UNIVERSITY OF TARTU

Faculty of Social Sciences

The Johan Skytte Institute of Political Studies

Sofia Samoilova

BEYOND SEAT SHARE: EVALUATING VOTING-POWER INDICES IN GOVERNMENT FORMATION

MA thesis

Supervisor: Martin Mölder, PhD

Tartu 2026

Authorship Declaration

I have prepared this thesis independently. All the views of other authors, as well as data from literary sources and elsewhere, have been cited.

Word count of the thesis: 20,145 words

Sofia Samoilova, 18.05.2026

Abstract

This thesis evaluates whether cooperative game-theoretic voting-power indices improve empirical explanations of government formation relative to models based solely on parliamentary seat share. While seat share is the dominant measure of bargaining power in the coalition formation literature, it captures numerical size rather than strategic necessity - two quantities that diverge substantially in fragmented multiparty legislatures. Four indices are examined: the Banzhaf Power Index and Shapley-Shubik Value, which compute pivotality across all winning coalitions, and the Deegan-Packel and Holler-Packel indices, which restrict attention to minimal winning coalitions only.

The analysis covers 4,939 party-election observations from 626 elections across 37 parliamentary democracies between 1945 and 2023, using nested logistic regression models evaluated through four complementary criteria: McFadden pseudo-R^2, likelihood ratio test, Bayesian Information Criterion, and area under the ROC curve.

The results are asymmetric. The all-coalitions family consistently outperforms seat share across both outcome variables and all four criteria. The minimal-winning-coalition family does not, because its restriction to minimal winning coalitions embeds an assumption about rational behaviour that the empirical record systematically violates. For researchers operationalising bargaining power in coalition research, the Banzhaf or Shapley-Shubik index is the appropriate structural measure; seat share, while interpretable, is a less precise substitute.

Table of Contents

List of Figures

List of Tables

INTRODUCTION

In the 2010 Belgian federal elections, the New Flemish Alliance obtained 17.4% of parliamentary seats, which appeared to be the largest share of any single party yet spent 541 days outside government while smaller parties with a fraction of its numerical weight entered and exited successive coalition negotiations (Deschouwer, 2012). In the 1973 Irish general election, the Labour Party secured just 13.7% of Dáil seats and still entered government, receiving a cabinet share that exceeded its numerical contribution by a considerable margin (Mair, 1987, pp. 82-84). Surprisingly, these episodes are not anomalies. Across European parliamentary democracies, the relationship between a party's share of parliamentary seats and its probability of entering government is systematically non-proportional: numerical size and political leverage diverge with sufficient regularity to constitute a fundamental explanatory challenge for the coalition formation literature (Martin & Stevenson, 2001, pp. 33-35; Laver & Schofield, 1998, pp. 193-200; Felsenthal & Machover, 2005, pp. 486-488).

The dominant response to this puzzle in comparative politics has been to treat seat share as the primary measure of a party's bargaining power in coalition negotiations (Gamson, 1961; Riker, 1962; Warwick & Druckman, 2006, p. 635). This is, on its face, a reasonable starting point. Seat share is directly observable, requires no modelling assumptions, and is available for virtually every parliamentary democracy across extended time periods, making it specifically suited to the demands of large-n comparative research (Warwick & Druckman, 2006, p. 635). The empirical regularity known as Gamson's Law, that cabinet portfolios are allocated roughly proportionally to parties' seat contributions within governing coalitions, lends further credibility to this approach (Warwick & Druckman, 2006, p. 635; Bäck, Meier & Persson, 2009, pp. 10-12). Yet seat share measures numerical size, not strategic necessity (Felsenthal & Machover, 2005, pp. 486-488). A party may hold a substantial share of parliamentary seats while being entirely dispensable for the construction of any viable governing majority; a much smaller party may be indispensable if no majority can be assembled without it (Shapley & Shubik, 1954, pp. 787-788). In fragmented multiparty legislatures, these two quantities, namely numerical size and strategic pivotality, can differ substantially (Ansolabehere et al., 2005, pp. 550–551). When they do, seat share not only underperforms as a predictor of coalition outcomes but actively misrepresents where bargaining power actually resides (Felsenthal & Machover, 2005, pp. 486-488; De Winter & Dumont, 2006, pp. 180-182).

Cooperative game theory offers a formal solution to this measurement problem (Shapley & Shubik, 1954; von Neumann & Morgenstern, 1944). Voting-power indices, most prominently the Banzhaf Power Index (Banzhaf, 1965) and the Shapley-Shubik Value (Shapley & Shubik, 1954), operationalize bargaining power as pivotality: the frequency with which a party's participation is decisive for converting a losing coalition into a winning one (Felsenthal & Machover, 1998, pp. 191-196). A second family of indices, the Deegan-Packel Index (Deegan & Packel, 1978) and the Holler-Packel Index (Holler, 1982), restricts this logic further by computing power exclusively over minimal winning coalitions, assuming that rational office-seeking parties will avoid forming governments larger than strictly necessary (Riker, 1962; Holler & Packel, 1983, pp. 21-23). These four indices have been applied to the analysis of voting power in the European Union Council of Ministers (Felsenthal & Machover, 1997; 2001), corporate governance (Leech, 2002), and in scattered analyses of national parliamentary systems (Laver & Schofield, 1998; Ansolabehere et al., 2005). The theoretical case for their superiority over seat share as measures of bargaining power is well-developed (Felsenthal & Machover, 2005; Straffin, 1994). What has not been established is whether this theoretical advantage translates into better empirical predictions of which parties enter government, which is the gap this thesis addresses.

This gap between theoretical argument and systematic empirical evaluation is the research puzzle motivating this thesis. If voting-power indices truly capture something about the government formation process that seat share misses, this should be detectable as a consistent pattern across a large and varied sample of parliamentary democracies (Martin & Stevenson, 2001; Glasgow, Golder & Golder, 2012). If the indices do not outperform seat share, or if some indices outperform it while others do not, that finding is equally consequential, since it would identify which theoretical assumptions about coalition formation are empirically supported (Felsenthal & Machover, 1998, pp. 247-249). The existing literature has produced a compelling case for pivotality-based measurement (Felsenthal & Machover, 2005; Ansolabehere et al., 2005; Warwick & Druckman, 2006) while simultaneously relying, in much of its empirical practice, on the very seat-share proxy that theory identifies as inadequate (Gamson, 1961; Warwick & Druckman, 2006, p. 635; Bäck & Dumont, 2007, pp. 18-22). Resolving this issue requires a systematic comparative test that the literature has not yet provided.

The academic relevance of this question extends in two directions. Methodologically, the choice of how to measure bargaining power is not a technical detail but a theoretically loaded decision that shapes every subsequent inference about why particular coalitions form and how office benefits are distributed (Laver & Schofield, 1998, pp. 5-7; De Winter & Dumont, 2006, pp. 175-176). A measure that assigns zero influence on a pivotal small party and positive weight to a strategically redundant large party will produce systematically misleading predictions regardless of how sophisticated the statistical model surrounding it is (Felsenthal & Machover, 2005, pp. 486-488). Essentially, the question of which parties enter government is among the most consequential in democratic politics: it determines which interests gain access to executive power, which policy agendas are implemented, and how electoral outcomes are translated into governing authority (Müller & Strøm, 2000; Martin & Vanberg, 2020, pp. 325-326). Improving the measurement of bargaining power is therefore not merely a refinement of existing approaches but a precondition for more accurate explanation of how parliamentary democracy works (Martin & Stevenson, 2001, pp. 33-35). This thesis contributes to that improvement by providing a large-n cross-national comparative test of all four established voting-power indices against seat share as predictors of government entry, evaluated within a consistent analytical framework and across a dataset spanning 626 elections in 37 countries from 1945 to 2023.

The aim of this thesis is to evaluate whether cooperative game-theoretic voting-power indices improve empirical explanations of government formation relative to models based on parliamentary seat share alone, and to identify which indices, if any, produce consistent and meaningful predictive gains across a broad comparative sample of parliamentary democracies.

The central research question is: to what extent do formal voting-power indices improve the empirical prediction of government entry compared to models based solely on parliamentary seat share? This question is addressed through two subordinate questions. First, do all four established indices improve on seat share, or is the predictive advantage confined to particular families of indices? Second, is any predictive advantage robust across different operationalizations of the government participation outcome?

The theoretical framework developed in Chapter 1 generates two directional hypotheses. First, the all-coalitions family (Banzhaf and Shapley-Shubik) is expected to outperform seat share as a predictor of government entry, because it captures pivotality across the full range of winning coalitions, including the surplus configurations that are empirically common in European parliamentary democracies. Second, the minimal-winning-coalition family (Deegan-Packel and Holler-Packel) is expected not to outperform seat share, because its restriction to minimal winning coalitions embeds an assumption about coalition formation that the empirical record systematically violates. Both hypotheses are formally stated at the close of Section 1.5, following the review of empirical applications, and are tested in Chapters 2 and 3.

The empirical analysis employs a quantitative comparative research design. The data source is the ParlGov database (Döring & Manow, 2024), which provides systematic information on parliamentary elections, seat distributions, and cabinet compositions across European parliamentary and OECD democracies from 1945 to 2023. The four voting-power indices were computed through full combinatorial enumeration of all possible coalitions within each parliamentary election. The analytical sample, after removing elections in which a single party held an outright majority and applying computational feasibility thresholds, comprises 4,939 party-election observations drawn from 626 elections across 37 countries. The primary outcome variable is a binary indicator of whether a party entered the first cabinet formed after an election; a secondary outcome captures participation in any cabinet during the electoral term. Predictive performance is evaluated through nested logistic regression models, comparing a seat-share baseline against joint models that add each power index as an additional predictor, using four complementary criteria: McFadden pseudo-R^2, the likelihood ratio test, the Bayesian Information Criterion, and the area under the ROC curve.

The thesis proceeds in three substantive chapters. Chapter 1 develops the theoretical framework, moving from the logic of coalition government formation through the principal explanatory traditions in the literature to the measurement critique of seat share and the formal properties of the four voting-power indices. Chapter 2 presents the empirical analysis, covering the research design, dataset construction, and model specification. Chapter 3 reports the results, interprets the findings, and addresses the limitations and scope of the conclusions.

1. FROM SEAT SHARE TO PIVOTALITY: THEORETICAL FOUNDATION OF COALITION BARGAINING

1.1. Coalition Types and the Logic of Government Formation

Coalition governments constitute the dominant form of executive authority in contemporary parliamentary democracies (Martin & Vanberg, 2020, pp. 325-326). In most European political systems, single-party majority governments are the exception rather than the rule: proportional representation electoral systems tend to produce fragmented party systems in which no single party secures an absolute parliamentary majority (Lijphart, 2012, pp. 63-70; Gallagher, Laver, & Mair, 2011, pp. 308-309). Under these conditions, government formation requires cooperation among multiple political parties that jointly command a legislative majority, making coalition governments not merely occasional outcomes but the standard mechanism through which executive power is organized in parliamentary democracies (Martin & Stevenson, 2001, pp. 34-36; Müller & Strøm, 2000).

A coalition government can be defined as an executive arrangement in which two or more political parties agree to govern together by coordinating policy decisions and sharing cabinet positions (Martin & Vanberg, 2020, pp. 325-327; Laver & Schofield, 1998, pp. 1-6). These agreements typically involve formal or informal bargaining over policy priorities, office allocation, and legislative strategy (Strøm, 1990; Müller & Strøm, 2000). Coalition partners may differ substantially in their ideological positions, electoral strength, and strategic objectives, yet they cooperate to secure access to executive office and influence public policy (Budge & Laver, 1986, pp. 487–490; Strøm, 1990). As such, coalition governance reflects both the fragmentation of party systems and the strategic interdependence of political actors operating within them.

The process of government formation represents an important link between electoral outcomes and political authority (Laver & Schofield, 1998, pp. 1-3). Elections determine the distribution of parliamentary seats among parties, but they do not directly determine who governs (Lijphart, 2012, pp. 71-72). Instead, post-electoral bargaining transforms legislative seat shares into governing coalitions capable of commanding parliamentary confidence (Martin & Stevenson, 2001, pp. 33-34; Strøm, 1990). The actual composition of the executive therefore depends on strategic interactions among parties rather than on electoral results alone (Martin & Stevenson, 2001, pp. 34-36; Glasgow, Golder & Golder, 2012, pp. 249-251). Building on this, this thesis conceptualizes government formation as a post-electoral, institutionally embedded bargaining process in which political parties negotiate over cabinet composition and portfolio allocation in order to form a government capable of securing and maintaining parliamentary confidence. The institutional context within which this bargaining unfolds varies systematically across parliamentary systems, and these variations are discussed further in Section 1.2.

Coalition formation processes can produce a variety of government types depending on how parties combine to secure legislative support. The coalition literature typically distinguishes between several ideal types based on parliamentary support and composition (Riker, 1962; Laver & Schofield, 1998, pp. 70-75; Strøm, 1990). The minimal winning coalition (MWC), introduced by Riker (1962), controls a parliamentary majority such that removing any single member party causes it to fall below that threshold. According to Riker's size principle, rational office-seeking parties should prefer MWCs because they minimize the number of partners among whom the benefits of office must be shared. In contrast, surplus majority coalitions include more parties than are strictly necessary to command a majority. Such coalitions arise for various reasons, including the desire to increase government stability, manage internal party divisions, or broaden policy consensus (Laver & Schofield, 1998, pp. 75-78; Martin & Stevenson, 2001, pp. 35-36; Müller & Strøm, 2000), and their empirical frequency constitutes a direct challenge to purely size-based theories. Finally, a minority government is one in which the governing party or coalition does not control a parliamentary majority but remains in office because it is tolerated by non-government parties (Strøm, 1990), relying on ad hoc legislative support to govern.

Taken together, this typology illustrates that coalition outcomes vary not only in the number of parties involved but also in their relationship to parliamentary majorities and the distribution of bargaining power among members. Following any parliamentary election, the distribution of seats typically permits a large number of mathematically feasible governing coalitions, yet in practice only one ultimately forms (Laver & Schofield, 1998, pp. 25-30; Martin & Stevenson, 2001, pp. 33-35). Explaining why a particular combination of parties emerges from this broader set of feasible alternatives has therefore been a core challenge in comparative politics. Early theoretical approaches focused on coalition size, most notably through Riker's (1962) size principle, but extensive empirical research demonstrates that observed outcomes systematically deviate from these predictions (Martin & Stevenson, 2001, pp. 34-36; Glasgow, Golder & Golder, 2012, pp. 248-250; Dumont et al., 2024). These patterns underscore a fundamental distinction between numerical size and bargaining power and motivate the need for more refined measures of political influence, as discussed in the sections that follow.

1.2. Explanatory Approaches to Coalition Formation

Explaining which parties enter government following parliamentary elections requires consideration of a complex interaction of strategic, ideological, and institutional factors (Laver & Schofield, 1998, pp. 30-31; Martin & Stevenson, 2001, pp. 35-38). Coalition outcomes cannot be reduced to a single explanatory dimension: in practice, size, policy, and institutions interact, but the relative weight assigned to each distinguishes the major theoretical approaches to coalition research.

The distribution of seats in parliament constitutes the key baseline for analysing coalition formation. In classical coalition theory, bargaining power is assumed to derive directly from numerical strength: parties with more seats are expected to have greater leverage in coalition negotiations and a higher probability of entering government (Riker, 1962; Laver & Schofield, 1998; Strøm, 1990). Riker (1962) formalizes this as the size principle: rational, office-seeking parties should prefer minimal winning coalitions, since any coalition that controls a majority and loses majority status upon the removal of any single member minimizes the number of partners among whom office benefits must be shared. Based on the author's explanation, a coalition is winning if:

$$\sum_{i \in S} s_i \geq q,$$

And minimal winning if for every member party, its removal causes the coalition to lose its majority:

$$\forall j \in S : \sum_{i \in S\{j\}} s_i < q,$$

where S – coalition of parties;

s_i – seat share of party I;

q – majority threshold (typically >50% of seats).

Gamson's Law (1961) complements this by predicting that cabinet portfolios are allocated proportionally to parties' seat shares within the governing coalition:

$$Portfolio\ Share_i \approx \frac{s_i}{\sum_{j \in S} s_j}$$

where S – set of parties in the government;

s_i – seat share of party I in the coalition;

s_j – seat share of every coalition member $j \in S$.

Together, these approaches establish a size-based model of coalition bargaining in which larger parties are more likely to enter government, coalitions tend toward minimal winning arrangements, and the distribution of office reflects relative parliamentary strength.

However, extensive empirical research demonstrates that seat share alone is an imperfect predictor of coalition outcomes. Coalition formation frequently deviates from minimal winning logic: surplus coalitions and minority governments are common, and smaller parties often enter government despite limited numerical weight (Martin & Stevenson, 2001; Glasgow, Golder & Golder, 2012; Dumont et al., 2024). An important improvement is the Effective Number of Parties index (Laakso & Taagepera, 1979), which captures the distribution of parliamentary power across parties rather than its concentration in a single seat-share figure:

$$N = \frac{1}{\sum_{i=1}^{n} s_i^2}$$

where N – effective number of parties;

si – share of party i's seats/votes;

n – number of parties.

In systems with a low effective number of parties, power is concentrated among a small number of large actors, making coalition formation more predictable and more closely aligned with seat-share expectations. In highly fragmented systems, by contrast, even relatively small parties may become pivotal for assembling a majority, while larger parties can be excluded if alternative coalition configurations are viable (Martin & Stevenson, 2001). Bargaining power therefore depends not only on the seats a party holds, but on how seats are distributed across the entire system, which is a function that seat share alone cannot capture.

While numerical size provides the arithmetic foundation of coalition formation, policy preferences and ideological compatibility propose significant constraints on which coalitions are politically feasible. Parties are not purely office-seeking actors; they are also policy-seeking, and therefore aim to enter governments that allow them to influence outcomes in line with their ideological positions. Axelrod (1970, pp. 165-187) develops a foundational contribution in this regard, arguing that coalitions tend to form among ideologically adjacent parties. The concept of connected MWCs predicts that parties will prefer coalition partners contiguous in ideological space, thereby minimizing internal policy conflict. This logic builds on spatial models of party competition, in which ideological distance between parties raises the expected costs of cooperation (Downs, 1957). Empirically, ideological proximity is a strong predictor of coalition formation, and extreme or anti-system parties are often excluded from government even when they hold substantial seat shares (Cox, 1997; Laver & Shepsle, 1996, pp. 35-38; Golder, 2006). The policy-based perspective thus introduces an important qualification to size-based theories: while seat shares determine which coalitions are arithmetically possible, ideology constrains which are politically viable. However, ideologically suboptimal coalitions sometimes form due to strategic considerations or institutional constraints (Laver & Schofield, 1998, pp. 105-110; Martin & Stevenson, 2001, pp. 35-38), and the policy perspective does not directly address how bargaining power is distributed within the ideologically feasible set, reinforcing the need for more precise measures of structural influence once ideological constraints are taken into account (Felsenthal & Machover, 2005, pp. 487-490).

Beyond size and ideology, institutional constraints structure the rules and incentives under which parties negotiate, shaping outcomes independently of seat distributions (Laver & Shepsle, 1996, pp. 25-28; Müller & Strøm, 2000). The investiture rule is a central example: under positive parliamentarism, a prospective government must secure an explicit parliamentary majority through a formal investiture vote before taking office, which tends to encourage majority coalitions. Under negative parliamentarism, a government can assume office without an affirmative majority provided that no majority actively votes against it, which makes minority governments structurally viable and empirically common (Strøm, 1990; Persson & Tabellini, 2003, pp. 115-117). Institutions also determine who initiates government formation, how long negotiations can last, and what alternatives are available if they fail – factors that can advantage certain parties independently of their size or ideological position (Baron & Ferejohn, 1989, pp. 1191-1195; Laver & Shepsle, 1996, pp. 30–34). Like ideology, however, institutions condition how bargaining power is exercised rather than measuring it directly, and neither tradition provides a formal measure of how influence is distributed among parties within the feasible set of coalitions.

Taken together, size, policy, and institutions define the explanatory ground of coalition formation research. Two major intellectual traditions have developed distinct approaches to synthesizing them, and these are compared and discussed in Table 1.

Table 1

Comparison of the European Politics and Game-Theoretic Traditions in Coalition Research

Criterion	European Politics Tradition	Game-Theoretic Tradition
Epistemological approach	Inductive: theory built from observed coalition outcomes (Laver & Schofield, 1998, pp. 7–8)	Deductive: predictions derived from axioms of rational choice (Laver & Schofield, 1998, pp. 3–5)
Primary method	Cross-national empirical analysis of actual government outcomes (Warwick, 1996; Bäck & Dumont, 2007)	Formal cooperative game-theoretic modelling (von Neumann & Morgenstern, 1944)
Core assumption	Institutions, party systems, and historical context shape outcomes (Laver & Schofield, 1998, p. 8)	Actors are rational and seek to maximize their share of government payoffs (Riker, 1962, pp. 22–28)
Key explanatory concepts	Policy and ideological compatibility, formateur advantage, institutional constraints (Axelrod, 1970; Diermeier & Stevenson, 1999)	Pivotality, minimal winning coalitions, formal solution concepts (Riker, 1962; Banzhaf, 1965)
Measure of bargaining power	Seat share and contextual leverage in coalition negotiations (Martin & Stevenson, 2001)	Formal voting-power indices: Banzhaf Power Index and Shapley-Shubik value (Felsenthal & Machover, 2005)

Treatment of coalition size	One factor among many; surplus and minority governments explained by contextual factors (Strøm, 1990)	Central prediction: rational actors form minimal winning coalitions to maximize payoffs (Riker, 1962, pp. 32–46)
Generalizability	Limited: anchored to post-war European cases and a fixed data universe (Laver & Schofield, 1998, pp. 8–10)	High: applies to any weighted voting game regardless of country or time period (Shapley & Shubik, 1954)
Main limitation	Theory developed and tested on the same data universe, constraining independent testing (Laver & Schofield, 1998, p. 9)	Abstracts away ideology, institutional context, and formateur effects (Strøm, 1990; Glasgow, Golder & Golder, 2012)

Source: compiled by the author

The two traditions differ not only in their explanations of coalition formation, but in their epistemological and methodological foundations. For example, the European Politics tradition adopts an inductive approach, deriving theoretical insights from observed coalition outcomes and emphasizing the role of institutions, party-system characteristics, and historically contingent context (Laver & Schofield, 1998, pp. 7-9; Warwick, 1996; Bäck & Dumont, 2007, pp. 18-22). The game-theoretic tradition, by contrast, is deductive: it derives predictions from axioms of rational choice and formal cooperative game theory, assuming that political actors seek to maximize their share of government payoffs (von Neumann & Morgenstern, 1944; Riker, 1962; Laver & Schofield, 1998, pp. 3-6). The European tradition explains deviations from minimal winning coalitions, such as surplus or minority governments, through contextual variables, while the game-theoretic tradition treats the minimal winning coalition as the equilibrium outcome of rational behavior. Each tradition has complementary strengths: the European tradition offers empirical validity and sensitivity to real-world political dynamics, while the game-theoretic tradition achieves generalizability across any weighted voting system. However, both carry corresponding limitations: the former is anchored to a relatively narrow set of post-war West European cases (Laver & Schofield, 1998, pp. 9-11), while the latter abstracts away ideology, institutional context, and procedural advantages that are central to observed coalition formation (Strøm, 1990; Glasgow, Golder & Golder, 2012, pp. 249-252).

Most importantly, the comparison reveals a shared but insufficiently specified concept at the center of both approaches: bargaining power. Both traditions recognize that coalition outcomes depend on the distribution of influence among parties, yet operationalize it in fundamentally different ways (Laver & Schofield, 1998, pp. 5-7; Martin & Stevenson, 2001, pp. 33-35). The European tradition relies predominantly on seat share as a proxy; the game-theoretic tradition formalizes power as pivotality, meaning how often a party is decisive in transforming a losing coalition into a winning one, captured through voting-power indices such as the Banzhaf index and the Shapley-Shubik value (Shapley & Shubik, 1954, pp. 789-792; Felsenthal & Machover, 2005, pp. 485-490). The divergence between seat share and observed coalition outcomes, including the frequent inclusion of small parties and exclusion of larger ones, raises a fundamental question: how should bargaining power be measured? Addressing it requires moving beyond seat-based proxies toward more precise conceptualizations of political influence.

1.3. The Measurement Problem of Seat Share

The three explanatory traditions reviewed in the preceding section each suppose that parties possess differential bargaining power in coalition negotiations, yet they offer little consensus on how such power should be measured (Laver & Schofield, 1998, pp. 5-7; Martin & Stevenson, 2001, pp. 33-35). In practice, the dominant solution across the empirical coalition formation literature has been seat share. Its appeal is straightforward: seat share is directly observable from electoral results, requires no modelling assumptions, and is available for virtually all parliamentary systems across extended time periods, making it uniquely suited to the demands of large-n comparative research (Gamson, 1961, pp. 373-374; Riker, 1962; Warwick & Druckman, 2006, p. 635). These practical advantages explain both its initial adoption and its persistence across decades of coalition scholarship. However, they do not resolve a more fundamental theoretical problem: seat share measures numerical size, not strategic necessity, and the two can deviate significantly. A party may hold a large share of parliamentary seats while being entirely non-essential for assembling any viable majority, while a much smaller party may be necessary to every feasible governing coalition (Felsenthal & Machover, 2005, pp. 486-488; Ansolabehere et al., 2005, pp. 550-551). In such cases, seat share does not merely underperform as a measure of bargaining power but actively misrepresents the distribution of strategic influence among parties.

It is important to acknowledge what seat share does explain. The empirical regularity known as Gamson's Law is described by Warwick and Druckman (2006, p. 635) as perhaps the strongest finding in political science. Research across Western European parliamentary democracies have consistently confirmed that the proportionality relationship between seat contributions and portfolio allocation holds even when ministerial salience is taken into account (Warwick & Druckman, 2006, pp. 636-638; Bäck, Meier & Persson, 2009, pp. 10-12). Seat share is therefore not a conceptually arbitrary measure: within governing coalitions, it captures a real and stable regularity in how office benefits are distributed. The analytical problem this thesis addresses, however, operates at a different level. The question of how portfolios are divided among parties already inside a coalition is analytically distinct from the prior and more fundamental question of which parties enter the coalition in the first place (Martin & Stevenson, 2001, pp. 33-35; De Winter & Dumont, 2006, pp. 175-176). A measure can perform well at explaining within-coalition payoff division while simultaneously failing to explain why a particular set of parties constitutes the coalition rather than any of the feasible alternatives. Warwick and Druckman (2006, pp. 635-637) themselves document this tension: pivotality-based measures introduce systematic deviations from strict proportionality in portfolio allocation that correlate with parties' structural positions in the legislature, suggesting that something beyond seat share is at work. Ansolabehere et al. (2005, pp. 550-552) develop this further, demonstrating that the divergence between seat shares and pivotality-based voting weights is not a measurement error but a systematic feature of fragmented multiparty legislatures, where the threshold structure of majority rule generates fundamentally nonlinear relationships between a party's numerical size and its strategic influence (Felsenthal & Machover, 2005, pp. 486-488).

Three empirical patterns documented in the literature confirm that this divergence manifests in observable coalition outcomes, each illustrating a distinct way in which seat share systematically fails as a predictor of coalition membership. The first concerns the frequency of minority governments. Strøm (1990) documents that minority cabinets constituted approximately one third of all post-war Western European governments, which is a frequency far too high to be treated as anomalous within a framework that treats seat share as the primary determinant of coalition outcomes. Their distribution was not random: minority governments were significantly more common where positive investiture requirements were absent, legislative committee systems afforded opposition parties agenda-setting influence, and the electoral costs of triggering early elections were sufficiently high to make governing without a majority strategically preferable (Strøm, 1990; Müller & Strøm, 2000). If formal majority status does not translate mechanically into coalition composition, seat share cannot be the primary driver of government formation outcomes.

The second pattern is the systematic formation of surplus majority coalitions. Across Western European parliamentary democracies, governing coalitions have routinely included more parties than strictly necessary to command a legislative majority (Laver & Schofield, 1998, pp. 75–78; Müller & Strøm, 2000). This directly challenges size-based theories: if bargaining power derives from seat share and parties are rational office-seekers, they should resist surplus partners, since doing so reduces each member's portfolio payoff without any compensating gain in legislative viability (Riker, 1962). Surplus coalitions arise primarily from ideological considerations operating independently of seat arithmetic: when parties required for a minimal winning coalition are ideologically distant, incorporating an additional but proximate partner can reduce internal policy compromise costs sufficiently to offset the portfolio losses (Laver & Schofield, 1998, pp. 77-78; Axelrod, 1970; Martin & Stevenson, 2001, pp. 34-36). Because seat share captures only the numerical dimension of parliamentary presence, it offers no analytical leverage on this calculus.

The third pattern operates at the level of individual party outcomes. Small parties frequently enter government and secure portfolio allocations exceeding Gamson's proportionality benchmark, while larger parties are occasionally excluded or receive cabinet shares falling short of their numerical weight (Laver & Schofield, 1998, pp. 193-200; Martin & Stevenson, 2001, pp. 35-36). Warwick and Druckman (2006, pp. 637-640) demonstrate that residuals from proportionality are not randomly distributed but systematic: parties whose structural position renders them pivotal for assembling a viable majority consistently receive portfolio shares exceeding seat-share predictions, while strategically dispensable parties receive correspondingly less. Ansolabehere et al. (2005, pp. 553-558), analysing over 300 coalition governments across seventeen parliamentary democracies from 1946 to 2001, confirm that pivotality-based voting weights account for systematic patterns in portfolio allocation that raw seat shares leave unexplained. The divergence between seat share and coalition outcomes is therefore not incidental but constitutive: it reflects the systematic difference between numerical size and strategic necessity that seat share is unable to bridge (Felsenthal & Machover, 2005, pp. 486-488; De Winter & Dumont, 2006, pp. 180-182).

What these three patterns collectively reveal is a conceptual amalgamation embedded in size-based approaches. In treating seat share as a measure of bargaining power, such approaches collapse two analytically distinct properties into a single number: a party's numerical size and its strategic necessity within the specific configuration of seats produced by a given election

(Felsenthal & Machover, 2005, pp. 486-488). What determines a party's leverage in coalition bargaining is not the absolute number of seats it controls, but whether its participation is required to convert a losing coalition into a winning one (Shapley & Shubik, 1954, pp. 787-788; Riker, 1962). The gap between the two can be significant. A party holding 35% of seats may be structurally redundant if viable majorities can be assembled without it, while a party holding only 8% may be indispensable if no majority threshold can be crossed in its absence. Seat share registers these parties as having very different bargaining power; a pivotality-based approach would recognize that influence flows from structural necessity rather than numerical size and might assign them scores that look nothing like their seat shares (Ansolabehere et al., 2005, pp. 550-552).

This discrepancy is illustrated with particular clarity by what cooperative game theorists call the paradox of size (Felsenthal & Machover, 2005, pp. 487-488). Consider a legislature in which three parties hold 50, 49, and 1 seat respectively out of 100, with a majority threshold of 51. Measured by seat share, the largest party appears overwhelmingly dominant while the party holding a single seat appears virtually powerless. Yet in terms of pivotality, the three parties are perfectly equal: no majority can be formed without at least two of them, and none can be removed from a winning coalition without causing it to fall below the threshold (Shapley & Shubik, 1954). As Felsenthal and Machover (2005) emphasize, such configurations are not artificial extreme cases: they arise with considerable regularity in the fragmented multiparty legislatures that characterize proportional representation systems. Whether a party is pivotal depends not on its own seat share alone but on the entire distribution of seats across the legislature, a context-dependence that seat share, by construction, cannot capture.

The formal concept that shapes this distinction is the dummy party. Within cooperative game theory, a party is a dummy if its votes are never decisive for converting a losing coalition into a winning one, irrespective of how many seats it holds (Felsenthal & Machover, 2005, pp. 486-488; Shapley & Shubik, 1954, pp. 787-788). A dummy can command a substantial seat share and yet be formally irrelevant: its presence neither enables nor prevents any majority from forming. Seat share assigns positive bargaining weight to every party with nonzero parliamentary representation, including dummies, because it has no mechanism for distinguishing decisive seats from redundant ones. A pivotality-based measure correctly assigns dummies zero influence, reflecting the game-theoretic reality that a party which never determines any coalition outcome exercises no genuine bargaining power regardless of its size

(Felsenthal & Machover, 1998, pp. 420-422). The difference between the two approaches is therefore not one of degree but of conceptual category: where dummies exist, and they are not uncommon in real multiparty legislatures, relying on seat share is not a marginal imprecision but a qualitative misspecification of where bargaining power resides (De Winter & Dumont, 2006, pp. 180-182; Glasgow, Golder & Golder, 2012, pp. 249-251).

These limitations point toward a clear measurement requirement: an operationalization of bargaining power in coalition formation must capture pivotality directly, rather than treating numerical size as its proxy. The cooperative game theory literature has developed a family of formal indices designed to precisely do this, each grounded in a distinct definition of what it means for a party to be decisive and each embedding different assumptions about the coalition formation process itself. It is to these four indices that the following section turns.

1.4. Voting-Power Indices: Theoretical Foundations and Properties

The preceding section demonstrated that seat share measures numerical size rather than strategic necessity, and that the two can diverge substantially in fragmented multiparty legislatures. The formal concept needed to bridge this gap is pivotality: a party's influence in coalition bargaining derives not from how many seats it holds, but from how often its participation is decisive for converting a losing coalition into a winning one (Shapley & Shubik, 1954, pp. 787-788; Riker, 1962). The cooperative game theory literature has developed four principal indices designed to operationalize this concept, and they divide into two families based on a fundamental assumption: should bargaining power be measured across all winning coalitions, or only across minimal winning coalitions? The first family, comprising the Banzhaf Power Index (BPI) and the Shapley-Shubik Value (SSV), defines power across every coalition commanding a majority. The second, consisting of the Deegan-Packel Index (DPI) and the Holler-Packel Index (HPI), restricts attention exclusively to minimal winning coalitions (Felsenthal & Machover, 1998, pp. 191-194; Holler & Packel, 1983, pp. 21-23). This distinction is not a technical detail but a theoretically influenced choice, because each family embeds different assumptions about which coalitions rational actors will actually form and how payoffs will be distributed within them.

All four indices are grounded in cooperative game theory with transferable utility (Shapley & Shubik, 1954, pp. 787-788; von Neumann & Morgenstern, 1944, pp. 10-18). The central analytical object is the characteristic function v, which assigns a value to every possible coalition of players. In the parliamentary application, this function takes a strictly binary form: $v(S) = 1$ if coalition S controls a legislative majority, and $v(S) = 0$ if it does not, constituting what game theorists call a simple game (Felsenthal & Machover, 1998, pp. 13-15). N denotes the set of all parliamentary parties and $W \subseteq 2^N$ the set of all winning coalitions. The majority threshold is operationalized as the standard simple majority quota $q = \lfloor T/2 \rfloor + 1$, where T is the total number of parliamentary seats, so that $S \in W$ if and only if its combined seat count meets or exceeds q. Since v(S) is binary, every coalition in W is formally equivalent regardless of the size of its majority. What really matters is crossing the threshold, not the margin by which it is crossed, precisely the property that makes pivotality the theoretically appropriate basis for measuring bargaining power.

The BPI, developed by Banzhaf (1965) and axiomatized by Dubey and Shapley (1979), conceptualizes voting power as the frequency with which a party is pivotal across all winning coalitions (Felsenthal & Machover, 1998, pp. 210-215; Straffin, 1994, pp. 1155-1156). For every winning coalition in the legislature, the BPI checks whether removing a given party would cause the coalition to lose its majority. The more often a party is that critical swing member across all possible winning coalitions, the higher its BPI score. Party $i \in N$ is pivotal in coalition $S \in W$ if $S \setminus \{i\} \notin W$. The unnormalized Banzhaf score is the total count of swing positions across all winning coalitions:

$$\beta_i = \sum 1[S \setminus \{i\} \notin W],$$

where β_i - unnormalized Banzhaf score of party i, equal to its total count of swing positions across all winning coalitions;

S - a coalition, that is, any subset of parties drawn from the full set N;

W - the set of all winning coalitions, where a coalition S is winning if its combined seat total meets or exceeds the majority quota q;

S{i} - coalition S with party i removed;

1[...] - an indicator function equal to 1 if the condition in brackets holds and 0 otherwise.

The normalized index divides each party's swing count by the total number of swings across all parties, ensuring all values sum to one:

$$BPI_i = \frac{\beta_i}{\sum_{j \in N} \beta_j},$$

where BPI_i - normalized Banzhaf score of party i;

$\sum_{\{j \in N\}} \beta_j$ - sum of unnormalized Banzhaf scores across all parties j in N, serving as the normalizing constant.

The BPI treats every coalition in which a party is decisive as equally important regardless of coalition size, making it sensitive to the sheer frequency of pivotal positions across the full winning set (Straffin, 1994, p. 1156).

The SSV, introduced by Shapley and Shubik (1954), defines voting power differently. Rather than counting swing positions across all coalitions at once, it imagines parties joining a coalition one by one in a random order and asks: how often is party i the one whose entry tips the running seat total across the majority threshold? The more often a party is that decisive last addition, averaged across all possible joining orders, the higher its SSV score (Felsenthal & Machover, 1998, pp. 191-196; Owen, 1995, pp. 265-268). By construction, SSV values sum to one across all parties (Shapley & Shubik, 1954, pp. 788-789):

$$SSV_i = \frac{1}{|N|!} \sum_{\pi} 1[P_\pi(i) \notin W \land P_\pi(i) \cup \{i\} \in W],$$

where SSV_i - Shapley-Shubik value of party i;

|N| - total number of parties in the legislature, so that |N|! is the total number of possible orderings of N;

π - a permutation of N, representing one possible sequential order in which parties join a coalition;

$P_\pi(i)$ - the set of parties that precede party i in permutation π, constituting the coalition already assembled before i joins;

$P_\pi(i) \notin W$ - the predecessor set does not yet constitute a winning coalition on its own;

$P_\pi(i) \cup \{i\} \in W$ - adding party i to the predecessor set produces a winning coalition;

$\wedge$ - logical conjunction, requiring both conditions to hold simultaneously for party i to be decisive in ordering π;

$1/|N|!$ - the probability assigned to each permutation under the assumption that all orderings are equally likely.

The implied weight on a pivotal position in a coalition of k predecessors is $k!(|N| - k - 1)!/|N|!$, which is maximized when $k \approx |N|/2$, so the SSV places greater emphasis on pivotal positions within coalitions of intermediate size.

Both the BPI and SSV are grounded in the same definition of pivotality yet can assign substantially different power scores to the same party. The source of divergence is the weighting scheme: the BPI counts all swing positions equally regardless of coalition size, while the SSV weights each decisive position by the probability of its sequential ordering (Felsenthal & Machover, 1998, pp. 215-220; Laruelle & Valenciano, 2001, pp. 83-85). The choice between them reflects an assumption about whether the sequential structure of coalition formation is theoretically relevant to the distribution of bargaining power.

The second family proceeds from a direct challenge to a foundational premise of the first. Both the BPI and SSV compute power across all winning coalitions W, including surplus coalitions considerably larger than a bare majority. Deegan and Packel (1978) and Holler (1982) argue this is theoretically unjustifiable: if parties are rational and office-seeking, they will not willingly enter coalitions larger than the minimum required, since surplus partners receive cabinet benefits that could otherwise be distributed among fewer members (Riker, 1962; Deegan & Packel, 1978, pp. 149-151). Both indices in this family are therefore computed exclusively over $M \subseteq W$, the set of minimal winning coalitions, where a coalition S belongs to M if and only if $S \in W$ and for every $j \in S$: $S \setminus \{j\} \notin W$.

The DPI assigns power based on participation in minimal winning coalitions, weighting each party inversely by coalition size (Deegan & Packel, 1978, pp. 149-152). The index rewards parties that appear in many MWCs, and especially in small ones, because in a smaller coalition each member receives a larger share of the cabinet benefits:

$$DP_i = \frac{\sum_{S \in M, i \in S} \frac{1}{|S|}}{\sum_{jN} \sum_{S \in M, j \in S} \frac{1}{|S|}},$$

where DP_i -Deegan-Packel score of party i;

M - the set of all minimal winning coalitions;

$S \in M$, $i \in S$ - the subset of MWCs that contain party i;

|S| - the number of members in MWC S;

1/|S| - the payoff received by party i from belonging to coalition S, reflecting equal division of a unit payoff among all |S| members;

denominator - the sum of all 1/|S| scores across all parties $j \in N$ and all MWCs containing j, normalising all DPI values to sum to one.

The HPI also focuses exclusively on minimal winning coalitions but does not weight parties by coalition size. Instead, it simply counts how many MWCs each party belongs to (Holler, 1982, pp. 262-264; Holler & Packel, 1983, pp. 21-23):

$$HP_i = \frac{|\{S \in M: i \in S\}|}{\sum_{j \in N} |\{S \in M: j \in S\}|}$$

where HPI - Holler-Packel score of party i, also referred to as the Public Good Index;

$\{S \in M : i \in S\}$ - the set of all MWCs containing party i;

$|\{S \in M : i \in S\}|$ - the count of MWCs containing party i;

$\sum_{\{j \in N\}} |\{S \in M : j \in S\}|$ - the total number of MWC memberships summed across all parties $j \in N$, normalising all HPI values to sum to one.

The intuition is that the benefits of coalition membership are partly a public good and all members share in policy outcomes regardless of how many partners are present, so membership in a large MWC is treated as equally valuable as membership in a small one. The DPI and HPI share the MWC restriction and differ only in their payoff assumption: the DPI treats coalition payoffs as divisible and equally distributed among members, while the HPI treats them as a public good. Where parties systematically differ in whether they belong predominantly to small or large MWCs, the two indices will assign substantially different power scores. The choice between them reflects a theoretical position on whether the value of holding office is rival or non-rival across coalition partners (Holler & Packel, 1983, pp. 21-23; Felsenthal & Machover, 1998, pp. 247-249), which is a question that cannot be resolved on axiomatic grounds alone.

Taken together, the four indices cover the principal theoretical positions in the measurement of voting power. Within Family 1, the BPI and SSV disagree on whether pivotal positions should be weighted by ordering probability or counted equally. Within Family 2, the DPI and HPI disagree on whether payoffs within MWCs are divisible or public. The between-family disagreement concerns which coalitions are relevant in the first place, a choice that maps directly onto the theoretical dispute between Riker's (1962) size principle and the empirical regularity that surplus and minority governments form with considerable frequency in real parliamentary systems (Strøm, 1990; Laver & Schofield, 1998, pp. 75-78). In combination, these distinctions can produce substantially different power assignments for the same party in the same electoral setting, making the choice of index a theoretically consequential decision that requires empirical evidence to resolve (Felsenthal & Machover, 1998, pp. 247-249; Deegan & Packel, 1978, pp. 152-154). Whether the four indices perform differently in predicting actual cabinet entry, and which consistently outperforms seat share as a predictor of government participation, is the central empirical question this thesis addresses, and answering it first requires reviewing the existing empirical record, which is the task of the following section.

1.5. Empirical Application of Voting-Power Indices

The preceding section established the theoretical case for voting-power indices as more appropriate operationalisations of bargaining power than seat share. The question remains, however, whether this theoretical advantage translates into better empirical performance, and whether the choice between the four indices produces meaningfully different predictions in practice. Resolving these questions requires turning to the empirical record of how voting-power indices have been applied across three domains: the European Union Council of Ministers, corporate governance, and national parliamentary coalition formation. These domains differ in their structural proximity to parliamentary coalition bargaining and reviewing them in sequence reveals both what is known and what remains unresolved, establishing the specific research gap that motivates the empirical analysis in Chapter 2.

The most extensively developed empirical application of voting-power indices is the EU Council of Ministers, a body whose weighted voting structure has served as a natural laboratory for power index analysis since the foundational work of Banzhaf (1965). Felsenthal and Machover (1997; 2001) applied the SSV and BPI to successive rounds of EU enlargement and treaty reform, demonstrating that the voting weight assigned to each member state under qualified majority voting rules bears a systematically asymmetric relationship to its formal power as measured by either index. Small member states consistently hold voting power exceeding their population or seat weight, while the largest states are correspondingly disadvantaged relative to their demographic contributions (Felsenthal & Machover, 2001, pp. 94-97). The EU application establishes three important points: first, that the difference between nominal weight and formal power can be substantial and systematic rather than random; second, that this divergence has real consequences for how influence is distributed among institutional actors, specifically member state delegations explicitly used power index analyses in treaty negotiations over voting weight reallocations (Deemen & Hosli, 2002, pp. 340-343); and third, that the choice between the SSV and BPI can produce different ordinal rankings of actors' relative power, particularly in large weighted-voting bodies with complex threshold structures (Felsenthal & Machover, 1998, pp. 420-422; Widgrén, 1994, pp. 1148-1150). However, the Council of Ministers is a weighted voting body operating under fixed procedural rules, not a parliamentary legislature engaged in post-electoral coalition formation, so the EU evidence cannot speak directly to whether these indices predict which parties enter national governing coalitions.

The same structural limitation applies to applications in corporate governance, where voting-power indices have been used to analyse the distribution of influence among shareholders and bloc voters in firms with concentrated ownership structures (Leech, 2002). In this setting, the BPI and SSV have been applied to shareholder voting games to assess whether the distribution of formal voting power among large shareholders reflects their nominal equity stakes, and findings broadly parallel those from the EU context: the relationship between share ownership and formal voting power is nonlinear, and small but strategically positioned shareholders can exercise influence disproportionate to their equity weight (Leech, 2002; Roth, 1988). However, corporate voting games lack the ideological dimension that characterises parliamentary coalition bargaining, involve no uncertainty about the majority threshold, and produce no post-vote coalition agreement requiring sustained policy coordination. The conditions are therefore sufficiently different from multiparty parliamentary bargaining to limit inference across domains. Taken together, the EU and corporate governance applications establish a consistent finding that formal voting power and nominal weight diverge substantially and systematically, but neither domain can speak to whether this divergence has observable consequences for government formation in parliamentary systems.

The empirical applications most directly relevant to this thesis concern the use of voting-power indices in the analysis of national parliamentary coalition formation. The large-n comparative tradition in European coalition research has relied predominantly on seat share as its primary predictor of coalition outcomes, with voting-power indices used only in scattered applications (Bäck & Dumont, 2007, pp. 18-22). Laver and Schofield (1998, pp. 193-200) computed SSV scores for parties across a sample of Western European parliamentary systems and documented that SSV distributions diverge substantially from seat share distributions. Importantly, these divergences correlate with systematic deviations from Gamson's proportionality benchmark in portfolio allocation: parties whose SSV scores exceed their seat shares tend to receive cabinet allocations outperforming seat-share predictions, while parties whose SSV scores fall short receive correspondingly fewer portfolios. This evidence was extended and systematised by Ansolabehere et al. (2005, pp. 553-558), whose analysis of over 300 coalition governments across seventeen parliamentary democracies from 1946 to 2001 found that pivotality-based measures outperform raw seat shares in explaining the observed distribution of portfolios across coalition partners, with this additional explanatory contribution statistically robust and not reducible to a rescaling of the seat share variable. Warwick and Druckman (2006, pp. 637-640) further confirm that residuals from Gamson's proportionality benchmark follow a systematic pattern consistent with pivotality: parties whose structural positions render them pivotal for assembling viable majorities consistently receive portfolio shares exceeding seat-share predictions, while strategically dispensable parties receive correspondingly less, a pattern that holds after controlling for the salience-weighted importance of individual ministries. One illustrative national case demonstrates that the pivotality logic also generates testable expectations corresponding to observable patterns in coalition entry, not only portfolio allocation: the Irish Labour Party entered coalition government in 1973 despite holding only 13.7% of Dáil seats, a configuration in which its addition to the Fine Gael-led grouping was necessary to cross the majority threshold, and in which the party accordingly received a disproportionate cabinet share relative to its seat contribution (Mair, 1987, pp. 82-84). While this case does not constitute systematic evidence, it illustrates how a party's pivotal structural position can translate into both government entry and outsized portfolio returns in ways that seat share alone would not predict.

The existing literature has, however, left three important empirical questions unresolved. First, no study has conducted a large-n cross-national comparative test evaluating all four indices against seat share as predictors of coalition entry within the same analytical framework. The most systematic existing comparisons focus on minimal integer voting weights rather than the full set of established indices, and the large-n comparative tradition in European coalition research has similarly relied on seat share rather than formal power indices as its primary predictor (Bäck & Dumont, 2007, pp. 18-22), leaving open whether the DPI and HPI family's restriction to minimal winning coalitions produces empirically superior predictions relative to the all-coalitions approach of the SSV and BPI (Felsenthal & Machover, 1998, pp. 420-422). Second, the literature has not examined systematically whether the performance advantage of pivotality-based measures over seat share varies across institutional contexts, in particular whether it is conditioned by the degree of legislative fragmentation, the stringency of investiture requirements, or the presence of electoral thresholds that truncate the effective party system (Müller & Strøm, 2000; Martin & Stevenson, 2001, pp. 37-39). If the seat share versus pivotality distinction matters most where fragmentation is high and no single party approaches majority status, then the marginal predictive contribution of voting-power indices should vary systematically with fragmentation, which is a conditional hypothesis that no existing study has tested. Third, the existing applications largely treat the prediction of coalition entry and the explanation of portfolio allocation as equivalent tasks, using evidence from one to draw inferences about the other, without consistently maintaining the analytic distinction between which parties enter government and how benefits are divided among those who do (De Winter & Dumont, 2006, pp. 175-176; Martin & Stevenson, 2001, pp. 33-35). A measure can perform well at explaining within-coalition payoff division while simultaneously failing to explain why a particular set of parties constitutes the coalition in the first place, and conflating these two questions has obscured the degree to which existing evidence actually supports the use of voting-power indices as predictors of government entry specifically. These gaps converge on the central research puzzle this thesis addresses: if voting-power indices are theoretically superior to seat share as measures of bargaining power, why has this superiority not been subjected to systematic empirical analysis across the full set of established indices and the full range of coalition outcomes? Resolving this tension requires a direct comparative test that evaluates whether the theoretical advantages of voting-power indices translate into observable gains in predicting which parties actually enter government.

The theoretical gaps identified above point to two testable directional expectations that follow directly from the preceding analysis. The empirical record reviewed in this section, documenting the high frequency of surplus majority coalitions and minority governments across European parliamentary democracies, establishes that the core assumption underlying Family 2 indices, namely that rational parties form only minimal winning coalitions, is systematically violated in practice. Family 1 indices carry no such restriction and measure pivotality across the full range of winning coalitions, including the surplus configurations that constitute a substantial share of observed outcomes. This asymmetry generates the following two hypotheses, which are tested in Chapters 2 and 3*:*

H1: *the all-coalitions family (the Banzhaf Power Index and the Shapley-Shubik Value) will outperform seat share as predictors of government entry, because they capture pivotality across the full range of winning coalitions, including the surplus coalitions that are empirically common in the parliamentary democracies under study.*

H2: *the minimal-winning-coalition family (the Deegan-Packel Index and the Holler-Packel Index) will not outperform seat share, because their restriction to minimal winning coalitions embeds an assumption about coalition formation that is systematically violated in the empirical record.*

Testing these hypotheses requires moving from the theoretical domain to the empirical one, which is the task of Chapter 2.

2. VOTING-POWER INDICES AS PREDICTORS OF GOVERNMENT ENTRY: A CROSS-NATIONAL TEST

2.1. A Cross-National Quantitative Comparison across 37 Parliamentary Democracies

This thesis employs a quantitative comparative research design to evaluate whether cooperative game-theoretic voting-power indices improve the empirical explanation of government formation relative to models based on parliamentary seat share alone. The central empirical question, whether replacing or supplementing seat share with formal indices of pivotality produces better predictions of which parties enter government, is inherently comparative in two senses. Firstly, it requires comparison across a large number of cases, because the predictive performance of any measure can only be assessed reliably when variation in outcomes is observed across many elections and many party systems. Secondly, it requires comparison across alternative measures, testing whether different operationalisations of the same underlying concept, defined as bargaining power, yield different explanatory results. A quantitative approach is therefore not merely convenient but logically necessary: the thesis aims to evaluate the marginal contribution of power indices over seat share through formal model comparison, and this requires sufficient observations for statistical inference to be meaningful.

The choice of a large-n cross-national design over a focused case study reflects the nature of the research question. Case studies of individual parliaments, such as existing analyses of coalition bargaining in the Greek or Israeli Knesset, can demonstrate that power indices are technically computable and that they diverge from seat share in specific configurations, but they cannot establish whether this divergence translates into systematic predictive gains across varied institutional and party-system contexts (Koki & Leonardos, 2019). If voting-power indices truly capture a dimension of bargaining strength that seat share misses, this should be detectable as a consistent pattern across a wide range of parliamentary democracies and not merely in selected cases chosen because their party systems are known to produce unexpected index values. Conversely, if the predictive gain is small or absent across a broad comparative sample, that finding is itself substantive: it would suggest that the theoretical advantages of formal pivotality measures do not translate into observable improvements in explanatory power at the national level.

The population of cases is defined as national parliamentary elections in multiparty democracies for which complete seat-distribution data and cabinet-composition records are available. The research problem of this thesis is only meaningful in systems where coalition bargaining is structurally necessary, which requires three conditions to hold simultaneously: the executive must depend on parliamentary confidence, no single party must regularly command an outright majority, and governments must be formed through inter-party negotiation over a seat distribution that genuinely constrains the feasible coalition space. These conditions are most reliably met in established parliamentary democracies with proportional or mixed electoral systems and stable, institutionalised party systems, which is the precise population that ParlGov was designed to cover (Döring & Manow, 2024). Presidential and semi-presidential systems where the executive is independently elected are excluded because cabinet entry is not determined by parliamentary coalition bargaining in the same sense, making the outcome variable theoretically ill-defined in those contexts. Majoritarian systems that regularly produce single-party majorities are excluded by the data-cleaning filter described in Section 2.2, since elections in which one party already holds an outright majority present no coalition formation problem to which power indices are relevant.

The unit of analysis is the party-election observation: a single political party as it appears in a single national election. This choice reflects the structure of the research question. Government formation is a party-level outcome, and the predictors of interest are likewise defined at the party level for a given election. The party-election unit is standard in the comparative coalition literature, used most prominently by Martin and Stevenson (2001) in their foundational analysis of government formation across Western Europe, and it allows the dataset to pool variation both within countries over time and across countries at the same time point. The analysis therefore treats each party-election as an independent observation while acknowledging that this assumption is a simplification: parties in the same election are not statistically independent of one another, since their seat shares and index values are mechanically linked through the seat distribution, and the same party observed across multiple elections introduces serial dependence. The consequences of this simplification are discussed in Section 3.5.

The decision to use logistic regression as the primary estimator follows directly from the binary nature of both dependent variables. A party either enters government or it does not, and the logistic model estimates the log-odds of the positive outcome as a linear function of the predictors, producing fitted probabilities bounded between zero and one. Alternative approaches such as linear probability models were considered but rejected on the grounds that they can produce predicted probabilities outside the unit interval, which is a particular concern given the skewed base rates in the data as only 30.2% of observations are coded 1 on the primary outcome. Multilevel models with country or election random effects were also considered as a way of addressing the non-independence of observations within elections and within countries; however, since the goal of the analysis is to compare the predictive performance of alternative measures rather than to produce unbiased causal estimates of individual coefficients, the added complexity of multilevel specifications was not warranted for the primary goal of this analysis, which is comparative model evaluation rather than unbiased causal estimation of individual coefficients. Multilevel models with country or election random effects would improve the precision of standard errors by accounting for the clustering of party observations within elections and elections within countries, but they would not alter the rank ordering of fit statistics across specifications, which is the basis of the main findings. The four-criteria evaluation framework used throughout Chapter 3, combining R^2 gain, likelihood ratio test, BIC, and AUC, mitigates the consequences of this simplification by not relying solely on p-values. Clustered standard errors at the election level remain the recommended correction for future replications of this analysis. The nested model comparison framework described in Section 2.3, combined with the BIC penalty for complexity, provides adequate control for overfitting without requiring the full apparatus of hierarchical modelling.

The time period of analyzed cases is anchored to the post-war period from 1945 onwards: this marks the re-establishment of stable democratic institutions across most of the countries in the sample, and the post-war period has been used as a standard starting point for cross-national democratic data collection on precisely these grounds (Boix, Miller & Rosato, 2013). ParlGov coverage also becomes substantially completed and more consistent from 1945, reducing measurement error in key variables (Döring & Manow, 2024). Following the removal of parties with zero or missing seat counts, the sum of remaining party seat totals was verified to equal the recorded parliament size for all 626 elections in the retained sample, with no discrepancies detected. The full list of countries, together with the number of elections and party-election observations contributed by each, is provided in Appendix A. The resulting dataset contains 4,939 party-election observations drawn from 626 elections. This breadth is a deliberate methodological choice: the goal is to evaluate predictive performance across the full range of cases for which the analysis is feasible, rather than to restrict the sample to conditions under which power indices are expected to perform best.

2.2. Data Sources and Construction of the Analytical Dataset

The empirical analysis draws on the ParlGov database (Döring & Manow, 2024), a publicly available comparative dataset covering national elections, parliamentary seat distributions, party characteristics, and cabinet compositions across European parliamentary and OECD democracies. The 2024 release records elections and governments from the late nineteenth century through 2023. For the purposes of this thesis, the database supplies three categories of information: the seat distribution among parties following each election, the composition of every cabinet formed after each election, and unique identifiers linking parties, elections, and cabinets across all three tables. ParlGov is organised around three core relational tables. The party table records individual parties, their country of origin, and their ideological placement on a standard left-right scale. The election table records each parliamentary election and lists every party that won seats in parliament, including the number of seats obtained and the total size of the parliament. The cabinet table records every government formed, linked to the election that produced the parliament from which it emerged, and lists the parties that participated in each cabinet along with the cabinet's start date.

The analytical dataset was constructed from the raw ParlGov tables in R through the following sequence of cleaning, filtering, and merging operations, described in sufficient detail to permit replication. Step 1 restricted the election sample to national lower-house parliamentary elections, excluding European Parliament elections, regional elections, and any other sub-national or supranational contests. Step 2 collapsed the election table to one row per party per election. A small number of duplicate records arise in the raw data from alliance or list arrangements where the same party appears under multiple list names in the same election; these were identified by duplicate combinations of election identifier and party identifier, and seat counts were summed before collapsing to the party's canonical ParlGov identifier. Party seat counts were then used to compute each party's seat share as the ratio of its seats to the total seats recorded for that parliament (seat_share = seats / seats_total). Following the removal of parties recorded with zero seats or missing seat counts, the recorded seat totals were checked to verify that the sum of remaining party seat counts still adds up to the parliament size variable for all retained elections; where residual discrepancies arose from unaffiliated members or independents not assigned to any party, the parliament size variable was adjusted accordingly. Parties with zero seats or missing seat counts were removed from the sample, since they hold no weight in the weighted voting game from which power indices are computed.

Step 3 removed elections in which a single party held an outright parliamentary majority, defined as controlling more than fifty percent of seats without any coalition partner. In such cases the concept of a weighted voting game with genuine coalition bargaining is inapplicable, and power indices either cannot be meaningfully computed or collapse to trivial values. Step 4 addressed computational feasibility. Power indices are computed by full combinatorial enumeration of all possible subsets of parties holding seats in a given parliament, a procedure whose computational cost grows exponentially in the number of parties with seats. With n parties there are 2^n possible subsets; for $n = 20$ this is approximately one million subsets, and for $n = 30$ it exceeds one billion. To ensure computational feasibility, thresholds were specified in advance at 22 parties holding seats for the Banzhaf and Shapley-Shubik calculations, and 20 parties for the Deegan-Packel and Holler-Packel calculations. In practice, the maximum number of parties holding seats in any election in the retained sample is 20, recorded in the 1992 Italian and 1991 Polish elections, meaning these thresholds were not binding and no elections were excluded on computational grounds. All 626 elections were included in all four index calculations, and all observations in the final dataset have complete data on all four indices. After cleaning the data, the thesis uses the dataset with values described in Table 2.

Table 2

Descriptive Summary of the Analytical Dataset

Statistic	Value
Toal party-election observations	4,939
Number of elections	626
Number of countries	37
Covered time period	1945-2023
In first post-election cabinet (in_gov)	30.2% (n = 1,490)
In any cabinet during electoral term (in_gov_cycle)	35.8% (n = 1,769)

Note: Temporal coverage refers to the post-war analytical sample; a small number of pre-1945 elections recorded in ParlGov are excluded on the grounds described in Section 2.1. The first block of rows reports sample scope; the second block reports the proportion of party-election observations entering government under each outcome definition.

Source: Compiled by the author based on the ParlGov 2024 dataset.

As shown in Table 2, the dataset contains 4,939 party-election observations drawn from 626 elections across 37 countries, covering the period from 1945 to 2023. Government entry in the first post-election cabinet (in_gov) is recorded for 30.2% of observations, while participation in any cabinet during the electoral term (in_gov_cycle) is somewhat higher at 35.8%, reflecting parties that entered reshuffled or replacement governments without being in the original cabinet.

A structural variable, in_mwc, records whether a party belongs to at least one minimal winning coalition in the parliament produced by its election; 93.1% of observations are coded 1 on this measure, reflecting the fact that in most parliamentary configurations the large majority of parties with meaningful seat holdings belong to at least one viable majority subset. Yet structural eligibility and actual government entry diverge sharply: while nearly all parties belong to at least one MWC, only 30.2% enter the first post-election cabinet. Explaining this gap, and whether formal measures of pivotality help close it, is the central empirical question of the thesis.

Two binary outcome variables are used in the regression analysis. The primary dependent variable, in_gov, records whether a party entered the first cabinet formed after the election, identified by its start date within ParlGov's cabinet table; where two cabinets share the same start date, the lower cabinet identifier is used as a tiebreaker. The second dependent variable, in_gov_cycle, records whether a party entered any cabinet formed during the full electoral term. This broader operationalisation captures participation in cabinets that formed following government reshuffles, collapses, or the replacement of one coalition by another during the same legislative period and is included as a robustness check.

The descriptive contrast between in_mwc and in_gov is substantively informative. Table 3 cross-tabulates the two variables.

Table 3

Cross-tabulation of MWC Membership (in_mwc) and First-Cabinet Entry (in_gov)

	Not in cabinet (in_gov=0)	In cabinet (in_gov=1)	Row total
Not in any MWC (in_mwc=0): n	309	20	329
Not in any MWC (in_mwc=0): % within	93.9%	6.1%	100%

In at least one MWC (in_mwc=1): n	3,086	1,470	4,556
In at least one MWC (in_mwc=1): % within	67.7%	32.3%	100%
Column total	3,395	1,490	4,885

Note: Row percentages sum to 100% within each in_mwc category. Total N = 4,885; 54 observations are excluded due to missing values on in_gov. MWC = minimal winning coalition.

Source: Compiled by the author based on the ParlGov 2024 dataset.

Table 3 illustrates that MWC membership is a necessary but far from sufficient condition for cabinet entry: of the 329 parties belonging to no MWC, only 20 (6.1%) entered the first post-election cabinet, confirming that structural exclusion from all winning coalitions is an almost insuperable barrier to government participation. Among the 4,556 parties belonging to at least one MWC, however, only 32.3% entered government, meaning that the large majority of structurally eligible parties were nonetheless excluded from the specific cabinet that formed. The table makes visible the scale of the gap between structural eligibility and political realisation that the regression analysis is designed to investigate.

All data construction, merging, and cleaning operations were carried out in R using the tidyverse package (Wickham et al., 2019), which provided the filtering, grouping, and joining functions used across all dataset construction steps described above. The four voting-power indices were computed through full combinatorial enumeration of all 2^n subsets of parties within each election, implemented via custom functions in base R without reliance on any external package. For each election, party seat counts were assembled into an integer weight vector and the majority quota was set to floor (0.5 × total_seats) + 1. Subset sums, bitmasks, and subset sizes were computed iteratively across the weight vector; a coalition was classified as winning if its summed seat count met or exceeded the quota, and as minimal winning if it was winning and removing any single member caused it to fall below the threshold. The Banzhaf index was computed by counting, for each party, the winning coalitions of the remaining parties in which adding that party crosses the majority threshold, then normalising by the total swing count across all parties. The Shapley-Shubik value weighted each pivotal position by the coefficient $k!(n-k-1)!/n!$, computed via lgamma() for numerical stability, where k is the number of parties already assembled before the pivotal party joins. The Deegan-Packel index assigned each party the sum of $1/|S|$ across all minimal winning coalitions containing it, normalised to sum to one; the Holler-Packel index counted the number of minimal winning coalitions containing each party and normalised accordingly. All four indices are measured on a common [0,1] scale, directly comparable to seat share. Logistic regression models were estimated using base R's glm() function with family = binomial, and likelihood ratio tests were obtained via anova() with test = "Chisq"; McFadden pseudo-R^2 values were extracted using the pR2() function from the pscl package (Jackman, 2020). Logistic regression models were estimated using base R's glm() function with family = binomial, and likelihood ratio tests were obtained via anova() with test = "Chisq"; McFadden pseudo-R^2 values were extracted using the pR2() function from the pscl package (Jackman, 2020). All figures were produced using the ggplot2 package (Wickham, 2016), which is loaded as part of tidyverse.

2.3. Model Specification and Evaluation Criteria

The analytical strategy of this thesis is built around a systematic comparison of nested logistic regression models. For each combination of dependent variable and power index, three models are estimated. The index-only model regresses the dependent variable on the power index alone (DV ~ index), allowing direct assessment of how each index performs as a standalone predictor relative to seat share. The baseline model regresses the dependent variable on seat share alone (DV ~ seat_share), establishing the performance benchmark that the existing literature implicitly relies upon. The joint model adds the power index as a second predictor alongside seat share (DV ~ seat_share + index), asking specifically whether a power index adds explanatory value over and above what seat share already captures. By retaining seat share in both the baseline and the joint model, the joint specification isolates the contribution of the variation in each index that is independent of seat share — that is, the variation arising from the strategic structure of the parliament rather than from raw numerical size. If a power index were simply a rescaled version of seat share carrying no additional information, its inclusion in the joint model would produce no improvement in fit. Any improvement that does appear is therefore attributable to genuine structural information about pivotality that seat share cannot capture.

With two dependent variables (in_gov and in_gov_cycle) and four indices (Banzhaf, Shapley-Shubik, Deegan-Packel, Holler-Packel), the full analysis estimates twenty-four models organised into eight baseline-joint pairs, each accompanied by a corresponding index-only model. Each pair is estimated on the same analytical sample, ensuring that all model comparisons within a pair are made on identical data and that differences in fit cannot be attributed to differences in the observations included. Table 4 provides an overview of the full set of model specifications and the evaluation criteria applied to each pair.

Table 4

Overview of Estimated Models in the Analysis

DV	Index	Index-only model	Baseline Model	Joint Model	Evaluation Criteria
in_gov	BPI	DV ~ BPI	DV ~ seat_share	DV ~ seat_share + BPI	R^2 gain, LR test, ΔBIC, AUC gain
	SSV	DV ~ SSV		DV ~ seat_share + SSV	
	DPI	DV ~ DPI		DV ~ seat_share + DPI	
	HPI	DV ~ HPI		DV ~ seat_share + HPI	
in_gov_cycle	BPI	DV ~ BPI		DV ~ seat_share + BPI	
	SSV	DV ~ SSV		DV ~ seat_share + SSV	
	DPI	DV ~ DPI		DV ~ seat_share + DPI	
	HPI	DV ~ HPI		DV ~ seat_share + HPI	

Note: Each DV–index combination produces one index-only model, one baseline model, and one joint model, all estimated on the same analytical sample. Evaluation criteria applied to each baseline–joint pair: R^2 gain = difference in McFadden pseudo-R^2 between joint and baseline model; LR test = likelihood ratio test of the null hypothesis that the index coefficient is zero; ΔBIC = BIC of joint model minus BIC of baseline model; AUC gain = difference in area under the ROC curve between joint and baseline model. The same criteria are applied to compare index-only and baseline models.

Source: Compiled by the author based on the ParlGov 2024 dataset.

Each pair presented in Table 4 shares the same analytical sample, ensuring that model comparisons within each pair are made on identical data. This is particularly important given the computational thresholds described in Section 2.2, which produce slightly different sample sizes for the two index families; all comparisons are therefore made within families rather than across them.

Four criteria are used to evaluate model performance and compare specifications within each pair, and they are designed to be complementary rather than redundant. Each captures a different dimension of what it means for a model to fit the data better, and their convergence or divergence across the eight pairs is itself informative about the nature and robustness of any predictive advantage.

McFadden's pseudo-R^2 measures the proportional reduction in log-likelihood relative to an intercept-only null model and provides a summary measure of overall model fit analogous to R^2 in linear regression, though it is not directly interpretable as a proportion of variance explained (McFadden, 1974). The key quantity reported throughout Chapter 3 is the R^2 gain, the difference between the joint and baseline pseudo-R^2 values, which measures how much of the remaining unexplained variation in government entry is captured by adding the index to a model that already includes seat share. This quantity directly addresses the research question: a positive R^2 gain means the index adds predictive value over seat share; a zero or negative gain means it does not.

The likelihood ratio test assesses whether the improvement in fit produced by the joint model is statistically significant, testing the null hypothesis that the added index coefficient is zero under a chi-squared distribution with one degree of freedom. Statistical significance alone is, however, insufficient as an evaluation criterion. With nearly 5,000 observations, even trivially small improvements in log-likelihood can reach conventional significance thresholds, and a statistically significant LR test does not guarantee that the improvement is practically meaningful. The LR test is therefore interpreted jointly with the other criteria throughout Chapter 3, and cases where it reaches significance, but other criteria do not are flagged explicitly.

The Bayesian Information Criterion penalises model complexity by imposing a cost for each additional parameter estimated, equal to log(n) per parameter (Raftery, 1995). The key quantity is ΔBIC (the BIC of the joint model minus the BIC of the baseline) where negative values indicate that the improvement in fit outweighs the complexity penalty and positive values indicate that the added parameter is not earning its keep. Values of ΔBIC below -10 are conventionally interpreted as strong evidence in favour of the more complex model (Raftery, 1995), and this threshold is used as a benchmark throughout the analysis. The BIC criterion is particularly important in the present context because it guards against the possibility that the LR test reaches significance not because a power index is genuinely informative but simply because the large sample size amplifies even trivial effects.

Finally, the area under the ROC curve evaluates discriminative ability, or the model's capacity to assign higher predicted probabilities to government entrants than to non-entrants across all possible classification thresholds. An AUC of 0.5 indicates no discriminative ability beyond chance; an AUC of 1.0 indicates perfect discrimination. The AUC is particularly well suited to the present application because it is insensitive to the base rate of the outcome variable, making it directly comparable across specifications estimated on samples with different proportions of government entrants. It also evaluates the full ranking of predicted probabilities rather than performance at a single threshold, which means it captures genuine improvement in discrimination rather than artefacts of threshold choice. The key quantity reported is the AUC gain, the difference between the AUC of the joint model and the AUC of the baseline, with the sign indicating whether adding the index improves or worsens rank-order discrimination relative to seat share alone.

These four criteria are deliberately chosen to be sensitive to different potential weaknesses of a predictive model. The R^2 gain and LR test are sensitive to improvements in overall log-likelihood fit. The BIC penalises complexity and guards against overfitting. The AUC evaluates rank-order discrimination independently of threshold choice and base rate. A power index that clears all four criteria provides the strongest possible evidence of a meaningful and robust predictive contribution. An index that clears only some, for instance, reaching LR significance but increasing BIC, or improving R^2 but reducing AUC, provides a more ambiguous signal, and such cases are interpreted with corresponding caution in Chapter 3. The convergence of all four criteria for the first-family indices, and their consistent failure for the second-family indices, is what gives the main finding its evidential weight.

3. ANALYSIS AND RESULTS

3.1. Descriptive Statistics and Predictor Correlations

The analytical dataset contains 4,939 party-election observations drawn from 626 elections across 37 countries, covering the period from 1945 to 2023. Country coverage is uneven, reflecting both the depth of ParlGov's historical coverage and variation in the number of parties contesting elections across systems. The number of observations per country ranges from 13 to 397, with a median of 95, meaning that larger, more electorally active democracies with longer post-war records contribute substantially more observations than smaller or more recently democratised systems. Germany, Italy, the Netherlands, and the Nordic countries are among the most heavily represented, while newer democracies in Southern and Eastern Europe contribute fewer observations. This distributional imbalance is a consequence of the data source rather than a selection decision, and it means that results from the pooled analysis are implicitly weighted toward Western European parliamentary democracies with continuous democratic histories since 1945. This should be kept in mind when interpreting the generalisability of the findings.

The two dependent variables have markedly different base rates. Government entry in the first post-election cabinet (in_gov) is coded 1 for 30.2% of observations, while participation in any cabinet during the electoral term (in_gov_cycle) is somewhat higher at 35.8%, reflecting parties that entered reshuffled or replacement governments without being in the original cabinet. The difference between the two rates, approximately 5.6 percentage points, represents parties that were not part of the initial coalition but joined a successor or reshuffled government during the same legislative period. This pattern is most common in systems characterised by frequent government turnover, such as Italy and Belgium, where the composition of the executive can shift substantially within a single electoral term. The base rate of 30.2% for in_gov is moderately imbalanced, roughly one government entrant for every two and a half non-entrants, which is relevant for the interpretation of the AUC as the primary discriminative criterion, since AUC is robust to class imbalance in a way that accuracy-based measures are not.

As established in Section 2.2, 93.1% of observations belong to at least one minimal winning coalition (in_mwc), a structural property that stands in sharp contrast to the much more selective political outcome of actual cabinet entry. The gap between these two figures, which is 93.1% structurally eligible versus 30.2% entering government, is the fundamental empirical puzzle the analysis seeks to illuminate. It confirms that structural eligibility for government, in the formal game-theoretic sense of belonging to at least one viable majority coalition, is a necessary but very far from sufficient condition for cabinet entry. The large majority of the selection process operates through political mechanisms that the seat distribution alone cannot capture, and it is precisely this residual gap that the voting-power indices are asked to help explain.

All five predictors have identical means of 0.127, a mechanical consequence of all five being normalised to sum to one within each election. The distributions differ substantially, however, in their spread and upper tail. Seat share reaches a maximum of 0.497, just below the majority threshold, with a strongly right-skewed distribution: most parties hold modest shares, and a small number of large parties dominate each legislature. The Banzhaf index has a maximum of 0.998, indicating that in some elections a single party captures almost all swing positions while holding less than a majority of seats, directly illustrating the nonlinearity between seat share and pivotality discussed in Section 1.3. The Shapley-Shubik value has a somewhat lower maximum of 0.867, reflecting its weighting of pivotal positions by coalition-formation probability, which smooths extreme values relative to the flat Banzhaf count. The MWC-based indices are more compressed: Deegan-Packel reaches a maximum of 0.467 and Holler-Packel of 0.333, consistent with their restriction to minimal winning coalitions. The standard deviations follow the same ordering: Banzhaf is most dispersed at 0.149, followed by Shapley-Shubik at 0.145, while Deegan-Packel (0.092) and Holler-Packel (0.084) are considerably more compressed than seat share (0.131). This ordering matters for the regression analysis: indices with greater dispersion relative to seat share carry more potential information about variation in government entry, but as Section 3.2 documents, the MWC-based indices' greater compression does not explain their poor predictive performance, and their failure is substantive rather than a consequence of restricted range.

Table 5 reports the pairwise Pearson correlations among all five predictors.

Table 5

Pairwise Pearson Correlations Among All Five Predictors

	Seat Share	BPI	SSV	DPI	HPI
Seat Share	1.000				
BPI	0.891	1.000			
SSV	0.909	0.995	1.000		
DPI	0.779	0.843	0.844	1.000	
HPI	0.708	0.746	0.752	0.983	1.000

Note: All correlations are computed over the sample of 4,939 complete observations. Upper triangle omitted by symmetry.

Source: Compiled by the author based on the ParlGov 2024 dataset.

Several patterns emerge clearly. Within the first index family, the correlation between Banzhaf and Shapley-Shubik is 0.995, which is effectively a near-perfect linear relationship, meaning the two all-coalitions indices are empirically almost indistinguishable in this dataset. Within the second family, Deegan-Packel and Holler-Packel are similarly near-identical at 0.983. The correlations between seat share and the four indices follow a clear gradient: seat share correlates most strongly with Shapley-Shubik (0.909) and Banzhaf (0.891), somewhat less with Deegan-Packel (0.779), and least with Holler-Packel (0.708). This gradient reflects the theoretical distinction between the two index families: the all-coalitions indices track seat share most closely because large parties tend to be pivotal across many winning coalitions; the MWC-based indices diverge more because the minimal winning coalition structure can elevate small parties and suppress large ones relative to their seat weight. Critically, this gradient also implies that the MWC-based indices carry more information that is structurally independent of seat share than the all-coalitions indices do. If structural independence from seat share were sufficient for predictive gain, Deegan-Packel and Holler-Packel should outperform Banzhaf and Shapley-Shubik. However, they do not, as Section 3.2 documents, and this is the central empirical puzzle that will be addressed further.

3.2. Regression Results and Model Fit Across Eight Baseline-Joint Pairs

The results of the logistic regression analysis are summarised in Figure 1, which plots the predicted probability of government entry as a function of each index value, with seat share held at its median and the dashed grey line representing the seat-share-only baseline. Full model fit statistics for all twenty-four specifications, including McFadden R^2, AUC, ΔBIC, and LR test results, are reported in Appendix B. The results divide clearly along one axis: the two index families behave very differently, and this difference is consistent across both outcome variables and all four evaluation criteria.

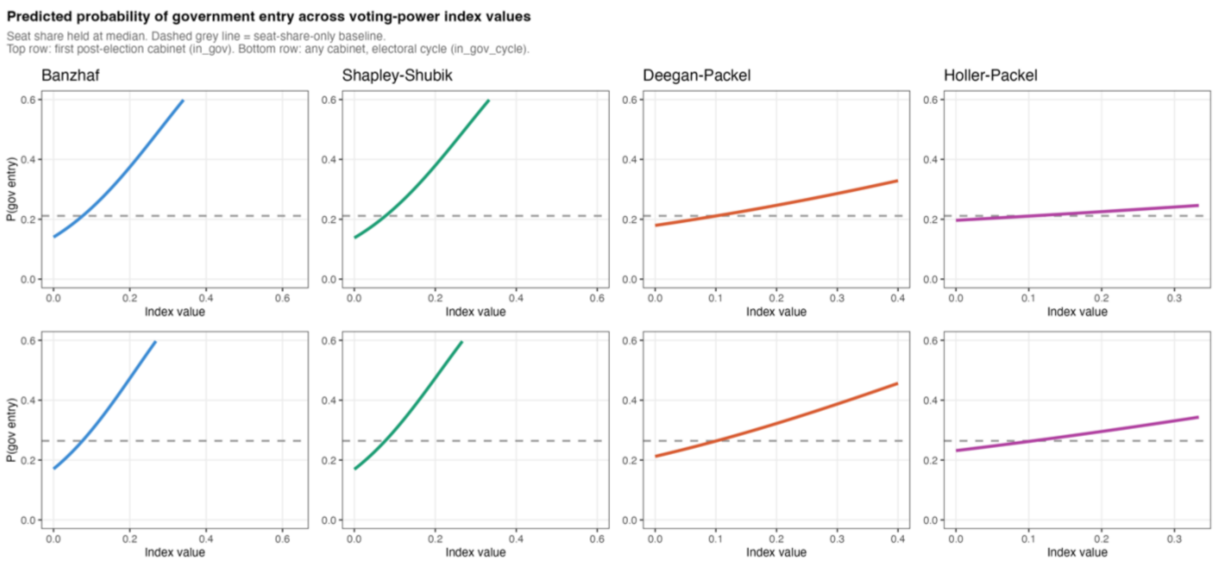

Figure 1. Predicted Probability of Government Entry Across Voting-Power Index Values

Note: Seat share held at median. Dashed grey line = seat-share-only baseline. Top row: first post-election cabinet (in_gov). Bottom row: any cabinet during the electoral term (in_gov_cycle). Full model comparison statistics on which these curves are based are reported in Appendix B.

Source: Compiled by the author based on the ParlGov 2024 dataset.

For in_gov, the seat-share baseline achieves a McFadden R^2 of 0.151 and an AUC of 0.789, a moderate level of predictive performance consistent with the established finding that seat share is a useful but insufficient predictor of government entry (Martin & Stevenson, 2001). The index-only columns reveal an important preliminary pattern before even considering the joint models. Banzhaf and Shapley-Shubik used alone achieve AUCs of 0.794 and 0.792 respectively, essentially matching and marginally exceeding seat share's 0.789 in isolation.

This indicates that the all-coalitions indices are at least as informative about government entry as seat share when used as standalone predictors, which is a finding that is consistent with the theoretical argument in Chapter 1 that pivotality captures something seat share misses, but that the two measures are strongly enough correlated that neither dramatically dominates the other when used alone. Deegan-Packel (0.731) and Holler-Packel (0.698) tell a strikingly different story: both perform noticeably worse than seat share even in isolation, before any conditioning on seat share has occurred. This confirms that the failure of the MWC-restricted family is not a multicollinearity artefact of the joint model, since these indices simply carry less information about government entry than seat share does, even when given the opportunity to predict on their own terms.

Turning to the joint models, adding Banzhaf to the seat-share baseline produces an R^2 gain of 0.026, an LR statistic of 153.3 ($p < 0.001$), a BIC reduction of 144.8 points, and an AUC gain of 0.008. All four criteria are cleared comfortably: the R^2 gain is substantively meaningful, the LR test is highly significant, the BIC reduction far exceeds the conventional threshold of 10 points, and the AUC improvement reflects genuine improvement in rank-order discrimination rather than a statistical artefact. Shapley-Shubik performs almost identically: R^2 gain of 0.025, BIC reduction of 139.4, AUC gain of 0.005, which is exactly what the near-perfect correlation between the two indices ($r = 0.995$) would lead us to expect. The near-identical performance of BPI and SSV across all criteria confirms that the two all-coalitions indices are effectively interchangeable as predictors in this dataset, and that the choice between them is unlikely to affect any substantive conclusion.

The second-family indices tell a sharply different story. Deegan-Packel generates an R^2 gain of only 0.002 (one thirteenth of the Banzhaf gain) and although the LR test reaches significance ($p < 0.001$), this reflects the large sample size rather than a practically meaningful improvement in fit. The BIC increases by 4.4 points, meaning the added complexity of including Deegan-Packel is not justified by the marginal improvement in log-likelihood. Most tellingly, the AUC falls from 0.789 to 0.787: adding Deegan-Packel to the model slightly worsens discriminative accuracy relative to seat share alone. Holler-Packel performs no better. It is the only index across all sixteen specifications for which the LR test fails to reach significance ($p = 0.113$), the BIC rises by 6.0 points, and the AUC again falls marginally. By every criterion, the second-family indices either add nothing or marginally harm predictive performance when combined with seat share.

The pattern for in_gov_cycle is nearly identical, with the first-family gains modestly larger and the second-family results equally consistent. Banzhaf produces an R^2 gain of 0.028 and a BIC reduction of 173.4 points; Shapley-Shubik produces gains of 0.027 and 162.1 respectively, both clearing all four criteria comfortably and by even wider margins than for the primary outcome. For in_gov_cycle, Deegan-Packel achieves a negative ΔBIC of −17.0, which would ordinarily suggest the joint model is preferred, but its AUC falls from 0.795 to 0.791, meaning the log-likelihood improvement does not translate into better rank-order discrimination. This divergence between BIC and AUC for Deegan-Packel in the in_gov_cycle specification is itself informative: it suggests that the index is capturing some systematic pattern in the data, enough to reduce the BIC, but that this pattern does not correspond to the variation in government participation that matters for discrimination between entrants and non-entrants. Holler-Packel's ΔBIC is negligible at −0.7 and its AUC again falls slightly to 0.792. The robustness of the pattern across both DVs rules out the possibility that the results are an artefact of how government participation is defined. Three conclusions follow from the full set of results in Figure 1 and Appendix B: Banzhaf and Shapley-Shubik add genuine, statistically robust, and practically confirmed predictive value over seat share for both outcome variables; Deegan-Packel and Holler-Packel do not; and even the best-performing indices produce modest absolute gains, leaving the large majority of variation in government entry unexplained, however, this point will be returned to in Section 3.4.

3.3. ROC Curves and Discriminative Accuracy

While the regression results in Section 3.2 establish whether adding a power index improves statistical fit, the ROC analysis addresses a more direct and, in some ways, more demanding question: does that improvement translate into better ability to distinguish government entrants from non-entrants across the full range of classification thresholds? A model can fit the data better in log-likelihood terms while still assigning nearly identical predicted probabilities to entrants and non-entrants, in which case the improvement is statistically detectable but practically meaningless. The AUC is the appropriate check on this because it evaluates the full ranking of predicted probabilities rather than performance at a single threshold, and because it is insensitive to the base rate of the outcome variable. Figures 2 and 3 present the ROC curves for in_gov and in_gov_cycle respectively. In both figures the seat-share baseline is plotted as a thick black line to make it clearly distinguishable from the index curves, the all-coalitions family (BPI, SSV) in solid blue lines, and the MWC family (DPI, HPI) in dashed red lines.

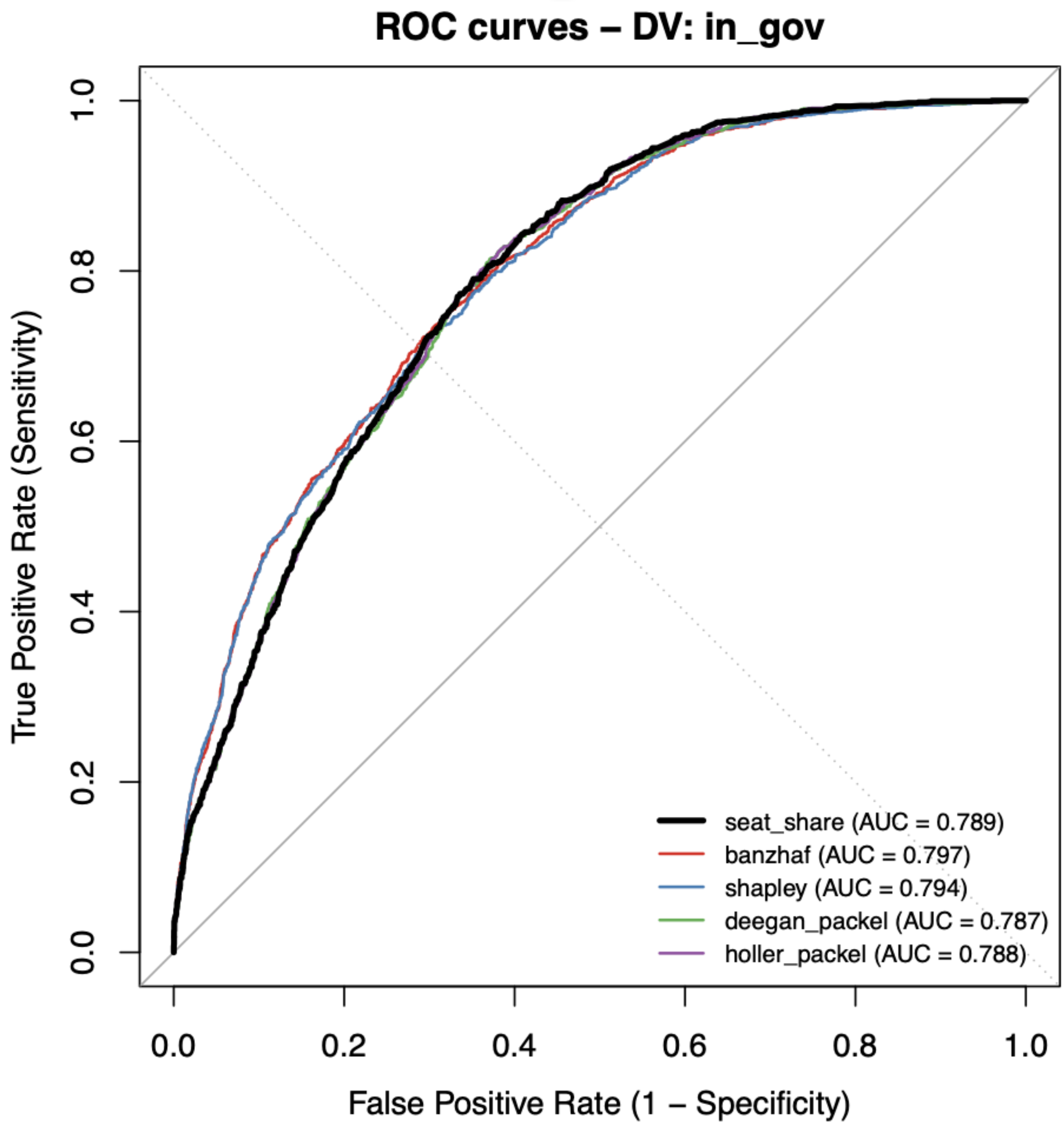

Figure 2. ROC Curves - DV = in_gov (first post-election cabinet)

Note: Thick black line = seat-share baseline. Solid blue = all-coalitions family (BPI, SSV). Dashed red = MWC family (DPI, HPI). AUCs shown in legend.

Source: Compiled by the author based on the ParlGov 2024 dataset.

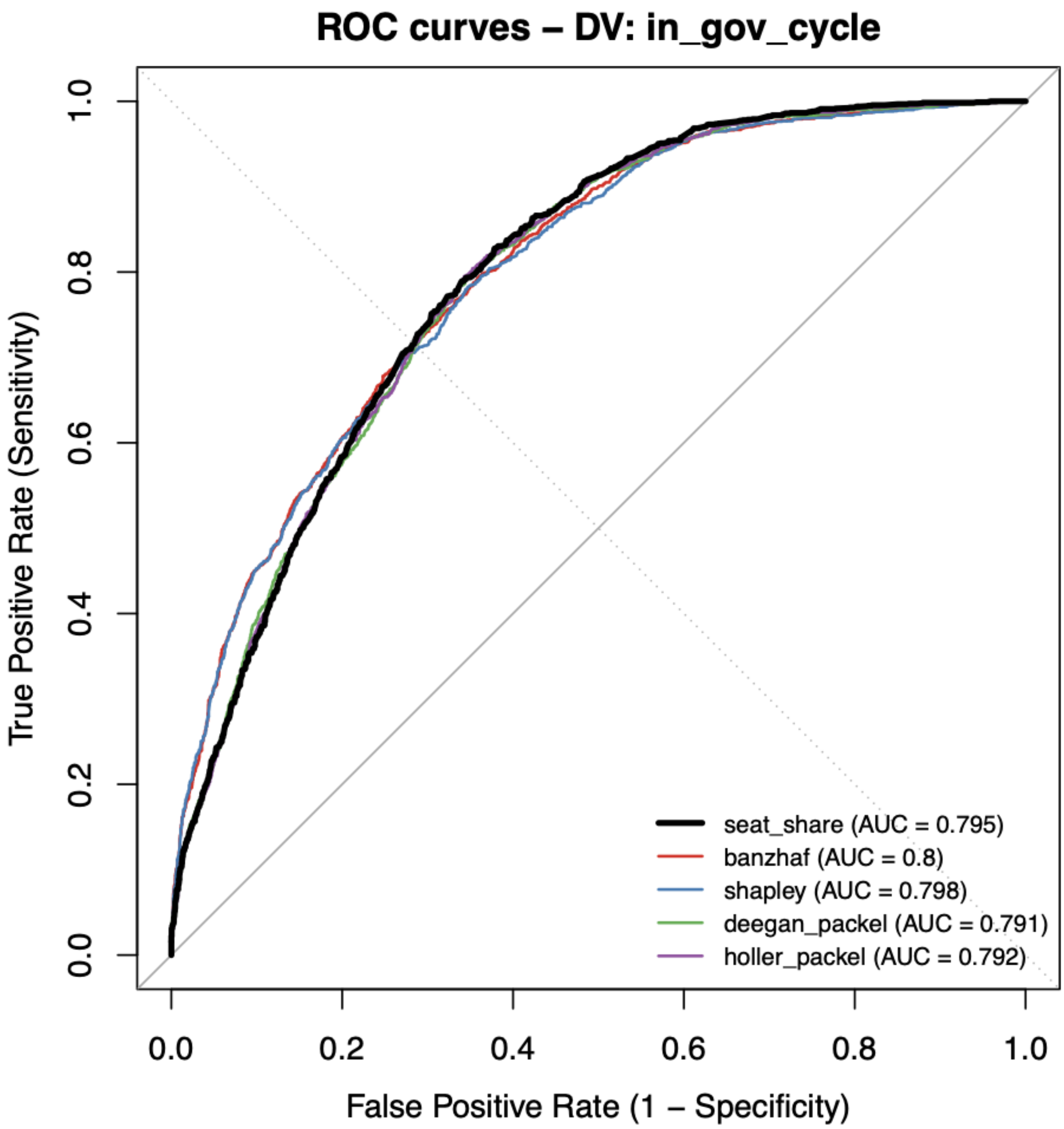

Figure 3. ROC Curves - DV = in_gov_cycle (any cabinet during electoral term)

Note: Thick black line = seat-share baseline. Solid blue = all-coalitions family (BPI, SSV). Dashed red = MWC family (DPI, HPI). AUCs shown in legend.

Source: Compiled by the author based on the ParlGov 2024 dataset.

For in_gov, the seat-share baseline achieves an AUC of 0.789, correctly ranking a randomly selected government entrant above a randomly selected non-entrant approximately 79% of the time, which is considerably better than chance but far from perfect. The shape of the baseline ROC curve is instructive: it rises steeply in the lower-left region of the plot, indicating that seat share is particularly effective at identifying the most likely government entrants at low false positive rates, but flattens progressively as the threshold is lowered and more parties are classified as entrants. This pattern reflects the underlying structure of the data: a small number of large parties with high seat shares are reliably identified as government entrants, while the bulk of the discrimination problem concerns the large middle range of parties with moderate seat shares and ambiguous government participation records.

The Banzhaf and Shapley-Shubik curves lie consistently above the seat-share baseline across virtually the full range of false positive rates, with AUCs of 0.797 and 0.794 respectively. The separation is modest in absolute terms but notable for its uniformity: the first-family indices do not simply outperform seat share at a single threshold or in a specific region of the ROC space but improve discrimination consistently across all thresholds. This uniformity is theoretically meaningful since it suggests that the additional information the all-coalitions indices carry over seat share is distributed across the full spectrum of parties rather than concentrated among a specific subset, such as only the smallest or only the largest parties. A party's pivot structure relative to its seat share affects its government entry probability at all levels of parliamentary representation, not only at the extremes.

The Deegan-Packel and Holler-Packel curves tell the opposite story. Both lie at or below the seat-share baseline across most of the ROC space, with AUCs of 0.787 and 0.788 - fractionally lower than the baseline. The direction of this result is more informative than its magnitude: the MWC-based indices do not merely fail to add discriminative value, they marginally reduce it. Adding them to a seat-share model moves the predicted probabilities of government entrants and non-entrants slightly closer together rather than further apart. This is consistent with the interpretation developed in Section 3.4: the variation these indices introduce relative to seat share reflects the structure of the minimal winning coalition game, which does not correspond to the political game that determines government entry. Their independent variation from seat share is therefore not signal but noise from the perspective of predicting which parties enter government.

Figure 3 replicates the same pattern for in_gov_cycle, with the seat-share baseline AUC rising marginally to 0.795. Banzhaf and Shapley-Shubik again lie above it at 0.800 and 0.798, while Deegan-Packel and Holler-Packel fall below at 0.791 and 0.792. The rank ordering of all five predictors is identical across both figures, and the direction of gain and loss holds without exception across both DVs. This replication across two different operationalisations of the outcome variable rules out the possibility that the pattern is an artefact of how government participation is defined and strengthens confidence that the ROC results reflect a genuine and consistent property of the data rather than a threshold-specific or outcome-specific finding. Taken together, the ROC curves confirm and sharpen the regression findings: the first-family indices provide a small but consistent and uniform discriminative advantage over seat share, while the second-family indices marginally reduce discriminative accuracy, and neither the gains nor the losses are confined to specific regions of the classification threshold space.

3.4. Discussion of the Results

The results reported in Sections 3.2 and 3.3 allow a direct answer to the research question: to what extent do formal voting-power indices improve empirical explanations of government formation compared to models based solely on party size? The answer is partial and asymmetric across the two index families, and the observed asymmetry is itself theoretically informative in ways that speak directly to the debates reviewed in Chapter 1.

Banzhaf and Shapley-Shubik consistently improve on seat share as predictors of government entry. The improvement is confirmed by all four evaluation criteria and holds across both dependent variables and all classification thresholds visible in the ROC curves. This is the core positive finding of the thesis: formal measures of strategic pivotality capture something about the government formation process that raw seat share misses, and this something is real enough to survive a complexity penalty, to replicate across two different operationalisations of the outcome, and to manifest as consistent improvement in rank-order discrimination rather than as a threshold-specific artefact. The theoretical interpretation connects directly to the argument developed in Section 1.3. Seat share measures how many seats a party holds; the first-family indices measure how often a party's participation is decisive for converting a losing coalition into a winning one. These two quantities are strongly but not perfectly correlated, and it is precisely in their deviations that the predictive gain resides. A party elevated by the pivot structure relative to its seat share- one whose participation is required for more winning coalitions than its numerical size alone would suggest- is more likely to enter government than seat share alone would predict. A party suppressed by the pivot structure is correspondingly less likely. The first-family indices pick up this variation; seat share cannot.

The confirmation of H1 therefore vindicates the core theoretical claim of the cooperative game theory tradition reviewed in Section 1.2: that bargaining power in coalition negotiations derives from structural necessity rather than numerical size, and that this distinction has observable consequences for which parties enter government. This is not a trivial finding. The dominant empirical practice in the coalition formation literature has been to use seat share as the primary measure of bargaining leverage, and the finding that a pivotality-based measure adds consistent predictive value over seat share, even in a deliberately sparse two-predictor model, demonstrates that the theoretical critique of seat share developed by Felsenthal and Machover (2005), Ansolabehere et al. (2005), and others has empirical purchase beyond the specific cases and portfolio allocation outcomes on which it has previously been evaluated.

The bounded nature of the gain matters equally, however, and deserves careful interpretation rather than being treated merely as a caveat. An R^2 improvement of 0.026 and an AUC gain of 0.008 are robust and consistent, but the joint model with Banzhaf still leaves more than 82% of the variation in in_gov unexplained. This residual is not a measurement failure: it accurately reflects the limits of what any structural measure computed from a seat distribution can explain about a process that is fundamentally political. Government formation takes place within the combinatorial space defined by the seat distribution, but it is not determined by that space. The ideological distances between potential partners impose constraints on which arithmetically feasible coalitions are politically viable (Axelrod, 1970; Laver & Schofield, 1998). The identity of the formateur and the procedural rules governing negotiation create asymmetric advantages that are invisible in a seat vector (Baron & Ferejohn, 1989; Laver & Shepsle, 1996). The strategic calculations of party leaders, the reputational costs of particular alliances, and the accumulated legacies of previous coalition agreements all shape the outcome of negotiations in ways that no index of pivotality can capture (Müller & Strøm, 2000; Martin & Stevenson, 2001). The modest predictive gains of the first-family indices should therefore be read not as evidence that pivotality is a weak determinant of coalition outcomes but as evidence that it is one determinant among several, operating alongside ideological and institutional factors that the present analysis deliberately holds in the residual. This framing has a direct implication for how the R^2 values reported in this chapter should be interpreted in a comparative context: the absolute level of fit is low because the models are intentionally sparse, not because seat share and pivotality exhaust the available explanatory information. Multivariate specifications that add left-right distance between potential coalition partners, a binary indicator for positive investiture requirements, and country fixed effects to absorb unobserved national-level heterogeneity consistently achieve substantially higher pseudo-R^2 values in the comparative coalition literature (Martin & Stevenson, 2001; Glasgow, Golder & Golder, 2012). The contribution of the present analysis is not to build the richest possible predictive model but to isolate the specific and previously untested contribution of pivotality over and above seat share under a controlled comparison. Incorporating these additional predictors alongside voting-power indices is the most important direction for future research, as discussed in the conclusion.

It is also worth reflecting on what the structure of the residual itself implies about the coalition formation process. The 82% of variation in government entry that the joint Banzhaf model leaves unexplained is not a homogeneous mass of noise: it is the accumulated effect of factors that operate conditionally and interactively rather than additively. Ideological distance, for instance, does not simply add a fixed increment to the probability of coalition entry independently of seat share and pivotality; it interacts with both, since a party that is pivotal but ideologically extreme may be excluded precisely because its structural indispensability makes it a more threatening rather than a more attractive coalition partner. Similarly, the formateur's identity introduces a directional asymmetry that a symmetric measure of pivotality cannot capture: the party tasked with initiating negotiations enjoys procedural advantages that translate into higher government entry probability independently of its pivot score, and these advantages are distributed unequally across parties in ways correlated with but not reducible to seat share (Baron & Ferejohn, 1989; Laver & Shepsle, 1996). The implication is that the residual is not simply waiting to be explained by adding more variables additively to the present specification. A fuller model would need to incorporate interaction terms between pivotality and ideological distance, and between pivotality and formateur status, to capture the conditional logic through which structural position translates into actual government entry. The modest but robust predictive gain of the first-family indices reported here should therefore be understood as a lower bound on the contribution of pivotality in a correctly specified model, not as its ceiling.

The failure of Deegan-Packel and Holler-Packel requires a more careful explanation because it is not what a naive reading of the theoretical argument would predict. These indices have lower correlations with seat share than the first-family indices: 0.779 and 0.708 respectively, compared to 0.891 and 0.909, meaning they carry substantially more information that is genuinely independent of seat share. If predictive gain over seat share were simply a function of structural independence from it, the MWC-based indices should outperform Banzhaf and Shapley-Shubik. Instead, their greater independence translates into worse predictive performance, both in the joint models and when used in isolation. The explanation lies not in the quantity of information these indices carry independently of seat share but in its quality: the variation they introduce relative to seat share reflects their restriction to minimal winning coalitions, and that restriction embeds an assumption about the coalition formation process that is systematically violated in the empirical record.

By computing power exclusively over minimal winning coalitions, Deegan-Packel and Holler-Packel operationalise a version of Riker's (1962) size principle as a theoretical premise rather than testing it as an empirical claim. The assumption is that rational office-seeking parties will resist forming coalitions larger than the minimum required, since surplus partners reduce each member's share of cabinet benefits without compensating gains in legislative viability. But as documented in Section 1.3, surplus majority coalitions are empirically common across the Western European parliamentary democracies that make up the bulk of the ParlGov sample (Laver & Schofield, 1998, pp. 75-78; Müller & Strøm, 2000). They arise because ideology, institutional incentives, and the desire for policy stability create reasons to include additional partners even at a cost in portfolio allocation terms. By excluding surplus coalitions from the computation of power, the MWC-based indices assign zero weight to a large portion of the coalition space that is empirically relevant. The variation they introduce relative to seat share therefore reflects the structure of a game that does not correspond to how governments form, and their independent information moves predictions in the wrong direction rather than the right one.

The confirmation of H2 therefore does more than establish a negative result about two indices. It empirically adjudicates a theoretical debate that Chapter 1 identified as unresolvable on axiomatic grounds alone (Felsenthal & Machover, 1998, pp. 247-249). The choice between the all-coalitions and MWC-only approaches has previously been treated as a matter of competing prior assumptions about rational behaviour, with the game-theoretic literature divided between those who follow Riker in treating the minimal winning coalition as the equilibrium outcome and those who argue that the full coalition space is the appropriate domain for measuring power. The present analysis provides systematic empirical evidence that the all-coalitions assumption is more consistent with observed government formation patterns across a broad cross-national sample. For researchers making practical decisions about how to operationalise bargaining power in coalition research, this is a meaningful result: Banzhaf or Shapley-Shubik is the appropriate choice; Deegan-Packel and Holler-Packel are not, at least for the purpose of predicting government entry.

A final observation concerns what the results collectively imply about the relationship between the European Politics and game-theoretic traditions reviewed in Table 1. The European Politics tradition has emphasised contextual, ideological, and institutional determinants of coalition outcomes while relying on seat share as its operational measure of bargaining leverage. The game-theoretic tradition has formalised bargaining power as pivotality while producing indices whose empirical performance had not been systematically evaluated against seat share across a large comparative sample. The present findings suggest that the two traditions are more complementary than competing: the first-family indices provide a structurally superior measure of the numerical dimension of bargaining leverage that the European Politics tradition has always acknowledged as important, while the large unexplained residual in the joint models confirms that the contextual and ideological factors the European Politics tradition emphasizes are not merely nuisances to be controlled for but central determinants of coalition outcomes in their own right. The implication is integrative rather than adjudicative: formal voting-power indices of the all-coalitions family should be incorporated alongside ideological distance measures and institutional indicators rather than replacing them. The game-theoretic and European Politics traditions share a common object of explanation but have approached it through separate methodological vocabularies; the present analysis provides a bridge by demonstrating that the formal pivotality concept, when operationalized through the appropriate family of indices, can be made to speak directly to the comparative coalition formation literature's core empirical questions. At the same time, the analysis reinforces a point that the European Politics tradition has long maintained: the seat distribution is the starting point for understanding coalition bargaining, not its endpoint. Strategic necessity and numerical size are jointly relevant, and neither tradition alone captures the full picture. By providing a systematic cross-national evaluation that situates formal indices within the same predictive framework as seat share, this thesis offers a methodological basis for future research that draws on both traditions simultaneously, using first-family voting-power indices as the structural component of a richer multivariate model that also incorporates the ideological and institutional variables that the European Politics tradition has developed with considerable empirical sophistication.

3.5. Limitations and Scope of Conclusions

The findings reported in this chapter are subject to a set of limitations that condition the scope of the conclusions that can be drawn. They do not undermine the core results, but they define the boundaries within which those results should be interpreted.

The most consequential limitation follows directly from the analytical design. The models include only seat share and a single power index, omitting the full range of factors known to influence government formation. Ideological proximity is the most important absent predictor: research consistently documents that ideological distance is a strong independent predictor of coalition entry over and above size considerations (Axelrod, 1970; Laver & Schofield, 1998; Glasgow, Golder & Golder, 2012). The formateur's identity and the institutional rules governing negotiation are equally important omitted factors: under different investiture rules and formateur selection procedures, the same seat distribution can produce very different coalition outcomes (Baron & Ferejohn, 1989; Laver & Shepsle, 1996). The unexplained residual of more than 80% in the joint models is substantially attributable to these and other omitted factors. The R^2 values and AUC scores reported in this chapter should not be read as an assessment of how well government formation can be explained in principle, but as a controlled comparison of two structural measures under a deliberately sparse specification. What this design identifies is their relative contribution in isolation, precisely the quantity relevant to the research question, and any richer model incorporating ideology and institutional variables alongside voting-power indices would be expected to explain substantially more of the variation in government entry.

The logistic regression models treat all 4,939 observations as statistically independent, an assumption violated in two directions. First, parties within the same election are jointly determined by a single post-electoral negotiation: their seat shares and index values are mechanically linked through the seat distribution, meaning the predictors are not independent across observations within the same election. Second, the same party appears across multiple elections over time, introducing serial dependence in both the outcome and the predictors. The practical consequence is that standard errors are likely underestimated throughout, and the likelihood ratio test p-values reported in Appendix B overstate precision to some degree. This concern is most acute for the second-family indices, where effect sizes are small and the distinction between significance and non-significance is consequential. The substantive conclusions, however, rest primarily on the pattern of R^2 gains, BIC differences, and AUC results rather than on p-values alone, and the first-family gains clear all four criteria by wide margins that are unlikely to be reversed by corrected standard errors. Addressing the independence assumption through clustered standard errors or multilevel specifications is an important refinement for future work but does not alter the direction or ranking of the findings reported here.

The computational thresholds described in Section 2.2 exclude the most fragmented elections from the index calculations, those with more than 22 parties holding seats for BPI and SSV, and more than 20 for DPI and HPI. These are precisely the configurations where pivotality and seat share diverge most sharply, since highly fragmented legislatures generate the largest number of configurations in which small parties are decisive and large ones are structurally redundant (Felsenthal & Machover, 1998, pp. 420-422; Ansolabehere et al., 2005, pp. 550-552). The reported gains for the first-family indices are therefore likely conservative lower bounds on the true advantage of pivotality-based measures in the full population of parliamentary elections: in the excluded highly fragmented elections, the structural difference between seat share and pivotality would be even more pronounced, and the first-family indices would be expected to add more predictive value than is visible in the retained sample. The exclusion is not random with respect to fragmentation, and this limits the generalisability of the findings to the range of party-system fragmentation observed in the retained sample.

The uniform simple majority threshold applied across all 37 countries does not capture the institutional diversity of investiture rules. Systems operating under negative parliamentarism face an effective winning threshold lower than 50%+1, meaning that minority governments are structurally viable without requiring a majority of seats (Strøm, 1990). Applying the simple majority threshold to these systems misclassifies some coalitions as non-winning when they are in practice viable governing arrangements, and vice versa. This approximation error is most consequential for the MWC-based indices, which are highly sensitive to the exact composition of the minimal winning set: a coalition that is classified as minimal winning under the simple majority threshold may not be minimal winning under the system-specific effective threshold, and the power scores assigned to parties are correspondingly distorted (Felsenthal & Machover, 1998, pp. 247-249). Correcting for this limitation would require system-specific threshold data that are not consistently available across the full scope of the sample, but doing so in a subset of systems for which such data exist would provide a useful robustness check on the second-family results.

Finally, the 37 ParlGov countries are predominantly established Western European parliamentary democracies with continuous democratic records since 1945. The findings are directly generalisable to this population. Whether the same pattern would hold in parliamentary systems outside it, including newer democracies in Central and Eastern Europe, where coalition norms and party system institutionalisation differ substantially from the Western European baseline, or parliamentary systems in other regions entirely, is an open question that extension to broader data sources would be required to answer. The replication of the core pattern across 37 countries and 626 elections spanning nearly eight decades provides a strong basis for confidence within this population, but the boundaries of generalisation should be respected. Taken together, these limitations suggest that the true advantage of first-family indices is likely somewhat larger than reported, that the conclusions about second-family indices are robust and if anything understated, and that the findings are bounded to the established parliamentary democracies covered by ParlGov.

A further limitation concerns the binary operationalisation of the dependent variable. Coding government participation as a dichotomous indicator, either a party enters cabinet or it does not, discards information about the degree of a party's involvement in governing coalitions. A small party holding one minor portfolio is treated identically to the largest partner controlling the most senior ministries, even though their structural positions, bargaining contributions, and policy influence differ substantially. This conflation is a consequence of the research design's focus on entry as the primary outcome, which follows the dominant convention in the comparative coalition literature (Martin & Stevenson, 2001; Glasgow, Golder & Golder, 2012), but it may understate the predictive contribution of voting-power indices in configurations where a pivotal party enters government with disproportionately large portfolio allocations relative to its seat share. If pivotality shapes not only the probability of entry but also the depth of a party's integration into the governing coalition, as Ansolabehere et al. (2005) and Warwick and Druckman (2006) suggest for portfolio allocation, then the binary coding of the outcome variable captures only the coarsest dimension of that relationship. An extension using the proportion of cabinet portfolios received as a continuous outcome measure, or disaggregating by ministerial salience, would allow a more fine-grained assessment of whether formal voting-power indices explain not just which parties enter government but how extensively they do so. Such an extension would also allow a direct test of whether the all-coalitions indices outperform seat share in both the entry and the allocation dimensions simultaneously, which is a question the present analysis is not designed to answer but which follows naturally from the findings reported here.

CONCLUSION

The starting point of this thesis was a puzzle that the coalition formation literature has long acknowledged but insufficiently resolved: the relationship between a party's share of parliamentary seats and its probability of entering government is systematically non-proportional. Large parties are sometimes excluded from office entirely; small parties are sometimes indispensable to any viable governing majority. The dominant empirical response to this puzzle has been to treat seat share as the primary measure of bargaining power in coalition negotiations - a choice that is practically convenient but theoretically problematic, since seat share measures numerical size rather than strategic necessity. The objective of this thesis was to evaluate whether cooperative game-theoretic voting-power indices, which operationalise bargaining power as pivotality rather than size, improve empirical explanations of government formation relative to models based on seat share alone.

The theoretical framework developed in Chapter 1 established why this question matters. Seat share and pivotality are related but distinct concepts: a party's strategic leverage in coalition bargaining derives not from how many seats it holds but from how often its participation is decisive for converting a losing coalition into a winning one. Four established indices operationalise this concept in different ways. The Banzhaf Power Index and the Shapley-Shubik Value compute pivotality across all winning coalitions, including surplus coalitions larger than the minimum required. The Deegan-Packel Index and the Holler-Packel Index restrict attention exclusively to minimal winning coalitions, embedding the assumption that rational office-seeking parties will avoid surplus partners. This distinction between the two families carries direct empirical implications: if surplus coalitions are empirically common, which the historical record of Western European parliamentary democracies confirms they are, the all-coalitions family should better reflect the strategic environment in which actual coalition bargaining takes place. Chapter 1 identified this as a theoretical debate that could not be resolved on axiomatic grounds alone and required systematic empirical evidence to adjudicate.

The empirical analysis in Chapter 2 addressed this through a quantitative comparative research design. All four indices were computed for every parliamentary election in the analytical sample and evaluated against a seat-share baseline through nested logistic regression models estimated on 4,939 party-election observations drawn from 626 elections across 37 parliamentary democracies from 1945 to 2023. The data source was the ParlGov database, which provided systematic information on seat distributions and cabinet compositions across the full sample. Two outcome variables were used, cabinet entry in the first post-election government and participation in any cabinet during the electoral term, and predictive performance was assessed through four complementary criteria: McFadden pseudo-R^2, the likelihood ratio test, the Bayesian Information Criterion, and the area under the ROC curve.

The central research question asked to what extent formal voting-power indices improve the empirical prediction of government entry compared to models based solely on parliamentary seat share. The answer is clear but asymmetric. The all-coalitions family (Banzhaf and Shapley-Shubik) consistently and robustly improves on seat share across both outcome variables and all four evaluation criteria. The improvement is modest in absolute terms: the best-performing index raises the area under the ROC curve by approximately 0.008 and reduces the Bayesian Information Criterion by nearly 145 points relative to the seat-share baseline. But it is genuine since it survives a complexity penalty, replicates across both dependent variables, and manifests as consistent improvement in discriminative accuracy across the full range of classification thresholds. The MWC-restricted family (Deegan-Packel and Holler-Packel) does not improve on seat share. These indices either leave predictive performance unchanged or marginally reduce discriminative accuracy when added to a seat-share model, and they perform worse than seat share even when used in isolation. The two hypotheses derived from the theoretical framework are therefore both confirmed: H1, that the all-coalitions family would outperform seat share, is supported; H2, that the MWC-restricted family would not, is equally supported.

The asymmetry between the two families is itself the most theoretically informative finding of the thesis. The failure of Deegan-Packel and Holler-Packel is not a measurement artefact since these indices carry more information structurally independent of seat share than the all-coalitions indices do, yet this independence translates into worse rather than better predictive performance. The explanation is substantive: by restricting attention to minimal winning coalitions, these indices mischaracterise the coalition space that actual negotiators operate within. Surplus coalitions are too common in the empirical record to be treated as theoretically irrelevant, and an index that ignores them captures variation in a formal game-theoretic structure that does not correspond to how governments form. The all-coalitions indices, by contrast, remain closer to empirical reality and produce predictions accordingly.

The research puzzle is therefore partially but genuinely resolved. Pivotality-based measurement does improve on seat share, and the improvement is not trivial, but it is consistent, replicable, and theoretically interpretable. The qualification is equally important: the joint model with the best-performing index still leaves more than 80% of the variation in government entry unexplained. This residual reflects the collective influence of ideological compatibility, formateur identity, institutional constraints, and contingent bargaining dynamics, factors that no index computed from a seat distribution can capture. Voting-power indices of the all-coalitions family are a better structural predictor of government entry than seat share, but they are one input into a richer explanatory framework rather than a complete account of coalition formation. As for limitations, the findings are directly generalisable to the established parliamentary democracies covered by ParlGov, and caution is warranted in extending them to newer or institutionally distinct systems outside this population.

Three directions for future research follow from these findings. The most important is the incorporation of ideological distance measures alongside the all-coalitions indices in a joint model, which would allow direct assessment of how structural pivotality, and ideological compatibility interact in determining which parties enter government. Second, the analysis should be extended to examine whether the predictive advantage of the all-coalitions indices varies with the degree of legislative fragmentation, since the divergence between seat share and pivotality is theoretically largest in highly fragmented party systems. Third, a unified analysis modelling both coalition entry and portfolio allocation using the same structural measures would provide a more complete picture of how the formal distribution of parliamentary influence shapes the full range of coalition outcomes.

The broader implication of this thesis is methodological. The gap between the theoretical case for pivotality-based measurement and the empirical practice of relying on seat share is not merely a technical inconvenience, as it reflects a genuine misalignment between how bargaining power is conceptualised and how it is operationalised in much of the coalition formation literature. This thesis demonstrates that closing that gap, at least partially and for the right family of indices, produces measurable improvements in the prediction of government entry. For researchers working in this tradition, Banzhaf or Shapley-Shubik is the appropriate operationalisation of structural bargaining leverage; seat share, while useful and interpretable, is a less precise substitute. That is a modest but empirically grounded contribution to the ongoing effort to bring measurement practice into closer alignment with theoretical foundations in the study of coalition government.

REFERENCE LIST

Ansolabehere, S., Snyder, J. M., Strauss, A. B., & Ting, M. M. (2005). Voting weights and formateur advantages in the formation of coalition governments. American Journal of Political Science, 49(3), 550–563.

Axelrod, R. (1970). Conflict of interest: A theory of divergent goals with applications to politics. Markham.

Bäck, H., & Dumont, P. (2007). Combining large-n and small-n strategies: The way forward in coalition research. West European Politics, 30(3), 467–501.

Bäck, H., Meier, H. E., & Persson, T. (2009). Party size and portfolio payoffs: The proportional allocation of ministerial posts in coalition governments. The Journal of Legislative Studies, 15(1), 10–34.

Banzhaf, J. F. (1965). Weighted voting doesn't work: A mathematical analysis. Rutgers Law Review, 19(2), 317–343.

Baron, D. P., & Ferejohn, J. A. (1989). Bargaining in legislatures. The American Political Science Review, 83(4), 1181–1206.

Boix, C., Miller, M., & Rosato, S. (2013). A complete data set of political regimes, 1800–2007. Comparative Political Studies, 46(12), 1523–1554.

Budge, I., & Laver, M. (1986). Office seeking and policy pursuit in coalition theory. Legislative Studies Quarterly, 11, 485–506.

Cox, G. W. (1997). Making votes count: Strategic coordination in the world's electoral systems. Cambridge University Press.

De Winter, L., & Dumont, P. (2006). Parties into government: Still many puzzles. In R. Katz & W. Crotty (Eds.), Handbook of party politics (pp. 175–188). SAGE Publications.

Deegan, J., & Packel, E. W. (1978). A new index of power for simple n-person games. International Journal of Game Theory, 7(2), 113–123.

Deemen, A., & Hosli, M. O. (2002). Effects of enlargement on efficiency and coalition formation in the Council of the European Union.

Deschouwer, K. (2012). *The politics of Belgium: Governing a divided society* (2nd ed.). Palgrave Macmillan.

Diermeier, D., & Stevenson, R. T. (1999). Cabinet survival and competing risks. American Journal of Political Science, 43(4), 1051–1068.

Döring, H., & Manow, P. (2024). ParlGov database. https://www.parlgov.org/

Downs, A. (1957). An economic theory of democracy. Harper.

Dubey, P., & Shapley, L. S. (1979). Mathematical properties of the Banzhaf Power Index. Mathematics of Operations Research, 4(2), 99–131.

Dumont, P., Grofman, B., Bergman, T., & Louwerse, T. (2024). Introduction: Old puzzles, conceptual vagueness and new developments in the study of coalition governments.

Felsenthal, D. S., & Machover, M. (1997). The weighted voting rule in the EU Council of Ministers, 1958–1995: Intentions and outcomes. Electoral Studies, 16(1), 33–47.

Felsenthal, D. S., & Machover, M. (1998). The measurement of voting power: Theory and practice, problems and paradoxes. Edward Elgar.

Felsenthal, D. S., & Machover, M. (2001). The Treaty of Nice and qualified majority voting. Social Choice and Welfare, 18(3), 431–464.

Felsenthal, D. S., & Machover, M. (2005). Voting power measurement: A story of misreinvention. Social Choice and Welfare, 25, 485–506.

Gallagher, M., Laver, M., & Mair, P. (2011). Representative government in modern Europe. McGraw-Hill.

Gamson, W. A. (1961). An experimental test of a theory of coalition formation. American Sociological Review, 26, 565–573.

Glasgow, G., Golder, M., & Golder, S. N. (2012). New empirical strategies for the study of parliamentary government formation. Political Analysis, 20(2), 248–270.

Golder, S. N. (2006). The logic of pre-electoral coalition formation. Ohio State University Press.

Holler, M. J. (1982). Forming coalitions and measuring voting power. Political Studies, 30(2), 262–271.

Holler, M. J., & Packel, E. W. (1983). Power, luck and the right index. Zeitschrift für Nationalökonomie / Journal of Economics, 43(1), 21–29.

Jackman, S. (2020). *pscl: Classes and Methods for R Developed in the Political Science Computational Laboratory* (Version 1.5.5) [R package]. https://github.com/atahk/pscl

Koki, C., & Leonardos, S. (2019). Coalitions and voting power in the Greek Parliament of 2012: A case-study. Homo Oeconomicus: Journal of Behavioral and Institutional Economics, 35(4), 295–313.

Laakso, M., & Taagepera, R. (1979). "Effective" number of parties: A measure with application to West Europe. Comparative Political Studies, 12, 3–27.

Laruelle, A., & Valenciano, F. (2001). Shapley-Shubik and Banzhaf indices revisited. Mathematics of Operations Research, 26(1), 89–104.

Laver, M., & Schofield, N. (1998). Multiparty government: The politics of coalition in Europe. University of Michigan Press.

Laver, M., & Shepsle, K. A. (Eds.). (1996). Making and breaking governments: Cabinets and legislatures in parliamentary democracies. Cambridge University Press.

Leech, D. (2002). Voting power in the governance of the International Monetary Fund. Annals of Operations Research, 109(1), 375-397.

Lijphart, A. (2012). Patterns of democracy (2nd ed.). Yale University Press.

Mair, P. (1987). The changing Irish party system: Organisation, ideology and electoral competition. Francis Pinter.

Martin, L. W., & Stevenson, R. T. (2001). Government formation in parliamentary democracies. American Journal of Political Science, 45(1), 33–50.

Martin, L. W., & Vanberg, G. (2020). Coalition government, legislative institutions, and public policy in parliamentary democracies. American Journal of Political Science, 64(2), 325–340.

McFadden, D. (1974) Conditional Logit Analysis of Qualitative Choice Behavior. Frontiers in Econometrics, 105-142.

Müller, W. C., & Strøm, K. (Eds.). (2000). Coalition governments in Western Europe. Oxford University Press.

Owen, G. (1995). Game theory (3rd ed.). Academic Press.

Persson, T., & Tabellini, G. (2003). The economic effects of constitutions. MIT Press.

Raftery, A. E. (1995). Bayesian model selection in social research. Sociological Methodology, 25, 111–163.

Riker, W. H. (1962). The theory of political coalitions. Yale University Press.

Roth, A. E. (Ed.). (1988). The Shapley value: Essays in honor of Lloyd S. Shapley. Cambridge University Press.

Shapley, L. S., & Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. American Political Science Review, 48(3), 787–792.

Straffin, P. D. (1994). Power and stability in politics. In R. J. Aumann & S. Hart (Eds.), Handbook of game theory with economic applications (Vol. 2, pp. 1127–1151). North-Holland.

Strøm, K. (1990). Minority government and majority rule. Cambridge University Press.

von Neumann, J., & Morgenstern, O. (1944). Theory of games and economic behavior. Princeton University Press.

Warwick, P. V. (1996). Coalition government membership in West European parliamentary democracies. British Journal of Political Science, 26(4), 471–499.

Warwick, P. V., & Druckman, J. N. (2006). The portfolio allocation paradox: An investigation into the nature of a very strong but puzzling relationship. European Journal of Political Research, 45(4), 635–665.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer. https://ggplot2.tidyverse.org

Wickham, H., et al. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Widgrén, M. (1994). Voting power in the EC decision making and the consequences of two different enlargements. European Economic Review, 38(5), 1153–1170.

APPENDICIES

Appendix A. Countries, Elections, and Party-Election Observations in the Analytical Sample

Country	Elections	Observations
Australia	27	138
Austria	23	105
Belgium	32	276
Bulgaria	11	65
Canada	13	67
Croatia	7	95
Cyprus	9	52
Czech Republic	9	57
Denmark	44	397
Estonia	9	56
Finland	30	254
France	25	188
Germany	29	240
Greece	7	44

Country	Elections	Observations
Hungary	4	25
Iceland	31	156
Ireland	23	161
Israel	25	301
Italy	19	237
Japan	12	95
Latvia	10	66
Lithuania	8	90
Luxembourg	21	125
Malta	3	13
Netherlands	29	305
New Zealand	12	74
Norway	25	177
Poland	8	68
Portugal	12	77
Romania	8	66

Country	Elections	Observations
Slovakia	9	63
Slovenia	10	94
Spain	10	125
Sweden	32	188
Switzerland	27	308
Turkey	6	32
United Kingdom	7	59
Total	**626**	**4,939**

Appendix B. Regression Model Comparison Results: Baseline vs. Joint Models

Index	R^2 baseline	R^2 joint	R^2 gain	LR stat	p-value	BIC baseline	BIC joint	ΔBIC	AUC gain
DV = in_gov (first post-election cabinet entry)									
Banzhaf	0.151	0.177	0.026	153.3	< 0.001	5,118.4	4,973.6	-144.8	+0.008
Shapley-Shubik	0.151	0.176	0.025	147.9	< 0.001	5,118.4	4,979.1	-139.4	+0.006
Deegan-Packel	0.151	0.153	0.002	12.9	< 0.001	5,118.4	5,114.0	+4.4	-0.002
Holler-Packel	0.151	0.152	0.001	2.5	0.113	5,118.4	5,124.4	+6.0	-0.001
DV = in_gov_cycle (any cabinet during electoral term)									
Banzhaf	0.158	0.186	0.028	181.9	< 0.001	5,404.4	5,231.0	-173.4	+0.006
Shapley-Shubik	0.158	0.184	0.027	170.6	< 0.001	5,404.4	5,242.3	-162.1	+0.003
Deegan-Packel	0.158	0.162	0.004	25.5	< 0.001	5,404.4	5,387.4	-17.0	-0.003
Holler-Packel	0.158	0.159	0.001	9.2	0.002	5,404.4	5,403.6	+0.7	-0.002

Non-exclusive licence to reproduce thesis and make thesis public

I, Sofia Samoilova

1. grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the digital archives of the University of Tartu until the expiry of the term of copyright, my thesis Beyond Seat Share: Evaluating Voting-Power Indices in Government Formation, supervised by Martin Mölder;
2. grant the University of Tartu a permit to make the thesis specified in point 1 available to the public via the web environment of the University of Tartu, including via the digital archives, under the Creative Commons licence CC BY NC ND 4.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright;
3. am aware of the fact that the author retains the rights specified in points 1 and 2;
4. confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

author's name

Sofia Samoilova
18.05.2026