Cabinet Survival and Competing Risks

Daniel Diermeier, *Northwestern University*
Randy T. Stevenson, *Rice University*

We attempt to resolve a recent controversy in the study of cabinet terminations pertaining to the shape of hazard rates. On the one hand, Warwick (1992b) provides evidence that cabinets are more likely to terminate the longer they are in office. Alt and King's (1994) analysis, on the other hand, suggests that hazard rates are constant over the life-time of a cabinet. This issue is of particular theoretical importance, since a constant hazard rate would add support to the nonstrategic model of cabinet termination due to Browne et al. (1986) while an increasing hazard rate would seem to favor Lupia and Strom's (1995) strategic approach. By applying a semi-parametric competing risk approach to data on cabinet durations, we are able to show that through its use of theory-based censoring the previous literature in effect analyzed only one mode in which cabinets terminate: the case where one cabinet is replaced by another without a new election. Once cabinet terminations that lead to chamber dissolutions with subsequent elections are analyzed directly, we can show that they are governed by a very different stochastic process. Hazard rates are not flat as in the case of replacements, but increase over the life of the government. Further the covariates governing replacement terminations fail to explain dissolution terminations. These findings add support to the strategic approach suggested by Lupia and Strom.

1. Introduction

Since the work of Browne et. al. (1984, 1986, 1988) and Strom (1985, 1988), the literature on cabinet duration has been characterized by two different approaches. One approach has been to identify a robust set of covariates that influence mean cabinet duration. The other approach has tried to determine the stochastic process that governs cabinet survival. Since Browne et. al., this second aspect has been pursued through the concept of "critical events," exogenous random shocks that destabilize an existing government. These random shocks are assumed to follow a Poisson process which implies a constant hazard rate for cabinet terminations. Cabinets are no more likely to fall today than they were yesterday.

The somewhat heated debate between these two approaches, however, subsided, when King et. al. (1990) presented a "unified" model of cabinet dissolutions that combined the attribute approach proposed by Strom (1985, 1988) and the events approach introduced by Browne et. al. (1986, 1988). King et. al.'s approach retained Browne et. al.'s constant hazard rate, but allowed it to depend on a set of time-invariant covariates, the attributes suggested by Strom. The unified model demonstrated that considerable insight can be gained from experimenting with nonnormal stochastic components

American Journal of Political Science, Vol. 43, No. 4, October 1999, Pp. 1051–1068

and focusing on the hazard rate as the critical structural component in the stochastic process.

Following this insight, Warwick and Easton (1992) pointed out that the unified model is but a special case of a whole class of survival models that relax the assumption of fixed hazards. In a series of works, Warwick (1992a, 1992b, 1992c, 1994) as well as Warwick and Easton (1992) consider different versions of time-varying hazard rates.

Lupia and Strom (1995) recently presented a game-theoretic model that sheds some new light on our understanding of cabinet termination. Lupia and Strom suggest that what makes an event critical depends on the strategic interaction between the current and potential members of a governing coalition. The key idea of their model is to interpret an "event" as a bargaining parameter that determines each party's outside option if the current government falls. Whether a cabinet terminates then depends on the attractiveness of this outside option compared to each party's payoff from the current government. Specifically, events correspond to informative public opinion polls or shared expectations about electoral outcomes that provide each party with information about what would happen if an election was held today. Whether an event leads to the fall of a government then depends on the relation between the expected electoral net benefits and the benefits and costs from a nonelectoral redistribution of power. Expected electoral gains thus do not necessarily lead to early elections. Rather, a party with favorable electoral prospects may extract advantages through renegotiations. Early elections only occur if these benefits cannot be captured given the current allocation of seats in the chamber.

This approach implies that in each round of bargaining one of three possible outcomes can occur: the status quo government persists, the chamber is dissolved and early elections are held, or the incumbent cabinet is directly replaced by a new one. So, either a government survives or it fails according to one of two termination modes, which we call *dissolution* and *replacement*. Lupia and Strom give necessary and sufficient conditions for each of these outcomes. This approach has important consequences for the empirical study of cabinet terminations. Rather than simply studying the failure of governments one now also has to take into account their mode of termination. As Lupia and Strom show, the factors that determine cabinet survival (irrespective of the termination mode) are in general different from those that decide whether cabinets end in dissolution or replacement. This suggests that different modes of cabinet termination are governed by different stochastic processes. Lupia and Strom, however, do not indicate how the strategically relevant parameters can be operationalized and measured.

Lupia and Strom argue that their model implies increasing pooled hazard rates. In contrast, a constant hazard rate would add support to the nonstrategic model of cabinet termination due to Browne et. al.[1] Hazard rates could thus, in principle, be used to empirically evaluate the two competing research programs. A look at the empirical efforts in the field, however, seems to produce conflicting results. The main line of disagreement is between those who think that hazard rates rise over the life of a cabinet (Warwick and Easton 1992, Warwick 1994) and those who maintain that they are flat (King et. al. 1990; Alt and King 1994). As with any empirical endeavor, differences in methods and data have generated some of this disagreement, but in our opinion these effects are marginal. Rather, the main difference between these authors is really one of interpretation, not an actual discongruity in the reported results. More precisely, Warwick's (1992b, 1994) claim of increasing hazards is based on the statistical significance of an upward trend in their country specific and pooled samples, and no one really disputes that this finding is technically correct. What is in dispute, however, is the substantive significance of the results. Specifically, Alt and King (1994) show that, for Warwick and Easton's pooled sample, the slope of the hazard rate changes by only 10 percent around its mean value (from .027 to .033) over its entire range. Further, a similar examination of the country specific estimates given in Warwick (1992b) reveals the same pattern (indeed the change is an order of magnitude smaller). Substantively, Alt and King argue, these small changes, although statistically significant, are substantively equivalent to a constant rate. We interpret this to mean that while there is some evidence for increasing hazard rates in samples that do not distinguish between types of failure, all the available evidence suggests these increases are small enough to be considered substantively flat.

What then about Strom and Lupia's prediction that hazard rates should be increasing? Even if their prediction is not falsified statistically, the current literature can hardly be taken as a clear confirmation. Indeed, with such a small rise in the hazard, one may question the need for all the extra complexity of Lupia and Strom's strategic model when compared to Browne et. al.'s simple alternative.

In a recent paper Diermeier and Stevenson (1997) argue that an investigation of pooled hazard may not be the most powerful test of the Lupia and Strom model. Although Lupia and Strom do provide necessary and sufficient conditions for each termination mode, the empirical implications of their analysis remain unclear. To derive testable predictions Diermeier and Stevenson embed Lupia and Strom's deterministic model in a stochastic environment similar to Browne et al. (1984, 1986). This allows them to formulate the nonstrategic model due to Browne et al. as the null hypothesis and to test whether Lupia and Strom's model significantly improves upon

[1]For different view of the meaning of increasing hazard rates see Warwick (1994).

the nonstrategic approach. Cabinets make decisions about whether to continue the current government by comparing randomly arriving outside options with the expected utility of staying in office. Since the latter is decreasing as the next mandatory election approaches, both pooled and dissolution hazard rates will increase. This, however, may not be true for replacement hazards. Intuitively, this follows because a government that forms due to a replacement still faces the same mandatory next election as its predecessor, while a dissolution "resets the clock" for the new government and pushes the next mandatory election back in time.[2] Since the Lupia and Strom framework implies increasing hazard rates only for pooled failures and dissolutions, but not for replacement terminations, a finding of flat hazard rates for cabinet replacement is perfectly consistent with the Lupia and Strom model.

This clarification is useful in understanding previous empirical results because it implies that a constant (or downward sloping) replacement hazard rate will, when combined with an upward sloping curve for dissolutions (as Lupia and Strom's model predicts), result in a *pooled* hazard rate that will increase at a rate smaller than for dissolution separately—thus, making it harder to detect an increase in an empirical analysis of pooled termination types. This would allow us to account for the disagreements in the previous empirical literature. In samples that do not distinguish between termination types, increasing hazards are apparent, but are quite small.

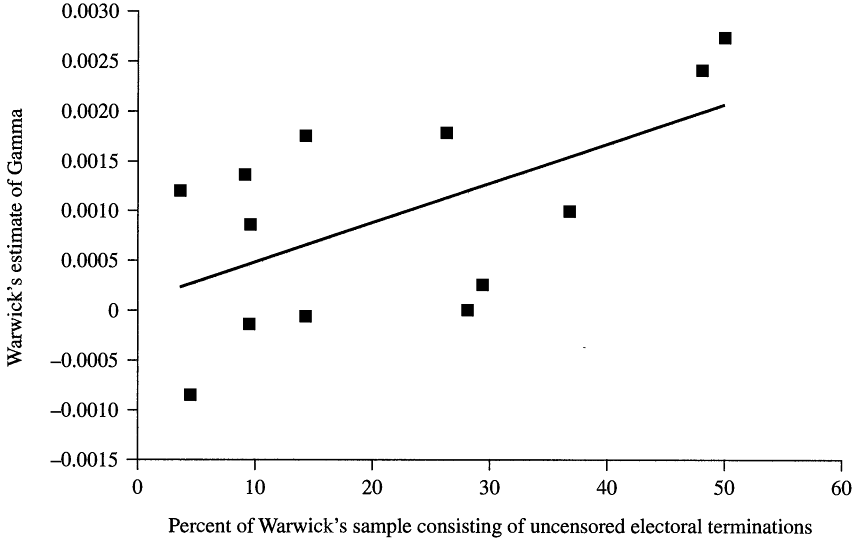
This tendency will be more pronounced if the sample contains mainly replacement failures or if dissolutions are censored randomly.[3] This can be illustrated by examining the country specific hazard rates that are reported by Warwick (1994).

In Figure 1, a measure of the slope of the hazard rate that was estimated by Warwick is given on the Y-axis and the percentage of failures (in each country) that terminated due to an election and that were uncensored is given on the X-axis. The positive relationship between the variables gives some preliminary support to the idea that the more the sample consisted of re-

[2]In order to compare our competing risk model with recent empirical studies of cabinet survival we need to ignore some of the subtleties of the Diermeier and Stevenson model. For instance, they point out that rather than estimating hazard rates in elapsed time one should use hazard rates in remaining time (until the next mandatory election). While theoretically sound, the distinction is of little practical relevance if hazard rates are monotone. Further, such an approach would make a comparison with the previous literature impossible, which has consistently used hazard rates in elapsed time.

[3]In King et al. (1990), most cabinet terminations that occur within twelve months of a regularly scheduled election are censored. The rationale is that these governments would have survived longer if the end of the "constitutional inter-election period" (CIEP) had not necessitated elections. This procedure, in a slightly different form, is also followed by Warwick and his colleagues. We will discuss the issue of censoring extensively in the next section.

Figure 1. Variation between Hazard Shape and Dissolution in 13 Countries



placement cabinet failures and censored dissolutions, the smaller the slope of the pooled hazard rate estimated by Warwick.[4]

The competing risk approach, however, is not only useful in reconciling disagreements in previous empirical studies. More importantly, it can be used to devise a more powerful tool in assessing the relative merit of the strategic approach due to Lupia and Strom. Diermeier and Stevenson (1997) show that while both pooled and dissolution hazards should be increasing, dissolution hazards may increase at a steeper rate. This would be the case, for instance, if replacement hazard rates are flat. Competing risk analysis allows us to measure dissolution hazards directly and thus test the strongest implication of the Lupia and Strom model.

In the next section we briefly introduce competing risk analysis and show that the censoring regime used by King et. al. (1990), Alt and King (1994) and Warwick and Easton (1992), and Warwick (1994) implicitly resembles a competing risk model of replacement failures. Next, we use competing risk analysis to estimate termination specific hazard rates and show

[4]Clearly this evidence is only impressionistic, given the small sample sizes involved and the influence of outliers on the slope of the line in Figure 1.

that the results of our analysis of the hazard for replacement strongly resemble those obtained by Warwick (1994) as well as King et. al. (1990). These hazard rates are flat and their list of covariates exhibit a significant influence on replacement risks. However, once we apply competing risk analysis to dissolutions, we can show that this termination mode is governed by a rather different stochastic process. Not only is the hazard rate for dissolutions increasing, which confirms the Lupia and Strom model, but the covariates suggested by previous models are less effective in explaining dissolutions rather than replacements.

2. Duration Models and Competing Risk Analysis

Duration or survival analysis refers to a class of statistical techniques used to analyze the duration of events such as the life-span of a cabinet.[5] In a case of competing risks, only the occurrence of the earliest risks will be observed. A government may have been defeated in the next election, but if the prime minister, anticipating the defeat, decided to step down or was toppled in a palace revolt, then the termination due to the electoral risk cannot be observed since the government has already left the sample.

We first give a brief overview of duration analysis and then introduce the competing risk model. Following standard practice we could simply specify the distribution of cabinet duration times using either the conditional density $f(t \mid X,\beta)$ or cumulative distribution function $F(t \mid X,\beta)$ and then maximize the corresponding likelihood function. However, with duration data it is usually more convenient to specify the *hazard rate* $\lambda(t \mid X,\beta)$ instead—the conditional probability that a cabinet terminates in period t given that it has survived so far.

The hazard rate is defined as

$$\lambda(t \mid X,\beta) = \frac{f(t \mid X,\beta)}{1 - F(t \mid X,\beta)} .$$

Defining the *survival function* as $S(t \mid X,\beta) = 1 - F(t \mid X,\beta)$ we can rewrite the previous equation as

$$\lambda(t \mid X,\beta) = \frac{f(t \mid X,\beta)}{S(t \mid X,\beta)} .$$

The survivor function can also be written as the exponential of the negative integrated hazard (Cox and Oakes 1984), or

[5]For a recent overview of duration models in political science see Box-Steffensmeier and Jones (1997). For applications in economics see Kiefer (1988) or Heckman and Singer (1984).

$$S(t \mid X, \beta) = \exp\left[-\int_0^t \lambda(u \mid X, \beta) du \right]. \tag{1}$$

Density function, hazard rate, and survival function thus contain the same statistical information. Which specification we use is simply a matter of mathematical convenience.

In order to investigate the different cabinet termination modes, however, we have to go beyond ordinary duration analysis and use a competing risk framework. Competing risk analysis is a generalized version of duration analysis that investigates multiple modes of termination or "risks."[6] The key idea is to introduce risk-specific hazard rates $\lambda_j(t \mid X_j, \beta_j)$ where $j = 1, \ldots, r$ indicates which risk is analyzed. In our context we will have two risks to consider: dissolutions and replacements. Note that $\lambda_j(t \mid X_j, \beta_j)$ implies risk-specific survival and density functions $S_j(t \mid X_j, \beta_j)$ and $f_j(t \mid X_j, \beta_j)$.

Formally the competing risk model can be specified in terms of a latent variable approach. Let T_j be a (latent) random variable corresponding to the duration until termination due to risk $j (j = 1, \ldots, r)$. Note that of these latent random variables only the smallest can be observed. For example, if we see a cabinet calling an early election we do not observe that it would have been replaced three weeks later.

In the Lupia and Strom model a cabinet's fate is completely determined by the costs and benefits to the ruling coalition and the outside party. Diermeier and Stevenson (1997) show that in a stochastic framework this implies that the modes of cabinet termination are conditionally independent.[7] Thus, if the independent variables are measures of these costs and benefits (and ruling out simultaneous failures which can just be recoded as a new category), the overall hazard, $\lambda(t \mid X, \beta)$, is simply

$$\lambda(t \mid X, \beta) = \sum_{j=1}^{r} \lambda_j(t \mid X_j, \beta_j).$$

We can then specify the likelihood function (e.g., Kalbfleisch and Prentice 1980). Given that all observations have completed their spells, the contribution to the likelihood function of a cabinet i that has failed due to risk j is simply the probability that i fails at any point in time due to risk j, given that it has not previously failed due to any other risk,

$$L_i = f_j(t_i \mid X_{ij}, \beta_j) \prod_{j \neq j'} S_{j'}(t_i \mid X_{ij'}, \beta_{j'}),$$

[6]For an overview of competing risk analysis see, e.g., David and Moeschberger (1978).

[7]We wish to thank an anonymous referee for pointing this out to us.

where the notation $j \neq j'$ simply means to take the product over all risks except j. We can then write

$$L_i = \lambda_j\left(t_i \mid X_{ij}, \beta_j\right) S\left(t_i \mid X_i, \beta\right).$$

Given (1) this can be rewritten as

$$L_i = \lambda_j\left(t_i \mid X_{ij}, \beta_j\right) \prod_{j=1}^{r} \exp\left[-\int_0^{t_i} \lambda_j\left(u \mid X_{ij}, \beta_j\right) du\right].$$

Then for the likelihood function of a sample of size n, we have

$$L = \prod_{i=1}^{n} \lambda_j\left(t_i \mid X_{ij}, \beta_j\right) \prod_{j=1}^{r} \exp\left[-\int_0^{t_i} \lambda_j\left(u \mid X_{ij}, \beta_j\right) du\right].$$

Now let T_j be the observed risk, i.e., $T_j = \min\{T_1, \ldots, T_r\}$, with corresponding failure times t_{ji}, $i = 1, \ldots, n_j$, where n_j is the number of failures due to risk j. Then we can rewrite the likelihood function as

$$L = \prod_{j=1}^{r} \prod_{i=1}^{n_j} \lambda_j\left(t_i \mid X_{ij}, \beta_j\right) \exp\left[-\int_0^{t_i} \lambda_j\left(u \mid X_{ij}, \beta_j\right) du\right]$$

$$L = \prod_{j=1}^{r} \prod_{i=1}^{n_j} \lambda_j\left(t_i \mid X_{ij}, \beta_j\right) S\left(t_i \mid X_{ij}, \beta_j\right).$$

Finally, if we define the following indicator function:

$d_{ij} = 1$ if i failed due to risk j
$d_{ij} = 0$ otherwise

we can write

$$L = \prod_{j=1}^{r} \prod_{i=1}^{n} \lambda_j\left(t_i \mid X_{ij}, \beta_j\right)^{d_{ij}} S\left(t_i \mid X_{ij}, \beta_j\right).$$

Given conditional independence the likelihood function factors into separate risks where failures due to alternative risks are simply treated as randomly censored (e.g., David and Moeschberger 1978; Kalbfleisch and Prentice 1980). That is, in this case competing risk analysis is equivalent to standard duration analysis plus random censoring.

This suggests an illuminating connection to the controversy on censoring in the study of cabinet durations (King et. al. 1990; Warwick and Easton 1992). In most duration analysis, censoring is used to solve a data problem; e.g., some observations may have left the sample before their termination mode can be observed. King et. al. (1990), however, introduced censoring based on theoretical considerations. They regard some cabinet terminations as less theoretically interesting or relevant than others. A large group of censored cabinets consists of cases where governments terminate in the year preceding regularly scheduled elections. While all governments have to stand reelection at some time, the period between mandated elections, the "constitutional inter-election period" (CIEP), varies from three years to five years. Studying kernel density estimates of government durations, King et. al. find a bump in the number of governments that fail about a year prior to the CIEP. King et. al. conjecture that these governments would have survived longer if the end of the CIEP had not neccessitated elections. Thus, they conclude that these failures (called "CIEP failures") should be treated differently from "real" terminations (e.g., those resulting from a vote of no-confidence in the legislature).[8] That is, the theoretically relevant failure times for some cabinets are not observed because these observations have already left the sample due to a previously occurring (and observed) uninteresting type of failure. The authors conclude that these theoretically uninteresting observations should be treated as censored.

Our derivation of the likelihood function in the competing risk case points out that King et. al.'s theoretically motivated censoring regime is equivalent to a competing risk analysis with two independent kinds of failure: "interesting failure" and "uninteresting failure." Indeed, since early dissolutions make up the largest group of censored cases in King et. al.'s study (as well as the various studies that have adopted a similar approach), estimates of these models should resemble competing risk models in which the analyzed risk is "nonelectoral" failure (i.e., where all electoral failures would be censored). Below, we show that this is indeed the case for our analysis of Lupia and Strom's "replacement" failure type.

With respect to dissolutions, Strom and Lupia's other failure type, there is less of a connection with previous efforts. Indeed, previous studies that have censored failures within a year of the CIEP suppress in the sample exactly the part of the failure distribution that, according to Lupia and Strom, will contain most of the dissolutions. In the Lupia and Strom framework there is nothing "theoretically uninteresting" about these failures at the end

[8]Warwick and Easton (1992) criticize some of the rationales for censoring put forward by King et al. (1990) and use a different criterion for censoring, but do not take issue with King et al.'s general approach of theoretically motivated censoring.

of the CIEP, since their theory predicts these failures to occur for the very same reasons and by the same process as dissolutions that occur at any other time in the CIEP. If we take Lupia and Strom's theoretical model as the guide for our empirical model, then we should construct the competing risk analysis of dissolutions by censoring only alternative risks such as replacements and technical failures, but not elections occurring within a year of the CIEP.

However, to demonstrate that our empirical results do not depend on this particular view of censoring we also estimate dissolution failures with "theory-based" censoring. While mainly intended as a robustness check, this analysis also allows us to investigate a substantively important implication of the Lupia and Strom framework. In their model the existence of the CIEP is an integral part of the strategic calculations that parties make. Since parties know that there is a maximum time until an election, they make dissolution decisions *throughout the term of the CIEP* based on this knowledge. Dissolution hazards are thus expected to increase throughout the life of the government and not just at the end of the CIEP. An increasing dissolution hazard rate with CIEP censoring would thus provide additional support to Lupia and Strom's theoretical prediction.

3. Estimation

The data used in our analysis are those in King et. al. (1990) and Warwick (1992b).[9] They include all the post-war cabinets of Austria, Belgium, Canada, Denmark, Finland, France (fourth republic), West Germany, Iceland, Ireland, Israel, Italy, Netherlands, Norway, Sweden, Portugal, Spain, and the United Kingdom.[10] In each case we coded whether a cabinet ended with a chamber dissolution or was replaced by a new cabinet without an intermediate election. Cabinets that reach the very end of the CIEP or fail due to technical reasons are randomly censored.[11]

Our goals are first to estimate the shape of the risk-specific hazard rates and second to identify a set of risk-specific covariates that significantly influence cabinet duration. From a theoretical point of view the most important part of the analysis pertains to the shape of dissolution hazards. Increasing dissolution hazards would add strong support to the strategic approach

[9]These data were originally assembled by Strom (1985). We wish to thank Kaare Strom, Gary King, and Paul Warwick for allowing us to use their data.

[10]Diermeier and Stevenson (1997) argue that Sweden and Norway should not be included in the data set, because the Lupia-Strom model strictly speaking does not apply to them. Here we include them in our analysis to make our results comparable to King et al. (1990) and Warwick (1994).

[11]A technical failure occurs when a government terminates for obviously nonpolitical reasons such as the death of the prime minister. Given conditional independence these randomly censored cabinets can simply be lumped with the alternative risk.

favored by Lupia and Strom (1995). A flat hazard would add support to a nonstrategic view in the tradition of Browne et. al. (1986).

Given these goals our estimation approach needs to be as flexible as possible. We adopt a *semi-parametric* approach that is now standard in the cabinet durations literature (e.g., Warwick 1992c, 1994). That is, while we specify the influence of covariates parametrically, we leave the underlying baseline hazard $\lambda_j^0(t)$ unspecified. Following Cox (1972) we assume that[12]

$$\lambda_j\left(t_i \mid X_j, \beta_j\right) = \lambda_j^0(t)\exp\left(\beta_j X_j\right).$$

Let X_j^1 and X_j^2 be two covariate vectors. Then

$$\frac{\lambda_j\left(t_i \mid X_j^1, \beta_j\right)}{\lambda_j\left(t_i \mid X_j^2, \beta_j\right)} = \exp\left(\beta_j\left(X_j^1 - X_j^2\right)\right).$$

That is, the possibly time-dependent baseline hazards cancel and covariate coefficients can be estimated without having to specify the baseline hazard precisely. Given the implied proportionality of the hazard rates, this approach is usually referred to as the *proportional hazard model.*[13]

Using standard maximum likelihood techniques we can then estimate the marginal effect of risk-specific covariates on the duration of cabinets and provide nonparametric estimates of the underlying baseline risk-specific hazard rates.

This leaves the question of which covariates to include in our analysis. The main purpose of our analysis is to demonstrate that new insights can be gained from a competing risk approach. We thus reanalyze the currently best-supported statisitically model, developed by Warwick (1994) in his recent book. Warwick finds the following empirical regularities.[14]

[12]We also estimated a parametric model based on the Weibull distribution (used by Warwick 1992b) with no qualitative differences.

[13]The assumption of proportional hazard rates is a common one in many applications of duration models including studies of cabinet duration. Although previous studies have examined the assumption statistically using similar data as those used here and have found no evidence of nonproportionality, they have not tested the assumption in the competing risks framework. Consequently, we did our own tests for proportionality for all the estimations which follow. This was done by generating plots of rescaled Schoenfeld residuals vs. survival times and testing them for constancy. This method was developed by Grambsch and Therneau (1994) and is advocated in Box-Steffensmeier and Zorn's (1998) recent review of proportional hazard models in political science. In all fifty-six tests (one for each variable in eight estimated equations), nonproportionality can be clearly rejected. The only variable that may appear dependent on time (and only in two of the eight estimation in which it appears) is the post-election variable (which indeed one might expect to have some temporal quality), but even this slight nonproportionality in the graphs can be rejected statistically.

[14]For operationalization of these variables see Warwick (1992c, 1994).

- *Majority Status:* Minority cabinets have a shorter expected life-span.[15]
- *Post-Election:* Cabinets that formed later in the constitutional inter-election periods tend to have a shorter duration.
- *Investiture:* If a constitution requires an investiture vote before the new cabinet can assume office, cabinets are of a shorter expected duration.
- *Returnability:* The higher the proportion of government parties that return to power after a termination, the shorter the cabinet's duration.
- *Left-Right Diversity:* This measures the government's diversity on a left-right scale using the European Manifestos Project's data (Budge, Robertson, and Hearle 1986). The higher the diversity, the lower the cabinet's life-expectancy.
- *Clerical-Secular Diversity:* This measure is based on Dodd's (1976) party scales. The higher this measure, the lower the cabinet's life-expectancy.
- *Regime Support Diversity:* This measure is also derived from Dodd (1976). Higher values are associated with lower life-spans.

Table 1 summarizes the estimation results.

First, the estimates of the covariates for the pooled model in Table 1 essentially replicate the results reported in Warwick (1994). Second, if one compares the estimates for the covariates in the replacement model with those in these pooled model, one finds they are quite similar as well. All the signs are the same and in only one case, *Post-Election,* does the statistical significance of the variables change. These similarities give credence to our suspicion that, due to the censoring of failures within a year of the CIEP, the pooled models that have been estimated in the literature are essentially models of cabinet replacement.

Turning to the dissolution models, we present estimates for two cases. First, we consider the case without "theory-based" censoring. That is, dissolutions that occur within a year of the end of the CIEP are not censored. Second, for robustness, we also consider the case where CIEP failures are censored. In both cases, dissolutions are not as well-explained by the covariates in the standard models as are replacements. Three variables (*Investiture Vote, Clerical-Secular Diversity,* and *Regime-Support Diversity*) that are significant in the replacement model drop to insignificance in the dissolution models. Moreover, *Post-Election* is more important in explaining dissolution than replacement. Finally, the sign of the *Regime Support Diversity* variable reverses between the models. These differences indicate that, at least with

[15] *Majority Status* is coded zero for minority cabinets and one otherwise.

Table 1. Cox Partial Likelihood Estimates for Combined and Independent Competing Risks*

Covariates	Pooled	Dissolution censoring CIEP failures	Dissolution not censoring CIEP failures	Replacement
Majority Status	−1.41 (6.77)	−1.60 (4.31)	−1.06 (4.36)	−1.36 (5.30)
Post-Election	−0.48 (−2.25)	−1.12 (−3.41)	−2.00 (−8.11)	−.16 (−0.71)
Investiture Vote	0.68 (3.78)	0.23 (0.81)	−0.06 (−0.30)	1.02 (4.23)
Returnability	1.31 (2.64)	1.06 (1.36)	1.10 (2.15)	1.52 (2.33)
Left-Right Diversity	0.27 (3.08)	0.33 (2.18)	0.20 (1.72)	0.24 (2.23)
Clerical-Secular Diversity	0.13 (2.00)	0.13 (1.23)	0.02 (0.23)	0.14 (1.73)
Regime Support Diversity	0.22 (3.55)	−.33 (−1.37)	−.19 (−1.21)	0.30 (4.24)
Log Likelihood	−778	−234	−467	−531
N (number of cases failing due to Risk)	268	268(55)	268 (124)	268 (117)

*T–statistics are in brackets. In the case of dissolutions with CIEP censoring and pooled failures we used Warwick's (1994) censoring regime. The estimates of pooled failures are thus comparable with Warwick (1994). Our estimates are very close to his but not exactly the same because of some differences in the sample.

respect to the influence of covariates, the process that generates cabinet replacements is different from the process producing cabinet dissolutions.[16]

These results are at least *prima facie* consistent with a bargaining approach to the study of cabinet survival where cabinets are viewed as the result of a bargain over policy and office perks. Such bargains are more difficult

[16]For robustness we also estimated the empirical model used in King et al. (1990). See Warwick (1992c, 1994) for a thorough critique of this model. Again we find that the pooled model estimated by King et al. essentially captures cabinet replacements. All signs are the same, and only *Majority Status* ceases to be significant in the case of replacements. In the case of dissolutions three variables important in determining replacements (*Investiture Vote, Polarization,* and the *Effective Number of Parties;* for operationalization of these variables see Alt and King 1984) fail to significantly impact dissolution, while *Majority Status* and the timing variable, *Post-Election,* are much more important in the case of dissolution than replacement.

to sustain in ideologically complex coalitions (these are the "diversity" variables in Table 1). Our estimates suggest that the higher the potential tensions over policy choices, the more likely that a cabinet will terminate in a replacement. Variables such as the post-election indicator, on the other hand, reflect the incentives to strategically call for early elections. Following the logic of the Lupia and Strom model they should be an important determinant of cabinet dissolutions, but less relevant in the case of replacements. This intuition is confirmed by our estimates. Note also that in both cases minority governments are less stable. Some of our findings are more puzzling. For example, why is *Returnabilty* relevant for both kinds of failures and not just for dissolutions?

Figures 2 and 3 provide our estimates of the hazard rates for pooled failures, replacements, and dissolutions. Figure 2 illustrates the case in which failures near the end of the CIEP were not censored, while Figure 3 provides the estimates for the case in which failures within one year of the CIEP are censored.

These figures show that hazard rates for the three classes of failures diverge begining after about 500 days, with the hazard for dissolutions rising more steeply than that for pooled failures and much more steeply than that for replacements, which are essentially flat.[17]

In order to test for whether the hazards slopes are statistically different from zero, we estimated the "slope coefficient" from the Weibull distribution for these hazard rates. If this parameter is less than 1, the estimated hazard is decreasing, if it equals 1 it is constant, and if it is greater than 1, it is increasing.[18] Our estimates for the slope parameters in Figure 2 were 1.35, 1.003, and 2.04 for pooled failures, replacements, and dissolutions, respectively. The estimate for replacements is not statistically different from one ($p = .963$). So, we cannot reject the null hypothesis of a flat replacement hazard. However, the estimates of the slope coefficients for both dissolutions and

[17]Hazard rates must be reported conditional on specified values for the covariates. In order to ensure that the magnitude of the reported hazard rates were comparable across models, the same values for these covariates were used in all the hazard estimates. These were the modal values for the dichotomous variables and the mean values for nondichotomous variables. The modal case with respect to the dichotomous variables was a majority cabinet with no investiture vote that did not form immediately following the election..

[18]The only drawback of using the Weibull distribution to estimate slopes is the assumption of monotonic hazard rates, but looking at all the figures this seems eminently reasonable. Alternatively, we could use estimates of the survival rates from the proportional hazards model (see Kalbfleisch and Prentice 1980). We estimated by both methods, and they were always in agreement. The parameteric versions are reported below. In addition, for each estimation, "log-log" plots were examined to determine if the data fit the Weibull model. In each case the use of the Weibull was justified. These plots, as well as all the other statistical tests mentioned here and elsewhere, are available from the authors.

Figure 2. Pooled Failures (circle), Elections (triangle), and Replacements (square), No Censoring of CIEP Failures

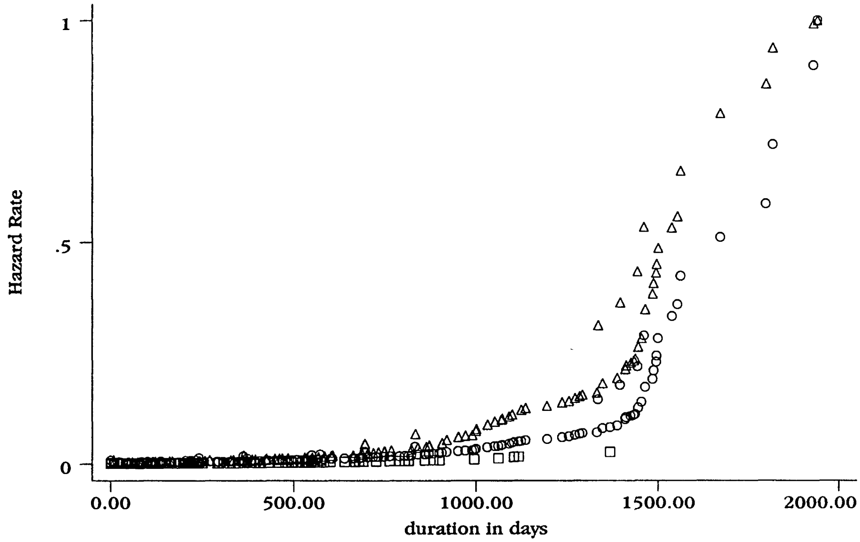
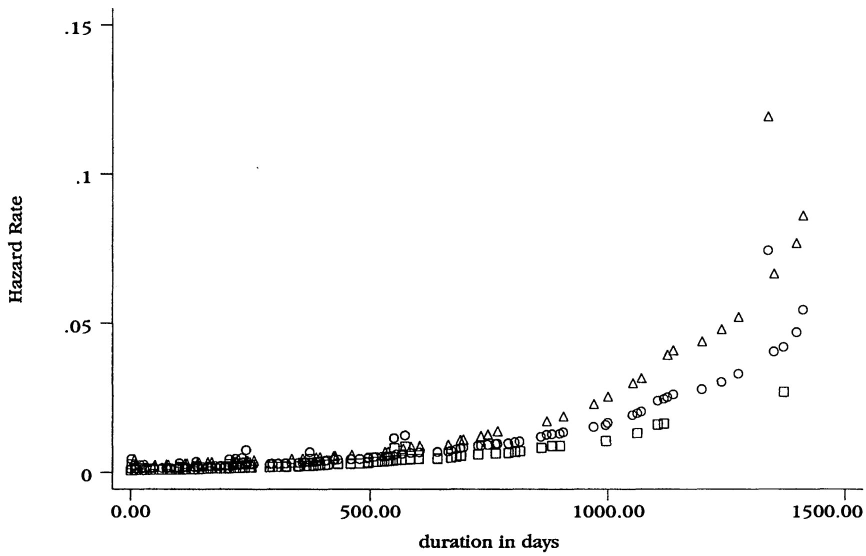


Figure 3. Pooled Failures (circle), Elections (triangle), and Replacements (square), with Censoring of CIEP Failures



pooled failures are significantly greater than 1 (2.04 and $p < .001$ for the case of dissolutions; 1.35 and $p < .001$ for the case of pooled failures). Note that, consistent with our expectations, dissolution hazards increase at a steeper rate than pooled or replacement hazards.

While we still find a divergence in the case with CIEP censoring (Figure 3), it is less pronounced. Turning to our estimates for the slope parameters, however, our substantive findings are robust. The estimates for the Weibull slope parameters in Figure 3 are 1.10, 1.003, and 1.42 for pooled failures, replacements, and dissolutions, respectively. Again for the case of replacements we cannot reject the null hypothesis of flat hazards ($p = .963$), while dissolution hazards are clearly increasing ($p < .001$). In the case of pooled hazards, we now no longer can reject the null hypothesis of flat hazards at conventional significance levels, although this is a close call ($p = .108$).[19] This finding is consistent with the previous literature and reflects the inconclusive emprirical debate over whether pooled hazards are increasing or not. There is some question of whether pooled hazard rates are in fact increasing. But in any case the slope of pooled hazards is rather flat, substantively equivalent to a constant hazard rate.

Fortunately, we need not rely on pooled hazards to assess the empirical merit of the Lupia and Strom model as compared to Browne et. al.'s nonstrategic alternative. Our estimates of dissolution failures provide a more powerful test than pooled or replacement hazards. As mentioned above, replacement hazards cannot be used to test the Lupia and Strom model, since their model is consistent with various replacement hazard shapes. Likewise, pooled hazards will be less conclusive if replacement rates are flat (as our estimates indicate). With respect to dissolutions, however, Lupia and Strom predict an increasing hazard rate, while Browne et. al. suggest this rate should be constant. This implication is confirmed in all our dissolution estimates.

The competing risk analysis thus reveals *two* stochastic processes that *differ* both with respect to the estimated parameters and the shape of the hazard rate. Thus, we conclude that the underlying dynamics of coalition duration are more subtle than indicated by previous single risk models. In particular our results suggest that the results reported by previous models are only relevant for the case of replacements. Since these failures are over represented in the sample, and most of the dissolution terminations are treated as censored in their analysis, they dominate the pooled hazard and produce

[19]For the King et al. model our findings are similar. In the uncensored case we have for pooled hazards an estimate of 1.44 ($p < .001$), for dissolution 2.22 ($p < .001$), and for replacment 1.09 ($p = .19$). In the censored case we find for pooled hazards 1.16 ($p = .009$), for dissolution 1.48 ($p < .001$), and for replacement 1.09 ($p = .19$). The only difference to the Warwick estimates is that now even in the case with CIEP censoring pooled hazards are increasing at statistically significant levels.

results similar to our replacement model. More importantly, however, competing risk analysis reveals a hidden stochastic process related to dissolutions. The dynamics of this process are very different from the dominant one governing replacements and may reflect the strategic decisions of government parties which take into account the likely costs and benefits from calling early elections.

4. Conclusions

The debate about the usefulness of game-theoretic models for the study of cabinet durations occupies center stage in current research on coalition governments. Lupia and Strom (1995) have suggested a bargaining model to explain cabinet terminations. Their analysis suggests that rather than focussing on cabinet survival *per se,* it is important to distinguish between cabinets that end in elections and those that are replaced without elections.

From an empirical point of view this corresponds to a competing risk analysis. Our analysis demonstrates that the stochastic processes that correspond to each failure type are indeed markedly different. Standard models of cabinet termination apply at best to replacement hazards, not to the risk of terminating in an early election. Moreover, while pooled and replacement hazards seem to be flat (or at best slightly increasing), this is not true for dissolution hazards. These hazard rates sharply increase as the next mandatory election approaches but also rise (moderately) when the end of the CIEP is still years away. One implication of this finding is that theory-based censoring (due to King et. al. 1990) essentially corresponds to analyzing replacement hazards only.

We can thus reconcile Alt and King's (1994) finding of constant pooled hazard rates with the Lupia and Strom model. Their analysis inadvertently focused on replacement hazards, while the Lupia and Strom model emphasizes dissolution hazards. Alt and King's analysis thus neither confirms nor disconfirms the Lupia and Strom model. Our finding of increasing dissolution hazards, however, adds empirical support to the strategic approach suggested by Lupia and Strom.

Manuscript submitted October 5, 1998.
Final manuscript received March 30, 1999.

REFERENCES

Alt, James, and Gary King. 1994. "Transfers of Governmental Power: The Meaning of Time Dependence." *Comparative Political Studies* 27:190–210.

Box-Steffensmeier, Janet M., and Bradford S. Jones. 1997. "Time is of the Essence: Event History Models in Political Science." *American Journal of Political Science* 41:1414–1461.

Box-Steffensmeier, Janet M., and Christopher Zorn. 1998. "Duration Models and Proportional Hazard in Political Science" Unpublished manuscript. Political Methodology Paper Archive.

Browne, Eric C., John P. Frendreis, and Dennis W. Gleiber. 1984. "An 'Events' Approach to the Problem of Cabinet Stability." *Comparative Political Studies* 17:167–197.

Browne, Eric C., John P. Frendreis, and Dennis W. Gleiber. 1986. "The Process of Cabinet Dissolution: An Exponential Model of Duration and Stability in Western Democracies." *American Journal of Political Science* 30:628–650.

Browne, Eric C., John P. Frendreis, and Dennis W. Gleiber. 1988. "Contending Models of Cabinet Stability: A Rejoinder." *American Political Science Review* 82:930–941.

Budge, Ian, David Robertson, and Derek Hearl. 1986. *ECPR Party Manifestos Project.* Oslo: Norwegian Social Science Data Archive.

Cox, D. R. 1972."Regression Models and Life-Tables (with discussion)." *Journal of the Royal Statistical Society B* 34:187–220.

Cox, D. R., and D. Oakes. 1984. *Analysis of Survival Data.* London: Chapman and Hall.

David, H. A., and M. L. Moeschberger. 1978. *The Theory of Competing Risks.* New York: Macmillan.

Dodd, Lawrence C. 1976. *Coalitions in Parliamentary Government.* Princeton: Princeton University Press.

Diermeier, Daniel and Randolph Stevenson. 1997. "Cabinet Termination and Critical Events." Unpublished manuscript. Stanford University.

Grambsch, Patricia M., and Terry M. Therneau. 1994. "Proportional Hazards Tests and Diagnostics Based on Weighted Residuals." *Biometrika* 81:515–526.

Heckman, James, and Burton Singer. 1984. "Economic Duration Analysis." *Econometrica* 24:63–132.

Kalbfleisch, J. D., and R. L. Prentice. 1980. *The Statistical Analysis of Failure Time Data.* New York: John Wiley and Sons.

Kiefer, N. 1988. "Economic Duration Data and Hazard Functions." *Journal of Economic Literature* 26:646–679.

King, Gary, James Alt, Nancy Burns, and Michael Laver. 1990. "A Unified Model of Cabinet Dissolution in Parliamentary Democracies." *American Journal of Political Science* 34:846–871.

Lupia Arthur, and Kaare Strom. 1995. "Coalition Termination and the Strategic Timing of Parliamentary Elections" *American Political Science Review* 89:648–669.

Strom, Kaare. 1985. "Party Goals and Government Performance in Parliamentary Democracies." *American Political Science Review* 79:738–754.

Strom, Kaare. 1988. "Contending Models of Cabinet Stability." *American Political Science Review* 82:923–930.

Warwick, Paul. 1992a. "Ideological Diversity and Government Survival in Western European Parliamentary Democracies." *Comparative Political Studies* 25:332–361.

Warwick, Paul. 1992b. "Rising Hazards:An Underlying Dynamic of Parliamentary Government." *American Journal of Political Science* 36:857–876.

Warwick, Paul. 1992c. "Economic Trends and Government Survival in West European Parliamentary Democracies." *American Political Science Review* 86:875–887.

Warwick, Paul. 1994. *Government Survival in Parliamentary Democracies.* Cambridge: Cambridge University Press.

Warwick, Paul, and Stephen T. Easton. 1992. "The Cabinet Stability Controversy: New Perspectives on a Classic Problem." *American Journal of Political Science* 36:122–146.