Estimating Interdependent Duration Models with an Application to Government Formation and Survival*

Jude C. Hays†
University of Illinois at Urbana-Champaign

Aya Kachi‡
University of Illinois at Urbana-Champaign
Princeton University (Visiting)

July 9, 2008

Draft prepared for the 25th (*Silver Edition!*) Society for Political Methodology Summer Conference

Abstract

This paper is part of a larger project in which we develop methods for estimating the causal effects of variables on (1) the duration of bargaining processes, broadly defined, and (2) the survival of bargained outcomes when both are jointly determined. There are many potential applications in political science including, but not limited to, the duration of war and survival of cease-fire agreements, coalition formation and government survival, and negotiations over and enforcement of international agreements. Our primary claim is that, in most cases, it is inappropriate to estimate the effects of variables on these two durations -the bargaining and the outcome- in isolation. Our argument is motivated by game theoretic models that show bargaining duration is correlated with the survival of bargained outcomes because players incorporate their beliefs about the survival of bargained outcomes into their decision-making at the bargaining stage. To address this problem, we develop, and examine the properties of two maximum likelihood estimators–a seemingly unrelated regresssions (SUR) estimator and a limited information maximum likelihood (LIML) estimator. We use both estimators to analyze the duration of government formation and survival in a sample of European parliamentary democracies over the period 1945 to 1998. We conclude that estimated effects based on single equation models of either government formation or survival, the predominant method of analysis in the existing literature, are likely biased because they fail to capture significant indirect effects generated by strategic and other forms of interdependence that link the two durations.

Scholarly interest in the empirical determinants of government formation and dissolution in par-

*The authors thank Christopher H. Achen, Seden Akcinaroglu, Janet Box-Steffensmeier, Damarys Canache, José Cheibub, Tom S. Clark, Songying Fang, Brian J. Gaines, Matt Golder, Sona Golder, Kosuke Imai, James Kuklinski, Matthew Lebo, Benjamin E. Lauderdale, Eduardo L. Leoni, Stephen Meserve, Burt L. Monroe, Kristopher W. Ramsay, Munro Richardson, Jacob Shapiro, Aaron B. Strauss, Milan Svolik, Dustin Tingley, Bonnie Weir and Christopher Zorn for helpful comments. Earlier versions of this paper were presented at the Comparative Politics Workshop at the University of Illinois, the 66th MPSA Annual National Conference, the New Faces in Political Methodology Conference at Pennsylvania State University and the Political Methodology Colloquium at Princeton University.

†Assistant professor, Department of Political Science, University of Illinois. E-mail: jchays@uiuc.edu; URL: https://netfiles.uiuc.edu/jchays/www/page.html.

‡Visiting Student Research Collaborator, Department of Politics, Princeton University, and Ph.D. student, Department of Political Science, University of Illinois. E-mail: akachi@princeton.edu; URL: https://netfiles.uiuc.edu/akachi2/home.

liamentary democracies is longstanding. Two of the most popular topics in this literature are explaining the durations of both coalition bargaining over ministerial portfolios and government survival. There is good reason for this focus. The failure of parliamentary parties to form governments quickly (e.g., the recent crisis in Belgium) and chronic government instability (e.g., Italy for much of the postwar period) have significant social costs and are viewed as symptoms of dysfunctional democracy.

We argue that the lengths of coalition bargaining and government survival are interdependent duration processes. Unfortunately, to this point, the two have been studied largely in isolation. This practice is inefficient at best. We can use information about the relationship between formation and dissolution to get more precise estimates of the determinants of each. At worst, the single equation studies suffer from multiple sources of bias. One potential problem is omitted variable bias in regressions that leave out the important "right-hand-side" duration. Simultaneity is a concern for studies that do connect government formation and dissolution in single equation models by putting variables like crisis duration or the number of cabinet formation attempts on the right-hand-side of government survival regressions. Finally, when government formation and survival are interdependent, single equation models fail to capture the indirect effects of variables that determine both durations. This is troubling given that most of the variables found in empirical studies of government formation also show up in government survival regressions.

Our paper is organized as follows. First, we briefly review the existing empirical duration research on government formation and survival. Our main concern is the failure, particularly in empirical work, to take the implications of interdependence seriously. Second, we discuss some empirical implications of the relevant theoretical studies of bargaining and the breaking of bargaining outcomes. Various sources of interdependence imply that empirical models of government formation and survival should be estimated jointly. Third, we estimate a joint model of government formation and survival using both SUR and LIML estimators. We find evidence of positive duration interdependence, and our results suggest that existing single equation studies produce biased estimates of counterfactual effects. We conclude with a discussion about the ubiquity and importance of interdependent duration processes across political science subfields and mention some possible theoretical and empirical refinements to the models we present.

1 Existing Studies of Coalition Bargaining and Cabinet Survival

Cabinet formation and government survival are among the most central topics in the comparative study of advanced industrial democracies. The quantitative empirical literature is already large and growing rapidly. Typically, the empirical studies explore a set of contextual and cabinet specific factors that determine both kinds of durations. The effects are estimated separately (e.g., for the coalition bargaining duration, Diermeier and Roozendaal 1998; Martin and Stevenson 2001; and for the cabinet survival, Warwick and Easton 1992; King and Alt 1994, and Diermeier and Stevenson 1999). This is not to say that interdependence has been completely ignored. King et al. (1990) and Warwick (1992), among others, put government formation duration, what they call crisis duration, and the number of formation attempts on the right-hand-side of their government survival models.

In the more theoretically oriented literature, Strøm et al. (1994) highlight the importance of cabinet termination and dissolution rules for government formation. Fearon (1998) also formalizes the effects of expected enforcement levels of bargained outcomes on the bargaining stage itself, in the context of international agreements. His formulation suggests that a longer shadow of the future can give states an incentive to bargain harder, delaying agreement in hope of getting a better deal. Diermeier, Eraslan and Merlo (2003) also formalize explicitly the interdependence of government formation bargaining and the bargained outcome -cabinet survival. The main purpose of Diermeier et al. (2003) is to analyze the conditions under which certain types of coalitions are formed. As an empirical matter, their interest lies in estimating the probability that a particular type of coalition is chosen: durations of bargaining and government survival still play important roles in their model, but those durations are not the primary focus of their analysis. In their model, the inefficient delay of bargaining is generated mainly by a stochastic factor, the state of the world that is either favorable or unfavorable for a formed cabinet to survive, while the inefficient delay in Fearon (1998) is mainly due to the dichotomous bargaining choices or uncertainty.

There are fewer theoretical studies of government termination. Laver and Shepsle (1996) stress that the ending of one cabinet begins the formation process for the next and that dissolution and formation are conceptually nonseparable, though their own emphasis is more on the making than breaking of governments. Lupia and Strøm (1995) show that majority governments may dissolve and call early elections when the expected payoff is high enough. Their model explains why a cabinet, which is an "equilibrium" of the earlier bargaining process, might find it worthwhile to terminate its tenure and call an election. All of these studies make important contributions, but fall short of the kind of systematic integration that we see as necessary.

Moreover, the simultaneity problem is obvious from the recursive structures of these models (even if a solution is not!). Bargaining and survival durations are clearly related, but the causal arrow points both ways. If we put one duration on the right hand side of a model explaining the other-as is frequently done in studies of government survival-our estimates will be biased by the reverse causal relationship. The clear empirical implication of these formal models is that we should not estimate coalition bargaining and government survival durations separately or naively put one duration on the right-hand-side of a single-equation regression that has the other duration on the left-hand-side. In the next section, we demonstrate a more appropriate empirical strategy and suggest a new estimator based on a bivariate Weibull distribution.

2 Empirical Strategies for Modeling Interdependent Durations

Our approach to modeling interdependent durations borrows heavily from the literature on estimating systems of equations. There are two major issues in our view. The first is modeling and estimating the covariance structure of disturbances across equations, and the second is identifying and estimating causal relationships between durations that are simultaneously determined. Most readers will be familiar with the basic methods associated with simultaneous equation modeling, but may be less acquainted with the main tool that we use to model covariance structures, the Farlie-Gumbel-Morgenstern copula, which we present in the next section. Before turning to that discussion, however, we briefly note some recent and related work in political science.

Somewhat surprisingly, once we go beyond the extensive endogenous selection and simultaneous equations literatures that employ multivariate normal distributions as the analytic tool of choice, the examples of cross-equation disturbance covariance modeling are relatively few in political science. There are a couple of notable exceptions of course. Boehmke et al. (2006), for example, start with Gumbel's bivariate exponential distribution to model non-random selection when the outcome of interest is a duration. They generalize one of the marginals from Gumbel's distribution, their duration distribution, to be a Weibull and use the marginal exponential distribution as the basis of a discrete choice / selection model. Their substantive application is the effect of leaders' decisions to go to war on their subsequent post-crisis tenure. Boehmke (2006) uses the same bivariate distribution (i.e., a joint Weibull-exponential) to model, as seemingly unrelated regressions (SURs), the timing of issue position taking in Congress with respect to NAFTA and the content of those positions. Quiroz Flores (2006) also uses Gumbel's bivariate exponential distribution to estimate a SUR model of the tenure of chief executive officers and the average tenure of their ministers.

Gumbel's bivariate exponential distribution comes from the copula we discuss next and use to derive a bivariate Weibull distribution. In our discussion, we stress that changing the marginal distributions inserted into the copula changes the properties of the resulting joint distribution. A bivariate Weibull distribution has a different range for the correlation coefficient, for example, than a bivariate exponential. Likewise, a joint Weibull-exponential distribution has its own unique range for the correlation coefficient. For each joint distribution, this and other properties have to be calculated anew from the underlying density and distribution functions. We demonstrate in Appendix 1. Later, we use the bivariate Weibull in an empirical analysis of government formation duration and survival.

2.1 The Bivariate Farlie-Gumbel-Morgenstern Copula

To estimate a SUR model for interdependent durations, we suggest using a bivariate Weibull distribution that, as far as we know, has not been employed in political science. Maximum likelihood (ML) estimation using this distribution offers a possible solution to the problems caused by unobservables (mainly inefficiencies) from which the existing literature may suffer.[1]

What we are suggesting here is, in a sense, a generalization of the Gumbel (Figure 1 and 2). In fact, scholars from other disciplines have attempted to generalize Gumbel to generate potentially useful bivariate Weibulls for the purpose of estimating models of correlated event durations. For example, Johnson, Evans and Green (1990) mention two different types of bivariate Weibulls that were originally derived in Hougaard (1986) and Lu and Bhattacharyya (1990) respectively. However, both of these bivariate Weibulls were generated from the *first* type of bivariate exponential distribution introduced in Gumbel (1960), and this *first* type suffers from a drawback: the correlation (ρ) is restricted to $-.40365 \leq \rho \leq 0$. Accordingly, the generated bivariate Weibulls also have correlations restricted to $0 \leq \rho \leq 1$ (for the Hougaard (1986) one), and $-.20 \leq \rho \leq .32$ (for the Lu and Bhattacharyya (1990) one). Since these distrutions are asymmetric, highly asymmetric in many cases, they are not ideally suited for empirical research in political science.

[Figure 1 ABOUT HERE]

[Figure 2 ABOUT HERE]

[1]Unobservables are much more pernicious in the selection model context because they are a pontential source of endogeneity and bias (see Boehmke et al. 2006).

Instead of generalizing the *first* type of the exponential distribution introduced in Gumbel (1960), we suggest generalizing the *second* type to derive a joint Weibull distribution. This *second* type belongs to the Farlie-Gumbel-Morgenstern family of distributions, and the following general characteristics are proven in Gumbel (1958):

> *A bivariate distribution function can be constructed from two marginal probability functions, $F(y_1)$ and $F(y_2)$, by the following copula with a constraint on the dependence parameter α, $-1 \leq \alpha \leq 1$;*
>
> $$F(y_1, y_2) = F(y_1)F(y_2)[1 + \alpha\{1 - F(y_1)\}\{1 - F(y_2)\}], \tag{1}$$
>
> *where $-1 \leq \alpha \leq 1$. The associated joint density function is given as*
>
> $$f(y_1, y_2) = f(y_1)f(y_2)[1 + \alpha\{2F(y_1) - 1\}\{2F(y_2) - 1\}]. \tag{2}$$

The joint cumulative and the joint density functions of the bivariate Weibull distribution, which can be obtained from the copula (1) and (2), are

$$F(y_1, y_2) = (1 - e^{-(\frac{y_1}{\theta_1})^{\lambda_1}})(1 - e^{-(\frac{y_2}{\theta_2})^{\lambda_2}})(1 + \alpha e^{-(\frac{y_1}{\theta_1})^{\lambda_1} - (\frac{y_2}{\theta_2})^{\lambda_2}})$$
$$f(y_1, y_2) = \frac{\lambda_1}{\theta_1}\frac{\lambda_2}{\theta_2}(\frac{y_1}{\theta_1})^{\lambda_1 - 1}(\frac{y_2}{\theta_2})^{\lambda_2 - 1}e^{-2[(\frac{y_1}{\theta_1})^{\lambda_1} - (\frac{y_2}{\theta_2})^{\lambda_2}]}[4\alpha - 2\alpha e^{(\frac{y_1}{\theta_1})^{\lambda_1}} - 2\alpha e^{(\frac{y_2}{\theta_2})^{\lambda_2}} + (1 + \alpha)e^{(\frac{y_1}{\theta_1})^{\lambda_1} + (\frac{y_2}{\theta_2})^{\lambda_2}}], \tag{3}$$

where $y_1 \geq 0$, $y_2 \geq 0$, $-1 \leq \alpha \leq 1$, $\theta_1 > 0$, $\theta_2 > 0$, $\lambda_1 > 0$ and $\lambda_2 > 0$. Again, α is a dependence parameter, which induces the correlation between y_1 and y_2, and the λs are shape parameters that determine the curvature of the distribution. θs are scale parameters. Note that this becomes the Gumbel (exponential) distribution when $\theta_1 = \theta_2 = 1$ and $\lambda_1 = \lambda_2 = 1$.

We make strong assumptions about the shape of the distributions when we use a bivariate exponential (i.e., a constant hazard rates). The Weibull distribution, by contrast, is much more flexible in the sense that it can take different shapes, including the exponential's when its shape parameters are set to 1. In addition, the bivariate Weibull distribution expands the possible range of ρ to $-.322409 \leq \rho \leq .322409$. Figure 3 plots the correlation ρ against the dependence parameter α when $\lambda_1 = \lambda_2 = \lambda$.[2] The three lines correspond with the cases where $\lambda = 3.29$, $\lambda = 1$ and $\lambda = .5$ respectively.

[Figure 3 ABOUT HERE]

We summarize some of the mathematical properties of this bivariate Weibull in Appendix 1 for those who are interested. These results provides the basis for maximum likelihood estimation.

2.2 A Seemingly Unrelated Regressions (SUR) Model for Interdependent Durations

The SUR model for interdependent durations takes the form

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta}_i + \mathbf{u}_i,\ i = 1, \cdots, M, \tag{4}$$

[2] As shown in Appendix 1, the scale parameter θ does not affect ρ.

where $\mathbf{y}_i$ is a vector of observations on a the dependent variable, $\mathbf{X}_i$ is a matrix of observations on the independent variables, $\boldsymbol{\beta}_i$ is a vector of coefficients, and $\mathbf{u}_i$ is a vector of disturbances, each for equation i. The FGM bivariate Weibull version of the SUR imposes the covariance structure implied by the marginals and copula on the matrix of disturbances $\mathbf{\Sigma}(\alpha)$ and gives the following likelihood for the case of $M = 2$;

$$\begin{aligned} L(y_1, y_2|\mathbf{X}, \boldsymbol{\beta}, \lambda_1, \lambda_2) &= f(y_1, y_2) \\ &= \frac{\lambda_1}{\theta_1}\frac{\lambda_2}{\theta_2}(\frac{y_1}{\theta_1})^{\lambda_1 - 1}(\frac{y_2}{\theta_2})^{\lambda_2 - 1}e^{-2[(\frac{y_1}{\theta_1})^{\lambda_1} - (\frac{y_2}{\theta_2})^{\lambda_2}]} \\ &\times [4\alpha - 2\alpha e^{(\frac{y_1}{\theta_1})^{\lambda_1}} - 2\alpha e^{(\frac{y_2}{\theta_2})^{\lambda_2}} + (1+\alpha)e^{(\frac{y_1}{\theta_1})^{\lambda_1} + (\frac{y_2}{\theta_2})^{\lambda_2}}], \end{aligned} \tag{5}$$

where the θ_i's, the scale parameters, are equal to $\mathbf{X}_i\boldsymbol{\beta_i}$ and the λ_i's are the shape parameter.

2.3 A Simultaneous Equations (SEQ) Model for Interdependent Durations

The main drawback of the SUR model is that it does not address the kinds of causal relationships that we are interested in evaluating–specifically, in our application, the causal effect of government survival on bargaining duration and vice versa. The SUR model allows us to estimate covariances that are consistent with various causal stories, but it will rarely provide us with a "smoking gun." Take Boehmke's (2006) analysis, for example. He finds evidence that the NAFTA positions taken by members of Congress correlate with amount of time it took them to announce their positions. He interprets this as evidence of strategic delay. In other words, the covariance between how long it took a legislator to announce his or her position and the content of the position was driven by the latter. This is certainly plausible, but another possibility is that the amount of time taken to consider the pros and cons of the trade agreement determined the policy position. This might be a popular interpretation among trade economists who tend to believe strongly that support for protectionism simply reflects one's ignorance of the benefits from trade. Academic proponents of free trade may not be too surprised to learn that the opponents of NAFTA disproportionately announced their opposition early into the debate, before giving the issues careful deliberation.

In this section, we derive a limited information maximum likelihood estimator (LIML) for the bivariate Weibull SEQ model and discuss how to get to the full information estimator (FIML). The LIML, an equation by equation estimator, is relatively easy to derive and implement when compared with the FIML, which is neither of these. The only wrinkle with the LIML is that the likelihood is the sum of two extreme value distributions, but this particular convolution, while tricky mathematically, was solved by Gumbel in the late 1940's. For the FIML, we need to obtain the implied covariances, densities, and distribution functions from the reduced-form equations and then impose this structure on the joint density given by the Fairlie-Gumbel-Morgenstern (FGM) copula. In our view, at the present time, the computational costs of the FIML make the LIML an extremely attractive option. We show experimentally via Monte Carlo simulations that this estimator has superior statistical properties to simpler, frequently-used alternatives.

We start by writing the structural version of our model as a set of simultaneous Weibull equations in log-linear form (see Box-Steffensmeier and Jones 2004, p. 26);

$$\begin{bmatrix} \ln y_1 \\ \ln y_2 \end{bmatrix} = \begin{bmatrix} 0 & \alpha_1 \\ \alpha_2 & 0 \end{bmatrix} \begin{bmatrix} \ln y_1 \\ \ln y_2 \end{bmatrix} + \begin{bmatrix} \beta_1 & \beta_2 & 0 \\ \beta_3 & 0 & \beta_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} \lambda_1^{-1} u_1 \\ \lambda_2^{-1} u_2 \end{bmatrix}, \tag{6}$$

where u_1 and u_2 are type-1 extreme value distributions scaled by the inverse of the shape parameters from their respective, underlying Weibull distributions. There is one common exogenous variable in this system of equations and one exclusive variable in each equation. Solving for the reduced form, we have

$$\begin{bmatrix} \ln y_1 \\ \ln y_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{1-\alpha_1\alpha_2} & \frac{\alpha_1}{1-\alpha_1\alpha_2} \\ \frac{\alpha_2}{1-\alpha_1\alpha_2} & \frac{1}{1-\alpha_1\alpha_2} \end{bmatrix} \begin{bmatrix} \beta_1 & \beta_2 & 0 \\ \beta_3 & 0 & \beta_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} \frac{1}{1-\alpha_1\alpha_2} & \frac{\alpha_1}{1-\alpha_1\alpha_2} \\ \frac{\alpha_2}{1-\alpha_1\alpha_2} & \frac{1}{1-\alpha_1\alpha_2} \end{bmatrix} \begin{bmatrix} \lambda_1^{-1} u_1 \\ \lambda_2^{-1} u_2 \end{bmatrix}, \tag{7}$$

or after multiplication

$$\begin{bmatrix} \ln y_1 \\ \ln y_2 \end{bmatrix} = \begin{bmatrix} \left(\frac{\beta_1}{1-\alpha_1\alpha_2} + \frac{\alpha_1\beta_3}{1-\alpha_1\alpha_2}\right) & \frac{\beta_2}{1-\alpha_1\alpha_2} & \left(\frac{\alpha_1\beta_4}{1-\alpha_1\alpha_2}\right) \\ \left(\frac{\alpha_2\beta_1}{1-\alpha_1\alpha_2} + \frac{\beta_3}{1-\alpha_1\alpha_2}\right) & \left(\frac{\alpha_2\beta_2}{1-\alpha_1\alpha_2}\right) & \frac{\beta_4}{1-\alpha_1\alpha_2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} \frac{\lambda_1^{-1}}{1-\alpha_1\alpha_2} u_1 + \frac{\alpha_1\lambda_2^{-1}}{1-\alpha_1\alpha_2} u_2 \\ \frac{\alpha_2\lambda_1^{-1}}{1-\alpha_1\alpha_2} u_1 + \frac{\lambda_2^{-1}}{1-\alpha_1\alpha_2} u_2 \end{bmatrix}, \tag{8}$$

which can be written more compactly as

$$\begin{bmatrix} \ln y_1 \\ \ln y_2 \end{bmatrix} = \mathbf{Gx} + \begin{bmatrix} v_{11} + v_{12} \\ v_{21} + v_{22} \end{bmatrix} = \mathbf{Gx} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \tag{9}$$

where $\mathbf{Gx}$ is the systematic component of the model and $\mathbf{v}$ is the stochastic part containing the reduced-form disturbances. Notice that v_1 and v_2 are weighted sums of two extreme value distributions with means

$$\mu_{v_1} = \frac{\lambda_1^{-1}\gamma}{1-\alpha_1\alpha_2} + \frac{\alpha_1\lambda_2^{-1}\gamma}{1-\alpha_1\alpha_2} \text{and } \mu_{v_2} = \frac{\alpha_2\lambda_1^{-1}\gamma}{1-\alpha_1\alpha_2} + \frac{\lambda_2^{-1}\gamma}{1-\alpha_1\alpha_2}, \tag{10}$$

and variances

$$\sigma_{v_1}^2 = \frac{\pi^2}{6}\left[\left[\frac{\lambda_1^{-1}}{1-\alpha_1\alpha_2}\right]^2 + \left[\frac{\alpha_1\lambda_2^{-1}}{1-\alpha_1\alpha_2}\right]^2\right] \text{and } \sigma_{v_2}^2 = \frac{\pi^2}{6}\left[\left[\frac{\lambda_2^{-1}}{1-\alpha_1\alpha_2}\right]^2 + \left[\frac{\alpha_2\lambda_1^{-1}}{1-\alpha_1\alpha_2}\right]^2\right] \tag{11}$$

and covariance

$$\begin{aligned}\text{cov}\,(v_1, v_2) &= E(v_1 v_2) - E(v_1)E(v_2) \\ &= \left[\frac{\alpha_2}{\lambda_1^2(1-\alpha_1\alpha_2)^2} + \frac{\alpha_1}{\lambda_2^2(1-\alpha_1\alpha_2)^2}\right]\frac{1}{6}\pi^2,\end{aligned} \tag{12}$$

where the parameter γ in (10) is the Euler-Mascheroni constant ($\gamma \approx 0.577$). The univariate distribution function for v_1 is

$$F(v_1) = 2e^{-v_1/2}K_1\left(2e^{-v_1/2}\right), \tag{13}$$

where $K_\alpha(\varphi)$ is a modified Bessel function of the second kind and order α (Gumbel 1947). Taking the derivative gives the density

$$f_{v_1}(v_1) = 2e^{-v_1}K_0\left(2e^{-v_1/2}\right). \tag{14}$$

This is everything we need for the LIML, but before turning to that estimator, it is worth noting the temptation to write down a likelihood function in terms of the structural model in equation (6). Starting with the unobserved stochastic component of the model $\mathbf{u}$, and assuming for the moment that the shape parameters are both 1, we have:

$$L(\mathbf{u}) = \exp\left(-\mathbf{u} - \exp\left(-\mathbf{u}\right)\right). \tag{15}$$

Following this logic, the likelihood for the observed dependent variables would involve the standard transformation by the Jacobian term giving:

$$L(\mathbf{y}) = |\mathbf{I} - \mathbf{A}|\exp\left[-\left[(\mathbf{I} - \mathbf{A}\mathbf{y}) - \mathbf{B}\mathbf{x}\right] - \exp\left(-\left[(\mathbf{I} - \mathbf{A}\mathbf{y}) - \mathbf{B}\mathbf{x}\right]\right)\right]. \tag{16}$$

The problem with this likelihood is that the parameters α_1 and α_2 are not identified unless they are constrained to take the same value. (This, for example, is roughly how the estimation of spatial lag models proceeds (Franzese and Hays 2007, Franzese and Hays 2008). In short, the way we identify, for example, (in the exactly identified case like we have above) is to compare the covariances of x_3, the variable that is exclusive to the equation for y_2, with both y_1 and y_2. This is the method of indirect estimation or indirect lest squares in the familiar linear additive context. It follows that, in order to estimate α_1 and α_2, we need to write the likelihood in terms of the reduced form model. The difficulty, as we can see from equation (8), is that the structure of the disturbances is much more complex -i.e., again both v_1 and v_2 are weighted sums of two extreme value distributions. However, as long as we are willing to ignore the information about the relationship between v_1 and v_2 , we can take Gumbel's convolution results for the two extreme value random variables and proceed with estimation. Note that, since all of the right-hand-side variables in the reduced-form model are exogenous, ignoring this information, while inefficient, does not make the LIML estimator inconsistent. The limited information likelihood is simply

$$\begin{aligned}f(v_1, v_2|\mathbf{x}, \mathbf{G}) &= f(v_1)f(v_2) \\ &= 2e^{-v_1}K_0\left(2e^{-v_1/2}\right) \times 2e^{-v_2}K_0\left(2e^{-v_2/2}\right).\end{aligned} \tag{17}$$

If we are not willing to ignore the relationship between v_1 and v_2 , then we need to take the densities, distributions, and covariance implied by the reduced-form SEQ model and plug these into the FGM copula. The joint density given by the FGM copula is

$$f(v_1, v_2) = f(v_1)f(v_2)\left[1 + \alpha\left\{2F(v_1) - 1\right\}\left\{2F(v_2) - 1\right\}\right]. \tag{18}$$

2.4 Monte Carlo Evaluations of the LIML and Naïve Estimators

Tables 1-8 present the results of several Monte Carlo experiments in which we evaluate the performance of three estimators: the LIML-SEQ estimator, the ML-AEDM estimator, and the ML-AIDM. The last two estimators are what political scientists currently use. ML-AEDM stands for maximum likelihood *assumed* exogenous duration model. This is the standard ML applied to a single equation that has an endogenous duration on the right-hand-side–by standard we mean that the estimator treats the endogenous duration as exogenous. In the context of our application (coalition bargaining duration and government survival), this is what most people are using when they put ”crisis duration”–the number of days to form a government–on the right-hand-side of a government survival model. ML-AIDM stands for ML *assumed* independent durations model. This is when the analyst fails to recognize that his or her duration of interest is linked in important ways to another or multiple other durations. The ML-AEDM suffers from simultaneity bias while the ML-AIDM suffers from omitted variable bias.

The experimental data is generated using the reduced form SEQ model. We are interested in the cases of positive reinforcing interdependence, negative reinforcing interdependence, and mixed interdependence for medium-sized and large samples. The results of our Monte Carlos are not too surprising. The bias of the LIML is much smaller than the biases of the other two estimators. Both ML-AEDM and ML-AIDM can get the sign wrong on average, although, in our experiments, this only happens to the ML-AEDM estimator. This is due to the parameter values we chose to generate the data. To get the ML-AIDM to give the wrong sign on average we need the common variable to have opposite signs in the two durations. We have not yet included this as one of our cases.

The standard error estimates for the LIML are unacceptably overconfident. This is because we report the standard calculation (i.e., negative expected inverse-Hessian) and this calculation ignores the cross-equation correlation in the disturbances. In our application below, we address this by bootstrapping the structural disturbances.

One pattern that is a bit surprising is it seems like the estimators have much more difficulty assessing (accurately) the strength of negative relationships compared with positive ones. We are pretty certain, however, that this is an artifact of the experimental design, in particular, the fact that all our X's take positive values and we have no constants in the model. This means when some of the parameters take negative values the observations are more clustered at the lower boundary (the durations have a lower bound of zero), which makes estimation more difficult than it would be otherwise.

3 Application:
The Determinants of Coalition Bargaining and Cabinet Survival Durations

In this section, we examine the empirical determinants of government formation and survival durations. We start with a description of the explanatory variables, provide some descriptive statistics, and then present the bivariate SUR, LIML-SEQ, ML-AEDM, and ML-AIDM results.

3.1 Description of Explanatory Variables

Table 9 is the list of variables that are (theoretically) relevant for our purposes. Most of these are common in the existing literature. They are conceptualized and measured in the following ways:

[Table 9 ABOUT HERE]

- **Investiture:** Investiture is a dummy variable that indicates whether there is a legal requirement that the parliament approve any government, by vote, before it is invested with power. This could affect both the formation and survival durations. Because a cabinet has to be approved by the legislature, a party might have to form a coalition with other undesirable parties, with which the the party would never form a coalition otherwise. This could be a hurdle that increases the duration of the coalition formation process, and forming a coalition with an undesirable party could cause the government to fail relatively quickly.

- **Continuation:** Continuation is a dummy variable that indicates whether an incumbent government may stay in office, becoming the first proposer of a proto-coalition for the next government, without resigning. We expect that continuation rules shorten the formation process, by allowing the incumbent governments to start negotiating while they are still in office.

- **Number of effective parties:** The number of parties in parliament increases the possible combinations of coalition partners. The conventional wisdom is that larger numbers are associated with longer formation durations. But is it the number of parties itself that makes the bargaining process longer? It is not clear until we take into account the effects of polarization or ideological diversity among parties within parliament. We also expect that there might be a relationship between the number of parties and the survival of governments for the similar reasons.

- **Polarization:** The variable polarization captures the diversity of preferences of the parties in the parliament. It is measured as the percentage of seats held by extremist parties. There are two different ways to think about the effects of polarization. If the extremist parties in the parliament are somehow excluded from the coalition bargaining, for example, the bargaining could proceed relatively quickly among the rest of the parties. Larger share of seats held by extremists, however, could also indicate difficult bargaining if they are included in the bargaining process. We expect that higher polarization is likely to shorten the survival of cabinets, but it could also unify governing coalitions and make them more stable.

- **Returnability:** The variable *returnability* indicates the proportion of parties that stay in the new government after a collapse of the old one. It is measured as an overall proportion of government parties that entered the next government following a collapse or early termination. We expect that the formation process might be short for the high returnability counties. For example, if the same parties return, then they are familiar with each other's preferences, and also the future returnability would reduce the disutility from being in a short-lived cabinet. For the same reason, we expect that returnability would shorten the survival duration.

- **Post-election:** The variable *post-election* indicates whether cabinet negotiations began immediately after an election. We expect that it takes longer to form a cabinet in the post-election cases, because parties are less certain about other parties' preferences. For the same reason, post-election cabinets might also be short-lived. Of course, it is also possible that the first cabinets formed in a new parliament are the most "natural" and stable ones given the electoral results.

- **Caretaker:** The variable *caretaker* indicates whether the coalition is a caretaker government. Caretaker governments are typically in office for shorter periods of time.

- **Survival duration:** The inclusion of the *survival duration* variable in the formation model is based on formal bargaining models. Parties consider the expected length of each of the possible proto-coalitions in the formation process.

- **Previous defeat:** The variable *previous defeat* indicates whether the previous cabinet was defeated by vote of no confidence. Votes of non confidence typically lead to leadership changes and might therefore increase the time it takes for a new government to form.

- **Pre-electoral coalition:** This variable should relevant only for the "post-election" cabinets, which are formed right after elections and not in the middle of electoral terms following government collapses. *pre-electoral coalition* has not been widely used in the study of government formation (yet), but we expect that a pre-electoral coalition agreement would facilitate and shorten the coalition formation process -including the cases where coalition formation is almost immediate. In our model, we treat it as a country-specific factor. It is measured as the proportion of cabinets that had pre-electoral coalition agreements among all the cabinets formed in the given country during the time period under study. This variable exclusively affects the formation process.

- **Pre-electoral coalition×Number of effective parties:** We suspect that the existence of pre-electoral coalition agreements would have more substantial effects when the number of effective party is larger; for example, the effects of pre-electoral coalition would be trivial under a two-party system. This variable also exclusively affects the formation process.

- **Formation duration:** Typical studies of government termination include this variable, *formation duration* to explain the government survival. We also believe that there is a systematic relationship between the formation and the survival duration.

- **Maximum duration:** The variable maximum duration is the time between the beginning of a formed cabinet and the next scheduled election. This is the maximum time that the government could survive.We expect that maximum duration affects the survival duration, in the way that cabinets survive longer, on average, when there is a longer time period left until the next election.

3.2 Data

Our dataset consists of 475 cabinets from sixteen Western European countries -Austria, Belgium, Denmark, Finland, France (Fourth Republic), Germany, Iceland, Ireland, Italy, Luxembourg, The Netherlands, Norway and Sweden. The data run between 1945 and 1998.

3.3 Descriptive Statistics

Interestingly, there is a positive relationship between the time it takes for government formation and the length of government survival in our sample (see Figure 4). Governments that formed in less than fifty days survived, on average, 580 days whereas coalitions that took more than 100 days to reach agreement lasted 818 days. This is a bit perplexing since we might expect long delays in government formation to be indicative of the inability of parliamentary parties to work together effectively (King et al. 1990). There is another way to look at this relationship, however. Parties that anticipate long-lasting governments may bargain harder over coalition agreements since these "contracts" will determine the balance of executive power, distribution of benefits from holding office, and overall course of policy for a significant period of time into the future. Is this relationship spurious or causal? And what implications, if any, does this interdependence have for empirical analyses of government formation and survival durations? We explore these questions with univariate and bivariate regressions in the next section. Descriptive statistics for the covariates organized by country, party structure, and cabinet attributes are provided in Tables 10, 11 and 12 respectively.

[Figure 4 ABOUT HERE]

[Tables 10, 11 and 12 ABOUT HERE]

3.4 Univariate Estimates and Bivariate SUR Results

Recall from Section 2 that the parameters of the model, θ, λ, and α, are constrained: $\theta > 0$, $\lambda > 0$, and $-1 \leq \alpha \leq 1$. For estimation, we work with an unconstrained parameter vector that is mapped into a constrained parameter space using the following link functions:

$\theta = e^{X\beta}$, $\lambda = 100(\frac{1}{1+e^{\kappa}})$, and $\alpha = 2(\frac{1}{1+e^{\gamma}} - 0.5)$,

where X is a matrix of independent variables, β is a vector of coefficients on these variables, and κ and γ are constants.

Table 13 contains estimates for our univariate (single equation) models of government formation and survival. Our univariate results are mostly consistent with those found in existing studies. With respect to the country attributes, we find that continuation rules have a statistically significant and negative effect on formation duration. We find that legislative polarization and ideological diversity increase the time it takes for government formation. We also find that post-election cabinets, caretaker cabinets, and cabinets that follow previously defeated governments form more slowly than those following governments that collapse for other reasons. Table 13 also provides the estimates for our univariate models of government survival duration. In short, our results indicate that investiture, polarization, returnability, and caretaker status all reduce government survival duration while post-election formation, majority status, and maximum possible duration increase survival time.

[Table 13 ABOUT HERE]

[Table 14 ABOUT HERE]

Our bivariate results are given in Table 14. We include the full battery of regressors in each equation, less the endogenous duration variables. Our estimate for the interdependence parameter α is positive and statistically significant. However, the standard error estimates do not suggest large efficiency gains.

3.5 LIML, ML-AEDM, and ML-AIDM Results

These results are very similar to those in the previous section except that, to simplify the LIML estimator, we want to be exactly identified (rather than overidentified). This means that we only have one exclusive variable in each equation. The interesting finding that comes out of our new set of estimates is that the positive correlation between bargaining duration and government survival seems to be driven by the latter causing the former. We interpret this as strong evidence that parties anticipate the length of the future government's tenure and this affects how they bargain. This is the idea of strategic interdependence that comes out of the game theoretic literature on the topic developed Diermeier, Merlo and others. We do not find evidence of the reverse causal relationship. In other words, although there are studies that suggest that the duration of the formation processes have implications on the government survival, we find that this is not the case. These theories maintain that longer bargaining indicates the difficulty in reaching agreements among the coalition members in general and hence shorter lifetime of the formed government. The SUR model does not allow the analyst to easily distinguish these two possibilities empirically.

4 Conclusion

We conclude with a discussion about the ubiquity and importance of interdependent duration processes across political science subfields and mention some possible theoretical and empirical refinements to the models we present. First, starting with refinements, we hope to generalize the Farlie-Gumbel-Morgenstern family into a multivariate case. This would be useful for political science studies, for example, where there is a strategic dependence among the decisions taken by multiple actors, such as states. Second, the bivariate Weibull distribution presented here is based on the copula proven in Gumbel (1959). This family of bivariate distribution allows ρ to range only $|\rho| \leq \frac{1}{3}$ maximum (Long and Krzysztofowicz 1995). There is a further generalized version of the Farlie-Gumbel-Morgenstern, which is characterized by the following joint density (Long and Krzysztofowicz 1995). (We do not know what people call this family of distributions.)

$$h(y_1, y_2) = f(y_1)g(y_2)[1 + \gamma c(F(y_1), G(y_2))]$$

This is basically relaxing the restriction on the functional form of $c(F(y_1), G(y_2))$. The tradeoff is that the relaxation of restrictions could expand the rage of ρ, but once we deviate from the Farlie-Gumbel-Morgenstern form of bivariate distribution, the derivation for the constraints on the parameters become extremely cumbersome, and imposing such complicated constraints may make the estimation difficult.

Turning to theory, although the bargaining model developed by Diermeier et. al (2003) is mathematically rigorous and conceptually intuitive, there is one story missing. The *cake size* of their bargaining model is the expected government duration that is to be allocated among the proto-coalition members, and the cake size is determined by a stochastic factor, the state of the world. However, many studies suggest that parties are not only interested in staying in government. Strøm (1990), for example, mentions some cases where a party could decide not to join a coalition (even if the formateur offers the party to join the coalition), if the coalition's collective preference on certain policy issue would differ from the party's ideal position and if the party has some expectation that their policy stance is going to be closer to the public's ideal point in the near future. In this case, the party has an incentive to wait until the next election and win more seats and become the formateur to attempt to implement more preferable policies. We would argue that the closeness of a party's policy stance to the public's ideal point could be another type of *cake*, and the public opinion on certain issues can follow some stochastic sequence.

There is another possible way in which the policy stance of the parties could affect the coalition bargaining. Suppose the coalition bargaining is over specific positions within the cabinet, such as the minister of finance and the minister of defense, and the number of posts a coalition party could get is determined by the number of the seats they gained in the election. For example, if a party could gain only one post in the coalition, then the party might bargain over exactly which post it will gain. However, the problem is that the party's preference over the different posts within the cabinet could be determined by the issues that become salient in the near future. This uncertainty about the salience of the issue could be another stochastic factor. The *cake* in this story is the posts within the cabinet the parties gain.

One of the strengths of the bargaining model developed by Diermeier et al. (2003) is that it is applicable to a number of phenomena that interest political scientists. In the area of international cooperation, for example, there is an extensive literature on the bargaining process and the consequent contents of international agreements. At the same time, the issue of when and why states comply is another important topic. Scholars are aware that the compliance level —or time until the agreement is violated by one of the members— is endogenous to the bargaining process. The modified model we suggested above becomes more attractive when we shift our focus to the bargaining process where states care about the content as well as the compliance level of the agreements.

Another example is the relationship between the constitutional design process and the lifespan of constitutions. According to Elkins, Ginsburg and Melton 2007, for example, bargaining over the contents of newly written constitutions is likely to be correlated with the lifespan. For this case, the content of the constitutions is the central focus of the bargaining.

Appendix 1:
Mathematical properties of the bivariate Weibull distribution

In the following, we briefly describe some mathematical properties of the bivariate Weibull distribution presented above. Although it was already proven that any bivariate distribution that belongs to the Farlie- Gumbel-Morgenstern family satisfies the axioms of probability (Gumbel (1958); Gumbel (1960)), the following mathematical properties that lead to the derivation of ρ are useful.

As shown in Gumbel (1958) and Gumbel (1960), a bivariate distribution function can be constructed from two marginal probability functions, $F(y_1)$ and $F(y_2)$, by the following copula with a constraint on parameter α, $-1 \leq \alpha \leq 1$;

$$F(y_1, y_2) = F(y_1)F(y_2)[1 + \alpha\{1 - F(y_1)\}\{1 - F(y_2)\}], \tag{19}$$

where $-1 \leq \alpha \leq 1$.

The associated joint density function is given as follows;

$$f(y_1, y_2) = f(y_1)f(y_2)[1 + \alpha\{2F(y_1) - 1\}\{2F(y_2) - 1\}]. \tag{20}$$

Using the above findings, a bivariate Weibull distribution can be constructed from the following marginal distributions and densities;

$$F(y_i) = 1 - e^{-(\frac{y_i}{\theta_i})^{\lambda_i}},$$

$$f(y_i) = \frac{\lambda_i}{\theta_i}(\frac{y_i}{\theta_i})^{\lambda_i - 1} e^{-(\frac{y_i}{\theta_i})^{\lambda_i}}; i = 1, 2,$$

where $\lambda_i > 0$ and $\theta_i > 0$.

The joint probability and density are given by

$$F(y_1, y_2) = (1 - e^{-(\frac{y_1}{\theta_1})^{\lambda_1}})(1 - e^{-(\frac{y_2}{\theta_2})^{\lambda_2}})(1 + \alpha e^{-(\frac{y_1}{\theta_1})^{\lambda_1} - (\frac{y_2}{\theta_2})^{\lambda_2}}) \tag{21}$$

$$f(y_1, y_2) = \frac{\lambda_1}{\theta_1}\frac{\lambda_2}{\theta_2}(\frac{y_1}{\theta_1})^{\lambda_1 - 1}(\frac{y_2}{\theta_2})^{\lambda_2 - 1} e^{-2[(\frac{y_1}{\theta_1})^{\lambda_1} - (\frac{y_2}{\theta_2})^{\lambda_2}]}[4\alpha - 2\alpha e^{(\frac{y_1}{\theta_1})^{\lambda_1}} - 2\alpha e^{(\frac{y_2}{\theta_2})^{\lambda_2}} + (1+\alpha) e^{(\frac{y_1}{\theta_1})^{\lambda_1} + (\frac{y_2}{\theta_2})^{\lambda_2}}], \tag{22}$$

where $y_i \geq 0$, $-1 \leq \alpha \leq 1$, $\theta_i > 0$ and $\lambda_i \geq 0$.

The joint cdf must satisfy the following boundary conditions

$$\begin{cases} F(0, y_2) = F(y_1, 0) = 0 \\ F(\infty, \infty) = 1, \end{cases}$$

and the joint density has to be nonnegative, $f(y_1, y_2) \geq 0$. Any bivariate distribution that belongs to the Farlie-Gumbel-Morgenstern family satisfies these conditions.

Another condition a bivariate distribution always has to satisfy is Fréchet's inequality,

$$F(y_1, y_2) \leq F_1(y_i); i = 1, 2 \tag{23}$$

for all y_1 and y_2. Since the dependence of the joint distribution and density on y_1 and y_2 is symmetric, it is sufficient to explore the performance of one variable y_1. This applies to all the calculations in the rest of this appendix. From (21) it follows after a simplification that

$$\alpha e^{-(\frac{y_1}{\theta_1})^{\lambda_1}}(1-e^{-(\frac{y_2}{\theta_2})^{\lambda_2}}) \le 1. \tag{24}$$

In sum, the function (21) satisfies all the required axioms of probability function, under the conditions of $-1 \le \alpha \le 1$, $\theta_i > 0$ and $\lambda_i \ge 0$.

The followings are the relevant computations to derive the correlation coefficient ρ. By definition, the correlation coefficient of two random variables, Y_1 and Y_2 can be obtained as

$$\rho = \frac{E(y_1 y_2) - E(y_1)E(y_2)}{\sigma_{y_1}\sigma_{y_2}}. \tag{25}$$

For our marginal probabilities, $F(y_1)$ and $F(y_2)$, the means and variances are

$$\begin{aligned} E(y_i) &= \theta_i \Gamma(1+\frac{1}{\lambda_i}) = \theta_i \frac{1}{\lambda_i}\Gamma(\frac{1}{\lambda_i}) \\ Var(y_i) &= \theta_i^{\frac{2}{\lambda_i}}[\Gamma(1+\frac{2}{\lambda_i}) - \Gamma^2(1+\frac{1}{\lambda_i})]; \quad i=1,2. \end{aligned} \tag{26}$$

Now the only term we need to compute to obtain ρ is $E(y_1 y_2)$. From (15), the marginal densities are

$$f(y_i) = \int_0^\infty f(y_1,y_2)dy_i = \frac{\lambda_i}{\theta_i}(\frac{y_i}{\theta_i})^{\lambda_i - 1} e^{-(\frac{y_i}{\theta_i})^{\lambda_i}}; i=1,2, \tag{27}$$

which is, of course, the Weibull densities.

The conditional expectation of y_1 can be obtained as

$$\begin{aligned} E(y_1|y_2) &= \int_0^\infty y_1 f(y_1|y_2)dy_1 \\ &= \frac{1}{\lambda_1}\theta_1\Gamma(\frac{1}{\lambda_1})2^{-\frac{1}{\lambda_1}}e^{-(\frac{y_2}{\theta_2})^{\lambda_2}}[-2\alpha(2^{\frac{1}{\lambda_1}}-1) + e^{(\frac{y_2}{\theta_2})^{\lambda_2}}(2^{\frac{1}{\lambda_1}}(1+\alpha)-\alpha)], \end{aligned} \tag{28}$$

where

$$f(y_1|y_2) = \frac{f(y_1,y_2)}{f(y_2)}. \tag{29}$$

The expectation of the cross-product can be computed as

$$\begin{aligned} E(y_1 y_2) &= \int_0^\infty y_2 E(y_1|y_2) f(y_2) dy_2 \\ &= \frac{\theta_1}{\lambda_1}\frac{\theta_2}{\lambda_2}\Gamma(\frac{1}{\lambda_1})\Gamma(\frac{2}{\lambda_2})\big[1+\alpha(1-2^{-\frac{1}{\lambda_1}} - 2^{-\frac{1}{\lambda_2}} + 2^{-\frac{\lambda_1+\lambda_2}{\lambda_1\lambda_2}})\big]. \end{aligned} \tag{30}$$

Substituting (30) into (25), we get

$$\rho = \frac{2^{-\frac{\lambda1+\lambda2}{\lambda1\lambda2}}\left(2^{\frac{1}{\lambda1}}-1\right)\left(2^{\frac{1}{\lambda2}}-1\right)\alpha\Gamma\left[\frac{1}{\lambda1}\right]\Gamma\left[\frac{1}{\lambda2}\right]}{\lambda1\lambda2\sqrt{-\Gamma^2\left[1+\frac{1}{\lambda1}\right]+\Gamma\left[\frac{2+\lambda1}{\lambda1}\right]}\sqrt{-\Gamma^2\left[1+\frac{1}{\lambda2}\right]+\Gamma\left[\frac{2+\lambda2}{\lambda2}\right]}}. \tag{31}$$

Note that the scale parameter θ does not affect the dependence of y_1 and y_2, ρ. As mentioned in the previous sections, ρ is increasing in α and the maximum range is $-.322409 \leq \rho \leq .322409$ when $\lambda_1 = \lambda_2 = 3.29035$.

Appendix 2:
Log of the Weibull is type-I-extreme-value distributed

In the derivation of the Weibull FIML estimator, we claimed that logged weibull random variables are extreme-value distributed (type I) with the scale parameter being the inverse of the Weibull shape parameter.

Recall the density and distribution functions of the type-I extreme value distribution and the Weibull distribution;

$$\text{Extreme value} \begin{cases} f(u) = \frac{1}{b} e^{\frac{u}{b}} e^{-e^{\frac{u}{b}}} \\ F(u) = 1 - e^{-e^{\frac{u}{b}}}, \end{cases}$$
$$\text{Weibull} \begin{cases} f(y) = \frac{\lambda}{\theta} (\frac{y}{\theta})^{\lambda-1} e^{-(\frac{y}{\theta})^{\lambda}} \\ F(y) = 1 - e^{-(\frac{y}{\theta})^{\lambda}}, \end{cases}$$

where b and θ are the scale parameters and λ is the shape parameter.

Consider a random variable X scaled by θ and set $U = \ln(\frac{X}{\theta})$. The extreme value cdf can be transformed as follows;

$$\begin{aligned} F(u) &= F(\ln(\frac{x}{\theta})) \\ &= 1 - e^{-e^{\frac{\ln(\frac{x}{\theta})}{b}}} \\ &= 1 - e^{-(\frac{x}{\theta})^{\frac{1}{b}}}. \end{aligned} \tag{32}$$

Set $\lambda = \frac{1}{b}$, and we can see that X is Weibull-distributed with a shape parameter λ and a scale parameter θ. Also note that the scale parameter of the extreme value distribution, b, is the inverse of λ, the shape parameter of the Weibull.

References

Boehmke, Frederick J., Daniel S. Morey, and Megan Shannon. 2006. "Selection Bias and Continuous- Time Duration Models: Consequences and a Proposed Solution." *American Journal of Political Science* 50(1):192-207.

Boehmke, Frederick J. 2006. "The Influence of Unobservable Factors on Position Timing and Content in the NAFTA Vote." *Political Analysis* 14(4):421-438.

Box-Steffensmeier, Janet M. and Bradford Jones. 2004. *Event History Modeling* Cambridge Universityt Press.

Diermeier, Daniel, Hülya Eraslan and Antonio Merlo. 2002. "Coalition Government and Comparative Constitutional Design." *European Economic Review* 46:893-907.

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

Diermeier, Daniel and Peter van Roozendaal. 1998. "The Duration of Cabinet Formation Process in Western Multi-Party Democracies." *British Journal of Political Science* 28:609-626.

Diermeier, Daniel and Randy T. Stevenson. 1999. "Cabinet Survival and Competing Risks." *American Journal of Political Science* 43(4):1051-1068.

Diermeier, Daniel and Randy T. Stevenson. 2000. "Cabinet Termination and Critical Events." *American Political Science Review* 94(3):627-640.

Druckman, James and Michael F. Thies. 2002. "The Importance of Concurrence: The Impact of Bicameralism on Government Formation and Duration." *American Journal of Political Science* 46(4):760-771.

Elkins, Zachary, Thomas Ginsburg and James Melton. 2007. "The Lifespan of Written Constitutions." *Unpublished manuscript* University of Illinois at Urbana-Champaign, Urbana, Illinois

Fearon, James D. 1998. "Bargaining, Enforcement, and International Cooperation." *International Organization* 52(2):269-305.

Franzese, Jr., Robert J. and Jude C. Hays. 2007. "Spatial-Econometric Models of Cross-Sectional Interdependence in Political-Science Panel and TSCS Data." *Political Analysis* 15(2):140-164.

Franzese, Jr., Robert J. and Jude C. Hays. 2008. "Interdependence in Comparative Politics: Substance, Theory, Empirics, Substance." *Comparative Political Studies* 41(4/5):742-780.

Grofman, Bernard and Peter van Roozendaal. 1997. "Review Article: Modelling Cabinet Durability and Termination." *British Journal of Political Science* 27:419-451.

Gumbel, E.J. 1947. "The Distribution of the Range." *The Annals of Mathematical Statistics* 18(3):384-412.

Gumbel, E.J. 1959. "Multivariate Distributions with Given Margins." *Revista da Faculdade de Ciencias* 7(2):179-218.

Gumbel, E.J. 1960. "Bivariate Exponential Distributions." *Journal of the American Statistical Association* 55(292):698-707.

Hougaard, Philip. 1986. "A Class of Multivariate Failure Time Distributions." *Biometrika* 73:671-678

Johnson, Richard A., James W. Evans and David W. Green. 1990. "Some Bivariate Distributions for Modeling the Strength Proporties of Lumber." *Research Paper Forest Service*, United States Department of Agriculture: FPL-RP-575.

King, Gary, James E. Alt, Nancy Elizabeth Burns and Michael Laver. 1990. "A Unified Model of Cabinet Dissolution in Parliamentary Democracies." *American Journal of Political Science* 34(3):846-871.

King, Gary and James Alt. 1994. "Transfers of Governmental Power: The Meaning of Time Dependence." *Comparative Political Studies* 27(2):190-210

Laver, Michael. 1998. "Models of Government Formation." *Annual Review of Political Science* 1:1-25.

Laver, Michael. 2003. "Government Termination." *Annual Review of Political Science* 6:23-40.

Laver, Michael and Kenneth A. Shepsle. 1996. *Making and Breaking Governments: Cabinets and Legislatures in Parliamentary Democracies* Cambridge University Press.

Long, Dou and Roman Krzysztofowicz. 1995. "A Family of Bivariate Densities Constructed from Marginals." *Journal of the American Statistical Association* 90(430):739-746.

Lu, Jye-Chyi and Gouri K. Bhattacharyya. 1990. "Some New Constructions of Bivariate Weibull Models." *Annals of Institute of Statistical Mathematics* 42(3):543-559.

Lupia, Arthur and Kaare Strøm. 1995. "Coalition Termination and the Strategic Timing of Parliamentary Elections." *American political Science Review* 89(3):648-665.

Martin, Lanny ad Randy T. Stevenson. 2001. "Government Formation in Parliamentary Democracies." *American Journal of Political Science* 45(1):33-50.

Martin, Lanny ad George Vanberg. 2003. "Wasting Time? The Impact of Ideology and Size on Delay in Coalition Formation." *British Journal of Political Science* 33:323-344.

Quiroz Flores, Alejandro. 2006. "The Power of Remove as Political Survival: Simultaneous Equations in Survival Models." *Working Paper* Department of Politics, New York University. New York, NY.

Roozendaal, Peter van. 1997. "Government Survival in Western Multi-party Democracies." *European Journal of Political Research* 32:71-92.

Shipan, Charles R. and Craig Volden. 2006. "Bottom-Up Federalism: The Diffusion of Antismoking Policies from U.S. Cities to States." *American Journal of Political Science* 50(4):825-843.

Strøm, Kaare. 1990. *Minority Government and Majority Rule Cambridge* Cambridge University Press.

Strøm, Kaare, Ian Budge and Michael J. Laver. 1994. "Constraints on Cabinet Formation in Parliamentary Democracies." *American Journal of Political Science* 38(2):303-335.

Warwick, Paul V. 1992. "Rising Hazards: An Underlying Dynamic of Parliamentary Government." *American Journal of Political Science* 36(4):857-76.

Warwick, Paul V. 1992. "Economic Trends and Government Survival in West European Parliamentary Democracies." *American Political Science Review* 86(4):875-87.

Warwick, Paul V and Stephen T. Easton. 1992. "The Cabinet Stability Controversy: New Pwespectives on a Classic Problem." *American Journal of Political Science* 36(1):122-46.

5 Tables and Figures

Table 1: Monte Carlo Simulation Results (1): Sample Size=540, Trials=100

	Equation 1 (Formation)			**Equation 2 (Survival)**			**Interdependence**	
	Common (β_1)	Exclusive (β_2)	Instrument	Common (β_3)	Exclusive (β_4)	Instrument	$Y_2 \rightarrow Y_1$ (α_1)	$Y_1 \rightarrow Y_2$ (α_2)
Reduced Form								
True Parameters	2.00	1.33	0.67	2.00	0.67	1.33		
Estimated Parameters (LIML)	1.87	1.20	0.54	1.84	0.55	1.23		
True SD (LIML)	0.18	0.18	0.18	0.20	0.18	0.19		
Estimated SE (LIML)$^{(c)}$	0.07	0.07	0.07	0.07	0.07	0.07		
Structural Form								
True Parameters	**1.00**	**1.00**		**1.00**	**1.00**		**0.50**	**0.50**
Estimated Parameters (LIML)	1.07	0.96		0.98	0.98		0.44	0.46
Estimated Parameters (AEDM)$^{(a)}$	0.78	0.99		0.77	0.99		0.66	0.66
Estimated Parameters (AIDM)$^{(b)}$	2.94	2.28		2.94	2.27		NA	NA
True SD (AEDM)	0.02	0.07		0.05	0.02		0.05	0.08
Estimated SE (AEDM)$^{(c)}$	0.02	0.06		0.05	0.02		0.05	0.06
True SD (AIDM)	0.10	0.09		0.10	0.10		NA	NA
Estimated SE (AIDM)$^{(c)}$	0.04	0.04		0.04	0.04		NA	NA

(a) AEDM = Assumed Endogenous Duration Model

(b) AIDM = Assumed Independent Durations Model

(c) The estimated standard errors for the reduced form parameters in the *LIML* models were calculated as the negative expected inverse Hessian. The estimated standard errors for the parameters in *AEDM* and *AIDM* models were calculated using the ML standard error formula.

Table 2: Monte Carlo Simulation Results (2): Sample Size=540, Trials=100

	Equation 1 (Formation)			Equation 2 (Survival)			Interdependence	
	Common (β_1)	Exclusive (β_2)	Instrument	Common (β_3)	Exclusive (β_4)	Instrument	$Y_2 \to Y_1$ (α_1)	$Y_1 \to Y_2$ (α_2)
Reduced Form								
True Parameters	1.20	0.80	0.40	0.40	0.80	-0.40		
Estimated Parameters (LIML)	1.14	0.75	0.33	0.30	0.69	-0.52		
True SD (LIML)	0.08	0.08	0.07	0.09	0.08	0.09		
Estimated SE (LIML)$^{(c)}$	0.07	0.07	0.07	0.07	0.07	0.07		
Structural Form								
True Parameters	**1.00**	**1.00**		**1.00**	**1.00**		**0.50**	**-0.50**
Estimated Parameters (LIML)	1.00	1.00		1.12	0.92		0.48	-0.71
Estimated Parameters (AEDM)$^{(a)}$	0.35	1.35		1.24	-0.23		0.77	0.98
Estimated Parameters (AIDM)$^{(b)}$	1.66	1.26		0.39	0.80		NA	NA
True SD (LIML)$^{(d)}$	0.10	0.08		0.26	0.13		0.11	0.16
True SD (AEDM)	0.04	0.04	0.04	0.04			0.08	0.05
Estimated SE (AEDM)$^{(c)}$	0.03	0.05	0.04	0.03			0.06	0.05
True SD (AIDM)	0.04	0.04		0.06	0.06		NA	NA
Estimated SE (AIDM)$^{(c)}$	0.04	0.04		0.04	0.04		NA	NA

(a) AEDM = Assumed Endogenous Duration Model

(b) AIDM = Assumed Independent Durations Model

(c) The estimated standard errors for the reduced form parameters in the *LIML* models were calculated as the negative expected inverse Hessian. The estimated standard errors for the parameters in *AEDM* and *AIDM* models were calculated using the ML standard error formula.

(d) We have not estimated the structural form standard errors for LIML yet.

Table 3: Monte Carlo Simulation Results (3): Sample Size=540, Trials=100

	Equation 1 (Formation)			**Equation 2 (Survival)**			**Interdependence**	
	Common (β_1)	Exclusive (β_2)	Instrument	Common (β_3)	Exclusive (β_4)	Instrument	$Y_2 \rightarrow Y_1$ (α_1)	$Y_1 \rightarrow Y_2$ (α_2)
Reduced Form								
True Parameters	0.67	1.33	-0.67	0.67	1.33	-0.67		
Estimated Parameters (LIML)	0.46	1.14	-0.90	0.41	1.11	-0.84		
True SD (LIML)	0.18	0.18	0.19	0.17	0.21	0.20		
Estimated SE (LIML)[c]	0.07	0.07	0.07	0.07	0.07	0.07		
Structural Form								
True Parameters	**1.00**	**1.00**		**1.00**	**1.00**		**-0.50**	**-0.50**
Estimated Parameters (LIML)	0.81	0.46		0.75	0.45		-0.85	-0.75
Estimated Parameters (AEDM)[a]	-0.59	1.36		1.23	-0.59		1.35	1.24
Estimated Parameters (AIDM)[b]	0.81	1.43		0.76	1.46		NA	NA
True SD (LIML)[d]	0.28	0.20		0.26	0.20		0.26	0.21
True SD (AEDM)	0.02	0.04		0.03	0.02		0.05	0.04
Estimated SE (AEDM)[c]	0.02	0.05		0.04	0.02		0.05	0.04
True SD (AIDM)	0.12	0.12		0.13	0.13		NA	NA
Estimated SE (AIDM)[c]	0.04	0.04		0.04	0.04		NA	NA

(a) AEDM = Assumed Endogenous Duration Model

(b) AIDM = Assumed Independent Durations Model

(c) The estimated standard errors for the reduced form parameters in the *LIML* models were calculated as the negative expected inverse Hessian. The estimated standard errors for the parameters in *AEDM* and *AIDM* models were calculated using the ML standard error formula.

(d) We have not estimated the structural form standard errors for LIML yet.

Table 4: Monte Carlo Simulation Results (4): Sample Size=2700, Trials=100

	Equation 1 (Formation)			**Equation 2 (Survival)**			**Interdependence**	
	Common (β_1)	Exclusive (β_2)	Instrument	Common (β_3)	Exclusive (β_4)	Instrument	$Y_2 \rightarrow Y_1$ (α_1)	$Y_1 \rightarrow Y_2$ (α_2)
Reduced Form								
True Parameters	2.00	1.33	0.67	2.00	1.33	0.67		
Estimated Parameters (LIML)	1.86	1.21	0.55	1.87	1.21	0.53		
True SD (LIML)	0.08	0.08	0.09	0.08	0.08	0.09		
Estimated SE (LIML)$^{(c)}$	0.03	0.03	0.03	0.03	0.03	0.03		
Structural Form								
True Parameters	**1.00**	**1.00**		**1.00**	**1.00**		**0.50**	**0.50**
Estimated Parameters (LIML)	1.01	0.97		1.05	0.97		0.46	0.44
Estimated Parameters (AEDM)$^{(a)}$	0.66	0.78		0.99	0.66		0.78	0.99
Estimated Parameters (AIDM)$^{(b)}$	2.94	2.28		2.94	2.27		NA	NA
True SD (LIML)$^{(d)}$	0.16	0.09		0.18	0.09		0.06	0.07
True SD (AEDM)	0.01	0.03		0.02	0.01		0.04	0.02
Estimated SE (AEDM)$^{(c)}$	0.01	0.03		0.02	0.01		0.03	0.02
True SD (AIDM)	0.04	0.04		0.04	0.04		NA	NA
Estimated SE (AIDM)$^{(c)}$	0.02	0.02		0.02	0.02		NA	NA

(a) AEDM = Assumed Endogenous Duration Model

(b) AIDM = Assumed Independent Durations Model

(c) The estimated standard errors for the reduced form parameters in the *LIML* models were calculated as the negative expected inverse Hessian. The estimated standard errors for the parameters in *AEDM* and *AIDM* models were calculated using the ML standard error formula.

(d) We have not done the estimation of the structural form standard errors for LIML yet.

Table 5: Monte Carlo Simulation Results (5): Sample Size=2700, Trials=100

	Equation 1 (Formation)			**Equation 2 (Survival)**			**Interdependence**	
	Common (β_1)	Exclusive (β_2)	Instrument	Common (β_3)	Exclusive (β_4)	Instrument	$Y_2 \rightarrow Y_1$ (α_1)	$Y_1 \rightarrow Y_2$ (α_2)
Reduced Form								
True Parameters	1.20	0.80	0.40	0.40	0.80	-0.40		
Estimated Parameters (LIML)	1.14	0.74	0.33	0.29	0.69	-0.52		
True SD (LIML)	0.04	0.04	0.04	0.04	0.04	0.04		
Estimated SE (LIML)$^{(c)}$	0.03	0.03	0.03	0.03	0.03	0.03		
Structural Form								
True Parameters	**1.00**	**1.00**		**1.00**	**1.00**		**0.50**	**-0.50**
Estimated Parameters (LIML)	1.00	0.99		1.08	0.92		0.49	-0.70
Estimated Parameters (AEDM)$^{(a)}$	0.35	1.36		1.23	-0.24		0.78	0.99
Estimated Parameters (AIDM)$^{(b)}$	1.66	1.26		0.39	0.80		NA	NA
True SD (LIML)$^{(d)}$	0.05	0.03		0.11	0.06		0.06	0.07
True SD (AEDM)	0.02	0.02		0.02	0.01		0.03	0.02
Estimated SE (AEDM)$^{(c)}$	0.01	0.02		0.02	0.01		0.03	0.02
True SD (AIDM)	0.02	0.02		0.03	0.02		NA	NA
Estimated SE (AIDM)$^{(c)}$	0.02	0.02		0.02	0.02		NA	NA

(a) AEDM = Assumed Endogenous Duration Model

(b) AIDM = Assumed Independent Durations Model

(c) The estimated standard errors for the reduced form parameters in the *LIML* models were calculated as the negative expected inverse Hessian. The estimated standard errors for the parameters in *AEDM* and *AIDM* models were calculated using the ML standard error formula.

(d) We have not done the estimation of the structural form standard errors for LIML yet.

Table 6: Monte Carlo Simulation Results (6): Sample Size=2700, Trials=100

	Equation 1 (Formation)			Equation 2 (Survival)			Interdependence	
	Common (β_1)	Exclusive (β_2)	Instrument	Common (β_3)	Exclusive (β_4)	Instrument	$Y_2 \rightarrow Y_1$ (α_1)	$Y_1 \rightarrow Y_2$ (α_2)
Reduced Form								
True Parameters	0.67	1.33	-0.67	0.67	1.33	-0.67		
Estimated Parameters (LIML)	0.44	1.12	-0.87	0.45	1.12	-0.88		
True SD (LIML)	0.08	0.07	0.08	0.09	0.10	0.09		
Estimated SE (LIML)$^{(c)}$	0.03	0.03	0.03	0.03	0.03	0.03		
Structural Form								
True Parameters	**1.00**	**1.00**		**1.00**	**1.00**		**-0.50**	**-0.50**
Estimated Parameters (LIML)	0.79	0.43		0.80	0.44		-0.78	-0.79
Estimated Parameters (AEDM)$^{(a)}$	-0.59	1.35		1.24	-0.59		1.36	1.23
Estimated Parameters (AIDM)$^{(b)}$	0.78	1.45		0.79	1.44		NA	NA
True SD (LIML)$^{(d)}$	0.13	0.09		0.12	0.10		0.09	0.09
True SD (AEDM)	0.01	0.02		0.02	0.01		0.02	0.02
Estimated SE (AEDM)$^{(c)}$	0.01	0.02		0.02	0.01		0.02	0.02
True SD (AIDM)	0.06	0.05		0.05	0.06		NA	NA
Estimated SE (AIDM)$^{(c)}$	0.02	0.02		0.02	0.02		NA	NA

(a) AEDM = Assumed Endogenous Duration Model

(b) AIDM = Assumed Independent Durations Model

(c) The estimated standard errors for the reduced form parameters in the *LIML* models were calculated as the negative expected inverse Hessian. The estimated standard errors for the parameters in *AEDM* and *AIDM* models were calculated using the ML standard error formula.

(d) We have not done the estimation of the structural form standard errors for LIML yet.

Table 7: Monte Carlo Simulation Results (7): Sample Size=540, Trials=100

	Equation 1 (Formation)			Equation 2 (Survival)			Interdependence	
	Common (β_1)	Exclusive (β_2)	Instrument	Common (β_3)	Exclusive (β_4)	Instrument	$Y_2 \rightarrow Y_1$ (α_1)	$Y_1 \rightarrow Y_2$ (α_2)
Reduced Form								
True Parameters	-2.00	-1.33	-0.67	-2.00	-0.67	-1.33		
Estimated Parameters (LIML)	-2.16	-1.44	-0.78	-2.16	-0.74	-1.47		
True SD (LIML)	0.16	0.18	0.18	0.17	0.18	0.20		
Estimated SE (LIML)$^{(c)}$	0.07	0.07	0.07	0.07	0.07	0.07		
Structural Form								
True Parameters	**-1.00**	**-1.00**		**-1.00**		**-1.00**	**0.50**	**0.50**
Estimated Parameters (LIML)	-1.01	-1.05		-1.05		-1.07	0.53	0.51
Estimated Parameters (AEDM)$^{(a)}$	0.61	-0.51		-0.66		0.60	-0.52	-0.65
Estimated Parameters (AIDM)$^{(b)}$	-1.80	-1.12		-1.81		-1.11	NA	NA
True SD (LIML)$^{(d)}$	0.33	0.20		0.33		0.21	0.10	0.10
True SD (AEDM)	0.02	0.06		0.04		0.02	0.06	0.05
Estimated SE (AEDM)$^{(c)}$	0.02	0.06		0.04		0.02	0.06	0.04
True SD (AIDM)	0.08	0.07		0.07		0.08	NA	NA
Estimated SE (AIDM)$^{(c)}$	0.04	0.04		0.04		0.04	NA	NA

(a) AEDM = Assumed Endogenous Duration Model

(b) AIDM = Assumed Independent Durations Model

(c) The estimated standard errors for the reduced form parameters in the *LIML* models were calculated as the negative expected inverse Hessian. The estimated standard errors for the parameters in *AEDM* and *AIDM* models were calculated using the ML standard error formula.

(d) We have not done the estimation of the structural form standard errors for LIML yet.

Table 8: Monte Carlo Simulation Results (8): Sample Size=540, Trials=100

	Equation 1 (Formation)			**Equation 2 (Survival)**			**Interdependence**	
	Common (β_1)	Exclusive (β_2)	Instrument	Common (β_3)	Exclusive (β_4)	Instrument	$Y_2 \rightarrow Y_1$ (α_1)	$Y_1 \rightarrow Y_2$ (α_2)
Reduced Form								
True Parameters	-0.67	-1.33	0.67	-0.67	-1.33	0.67		
Estimated Parameters (LIML)	-0.88	-1.56	0.46	-0.90	-1.51	0.45		
True SD (LIML)	0.18	0.17	0.18	0.21	0.19	0.19		
Estimated SE (LIML)$^{(c)}$	0.07	0.07	0.07	0.07	0.07	0.07		
Structural Form								
True Parameters	**-1.00**	**-1.00**		**-1.00**	**-1.00**		**-0.50**	**-0.50**
Estimated Parameters (LIML)	-1.16	-1.42		-1.15	-1.38		-0.31	-0.29
Estimated Parameters (AEDM)$^{(a)}$	-0.69	-1.02		-0.75	-0.69		-1.02	-0.75
Estimated Parameters (AIDM)$^{(b)}$	0.13	-0.52		0.14	-0.51		NA	NA
True SD (LIML)$^{(d)}$	0.24	0.15		0.27	0.17		0.12	0.12
True SD (AEDM)	0.02	0.05		0.04	0.02		0.05	0.04
Estimated SE (AEDM)$^{(c)}$	0.02	0.05		0.04	0.02		0.05	0.04
True SD (AIDM)	0.14	0.15		0.15	0.15		NA	NA
Estimated SE (AIDM)$^{(c)}$	0.04	0.04		0.04	0.04		NA	NA

(a) AEDM = Assumed Endogenous Duration Model

(b) AIDM = Assumed Independent Durations Model

(c) The estimated standard errors for the reduced form parameters in the *LIML* models were calculated as the negative expected inverse Hessian. The estimated standard errors for the parameters in *AEDM* and *AIDM* models were calculated using the ML standard error formula.

(d) We have not done the estimation of the structural form standard errors for LIML yet.

Table 9: Variables That Might Affect Each Duration

	Coalition Formation Duration		Government Survival Duration
Country attributes:			
	Investiture		Investiture
*	Continuation		
*	Pre-election coalition		
*	Pre-election coalition×Effective parties		
Party structure attributes:			
	Effective parties		Effective parties
	Polarization		Polarization
	Returnability		Returnability
	Ideological diversity		Ideological diversity
Cabinet attributes:			
	Post-election		Post-election
	Caretaker		Caretaker
	Survival duration		Formation duration
*	Previous defeat	*	Maximum duration
		*	Majority status

* indicates exclusive variables.

Figure 1: Weibull Margins: Varying the Scale Parameter

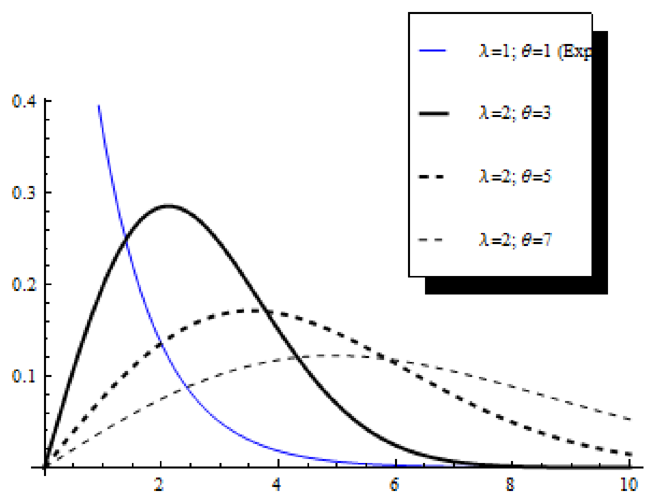

Table 10: Descriptive Statistics of the Variables Included in the Model -by Country (1)

	Durations		*Country attributes*		
	Formation [avg days]	Survival [avg days]	Investiture	Continuation	Pre-electoral coalition [%]
Austria	46 (47)	835 (460)	No	No	71 (45)
Belgium	50 (51)	543 (492)	Yes	No	59 (49)
Denmark	18 (16)	614 (350)	No	Yes	33 (47)
Finland	31 (38)	393 (378)	No	No	14 (35)
France	14 (12)	142 (131)	Yes	No	71 (45)
Iceland	42 (45)	811 (511)	No	No	44 (50)
Ireland	22 (16)	887 (495)	Yes	No	50 (50)
Italy	33 (31)	270 (215)	Yes	No	31 (46)
Luxembourg	33 (25)	1124 (649)	No	No	33 (47)
Netherlands	74 (57)	752 (528)	No	No	38 (49)
Norway	11 (12)	758 (401)	No	Yes	62 (49)
Portugal	38 (42)	485 (481)	Yes	No	78 (41)
Spain	37 (31)	1041 (312)	Yes	No	100 (0)
Sweden	10 (10)	740 (434)	N to Y	Yes	41 (49)
United Kingdom	6 (10)	911 (543)	No	No	14 (35)
Germany	30 (28)	717 (510)	Yes	No	93 (26)

Data source: Warwick (1994), Golder (2005), Keesing's World News Archives

Figure 2: Weibull Margins: Varying the Shape Parameter

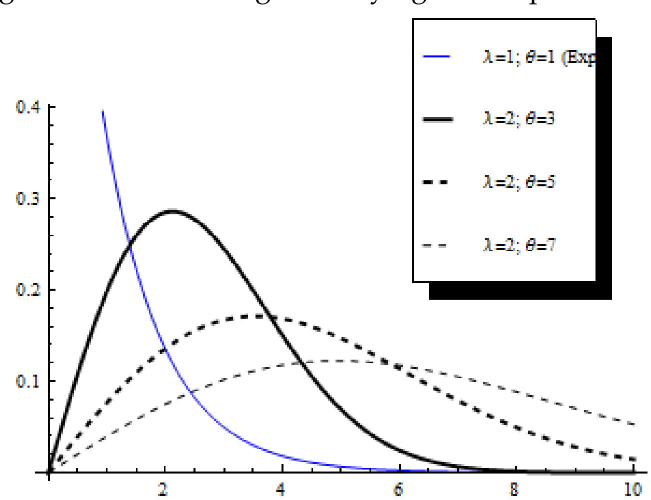

Table 11: Descriptive Statistics of the Variables Included in the Model -by Country (2) (contd.)

	Party structure attributes				
	Effective parties [Mean]	Polarization [%]	Returnability [%]		Ideological diversity
Austria	2.51 (0.50)	6.0 (9.4)	88	(32)	1.5
Belgium	4.86 (2.12)	12.0 (6.5)	79	(41)	3
Denmark	4.48 (0.78)	15.1 (7.6)	69	(46)	2.5
Finland	5.03 (0.30)	21.6 (3.8)	64	(48)	3.5
France	6.39 (1.23)	32.2 (3.2)	78	(41)	2.5
Iceland	3.82 (0.64)	16.4 (2.9)	42	(49)	2
Ireland	2.87 (0.37)	5.6 (22.1)	17	(38)	1.5
Italy	3.95 (1.04)	32.6 (7.6)	82	(38)	3
Luxembourg	3.37 (0.45)	6.1 (3.6)	67	(47)	2
Netherlands	4.66 (0.80)	6.6 (4.5)	72	(45)	3
Norway	3.48 (0.63)	7.5 (8.8)	24	(43)	3
Portugal	3.46 (0.72)	14.1 (4.3)	67	(47)	2.5
Spain	2.68 (0.19)	5.2 (2.6)	100	(0)	2.5
Sweden	3.40 (0.38)	5.0 (2.7)	60	(49)	2.5
United Kingdom	2.13 (0.10)	0.7 (0.9)	67	(47)	1.5
Germany	2.73 (0.49)	0.2 (1.0)	82	(38)	3

Data source: Warwick (1994), Golder (2005), Keesing's World News Archives

Table 12: Descriptive Statistics of the Variables Included in the Model -by Country (3) (contd.)

	Cabinet attributes				
	Post-election [Mean %]	Caretaker [Mean %]	Previous defeat [Mean %]	Maximum survival [days]	Majority status [Mean %]
Austria	72 (46)	0 (0)	4 (20)	1198 (349)	96 (20)
Belgium	50 (51)	6 (23)	11 (32)	1195 (308)	89 (32)
Denmark	70 (47)	0 (0)	30 (47)	1297 (290)	15 (36)
Finland	31 (47)	13 (34)	11 (31)	831 (485)	71 (46)
France	11 (31)	4 (19)	57 (50)	1138 (460)	57 (50)
Iceland	71 (46)	8 (28)	4 (20)	1245 (336)	83 (38)
Ireland	77 (43)	0 (0)	14 (35)	1601 (349)	50 (51)
Italy	20 (40)	7 (26)	14 (35)	1148 (486)	56 (50)
Luxembourg	72 (46)	0 (0)	6 (24)	1388 (568)	100 (0)
Netherlands	69 (47)	10 (31)	34 (48)	1230 (503)	90 (31)
Norway	55 (51)	0 (0)	7 (26)	1107 (420)	31 (47)
Portugal	50 (51)	25 (44)	20 (41)	1100 (430)	65 (49)
Spain	88 (35)	0 (0)	0 (0)	1326 (241)	63 (52)
Sweden	70 (47)	0 (0)	4 (19)	1125 (413)	26 (45)
United Kingdom	71 (46)	0 (0)	8 (28)	1485 (528)	79 (41)
Germany	58 (50)	4 (20)	4 (20)	1103 (419)	100 (0)

Data source: Warwick (1994), Golder (2005), Keesing's World News Archive

Table 13: Univariate Weibull Results: The Results When Government Formation and Survival Duration Are Estimated Separately

Formation duration ($\mathbf{y}_1$)			Cabinet survival ($\mathbf{y}_2$)		
Variable	**Coefficient**	**(S.E.)**	**Variable**	**Coefficient**	**(S.E.)**
$\boldsymbol{\theta}_1$ (Scale parameter 1)			$\boldsymbol{\theta}_2$ (Scale parameter 2)		
Country attributes:			***Country attributes:***		
Investiture	-0.064	(0.108)	**Investiture**	-0.179**	(0.069)
Continuation	-0.965**	(0.133)			
Pre-electoral coalition	-0.174	(0.524)			
Pre-electoral coalition ×Effective parties	0.251†	(0.141)			
Party structure attributes:			***Party structure attributes:***		
Effective parties	0.004	(0.055)	**Effective parties**	-0.027	(0.027)
Polarization	0.921*	(0.490)	**Polarization**	-1.579**	(0.256)
Returnability	0.300	(0.275)	**Returnability**	-0.351*	(0.174)
Ideological diversity	0.247**	(0.095)	**Ideological diversity**	-0.025	(0.063)
Cabinet attributes:			***Cabinet attributes:***		
Post-election	0.855**	(0.205)	**Post-election**	0.315**	(0.089)
Caretaker	0.274	(0.215)	**Caretaker**	-1.042**	(0.148)
Previous defeat	0.366**	(0.135)	**Maximum duration**	0.456**	(0.104)
			Majority status	0.175*	(0.071)
Constant	1.728**	(0.313)	**Constant**	6.280**	(0.222)
$\boldsymbol{\lambda}_1$ (Shape parameter 1)			$\boldsymbol{\lambda}_2$ (Shape parameter 2)		
Constant	0.992**	(0.037)	**Constant**	1.456**	(0.055)
Log-likelihood (Sum): -5386.36					
$N = 475$					

Data source: Warwick (1994), Golder (2005), Keesing's World News Archive

Significance levels : † : 10% $*$: 5% $**$: 1%

Figure 3: Correlation versus dependence plot for BVW ($\lambda_1 = \lambda_2 = \lambda = 3.29, 1, 0.5$)

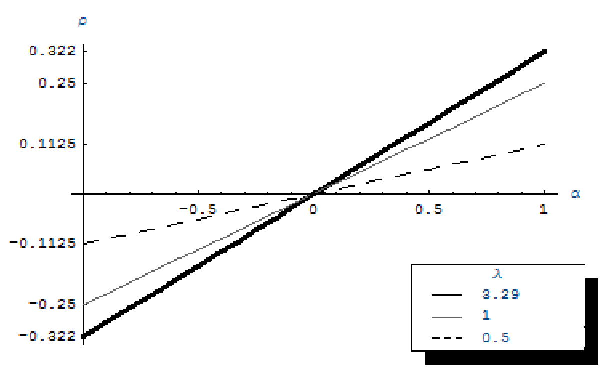

Table 14: Bivariate Weibull Results for Government Formation and Survival Duration (SUR)

Formation duration ($\mathbf{y}_1$)			Cabinet survival ($\mathbf{y}_2$)		
Variable	**Coefficient**	**(S.E.)**	**Variable**	**Coefficient**	**(S.E.)**
$\boldsymbol{\theta}_1$ (Scale parameter 1)			$\boldsymbol{\theta}_2$ (Scale parameter 2)		
Country attributes:			***Country attributes:***		
Investiture	-0.050	(0.108)	**Investiture**	-0.187**	(0.069)
Continuation	-0.975**	(0.133)			
Pre-electoral coalition	-0.177	(0.519)			
Pre-electoral coalition ×Effective parties	0.227	(0.140)			
Party structure attributes:			***Party structure attributes:***		
Effective parties	0.010	(0.056)	**Effective parties**	-0.029	(0.027)
Polarization	0.987*	(0.493)	**Polarization**	-1.600**	(0.257)
Returnability	0.299	(0.274)	**Returnability**	-0.345*	(0.175)
Ideological diversity	0.240*	(0.095)	**Ideological diversity**	-0.032	(0.063)
Cabinet attributes:			***Cabinet attributes:***		
Post-election	0.896**	(0.206)	**Post-election**	0.310**	(0.089)
Caretaker	0.259	(0.216)	**Caretaker**	-1.036**	(0.148)
Previous defeat	0.381**	(0.135)	**Maximum duration**	0.452**	(0.103)
			Majority status	0.174*	(0.071)
Constant	1.705**	(0.313)	**Constant**	6.313**	(0.222)
$\boldsymbol{\lambda}_1$ (Shape parameter 1)			$\boldsymbol{\lambda}_2$ (Shape parameter 2)		
Constant	0.990**	(0.0367)	**Constant**	1.454**	(0.055)
$\tilde{\boldsymbol{\alpha}}$					
Constant	0.259*	(0.134)			
Log-likelihood: -5384.49					
$N = 475$					

Data source: Warwick (1994), Golder (2005), Keesing's World News Archive

Significance levels : $\dagger$: 10% $*$: 5% $**$: 1%

(a) The parameters of the model -θ, λ, and α- are constrained: $\theta > 1$, $\lambda > 0$, and $-1 \leq \alpha \leq 1$. However, for estimation, we work with an unconstrained parameter vector that is mapped into a constrained parameter space using link functions. We report both the constrained α and unconstrained $\tilde{\alpha}$ parameter estimates. Note that both estimates depend on the number of attempts variable. The link function is $\alpha = 2(\frac{1}{1+e^{-X\gamma}} - 0.5)$, where, in general, $\tilde{\alpha}$ is $X\gamma$, X is a vector of ones, and γ is a vector of coefficients on the variables included in α (in this case, there is only the constant).

Table 15: Estimation Results for the Government Formation and Survival Duration Using the LIML, AEDM and AIDM Models with Exponential Hazards

	LIML Reduced Form	LIML Structural Form	AEDM	AIDM
DV: Formation Duration				
Constant	-1.537**	-4.450	-2.069**	-1.217**
	(0.534, 0.532)		(0.680)	(0.47)
Investiture	0.282†	0.604	0.366*	0.312†
	(0.165,0.118)		(0.166)	(0.163)
Continuation	-0.672**	-0.656	-0.681**	-0.676**
	(0.212,0.211)		(0.212)	(0.212)
Maximum Duration	0.270			
	(0.217,0.204)			
Effective Parties	0.208**	0.292	0.217**	0.206**
	(0.063,0.072)		(0.063)	(0.063)
Polarization	0.694	2.355	0.998	0.810
	(0.685,0.857)		(0.672)	(0.671)
Ideological Diversity	0.358*	0.284	0.333*	0.318*
	(0.144,0.284)		(0.140)	(0.140)
Returnability	0.708	0.944	0.756	0.702
	(0.467,0.944)		(0.466)	(0.467)
Post-Election	0.961**	0.774	1.050**	1.127**
	(0.214,0.774)		(0.174)	(0.169)
Caretaker	-0.012	0.449	0.112	-0.034
	(0.326,0.449)		(0.337)	(0.325)
Government Survival		0.517*	0.122†	
		(0.211, —)	(0.071)	
DV: Government Survival				
Constant	5.637**	5.708	5.636**	5.623**
	(0.521,0.327)		(0.509)	(0.512)
Investiture	-0.623**	-0.636	-0.630**	-0.618**
	(0.167,0.183)		(0.164)	(0.164)
Continuation	-0.031			
	(0.216,0.191)			
Maximum Duration	0.522*	0.510	0.512*	0.523*
	(0.217,0.155)		(0.217)	(0.217)
Effective Parties	-0.162*	-0.172	-0.170**	-0.162*
	(0.067,0.045)		(0.067)	(0.067)
Polarization	-3.214**	-3.246	-3.333**	-3.206**
	(0.785,0.558)		(0.787)	(0.783)
Ideological Diversity	0.143	0.127	0.122	0.140
	(0.139,0.114)		(0.138)	(0.137)
Returnability	-0.457	-0.489	-0.499	-0.435
	(0.439,0.228)		(0.412)	(0.412)
Post-Election	0.361†	0.317	0.277	0.359†
	(0.213,0.204)		(0.222)	(0.213)
Caretaker	-0.893**	-0.892	-0.893**	-0.887**
	(0.326,0.245)		(0.324)	(0.324)
Formation Duration		0.046	0.073	
		(0.103, —)	(0.054)	

Variables in blue are instrumental (or the equation-specific) variables. For the LIML reduced form standard errors, we have not corrected for the cross-equation covariance structure of the disturbances. The bootstrapped standard errors (The second entry of the parentheses below the coefficients) adjusts for the cross-equation covariance structure of the disturbances.

Figure 4: Government Formation and Duration

Time for Formation	Average Survival	Standard Deviation	Minimum	Maximum
Less than 50 days	580	481	1	1818
Between 50 and 100 days	649	515	10	1941
More than 100 days	818	529	36	1616

Data source: Warwick (1994), Golder (2005), Keesing's World News Archive