Systematically Dependent Competing Risks and Strategic Retirement *

American Journal of Political Science 53 (3)

Kentaro Fukumoto †

March 11, 2009

Abstract

In many applications of survival analysis, the risk of an event occurring for one reason is dependent on the risk of the same event occurring for another reason. For example, when politicians suspect they might lose an election, they may strategically choose to retire. In such situations, the often-used multinomial logit model suffers from bias and underestimates the degree of strategic retirement, for example, to what extent poor prior electoral performance diminishes electoral prospects. To address this problem, the present paper proposes a systematically dependent competing risks (SDCR) model of survival analysis. Unlike the frailty model, the SDCR model can

 Earlier versions of this paper were presented at the 63rd annual national conference of the Midwest Political Science Association, Chicago, IL, USA, April 7-10, 2005, the 22nd Annual Summer Meeting of the Society for Political Methodology, Tallahassee, FL, USA, July 21-23, 2005 and workshops at Harvard University, April and May, 2005. Thanks to Janet M. Box-Steffensmeier and Bradford S. Jones for making their data public, Barry C. Burden, Leif-Eric Easley, Andy Eggers, Ben Goodrich, Mike Kellermann, Jeffrey Lazarus and three anonymous referees for their comments and help. This research is supported by Japan Society for the Promotion of Science.

†Kentaro Fukumoto is professor of political science, Gakushuin University, 1-5-1 Mejiro, Toshima, Tokyo, 171-8588 Japan (Kentaro.Fukumoto@gakushuin.ac.jp).

also deal with more than two risks. Monte Carlo simulation demonstrates how much the SDCR model reduces bias. Reanalysis of data on U.S. congressional careers (Box-Steffensmeier and Jones, 2004) documents the strategic retirement of representatives, indicating that electoral pressure is more effective at turning out incumbents than previously recognized.

Running Head: Systematically Dependent Competing Risks

Keywords: survival analysis, duration analysis, event history analysis, competing risks, multinomial logit, frailty model, Congress, strategic retirement, incumbents.

INTRODUCTION

Survival analysis (also known as duration analysis or event history analysis) has become a popular tool for analyzing whether and when an event happens (Box-Steffensmeier and Jones, 2004). For example, what determines whether and when a U.S. House member leaves office? Survival analysis answers both questions at the same time, essentially treating members who have not left as "censored."

In the case of House members leaving office, it is clear that there are a number of distinct ways in which a congressional career may end. In an often-cited work on survival analysis, Box-Steffensmeier and Jones (1997) argue that representatives leave the House for four reasons (i.e. competing risks): ambition (running for other public offices such as senator or governor), retirement (declining to run for some other reason), electoral defeat in a primary election, or electoral defeat in the general election. As a way of estimating these competing risks, they use the multinomial logit (hereafter, MNL) model to demonstrate that some explanatory factors (e.g., whether a state's Senate seat has opened up) may raise the possibility that a lawmaker exits for one reason (such as ambition) but not for the others (such as retirement or election competition). This competing risks approach provides a finer-grained analysis of *how* a particular event occurs, not just *when* and in association with what factors.

Here is the problem: most studies (implicitly) assume that competing risks are independent, though this assumption is sometimes not reasonable. For example, lawmakers who suspect they might lose a reelection bid might strategically choose to retire in order to avoid a costly and ultimately futile electoral campaign (Cox and Katz, 2002; Jacobson and Kernell, 1983). Thus, the actual rate of incumbent defeat is a deceptively low indicator of the effect of general election success on career longevity. In such situations, the MNL model suffers from bias and underestimates the degree of strategic retirement, for example, the extent to which a poor prior vote margin makes electoral prospects worse. A conventional method to address the problem of dependence among competing risks is the frailty model, which can be called a *stochastically* dependent competing risks model as explained below. Unfortunately, it is difficult for the frailty model to identify more than two risks due to computational burden (Gordon, 2002).

This paper proposes a *systematically* dependent competing risks (hereafter, SDCR) model, which is robust against bias and can model any number of risks. In a nutshell, the model assumes that a set of independent variables affects one hazard (say, retirement from the House) in the same way that it affects others (e.g., defeat in primary and general elections) and estimates all hazards simultaneously. This approach enables the model to estimate accurately the interrelationship of hazards.

While the main purpose of this article is methodological, it also contributes to legislative studies by redressing bias in existing analyses of the duration of congressional careers. Legislative scholars have typically focused on only one mode of departure from Congress (for ambition, see Black, 1972; Brace, 1984; Copeland, 1989; Rohde, 1979; for retirement, see Brace, 1985; Groseclose and Krehbiel, 1994; Hall and Houweling, 1995; Hibbing, 1982; Theriault, 1998). Frantzich (1978) classifies six reasons for leaving the House but he does not consider the relationship among them. Kiewiet and Zeng (1993) examine ambition for higher office and retirement together, while Cox and Katz (2002), Jacobson and Kernell (1983) and Wolak (2007) relate retirement to electoral failure. Finally, Box-Steffensmeier and Jones (1997) examine four means of exit from Congress. While these analyses have all contributed to our understanding of the duration of congressional careers, it is known that the bias problem plagues the empirical results (Cox and Katz, 2002).

On the one hand, when scholars analyze whether legislators are reelected or not, they usually drop retirees, since they did not enter the election. Some of those retirees may have chosen not to stand for election because, having received a narrow margin of victory in a previous contest and expecting a high-quality challenger to enter the race, they fear they might lose if they run again. By omitting retirees, studies likely underestimate the extent to which unpromising prior electoral margins precipitate the end of congressional careers. On the other hand, when scholars have studied whether incumbents run again or not, the ultimate electoral results of course can not be used because they are *post hoc* and at any rate only available for those who decide not to retire. This study addresses both problems, effectively using data from all stages of the electoral process (from the decision to run through the results of general elections) to produce better estimates of the dependent risks involved in the end of congressional careers.

The substantive message of the present paper is that electoral challenges are more effective in unseating incumbents than is conventionally thought. Even when the rate of actual incumbent defeat is low, lawmakers may be strategically retiring so as to avoid electoral defeat. Dissatisfied constituents may succeed in replacing legislators not by failing to choose them but by discouraging them from running in the first place. This may seem an obvious possibility, but existing literature has not been conclusive on this point (Wolak, 2007). The present article claims that the substantial turnover of American elections is higher, and thus citizens are more effective at holding representatives to account, than has been recognized.

The remainder of the paper is organized as follows. The first section introduces a SDCR model of survival analysis, using the running example of the U.S. congressional career. The next section demonstrates Monte Carlo simulation to show how well the SDCR model addresses the bias to which the MNL model is prone. The third section reanalyzes the Box-Steffensmeier and Jones (2004) data and illustrates that the SDCR model reveals strategic retirement, which the original analysis underestimates. The penultimate section discusses limits of the SDCR model and elaborates on how this approach can be applied to other research topics. The concluding section summarizes the methodological and empirical findings of the article.

MODEL

Survival analysis is the appropriate framework for studying when and how legislators' political careers come to an end. The present section employs a discrete time approach to survival analysis, although a general model using continuous time will be developed subsequently.

Each observation is a legislator, indexed by i, at electoral term t. The dependent random variable Y represents the fate of the legislator at the end of the term. If a legislator decides not to run by the end of the term, $y = 1$. When an incumbent runs but is not reelected at primary, $y = 2$. In the case where the candidate wins the primary but loses in the general election, $y = 3$. Finally, if the incumbent survives the general election, the term is said to be "censored" and $y = 4$.

Serial Competing Risks

In a serial competing risks model, a subject who survives one type of risk may be subsequently exposed to the next. For example, a lawmaker who decides to run for an election (in which case one knows that $y > 1$) is at risk of losing in the primary election. Once the incumbent wins the primary ($y > 2$), the candidate is then at risk of losing the general election. On the other hand, a candidate who loses the primary ($y = 2$) never faces the risk of general election defeat. No candidate is at risk of losing both primary and general elections at the same time. Thus, there is a serial order in which a subject is at each risk.

In the discrete time model, an observation of a legislator i at term t appears in the dataset only if the legislator survives the previous terms $u(< t)$. Thus, the hazard at term t is the same as the probability that an event occurs during term t. The quantity of interest is not the *marginal* hazard for risk r ($Pr(y = r|t), r \in \{1, 2, 3\}$) but the *conditional* hazard where a subject survives preceding risks $s(< r)$: $h_r(t) = Pr(y = r|t, y \geq r)$. The marginal hazards are thus

$$Pr(y = 1|t) = h_1(t)$$
$$Pr(y = 2|t) = (1 - h_1(t)) \times h_2(t) \tag{1}$$
$$Pr(y = 3|t) = (1 - h_1(t)) \times (1 - h_2(t)) \times h_3(t)$$

and the survival function is

$$Pr(y = 4|t) = (1 - h_1(t)) \times (1 - h_2(t)) \times (1 - h_3(t)). \tag{2}$$

Since these conditional hazards fall between 0 and 1, a binomial logistic model is employed. The log odds of the conditional hazard for risk r ($r \in \{2, 3\}$; $r = 1$ is addressed later) is parameterized by the log odds of a baseline hazard for risk r ($\lambda_r^0(t)$) plus a linear predictor ($g_r(\cdot)$) of time varying covariates ($x(t)$, not including the constant term):

$$h_r(t) = \frac{\lambda_r^0(t) \exp(g_r(x(t)))}{1 + \lambda_r^0(t) \exp(g_r(x(t)))} \quad (r \in \{2, 3\}). \tag{3}$$

Baseline hazards and linear predictors will be discussed further below.

Parallel Competing Risks

In a parallel competing risks model, a subject is exposed to several risks simultaneously. For instance, Box-Steffensmeier and Jones (2004) decompose an incumbent's choice not to run ($y = 1$) into two distinct types: bypassing an election in order to seek an alternative office (which they refer to as "ambition" ($y = 1(a)$)) and retirement for other reasons ($y = 1(b)$). There is no necessary order in which these events take place, that is, a lawmaker is at both risks simultaneously. Let the marginal hazard of ambition and that of retirement be denoted by $h_{1(a)}(t)$ and $h_{1(b)}(t)$, respectively.

$$\begin{aligned} h_{1(a)}(t) &= Pr(y = 1(a)|t) \\ h_{1(b)}(t) &= Pr(y = 1(b)|t) \\ h_1(t) &= Pr(y = 1(a) \cup y = 1(b)|t) \\ &= Pr(y = 1(a)|t) + Pr(y = 1(b)|t) - Pr(y = 1(a) \cap y = 1(b)|t) \\ &= h_{1(a)}(t) + h_{1(b)}(t). \qquad (\because \text{the two events are mutually exclusive}) \end{aligned} \tag{4}$$

If one assumes that these two hazards are independent of each other conditioned on covariates, one can use a single MNL model.[1] In the notation of the present paper, the competing risks of retirement and ambition can be written:

$$\begin{aligned} h_{1(a)}(t) &= \frac{\lambda^0_{1(a)}(t)\exp(g_{1(a)}(x(t)))}{1 + \lambda^0_{1(a)}(t)\exp(g_{1(a)}(x(t))) + \lambda^0_{1(b)}(t)\exp(g_{1(b)}(x(t)))} \\ h_{1(b)}(t) &= \frac{\lambda^0_{1(b)}(t)\exp(g_{1(b)}(x(t)))}{1 + \lambda^0_{1(a)}(t)\exp(g_{1(a)}(x(t))) + \lambda^0_{1(b)}(t)\exp(g_{1(b)}(x(t)))}. \end{aligned} \tag{5}$$

Dependence between Competing Risks

A common method of analyzing competing risks is the MNL model, where all risks are regarded as parallel. Suppressing the time indicator, the model is expressed as:

$$Pr(y = r) = \frac{\lambda^0_r \exp(g_r(x))}{1 + \sum_r \lambda^0_r \exp(g_r(x))},$$

[1] By contrast, Rohde (1979)'s study of ambition excludes members who retire. But since those who eventually retire are also at the risk of ambitious exit (i.e. leaving for higher office), his method is not appropriate.

where $r \in \{1(a), 1(b), 2, 3\}$. Box-Steffensmeier and Jones (2004, 173-74) and Kiewiet and Zeng (1993) employ this model.[2] This model assumes independence of irrelevant alternatives, namely, that competing risks are independent of each other. However, one competing risk may be dependent on another. For instance, those incumbents whose prior vote margin was slim may expect that they would lose in a reelection bid if they dared to run. They may therefore choose to retire before the election in order to avoid financial cost and humiliation. Thus, the retirement hazard ($h_{1(b)}$) should be dependent on the general election hazard (h_3). Therefore, if one uses the MNL estimator, the coefficients of election hazard covariates such as prior margin will be biased toward zero (for another explanation, see Brace, 1985, 111-2).

The most common way to deal with dependent competing risks is the frailty model (Box-Steffensmeier and Jones, 2004; Gordon, 2002). The model of general election and retirement hazards can be written as follows:

$$\begin{aligned} g_3(x) &= x_3\beta_3 + \nu_3 \\ g_{1(b)}(x) &= x_{1(b)}\beta_{1(b)} + \nu_{1(b)} \\ (\nu_3, \nu_{1(b)}) &\sim p_{31(b)}(\nu_3, \nu_{1(b)}), \end{aligned}$$

where x_3 (or $x_{1(b)}$) is a vector of covariates for hazard h_3 (or $h_{1(b)}$), the ν's (frailties) are random variables independent of covariates but not of each other and $p_{31(b)}$ is a bivariate normal distribution. Conditioned on these covariates and random variables, conditional hazards h_3 and $h_{1(b)}$ are independent. In that sense, one can call the frailty model a *stochastically* dependent competing risks model. To estimate parameters, one needs to integrate out the frailties

[2] Box-Steffensmeier and Jones (2004, 169-72) also introduce multiple binomial logistic regressions using all observations for independent competing risks. Such approaches are not appropriate, however, because they do not constrain the data generation process so that only one event is observed.

through Markov Chain Monte Carlo methods or numerical integration. In cases with more than two risks (like the running example of congressional careers, where subjects are at four risks), however, it is very difficult to identify estimates even by integrating a multivariate normal distribution (Gordon, 2002).[3]

To address these limitations, this article proposes a *systematically* dependent competing risks model, where the hazard for one risk is conditioned on the same linear combination of covariates used to estimate the other hazard(s). For the two risks of the above example,

$$\begin{aligned} g_3(x) &= x_3\beta_3 \\ g_{1(b)}(x) &= x_{1(b)}\beta_{1(b)} + \rho_{31(b)}g_3(x) \\ &= x_{1(b)}\beta_{1(b)} + \rho_{31(b)}(x_3\beta_3), \end{aligned}$$

where $\rho_{31(b)}$ is the dependence parameter expressing the extent to which hazard h_3 affects hazard $h_{1(b)}$. x_3 affects $h_{1(b)}$ in the same way (which is expressed as β_3) as it does h_3.[4] There

[3]For two serial risks, you can use the probit model with sample selection (Van de Ven and Van Praag, 1981) instead of the frailty model. For more than two risks, however, this does not work because of the same identification problem of the multivariate normal distribution. For parallel risks, you may substitute the multinomial probit or the random utility maximization nested logit (Heiss, 2002; Maddala, 1983, 67-73) for the frailty model. The latter assumes common correlation among the parallel risks, which may not be reasonable for actual political science data. Finally, all these models fail to deal with serial and parallel risks simultaneously or take into account systematic dependence among them.

[4]To put it differently, the SDCR model is restricted as follows:

$$\begin{aligned} g_3(x) &= x_3\beta_3 \\ g_{1(b)}(x) &= x_{1(b)}\beta_{1(b)} + x_3\beta_{31(b)} \\ \exists\rho_{31(b)} : \beta_{31(b)} &= \rho_{31(b)}\beta_3 \end{aligned}$$

are two points of difference between the two competing risks models. Dependence among risks in the SDCR model is systematic in the sense that it is a relationship between observed covariates instead of random variables. The second, but more important, departure is that the SDCR model can deal with more than two risks, which is practically impossible for the frailty model.

Though there can be various model specifications of four risks of congressional career

---

Gordon (2002) also introduces "systematic dependence" but with a different meaning than in the present paper. Gordon means that ν_3 and $\nu_{1(b)}$ are systematically related as, for example, $\nu_3 = \nu_{1(b)}^{-1}$. In the SDCR model, no random variables are used in the first place.

The reader may wonder why it is not sufficient just to include the same independent variables for both hazards as

$$g_3(x) = x\beta_{x3} + z\beta_{z3}$$
$$g_{1(b)}(x) = x\beta_{x1(b)} + z\beta_{z1(b)}.$$

In fact, this model is often employed. It should be clear that a model in which two risks are dependent simply because they are each determined by the same variables is quite different from one in which the dependence is the result of a direct effect of one on the other (this study deals with the latter case). The former model implies that, following a certain amount of change in one risk (say, $\Delta g_3 = 1$), the other risk ($g_{1(b)}$) increases (or decreases) by $\beta_{x1(b)}/\beta_{x3}$ due to x but by $\beta_{z1(b)}/\beta_{z3}$ due to z, where the former is not necessarily the same as the latter. But if one risk (g_3) really affects the other risk ($g_{1(b)}$), the relationship between them should be constant whichever covariate changes (c.f. Figure 4) and the value of $\beta_{x1(b)}/\beta_{x3} = \beta_{z1(b)}/\beta_{z3} = \rho_{31(b)}$ is exactly the quantity of interest in the study of strategic retirement. Thus, this simple model is not appropriate for capturing dependence between risks. Moreover, it does not make (statistically) clear to what extent the two hazards are dependent, whereas the SDCR model directly expresses it as $\rho_{31(b)}$.

termination, this article assumes that, as strategic politician theory argues (Cox and Katz, 2002; Jacobson and Kernell, 1983; Lazarus, 2005), general election hazard (h_3) increases the other three hazards and that ambition hazard ($h_{1(a)}$) decreases retirement hazard ($h_{1(b)}$).

$$\begin{aligned} g_{1(a)}(x) &= x_{1(a)}\beta_{1(a)} + \rho_{31(a)}(x_3\beta_3) \\ g_{1(b)}(x) &= x_{1(b)}\beta_{1(b)} + \rho_{31(b)}(x_3\beta_3) + \rho_{1(a)1(b)}(x_{1(a)}\beta_{1(a)}) \\ g_2(x) &= x_2\beta_2 + \rho_{32}(x_3\beta_3) \\ g_3(x) &= x_3\beta_3. \end{aligned} \tag{6}$$

When lawmakers anticipate their electoral odds are bad (i.e. when $g_3(x)$ is high), they are (i) more likely to strategically choose not to run in order to avoid the costs of a losing campaign ($\rho_{32} > 0$), (ii) more likely to be defeated by a relatively stronger candidate in the primary ($\rho_{31(b)} > 0$), and (iii) more likely to take a risk on another office ($\rho_{31(a)} > 0$). If a representative has a good chance of being promoted to higher office (i.e. when $g_{1(a)}(x)$ is high), this lawmaker may be less likely to leave politics for other reasons ($\rho_{1(a)1(b)} < 0$). Note that the SDCR model can take into account not only dependence between serial risks (ρ_{32}) but also dependence between parallel risks ($\rho_{1(a)1(b)}$) as well as dependence between serial risks and parallel risks ($\rho_{31(a)}, \rho_{31(b)}$). Moreover, since the maximum likelihood method is employed, estimation is both easier and faster than the frailty model.

If conditional hazards do not share any parameters, the above suggests use of nested (multinomial) logistic regression. Then, commercial software packages can estimate parameters by applying logistic regression for every risk separately, to only those observations that face a particular risk. For example, for the general loss hazard, only legislators who run for general election would be considered.[5] However, this is not the case for congressional careers.

[5]Box-Steffensmeier and Jones (1997) take this strategy, though in a somewhat unsatisfactory way: they do not mention serial competing risks, nor do they omit primary election

Conditional hazards share some parameters (β_r). So when researchers omit censored observations whose survival is dependent on the shared parameters, estimates of the parameters are biased. A better approach is therefore to use all observations in maximizing likelihood.

Some technical caveats are in order. In the following dependence model, more than two variables in x_r should be excluded from x_s for the reason of identification (not vice versa).[6]

$$\begin{aligned} g_r(x) &= x_r\beta_r \\ g_s(x) &= x_s\beta_s + \rho_{rs}(x_r\beta_r). \end{aligned}$$

losers when they study general election defeat.

[6]Let us decompose covariates into shared variables (w) and the others (z_r and z_s).

$$\begin{aligned} g_r(x) &= w\beta_{wr} + z_r\beta_{zr} \\ g_s(x) &= w\beta_{ws} + z_s\beta_{zs} + \rho_{rs}g_r(x) \\ &= w\beta^*_{ws} + z_s\beta_{zs} + z_r(\rho_{rs}\beta_{zr}), \end{aligned}$$

where $\beta^*_{ws} = \beta_{ws} + \rho_{rs}\beta_{wr}$. If there are no variables in z_r, ρ_{rs} can be any value and β^*_{ws} is determined accordingly. This is the multicollinearity problem. When z_r contains only one variable, this model is substantially the same as a simple model,

$$g_s(x) = w\beta^*_{ws} + z_s\beta_{zs} + z_r\beta^*_{zr}$$

and ρ_{rs} is determined as β^*_{zr}/β_{zr}. By contrast, since β_{zs} has no linear relationship with other parameters, even when z_s has only a small number of variables, the multicollinearity problem is avoided. Monte Carlo simulation (not reported) bears out these points. The omission of some variables in the model for $g_r(x)$ from the model for $g_s(x)$ is reminiscent of the exclusion restriction condition in the Heckman selection model which is also required to avoid the multicollinearity problem.

As in the case of the multicollinearity problem in OLS or the exclusion restriction condition in the Heckman selection model, deciding which variables to exclude from or include in the present model requires judgment and insight into the particular application. The present paper recommends the following general rule:

Exclusion Rule: Exclude from x_s those variables (say, x_{rs}) that significantly affect h_r, when they are expected to affect h_s indirectly through ρ_{rs}.

The sections below will provide examples.

Likewise, for every pair of hazards, at most one hazard may have a dependence parameter. For instance, in the current application, it is assumed that the general election hazard induces the retirement hazard, not vice versa. Once incumbents survive the retirement hazard, there is no reason to think that they would be more or less vulnerable in the general election solely due to their retirement hazard. This model therefore forces analysts to carefully and explicitly consider the direction of causality among competing risks in a way the frailty model does not. While somewhat restrictive, this may improve research design and make results easier to interpret.

Estimation

The SDCR model employed here follows Box-Steffensmeier and Jones's (2004) measure of time dependency, namely, the natural log of the number of terms plus the constant term.

$$\log \lambda_r^0(t) = \alpha_r + \tau_r \log(t). \tag{7}$$

The likelihood of a set of parameters, θ (the α's, the τ's, the β's and the ρ's), for legislator i's exit type y at term t is denoted by $\mathscr{L}_i(\theta|y_{it}, t)$:

$$\mathscr{L}_i(\theta|y_{it}, t) \propto Pr(y_{it}|t, \theta).$$

The likelihood is obtained by substituting equations (1) through (7) with $Pr(y|t,\theta)$ (the detailed likelihood function is described in the Appendix). Suppose a legislator i is observed up to the T_ith term and there are n legislators. Then, total likelihood $\mathscr{L}$ is

$$\mathscr{L} = \prod_{i=1}^{n} \prod_{t=1}^{T_i} \mathscr{L}_i(\theta|y_{it}, t).$$

The goal of maximum likelihood estimation is to find the parameters that maximize this expression. Since T_i observations of the same legislator i are not independent of each other due to unobserved and omitted idiosyncratic factors of the legislator, analysts need to use robust standard errors. See the Appendix again for details on the approach taken to obtain robust standard errors.

Generalization

This subsection generalizes the dependent competing risks model to include both *systematically* dependent competing risks (SDCR model) and *stochastically* dependent competing risks (frailty model) and to account for continuous time. First, one can consider a partial SDCR model:

$$g_r(x) = x_{r0}\beta_{r0} + x_{rs}\beta_{rs}$$
$$g_s(x) = x_s\beta_s + \rho_{rs}(x_{rs}\beta_{rs}),$$

where x_{r0} denotes covariates that appear in the model for risk r but not the model for risk s, x_{rs} denotes covariates that appear in the model for risk r that also appear in the model for risk s, and covariates in x_s may or may not appear in x_{rs} or x_{r0}. If x_{r0} is empty, the model is reduced to the full SDCR model in the previous sections.

It is also easy to synthesize partial SDCR and frailty models into a general dependent competing risks model:

$$
\begin{aligned}
g_r(x) &= x_{r0}\beta_{r0} + x_{rs}\beta_{rs} + \nu_r \\
g_s(x) &= x_s\beta_s + \rho_{rs}(x_{rs}\beta_{rs}) + \nu_s \\
(\nu_r, \nu_s) &\sim p_{rs}(\nu_r, \nu_s).
\end{aligned}
$$

The same problems for estimation and identification that Gordon (2002) mentions arise here.

It is now clear when researchers should use SDCR model or frailty model. On the one hand, if they know what factors produce dependence among risks, one lets these risks share these same factors as covariates, as in x_{rs} in the above notation. This was the approach taken in the previous section addressing the SDCR estimation of congressional career duration. On the other hand, when a researcher is not sure what factors produce dependence among risks or "an analyst has reason to suspect the presence of stochastic dependence among risks due to the prevalence of unobserved heterogeneity" (Gordon, 2002, 214), he or she should put unknown factors or "unobserved heterogeneity" into the frailty model as random effects like the ν's. If the joint distribution of the two random variables (p_{rs}) follows a bivariate normal, its variance covariance matrix and dependence parameter ρ_{rs} represent how strongly and in which direction the two risks are correlated conditioned on covariates:

$$
\begin{aligned}
\frac{\partial g_s(x)}{\partial x_{rs}} &= \rho_{rs}\beta_{rs} \\
E\left(\frac{\partial g_s(x)}{\partial \nu_r}\right) &= \frac{\text{Cov}(\nu_r, \nu_s)}{\text{Var}(\nu_r)}.
\end{aligned}
$$

In a continuous time model, hazards have an open lower bound of zero but no upper bound (unlike in the case of the discrete time model, where hazards are probabilities and thus bounded above at 1), the present paper parameterizes log (not log odds as in the discrete time model) of marginal (not conditional as in the discrete time model) hazard by the log of the baseline hazard plus a linear predictor:

$$h_r(t) = \lambda_r^0(t)\exp(g_r(x(t)|\nu_r))$$

$$\nu = (\nu_1, \ldots, \nu_{\bar{r}}) \sim p_\nu(\nu)$$

and $\bar{r}$ is the number of risks. This is exactly the proportional hazard model. You can assume any kind of continuous distribution for the baseline hazard, $\lambda_r^0(t)$ (and you can do so non-parametrically in the Cox model). The argument above for $g_r(x(t))$ in the discrete time model holds in the continuous time model, as well. The estimation procedure is the same as the stochastic competing risks models for continuous time (Gordon, 2002).

MONTE CARLO SIMULATION

Setup

Monte Carlo simulation in this section demonstrates that, when the true data generation process is what the SDCR model assumes (as described in the Appendix), the MNL model produces biased estimates. It is shown that the coefficient estimates of a few independent variables on general election hazard are biased toward zero because strategic retirement is not considered. The setup of the Monte Carlo simulation is constructed so that it is easy to compare with the estimated parameters one obtains by applying the SDCR model to congressional data, as will be seen in the next section. The number of observations is $5,399$. The independent variables are $\log(\tilde{t})$, Vote, Redistrict, Scandal, Dummy1, and Dummy2. Logarithm of the term, $\log(\tilde{t})$, is sampled from a normal distribution with mean zero and variance 0.6. Vote is sampled from a uniform distribution between -50 and 50. Suppose that a random variable $b(m)$ is equal to $1 - m$ with the probability of m and is equal to $-m$ with the probability of $1 - m$. Redistrict, Scandal, Dummy1, and Dummy2 are sampled from $b(0.02), b(0.013), b(0.2)$ and $b(0.4)$, respectively. Since averages are almost zero for all sampled independent variables, it is easy to calculate approximate averages of conditional probabilities of an event by using the constant term only (e.g. $\bar{h}_3 = (1+\exp(-\tilde{\alpha}_3))^{-1}$). These independent variables are sampled only once through the whole process of this simulation.

Let $\tilde{\alpha}_{1(a)} = -3.4, \tilde{\alpha}_2 = -4.4, \tilde{\alpha}_3 = -3.5, \tilde{\tau}_{1(a)} = 0.6, \tilde{\tau}_{1(b)} = 1.3, \tilde{\tau}_2 = 0.3, \tilde{\tau}_3 = -0.3, \tilde{\rho}_{1(a)1(b)} = 0.08, \tilde{\rho}_{31(a)} = -0.01, \tilde{\rho}_{31(b)} = 0.18, \tilde{\rho}_{32} = 0.3, \tilde{x}_3 =$(Vote, Redistrict, Scandal, Dummy1, Dummy2) and $\tilde{\beta}_3 = (b_{\text{vote}}, 1.56, 3.16, 0, 0)$. The values of b_{vote} and $\tilde{\alpha}_{1(b)}$ are manipulated as explained shortly. Since the coefficients of Vote, Redistrict and Scandal in $\tilde{\beta}_3$ are not zero, they are excluded from $\tilde{x}_{1(a)} = \tilde{x}_2 =$ (Dummy1, Dummy2). Suppose $\tilde{\beta}_{1(a)} = (1.3, -1.4)$ and $\tilde{\beta}_2 = (-0.2, 0.2)$. Not only Vote, Redistrict and Scandal but also Dummy1 and Dummy2 are excluded from $\tilde{x}_{1(b)} = \emptyset$, and thus $\tilde{\beta}_{1(b)} = \emptyset$.

A single iteration of the simulation is performed as follows. Following the SDCR model, the dependent variable Y is sampled. Then, point estimates of each parameter, their standard errors and the 95 % confidence interval are calculated with both the SDCR and MNL estimators. In the MNL analysis, all the independent variables are used as the covariates for all of the four events. This trial is repeated $1,000$ times for a set of values of $\tilde{\alpha}_{1(b)}$ and b_{vote}. The simulation produces estimates using values of $\tilde{\alpha}_{1(b)}$, which determines the average probability of retirement, ranging from -4 to 1 (at one-point intervals) and b_{vote}, the coefficient of Vote on general election hazard, taking the value of -0.06 or -0.1.[7]

[7]The author wrote the code and estimated parameters of the model using the open source statistical software package `R`. The code is available at `http://www-cc.gakushuin.ac.jp/~e982440/SDCR/`. For the MNL model, `multinom` command in `nnet` (neural net) library is used. Robust standard errors are not calculated because all observations are sampled independently.

Results

Though it is usually impossible to compare parameters between different models, the SDCR model and the MNL one are the same as for the general electoral defeat probability conditioned on their primary victory ($h_3 = (1 + \exp(-(\tilde{\alpha}_3 + \tilde{\tau}_3 \log(t) + \tilde{x}_3\tilde{\beta}_3))^{-1})$). If average estimates of $\tilde{\beta}_3$ by the MNL is different from $\tilde{\beta}_3$, the MNL estimator is biased (this does not hold for $\tilde{\beta}_{1(a)}, \tilde{\beta}_{1(b)}$ and $\tilde{\beta}_2$). Among parameters in $\tilde{\beta}_3$, this paper focuses on the coefficient estimates of Vote, Redistrict and Scandal which are included in $\tilde{x}_3$ but excluded from $\tilde{x}_{1(a)}, \tilde{x}_{1(b)}$ and $\tilde{x}_2$.

Figure 1 illustrates that larger values of $\tilde{\alpha}_{1(b)}$ or b_{vote} leads to worse bias of the estimate of b_{vote} as predicted. The upper panel is the case where the true value of b_{vote} is -0.06. The vertical axis represents estimate values of b_{vote}, while the horizontal axis corresponds to the values of $\tilde{\alpha}_{1(b)}$. The solid line connects averages of $1,000$ estimates of b_{vote} by the SDCR model for every value of $\tilde{\alpha}_{1(b)}$. The shaded area is the 95 % confidence interval of the $1,000$ estimates, namely the area between the 2.5 % quantile value and the 97.5 % quantile value. The dashed line and the patterned area indicate the same for the MNL model. While the average of the SDCR estimates (solid line) is almost equal to the true value for all levels of $\tilde{\alpha}_{1(b)}$, the average of the MNL estimates (dashed line) is above the true value, and increasingly so as $\tilde{\alpha}_{1(b)}$ gets large. Deviation of the dashed line from the solid line illustrates the amount of bias. The MNL biases the estimate of b_{vote} toward zero.

[Figure 1 about here]

The lower panel shows the case of $b_{\text{vote}} = -0.1$. Bias is severe here and even the 95 % confidence interval (patterned area) does not contain the true value when $\tilde{\alpha}_{1(b)} \geq -2$. Moreover, when $\tilde{\alpha}_{1(b)}$ is not smaller than zero, the 95 % confidence intervals of both models (shaded area and patterned one) no longer overlap.

Figure 2 displays the same results for the coefficients of Redistrict and Scandal when

$b_{\text{vote}} = -0.06$. The SDCR estimates are unbiased. But, the larger $\tilde{\alpha}_{1(b)}$ is, the more biased the MNL estimates are toward zero. When $\tilde{\alpha}_{1(b)}$ is 1, the coefficient estimates of Redistrict and Scandal can be negative with the probability of 3 % and 1 %, respectively.

[Figure 2 about here]

In Figure 3, the vertical axis measures how often the true coefficient values of Vote ($b_{\text{vote}} = -0.06$ or -0.1), Redistrict or Scandal fall in their 95 % confidence intervals computed from their point estimates and standard errors based on the normal distribution. The horizontal axis shows the values of $\tilde{\alpha}_{1(b)}$. It is not surprising that roughly 95 % of the SDCR estimates are in their 95 % confidence intervals because the SDCR model is exactly the data generation process. Rather, what one should pay attention to is that the MNL coverage rates (except for Redistrict) become worse as $\tilde{\alpha}_{1(b)}$ becomes larger , especially when the effect of Vote on general election hazard is strong ($b_{\text{vote}} = -0.1$).

[Figure 3 about here]

The simulation indicates the extent to which dependence among competing risks can lead to bias in MNL estimates, and suggests that scholars should instead consider an SDCR model when such dependence is suspected.

REANALYSIS

Data

This section continues to compare the SDCR and MNL models by applying both models to the U.S. congressional career data studied by Box-Steffensmeier and Jones (2004).[8] It is

[8] Their data and STATA code were downloaded from `http://www.u.arizona.edu/~bsjones/eventhistory.html` on June 29, 2005, with

comprised of information pertaining to the career path for every House member elected in each freshman class from 1950 to 1976. Each member of the House was tracked from the first reelection bid until the last term served in office. The data does not contain elections of and after the Republican revolution in 1994. There are 5,399 observations. For the purpose of model comparison, dependent and independent variables (including the baseline hazard form) are the same between this paper and Box-Steffensmeier and Jones (2004, Table 10.6), though some covariates are dropped as noted below. The dependent variable is realized risk (y) which takes five values: ambition, retire, primary, general or reelection. Duration time is the number of terms they have served up to the time of observation (t). The independent variables (x) are as follows:

Party: Whether or not the incumbent is a Republican.

Redistrict: Whether or not the incumbent's district was substantially redistricted.

Scandal: Whether or not the incumbent was involved in scandal.

Open Gub.: Whether or not there was an open gubernatorial seat available during the election cycle.

Open Sen.: Whether or not there was an open U.S. Senatorial seat available during the election cycle.

Leadership: Whether or not the incumbent had a leadership position in the House.

Age: The incumbent's age (in years) at each election cycle.

Prior Margin: The percentage vote margin (over the next best finishing candidate) the incumbent (or the party) received in the previous election.

This choice of independent variables is not exhaustive but is reasonable in the context of existing scholarship (Black, 1972; Brace, 1984, 1985; Copeland, 1989; Cox and Katz, 2002; Frantzich, 1978; Groseclose and Krehbiel, 1994; Hall and Houweling, 1995; Hibbing,

---

thanks to the authors for making these available. As of November 2008, they can be found at `http://psfaculty.ucdavis.edu/bsjjones/eventhistory.html`.

1982; Kiewiet and Zeng, 1993; Rohde, 1979; Theriault, 1998). Since the aim here is model comparison rather than interpretation of independent variables' effects, and since predicted effects of many variables are obvious, this article does not elaborate on them but refers interested readers to the original analyses (definition of these variables is described in Box-Steffensmeier and Jones (1997, 1448-49) and Box-Steffensmeier and Jones (2004, 109, 171)).

Box-Steffensmeier and Jones (2004, Table 10.6) employ all of the above covariates for each risk in the MNL model. Doing the same with the SDCR model would produce multicollinearity problems. Thus, the present paper excludes Redistrict, Scandal and Prior Margin from $x_{1(a)}, x_{1(b)}$ and x_2. Why? According to the congressional studies literature and the results of the original analyses cited above, it is expected that Redistrict and Scandal increase the general election hazard (h_3) (Ansolabehere, Snyder, and Stewart, 2000; Brace, 1985; Cox and Katz, 2002; Frantzich, 1978; Hall and Houweling, 1995; Kiewiet and Zeng, 1993) while Prior Margin decreases it (Bond, Covington, and Fleisher, 1985; Cox and Katz, 2002; Hibbing, 1982; Jacobson and Kernell, 1983). Furthermore, the general election hazard is expected to increase the primary election hazard (h_2), the retirement hazard ($h_{1(b)}$) and the ambition hazard ($h_{1(a)}$). Thus, for example, although Prior Margin will decrease retirement hazard, the (indirect) effect of Prior Margin on retirement hazard is already taken into consideration by way of hazard dependence parameter $\rho_{31(b)}$. Moreover, no substantive theory argues that Prior Margin has a direct effect on the retirement hazard regardless of the general election hazard. Therefore, we should exclude Prior Margin from the retirement hazard covariates (x_2).[9] A similar argument applies for Redistrict, Scandal, the primary

[9] Since this issue is analogous to that of an instrumental variable, the following explanation of an instrumental variable will be helpful for understanding the issue at hand: "[i]t is still the case that z [read instrument variable or Prior Margin] and y [dependent variable or retirement hazard] will be correlated, but the only source of such correlation is the indirect path of z being correlated with x [independent variable or electoral hazard], which in turn

election hazard and the ambition hazard. There is no other path from these three covariates to the three hazards than through the general election hazard.

By contrast, Age is supposed to have both direct and indirect effects on the ambition and retirement hazards. Once controlling for seniority (by baseline hazard), Age is expected to raise the general election hazard because older candidates are likely to be less attractive to their constituency since voters expect them to serve fewer future terms than younger candidates with the same seniority. Thus, Age will have an indirect effect on ambition and retirement via $\rho_{31(a)}$ and $\rho_{31(b)}$. In addition, according to previous studies (Brace, 1985; Groseclose and Krehbiel, 1994; Hall and Houweling, 1995; Kiewiet and Zeng, 1993; Theriault, 1998), irrespective of congressional electoral prospects and controlling for seniority, older legislators are more likely to retire for not electoral but physical or mental reasons and to have shorter time ahead to enjoy other prestigious offices. Therefore, Age is expected to have a direct effect on ambition and retirement, and is kept in $x_{1(a)}$ and $x_{1(b)}$.[10]

Open Gub. as well as Open Sen. are expected to facilitate ambitious exit from the House ($h_{1(a)}$) (Black, 1972; Copeland, 1989; Rohde, 1979) which will decrease retirement ($h_{1(b)}$) and, thus, are purged from $x_{1(b)}$. There is no other route by which these two independent variables are expected to affect retirement hazard than through ambition hazard.

Finally, Box-Steffensmeier and Jones (2004, 171) do not report the estimate of Leadership coefficient on defeats in primaries and note that "there were almost no cases where a member in the leadership lost in primary". Thus, the variable Leadership is dropped from x_2 in order to avoid unstable estimates.

determines y. The more direct path of z being a regressor in the model for y is ruled out" (Cameron and Trivedi, 2005, 97, bracketed text added).

[10]Excluding Age does not substantially change coefficient estimates for the other variables.

Results

Table 1 compares the results of the MNL model (which is almost the same as Box-Steffensmeier and Jones (2004, Table 10.6)) and the SDCR model.

[Table 1 about here]

The most important quantities of interest are the between-hazards dependence parameters (the ρ's). In the second to last row of the lower panel, the estimates of ρ_{32}, $\rho_{31(b)}$ and $\rho_{31(a)}$ are presented from the left. The last line shows $\rho_{1(a)1(b)}$. Only the dependence of the retirement hazard on the general election hazard ($\rho_{31(b)}$) is significantly different from zero (with the expected positive sign). But this measurement is not applicable to the MNL model. In order to compare estimated degree of strategic retirement between the two models, the current study pays attention to the ratio of the retirement hazard change to the general election hazard change, namely, the derivative of the retirement hazard with respect to the general election hazard ($\frac{\partial h_{1(b)}}{\partial h_3}$). This paper refers to this as "strategic retirement responsiveness."[11] Figure 4 compares strategic retirement responsiveness by both models. The figure was produced as follows. All covariates are set at zero (for dichotomous covariates and $\log(t)$) or mean value (for a continuous covariate, namely, 51.4 for Age) except for Prior Margin (x_{PM}), which moves from its empirical minimum (0) to maximum (100) by 0.1. The general election hazard ($h_3(x_{PM})$) and the retirement one ($h_{1(b)}(x_{PM})$) are calculated for each value of Prior Margin using the SDCR or MNL parameter point estimates. Since the retirement hazard does not have Prior Margin as one of its covariates ($x_{1(b)}$), it changes only because of changes in the general election hazard. The corresponding strategic retirement responsiveness at $h_3(x_{PM})$ is approximated by $\frac{h_{1(b)}(x_{PM}+0.1)-h_{1(b)}(x_{PM})}{h_3(x_{PM}+0.1)-h_3(x_{PM})}$. Besides

[11]But note that, since the MNL model assumes that no hazard is dependent on the other ones, it makes no sense to calculate the retirement hazard conditioned on the general election hazard in that setting.

these baseline curves, the lines for the cases where Redistrict is one or Scandal is one are also drawn similarly. Note that the lines of the SDCR model for the three cases overlap because, unlike the MNL model, the relationship between the general election hazard and the strategic retirement responsiveness does not depend on covariate values (footnote 4). The difference of the two models is clear: for most of the range of general election hazard values, the strategic retirement responsiveness of the MNL model is smaller than that of the SDCR model. Moreover, in the MNL model, after the general election hazard surpasses around 0.2, the strategic retirement responsiveness is less than zero, which means that the retirement hazard decreases; this is obviously not reasonable.

[Figure 4 about here]

Empirically, 5.8 % of those who ran for general election lost, while the SDCR model implies that 8.6 % ($\bar{h}_3(y = 1(b))$) of those who retired would lose if they ran, an increase of one and half times. Thus, the SDCR model supports the hypothesis that those who foresee electoral loss retire.

Another focus of this study is the coefficient of covariates included for the general election hazard but not for the other hazards. The SDCR estimates of Redistrict and Scandal coefficients are larger than the original analysis (the MNL model), while the SDCR estimate of Prior Margin coefficient is almost the same as the original analysis. The empirical average of $\hat{\alpha}_{1(b)} + \hat{\tau}_{1(b)} \log(t) + \hat{g}_{1(b)}(x_{1(b)})$ is -3.4. If one assumes the SDCR estimates are the true values, the situation is almost equivalent to the case of $\alpha_{1(b)} = -3.4$ in Figures 1 (the upper panel) and 2, where it is shown that bias of the MNL model is almost negligible and the 95 % confidence intervals of the MNL model are as wide as they should be.[12] Thus, the

[12]Let $\bar{G}_r$ be equal to the empirical average of $\hat{\alpha}_r + \hat{\tau}_r \log(t) + \hat{g}_r(x_r)$ and V_r be equal to the variance of $x_r\hat{\beta}_r$. It turns out that $\bar{G}_{1(a)} = -3.4, \bar{G}_{1(b)} = -3.4, \bar{G}_2 = -4.4, \bar{G}_3 = -3.5$. These are the almost same as the values of the $\tilde{\alpha}$'s in the Monte Carlo setup. Besides, it is shown

SDCR model suggests that the original analysis underestimates coefficients of Redistrict and Scandal but not that of Prior Margin because few lawmakers retire.

Some readers may question whether the SDCR model fits the data much better than the MNL model. Table 2 reports -2 log likelihood, Akaike information criteria (AIC) and Bayesian information criteria (BIC) of every model. Since the MNL and the SDCR model differ in the number of parameters, it is better to use AIC or BIC. The (full) SDCR model (the second row) is better than the MNL model in terms of BIC, though not in terms of AIC. Since three of the four dependence parameters (ρ's) in the (full) SDCR model are not significant, a simple SDCR model is estimated where no dependence parameter other than $\rho_{31(b)}$ is used and all independent variables are included except for Redistrict, Scandal and Prior Margin excluded from $x_{1(b)}$. This simple SDCR model (the third row) fits better than the MNL model not only in terms of AIC or BIC but also log likelihood, even if the former uses fewer parameters than the latter (note that the former is not nested within the latter). Finally, an unrestricted model is compared which has no dependence parameter and uses all independent variables. Again, in terms of BIC, both the full and simple SDCR models are better than the unrestricted model, though not in terms of AIC. Taking into account that both SDCR models contain fewer parameters than the unrestricted model, the SDCR models explain the data more efficiently. In sum, at least in terms of BIC, both SDCR models fit the data better than the MNL model and the unrestricted model.

[Table 2 about here]

---

that $V_{1(a)} = 0.73, V_{1(b)} = 0.49, V_2 = 0.02, V_3 = 2.50$. These are close to the variance of $\tilde{x}_r\tilde{\beta}_r$ in the Monte Carlo simulation (0.74, 0, 0.02 and 3.03, respectively, where $b_{\text{vote}} = -0.06$) except $V_{1(b)}$ and V_3. Though $\tilde{x}_{1(b)} = \emptyset$ in the simulation, the variance of $\tilde{\tau}_{1(b)} \log(\tilde{t})$ in the simulation, 0.99, is nearly equal to that of $\hat{\tau}_{1(b)} \log(t) + x_{1(b)}\hat{\beta}_{1(b)}$ of the real data, 1.05. In the Monte Carlo setup, the variance of $\log(\tilde{t})$ is 0.6 and approximately as large as that of this real data. Therefore, the real data and the Monte Carlo simulation are comparable.

While one could argue whether SDCR is cost-efficient in terms of its slightly better fit with the data, it is important to remember that the quantity of interest of this study is dependence among competing risks. The ρ's in both the full and simple SDCR models express this quantity directly, while the unrestricted model and the MNL model ignore it.

Finally, the coefficients of Age for ambition and retirement are significantly different from zero. Age has direct effects on ambition and retirement other than indirect effects caused by their dependence on the general election hazard. Age facilitates retirement not just because Age dampens electoral prospects but also because Age itself makes continued congressional service physically difficult. This result justifies inclusion of Age in $x_{1(a)}$ and $x_{1(b)}$.

DISCUSSION

Limits of the SDCR Model

Should scholars always employ the SDCR model instead of the MNL model? Not necessarily. First, the MNL model assumes that every hazard is independent of the others. Thus, the MNL model is preferred when researchers affirm this assumption. Second, given a dependent variable and a set of independent variables, there is only one MNL model, though there can be various kinds of SDCR models corresponding to hazard dependence structure. When scholars have no substantive theory that suggests dependence among hazards, they cannot specify which SDCR model to use and may turn to the MNL model.[13] Third, as the Monte Carlo simulation (Figures 1 through 3) implies, the smaller the retirement hazard ($\tilde{\alpha}_{1(b)}$), that is, the smaller a "downstream" hazard dependent on an "upstream" hazard (such as

[13] In this situation, choice of the MNL model as a kind of "null" or "focal" model seems straight-forward but is a matter of convenience, not substance. The assumptions of the SDCR model are not stronger than those of the MNL model.

electoral hazard), the less problematic the MNL estimates and, therefore, the less helpful the SDCR model.

But if none of the three conditions above is met, researchers would benefit from employing the SDCR model. The SDCR model enables and prompts students to take into consideration strategic behavior and quantify it. Though, to date, no ready-made software can implement the SDCR model, the author's code is available for the open source statistical software package `R` (see footnote 7). Even if scholars wish to study models different from that of the present paper (e.g. with different risk dependence structure, different number of risks), it is easy to adapt the present code for their models (for use with other statistical software packages).

Practically speaking, the present paper recommends that scholars estimate the MNL model first, for exploratory reasons. When covariates x_{rs} are significant for h_r and h_s and a reasonable theory supports that the upstream hazard h_r (e.g. electoral defeat) causes the downstream hazard h_s (e.g. retirement), researchers might omit x_{rs} from x_s and estimate the SDCR model ("Exclusion Rule" mentioned above). Only if students strongly suspect that some of these covariates also have direct effects on the downstream hazard, they might retain those covariates in the downstream hazard. Although the rule above may seem abstract, finding variables to exclude in the SDCR model is admittedly as difficult as exclusion restriction or instrumental variables.

In addition, the SDCR model is biased in case of omitted variables, because it does not take into consideration omitted variables. But the MNL model is biased toward zero compared with the SDCR models. This is because the MNL estimates suffer from omitted variables as well as dependence between risks, the latter of which the SDCR model addresses. Monte Carlo simulation (not reported, though available upon request) confirms this.

Potential Applications

In the field of political science, there are many other promising applications for which the SDCR model might be useful. A few of these are discussed below.

Diermeier and Stevenson (1999, 2000) and Lupia and Strøm (1995) improve upon earlier studies of cabinet duration by distinguishing between two types of cabinet termination in parliamentary systems: replacement by another cabinet without an election and chamber dissolution with a subsequent election. But, when a cabinet expects that it will be replaced, it may dissolve the legislature and call an election. Gordon (2002) applies his stochastically dependent competing risks model to this case but fails to find significant dependence. It may be the case, however, that the two hazards are systematically dependent, in which cases those estimates are biased.

The U.S. Senate tends to confirm most federal court judges and executive officials the President nominates. But the Senate sometimes rejects nominees, or adjourns without any decision. In some cases, nominees are withdrawn (Bond, Fleisher, and Krutz, 2002). When the confirmation hazard is high, rejection or withdrawal hazards should be low. Moreover, a President who anticipates senatorial rejection may prefer to withdrawal a candidate rather than pay the political costs of a defeat.

Cases worthy of study are also found in international relations. If countries involved in a war realize that an imposed settlement is likely but costly, they may hasten to stop the war by negotiation (cf. Slantchev, 2004). Therefore, it seems likely that an imposed settlement hazard raises the negotiated settlement hazard. Or, in an international crisis that ends with either back-down or an attack (cf. Fearon, 1994), the lower the back-down hazard, the higher the attack hazard. The SDCR model may be productively applied to test these hypotheses.

CONCLUSION

Competing risks models of survival analysis study whether, when and how an event happens in regards to a particular subject. Sometimes, these risks are dependent on each other. The SDCR model proposed in this article enables us to estimate more than two risks, which is almost impossible for the widely-used frailty model (i.e., a stochastically dependent competing risks model). The SDCR model also allows the researcher to separately analyze direct and indirect effects of covariates and distinguish between serial risks from parallel ones. Moreover, Monte Carlo simulation demonstrated that another often-employed method, the MNL model, would result in underestimation of parameters in the case where data is generated in the way the SDCR model assumes. Finally, the present paper combines both types of dependent competing risks model into a general one and extends it into continuous time.

This article has relevance to American politics as well by applying the SDCR model to data on the career paths of members of the U.S. House of Representatives. The main substantive contribution is to reveal that lawmakers strategically retire so as not to incur electoral defeat, and to estimate this effect by taking into account bias to which the often-employed MNL model is vulnerable.[14] The present analysis showed that, if they ran, retirees would be more likely defeated than the actual candidates. Thus, even if observed incumbent defeats are rare, voters may succeed in replacing legislators by forcing retirement. In this sense, American democracy works better in terms of popular control of the government than previously thought. This finding has relevance for policy debates on measures designed to

[14]It is true that scholars have demonstrated strategic retirement before, for example, by indicating that the same independent variables increase both the electoral hazard and the retirement hazard. But these scholars do not show how much the electoral hazard increases the retirement hazard because they do not model strategic retirement explicitly. The contribution of this paper is to make strategic retirement clear by quantifying it.

counter incumbency advantage, such as term limits and campaign spending legislation.

This study shows that the SDCR model can be useful when scholars want to understand the factors influencing distinct and interrelated ways in which a subject may be terminated (whether that subject is a legislator, a cabinet, a nominee, or a war). In addition to demonstrating the SDCR model's usefulness for explaining the end of congressional careers, the article offers a template for future applications in other promising areas of research.

APPENDIX: The Likelihood Function and the Robust Standard Error

Let $\theta = (\alpha_{1(a)}, \alpha_{1(b)}, \alpha_2, \alpha_3, \tau_{1(a)}, \tau_{1(b)}, \tau_2, \tau_3, \beta_{1(a)}, \beta_{1(b)}, \beta_2, \beta_3, \rho_{1(a)1(b)}, \rho_{31(a)}, \rho_{31(b)}, \rho_{32})$. The likelihood $\mathscr{L}_i(\theta|y_{it}, t, x_i(t))$ is proportional to $Pr(y_{it}|t, x_i(t), \theta)$ where

$$
\begin{aligned}
Pr(y_{it} = 1(a)|t, x_i(t), \theta) &= \frac{\exp(G_{1(a)}(t))}{1 + \exp(G_{1(a)}(t)) + \exp(G_{1(b)}(t))} \\
Pr(y_{it} = 1(b)|t, x_i(t), \theta) &= \frac{\exp(G_{1(b)}(t))}{1 + \exp(G_{1(a)}(t)) + \exp(G_{1(b)}(t))} \\
Pr(y_{it} = 2|t, x_i(t), \theta) &= \frac{1}{1 + \exp(G_{1(a)}(t)) + \exp(G_{1(b)}(t))} \times \frac{\exp(G_2(t))}{1 + \exp(G_2(t))} \\
Pr(y_{it} = 3|t, x_i(t), \theta) &= \frac{1}{1 + \exp(G_{1(a)}(t)) + \exp(G_{1(b)}(t))} \times \frac{1}{1 + \exp(G_2(t))} \times \frac{\exp(G_3(t))}{1 + \exp(G_3(t))} \\
Pr(y_{it} = 4|t, x_i(t), \theta) &= \frac{1}{1 + \exp(G_{1(a)}(t)) + \exp(G_{1(b)}(t))} \times \frac{1}{1 + \exp(G_2(t))} \times \frac{1}{1 + \exp(G_3(t))}
\end{aligned}
$$

$$
\begin{aligned}
G_{1(a)}(t) &= \alpha_{1(a)} + \tau_{1(a)} \log(t) + x_{i,1(a)}(t)\beta_{1(a)} + \rho_{31(a)}(x_{i,3}(t)\beta_3) \\
G_{1(b)}(t) &= \alpha_{1(b)} + \tau_{1(b)} \log(t) + x_{i,1(b)}(t)\beta_{1(b)} + \rho_{31(b)}(x_{i,3}(t)\beta_3) + \rho_{1(a)1(b)}(x_{i,1(a)}(t)\beta_{1(a)}) \\
G_2(t) &= \alpha_2 + \tau_2 \log(t) + x_{i,2}(t)\beta_2 + \rho_{32}(x_{i,3}(t)\beta_3) \\
G_3(t) &= \alpha_3 + \tau_3 \log(t) + x_{i,3}(t)\beta_3.
\end{aligned}
$$

Robust estimation of the variance covariance matrix is (Box-Steffensmeier and Zorn, 2002, 1090):

$$\hat{V}_R(\theta) = \hat{V}_{MLE} \sum_{i=1}^{n} \left[\left(\sum_{t=1}^{T_i} \frac{\partial \log \mathscr{L}_i}{\partial \theta} \right)' \left(\sum_{t=1}^{T_i} \frac{\partial \log \mathscr{L}_i}{\partial \theta} \right) \right] \hat{V}_{MLE},$$

where $\hat{V}_{MLE}$ is the maximum likelihood estimate of the variance covariance matrix. The square root of $\hat{V}_R(\theta)$'s diagonal evaluated where θ is equal to its maximum likelihood estimate gives robust standard errors.

References

Ansolabehere, Stephen, James M. Snyder, and Charles Stewart, Jr. 2000. "Old Voters, New Voters, and the Personal Vote: Using Redistricting to Measure the Incumbency Advantage." *American Journal of Political Science* 44 (1): 17-34.

Black, Gordon. 1972. "A Theory of Political Ambition: Career Choices and the Role of Structural Incentives." *American Political Science Review* 66 (1): 144-59.

Bond, Jon R., Cary Covington, and Richard Fleisher. 1985. "Explaining Challenger Quality in Congressional Elections." *Journal of Politics* 47 (2): 510-29.

Bond, Jon R., Richard Fleisher, and Glen S. Krutz. 2002. "Maligned Neglect: How Senate Consideration of Presidential Nominations Has Changed.". Prepared for Presentation at the 98th Annual Meeting of the American Political Science Association, Boston, MA, August 29 - September 1, 2002.

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 (4): 1414-61.

Box-Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists.* Cambridge, UK: Cambridge University Press.

Box-Steffensmeier, Janet M., and Christopher Zorn. 2002. "Duration Models for Repeated Events." *Journal of Politics* 64 (4): 1069-94.

Brace, Paul. 1984. "Progressive Ambition in the House: A Probabilistic Approach." *Journal of Politics* 46 (2): 556-71.

Brace, Paul. 1985. "A Probabilistic Approach to Retirement from the U.S. Congress." *Legislative Studies Quarterly* 10 (1): 107-23.

Cameron, A. Colin, and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.

Copeland, Gary W. 1989. "Choosing to Run: Why House Members Seek Election to the Senate." *Legislative Studies Quarterly* 14 (4): 549-65.

Cox, Gary W., and Jonathan N. Katz. 2002. *Elbridge Gerry's Salamander: The Electoral Consequences of the Reapportionment Revolution*. Cambridge, UK: Cambridge University Press.

Diermeier, Daniel, and Randolph T. Stevenson. 2000. "Cabinet Terminations and Critical Events." *American Political Science Review* 94 (3): 627-40.

Diermeier, Daniel, and Randy T. Stevenson. 1999. "Cabinet Survival and Competing Risks." *American Journal of Political Science* 43 (4): 1051-68.

Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88 (3): 577-92.

Frantzich, Stephen E. 1978. "Opting Out: Retirement from the House of Representatives, 1966-1974." *American Politics Quarterly* 6 (3): 251-73.

Gordon, Sanford. 2002. "Stochastic Dependence in Competing Risks." *American Journal of Political Science* 46 (1): 200-17.

Groseclose, Timothy, and Keith Krehbiel. 1994. "Golden Parachutes, Rubber Checks, and Strategic Retirements from the 102d House." *American Journal of Political Science* 38 (1): 75-99.

Hall, Richard L., and Robert P. van Houweling. 1995. "Avarice and Ambition in Congress: Representatives' Decisions to Run or Retire from the U.S. House." *American Political Science Review* 89 (1): 121-36.

Heiss, Florian. 2002. "Structural Choice Analysis with Nested Logit Models." *Stata Journal* 2 (3): 227-52.

Hibbing, John R. 1982. "Voluntary Retirement from the U.S. House of Representatives: Who Quits?" *American Journal of Political Science* 26 (3): 467-84.

Jacobson, Gary C., and Samuel Kernell. 1983. *Strategy and Choice in Congressional Elections*. 2nd ed. New Haven: Yale University Press.

Kiewiet, D. Roderick, and Langche Zeng. 1993. "An Analysis of Congressional Career Decisions, 1947-1986." *American Political Science Review* 87 (4): 928-41.

Lazarus, Jeffrey. 2005. "Unintended Consequences: Anticipation of General Election Outcomes and Primary Election Divisiveness." *Legislative Studies Quarterly* 30 (3): 435-61.

Lupia, Arthur, and Kaare Strøm. 1995. "Coalition Termination and the Strategic Timing of Parliamentary Elections." *American Political Science Review* 89 (3): 648-65.

Maddala, G. S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge, UK: Cambridge University Press.

Rohde, David. 1979. "Risk-Bearing and Progressive Ambition: The Case of Members of the United States House of Representatives." *American Journal of Political Science* 23 (1): 1-26.

Slantchev, Branislav L. 2004. "How Initiators End Their Wars: The Duration of Warfare and the Terms of Peace." *American Journal of Political Science* 48 (4): 813-29.

Theriault, Sean M. 1998. "Moving Up or Moving Out: Career Ceiling and Congressional Retirement." *Legislative Studies Quarterly* 23 (3): 419-33.

Van de Ven, Wynand P. M. M., and Benard M. S. Van Praag. 1981. "The Demand for Deductibles in Private Health Insurance: A Probit Model with Sample Selection." *Journal of Econometrics* 17 (2): 229-52.

Wolak, Jennifer. 2007. "Strategic Retirement: The Influence of Public Preferences on Voluntary Departures from Congress." *Legislative Studies Quarterly* 32 (2): 285-308.

	General			Primary			Retire			Ambition		
	Est.	S.E.		Est.	S.E.		Est.	S.E.		Est.	S.E.	
MNL Model												
Party	−0.18	0.12		−0.11	0.26		0.18	0.14		0.31	0.15	*
Redistrict	1.37	0.32	**	1.42	0.58	*	1.36	0.27	**	1.55	0.31	**
Scandal	2.62	0.35	**	3.11	0.44	**	1.09	0.42	**	−13.00	0.15	**
Open Gub.	0.12	0.15		0.20	0.31		0.04	0.17		0.49	0.15	**
Open Sen.	−0.21	0.21		−0.42	0.44		0.07	0.20		1.01	0.15	**
Leadership	−0.60	0.48		−21.73	0.3	**	−0.39	0.29		−1.53	1.02	
Age	0.04	0.01	**	0.04	0.02	*	0.08	0.01	**	−0.06	0.01	**
Prior Margin	−0.06	0.01	**	0.00	0.01		−0.01	0.00	**	0.00	0.00	
$\log(t)$	−0.26	0.09	**	−0.06	0.2		0.53	0.12	**	0.51	0.11	**
Constant	−3.13	0.36	**	−6.15	0.75	**	−7.89	0.49	**	−1.14	0.39	**
SDCR Model												
Party	−0.18	0.13		−0.18	0.25		0.17	0.15		0.32	0.15	*
Redistrict	1.56	0.35	**									
Scandal	3.16	0.44	**									
Open Gub.	0.11	0.16		0.05	0.31					0.48	0.16	**
Open Sen.	−0.23	0.22		−0.26	0.44					1.00	0.16	**
Leadership	−1.05	0.68					0.01	0.45		−1.60	1.03	
Age	0.04	0.01	**				0.07	0.01	**	−0.07	0.01	**
Prior Margin	−0.06	0.01	**									
$\log(t)$	−0.30	0.14	*	0.30	0.17		0.59	0.12	**	0.56	0.13	**
Constant	−3.37	0.73	**	−4.75	0.47	**	−7.81	0.49	**	−1.04	0.47	*
γ General				0.30	0.19		0.18	0.06	**	−0.01	0.05	
γ Ambition							0.08	0.18				

Note: $N = 5399$, log likelihood (MNL)=−2972.58, log likelihood (SDCR) =−2997.85
** $p < 0.01$, * $p < 0.05$, S.E.=Robust standard error.

Table 1: The Results of the Two Competing Risks Models of Congressional Career Paths

Model	−2 Log Likelihood	AIC	BIC
MNL	5945.2	6025.2	6288.9
Full SDCR	5995.7	6057.7	6262.1
Simple SDCR	5940.8	6016.8	6267.4
Unrestricted	5933.6	6013.6	6277.4

Table 2: Comparison of Model Fits

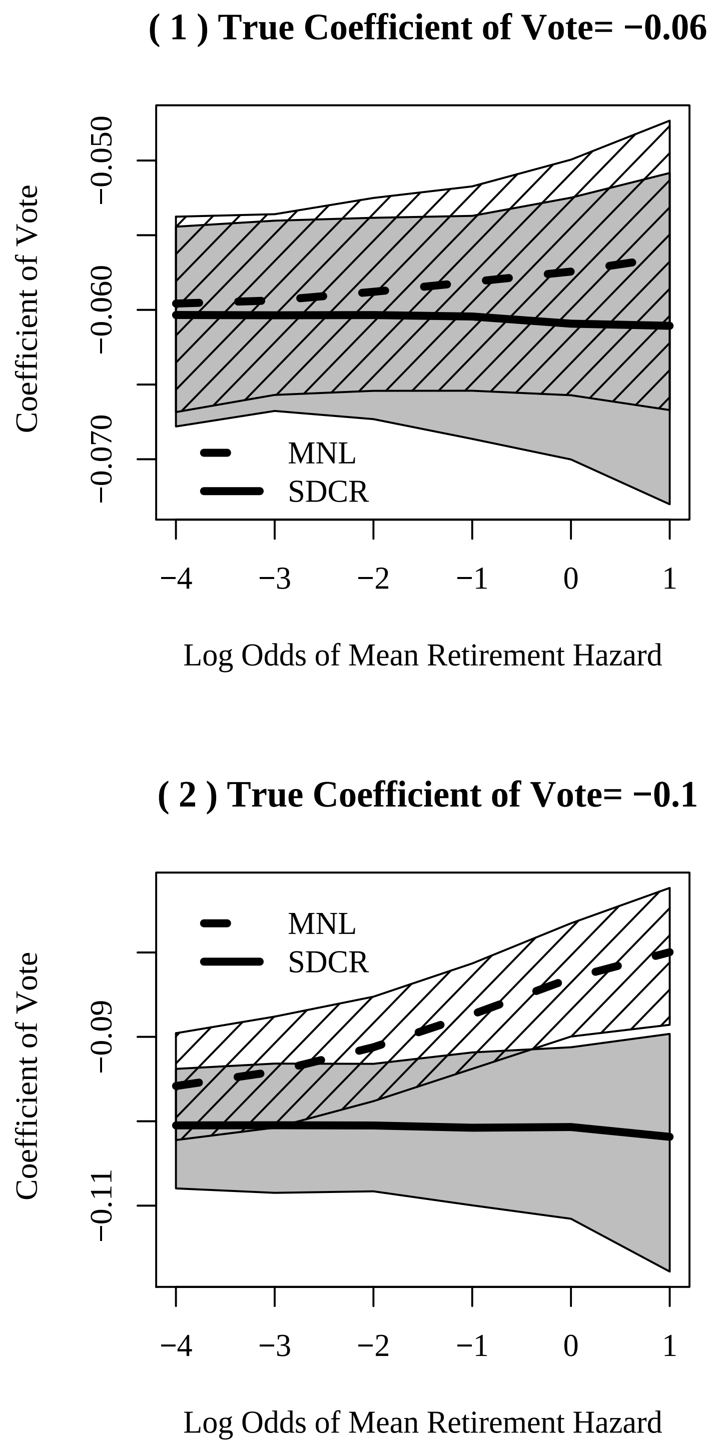
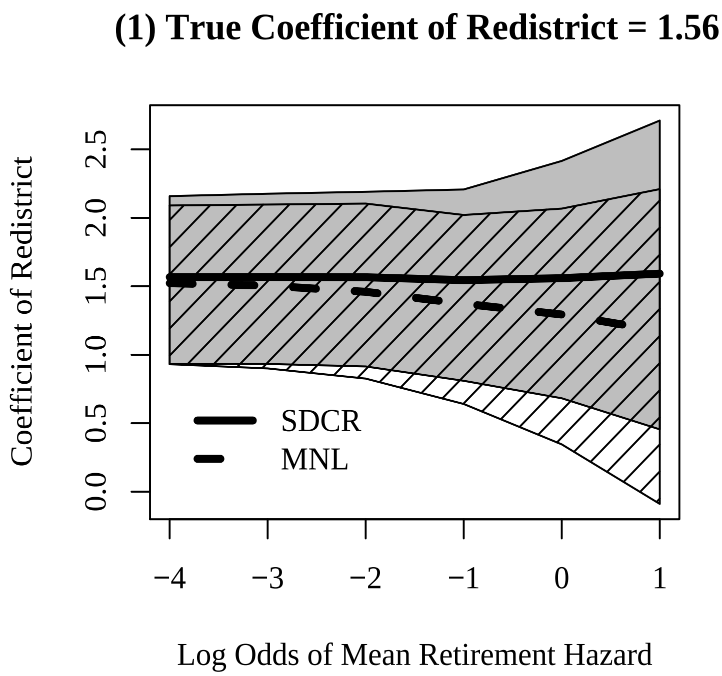


Figure 1: The Larger the Retirement Level ($\alpha_{1(b)}$) or the Larger the True Coefficient of Vote (b_{vote}), the More Biased Estimated Coefficient of Vote ($\hat{b}_{\text{vote}}$)



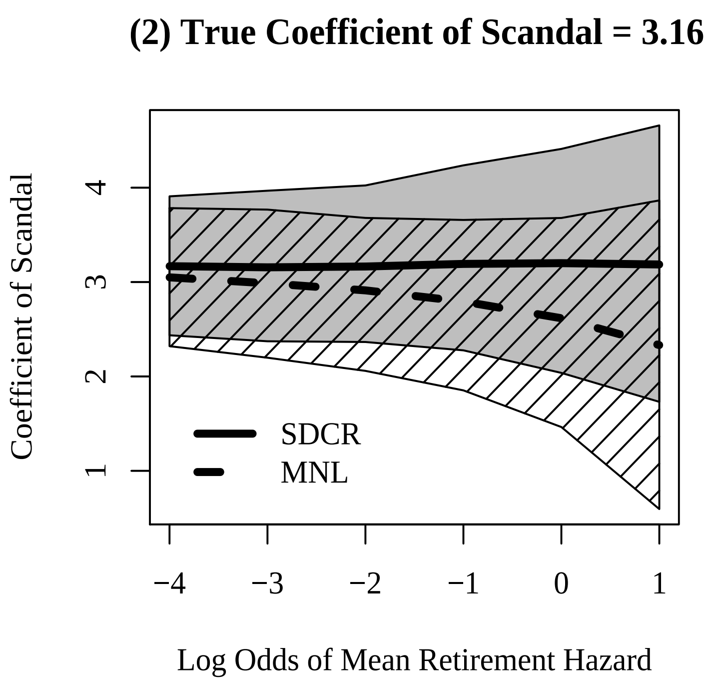


Figure 2: The Larger the Retirement Level ($\alpha_{1(b)}$), the More Biased Estimated Coefficients of Redistrict and Scandal

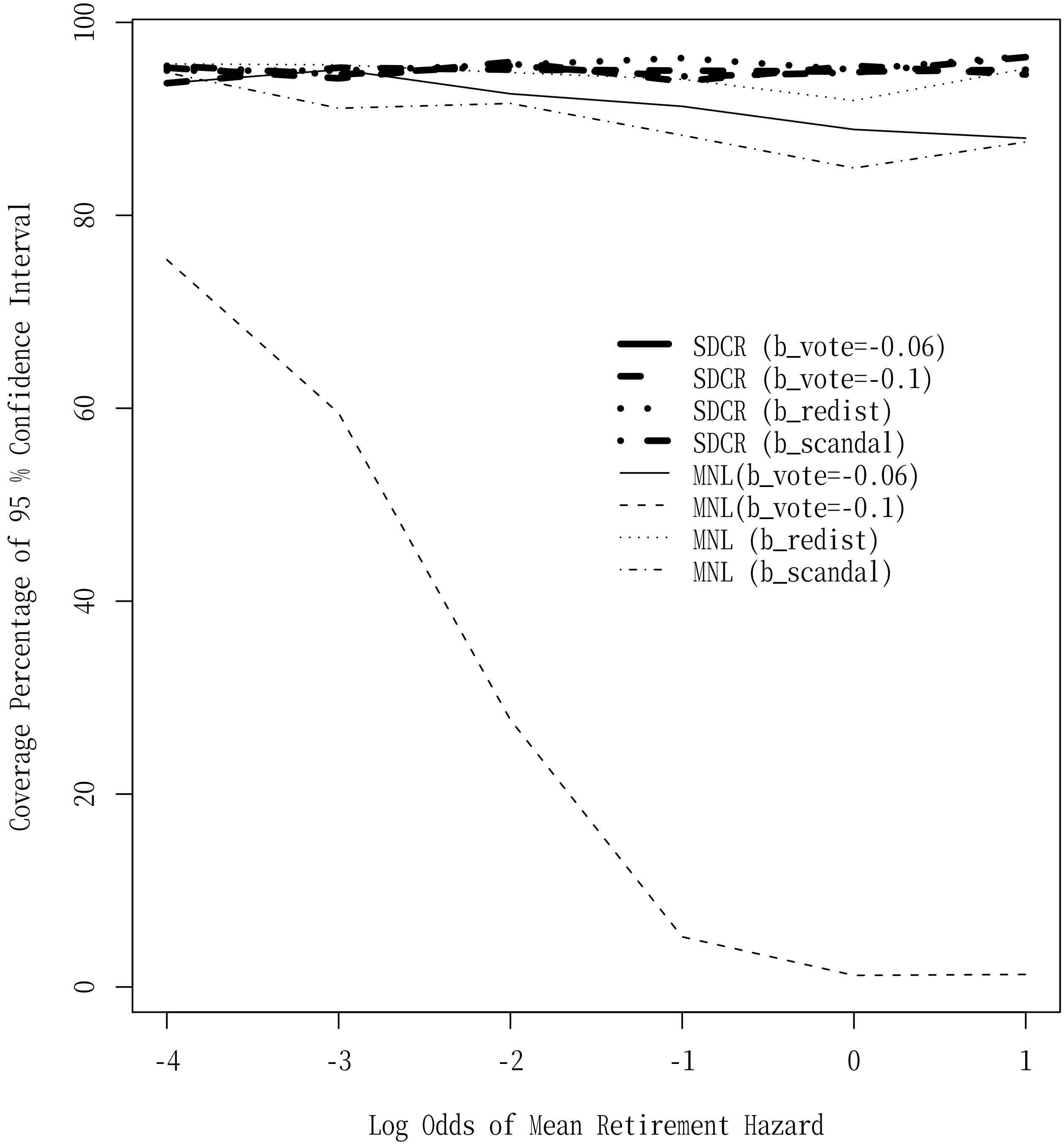


Figure 3: How Often (%) Are the True Coefficients in Their 95 % Confidence Intervals

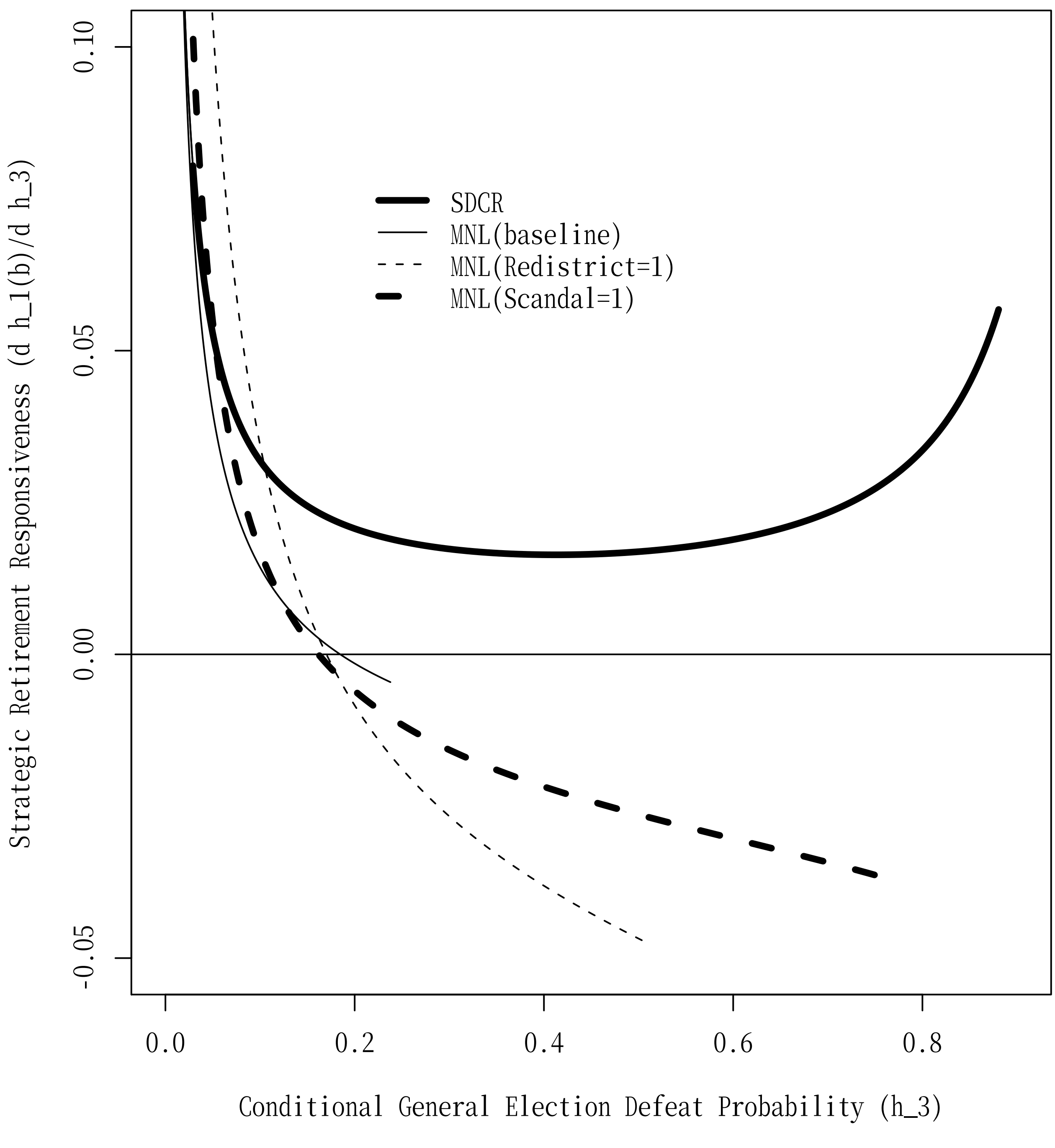


Figure 4: The Larger the Conditional General Election Defeat Probability (h_3), the Larger Strategic Retirement Responsiveness ($\frac{\partial h_{1(b)}}{\partial h_3}$)