On Selection of Formateurs during Coalition Formation in Multiparty Parliamentary Democracies

Ehud Menirav*

2 December 2004

Abstract

Models of government formation in parliamentary democracies typically employ rules that govern selection of a party that proposes to form a government (*formateur* selection rules). The choice of a particular rule fundamentally affects the predictions and outcomes of political economy models. Yet, recent evidence questions the empirical consistency of these selection rules. This paper proposes a new selection rule that demonstrates reasonable axiomatic properties and is empirical consistent. The new rule also generalizes the most commonly used selection rules. Finally, evidence contradicting results found in the literature is presented; this evidence reveals empirical inconsistencies inherent in existing selection rules.

JEL classification: D72, H19, C14
Keywords: Formateur selection rule, coalition formation

* The Eitan Berglas School of Economics, Tel-Aviv University, P.O.B. 39040, Tel-Aviv 69978, Israel. Tel.: +972-3-640-7983; Fax: +972-3-640-9908. E-mail address: menirav@post.tau.ac.il.
The paper is based in part on a chapter of my Ph.D. dissertation written under the supervision of Prof. Elhanan Helpman, Tel-Aviv University. I wish to thank Elhanan Helpman for his guidance, Antonio Merlo for sharing his data with me in addition to Tomer Blumkin, Isaac Meilijson, Tymofiy Mylovanov, Santiago Sanchez-Pages and Burkhard Schipper, seminar participants at Bonn University, Tel-Aviv University and the ASSET annual meeting, 2004, for helpful comments and suggestions.

1. Introduction

Selection of a *formateur* (i.e., the party that proposes a government for parliamentary approval) is a crucial issue in non-cooperative game theoretic modeling of government formation in multiparty parliamentary democracies.[1] Typically, governments are not directly chosen by the electorate but appointed after bargaining between the parties sitting in parliament. Usually, several different coalitions may be constructed, with each leaning toward a different policy. Therefore, the selection of a formateur fundamentally affects the set of equilibrium outcomes (see Baron, 1991).

The most common approach taken in the literature is to assume that the head of state (e.g., monarch) mandates a formateur according to a given selection rule. So far, the literature has focused on two formateur selection rules: the deterministic selection rule (hereinafter DSR) and the probabilistic selection rule (hereinafter PSR). According to the DSR, suggested by Austen-Smith and Banks (1988), the formateur is the largest party; if it fails to form a government, the task is assigned to the second-largest party and so on until a government is formed. According to the PSR, originally proposed by Baron and Ferejohn (1989), each party represented in parliament may be selected as formateur with a probability proportional to the respective party's seat share (weight).[2] If a formateur fails to form a government, a new formateur (possibly the same party) is selected based on the same probabilities, a process that continues until a government is formed.

Obviously, these selection rules ignore numerous other factors likely to influence formateur selection. However, in order to simplify and construct analytically tractable models, the literature employs selection rules based on party weight exclusively. These selection rules do not mean that the decision-maker tosses a die exhibiting the relevant probabilities and then constructs a government according to the toss results. Instead, the position taken in the literature (and in this paper as well) is that at the conclusion of a complex decision-making process, we can describe selection of a formateur as though it were the outcome of the application of prescribed selection rules based on the distributions of party weights.

Diermeier and Merlo (2004) have questioned the empirical consistency of these selection rules and shown that a weighted average of the two selection rules is more in agreement with the data. Moreover, I show that these rules as well as any weighted average have some undesirable properties; most salient is the rules' completely insensitivity to the distribution of weights among the parties represented in parliament. To illustrate, consider Example 1, where the weights of three parties are reported in three hypothetical situations.

Example 1:

Party	A	B	C
Weight in situation 1	0.34	0.335	0.325
Weight in situation 2	0.49	0.335	0.175
Weight in situation 3	0.49	0.48	0.03

Consider first situations 1 and 2. It is reasonable to expect that in situation 1, because the parties' weights are almost identical, any of the parties might become formateur; in situation 2, it is much more likely that party A will be assigned that responsibility whereas party B's prospects appear to decline. However, both the DSR and the PSR indicate that party B's prospect to become formateur is identical in both situations because it retains both its weight and second-largest position. In another illustration, consider situations 2 and 3 where Party A has the same weight in both. Party A's probability of becoming formateur nevertheless appears to be much higher in situation 2 than in situation 3, where it has a close contender. Again, as both the DSR and the PSR are distribution-insensitive, they fail to response to such situations and indicate that party A's prospect to become formateur is unaltered.

Accordingly, it seems that the search for a new formateur selection rule is a timely task from both empirical and theoretical perspectives. This paper takes a step in this direction as it proposes a distributional sensitive rule that has reasonable properties and is in agreement with the data. The functional form of the suggested rule has strong roots in economics but in a different context, having been frequently employed in the rent-seeking literature as a *contest success function*. Moreover, the suggested rule generalizes the most commonly used selection rules. In addition, I develop precise empirical tests to assess the consistency of selection rules; application of these tests results in findings contradicting those that have appeared in the literature to date.

The paper is organized as follows. Section 2 describes the new selection rule and explores its properties. Section 3 describes the data while Section 4 tests the empirical consistency of the selection rules; Section 5 concludes and compares the findings with the literature.

2. A Generalized Formateur Selection Rule

Let $N = \{1, 2, \ldots, n\}$ be the set of parties represented in parliament and let w_i denote the weight of party $i \in N$ (i.e., party i's number of seats divided by the total number of seats of the parties in N). Without loss of generality, assume that parties are arranged in descending order of their weights: $w_1 \geq w_2 \geq \ldots \geq w_n$. Let $w = (w_1, w_2, \ldots, w_n)$ denote the vector of parties' weights. The generalized formateur selection rule (hereinafter GSR) is given by

$$\pi_i(w) = \frac{w_i^p}{\sum_{j \in N} w_j^p} \qquad \text{for } p > 0, \tag{1}$$

where $\pi_i(w)$ is party i's probability of becoming formateur, and p is a parameter that measures the GSR's sensitivity to the distribution of party weights in parliament.

The functional form in (1) has been used extensively as a *contest success function* in the rent-seeking literature (see, e.g., Tullock, 1980) where w_i is interpreted as the "effort" or "strategic endowment" of player i. Skaperdas (1996) has shown that the functional form (1) is the only continuous functional form satisfying the following six axioms, which I denote as *Probability Distribution Function*, *Anonymity*, *Homogeneity*, *Weight Incentive*, *Subgroup Consistency* and

Independence from Non-Subgroup Alternatives. In the following, I briefly present the six axioms and some of the properties they induce; Afterwards, I illustrate the role of the parameter p and show that the GSR generalizes both the PSR and the DSR.

The first axiomatic property is that the probability of becoming formateur should satisfy the requirements of a probability distribution function: $\sum_{i\in N} \pi_i(w) = 1$, $\pi_i(w) \geq 0$ for all $i \in N$ and all w (with some abuse of notation, I continue using $\pi_i(w)$ for party i's probability of becoming formateur both for the general as well as the GSR case). The second axiom (*Anonymity*) states that each party's probability of becoming formateur does not depend on its identity or on the identities of the other parties but only on the parties' weights. The Anonymity property also implies that if m parties have identical weights in a given a distribution of parties weights, their probabilities of becoming formateur are all equal. *Homogeneity* of degree zero is the third axiom. Formally, $\pi_i(w) = \pi_i(\lambda w)$, where $\lambda w = (\lambda w_1, \lambda w_2, \ldots, \lambda w_n)$. The implication of this axiom is that the ratio of the probabilities of any two parties to become formateur depends on the ratio of their weights. It also implies that we can employ either number of seats or weight as the party's endowment.

According to the fourth axiom (*Weight Incentive*), each party's probability of becoming formateur is increasing in its own weight and decreasing in every other party's weight. This property has three interesting implications. First, combined with the Anonymity, Weight Incentive implies monotonicity in weight; namely, given a distribution of weights, $\pi_i(w) < \pi_j(w)$ if and only if $w_i < w_j$. Second, it implies a *Transfer Property* – other things being equal, the transfer of weights from party j to party i will increase party i's probability and decrease party j's probability of becoming formateur.[3] Third, it opens the possibility for the selection rule to be distribution-sensitive when the difference in probabilities due to changes in weights depends on the parties' weights. For the GSR case, it implies that the GSR is distribution-sensitive when $p \neq 1$. To verify this, consider the effect of a negligible transfer of weight from party j to party l on the probability that party i (which is not included in this transfer, $i \neq j \neq l$) will be chosen as formateur, given by

$$\frac{\partial \pi_i(w)}{\partial w_l} - \frac{\partial \pi_i(w)}{\partial w_j} = \frac{p \cdot \pi_i(w)}{\sum_{h \in N} w_h^p} \cdot (w_j^{p-1} - w_l^{p-1}) \,. \tag{2}$$

The right-hand side of (2) does not equal zero whenever $p \neq 1$ and $w_l \neq w_j$, implying the distribution-sensitivity of the GSR.

The distribution-sensitivity of the GSR takes a special form that may be referred to as the *Transfer-Sensitivity Property*. To illustrate this property, I define a progressive [regressive] transfer as a transfer of weight from a party with more [less] weight to a party with less [more] weight, without changing the order between the two parties' weights (i.e., after a progressive transfer the recipient party's weight is still lower than the donor party's weight). According to (2), if $p > 1$, then any progressive [regressive] transfer will increase [decrease] the probability of every party that does not participate in the transfer to become formateur. When $0 < p < 1$, the effect runs in the opposite direction. Following Example 1, I argue that a progressive [regressive] transfer should increase [decrease] π_i for every party i that does not participate in the transfer (i.e., I argue that we should expect that $p > 1$). To illustrate, note that situation 2 [1] is obtained from situation 1 [2] (see Example 1) by a regressive [progressive] transfer from party C [A] to party A [C]; thus, party B's probability of becoming formateur decreases [increases].

The fifth axiom is *Subgroup Consistency*. Suppose that for some reason we know that the formateur has to come from a nonempty subset of parties $M \subseteq N$. Subgroup Consistency demands that the selection among the subset of parties should be qualitatively similar to that of the selection among all parties. Formally, $\pi_i^M(w) = \pi_i(w)/[\sum_{j \in M} \pi_j(w)]$ $\forall\ i \in M$ and $\forall\ M \subseteq N$, where $\pi_i^M(w)$ denotes the probability of party $i \in M$ to be selected formateur among the subset M. The sixth axiom (*Independence from Non-Subgroup Alternatives*) demands that $\pi_i^M(w)$ should not depend on the weights of parties that are not included in subset M.

Consider now the role of the parameter p of the GSR. When $p = 1$, the GSR indicates that each party may become formateur with a probability that is equal to the respective party's weight;

thus, it is equivalent to the PSR. When p reaches infinity, the GSR demands that the formateur be the largest party.[4] In the following I focus solely on either the *first* formateur (i.e., the formateur that first attempts to form a government) or the *last* formateur (i.e. the formateur that succeeds in forming a government) and not the entire sequence of formateurs. Therefore, in practical terms, I consider the DSR to be a special case of the GSR when p reaches infinity. When $p < \infty$, each party has a positive probability of becoming formateur if and only if it has a positive weight.

Consider now the effect of a change in p on π_i as given by the following derivative:

$$\partial\pi_i(w)/\partial p = \pi_i \cdot \sum_{j=1}^{n} \pi_j \cdot [\ln(w_i) - \ln(w_j)].$$

The summation term is a weighted average of the log differences; its sign determines the derivative's sign. Accordingly, if $w_1 > w_n$ (i.e., not all parties obtain the same weight), then $\partial\pi_1(w)/\partial p > 0$ because the summation term must be positive given that $\ln(w_1) \geq \ln(w_j)$ for all j with strict inequality for at least $j = n$. Correspondingly, $\partial\pi_n(w)/\partial p < 0$ because the summation term must be negative. For parties of intermediate size, the effect of p on π is not monotonic: π may rise for a small enough p, but it will eventually decline as p rises. Figure 1 illustrates the effect of p on π for the case where $w_1 = 0.42$, $w_2 = 0.38$ and $w_3 = 0.2$.

[Figure 1 about here]

3. The Data

The empirical analysis is based on government coalitions formed *immediately after general elections* in cases where no single party won a majority in 12 parliamentary democracies in the period 1945-2004: Austria (1949-2002), Belgium (1946-2003), Denmark (1945-2001), Finland (1945-1999), Germany (1949-2002), Iceland (1946-2003), Ireland (1948-2002), Israel (1949-1992, 2003), Luxembourg (1945-1999), The Netherlands (1946-2003), Norway (1961-2002) and Sweden (1948-2002).[5] Situations where a single party wins a majority are assumed to be straightforward as the majority party is in a position to form any type of government and is thus automatically treated as formateur.[6] I opt to use data about governments formed after elections in order to free the empirical results from dependence on any specific model of government termination. Otherwise, one must consider why the government was terminated and the implications of those reasons when transferring the mandate to a new formateur.[7]

To illustrate the analytical problems raised by new governments formed in the absence of elections, we take two examples from Israel. Consider first Israel's Twenty-Second government (1986-1988). That government was formed within the framework of a rotation agreement reached between the two largest parties in parliament devised during formation of the previous government. Clearly, treating formation of the two governments as *independent* is inappropriate.[8] Next, consider the Ben-Gurion government in power from 1950 to 1951. In this case, termination was motivated by the desire of the sitting prime minister (Ben-Gurion) to revise his cabinet. As this step was initiated by the prime minister, Israel's president (the state's head) immediately mandated Ben-Gurion's party as the new formateur. Inclusion of these cases in the analysis may introduce a problem known in the literature as *inflated number of observations*.[9] Hence, I restrict the analysis to governments formed subsequent to general elections, instances where it is more realistic to assume that a new, independent coalition-formation game will commence.[10]

For each country and government formed, the dataset contains the parliamentary weight each party possessed as well as the identity of the first and last formateur.[11] From a theoretical perspective, the identities of the first and last formateur are usually the same because in the majority of models employing selection rules, the first formateur succeeds in forming a coalition in equilibrium. In reality, however, this is not always the case, making the distinction between the two types of formateur empirically relevant.

The decision regarding which type of formateur to use (the first or the last) can be affected by the motive inspiring the analysis. If we focus on the *process* of new government formation, we need to identify the first formateur. Any inference drawn from formateur selection after the first formation attempt would then be conditional on the choice of the bargaining model used in the government formation process. As such, the process may introduce factors that prevent the formateur from forming a government (see Diermeier and Merlo, 2004). However, if we focus on the *outcome* of new government formation, we need to identify the last formateur. Obviously, in each instance of government formation the conclusions regarding the consistency of the selection rules will be valid only for the first/last formateur but not for the entire process, which potentially includes other formateurs.

For Germany, Ireland and Norway, I assumed that the first and last formateur (the incoming chancellor's and prime minister's parties, respectively) are identical by definition because an institutionalized system of formateurs has never been used, and because governments are formed by freestyle bargaining.[12] Moreover, it turns out that in Israel, Luxembourg and Sweden, too, the identities of the first and last formateurs were always the same. In Denmark, Finland and Iceland there are five cases each, in The Netherlands there are four cases and in Austria and Belgium there is one case each where the first and the last formateur are not the same.

The data sources employed were: Muller (2000, 2003) for Austria; De Winter et al. (2003) for Belgium; Thomas (1982), Skjæveland (2003) and Damgaard (2003) for Denmark; Nyholm (1982), Berglund (1995), Nurmi and Nurmi (2001) and Nousiainen (2003) for Finland; Norpoth (1982) and Saalfeld (2003) for Germany; Indrioason (2003) for Iceland; Mitchell (2003) for Ireland; Korn and Shapira (1997) for Israel; Dumont and De Winter (2003) for Luxembourg; Narud and Strom (2003) and Valen (2003) for Norway; Timmermans and Andeweg (2003) for The Netherlands; and Bergman (2003) for Sweden. Data for recent years (i.e., Austria, Germany, Ireland and Sweden in 2002, Belgium in 1999 and 2003, Denmark and Norway in 2001, Iceland and Israel in 2003, Luxembourg in 1999 and The Netherlands in 2002 and 2003) were elaborated by use of *Keesing's Record of World Events* (1999-2003), local newspapers and official Internet websites. Diermeier and Merlo's (2004) dataset was also used to obtain the first formateur's identity for Belgium (1946-1992), Denmark (1945-1953), Finland (1945-1999) and The Netherlands (1946-1994).

4. Empirical Analysis

The empirical analysis is organized as follows. First, nonparametric goodness-of-fit tests are employed to determine just how well the observed data fit the hypothesized distribution of formateurs according to each of the formateur selection rules. Each country's data is divided into two mutually exclusive categories; comparison of the observed and predicted frequencies of the formateurs found in each category is then computed. The allocation to the categories is arranged in two different partitions. In the first partition (see subsection 4.1), one category contains the largest party in each government about to be formed, with the other containing the remaining parties. In the second partition (see subsection 4.2), one category is composed of parties having weights no more than $\overline{w}$ ($\overline{w}$ denotes the upper weight bound of the first category) and the other of parties having weights above $\overline{w}$. Second, parameter p of the GSR is estimated by applying the maximum likelihood principle, followed by hypothesis testing of the DSR and the PSR (see subsection 4.3).

4.1 The largest versus the remaining parties

Let $T = \{1, 2, \ldots, t\}$ be the set of parliaments and let N^k *and* w^k be the set of parties and the vector of their weights in parliament $k \in T$, respectively. Let $v_{i,k} = 1$ if party $i \in N^k$ in government $k \in T$ is the formateur, and 0 otherwise; let o_1 and o_2 denote the number of *observations* (formateurs) out of t governments formed belonging to the first (largest party) and the second category (remaining parties), respectively. Accordingly, $o_1 = \sum_{k\in T} v_{1,k}$ and $o_2 = t - o_1$. Similarly, let e_1 and e_2 denote the *expected* number of formateurs in each category according to the operative formateur selection rule.

The GSR is evaluated by using the chi-square goodness-of-fit test (hereinafter chi-square test) as well as a modified version of the test suggested here for the nonidentically distributed case (hereinafter the modified chi-square test). The (unmodified) chi-square statistic is given by

$$\sum_{i=1}^{2} (o_i - e_i)^2 / e_i \, , \tag{3}$$

with the expected numbers of formateurs given by $e_1 = \sum_{k \in T} \pi_1(w^k)$ and $e_2 = t - e_1$. Critical values of the test are given by the chi-square distribution with one degree of freedom. Under the null hypothesis, the observed and expected numbers of formateurs in each category should be fairly close, within the limits of random error. If they are not sufficiently close, the chi-square statistic computed will be relatively high (more than 3.84 at the 5% significance level). Such a result will indicate that the GSR with the given value of p is inconsistent with the data.

Although the chi-square test is popular, it assumes that $v_{1,1}, v_{1,2}, \ldots, v_{1,t}$ are independent and *identically* distributed. Yet, according to the GSR (and the PSR as a special case), the identical distribution assumption is inappropriate because the probability of observing $v_{1,k} = 1$, which is equal to $\pi_1(w^k)$, is not identical for each coalition formation; instead, it depends on the distribution of weights in each parliament $k \in T$. To confirm the effect on the chi-square statistic, let $\bar{\pi}$ denote the identical (and average) probability that the largest party is selected as formateur. Accordingly, the chi-square statistic in (3) may be reformulated (substituting $o_2 = t - o_1$, $e_2 = t - e_1$ and $e_1 = t\bar{\pi}$) as:

$$\sum_{i=1}^{2} \frac{(o_i - e_i)^2}{e_i} = \frac{(o_1 - e_1)^2}{e_1 e_2 / t} = z^2 \, ,$$

where $z \equiv \dfrac{o_1 - t\bar{\pi}}{\sqrt{t\bar{\pi}(1-\bar{\pi})}}$. Using *the central limit theorem*, z has an asymptotic standard normal distribution because the expectation and variance of o_1 in the identically distributed case are equal to $t\bar{\pi}$ and $t\bar{\pi}$ $(1 - \bar{\pi})$, respectively. Hence, z^2 has an asymptotic chi-square distribution with one degree of freedom.

However, when the probabilities that the largest party is selected as formateur in each coalition government are not identical, the expectation and variance of o_1 are equal to $\sum_{k\in T} \pi_1(w^k) = t\bar{\pi}$ and $\sum_{k\in T} \pi_1(w^k)\cdot[1 - \pi_1(w^k)]$, respectively. Although the expectation of o_1 is the same as in the identically distributed case, the variance is not. Hence, without some modification, the statistic does not display an asymptotic chi-square distribution. Therefore, define the modified chi-square statistic as

$$\left(\frac{o_1 - \sum_{k\in T} \pi_1(w^k)}{\sqrt{\sum_{k\in T} \pi_1(w^k)[1-\pi_1(w^k)]}} \right)^2 = \frac{(o_1 - e_1)^2}{\sum_{k\in T} \pi_1(w^k)[1-\pi_1(w^k)]} . \tag{4}$$

The Lyapunov's central limit theorem for the nonidentically distributed case (see Breiman, 1968, pp. 186-188) can be employed to show that the modified chi-square statistic converges to a chi-square distribution with one degree of freedom under the condition that the probabilities $[\pi_1(w^k), \forall k \in T]$ do not approach either 1 or 0 (see Appendix for the formal conditions). Notably, it can be shown that the modified chi-square statistic is always larger than the chi-square statistic; therefore, the chi-square statistic will be conservative regarding rejection of the null hypothesis.[13] Stated differently, if we do not use the above modification, one might conclude that a selection rule is in agreement with the data when, in effect, it is not.

The DSR is evaluated in two ways: one strict (i.e., determining if the largest party is the formateur in every government formed), the other loose, allowing deviations from the rule. We do not reject the hypothesis that the DSR closely approximates the data when we do not reject the null hypothesis that $\bar{\pi} \geq 0.95$, at a 5% significance level (i.e., loosely).

Use of the chi-square test with the DSR is not an option because chi-square test results are reasonable only when the sample size is sufficiently large given that the sampling distribution of the test statistic is only *asymptotically* chi-square. Although it is difficult to provide hard and fast rules, in practice, the chi-square test works fairly well when the expected number in each category is 5 or more (see Sheskin, 2000). When the expected numbers are smaller, the probabilities associated with the chi-square distribution may not be sufficiently close to the test's sampling probability distribution for appropriate inferences to be made. Under the loose evaluation of the DSR, $e_2 = 0.05 \cdot t$, which yields less than 5 because $t < 100$ for *all* the countries analyzed. Under such circumstances, the binominal distribution is preferable.

The binominal distribution enables us to specify the one-sided upper limit of the 95% confidence interval for $\overline{\pi}$. This upper limit (see Neter, Wasserman and Whitmore, 1993), denoted here by $u(o_1)$, is given by:

$$u(o_1) = \frac{(o_1 + 1) \cdot F_{[2(o_1+1),2(t-o_1),0.95]}}{(o_1 + 1) \cdot F_{[2(o_1+1),2(t-o_1),0.95]} + (t - o_1)},$$

where $F_{[2(o_1+1),2(t-o_1),0.95]}$ is the inverse of the F probability distribution with cumulative probability of 0.95, $2 \cdot (o_1 + 1)$ degrees of freedom in the numerator and $2 \cdot (t - o_1)$ degrees of freedom in the denominator.[14] By definition, if $u(o_1) < 0.95$, we can reject the hypothesis that the DSR closely approximates the data at the 5% significance level.

Table 1 reports the observed and expected fraction of instances where the largest party is chosen as the first and the last formateur for each country in the sample. The last column in Table 1 reports the expected fraction of instances the largest party should be chosen according the GSR for selected values of p as well as the associated variance of the probabilities of that party becoming a formateur.

[Table 1 about here]

As we can see from Table 1, no case – with the exception of Israel – ever arises where the largest party is always selected as *both* the first and the last formateur. In Austria, the largest party has always been selected as first formateur although in one case the largest party was not the last formateur. Luxembourg also seems to have a very high tendency for the largest party as formateur. In other countries, with the exception of Norway, the largest party was appointed formateur in the majority of instances (between 50% and 83%) whereas in Norway, it was selected only in a minority of cases.

The expected fraction of formateurs increases as p increases. However, for $p = 1$, the expected fractions are too low (excluding Norway) and for relatively high values of p (e.g., $p = 20$), they are too high (excluding Austria and Israel). That is, Table 1 suggests that the PSR usually underestimates the largest party's probability of becoming formateur, whereas the DSR overestimates that probability in most cases.[15] Intuitively, Table 1 also implies that the GSR, which uses intermediate values of p, may fit the data.

To confirm this supposition, consider first the looser interpretation of the DSR, tested by the binominal-test. As the values of $u(o_1)$ in Table 1 show, the DSR may be a reasonable approximation for Austria, Israel, Luxembourg and The Netherlands (for last formateur only). However, for other cases, the DSR is inconsistent with the data even in our looser formulation. The Swedish case is a boundary case because although we reject the null hypothesis that the DSR reasonably approximates the data at the 5% significance level, we do not reject that hypothesis at lower significance levels.

We now turn to the chi-square tests (see Table 2). When the working rule of the chi-square test is satisfied (i.e., $e_1 > 5$ and $e_2 > 5$), both the chi-square statistic and the significance level (in parenthesis) are given; otherwise, only chi-square statistics are shown. As readily perceived from Table 2, with the exception of Ireland and Norway, we can reject the null hypothesis that the PSR correctly approximates the data for every country (see the results for $p = 1$).[16] Table 2 also indicates that for each country, there are values of p for which we cannot reject the null hypothesis that the GSR correctly approximates the data. We do find some variation in the GSR's consistency in different countries. While we do not reject the null hypothesis that the GSR is consistent with the data for Belgium, Denmark, Finland, Iceland, Ireland and Sweden when we obtain relatively low values of p (e.g., $p = 3$ and 4), we cannot reject the GSR (intuitively, but not formally) for Austria, Israel and Luxembourg for relatively high values of p (e.g., $p = 20$). When comparison of the results for the first and last formateur is reasonable (i.e., when the two formateurs are not necessarily the same, the case in Austria, Belgium, Denmark, Finland, Iceland and The Netherlands), the qualitative results are almost identical.[17]

[Table 2 about here]

Moreover, it turns out that no significant effect on the results is observed when using the modified chi-square statistic. The reason for this is that the variance of the probabilities of the largest party to be selected as formateur is quite low – less than 0.04 (see Table 1). Hence, use of average probability in the unmodified chi-square statistic is reasonable.

4.2 Partition according to weight

We now turn to the second partition, which is based on weight distribution. In this partition, the Kolmogorov-Smirnov goodness-of-fit test (hereinafter KS-test) and a special procedure based on the chi-square test serve as the nonparametric goodness-of-fit tests.

Put briefly, the KS-test involves specifying the *cumulative* frequency distribution of formateurs that would occur given the theoretical distribution specified by the GSR and compares that distribution with the observed cumulative frequency distribution in the sample. Let $w_{i,k}$ denote the weight of party $i \in N^k$ in parliament $k \in T$. The theoretical distribution, denoted by $G(\overline{w}, p)$, with $\overline{w}$ as the upper bound of weights, is given by $\sum_{k\in T}\sum_{i\in N^k} \pi_i(w^k)\cdot d_{i,k} / t$, where $d_{i,k} = 1$ if $w_{i,k} \leq \overline{w}$ and 0 otherwise. The observed distribution, denoted by $S(\overline{w})$, is given by $\sum_{k\in T}\sum_{i\in N^k} v_{i,k}\cdot d_{i,k} / t$. The KS-statistic ($K$) represents the largest of the deviations found between $S(c)$ and $G(c, p)$; it is formally given by $K = max_{\overline{w}\in[0,1]} |S(\overline{w}) - G(\overline{w}, p)|$. Under the null hypothesis that the sample has been drawn from the specified theoretical distribution, we expect that for every value of $\overline{w}$, the two distributions $S(\overline{w})$ and $G(\overline{w}, p)$ will be fairly close, within the limits of random error. If they are not sufficiently close, this indicates that the specified theoretical distribution does not adequately approximate the data. Critical values of K, denoted by $K_\alpha(t)$, depend on the number of governments (t) and the level of significance (α) (for more details see Massey, 1951 and Sheskin, 2000).

Chi-square tests are used under a setup where the first category contains parties having weights no more than $\overline{w}$ and the second the remaining parties. Formally, $e_1 = t \cdot G(\overline{w}, p)$, $e_2 = t - e_1$, $o_1 = t \cdot S(\overline{w})$ and $o_2 = t - o_1$. The modified chi-square statistic for the nonidentically distributed case in this partition is formulated similarly to (4), and given by:[18]

$$\frac{(o_1 - e_1)^2}{\sum_{k \in T} [\sum_{i \in N^k} \pi_i(w^k) \cdot d_{i,k}] \cdot [1 - \sum_{i \in N^k} \pi_i(w^k) \cdot d_{i,k}]}. \tag{5}$$

The critical values for both the modified and unmodified statistics are again given by the chi-square distribution with one degree of freedom.

Although the chi-square test is particularly appropriate for discrete data (as in the previous partition), it is sometimes used with continuous data (like the data in the current partition). The chi-square test has a serious drawback in such cases as the results (and conclusions) may be influenced by where we divide the categories (i.e., the test is sensitive to the value of $\overline{w}$). Therefore, a *Matlab* program was constructed that considers all plausible values for $\overline{w}$ that maintain the working rule for at least 5 expected observations in each category and locates the maximum value of the chi-square statistics (hereinafter denoted by chi-max and modified chi-max). Relatively high (more than 3.84 at the 5% significance level) values of the chi-max or modified chi-max indicate that the GSR (at the given value of p) is inconsistent with the data.

Table 3 reports the KS-test findings for different values of p for every country in the dataset. The KS-test results underscore the validity of the previous findings. Here the PSR is inconsistent with the data for Belgium, Denmark, Finland, Iceland (using the first formateur data), Israel, Luxembourg, The Netherlands and Sweden. Again, the Irish and Norwegian data are consistent with the PSR. Regarding the DSR, the KS-test does not allow us to reject the hypothesis that the DSR – in our looser definition – approximates the Finish and Swedish data. It should be recalled, however, that the hypothesis was outright rejected in the previous test. Again, there are values of p for which we do not reject the null hypothesis that the GSR correctly approximates the data. Moreover, the results strengthen the hypothesis that the GSR, with low values of p, fits the data for Denmark and Norway because we reject the hypothesis that the GSR is consistent with the data when employing $p = 6$, 7, 10 for Denmark and $p \geq 4$ for Norway.

[Table 3 about here]

Next, consider the results regarding the modified chi-max (see Table 4). In this case, use of the modified chi-max statistic has a more significant effect on the results than in the previous case. The reason for this result is that the variance of the probabilities of the parties from the first category to be selected formateur is relatively high, and can reach 0.27. To save space, the results of the unmodified chi-max are not reported here. Instead, Table 4 reports cases (in bold) where rejection of the null hypothesis occurs only when using modified chi-max statistics and not when using chi-max statistics.

[Table 4 about here]

As Table 4 reveals, we reject the null hypothesis that the PSR correctly approximates the data for every country excluding Norway (see the results for $p = 1$). Importantly, this test provides initial evidence that the PSR is inconsistent with the Irish data. Accordingly, excluding Norway, we now have some statistical evidence that the PSR is inconsistent with the data for every country in the dataset. Regarding Norway, the begged-for conclusion is that the PSR is a reasonable approximation.

The chi-max tests show that the GSR is inconsistent with the Belgian data (for p = 2), the Danish data (for p = 4, 5) and the Icelandic data (for p = 2, 6, 7). Other results reinforce the previous findings.

4.3 Maximum Likelihood Estimation of *p*

Maximum likelihood estimation provides a means for choosing an asymptotically efficient and consistent estimator of p displaying an asymptotically normal distribution.[19] The technique's logic is best illustrated by the following. Let $f^k \in N^k$ denote the (ex-post) formateur in parliament $k \in$ T; and let $\pi_i(w^k;p)$ denote the GSR probability of party $i \in N^k$ becoming formateur in parliament $k \in T$ with an as yet unknown parameter p. By construction, $\pi_{f^k}(w^k;p)$ is the joint probability that the formateur f^k will be selected and that each of the other parties will not be selected formateur in parliament $k \in T$. Assuming that government formations are independent, $L(p)=\prod_{k\in T}\pi_{f^k}(w^k;p)$ denotes the probability of observing a particular sample of formateurs, with a yet unknown parameter p. The maximum likelihood estimate of p (hereinafter denoted by p_{ML}) makes the sample of formateurs most probable [i.e., p_{ML} = arg max $L(p)$].

Because the logarithmic function is monotonically increasing and easier to work with, I maximize ln $L(p)$ instead of $L(p)$. The necessary condition for maximizing ln $L(p)$ for an interior solution is formally given by

$$\left.\frac{\partial \ln L(p)}{\partial p}\right|_{p=p_{ML}} = \sum_{k\in T}\sum_{i\in N^k} \ln\left(\frac{w_{f^k,k}}{w_{i,k}}\right)\cdot \pi_i(w^k;p_{ML}) = 0 .$$

The necessary condition is highly nonlinear; thus, the maximum likelihood estimator must be numerically sought out. To satisfy an interior maximum, the second derivatives of the ln $L(p)$ should be negative at $p = p_{ML}$. Formally:

$$\left.\frac{\partial^2 \ln L(p)}{\partial p^2}\right|_{p=p_{ML}} = \sum_{k\in T}\sum_{i\in N^k}[\ln(\frac{w_{f^k,k}}{w_{i,k}})\cdot \pi_i(w^k;p_{ML})\cdot \sum_{j\in N^k}\ln(\frac{w_{i,k}}{w_{j,k}})\cdot \pi_j(w^k;p_{ML})] < 0.$$

Hypothesis testing is based on confidence intervals, which provide a range of plausible values of p. Therefore, it stands to reason that if a hypothesized value of p does not fall within the range of plausible values, then the data are inconsistent with the hypothesis and the null hypothesis should be rejected. Formally, the test statistic for the hypothesis that the PSR (DSR) closely approximates the data is given by

$$Z = \frac{p_{ML} - p_0}{\sqrt{V_H(p_{ML})}},$$

where $V_H(p_{ML}) = \left[-\left.\frac{\partial^2 \ln L(p)}{\partial p^2}\right|_{p=p_{ML}} \right]^{-1}$

is the Hessian estimator for the asymptotic variance of p_{ML}, and p_0 is the null hypothesis value of p. I use $p_0 = 1$ for the PSR and $p_0 = 100$ for the DSR (in practical terms, $p \geq 100$ indicates that the DSR correctly approximates the data). If the maximum likelihood estimator is sufficiently different from 1, causing |Z| to be relatively high (above 1.96 at the 5% significance level for a two-tailed test), it indicates that the PSR does not approximate the data appropriately. Similarly, if p_{ML} is sufficiently *less than* 100, causing Z to be relatively low (below –1.645), this result will indicate that the DSR is inconsistent with the data.

Table 5 presents the maximum likelihood estimations of p. The maximum likelihood principle cannot be directly employed for Israel and Austria for first formateurs as the formateur has always been the largest party, suggesting that p_{ML} should equal infinity in these cases. Moreover, the principle is also invalid for the Finnish data as there are two cases where the first or last formateur is not a political party (i.e., it has zero weight, making the log-likelihood constant zero).

[Table 5 about here]

Table 5 indicates that with the exception of Israel, Austria (for first formateur) and Finland, which are not part of the maximum likelihood estimation, we should reject the null hypothesis that the DSR correctly approximates the data for every country. This result is equivalent to the strict evaluation of the DSR. It is worth mentioning that here, too, the Norwegian data can be considered as generated by the PSR. Again, the results reinforce the previous findings and suggest point estimates for p for each country. As readily seen, these point estimates are relatively small (below 12), strengthening the desirability of the GSR.

5. Conclusion

In this paper, I proposed a new formateur selection rule, which I call the generalized selection rule (GSR); the PSR as well as the DSR are considered special cases of this rule. I first showed that the GSR is more appealing than the previous formateur selection rules because it is distributional sensitive. I then provide careful statistical tests that show that the GSR is in greater agreement with the data from 12 parliamentary democracies over the period 1945-2004. Specifically, I developed a modified chi-square test that takes into account the fact that observations are taken from dissimilar distributions. I used two types of formateurs in the empirical analysis - the first and last formateur in any coalition formation – and found that the qualitative results are almost identical.

The empirical analysis shows that the DSR (i.e., the GSR with $p \to \infty$) provides a reasonable approximation for Austria, Israel and Luxembourg. There is also some evidence implying that the Dutch data for last formateur is closely approximated by the DSR. These findings weaken Diermeier and Merlo's (2004) argument that the DSR is inconsistent with the data. However, Austria, Israel and Luxembourg are not part of the Diermeier and Merlo dataset; thus, this paper enlarges the literature's empirical coverage and shows that the DSR's consistency is country specific.

Moreover, it was found that the PSR (i.e., the GSR with $p = 1$) is fairly close to the Norwegian data, but not to any other country in the dataset. This finding is a major contribution because Diermeier and Merlo argue that the PSR is consistent with the data for 8 out of the 11 countries in their dataset. Specifically, for 7 of the countries that are included in both datasets (i.e., Belgium, Denmark, Finland, Germany, Iceland, Ireland and The Netherlands), the finding here – rejection of the PSR – contradicts the Diermeier and Merlo findings. This incongruity reflects, in the main, a refinement of the statistical approach although it may also reflect some differences in the data.[20]

Importantly, in each of the countries tested (including those where either the PSR or the DSR provides a reasonable approximation of the data), the GSR, with intermediate values of p (i.e., neither 1 nor ∞), correctly approximates the data. For example, the GSR with $p = 2$ is a reasonable approximation of the Icelandic, Danish and German data according to each of the statistical tests considered in this paper.

To conclude, the GSR proposed in this paper appears to display more desirable properties than do the previous rules applied. Furthermore, because the information requirement for the GSR is quite minimal, it is practical for use in conjunction with several theoretical models.

Notes

[1] The term *formateur* denotes either a person or a party that proposes a government for parliamentary approval (see, e.g., Laver and Schofield, 1990).

[2] Baron and Ferejohn (1988) loosely comment that selection probabilities should be "related to the seat share." Nonetheless, in the illustration they present, they use proportional probabilities, a step that connects them to the stated selection rule (see also Diermeier and Merlo, 2004).

[3] To see this for the GSR case, note that

$$\frac{\partial\pi_i(w)}{\partial w_i}-\frac{\partial\pi_i(w)}{\partial w_j}=\frac{p\cdot\pi_i(w)\cdot[1-\pi_i(w)]}{w_i}+\frac{p\cdot\pi_i(w)\cdot\pi_j(w)}{w_j}>0,\ \forall i,j\in N,\ i\neq j.$$

[4] Formally, $\lim_{p\to\infty}\pi_i(\boldsymbol{w})=\lim_{p\to\infty}1/[\sum_{j\in N}(w_j/w_i)^p]=\begin{cases}1/g & if \quad i\leq g\\ 0 & if \quad i>g\end{cases}$, where g is the number of the largest parties, all with the same weight. Note that $\lim_{p\to\infty}(w_j/w_i)^p$ is equal to 1 if $w_i = w_j$, to zero if $w_i > w_j$ and to infinity if $w_i < w_j$. This means that if $w_1 > w_2$ (i.e., $g = 1$), then $\lim_{p\to\infty}\pi_1(w)=1$.

[5] Cases where a single party won a majority (and are thus omitted from the dataset): Austria in the 1966 and 1971-1979 elections, Belgium in the 1950 elections, Germany in the 1953-1957 elections, Ireland in the 1957, 1965, 1969 and 1977 elections, Luxembourg in the 1954 elections, Norway in the 1945-1957 elections and Sweden in the 1968 elections. I also exclude three governments formed in Israel during 1996-2003, a period of direct election of the prime minister. The associated legislation had stipulated that the (party of the) prime minister-elect must be chosen as formateur.

[6] In Diermeier, Eraslan and Merlo's (2003) sample of 9 West European countries for the period 1947–1997, there are 22 government formations where one party controlled a majority. In these cases, the majority party was always selected as formateur.

[7] Typical cabinet termination mechanisms that do not call for new elections may be the death of the prime minister, prime ministerial resignation for non-political reasons, voluntary expansion of the ruling coalition, cabinet defeats by the parliamentary opposition, conflicts between coalition partners, intra-party conflict and other more technical events (e.g., a cabinet resignation at the time a new state's head is installed in office).

[8] Inclusion of such government coalitions may generate sufficient variation to reject the DSR and accept the PSR in its place.

[9] This problem invalidates nonparametric goodness-of-fit tests. Moreover, inclusion of governments wherein the formateur is the largest party may generate an illusion that strengthens the DSR's

consistency with the data. If the formateur is not the largest party, inclusion of such governments in the sample can introduce sufficient variation to enable acceptance of the PSR.

[10] These examples, which reflect the dependence of new government formation on termination of the previous government, are not unique to Israel. In Denmark, for example, the government of the Social Democrat Hans Hedtoft, 1953-1955, was terminated when he died in office. The Social Democrat Hans C. S. Hansen took over as prime minister until the conclusion of Hedtoft's term in 1957, when new elections were held. Treating the two government formations as independent is dubious. Nor was such a case unique to Danish history. The same occurred during the second government of Hans C. S. Hansen, 1957-1960, when he as well died in office. His Social Democrat successor, Viggo Kampmann, took over as prime minister while forming a government for about 9 months, until a new government was formed after general elections. The second government of Viggo Kampmann, 1960-1962, was likewise terminated as he was forced to retire after a set of serious heart attacks. Social Democrat Jens Otto Krag took over as prime minister for the remainder of Kampmann's term, until the 1964 elections (see Thomas, 1982). For systematical documentation of coalition termination in Western Europe, see Muller and Strom (2003).

[11] Following Saalfeld (2003), members appointed for Berlin are excluded from party weights in the German data over the period 1949-1990. Moreover, following Norpoth (1982) and Saalfeld (2003) but not Diermeier and Merlo (2004), I treat the Christian Democratic Union and the Christian Social Union in Bavaria as one party because they have always acted in unison, as one parliamentary bloc.

[12] The last formateur need not always be the tentative prime minister's party. It may be that one party takes the roll of the formateur, whereas another nominates the prime minister. For example, in 1948 in The Netherlands, J.R.H. Van Schaik (Catholic People's Party) formed the government but W. Drees (Labor Party) became the prime minister, with Van Schaik serving as vice premier (see Lijphart, 1975, p. 136). Excluding the above case, I assume that the identities of the last formateur and the prime minister's party are the same.

[13] To see this, note that the numerator is the same in both forms of the test statistic whereas the denominator (i.e., the variance) is smaller in the modified chi-square statistic: $\sum_{k\in T}\pi_1(w^k)-\sum_{k\in T}[\pi_1(w^k)]^2<t\bar{\pi}-t\bar{\pi}^2 \Leftrightarrow \sum_{k\in T}[\pi_1(w^k)-\bar{\pi}]^2>0$. That is, the above is true whenever there is variation in $\pi_1(w^k)$ across different k's.

[14] If $t = o_1$ then $u(o_1) = 1$.

[15] This may be the reason why a simple mix of the DSR and PSR may be more consistent with the data (see Diermeier and Merlo, 2004).

[16] This is true with slight abuse of the working rule for Finland, Luxembourg and Norway (see Table 2, note "a").

[17] In Finland [Belgium], for $p = 2$, we do *not* reject the hypothesis that the GSR is consistent with the data where last [first] formateur is used, whereas we reject this hypothesis for the first [last] formateur. In The Netherlands, for $p = 3$ or 4, we do *not* reject the hypothesis that the GSR is consistent with the data where the first formateur is used, whereas we reject this hypothesis for the last formateur.

[18] A sufficient condition for using Lyapunov's central limit theorem in this procedure is that $\sum_{k\in T}[\sum_{i\in N^k}\pi_i(w^k)\cdot d_{i,k}]\cdot[1-\sum_{i\in N^k}\pi_i(w^k)\cdot d_{i,k}]$ is relatively large for large samples.

[19] The consistency property means that the true parameter value generated by the data is recovered asymptotically for sufficiently large samples. The efficiency property means that the lowest possible variance of parameter estimates is achieved asymptotically. Obviously, use of large sample properties in a sample with between 11 to 22 observations may be questionable. This is why maximum likelihood estimation is employed as a complementary procedure to nonparametric tests.

[20] Regarding differences in the data, in addition to the use of the first and last formateur, here I use only governments that were formed immediately following elections whereas Diermeier and Merlo use data on the first formateur only in every government coalition formed. Regarding the time

frames, in this paper I extend the data on government coalitions constructed consecutively up to 2004, beyond the time limits of the Diermeier and Merlo database. Based on the analysis I conducted of the Diermeier and Merlo data (details of the analysis can be obtained from the author upon request), it appears that after taking the chi-square working rule throughout their paper into account, I conclude that the PSR is less attractive than Diermeier and Merlo argue. In the revised version, the PSR can be considered a good approximation for only 4 (Finland, Germany, Iceland and Norway) out of the 11 countries in their sample. The results regarding Belgium, France, Ireland and The Netherlands alter the conclusion to be reached regarding the PSR: We now reject this rule. Regarding Finland, Germany and Iceland – as indicated in the present paper – differences in the data (especially regarding Germany, see footnote 10, above) and use of adjustments to nonidentical distributions indicate that the PSR is rejected for these countries as well.

Appendix: Lyapunov's theorem for the nonidentically distributed case

Let $X_1, X_2, \ldots, X_t$ be independent, where $X_k = v_{1,k} - \pi_1(w^k)$, $k \in$ T. Recall that $v_{1,k}$ takes the value 1 with probability of $\pi_1(w^k)$ and zero otherwise. Accordingly, $E(X_k) = 0$, $E(X_k^2) \equiv \mathrm{Var}(X_k) = \pi_1(w^k)\,[1 - \pi_1(w^k)] < \infty$ and $E(|X_k|^3) = [1 - \pi_1(w^k)]^3 \cdot \pi_1(w^k) + \pi_1(w^k)^3 \cdot [1 - \pi_1(w^k)] = [1 - \pi_1(w^k)] \cdot \pi_1(w^k) - 2\pi_1(w^k)^2 \cdot [1 - \pi_1(w^k)]^2 < \infty$. Let $E(X_k^2) = \sigma_k^2$ and $S_t^2 = \sum_{k\in T} \sigma_k^2$.

The condition for applying Lyapunov's central limit theorem for the nonidentically distributed case is the following (see Breiman, 1968, p. 186): $\lim_{n\to\infty} \frac{1}{S_t^3} \sum_{k\in T} | X_k |^3 = 0$. Namely, the left-hand side of the above condition, which is always positive by definition, is relatively small when the sample is large. Using the above notation, this condition can be reformulated as:

$$\frac{1}{S_t^3}\sum_{k\in T} | X_k |^3 = \frac{1}{\sqrt{\sum\limits_{k\in T} \pi_1(w^k)\cdot[1-\pi_1(w^k)]}} - \frac{2\sum_{k\in T}[\pi_1(w^k)]^2 \cdot [1-\pi_1(w^k)]^2}{\{\sum\limits_{k\in T}\pi_1(w^k)\cdot[1-\pi_1(w^k)]\}^{3/2}} < \frac{1}{\sqrt{\sum\limits_{k\in T} \pi_1(w^k)\cdot[1-\pi_1(w^k)]}}.$$

Accordingly, a sufficient condition for using Lyapunov's central limit theorem is that $\sqrt{\sum_{k\in T} \pi_1(w^k)\cdot[1-\pi_1(w^k)]}$ is relatively large for large samples; that is, $\pi_1(w^k)$ equal 1 or 0 for only a few k's.

References

Austen-Smith, David, and Jeffrey Banks. 1988. "Elections, Coalitions, and Legislative Outcomes." *American Political Science Review* 82(2): 405-22.

Baron, David P. 1991. "A Spatial Bargaining Theory of Government Formation in Parliamentary Systems." *American Political Science Review* 85(1): 137-64.

Baron, David P., and John A. Ferejohn. 1989. "Bargaining in Legislatures." *American Political Science Review* 83(4): 1181-206.

Breiman, Leo. 1968. *Probability*. Reading, Mass.: Addison-Wesley.

Berglund, Sten. 1995. "The Finnish Parliamentary Election of 1995." *Electoral Studies* 14(4): 461-4.

Bergman, Torbjorn. 2003. "Sweden: When Minority Cabinets Are the Rule and Majority Coalitions the Exception." In *Coalition Governments in Western Europe,* ed. Wolfgang C. Muller and Kaare Strom. Oxford, UK: Oxford University Press, 192-230.

Damgaard, Erik. 2003. "Denmark: The Life and Death of Government Coalitions." In *Coalition Governments in Western Europe,* ed. Wolfgang C. Muller and Kaare Strom. Oxford, UK: Oxford University Press, 231-63.

De Winter, Lieven, Arco Timmermans, and Patrick Dumont. 2003. "Belgium: On Government Agreements, Evangelists, Followers and Heretics." In *Coalition Formation Coalition Formation in Western Europe*, ed. Wolfgang C. Muller and Kaare Strom. Oxford, UK: Oxford University Press, 300-55.

Diermeier, Daniel, Hülya Eraslan, and Antonio Merlo. 2003. "A Structural Model of Government Formation." *Econometrica* 71(1): 27-70.

Diermeier, Daniel, and Antonio M. Merlo. 2004. "An Empirical Investigation of Coalitional Bargaining Procedures." *Journal of Public Economics* 88(3-4): 783-97.

Dumont, Patrick, and Lieven De Winter. 2003. "Luxembourg: Stable Coalitions in a Pivotal Party System." In *Coalition Governments in Western Europe*, ed. Wolfgang C. Muller and Kaare Strom. Oxford, UK: Oxford University Press, 399-432.

Greene, William H. 2000. *Econometric Analysis*. 4th ed. International ed. New Jersey: Prentice-Hall.

Indrioason, Indridi H. 2003. "A Theory of Coalition and Clientelism: Coalition Politics in Iceland 1945-2000." *European Journal of Political Research*. Forthcoming.

Korn, Dani, and Boaz Shapira. 1997. *Coalition Politics In Israel*. Tel-Aviv: Zmora-Bitan. [Hebrew]

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

Lijphart, Arend. 1975. *The Politics of Accommodation: Pluralism and Democracy in the Netherlands*. 2nd ed. Berkeley: University of California Press.

Massey, Frank J. 1951. "The Kolmogorov-Smirnov Test for Goodness of Fit." *American Statistical Association Journal* 46(253): 68-78.

Merlo, Antonio. 1997. "Bargaining over Governments in a Stochastic Environment." *Journal of Political Economy* 105(1): 101-31.

Mitchell, Paul. 2003. "Ireland: From Single-Party to Coalition Rule." In *Coalition Governments in Western Europe*, ed. Wolfgang C. Muller and Kaare Strom. Oxford, UK: Oxford University Press, 126-57.

Muller, Wolfgang C. 2000. "The Austrian Election of October 1999: a Shift to the Right." *West European Politics* 23(3): 191-200.

Muller, Wolfgang C. 2003. "Austria: Tight Coalitions and Stable Government." In Coalition Governments in Western Europe, ed. Wolfgang C. Muller and Kaare Strom. Oxford, UK: Oxford University Press, 86-125.

Muller, Wolfgang C. and Kaare Strom. 2003. *Coalition Governments in Western Europe*. Oxford, UK: Oxford University Press.

Narud, Hanne Marthe, and Kaare Strom. 2003. "Norway: A Fragile Coalition Order." In *Coalition Governments in Western Europe*, ed. Wolfgang C. Muller and Kaare Strom. Oxford, UK: Oxford University Press, 158-91.

Neter, John, William Wasserman, and G. A. Whitmore. 1993. *Applied Statistics.* 4th ed. Boston, MA: Allyn and Bacon.

Norpoth, Helmut. 1982. "The German Federal Republic: Coalition Government at the Brink of Majority Rule." In *Government Coalitions in Western Democracies*, ed. Eric C. Browne and John Dreijmanis. New York: Longman, 7-31.

Nousiainen, Jaakko. 2003. "Finland: The Consolidation of Parliamentary Governance." In *Coalition Governments in Western Europe,* ed. Wolfgang C. Muller and Kaare Strom. Oxford, UK: Oxford University Press, 264-99.

Nurmi, Hannu, and Lasse Nurmi. 2001. "The 1999 Parliamentary Elections in Finland." *Electoral Studies* 20(1): 147-57.

Nyholm, Pekka. (1982). "Finland: A Probabilistic View of Coalition Formation." In *Government Coalitions in Western Democracies*, ed. Eric C. Browne and John Dreijmanis. New York: Longman, 71-108.

Saalfeld, Thomas. 2003. "Germany: Stable Parties Chancellor Democracy, and the Art of Informal Settlement." In *Coalition Governments in Western Europe*, ed. Wolfgang C. Muller and Kaare Strom. Oxford, UK: Oxford University Press, 32-85.

Sheskin, David J. 2000. *Handbook of Parametric and Nonparametric Statistical Procedures*. 2nd ed. Boca Raton: Chapman & Hall/CRC.

Siegel, Sidney, and N. John Castellan Jr. 1988. *Nonparametric Statistics for the Behavioral Sciences.* 2nd ed. International ed. Singapore: McGraw-Hill.

Skaperdas, Stergios. 1996. "Contest Success Functions." *Economic Theory* 7(2): 283-90.

Skjæveland, Asbjørn. 2003. *Government Formation in Denmark 1953-1998.* Department of Political Science and University of Aarhus.

Timmermans, Arco, and Rudy B. Andeweg. 2003. "The Netherlands: Still the Politics of Accommodation?" In *Coalition Formation Coalition Formation in Western Europe*, ed. Wolfgang C. Muller and Kaare Strom. Oxford, UK: Oxford University Press, 355-98.

Thomas, Alastair H. 1982. "Denmark: Coalitions and Minority Governments." In *Government Coalitions in Western Democracies*, ed. Eric C. Browne and John Dreijmanis. New York: Longman, 109-41.

Tullock, Gordon. 1980. "Efficient Rent Seeking." In *Toward a Theory of Rent-Seeking Society*, ed. James .M. Buchanan, Robert .D. Tollison and Gordon Tullock. College Station: Taxes A&M University Press, 97-112.

Valen, Henry. 2003. "The Storting Election in Norway, September 2001." *Electoral Studies* 22(1): 179-85.

Figure 1: Effect of p on $\pi(w)$ for $w_1 = 0.42$, $w_2 = 0.38$ and $w_3 = 0.2$

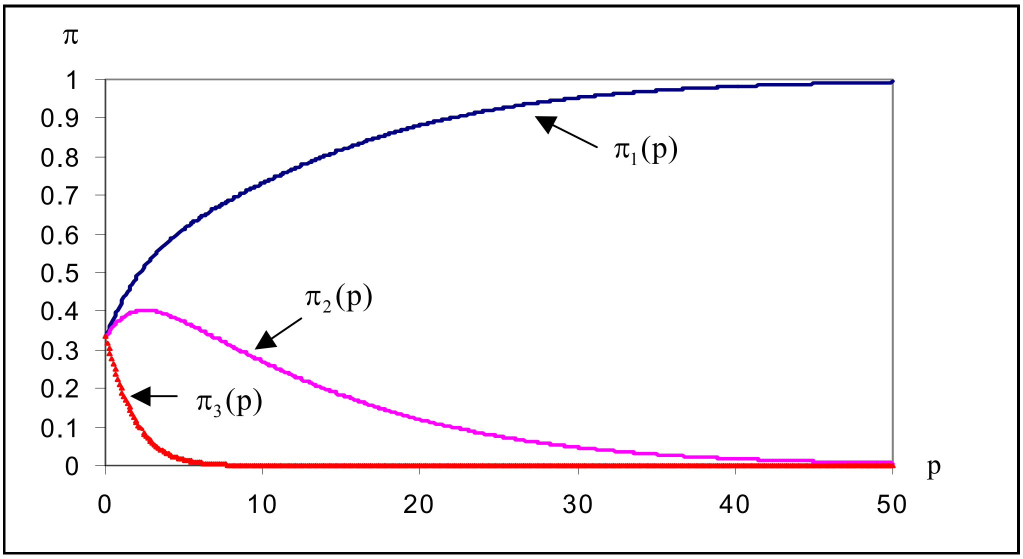

Table 1: Observed and expected formateur frequencies in the largest party category

	t	First formateur		Last formateur		Expected frequencies (e_1/t) and variance of formateurs in the largest party category				
		o_1/t	$u(o_1)$	o_1/t	$u(o_1)$	$p = 1$	$p = 2$	$p = 4$	$p = 7$	$p = 20$
Austria	13	1.00	1.000	0.92	0.996	0.44 (0.002)	0.52 (0.002)	0.60 (0.005)	0.68 (0.015)	0.84 (0.027)
Belgium	18	0.61	0.801*	0.6	0.844*	0.31 (0.013)	0.44 (0.016)	0.60 (0.020)	0.74 (0.021)	0.93 (0.009)
Denmark	22	0.68	0.840*	0.64	0.804*	0.36 (0.002)	0.59 (0.007)	0.82 (0.012)	0.93 (0.009)	0.99 (0.002)
Finland	16	0.63	0.822*	0.50	0.721*	0.27 (0.000)	0.37 (0.002)	0.50 (0.011)	0.63 (0.029)	0.81 (0.036)
Germany	13	0.77	0.934*	0.77	0.934*	0.46 (0.002)	0.55 (0.002)	0.63 (0.007)	0.71 (0.016)	0.87 (0.025)
Iceland	18	0.61	0.801*	0.72	0.884*	0.38 (0.001)	0.52 (0.005)	0.71 (0.012)	0.85 (0.015)	0.96 (0.005)
Ireland	13	0.62	0.834*	0.62	0.834*	0.47 (0.001)	0.64 (0.005)	0.80 (0.012)	0.89 (0.011)	0.98 (0.003)
Israel	14	1.00	1.000	1.00	1.000	0.37 (0.002)	0.63 (0.008)	0.82 (0.029)	0.88 (0.031)	0.93 (0.021)
Luxembourg	12	0.92	0.996	0.92	0.996	0.39 (0.002)	0.51 (0.008)	0.66 (0.017)	0.78 (0.019)	0.94 (0.010)
Netherlands	18	0.78	0.920*	0.83	0.953	0.42 (0.005)	0.65 (0.013)	0.86 (0.019)	0.94 (0.011)	0.99 (0.001)
Norway	11	0.36	0.650*	0.36	0.650*	0.33 (0.006)	0.47 (0.011)	0.60 (0.018)	0.69 (0.022)	0.85 (0.018)
Sweden	17	0.82	0.950*	0.82	0.950*	0.45 (0.001)	0.70 (0.002)	0.93 (0.001)	0.99 (0.000)	1.00 (0.000)

Note: * Rejection of the DSR according to the binominal-test at the 5% significance level. The numbers in parentheses are the variance of the probabilities of the largest party becoming a formateur.

Table 2: Chi-square statistics for the GSR – largest party versus the rest

	p = 1	*p* = 2	*P* = 3	*p* = 4	*p* = 5	*p* = 6	*p* = 7	*p* = 10	*p* = 20
First formateur									
Austria	16.3*	12.1*	10.13*	8.76*	7.70	6.84	6.13	4.62	2.52
	(0.000)	(0.001)	(0.001)	(0.003)					
Belgium	7.52*	2.25	0.51	0.01	0.14	0.66	1.46	5.08	26.93
	(0.006)	(0.134)	(0.475)	(0.905)	(0.707)	(0.417)			
Denmark	9.85*	0.79	0.34	3.1	7.79	13.86	21.01	57.49	181.32
	(0.002)	(0.373)	(0.560)						
Finland	10.18*[a]	4.47*	2.17	0.99	0.36	0.07	0.00	0.46	3.62
	(0.001)	(0.034)	(0.141)	(0.321)	(0.548)	(0.789)	(0.985)		
Germany	5.10*	2.52	1.59	1.05	0.67	0.40	0.21	0.00	1.04
	(0.024)	(0.112)	(0.207)						
Iceland	4.18*	0.55	0.04	0.95	2.65	4.9	7.55	17.39	66.18
	(0.041)	(0.457)	(0.835)	(0.330)					
Ireland	1.16	0.04[a]	0.97	2.6	4.75	7.42	10.6	23.28	92.69
	(0.281)	(0.845)							
Israel	23.89*	8.25*	4.53	3.16	2.52	2.16	1.92	1.52	1.05
	(0.000)	(0.004)							
Luxembourg	14.25*[a]	7.79*	4.99	3.44	2.46	1.78	1.3	0.47	0.07
	(0.000)	(0.005)							
Netherlands	16.58*	6.85*	3.77	2.36	1.55	1.03	0.67	0.12	0.69
	(0.000)	(0.009)	(0.052)	(0.125)	(0.213)	(0.310)	(0.412)		
Norway	0.16[a]	4.08	12.19	23.48	36.82	51.48	67.25	123.17	528.48
	(0.691)								
Sweden	9.45*	1.22	0.10	2.58	8.98	21.51	44.56	287.11	61982
	(0.002)	(0.270)							

Table 2 (cont.)

	$p = 1$	$p = 2$	$P = 3$	$p = 4$	$p = 5$	$p = 6$	$p = 7$	$p = 10$	$p = 20$
Last formateur [b]									
Austria	12.1* (0.001)	8.54* (0.003)	6.88* (0.009)	5.74* (0.017)	4.84	4.13	3.54	2.31	0.70
Belgium	10.57* (0.001)	3.90* (0.048)	1.41 (0.236)	0.36 (0.549)	0.01 (0.905)	0.09 (0.766)	0.45	2.71	18.31
Denmark	7.26* (0.002)	0.21 (0.373)	1.14 (0.56)	5.39	11.83	19.95	29.40	63.93	239.06
Finland	4.26* [a] (0.039)	1.13 (0.288)	0.19 (0.662)	0.00 (0.963)	0.21 (0.650)	0.65 (0.418)	1.26 (0.262)	3.49	11.24
Iceland	9.1* (0.003)	2.85 (0.092)	0.59 (0.441)	0.00 (0.945)	0.26	1.00	2.09	6.66	31.13
Netherlands	20.92* (0.000)	9.54* (0.002)	5.84* (0.016)	4.07* (0.044)	3.01 (0.083)	2.29 (0.130)	1.77 (0.184)	0.79	0.03

Notes: * Rejection of the GSR for a given p at the 5% significance level. The numbers in parentheses are the significance levels of the test, which are given only when the working rule for the chi-square test is satisfied.

[a] The expected number of observations in the first category (largest party) is only 4.33 for Finland and 4.63 for Luxembourg. The expected number of observations in the second category is only 4.66 for Ireland and 4.65 for Norway (i.e., each is less than 5).

[b] The results regarding the last formateur in Germany, Ireland, Israel, Luxembourg, Norway and Sweden are omitted because they are identical to those regarding the first formateur.

Table 3: KS-test of the GSR – testing the distribution of formateurs according to weight

	$p = 1$	$p = 2$	$p = 3$	$p = 4$	$p = 5$	$p = 6$	$p = 7$	$p = 10$	$p \to \infty$ (DSR)	$K_{0.05}(t)$
					First formateur					
Austria	0.253	0.181	0.174	0.168	0.162	0.156	0.150	0.134	0.000	0.361
Belgium	0.380*	0.237	0.158	0.130	0.143	0.154	0.162	0.182	0.222	0.309
Denmark	0.315*	0.116	0.111	0.186	0.233	0.262	0.280	0.304*	0.318*	0.281
Finland	0.430*	0.306	0.222	0.157	0.105	0.063	0.079	0.136	0.250	0.328
Germany	0.356	0.269	0.228	0.200	0.177	0.156	0.138	0.131	0.231	0.361
Iceland	0.312*	0.211	0.136	0.144	0.186	0.214	0.232	0.26	0.333*	0.309
Ireland	0.310	0.133	0.143	0.196	0.233	0.261	0.283	0.324	0.385*	0.361
Israel	0.527*	0.277	0.165	0.113	0.086	0.07	0.06	0.043	0.000	0.349
Luxembourg	0.437*	0.316	0.240	0.192	0.158	0.133	0.113	0.079	0.083	0.375
Netherlands	0.346*	0.219	0.163	0.133	0.115	0.124	0.134	0.148	0.167	0.309
Norway	0.185	0.257	0.387	0.453*	0.486*	0.505*	0.517*	0.533*	0.545*	0.391
Sweden	0.375*	0.200	0.096	0.124	0.151	0.164	0.17	0.175	0.176	0.318
					Last formateur [a]					
Austria	0.199	0.181	0.174	0.168	0.162	0.156	0.15	0.134	0.077	0.361
Belgium	0.435*	0.292	0.214	0.165	0.143	0.154	0.162	0.182	0.222	0.309
Denmark	0.322*	0.116	0.155	0.230	0.279	0.307*	0.325*	0.349*	0.364*	0.281
Finland	0.385*	0.273	0.196	0.134	0.125	0.125	0.137	0.169	0.313	0.328
Iceland	0.304	0.167	0.112	0.104	0.145	0.17	0.187	0.217	0.278	0.309
Netherlands	0.418*	0.243	0.163	0.133	0.115	0.101	0.090	0.067	0.083	0.309

Notes: * Rejection of the GSR for a given value of p at the 5% significance level.

[a] The results regarding the last formateur in Germany, Ireland, Israel, Luxembourg, Norway and Sweden are omitted because they are identical to those regarding the first formateur.

Table 4: Modified chi-max results for the GSR

	$p = 1$	$p = 2$	$p = 3$	$p = 4$	$p = 5$	$p = 6$	$p = 7$	$p = 10$	$p = 20$
					First formateur				
Austria	**5.07***	2.64	2.35	2.16	2.00	1.86	1.73	1.42	0.90
	(0.024)								
Belgium	15.33*	7.55*	3.46	5.36	7.15	8.99	10.98	18.54	103.07
	(0.000)	(0.006)							
Denmark	10.88*	1.89	1.6	**4.45$^{*\,a}$**	**11.20$^{*\,a}$**	24.26	50.90	368.55	>10000
	(0.001)	(0.169)	(0.206)	**(0.035)**	**(0.001)**				
Finland	13.89*	6.79*	3.72	2.02	0.83	0.28	0.52	1.89	12.77
	(0.000)	(0.009)	(0.054)	(0.155)	(0.395)	(0.598)	(0.470)		
Germany	9.39*	2.91	1.38	0.88	0.52	0.87	1.31	3.08	12.02
	(0.002)	(0.088)	(0.240)	(0.339)	(0.472)	(0.351)	(0.253)	(0.079)	
Iceland	7.64*	**5.39$^{*\,a}$**	2.62	1.81	2.43	**4.68$^{*\,a}$**	**4.76$^{*\,a}$**	2.50	13.09
	(0.006)	**(0.020)**	(0.106)	(0.179)	(0.119)	**(0.030)**	**(0.029)**		
Ireland	6.18*	0.61	0.23	0.48	1.45	2.44	3.71	9.73	98.49
	(0.013)	(0.435)	(0.631)						
Israel	18.32*	**5.47***	2.97	1.95	1.49	1.03	0.88	0.59	0.25
	(0.000)	**(0.019)**	(0.085)						
Luxembourg	10.86*	6.92*	**4.65***	3.46	2.56	2.15	1.86	1.30	0.57
	(0.001)	(0.009)	**(0.031)**	(0.063)					
Netherlands	13.45*	**6.90***	**4.81***	3.79	3.15	2.70	2.34	1.63	0.70
	(0.000)	**(0.009)**	**(0.028)**						
Norway	0.39	1.40	**8.10***	19.90	30.83	43.35	57.99	123.45	1219.50
	(0.532)	(0.237)	**(0.004)**						
Sweden	13.58*	**4.75***	1.91	0.84	0.39	0.19	0.09	0.01	0.00
	(0.000)	**(0.029)**	(0.167)						

Table 4 (cont.)

	$p = 1$	$p = 2$	$p = 3$	$p = 4$	$p = 5$	$p = 6$	$p = 7$	$p = 10$	$p = 20$
Last formateur [b]									
Austria	**5.07*** **(0.024)**	2.64	2.35	2.16	2.00	1.86	1.73	1.42	0.90
Belgium	20.17* (0.000)	11.51* (0.001)	6.24	5.36	7.15	8.99	10.98	18.54	103.07
Denmark	10.88* (0.001)	1.89 (0.169)	1.6 (0.206)	**6.73*** [a] **(0.009)**	**11.20*** [a] **(0.001)**	24.26	50.90	368.55	>10000
Finland	10.6* (0.001)	5.93* (0.015)	3.10 (0.078)	1.56 (0.212)	0.67 (0.414)	0.19 (0.665)	0.52 (0.470)	1.89	12.77
Iceland	7.98* (0.005)	3.86	1.77	1.81	1.27 (0.261)	2.80 (0.095)	**4.76*** [a] **(0.029)**	8.21	32.92
Netherlands	15.86* (0.000)	7.13* (0.008)	**4.81*** **(0.028)**	3.79	3.15	2.70	2.34	1.63	0.70

Notes: * Rejection of the GSR for a given p at the 5% significance level. The numbers in parentheses are the significance levels of the test that are given only when less than half of the probabilities are either above 0.95 or below 0.05 in order to be consistent with Lyapunov's theorem. Results in bold indicate that rejection of the null hypothesis occurs only when using the modified chi-max statistic and not when using the chi-max statistic.

[a] Rejection of the null hypothesis occurs only when using the modified chi-max statistic, not in any pervious tests.

[b] The results regarding the last formateur in Germany, Ireland, Israel, Luxembourg, Norway and Sweden are omitted because they are identical to those regarding the first formateur.

Table 5: Maximum likelihood estimation of the GSR

	p_{ML}	$Std_H(p_{ML})$	Statistic of the test for the DSR: H_0: $p_{ML} \leq 100$	Statistic of the test for the PSR: H_0: $p_{ML} = 1$
		First formateurs		
Belgium	4.8344	1.5171	-62.730*	2.528*
Denmark	2.4372	0.5227	-186.669*	2.750*
Germany	10.6608	5.6866	-15.711*	1.699
Iceland	3.0518	0.9039	-107.250*	2.270*
Ireland	2.6488	1.0390	-93.716*	1.587
Luxembourg	9.9123	4.8030	-18.757*	1.856
Netherlands	2.7494	0.8354	-116.414*	2.094*
Norway	1.3148	0.4738	-208.295*	0.665
Sweden	3.0712	0.7820	-123.957*	2.649*
		First formateurs [a]		
Austria	11.6380	6.1573	-14.351*	1.728
Belgium	6.1000	1.9470	-48.227*	2.619*
Denmark	2.6026	0.5614	-173.486*	2.855*
Iceland	3.0497	0.9034	-107.321*	2.269*
Netherlands	4.0324	1.4519	-66.096*	2.088*

Notes: * Rejection of the PSR / DSR at the 5% significance level.

[a] The results regarding the last formateur in Germany, Ireland, Luxembourg, Norway and Sweden are omitted because they are identical to those regarding the first formateur.