Cure Models for Political Science Research[†]

Janet M. Box–Steffensmeier (corresponding author)
Department of Political Science
Ohio State University
Columbus, OH 43210–1373
steffensmeier.2@osu.edu

Roman Ivanchenko
Department of Political Science
Ohio State University
Columbus, OH 43210–1373
ivanchenko.1@osu.edu

Christopher Zorn
Department of Political Science
University of South Carolina
Columbia, SC 29208
zorn@sc.edu

Version 4.0
March 11, 2006
ABSOLUTELY NOT FOR CITATION

[†]Authors are listed alphabetically. Thanks to John Freeman and Dan Reiter for many helpful comments and discussions. Box–Steffensmeier also thanks the National Science Foundation for support via a Mid–Career Methodological Opportunities Grant, number SES–0083418, and Zorn thanks the John M. Olin Foundation for a Faculty Fellowship supporting this research.

Abstract

This is a test of the Emergency Abstract Placeholder System. This is only a test. The authors of this paper, in voluntary cooperation with LaTeX, have developed this system to act as a placeholder in lieu of an abstract. If this had been an actual abstract, this text would contain information about the content of the paper to follow. This concludes this test of the Emergency Abstract Placeholder System.

Introduction

Political scientists have become increasingly interested in longitudinal processes and their associated dynamics. At the same time, over the past several decades the discipline has seen large, longitudinal databases in comparative politics, international relations, and other fields become increasingly available. As a result of these developments, the pages of political science journals are increasingly filled with applications using duration models[1] to answer not only new and interesting questions across subfields, but also to reexamine and provide new insights into fundamental questions that have long intrigued the discipline.

But the widespread application of duration models requires analysts to carefully assess the assumptions made in their use, and to consider important variants that can lead to more accurate modeling of longitudinal political phenomena. In this paper we discuss one such characteristic of all widely–used duration models – that every observation in the data will eventually experience the event of interest – and describe a class of models, known as "cure" or "split population" models, that allow the researcher to relax this strong and often untenable assumption. ((The cure approach will allow making more accurate inferences in a variety of political science applications. For example, analyzing duration of a legislative response to judicial decisions by the US Supreme Court or a state supreme court using regular event-history methods implies that eventually a relevant legislative body will review all judicial decisions. In other words, regardless of when or how many cases a supreme court decides on, all of them are subject to legislative review. The assumptions behind the hazard function formulation suggest that even if a case was reviewed by the Supreme Court in 1805 and elicited no Congressional response in over 200 years, there is still positive probability of congressional review in the next few decades. Similarly, analyzing duration to war between different states using the regular Cox model or a parametric model, assumes that given a long

[1]Duration models are also referred to as survival models (in the medical and biostatistics literature), event history models (in sociology and criminology) and reliability models (in engineering).

enough period of time, any two nations will engage in a bloody military conflict against each other. Therefore, maybe not now, but sometime in the future, the Republic of Moldova will ignore the disapproval of the world community and commit its troops to battle against the Democratic People's Republic of Korea. The theoretical implausibility of these statements is obvious. Ignoring these considerations is likely to lead to inaccurate inferences and policy decisions. The methodological solution to such problems will be made clear in this paper.)) More specifically, in section 2, we present the mathematical details of a general parametric cure model followed by the description of semiparametric split population models. In section 3, we go on to illustrate the usefulness of such models with two applications, both of which involve processes in which, as a theoretical matter, we would not expect all observations to experience the event of interest. The first is congressional responses to Supreme Court decisions; the second example considers the democratic peace. Section 4 concludes with a discussion of other potential applications, and a review of issues relating to estimation and software.

Cure Models for Duration Data

Cure models originated in biostatistics (e.g. Boag 1949; Berkson and Gage 1952), in response to a common problem in clinical trials. Often, such studies are interested in assessing the effect of some treatment on the time until a patient suffers mortality or morbidity from some cause. If the treatment is particularly effective, the subjects may never experience a relapse – instead, the treatment has "cured" them of the malady. Similarly, in sociology the split population model "splits" the population into groups that experience the event of interest and those that do not. Under such circumstances, standard duration models require that the distribution for the density that makes up the hazard integrate to one, i.e., that all subjects in the study eventually experience the event of interest. In contrast, the intuition behind cure models is to allow part of the population to never experience the event of interest. As

we discuss below, this is typically accomplished through a mixture of a standard hazard density and a point mass at zero (Maller and Zhao 1996). That is, cure models estimate an additional parameter (or parameters) for the probability of an eventual failure, which can be less than one for some portion of the data, in as well as estimating the hazard of failure at any particular point in time.[2]

Cure models have seen widespread applications and extension in recent years in biostatistics and epidemiology, c(e.g. Farewell 1982; Aalen 1988, 1992; Kuk and Chen 1992; Longini and Halloran 1996; Tsodikov 1998; Tsodikov et. al. 1998; Sy and Taylor 2000). Such models have also been used in criminology, to model post-release recidivism rates among prisoners (e.g. Maltz and McCleary 1977; Schmidt and Witte 1989; Chung et. al. 1991), in sociology, to model family related events (e.g. Yamaguchi and Ferguson 1995; Diekmann and Engelhardt 1999), in labor economics, to model duration and instances of unemployment (e.g. Swaim and Podgurski 1994, Mavromaras and Orme 2004), in health economics, to model the effect of policy instruments on individual decisions to start or quit smoking (e.g. Douglas and Hariharand 1994; Nicholas 2002), and in finance, to analyze bank failure (e.g. Cole and Gunther 1995; Hunter et al 1996; DeYoung 2003). In all instances, the motivation for the models is identical: to allow for the possibility that some subset of the data will never experience the event of interest. Beck (1998, 203) provides a clear and intuitive summary of how cure models work, using the example of criminal recidivism: ((the likelihood of recidivating is a product of two components. The first component accounts for the observed returns to prison and is computed as the density of the duration to return to prison weighted by the probability that a person is a recidivist. The second component is the sum of the probability

[2]Split–population models can also be motivated as addressing a simple form of heterogeneity. Such heterogeneity is relatively common in social scientific studies, as well as in other areas where duration models are widely used (e.g. biometric and medical studies). Mendenhall and Hader (1958) and Harris et. al. (1981), for example, allow for two subpopulations where each has a different hazard rate but neither is zero. More generally, "frailty" models (e.g. Vaupel et. al. 1979; Vaupel and Yashin 1985) allow for more general forms of heterogeneity among individuals in the data.

that a person is not a recidivist and the probability that the return to prison is censored weighted by the probability that this person is a recidivist.))

> "With a probit model to estimate whether an individual is a recidivist, it is easy to write down the likelihood for, and then estimate, a split population duration model. All we need to do is write down the density of observed times of return to prison and the probability of a former prisoner never returning. The former is just the product of the probability that the individual is a potential recidivist times the density of returning at an particular time (using, say, the Weibull); the latter is just the sum of the probabilities that the individual is not a potential recidivist and the probability that he is but his return time is censored (again using the Weibull)" (Beck 1998, 203).

Parametric Models

We begin our discussion with a standard parametric model for continuous–time duration data.[3] We assume the duration defined by the event of interest $T_i > 0$ has a density function $f(T_i|\mathbf{X}_i, \theta)$ with $\mathbf{X}_i$ a k–dimensional vector of covariates and θ a parameter vector to be estimated. The corresponding cumulative distribution function is then $F(T_i|\mathbf{X}_i, \theta) = \mathbf{Pr}(\mathbf{T_i} \leq \mathbf{t_i}|\mathbf{X_i}, \theta), \mathbf{t_i} > \mathbf{0}$, where t_i represents the duration defined by the end of the "follow–up" period for observation i. We also define R_i to be the observable indicator of "failure," such that $R_i = 1$ when failure is observed and $R_i = 0$ otherwise. The associated survival function (defined as the conditional probability of survival to time T and denoted $S(T_i|\mathbf{X}_i, \theta)$ is then equal to $1 - F(T_i|\mathbf{X}_i, \theta)$. We can then write the hazard rate as:

$$h(T_i|\mathbf{X}_i, \theta) = \frac{f(T_i|\mathbf{X}_i, \theta)}{S(T_i|\mathbf{X}_i, \theta)} \tag{1}$$

[3]This section draws extensively on Schmidt and Witte (1989).

This value is akin to the conditional probability of a failure at time T given that no failure has occurred prior to T (cf. Box–Steffensmeier and Jones 1997).

Here we consider a model for the duration T which splits the sample into two groups, one of which will eventually experience the event of interest (i.e., "fail") and the other which will not. To do so, define a latent (unobserved) variable Y such that $Y_i = 1$ for those observations that will eventually fail and $Y_i = 0$ for those that will not; define $\Pr(Y_i = 1) = \delta_i$. The corresponding conditional density and distribution functions are then defined as:

$$f(T_i|Y_i = 1) = g(T, \theta) \tag{2}$$

$$F(T_i|Y_i = 1) = G(T, \theta), \tag{3}$$

while the corresponding density $f(T_i|Y_i = 0)$ and cdf $F(T_i|Y_i = 0)$ are undefined.[4]

For those observations that experience the event of interest during the observation period, we observe both $R_i = 1$ and their duration T_i. Since these observations also necessarily belong to the group in which $Y_i = 1$, we can write the unconditional density for these observations as:

$$\begin{aligned} L_i|R_i = 1 &= \Pr(Y_i = 1)\ Pr(T_i \leq t_i|Y_i = 1, \mathbf{X}_i, \theta) \\ &= \delta_i g(T_i|\mathbf{X}_i, \theta) \end{aligned} \tag{4}$$

Intuitively, this illustrates that the observed duration is a function of two components: the probability that the observation would be among those that would eventually experience the event of interest (that is, $\Pr(Y_i = 1) \equiv \delta_i$) and, conditional on $Y_i = 1$, the conditional hazard of failure at time T_i. In contrast, for those observations in which we do not observe an event (that is, where $R_i = 0$), this fact may be due to two possible conditions. On one hand, it is

[4]Because $Y_i = 0$ implies that the observation will never experience the event of interest (and thus the duration will never be observed), the probabilities for $f(T_i|Y_i = 0)$ and $F(T_i|Y_i = 0)$ cannot be defined.

possible that the observation in question is among those that will never experience the event defining the duration (that is, $Y_i = 0$). It may also be the case, however, that the observation will experience the event, but simply did not do so during the observation period (that is, $Y_i = 1$ but $T_i > t_i$). If, as is routinely the case, we assume that censoring is uninformative, the contribution to the likelihood for observations with $R_i = 0$ is therefore:

$$\begin{aligned} L_i|R_i = 0 &= \Pr(Y_i = 0) + \Pr(Y_i = 1)\Pr(T_i > t_i|Y_i = 1, \mathbf{X}_i, \theta) \\ &= (1 - \delta_i) + \delta_i S(T_i|\mathbf{X}_i, \theta) \end{aligned} \tag{5}$$

Combining these values for each of the respective sets of observations, and assuming independence across observations, the resulting likelihood function is:

$$\mathbf{L} = \prod_{i=1}^{N} [\delta_i g(T_i|\mathbf{X}_i, \theta)]^{R_i} [(1 - \delta_i) + \delta_i S(T_i|\mathbf{X}_i, \theta)]^{(\mathbf{1}-\mathbf{R_i})} \tag{6}$$

with the corresponding log-likelihood:

$$\ln \mathbf{L} = \sum_{i=1}^{N} R_i \{\ln(\delta_i) + \ln[g(T_i|\mathbf{X}_i, \theta)]\} + (1 - R_i) \ln[(1 - \delta_i) + \delta_i S(T_i|\mathbf{X}_i, \theta)] \tag{7}$$

The probability δ_i is typically modeled as a logit function, although other specifications (e.g. probit, complimentary log–log, etc.) are also possible, and can include a vector of explanatory variables (say, $\mathbf{Z_i}$):[5]

$$\delta_i = \frac{\exp(\mathbf{Z_i}\gamma)}{1 + \exp(\mathbf{Z_i}\gamma)} \tag{8}$$

The hazard function for those who do experience the event may be any of the commonly–used parametric distributions (e.g. exponential, Weibull, log-logistic, etc.).[6]

[5]Note that this model is identified even when the variables in δ_i are identical to those in the model of duration. This means that one can test for the effects of the same set of variables on both the incidence of failure and the duration associated with it (Schmidt and Witte 1989).

[6]Scholars are actively working on semi- and nonparametric approaches for estimating the latency, but these efforts have been hampered by identifiability problems. Particularly promising, in our opinion, is

Semiparametric Cure Models

Over the last several years, there has been an increased interest and use of semiparametric statistical models (Kuk and Chen 1992; Taylor 1995; Peng 2003a; Peng and Carriere 2002; Peng, Dear, Carriere 2001; Li and Taylor 2002). Just as the semiparametric estimation techniques came to dominate the parametric survival analysis, practitioners interested in failure data with a proportion that will never fail became interested in departing from the stringent assumptions associated with the use of the parametric models cure models. Cox's partial likelihood method that allows estimating the non-cured fraction of observations without specifying the distribution of the baseline hazard became a plausible alternative to the parametric cure models (Peng and Carriere 2002).

The most appealing aspect of the semiparametric survival models is doing away with imposing parametric assumptions on the hazard rate:

$$h(T_i|\mathbf{X}_i, \theta) = h_0(T_i)\exp(\xi_\mathbf{i}) \tag{9}$$

In this equation, the hazard rate depends on the arbitrary, unspecified baseline hazard, $h_0(T_i)$, and on the function of X and θ, $\xi_\mathbf{i}$. A similar transformation is done to the survival function:

$$S(T_i|\mathbf{X}_i, \theta) = S_0(T_i)^{\exp(\xi_\mathbf{i})} \tag{10}$$

Subject is survival depends on the unspecified baseline survival rate, $S_0(T_i)$, and on the function of covariates, $\xi_\mathbf{i}$. Consequently, we can rewrite the likelihood function to include the unspecified baseline hazard and survival functions. However, the researchers specializing

recent work by Sy and Taylor (2000) who use the EM algorithm and a zero-tail constraint to reduce the near non-identifiability of the problem. There simulation shows that their methods "are competitive to the parametric methods under ideal conditions and are generally better when censoring from loss to follow-up is heavy" (2000, 227). See also work by Kuk and Chen (1992), Taylor (1995), and Peng and Dear (2000). **NOTE: This is no longer accurate: see, e.g., Peng (2003) and the next section.**

in the area of semiparametric cure models tend to include an additional parameter, c_i (Peng and Dear 2000; Peng 2003a). This parameter serves as an indicator function: it takes on the value of 1 if an observation will never experience failure (it is cured), and it takes on the value of 0 otherwise. Of course, whenever $R_i = 1$, $c_i = 0$. Thus, the indicator only affects the part of the likelihood that is responsible for the censored cases. Notice the difference between δ_i and c_i. The former is the probability of never experiencing the event, and the latter is the indicator of belonging to either cured or a non-cured group. The resulting likelihood function is:

$$\mathbf{L} = \prod_{i=1}^{N} [\delta_i g(T_i|\mathbf{X}_i, \theta)]^{R_i} [(1-\delta_i)^{c_i} + \{\delta_i S(T_i|\mathbf{X}_i, \theta)\}^{\mathbf{1-c_i}}]^{(\mathbf{1-R_i})} \tag{11}$$

An obvious problem with this set-up is the absence of data on c_i. To overcome this problem, the EM algorithm is used[7]. The E step of the algorithm calculates the expected value of c for each observation. The M step step splits estimation into two parts and allows separate modelling of δ_i probabilities (as shown above) and modelling of the duration part of the analysis using the partial likelihood method.[8]

In a recent study, Peng and Carriere (2002) compared performance of the parametric and semiparametric cure models. They concluded that even in cases when the underlying hazard rate follows a particular distribution both types of models produce comparable results. However, given that the researchers are unable to determine the distribution underlying the probability of failure, the use of the semiparametric cure models is all the more appropriate.

Unfortunately, the EM procedures do not usually appear in the popular statistical software packages. Additionally, the conventional EM estimation procedure does not easily produce standard errors. The recent efforts to obtain the information matrix for the semi-

[7]Lu and Ying (2004) proceed down the path of the Generalized Estimating Equations. They construct an algorithm that computes parameters sequentially per each iteration

[8]For a more extensive exposition of the expectation and maximization steps see Peng and Dear (2000) and Peng (2003a).

parametric cure models required a lot of computing power and considerable programming skills. Peng and Dear (2000) suggest obtaining the information matrix using Louis's (Louis 1982) approximation. To do so they had to resort to multiple imputations. Sy and Taylor (2000) had to create a separate routine for the estimation of the standard errors, and Peng (2003a) used the bootstrap method which can be computationally demanding if with complicated models (as is the case). Thus, it may appear that, although estimating the semiparametric cure models allows to relax the assumptions about the distributional nature of the underlying hazard rate, it comes at an expense of the computational intensity and not yet user-friendly algorithms. Fortunately, it is possible to avoid the problems associated with the EM algorithm and elusive standard errors. To do so, one has to take advantage of the equivalence between the counting process models and survival models. This would allow working with the semiparametric cure models using regular statistical packages.

For the sake of completeness, it is necessary to mention the issues of identifiability often surrounding semiparametric cure models. The main difficulty arises from the fact that the censored observations with survival times larger than the greatest time-to-event period are not directly distinguishable from the cured observations (which also indicates the presence of cured observations). This suggests that some observations that do not belong to the cured group will have a zero probability of experiencing the event (Peng 2003a), which results in an improper distribution. Visually, the right tail of the survival probability function will not be zero: it will be flat after the largest time-to-event period. Fortunately, it is relatively easy to correct this problem. Taylor (1995) suggests imposing a constraint of zero probability of survival for all censored patients whose survival time is greater than the largest time-to-event period. Visually, it would result in a drop of the probability of survival to zero, when it is past the largest time-to even period. Sy and Taylor (2000) argue that this approach virtually eliminates the identifiability problem. Li et al. (2001) point out that when the

probability of being uncured is modelled as a logit with covariates, the Taylor approach results in identifiability. [9]

Incorporating Time-Varying Covariates

This is where we talk about how one can estimate a cure model with time-varying covariates by taking advantage of the equivalence between Poisson/counting process models and survival analysis. Specifically, a "zero-inflated" Poisson model (cf. Lambert 1992; Zorn 1998) (which is really nothing more than a Poisson model mixed with a point mass at zero) which includes period-specific dummies can exactly estimate a discrete-time cure model with time-varying covariates, in the same fashion that a standard Poisson GLM with such dummies can replicate a Cox survival model.

Some relevant bullet points (from my ICPSR notes) include:

- Cox / Poisson equivalence (Whitehead 1980; Laird and Oliver 1981).
- "Zero–Inflated" Poisson model:
  - Latent probability of only observing zeros $= p_i^*$.
  - Binary realization of p_i^* is $p_i \in \{0, 1\}$.
  - Underlying count variable $Y_i^* \sim \text{Poisson}(\lambda_i)$ only observed if $p_i = 1$.

Probability of a zero outcome:

$$\begin{aligned} \Pr(Y_{it} = 0) &= \Pr(p_{it} = 0) + [\Pr(p_{it} = 1) \times \Pr(Y_{it}^* = 0)] \\ &= (1 - p_{it}^*) + p_{it}^*[\exp(-\lambda_{it})] \end{aligned}$$

[9] Peng (2003b) offers an interesting alternative. Rather then using the Taylor approach of constraining the survival probability past the largest time-to-event period to 0, he suggests smoothing out the right tail by parametricaly modelling this survival part. His "completion of the tail," smooth decrease of the survival probability to zero, results in a proper distribution.

Probability of a non–zero outcome:

$$\begin{aligned} \Pr(Y_{it} = y) &= \Pr(p_{it} = 1) \times \Pr(Y_{it}^* = y) \\ &= p_{it}^* \times \frac{\exp(-\lambda_{it})\lambda_{it}^y}{y!} \end{aligned}$$

with

$$\mathrm{E}(Y_{it}^*) \equiv \lambda_{it} = \exp(\mathbf{X}_{it}\boldsymbol{\beta})$$

and

$$\Pr(p_{it} = 1) = \frac{1}{1 + \exp(-\mathbf{Z}_{it}\gamma)} \text{ or } \Phi(\mathbf{Z}_{it}\gamma)$$

The likelihood and so forth are straightforward. Estimation can be a bit problematic, particularly if there is significant overlap between $\mathbf{X}$ and $\mathbf{Z}$, but in general is pretty well–behaved.

Practical Considerations

The use of cure models such as that outlined above should be considered whenever all observations cannot reasonably be assumed to "fail" at some point in the future. A particularly useful property of cure models is that they allow for separate estimation of the influence of covariates on the probability of experiencing the event from their effect on the time until the event of interest occurs for those observations that do experience the event. That is, covariates can have an independently positive or negative influence, or no effect at all, on both the incidence and the latency of an event. This fact makes cure models more flexible than other duration models; one may find that a particular covariate affects incidence but not latency, or vice versa. Such interpretations are not available with other duration models.

Analysts should thus consider using cure models whenever there is a theoretical reason to suspect that not all observations will eventually "fail." Such an assessment is relatively

straightforward and more common than the political science literature currently reflects. Political scientists have simply not been routinely asking whether all observations are expected to eventually fail, but they should. For example, in the study of state policy adoption, an area in political science with widespread use of duration models, one would not expect that all states would adopt particularly policies, whether they are substantively about lotteries, abortion, etc.

In addition to theoretical considerations, one can empirically look at the data to get a general sense of the need for relaxing the assumption made by all other duration models, i.e., that eventually all observations experience the event. By plotting a Kaplan Meier (KM) figure of the survivor function versus time (Kaplan and Meier 1958), the analyst will gain a sense of whether observations in the data exist that will not experience the event of interest. Price (1999) and Sy and Taylor (2000) illustrate the use of a KM survival curve to empirically assess the need for a split population model. If it "shows a long and stable plateau with heavy censoring at the tail," there is strong reason to suspect that there is a subpopulation that will not experience the event (Sy and Taylor 2000, 227; see also Peng et al. 2001).

Several issues arise in the estimation and interpretation of cure models. Note, for example, that when $\delta_i = 1 \forall i$ (that is, when all observations will eventually fail), the likelihood reduces to that for a standard duration model with censoring. However, testing for $\delta_i = 1$ is a case of a boundary condition[10] and thus standard asymptotic theory does not apply (Price 1999). Maller and Zhou (1996) offer a corrected likelihood–ratio test for the proposition that all observations will eventually experience the event of interest. Similarly, Peng et al (2001) examine the likelihood–ratio test in gamma, Weibull, and log-normal cure models, and provide corrections for those times when the assumptions of the test may not be completely appropriate (light censoring and/or large hazard rate). Morbiducci et al. (2001) offer examination of the log-odds residuals, whose distribution converges to the logistic dis-

[10]Note that this does not correspond to the case of $Z_i\gamma = 0$ (which yields $\delta_i = 0.5$).

tribution, when considering parametric cure models. If after estimating a regular (not cure) parametric model, the log-odds residuals are distributed bimodaly, rather than resembling a logistic distribution, it signifies the presence of two distinct groups.One groups experiences events early, and the other one is likely to experience the event late or not at all, which suggests that one should consider estimating a cure model. Issues of goodness-of-fit for split population models are an important, but currently ongoing, area of research (e.g., Sy and Taylor 2000). Finally, one can also test $H_0 : \delta = 1$ (i.e., if the assumption that all observations will eventually fail is true) statistically. If so, the equation reduces to the standard general duration model with censoring; that is, a (e.g.) Weibull model is a special case of the Weibull cure model.

Similarly to the argument above, a researcher who decides to estimate a parametric cure model may choose to specify the latency probability as Generalized F. Peng et al (1998) argue that because Generalized F nests within itself a whole range of distributions (normal, extended generalized gamma, extreme value etc.) one can tests its parameters to determine if a particular distribution should be used.[11] If the tests show that the Extended Generalized Gamma (EGG) distribution should be used, then one can estimate the EGG cure model. Itself, the EGG nests such distribution as Weibull, and Log-Normal, and it is possible to use regular likelihood ratio tests and AIC to select the most appropriate distribution. Due to the boundary problems, such comparison between the Generalized F and the distributions nested within the EGG is not appropriate (Peng et al 1998). Unfortunately, estimating the Generalized F model suffers from considerable computing complexities.

Morbiducci et al. (2001) suggests the use of the log-odds residuals to evaluate the goodness of fit. They suggest extracting the log-odds residuals, correcting them due to the presence of a cured fraction of observation, and visually examining them. If the plot of

[11]The standard likelihood ratio test may not be appropriate. Peng et al (1998) propose a simulated empirical chi-square distribution to conduct their test of the Extended Generalized Gamma.

the residuals resembles the shape of the logistic distribution, then the parametric model is likely to be appropriate. If the plot of the log-odds residuals does not correspond to the logistic distribution, one want to consider using a different parametric specification for the cure model.

(We probably need more here...).

Applications

Congressional Responses to Supreme Court Decisions, 1979–88

Our first example illustrates the potential split–population models offer for providing greater insight into political processes than do conventional duration models. We take as our example here one aspect of the separation of powers: specifically, the issue of Congressional responses to decisions of the U.S. Supreme Court, in the form of bills, hearings, or other kinds of formal actions taken in response to Court decisions. Such responses have two signature characteristics: they are typically taken to modify or reverse the Court's decisions, and they are relatively rare. In our data,[12] for example, only 132 of the 7033 decisions under scrutiny (1.9 percent) were the target of Congressional responses. But while it is unlikely that most Court decisions will ever be subject to Congressional scrutiny, scholars remain interested in those cases which will, and in the conditions under which those responses occur. Split-population models are ideal for this kind of analysis, in that they allow us to separate the effects of case-specific independent variables on the probability of the case ever being subject to Congressional scrutiny (the incidence) from the timing of that response (the

[12]Specifically, we examine Congressional responses to the decisions of the Warren and Burger Courts, i.e., the 1953 to 1985 terms, taken during the 96th-100th Congresses (1979-1988), as reported in Eskridge (1991). There are 7157 such cases; omitting 124 cases because of missing data (mostly on the *Liberal Decision* variable) yields a total of 7033 cases for analysis. For a more thorough analysis of such responses, see Zorn and Caldeira 1995; for a similar split-population analysis of successful Congressional overrides of Supreme Court statutory decisions, see Hettinger and Zorn 2005.

latency). Theoretically, there is a strong case for using the split population model since we do not expect Congress to respond to every Court decision with hearings, bills, etc.

Our dependent variable is the duration in years between the decision of the Court and the first Congressional response.[13] As noted above, only 2 percent of the cases in our data experience such responses; the remainder are censored. Figure 1 shows the Kaplan Meier survival curve. We see that the survival curve quickly levels off at a nonzero value at its right extreme, which provides empirical evidence that there is a substantial portion of the population that will not experience the event, i.e., that is "immune."

We examine the influence of a number of independent variables on the hazard of a response, including the year of the decision itself; the presence (coded 1) or absence (coded 0) of disagreement in the lower court decision, alteration of precedent by the Supreme Court, a declaration of unconstitutionality, or a liberal policy decision by the Court; the number of briefs amicus curiae filed on the merits in the case; and a series of indicator variables for the nature of the losing party to the case (federal, state, and local governments, businesses, class action litigants, and natural persons, with non-profit groups omitted as the baseline category).[14] In general, these variables indicate the salience of the case to members of Congress, either through their inherent importance in the constitutional system or their impact on important constituent groups or actors in the Congressional arena. We estimate two models: a standard log–logistic hazard model, and a split–population model that is log–logistic in the duration and uses a probit link for the probability of no response. Results of these estimates are presented in Table 1.

The standard log-logistic model results indicate that only two variables (year of decision and amicus briefs) significantly affect the hazard of a response, though several others (lower

[13]Here we examine only the time to the first event, even though in some cases more than one response occurs. See Box-Steffensmeier and Zorn (2002) on the issue of models for repeated events.

[14]A thorough discussion of these covariates and their expected effects can be found in Zorn and Caldeira (1995).

court disagreement, declaration of unconstitutionality, and state and natural person losing parties) are of marginal significance ($p < .10$, two-tailed). The estimated $\hat{\sigma}$ parameter, which (as in the Weibull model) indicates the extent of duration dependence, is not significantly different from 1.0 ($z = 0.84$), indicating that the hazard of a Congressional response, conditional on the independent variables and coefficient estimates, remains relatively constant over time. More important than the individual variable results, however, is the fit of the model to the data. The combination of high censoring/low hazards and the assumption that all observations will eventually "fail" results in a predicted median survival time for this model of nearly 710 years, clearly a suboptimal fit to the data.

In contrast, the cure model presented in columns 2 and 3 presents a somewhat different picture of the Congressional response data. The first column indicates the probability of a case being essentially "immune;" i.e., of its never being addressed by Congress, while the second shows the effects of the covariates on the (log of the) duration until such a response occurs, given that the case is among those for which a response is possible. The results are revealing: in most instances, we expect (and find) that the signs in the two parts of the model will be the same (i.e., variables which decrease the probability of a response also serve to increase the duration until such a response is forthcoming). So, for example, the presence of amicus curiae briefs both significantly increases the probability of a Congressional response, and decreases the length of time until that response occurs. At the same time, other variables appear to work at cross-purposes: liberal decisions by the Court, for example, are both less likely to be addressed, but also see responses more rapidly than do conservative cases, when they occur. Likewise, more recent decisions are both more likely to go ignored by Congress, but are also addressed more rapidly when such responses occur, than are older cases, though this result is likely due to older cases having greater "exposure" to response than more recent decisions.

Also important is the improvement in model fit gained by the split–population model. In contrast to the standard model, the split-population model predicts that the average long-term probability of a response is 0.675, while the median predicted survival time is reduced by 50 percent (to 355 years). It is clear that the split-population model fits the data better than the standard duration model. The differences in these models are illustrated in Figure 3, which plots the predicted forty-year survival probabilities for a "median" case[15] for each of the two models. Conditional on a case being part of the population that experiences the event, i.e., the "nonimmune population," the estimated survival rates are significantly less than for the general model, suggesting that separation of likely from unlikely cases for response provides estimates that yield better leverage on the long-term probability of Congressional action.

International Conflict

This is where the international conflict stuff goes, including a discussion of Table .....

I suspect I/we can write this up quickly – I use it as an example at ICPSR when I teach there, and the data certainly ought to be familiar enough to us ... :-)

This is an especially nice substantive example of the cured–fraction approach, I think, and of the transparency between counting process models (like the Poisson) and survival analysis.

I should also note that we have cool figures for all of these; they're just not incorporated into this version of the paper yet....

Frailty Models with Cured Fraction (this is mentioned in conclusion)

(If this section stays, I will add cited works to the bibliography).

[15] That is, a 1972 decision with one amicus brief and zeros on all other independent variables.

A natural extension of a survival cure model is to model the cured fraction by including heterogeneity. That is, rather than assuming that the only source of heterogeneity is based on whether an observation belongs to a cured or a susceptible group, a researcher may consider a possibility that the members of the cured fraction are not homogenous. Thus, a complete picture of heterogeneity (frailty) needs to include both, differences in susceptibility and differences within the susceptible group. Unfortunately, this may create a problem. Whenever one is modeling the presence of a cured fraction, he or she assumes that certain observations will have zero probability of experiencing the event. In other words, there has to be a point probability mass at zero and a continuous part that accounts for heterogeneity of those who do experience the event. Since most popular frailty distributions, gamma, inverse normal, and positive stable, are continuous, they do not allow for a probability (a point mass) of no risk (Price and Manatunga, 2001). One way of solving this probelm is offered by Aalen (1992), who, using Hougaard's (1986) findings, argues that employing the compound Poisson distribution as the distribution of the frailty term allows modelling this combination of discrete-continuous random effect. **(An obscene amount of math can be placed here)** The compound Poisson distribution is obtained by assuming that one's frailty depends on the sum of individual shocks, where the total number of shocks is Poisson distributed. Each shock is gamma distributed. If a case belongs to the cured fraction, then it experiences zero gamma distributed shocks. If a case does not belong to the cured fraction, then it does experience a particular number of gamma distributed shocks. Aalen (1986) suggests that one can think of the compound Poisson distribution as being a subject to a random number of shocks (Poisson distribution), where each shock is also random (gamma distributed). This method allows modelling a discrete probability mass of zero risks and a continuous part of the non-cured heterogeneous fraction. Presently the work in this area is being done by Aalen and Tretly (1999), Moger et al. (2004), and Moger and Aalen (2005).

Conclusion

In summary, cure models offer the potential for substantial improvements in the manner in which political scientists study duration data. In many cases in the social sciences, it is unrealistic to believe that all observations will eventually experience the event of interest. At a minimum, scholars need to ask whether the strong assumption of standard duration models apply. That is, is it reasonable to expect that eventually all observations will experience the event of interest? If not, a model accounting for a cured fraction is needed.

(**This is mentioned above**): Also, innovative modeling research by Price (2000) accounts for the fact that of those who do experience the event are still a heterogeneous group. That is, the split population model accounts for heterogeneity by separating those who experience the event and those who do not. In addition, Price accounts for remaining heterogeneity among the subpopulation that experiences the event through the use of a frailty model. She refers to this innovation as a frailty-cure model (see also Firth et al. 1999; Wienke et al. 2003).)

Also: Spatial stuff (Banerjee and Carlin 2004).

References

Banerjee, Sudipta, and Bradley P. Carlin. 2004. "Parametric Spatial Cure Rate Models for Interval–Censored Time-to-Relapse Data." *Biometrics* 60(1):268–75.

Beck, Nathaniel. 1998. "Modelling Space and Time: The Event History Approach," in E. Scarbrough and E. Tanenbaum, eds., *Research Strategies in the Social Sciences.* New York: Oxford University Press, 191–213.

Beck, Nathaniel, Jonathan N. Katz and Richard Tucker. 1998. "Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable." *American Journal of Political Science* 42(October):1260-88.

Berkson, J. and R. P. Gage. 1952. "Survival Curve for Cancer Patients Following Treatment." *Journal of the American Statistical Association* 47:501–15.

Boag, J. W. 1949. "Maximum Likelihood Estimates of the Proportion of Patients Cured by Cancer Therapy." *Journal of the Royal Statistical Society, Series B* 11(1):15–44.

Box–Steffensmeier, Janet M. and Bradford Jones. 1997. "Time is of the Essence: Event History Models in Political Science." *American Journal of Political Science* 41(October):1414–61.

Box–Steffensmeier, Janet M., Peter Radcliffe, and Brandon Bartels. 2005. "The Incidence and Timing of PAC Contributions to Incumbent U.S. House Members, 1993–94." *Legislative Studies Quarterly* 30:549–79.

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

Clark, David H., and Patrick M. Regan. 2003. "Opportunities to Fight: A Statistical Technique for Modeling Unobservable Phenomena." *Journal Conflict Resolution* 47(February):94-115.

Cole, Rebel A. and Jeffrey W. Gunther. 1995. "Separating the Likelihood and Timing of Bank Failure." *Journal of Banking and Finance* 19:1073-1089.

Chung, Ching-Fan, Peter Schmidt, and Ann D. Witte. 1991. "Survival Analysis: A Survey." *Journal of Quantitative Criminology* 7(1):59–98.

DeYoung, Robert. 2003. "The Failure of New Entrants in Commercial Banking Markets: A Split-Population Duration Analysis." *Review of Financial Economics* 12:7-33.

Diekmann, Andreas and Henriette Engelhardt. 1999. "The Social Inheritence of Divorce: Effects of Parent's Family Type in Postwar Germany." *American Sociological Review* 64:783-793.

Douglas, Stratford and Govind Hariharand. 1994. "The Hazard of Starting Smoking: Estimates from a Split Population Duration Model." *Journal of Health Economics* 13:213-230.

Farewell, V. T. 1982. "The Use of Mixture Models for the Analysis of Survival Data with Long-Term Survivors." *Biometrics* 38:1041–46.

Firth, David, Clive Payne, and Joan Payne. 1999. "Efficacy of Programmes for the Unemployed: Discrete Time Modelling of Duration Data from a Matched Comparison Study." *Journal of the Royal Statistical Society* Part 1: 111–20.

Greene, William H. 2003. *Econometric Analysis, 5th Ed.* Upper Saddle River, NJ: Prentice-Hall.

Harris, C.M., Kaylan, A.R., and Maltz, M.D. 1981. "Refinements in the Statistics of Recidivism Measurement." In Fox, J.A., ed. *Models in Quantitative Criminology.* New York: Academic Press.

Hettinger, Virginia, and Christopher Zorn. 2005. "Explaining the Incidence and Timing of Congressional Responses to the U.S. Supreme Court." *Legislative Studies Quarterly* 30(February):5–28.

Hunter, W. C., J. A. Verbrugge, and D. A. Whidbee. 1996. "Risk Taking and Failure in De Novo Savings and Loans in the 1980s." *Journal of Financial Services Research* 10:235-272.

Kuk, A. Y. C. and C. H. Chen. 1992. "A Mixture Model Combining Logistic Regression with Proportional Hazards Regression." *Biometrika* 79:531–41.

Laird, Nan, and Donald Oliver. 1981. "Covariance Analysis of Censored Survival Data Using Log-Linear Analysis Techniques." *Journal of the American Statistical Association* 96(June):231-40.

Lancaster, Tony. 1990. *The Econometric Analysis of Transition Data.* New York: Cambridge.

Li, Chin Shang and Jeremy M. G. Taylor. 2002. "A Semi-Parametric Accelerated Failure Time Cure Model." Statistics in Medicine 21(November):3235–47.

Li, Chin Shang, Jeremy M. G. Taylor, and Judy P. Sy. 2001. "Identifyability of Cure Models." *Statistics and Probability Letters* 54: 389-95.

Lu, Wenbin, and Zhiliang Ying. 2004. "On Semiparametric Transformation Cure Models." *Biometrika* 91(2):331–43.

Louis, Thomas A. 1982. "Finding the Observed Information Matrix Using the EM Algorithm." *Journal of the Royal Statistical Society, Series B.* 44(2):226-233

Maller, R. A., and Zhou, S. 1992. "Estimating the Proportion of Immunes in a Censored Sample." *Biometrika* (79):731–39.

Maller, R. A., and S. Zhou. 1996. *Survival Analysis with Long-Term Survivors.* New York: Wiley.

Maltz, M.D., and R. McCleary. 1977. "The Mathematics of Behavioral Change: Recidivism and Construct Validity." *Evaluation Quarterly* 1(August): 421–38.

Mavromaras, Christos G., and Chris D. Orme. 2004. "Temporary Layoffs and Split Population Models." *Journal of Applied Econometrics* 19:49-67.

Mendenhall, W., and Hader, R.J. 1958. "Estimation of Parameters of Mixed Exponentially Distributed Failure Time Distributions from Censored Life Data." *Biometrika* 45: 504–20.

Morbiducci, Marta, Alessandra Nardi, and Carla Rossi. 2003. "Classification of "Cured" Individuals in Survival Analysis: The Mixture Approach to the Diagnostic–Prognostic Problem." *Computational Statistics and Data Analysis* 41: 515-529

Nicholas, Angel Lopez. 2002. "How Important are Tobacco Prices in the Propensity to Start and Quit Smoking? An Analysis of Smoking Histories from the Spanish National Health Survey." *Health Economics* 11:521-535.

Peng, Yingwei. 2003a. "Fitting Semiparametric Cure Models." *Computational Statistics and Data Analysis* 41:481–90.

Peng, Yingwei. 2003b. "Estimating Baseline Distributions in Proportional Hazards Cure Models." *Computational Statistics and Data Analysis* 42:187–201.

Peng, Yingwei, and K. C. Carriere. 2002. "An Empirical Comparison of Parametric and Semiparametric Cure Models." *Biometrical Journal* 44:1002–14.

Peng, Yingwei, and Keith B. G. Dear. 2000. "A Nonparametric Mixture Model for Cure Rate Estimation." *Biometrics* 56(March):237–43.

Peng, Yingwei, Keith B. G. Dear, K. C. Carriere. 2001. "Testing for the Presence of Cured Patients: A Simulation Study." *Statistics in Medicine* 20: 1783–96.

Peng, Yingwei, Keith B. G. Dear, J.W. Denham. 1998. "A Generalized F Mixture Model for Cure Rate Estimation." *Statistics in Medicine* 17: 813-830.

Schmidt, Peter and Anne D. Witte. 1989. "Predicting Recidivism Using 'Split-Population' Survival Time Models." *Journal of Econometrics* 40(1):141–59.

Sposto, Richard. 2002. "Cure Model Analysis in Cancer: An Application to Data from the Children's Cancer Group." *Statistics in Medicine* 21(2):293-312.

Swaim, Paul and Michael Podgurski. 1994. "Female Labor Supply Following Displacement: A Split Population Model of Labor Force Participation and Job Search." *Journal of Labor Economics* 12:640-656.

Sy, Judy P., and Jeremy M.G. Taylor. 2000. "Estimation in a Cox Proportional Hazards Cure Model." *Biometrics* 56(March):227–36.

Taylor, Jeremy M.G. 1995. "Semi-Parametric Estimation in Failure Time Mixture Models." *Biometrics* 51:899–907.

Thall, P. F. 1988. "Mixed Poisson Regression Models for Longitudinal Interval Count Data." *Biometrics* 44:197–209.

Therneau, Terry M. 1997. "Extending the Cox Model." *Proceedings of the First Seattle Symposium in Biostatistics*. New York: Springer–Verlag.

Tsodikov, A. 1998. "A Proportional Hazards Model Taking Account of Long-Term Survivors." *Biometrics* 54:1508–15.

Tsodikov, A., M. Loeffler and A. Yakovlev. 1998. "A Cure Model with Time-Changing Risk Factor: An Application to the Analysis of Secondary Leukaemia." *Statistics in Medicine* 17:27–40.

Vaupel, J. W., K. G. Manton, and E. Stallard. 1979. "The Impact of Heterogeneity in Individual Frailty Models and the Dynamics of Mortality." *Demography* 16:439–54.

Vaupel, J. W., and A. I. Yashin. 1985. "The Deviant Dynamics of Death in Heterogeneous Populations." In *Sociological Methodology* (E. F. Borgatta, ed.) 179–211. San Francisco: Jossey-Bass.

Weinke, Andreas, Paul Lichtenstein, and Anatoli I. Yashin. 2003. "A Bivariate Frailty Model with a Cure Fraction for Modeling Familial Correlations in Diseases." *Biometrics* 59(December):1178–83.

Yamaguchi, Kazuo and Linda R. Ferguson. 1995. "The Stopping and Spacing of Childbirths and Their Birth-History Predictors: Rational-Choice Theory and Event-History Analysis." *American Sociological Review* 60:272-298.

Zorn, Christopher. 1998. "An Analytic and Empirical Examination of Zero-Inflated and Hurdle Poisson Specifications." *Sociological Methods and Research* 26(February):368–400.

Table 1: Models of Congressional Responses to Supreme Court Decisions, 1979-1988

	Standard	Cure Model	
Variable	Log-Logistic	Pr(No Response)	Duration
(Constant)	15.827**	-8.074*	18.190**
	(1.109)	(3.645)	(1.612)
Year of Decision	-0.128**	0.105*	-0.169**
	(0.013)	(0.043)	(0.020)
Lower Court Disagreement	-0.349	-1.900*	0.300
	(0.203)	(0.779)	(0.321)
Declaration of Unconstitutionality	1.002	-1.881	1.512
	(0.578)	(2.309)	(0.904)
Liberal Decision	-0.121	1.869**	-0.794*
	(0.206)	(0.591)	(0.332)
Number of Amicus Curiae Briefs	-0.083**	-0.169**	-0.054*
	(0.021)	(0.065)	(0.023)
Federal Government Loser	-0.230	-3.338	0.599
	(0.285)	(2.144)	(0.496)
State Government Loser	0.603	-2.639	1.509*
	(0.348)	(1.972)	(0.622)
Local Government Loser	-0.101	-6.059**	1.066
	(0.381)	(2.134)	(0.558)
Business Loser	-0.264	-1.250*	0.237
	(0.235)	(0.518)	(0.357)
Class Action Loser	-0.480	-1.315	-0.041
	(0.356)	(1.364)	(0.456)
Natural Person Loser	0.585	1.943**	-1.218
	(0.309)	(0.643)	(0.713)
$\hat{\alpha}$	1.099**	1.117**	
	(0.129)	(0.138)	
lnL	-730.31	-707.27	

Note: $N = 7033$. Cell entries are MLEs; standard errors are in parentheses. One asterisk indicates $p < .05$, two indicate $p < .01$ (two-tailed). See text for details.

Table 2: Cox and Cure Models of International Conflict, 1950–1985

		Cure Model	
Variable	Cox Model	Cure	Hazard
(Constant)	–	-0.172	-3.815
		(0.741)	(0.230)
Democracy	-0.382	0.567	-0.416
	(0.110)	(0.813)	(0.173)
Contiguity	0.949	-3.189	0.864
	(0.154)	(1.946)	(0.201)
Capability Ratio	-0.223	0.994	0.083
	(0.079)	(0.361)	(0.053)
Economic Growth	-3.695	-11.030	-5.168
	(1.169)	(5.712)	(1.437)
Alliance	-0.348	-15.864	-0.743
	(0.147)	(5.548)	(0.179)
Trade	-3.229	15.158	-16.474
	(10.621)	(27.135)	(14.798)

Note: $NT = 20448$. Entries are coefficient estimates; standard errors are in parentheses. Cure model findings do not include fixed effects by period; see text for details.

Table 3: Weibull Models of Democratic Survival, 1960-1992

		Cure Model	
	Weibull	Cured	
Variable	Model	Fraction	Hazard
(Constant)	8.10	-24.4	0.93
	(2.73)	(8.1)	(1.20)
Conflict (MIDs)	-0.25	0.40	0.32*
	(0.33)	(0.57)	(0.14)
Regional Democracy	4.07	-5.60	0.77
	(2.87)	(6.02)	(1.05)
ln(GDP Per Capita)	-1.86*	3.33*	-0.53*
	(0.42)	(1.11)	(0.19)
Economic Growth	-0.39	-2.28	-0.85*
	(0.90)	(2.03)	(0.41)
Islam	0.018*	-0.006	0.008*
	(0.008)	(0.018)	(0.002)
Shape Parameter	1.21	0.90	
	(0.21)	0.19	

Note: $NT = 1277$. Entries are coefficient estimates; standard errors are in parentheses. Asterisks indicate $p < .05$ (one-tailed). See text for details.

Figure 1: Kaplan–Meier Survival Estimate, International Conflict, 1950–1985

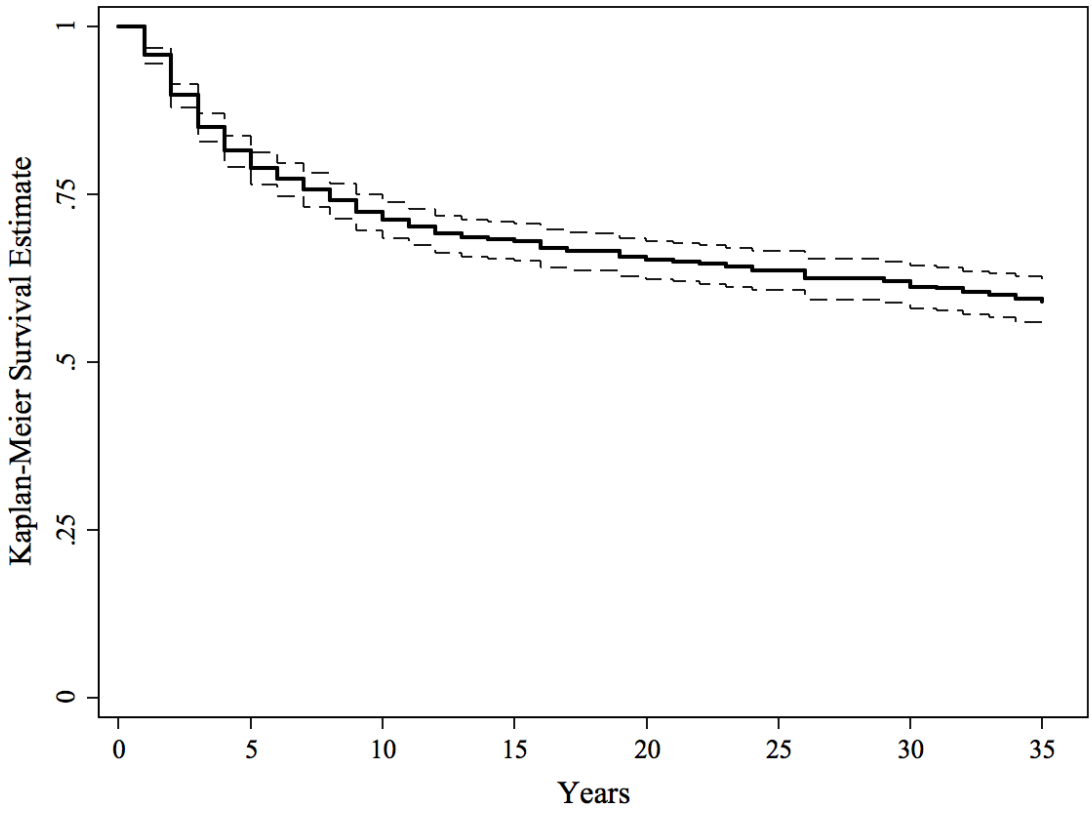

Figure 2: Kaplan–Meier Survival Estimate, Democratic Survival, 1960–1992