disband its party in s_{-j} only if minority penalty α is sufficiently high to deter party maintenance. On the other hand, under $\overline{\sigma}^{PL}$ interest group j supports its party and bears the minority penalty, which cannot be too high in order to provide incentive for party maintenance. For a given minority penalty α, the two profiles cannot both be equilibria. The lower bound $\underline{\beta}$ on the entry penalty ensures that these costs are high enough to prevent interest groups that are not represented by a party in centrist state s_0 from forming a new party. Note that such histories occur on the equilibrium path only under $\underline{\sigma}^{PL}$. The upper bound $\overline{\beta}$ on the entry penalty ensures that these costs are low enough that, under $\overline{\sigma}^{PL}$, non-centrist interest group j is willing to form a new party in preference state s_j, in those histories off the equilibrium path in which this interest group is not represented by a party. Note that for such histories under $\underline{\sigma}^{PL}$, interest group j never bears entry penalties since no party representing interest group $-j$ ever contests elections in preference state s_j.[10]

3 The Dynamics of Party Entry and Exit: Emprical Findings

The key empirical prediction of our model is that that more disproportional electoral systems should experience less churn as parties are less likely to enter and exit elections in these systems. As noted earlier, existing empirical studies cannot be used to evaluate this observation, which is new to the literature. To that end, our goal in this section is to estimate the relationship between the disproportionality of electoral systems and the variability in the number of active parties using cross-country elections data. As we detail below, the correlations that we uncover are consistent with our theoretical findings, and they are strikingly robust. We face two main measurement issues: first, we require a concise measure of the proportionality of an electoral system, which is determined by institutional characteristics such as electoral laws in a potentially complex manner, and second, we require an appropriate and objective measure of party entry and exit.

[10]Condition (2) does not play a role in the proof of Proposition 2, but is included in order to establish that the equilibria $\underline{\sigma}^{PL}$ and $\overline{\sigma}^{PL}$ can exist under plurality under parametric restrictions that ensure that σ^{PR} is the unique equilibrium under proportional representation. That the conditions of Proposition 2 can be met for some parameter values can be shown by example: the neighbourhood of the point with $\delta \approx 1$, $\underline{c} = \overline{c} = \frac{1}{4}$, $\underline{p} = p = \frac{1}{4}$, $\beta = \frac{1}{9}$, $\overline{u} - u = 1$ and $u - \underline{u} = \frac{3}{2}$ contains an open set of parameters for which the conditions of Proposition 2 are met and for which $\underline{\alpha} < \underline{p} - \beta$ and $\overline{\alpha} > 0$, so that both $\overline{\sigma}^{PL}$ and $\underline{\sigma}^{PL}$ can be equilibria at those parameters, depending on the value of α.

3.1 Measuring the Proportionality of an Electoral System

It is well-known that no perfectly satisfactory measure of the proportionality of an electoral system exists. Our approach is not to choose among various candidates, but to present our results for three measures used in the extensive empirical literature on Duverger's Law. As we detail below, our qualitative results are mostly invariant to the choice of of any one of these measures. The simplest alternative is the binary measure of the electoral formula of a country's lower house, as proposed by Persson and Tabellini (2005). An election in which a country elects its lower house exclusively through plurarlity rule is coded as a 1, and otherwise, elections are coded as a 0. Following their suggestion, we obtain a single measure for a given country by averaging this binary variable over all observed elections.[11] We treat this, which we refer to as the *Majoritarian Dummy*, as our baseline measure of proportionality.

As an alternative, we follow Taagepera and Shugart (1989) and measure the proportionality of an electoral system by its *effective district magnitude*: the total number of legislators directly elected in electoral districts divided by the total number of electoral districts.[12] This measure is directly determined by a country's electoral institutions, and it is well established that more proportional electoral systems are associated with higher effective district magnitudes.

Finally, we supplement our analysis with a third measure of the proportionality of an electoral system by using the least squares index of Gallagher (1991). This index, which has been used in empirical analyses of electoral systems, is a measure of the difference between parties' vote and seat shares in a given election.[13] In perfectly proportional electoral systems, parties' seat shares should be identical to their vote shares, while in less proportional systems front-running parties typically have seat shares exceeding their vote shares and lagging parties have seat shares well below their share of the votes. Formally, for a given election t in country c with J_{ct} total parties, let p_{jct} be the vote share that party j receives, and let s_{jct} be the seat share that party j wins in the legislature. Then the *disproportionality*

[11] In all but one country in our sample, this binary variable does not change over time.

[12] Effective district magnitude can differ from average district magnitude, which is defined as the total number of legislative seats divided by the number of electoral districts. Taagepera and Shugart (1989) argue that effective district magnitude is the superior measure of the proportionality of an electoral system. To the extent that a legislature does not feature at-large seats, these measures are identical.

[13] See also Lijphart (1994) and Taagepera and Grofman (2003).

index for this election is given by

$$g_{ct} = \sqrt{\frac{1}{2}\sum_{j=1}^{J_{ct}} (p_{jct} - s_{jct})^2} \tag{4}$$

where g_{ct} is an index that ranges from 0 to 1 with increasing values corresponding to more disproportional elections. Because disproportionality is a property of the electoral institutions of country, it should should not vary either by electoral district or by election. Hence, we aggregate district electoral outcomes and compute the disproportionality index at the national level. Furthermore, we average the disproportionality index over all elections for a different country, i.e.,

$$G_c = \frac{1}{T_c}\sum_{t=1}^{T_c} g_{ce} \tag{5}$$

where T_c is the total number of elections that we observe for country c. G_c constitutes an alternative measure of the (dis)proportionality of the electoral system of country c. The use of this measure for our analysis warrants an important caveat: since the numbers of active parties in a given country enters into G_c, it is generally agreed that measures of district magnitude are cleaner proxies for electoral systems than disproportionality indices (see, for example, Ordeshook and Shvetsova (1994)).

3.2 Measuring the Entry and Exit of Parties

Finding a measure of party participation decisions using cross-country elections data is difficult, and on this we cannot be guided by the existing literature, in which no such measures have been proposed. A natural idea is to use the variability of the effective number of parties (Taagepera and Shugart (1989)) over time as an indicator of party entry and exit. The key problem with this procedure is that participation decisions are made at the party level, whereas the effective number of parties abstracts from party identity. For instance, suppose parties A and B each won half of the votes and seats in the previous election, then B disbanded and party C formed, and then A and C each win half of the votes in the current election. Then we would measure no change in the effective number of parties in spite of the fact that one entry and one exit clearly occurred. Another simple idea is to use the party identification information contained in electoral records to track individual parties across electoral histories. However, these records list dozens of fringe parties in each country, many collecting just a handful of votes across all districts. Therefore, we opt to identify competitive parties by using vote thresholds. This brings the additional difficulty of aggregating electoral results across a country's districts: electoral systems differ in their number of districts (with more proportional systems having less districts on average than plurality systems) and parties may be active in some districts and not others. This can be the case if, for instance, a party's support is regional in nature. Alternatively, a successful entry in a few districts may be a launching pad for a new national party.

We construct our measure of party entry and exit as follows. For any election t in country c, we denote the number of electoral districts D_{ct}, where district d contributes a fraction σ_{dct} of the total seats in the national legislature. A party is said to have *entered* in district d in election t if its vote share in that district in $t-1$ was less than some threshold λ and its vote share in that district in t was greater than λ. Party exit is defined similarly. Let n_{dct} and x_{dct} represent, respectively, the total number of entering and exiting parties in district d during election t in country c. The *total number of entries* N_{ct} in a given election is obtained by summing over all districts as

$$N_{ct} = \sum_{d=1}^{D_{ct}} n_{dct} \cdot \sigma_{dct}, \tag{6}$$

and the *total number of exits* X_{ct} can be defined similarly as

$$X_{ct} = \sum_{d=1}^{D_{ct}} x_{dct} \cdot \sigma_{dct}. \tag{7}$$

We weigh the number of entries in each district by that district's size in order to correct for the variability in the number of electoral districts across electoral systems. For example, Israel, which is considered to have an electoral system that is almost perfectly proportional, has a single electoral district, so one entry is recorded if a new party collects a share of λ of votes at the national level. The United Kingdom, on the other hand, has all legislators elected by plurality rule in over six hundred electoral districts, so that one entry is recorded if a new party collects a share λ of votes in every district. The emergence of a regional party that collects the threshold share of votes in, say, half of the country's districts, would be recorded as half an entry. In the absence of weighing district-level party entries and exits, the variability in party structures in plurality rule systems would be dramatically overstated.[14] Finally, the *total net party movements* in an election (i.e., the total amount of partisan churn), M_{ct}, is simply defined as the sum of entries and exits as

$$M_{ct} = N_{ct} + X_{ct}.$$

3.3 Sample

We construct these variables from the CLE, which contains detailed information on the identities of all parties that participated in a large number of elections in many countries since 1945.[15] For each party and election, the CLE documents the number of votes that each party received in each district of a given election and the number of legislative seats that they were awarded. With this information, it is straightforward to construct the measures described above. For consistency, we restrict our analysis to only elections in the CLE for which all three of our measures of electoral systems are available. A tabulation of these elections is given in Table 1, and summary statistics with a participation threshold of $\lambda = 5\%$ are presented in Table 2. Each of the 454 elections in our dataset features an average of 1.28 million votes cast for 208 seats across 80 districts. Each election features an average of 3.59 parties, 0.71 of which are new entrants (as defined above) and 0.72 of which are new exits (as defined above). It is useful to note the large standard deviations of all variables relative to the means. These reflect not only cross sectional variation in the dataset but also substantial longitudinal variation in the numbers of parties, entries and exits. Because all countries do not hold elections at the same frequency (and several countries were formed or

[14] A number of measures of party nationalization have been proposed for use in conjuction with the Constituency Level Elections Archive (CLEA) dataset. These measures are inappropriate for our analysis because each one fails to satisfy at least one of the following properties: (1) the measures account for the absolute popularity of a parties, (2) the measures vary by party, and (3) the measures can be compared over time. All three of these properties are necessary to compute measures of party entry and exit for our analysis.

[15] The CLE unfortunately does not contain data on all democratic elections since 1945. Indeed, no single source does. We use only those elections contained in the CLE for our analysis and do not supplement our dataset with data from other sources in order to maintin consistent reporting. We replicated our analysis using a similar (thought not identical) sample of elections from the CLEA data set and obtained similar results. We report results using only the CLE because this is the dataset that has been primarily used to construct disproportionality indices (Gallagher and Mitchell (2005)).

ceased to exist since 1945) our data set constitutes an unbalanced panel.

We illustrate the relationship between the three alternative measures of disproportionality in Figure 1. Countries that are classified under plurality rule by the binary measure are shown as solid dots, and those classified under proportional representation are shown as hollow dots. On the axes, we plot the effective district magnitude against the average disproportionality index for each of the the countries in our sample. Countries with plurality rule, as defined by the first measure, have very small effective district magnitudes and very high disproportionality scores. Moreover, countries with lower effective magnitudes are associated have higher disproportionality scores, as is well known.

3.4 Results

We present two main sets of empirical results. First, we conduct static tests of Duverger's Law that explore the relationship between the proportionality of electoral systems and the number of parties that compete in elections. These tests replicate the traditional results in the literature. Second, we conduct dynamic tests of Duverger's Law that explore the the relationship between the proportionality of electoral systems and the *change* in the numbers of parties that compete in elections. These novel dynamic tests constitute our main empirical results. In all of these tests, we use a participation threshold of $\lambda = 5\%$. We conclude by showing that our main results are robust to different choices of λ.

We estimate these relationships using four different categories of control variables:

1. Decade fixed effects, in order to control for slowly varying global determinants of partisan political activity.[16]

2. Regional fixed effects for European countries, African countries, and former republics of the USSR in order to absorb any regional determinants of political activity.

3. Flexible controls for the number of districts in an election.[17]

[16] Our decade dummies are defined for the periods 1940-49, 1950-59, ... , 2000-2009. We replicated our analysis defining decade dummies for all possible periods (e.g., 1948-1957, ...) and obtained results that were statistically indistinguishable from those presented.

[17] In the results presented, we include polynomials of all orders up to 6 in D_{ct} and $\log D_{ct}$. As a robustness check, we replicated our analysis with polynomials of all orders up to 10 and obtained qualitatively similar and precise estimates of our coefficients of interest.

4. Flexible controls for the number of parties in an election.[18]

For the static tests, we use the first three sets of control variables (the number of parties is the dependent variable in these tests), and for the dynamic tests, we use all control variables. We specify all continuous variables in logarithms for all tests.[19]

Table 3 contains results from the traditional, static tests of Duverger's Law. Using all three measures of proportionality, we uncover statistically significant relationships between proportionality and the number of parties that compete in elections, thus replicating known empirical results.

Tables 4-6 contain our main empirical results of the dynamic relationship between the proportionality of electoral systems and the number of parties that compete in elections. For each of these tables, the dependent variables are the total number of entries, the total number of exits and the total number of movements respectively. The explanatory variable of interest in all regressions is one of the three measures of proportionality. Because these measures do not vary within countries with fixed electoral systems by construction, we cluster our standard errors at the country level to account for any induced multicollinearity. Since we cannot measure entry and exit for the first election observed in each country, we estimate these regressions on a sample of 411 elections.

We consider three different specifications: (1) no control variables, (2) decade fixed effects, regional fixed effects and flexible controls for the number of districts, and (3) all of those controls plus flexible controls for the number of parties. The first specification provides a raw correlation between proportionality and party dynamics, and the remaining two specifications show that this correlation is not simply an artifact of a variety of confounders. We present results with and without flexible controls for the number of parties because the inclusion of these controls may adversely affect the interpretation of the relationship of interest when using the Gallagher disproportionality index.

In Table 4, we present results for party entry. In support of our theoretical results, we

[18] We specify flexible controls for the number of parties in an analogous manner to the number of districts.

[19] Specifying continuous variables in logarithms mitigates measurement error by ensuring that electoral systems with many parties (which tend to be more proportional, per the static results) do not simply exhibit a large amount of partisan churn by construction. Rather, any such relationship between proportionality and partisan churn should be interpreted as independent of the total number of parties. We provide further support for this interpretation by flexibly controlling for the number of parties in some specifications. Because elections may feature zero entries or exits, we transform these variables as $\log(1+x)$ in order to conserve data.

find a robust positive relationship between proportionality and party entry using all three measures. In Table 4, we present analogous results for party exit. We find the same predicted positive relationship using the first two measures of proportionality. However, our results are less robust – though broadly consistent – when using the Gallagher index. For the reasons given above, we should be less confident in those results *a priori*. In Table 6, we present similarly robust results for total party movements, or partisan churn. In all regressions, as we specify successively richer sets of controls, we are able to explain an increasing amount of the variation in partisan entry, exit and churn (note the increases in R^2). However, our estimates of the relationship between proportionality and these variables do not systematically change in a statistically discernable manner. We interpret this as robust evidence that is consistent with the dynamic predictions of our model.

Finally, we replicate our entire analyis using alternative party inclusion thresholds (λ) in order to establish that our qualitative results are not driven by the choice of any particular threshold. We present our central result – the estimated relationship between proportionality, as measured by the majoritarian dummy and effective district magnitude, and partisan churn conditional on all controls – for $\lambda = 1, 2, ..., 10\%$ in Figure 2. Our estimated relationships are statistically significantly different from zero at the 5% level for all values of λ, which again points to the robustness of our findings.[20]

4 Conclusion

This paper presents a novel dynamic reinterpretation of Duverger's Law. We construct a minimal but transparent dynamic model that establishes that (*i*) static Duverger predictions on the comparative number of parties under plurality rule and proportional representation can be reversed when intertemporal incentives are taken into account and (*ii*) a unique dynamic prediction can be recovered if we focus our attention on the comparative variation in the number of parties over time across electoral systems. We finds robust empirical support in favor of the latter prediction.

Since party formation and maintenance decisions are typically made on a national level,

[20]The fact that our estimates appear to converge towards zero for higher values of λ is consistent with the fact that higher inclusion thresholds will mechanically attenuate coefficient estimates. Intuitively, a higher inclusion threshold reduces the amount of variation in the dependent variable (as $\lambda \to 1$, $\beta \to 0$ by construction).

the dynamic predictions of our model can only be verified appropriately with cross-country elections data. Further, since electoral systems rarely change within countries, this hinders attempts to attribute a causal effect of electoral systems on the evolution of the number of national parties. We consider the time-series correlations uncovered in this paper sufficiently novel, interesting and robust that the lack of a causal interpretation does not present a critical concern. However, we make a broader contribution in that we point to the interest of studying the comparative intertemporal properties of electoral systems. In future work, related questions along these lines may be amenable to causal inference.

References

Anagol, S. and T. Fujiwara (2015). The runner-up effect. *Journal of Political Economy*, forthcoming.

Austen-Smith, D. (2000). Redistributing income under proportional representation. *Journal of Political Economy 108*(6), 1235–1269.

Austen-Smith, D. and J. Banks (1988). Elections, coalitions, and legislative outcomes. *American Political Science Review 82*(02), 405–422.

Baron, D. P. and D. Diermeier (2001). Elections, governments, and parliaments in proportional representation systems. *The Quarterly Journal of Economics 116*(3), 933–967.

Benoit, K. (2002). The endogeneity problem in electoral studies: a critical re-examination of duverger's mechanical effect. *Electoral Studies 21*(1), 35–46.

Benoit, K. (2006). Duverger's law and the study of electoral systems. *French Politics 4*(1), 69–83.

Blais, A. and R. K. Carty (1991). The psychological impact of electoral laws: measuring duverger's elusive factor. *British Journal of Political Science*, 79–93.

Brancati, D. (accessed 2013). *Global Elections Dataset.* New York: Global Elections Database, http://www.globalelectionsdatabase.com.

Chhibber, P. and K. Kollman (1998). Party aggregation and the number of parties in india and the united states. *American Political Science Review*, 329–342.

Chhibber, P. and G. Murali (2006). Duvergerian dynamics in the indian states federalism and the number of parties in the state assembly elections. *Party Politics 12*(1), 5–34.

Cox, G. W. (1997). *Making votes count: strategic coordination in the world's electoral systems*, Volume 7. Cambridge Univ Press.

Crutzen, B. S. and N. Sahuguet (2009). Redistributive politics with distortionary taxation. *Journal of Economic Theory 144*(1), 264–279.

Diwakar, R. (2007). Duverger's law and the size of the indian party system. *Party Politics 13*(5), 539–561.

Duverger, M. (1951). *Les partis politiques.* Armand Colin.

Faravelli, M. and S. Sanchez-Pages (2012). (don't) make my vote count.

Feddersen, T., I. Sened, and S. Wright (1990). Rational voting and candidate entry under plurality rule. *American Journal of Political Science 34*, 1005–1016.

Feddersen, T. J. (1992). A voting model implying duverger's law and positive turnout. *American journal of political science*, 938–962.

Fey, M. (1997). Stability and coordination in duverger's law: A formal model of preelection polls and strategic voting. *American Political Science Review*, 135–147.

Fujiwara, T. (2011). A regression discontinuity test of strategic voting and duverger's law. *Quarterly Journal of Political Science 6*(3-4), 197–233.

Gaines, B. J. (1999). Duverger's law and the meaning of canadian exceptionalism. *Comparative Political Studies 32*(7), 835–861.

Gallagher, M. (1991). Proportionality, disproportionality and electoral systems. *Electoral studies 10*(1), 33–51.

Gallagher, M. and P. Mitchell (2005). *The politics of electoral systems.* Cambridge Univ Press.

Hamlin, A. and M. Hjortlund (2000). Proportional representation with citizen candidates. *Public Choice 103*(3-4), 205–230.

Herrera, H., M. Morelli, and T. R. Palfrey (2012). Turnout and power sharing.

Indridason, I. H. (2011). Proportional representation, majoritarian legislatures, and coalitional voting. *American Journal of Political Science 55*(4), 955–971.

Lijphart, A. (1994). *Electoral systems and party systems: A study of twenty-seven democracies, 1945-1990*. Oxford University Press.

Mershon, C. and O. Shvetsova (2013). The microfoundations of party system stability in legislatures. *The Journal of Politics 75*(04), 865–878.

Morelli, M. (2004). Party formation and policy outcomes under different electoral systems. *The Review of Economic Studies 71*(3), 829–853.

Myatt, D. P. (2007). On the theory of strategic voting. *The Review of Economic Studies 74*(1), 255–281.

Myerson, R. B. (1993). Effectiveness of electoral systems for reducing government corruption: a game-theoretic analysis. *Games and Economic Behavior 5*(1), 118–132.

Myerson, R. B. and R. J. Weber (1993). A theory of voting equilibria. *American Political Science Review 87*(01), 102–114.

Neto, O. A. and G. W. Cox (1997). Electoral institutions, cleavage structures, and the number of parties. *American Journal of Political Science*, 149–174.

Ordeshook, P. C. and O. V. Shvetsova (1994). Ethnic heterogeneity, district magnitude, and the number of parties. *American journal of political science*, 100–123.

Ortuno-Ortin, I. (1997). A spatial model of political competition and proportional representation. *Social Choice and Welfare 14*(3), 427–438.

Osborne, M. J. and A. Slivinski (1996). A model of political competition with citizen-candidates. *The Quarterly Journal of Economics 111*(1), 65–96.

Palfrey, T. (1989). A mathematical proof of duverger's law. In P. C. Ordeshook (Ed.), *Models of strategic choice in politics*. University of Michigan Press.

Persson, T. and G. E. Tabellini (2005). *The economic effects of constitutions*. MIT press.

Powell, G. B. (2000). *Elections as instruments of democracy: Majoritarian and proportional visions*. Yale University Press.

Reed, S. R. (2001). Duverger's law is working in italy. *Comparative Political Studies 34*(3), 312–327.

Riker, W. H. (1982). The two-party system and duverger's law: An essay on the history of political science. *The American Political Science Review*, 753–766.

Shugart, M. S. (2005). Comparative electoral systems research: the maturation of a field and new challenges ahead. In M. Gallagher and P. Mitchell (Eds.), *The politics of electoral systems*. Cambridge Univiversity Press.

Taagepera, R. and B. Grofman (2003). Mapping the indices of seats–votes disproportionality and inter-election volatility. *Party Politics 9*(6), 659–677.

Taagepera, R. and M. S. Shugart (1989). *Seats and votes: The effects and determinants of electoral systems*. Yale University Press.

A Appendix: Proofs

Proof of Proposition 1. Note that (1) implies that under proportional representation, forming (or maintaining, since $\overline{c} > \underline{c}$) a party is uniquely stage optimal in preference state s_0 for party j, irrespective of whether interest group $-j$ is represented by a party. Also, since $\overline{p} > \frac{1}{3}$, (1) implies that $\overline{c} \leq \overline{p}[\overline{u} - u]$, so that forming (or maintaining, since $\overline{c} > \underline{c}$) a party is uniquely stage optimal in preference state s_j for party j, irrespective of whether interest group $-j$ is represented by a party. Finally, since $\overline{c} > \underline{c}$, it follows that, for any state (s, ϕ) and any equilibrium σ^*, $V_j(s, \phi \cup \{j\}; \sigma^*) \geq V_j(s, \phi; \sigma^*)$. Hence, in any equilibrium under proportional representation, it must be that $\sigma_j^*(s, \phi) = 1$ for all states such that $s \in \{s_0, s_j\}$.

It remains only to determine interest groups' equilibrium actions in preference state s_{-j}. Fix an equilibrium σ^* and consider a state (s_{-j}, ϕ) such that $j \in \phi$. If interest group j disbands its party, its payoff is

$$V_j(s_{-j}, \phi; \sigma^*) = (1 - \overline{p})u + \overline{p}\underline{u} + \delta \mathbb{E} V_j(s', \{-j\}; \sigma^*)$$

If instead interest group j maintains its party, let $V^d(s_{-j}, \phi; \sigma^*)$ be its payoff. We have that

$$V_j^d(s_{-j}, \phi; \sigma^*) = \underline{p}\overline{u} + pu + \overline{p}\underline{u} - \underline{c} + \delta \mathbb{E} V_j(s', \{-j, j\}; \sigma^*).$$

By our results from above, we have that, for any $s \in \{s_0, s_j\}$,

$$V_j(s, \{-j\}; \sigma^*) = V_j(s, \{-j, j\}; \sigma^*) - [\overline{c} - \underline{c}],$$

so that $V_j(s_{-j}, \phi; \sigma^*) > V_j^d(s_{-j}, \phi; \sigma^*)$ if and only if (2) holds. Note that (2) also implies that in state (s_{-j}, ϕ) such that $j \notin \phi$, interest group j strictly prefers not to form a party. Hence, for any equilibrium σ^* under proportional representation, we have that $\sigma^* = \sigma^{PR}$. $\square$

Proof of Proposition 2. Define $\underline{\beta}$ and $\overline{\beta}$ such that

$$\underline{\beta}[\overline{u} - \underline{u}] \equiv \frac{1}{1 - \delta q} \left[\frac{1 - \overline{p}}{2} [\overline{u} - u] - \underline{c} \right] - \frac{1 - \delta \frac{1+q}{2}}{1 - \delta q} [\overline{c} - \underline{c}], \text{ and}$$
$$\overline{\beta}[\overline{u} - \underline{u}] \equiv \overline{p}[\overline{u} - u] - \overline{c} + \frac{\delta}{1 - \delta} \left[\frac{1 - \overline{p}}{2} [\overline{u} - u] - \underline{c} \right] + \frac{\delta(1 - q)}{1 - \delta} \frac{p - \underline{p}}{2} [\overline{u} - u].$$

Fix any equilibrium σ^*. First, note that since $\beta \geq 0$, under plurality as under proportional representation, (1) implies that maintaining an existing party is uniquely stage optimal in preference state s_0 for interest group j, irrespective of whether interest group $-j$ is represented by a party. Hence, by the arguments in the proof of Proposition 1, $\sigma_j^*(s_0, \phi) = 1$ whenever $j \in \phi$. Second, since $\alpha \geq 0$, (1) also implies that $\sigma_j^*(s_j, \phi) = 1$ whenever $j \in \phi$. Third, since no new party faces entry penalty β following entry when $\phi = \emptyset$, (1) also ensures that $\sigma_j^*(s, \emptyset) = 1$ is uniquely optimal when $s \in \{s_0, s_j\}$.

Now consider state $(s_0, \{-j\})$ and equilibrium σ^*. If interest group j does not form a party, its payoff is

$$\frac{1+\overline{p}}{2}u+\frac{1-\overline{p}}{2}\underline{u}+\delta\mathbb{E}V_j(s',\{-j\};\sigma^*),$$

while if interest group j forms a party, its payoff is

$$\left(\frac{1-\overline{p}}{2}-\beta\right)\overline{u}+\overline{p}u+\left(\frac{1-\overline{p}}{2}+\beta\right)\underline{u}-\overline{c}+\delta\mathbb{E}V_j(s',\{-j,j\};\sigma^*).$$

Hence, interest group j does not form a party if and only if

$$\begin{aligned}\overline{c}-\left[\frac{1-\overline{p}}{2}[\overline{u}-u]-\beta[\overline{u}-\underline{u}]\right] &\geq \delta\mathbb{E}\Big[V_j(s',\{-j,j\};\sigma^*)-V_j(s',\{-j\};\sigma^*)\Big] \\ &\equiv \delta\mathbb{E}\Delta V_j(s';\sigma^*) \qquad (8)\end{aligned}$$

Consider state (s_{-j},ϕ) such that $j\in\phi$ and such that $\sigma^*_{-j}(s_{-j},\phi)=1$. If interest group j maintains its party, its payoff is

$$(\underline{p}-\alpha+\beta\mathbb{I}_{-j\notin\phi})\overline{u}+pu+(\overline{p}+\alpha-\beta\mathbb{I}_{-j\notin\phi})\underline{u}-\underline{c}+\delta\mathbb{E}V_j(s',\{-j,j\};\sigma^*),$$

while if interest group j disbands its party, its payoff is

$$(1-\overline{p})u+\overline{p}\underline{u}+\delta\mathbb{E}V_j(s',\{-j\};\sigma^*).$$

Hence, under profile $\underline{\sigma}^{PL}$, it must be that

$$\underline{c}-\underline{p}[\overline{u}-u]+(\alpha-\beta)[\overline{u}-\underline{u}]\geq\delta\mathbb{E}\Delta V_j(s';\underline{\sigma}^{PL}), \qquad (9)$$

while under profile $\overline{\sigma}^{PL}$, it must be that

$$\underline{c}-\underline{p}[\overline{u}-u]+\alpha[\overline{u}-\underline{u}]\leq\delta\mathbb{E}\Delta V_j(s';\overline{\sigma}^{PL}). \qquad (10)$$

Fix a state (s_j,ϕ) such that $j\notin\phi$. Under $\underline{\sigma}^{PL}$, (1) ensures that the stage payoffs of interest group j are strictly positive when it forms a party, so that, by an argument in the proof of Proposition 1, $\underline{\sigma}^{PL}(s_j,\phi)=1$ is optimal. Under $\overline{\sigma}^{PL}$, interest group j forms a party in state (s_j, ϕ) with $j \notin \phi$ if and only if

$$\overline{p}[\overline{u} - u] - \overline{c} + (\alpha - \beta)[\overline{u} - \underline{u}] \geq -\delta \mathbb{E} \Delta V_j(s'; \overline{\sigma}^{PL}). \tag{11}$$

Note that (9), along $\underline{\sigma}_{-j}^{PL}(s_{-j}, \emptyset) = 1$ and the fact that $\overline{c} > \underline{c}$, implies that $\underline{\sigma}_j^{PL}(s_{-j}, \emptyset) = 0$ is optimal. Since the profile $\overline{\sigma}^{PL}$ is specified in all states except $(s_{-j}, \emptyset)$, a simple computation verifies whether either $\overline{\sigma}_j^{PL}(s_{-j}, \emptyset) = 0$ or $\overline{\sigma}_j^{PL}(s_{-j}, \emptyset) = 0$ are optimal. Actions in this state are irrelevant when verifying equilibrium incentives, since under $\overline{\sigma}^{PL}$ it can be reached only following deviations by two interest groups.

Hence, the relevant incentive constraints under $\underline{\sigma}^{PL}$ are (8) and (9), while the relevant incentive constraints under $\overline{\sigma}^{PL}$ are (8), (10) and (11). These can be further simplified through computation. First, note that

$$\begin{aligned} \Delta V_j(s_j; \overline{\sigma}^{PL}) &= \overline{c} - \underline{c} + \beta[\overline{u} - \underline{u}], \\ \Delta V_j(s_j; \underline{\sigma}^{PL}) &= \overline{c} - \underline{c}, \\ \Delta V_j(s_{-j}; \underline{\sigma}^{PL}) &= 0, \end{aligned}$$

so that we have that

$$\Delta V_j(s_{-j}; \overline{\sigma}^{PL}) = \frac{1}{1 - \delta \frac{1-q}{2}} \left[\underline{p}[\overline{u} - \underline{u}] - \alpha[\overline{u} - \underline{u}] - \underline{c} + \delta q \Delta V_j(s_0; \overline{\sigma}^{PL}) + \delta \frac{1-q}{2} \Delta V_j(s_j; \overline{\sigma}^{PL}) \right],$$

and that

$$\Delta V_j(s_0; \overline{\sigma}^{PL}) = \frac{1}{1 - \delta q} \left[\frac{1 - \overline{p}}{2} [\overline{u} - u] - \underline{c} + \delta \frac{1-q}{2} \Delta V_j(s_{-j}; \overline{\sigma}^{PL}) + \delta \frac{1-q}{2} \Delta V_j(s_j; \overline{\sigma}^{PL}) \right].$$

Further computation yields that

$$\begin{aligned} \delta \mathbb{E} \Delta V_j(s'; \overline{\sigma}^{PL}) = \frac{1}{1 - \delta \frac{1+q}{2}} \Bigg[& \delta \frac{1-q}{2} \left[\underline{p}[\overline{u} - u] - \alpha[\overline{u} - \underline{u}] - \underline{c} \right] \\ & + \delta q \left[\frac{1 - \overline{p}}{2} [\overline{u} - u] - \underline{c} \right] + \delta \frac{1-q}{2} \left[\overline{c} - \underline{c} + \beta[\overline{u} - \underline{u}] \right] \Bigg]. \end{aligned}$$

Similarly,

$$\Delta V_j(s_0;\underline{\sigma}^{PL}) = \frac{1}{1-\delta q}\left[\frac{1-\overline{p}}{2}[\overline{u}-u] - \underline{c} + \delta\frac{1-q}{2}\Delta V_j(s_j;\underline{\sigma}^{PL})\right],$$

and further computation yields that

$$\delta\mathbb{E}\Delta V_j(s';\underline{\sigma}^{PL}) = \frac{1}{1-\delta q}\left[\delta q\left[\frac{1-\overline{p}}{2}[\overline{u}-u] - \underline{c}\right] + \delta\frac{1-q}{2}\left[\overline{c}-\underline{c}\right]\right].$$

Evaluated at $\overline{\sigma}^{PL}$, (8) can be rewritten as

$$\beta[\overline{u}-\underline{u}] \geq \frac{1-\delta\frac{1-q}{2}}{1-\delta q}\left[\frac{1-\overline{p}}{2}[\overline{u}-u] - \underline{c}\right] - [\overline{c}-\underline{c}] + \frac{\delta\frac{1-q}{2}}{1-\delta q}\left[\underline{p}[\overline{u}-\underline{u}] - \alpha[\overline{u}-\underline{u}] - \underline{c}\right], \quad (12)$$

while evaluated at $\underline{\sigma}^{PL}$, it can be rewritten as

$$\beta[\overline{u}-\underline{u}] \geq \frac{1}{1-\delta q}\left[\frac{1-\overline{p}}{2}[\overline{u}-u] - \underline{c}\right] - \frac{1-\delta\frac{1+q}{2}}{1-\delta q}\left[\overline{c}-\underline{c}\right]. \quad (13)$$

A straightforward computation verifies that, for any α, the righthand side of (13) is strictly larger than the righthand side of (12), so that (12) holds whenever (13) holds.

Also, (9) can be rewritten as

$$\alpha[\overline{u}-\underline{u}] \geq \underline{p}[\overline{u}-u] - \underline{c} + \beta[\overline{u}-\underline{u}] + \frac{1}{1-\delta q}\left[\delta q\left[\frac{1-\overline{p}}{2}[\overline{u}-u] - \underline{c}\right] + \delta\frac{1-q}{2}\left[\overline{c}-\underline{c}\right]\right], \quad (14)$$

while (10) can be rewritten as

$$\alpha[\overline{u}-\underline{u}] \leq \underline{p}[\overline{u}-u] - \underline{c} + \frac{1}{1-\delta q}\left[\delta q\left[\frac{1-\overline{p}}{2}[\overline{u}-u] - \underline{c}\right] + \delta\frac{1-q}{2}\left[\overline{c}-\underline{c}+\beta[\overline{u}-\underline{u}]\right]\right]. \quad (15)$$

Finally, since the righthand side of (11) is increasing in α, it can be shown by computation to hold for all α if and only if

$$\beta[\overline{u}-\underline{u}] \leq \overline{p}[\overline{u}-u] - \overline{c} + \frac{\delta}{1-\delta}\left[\frac{1-\overline{p}}{2}[\overline{u}-u] - \underline{c}\right] + \frac{\delta(1-q)}{1-\delta}\frac{p-\underline{p}}{2}[\overline{u}-u], \quad (16)$$

That (13) holds follows since $\beta \geq \underline{\beta}$, and that (16) holds follows since $\beta \leq \overline{\beta}$. Hence, conditions (13) and (14) are sufficient for $\underline{\sigma}^{PL}$ to be an equilibrium, while (13), (15) and (16) are sufficient for $\overline{\sigma}^{PL}$ to be an equilibrium. Let $\check{\alpha}$ be the unique value of α such that (14) holds as an equality and define $\underline{\alpha} = \max\{\min\{\underline{p} - \beta, \check{\alpha}\}, 0\}$. Similarly, let $\hat{\alpha}$ be the unique value of α such that (15) holds as an equality and define $\overline{\alpha} = \min\{\max\{0, \hat{\alpha}\}, \underline{p} - \beta\}$. Hence, given any β satisfying (13), $\underline{\sigma}^{PL}$ is an equilibrium if $\alpha > \underline{\alpha}$, while $\overline{\sigma}^{PL}$ is an equilibrium if $\alpha < \overline{\alpha}$. These are sufficient conditions only, since our definition of $\underline{\alpha}$ and $\overline{\alpha}$ embeds the cases when these equilibria fails to exits. Furthermore, (14) and (15) imply that $\underline{\alpha} \geq \overline{\alpha}$, where the inequality is strict whenever $\underline{\alpha}, \overline{\alpha} \in (0, \underline{p} - \beta)$. □

B Appendix: Tables and Figures

Table 1: Data

Country	Years	Elections	Country	Years	Elections
Australia	1993-2001	4	Luxembourg	1945-2004	14
Austria	1945-2008	20	Malaysia	1959-1999	11
Belgium	1978-2007	18	Malta	1945-2003	16
Bermuda	1989-1998	3	Mauritius	1995-2000	2
Bolivia	1989-2002	6	Mexico	1997-2000	5
Botswana	1999	1	Netherlands	1948-2006	18
Bulgaria	1994-2005	4	New Zealand	1946-1999	22
Canada	1945-2000	26	Norway	1977-2005	8
Costa Rica	1953-2002	13	Poland	1991-2007	12
Cyprus	1991-1996	2	Portugal	1975-2005	12
Czech Republic	1996-2008	7	Romania	1992-2004	8
Estonia	1992-2007	5	Russia	1995-1999	2
Finland	1999-2007	3	Slovakia	1994-1998	2
France	1988-2007	5	South Africa	1994-1999	4
Germany	1990-2009	6	Spain	1977-2008	20
Greece	1946-2007	20	Sweden	1944-2006	20
Hungary	1990-2006	5	Switzerland	1947-2007	16
Iceland	1959-2007	13	Trin. & Tobago	1966-2002	10
Ireland	1948-1997	16	Turkey	1999-2002	6
Israel	1949-2003	16	United Kingdom	1945-2005	16
Italy	1948-2006	14	United States	1986-2000	16
Latvia	1993-2006	5	Venezuela	1958-1988	7

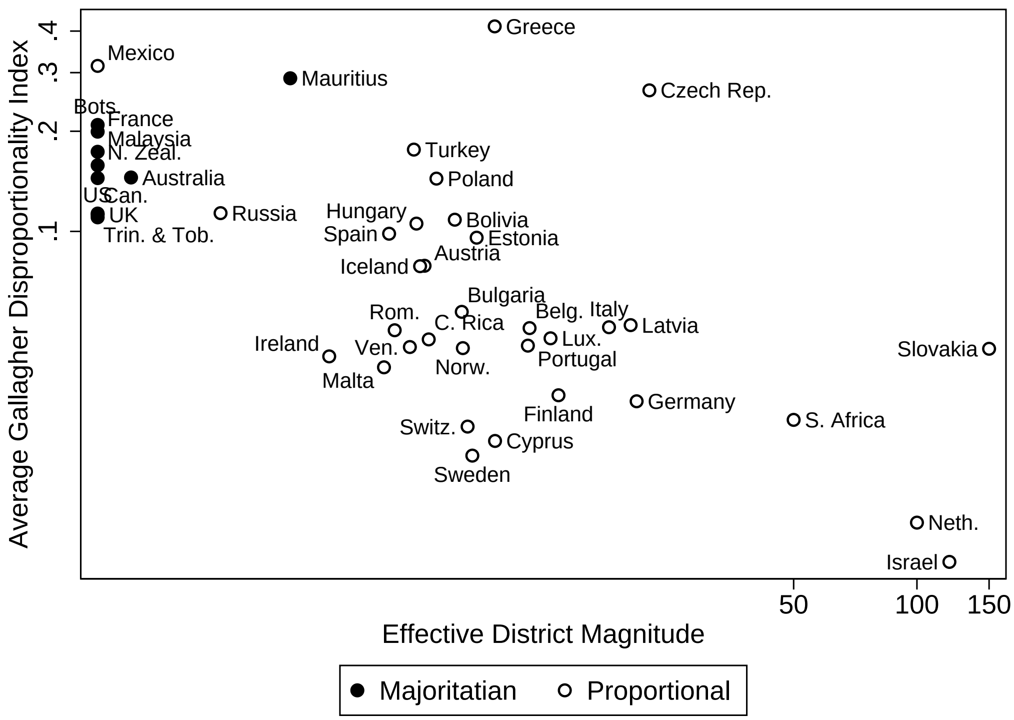
Notes: All data comes from the Constituency-Level Elections Dataset.

Table 2: Summary Statistics

Variable	Mean	Std. Dev.	Source
Number of Districts	79.61	148.84	CLE
Total Votes/District (millions)	1.28	2.51	CLE
Total Seats in Play	207.96	169.79	CLE
Effective District Magnitude	17.16	35.34	Authors' Calculations
Average Gallagher Disproportionality Index	0.10	0.09	Authors' Calculations
Majoritarian Dummy	0.23	0.41	Persson and Tabellini (2005)
Number of Parties	3.59	1.52	Authors' Calculations
Number of Entries	0.71	1.01	Authors' Calculations
Number of Exits	0.72	1.06	Authors' Calculations

Notes: $N = 454$. CLE corresponds to the Constituency-Level Elections Dataset. Number of parties, entries and exits are calculated with a 5% inclusion threshold.

Figure 1: Electoral Proportionality: Three Measures



Notes: In this figure, we present three alternative measures of electoral proportionality. Majoritarian electoral systems as defined by Persson and Tabellini (2005) are shown as solid dots. The Average Gallagher Dispropotionality Index for a given country is constructed by averaging the Gallagher Disproportionality Index for each election in the sample for each country. Both axes are in log scale.

Table 3: Static Tests of Duverger's Law

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Majoritarian Dummy	-0.23**	-0.17*				
	(0.09)	(0.10)				
Effective District Magnitude			0.09***	0.05**		
			(0.02)	(0.03)		
Average Gallagher					-2.76***	-2.77***
Disproportionality Index					(0.86)	(0.88)
Decade, Regional and District Number Controls Included?	N	Y	N	Y	N	Y
R^2	0.06	0.23	0.10	0.23	0.31	0.42
Number of Observations	454	454	454	454	454	454

Notes: Dependent variable is the total number of parties calculated with a 5% vote share inclusion threshold. Majoritarian dummy is obtained from Persson and Tabellini (2005). Average Gallagher Disproptionality Index for a given country is constructed by averaging the Gallagher Disproportionality Index for each election in the sample for each country. Flexible control for the number of districts and parties is achieved by including sixth order polynomials in those variables and in the log of those variables. All continuous variables are specified in logarithms. To conserve data, dependent variable is transformed as $\log(1+x)$. Robust standard errors clustered by country are presented in parentheses. *** - 1% significance level, ** - 5% significance level, * - 10% significance level

Table 4: Dynamic Tests of Duverger's Law: Entry

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Majoritarian Dummy	-0.21**	-0.34***	-0.28***						
	(0.08)	(0.09)	(0.09)						
Average District Magnitude				0.06**	0.11**	0.08***			
				(0.03)	(0.03)	(0.03)			
Average Gallagher Disproportionality Index							-0.72*	-1.01***	-0.11
							(0.43)	(0.32)	(0.75)
Decade, Regional, and District Number Controls Included?	N	Y	Y	N	Y	Y	N	Y	Y
Flexibly Controlled for Number of Parties?	N	N	Y	N	N	Y	N	N	Y
R^2	0.03	0.18	0.30	0.04	0.20	0.31	0.02	0.17	0.28
Number of Observations	411	411	411	411	411	411	411	411	411

Notes: Dependent variable is total number of party entries as computed according to equation (6) with a 5% inclusion threshold. Majoritarian dummy is obtained from Persson and Tabellini (2005). Average Gallagher Disproptionality Index for a given country is constructed by averaging the Gallagher Disproportionality Index for each election in the sample for each country. Flexible control for the number of districts and parties is achieved by including sixth order polynomials in those variables and in the log of those variables. All continuous variables are specified in logarithms. To conserve data, dependent variable is transformed as $\log(1+x)$. Robust standard errors clustered by country are presented in parentheses. *** - 1% significance level, ** - 5% significance level, * - 10% significance level

Table 5: Dynamic Tests of Duverger's Law: Exit

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Majoritarian Dummy	-0.19**	-0.29***	-0.30***						
	(0.08)	(0.10)	(0.10)						
Average District Magnitude				0.06*	0.07**	0.07**			
				(0.03)	(0.03)	(0.03)			
Average Gallagher Disproportionality Index							-0.50	-0.71*	-0.16
							(0.547	(0.43)	(0.71)
Decade, Regional, and District Number Controls Included?	N	Y	Y	N	Y	Y	N	Y	Y
Flexibly Controlled for Number of Parties?	N	N	Y	N	N	Y	N	N	Y
R^2	0.03	0.15	0.22	0.03	0.15	0.22	0.01	0.14	0.20
Number of Observations	411	411	411	411	411	411	411	411	411

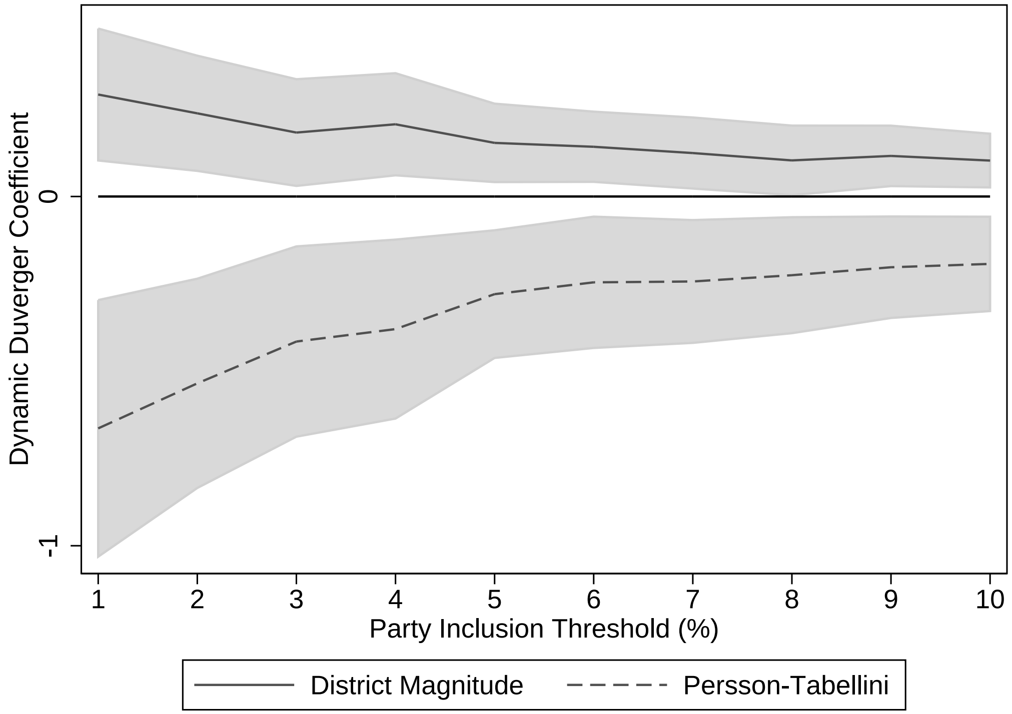
Notes: Dependent variable is total number of party exits as computed according to equation (7) with a 5% inclusion threshold. Majoritarian dummy is obtained from Persson and Tabellini (2005). Average Gallagher Disproptionality Index for a given country is constructed by averaging the Gallagher Disproportionality Index for each election in the sample for each country. Flexible control for the number of districts and parties is achieved by including sixth order polynomials in those variables and in the log of those variables. All continuous variables are specified in logarithms. To conserve data, dependent variable is transformed as $\log(1+x)$. Robust standard errors clustered by country are presented in parentheses. *** - 1% significance level, ** - 5% significance level, * - 10% significance level

Table 6: Dynamic Tests of Duverger's Law: Total Movements

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Majoritarian Dummy	-0.28**	-0.46***	-0.42***						
	(0.11)	(0.13)	(0.13)						
Average District Magnitude				0.09**	0.14***	0.12***			
				(0.04)	(0.04)	(0.04)			
Average Gallagher Disproportionality Index							-0.93	-1.23***	-0.01
							(0.60)	(0.48)	(0.79)
Decade, Regional, and District Number Controls Included?	N	Y	Y	N	Y	Y	N	Y	Y
Flexibly Controlled for Number of Parties?	N	N	Y	N	N	Y	N	N	Y
R^2	0.04	0.17	0.27	0.04	0.19	0.28	0.01	0.16	0.25
Number of Observations	411	411	411	411	411	411	411	411	411

Notes: Dependent variable is total number of party movements (entries + exits) as computed according to equations (6) and (7) with a 5% inclusion threshold. Majoritarian dummy is obtained from Persson and Tabellini (2005). Average Gallagher Disproptionality Index for a given country is constructed by averaging the Gallagher Disproportionality Index for each election in the sample for each country. Flexible control for the number of districts and parties is achieved by including sixth order polynomials in those variables and in the log of those variables. All continuous variables are specified in logarithms. To conserve data, dependent variable is transformed as $\log(1+x)$. Robust standard errors clustered by country are presented in parentheses. *** - 1% significance level, ** - 5% significance level, * - 10% significance level

Figure 2: Robustness: Alternative Inclusion Thresholds



Notes: In this figure, we present coefficient estimates from regressions of total party movements on the Majoritarian dummy from Persson and Tabellini (2005) and on effective district magnitude respectively. Total party movements (entries + exits) are computed according to equations (6) and (7) with various inclusion thresholds of 1,2,...,10% vote share. Each regression includes decade and regional controls along with flexible controls for the number of districts and parties in each election. The shaded regions correspond to 95% confidence intervals from robust standard errors that are clustered by country.