Web Appendix for
"Rethinking Third Party Interventions into Civil Wars: An Actor-Centric Approach"

Michael G. Findley & Tze Kwang Teo
Department of Political Science
University of Illinois at Urbana-Champaign
mfindley@uiuc.edu   tzeteo@uiuc.edu

A Coding Rules

Our outcome variables, *intervention on the side of the government* and *intervention on the side of the opposition*, are coded based on the "target of intervention" variable in Regan's dataset (2002). "Target" is originally a nominal variable coding three types of intervention: intervention on the side of the government, on the side of the opposition, and "neutral" interventions. We do not model the latter.

Rivalry with the civil war country is a dummy variable, coded as 1 if the potential intervener is a rival of the civil war country during the period of observation, according to Klein, Goertz, and Diehl's (2006) criteria.

Rivalry between potential and actual intervener are two dummy variables indicating rivalry between the potential intervener and actual intervener(s) on the side of the government, and on the side of the opposition likewise. The criteria and source of data for rivalry are the same as above.

Alliance with the civil war country is a dummy variable, coded as 1 if the potential intervener is allied to the civil war country via a formal defense or neutrality pact, or entente during the period of observation. Alliance data are obtained from the Correlates of War (COW) Alliance Data (Gibler and Sarkees 2004).

Alliance between potential and actual intervener are two dummy variables indicating alliance ties between the potential intervener and actual intervener(s) on the side of the government, and on the side of the opposition likewise. The criteria and source of data for alliance relations are the same as above.

Geographical Contiguity is a dummy variable, coded as 1 if the potential intervener and civil war country are geographically land contiguous to one another. Contiguity data are obtained from the COW Contiguity Data (Stinnett, Tir, Schafer, Diehl, and Gochman 2002).

Colonial History is a dummy variable, coded as 1 if the potential intervener was previously the colonizer of the civil war country. The data are obtained from the Issue Correlates of War (ICOW) Colonial History Data (Hensel 1999).

Major Power status is a dummy variable, coded as 1 if the potential intervener is designated by the COW Project as a major power during the period of observation.

Cold War is a dummy variable, coded as 1 if the conflict began before 1989.

Same Region is a dummy variable, coded as 1 if the potential intervener and civil war country are located in the same region, as defined by the COW Project.

Capability Ratio is the ratio of the potential intervener's Composite Indicator of National Capability (CINC; Singer 1987) score to the civil war country's CINC score. The base-10 logarithm transformation is taken. Version 3 of the CINC data is used.

Joint Democracy is a dummy variable, coded as 1 if both the potential intervener and civil war country score 6 or higher on the Polity scale. Regime type data are obtained from the Polity 4 project (Marshall and Jaggers 2004).

Ethnic: Opposition groups are coded as ethnic based on Gurr's Minority At Risk classification (1993). The data are obtained from Regan (2002).

Ideology: Opposition groups are coded by Regan (ibid.) as being ideological if the conflict revolves around contesting the dominant political or economic ideology.

Fatalities: Data on the levels of fatalities incurred over the course of the civil conflict are obtained from the Uppsalla/PRIO Battle Deaths Data (Lacina and Gleditsch 2005). The base-10 logarithm transformation is taken.

Refugees: Data on the number of refugees fleeing from the civil war country are obtained from Moore and Shellman's (2004) dataset on forced migration. The base-10 logarithm transformation is taken.

B "Mixture Cure" Survival Models

Standard survival models assume that *all* cases will eventually "fail", given enough time (Box-Steffensmeier and Jones 2004, 148–149); right-censored cases are simply those that have *yet to* fail. In the context of our analyses, this is tantamount to assuming that each and every outside state will intervene in the civil war sooner or later, should the conflict persist long enough. We contend, however, that some countries will never interfere in the affairs of other states because they are either unwilling to, or incapable of doing so. This suggests there are in fact two subpopulations within the total population of outside states—one comprising states that will eventually intervene, and another comprising states that will never do so. These two subpopulations are, using biostatistics terminology, the "uncured" and "cured fractions" respectively.

The existence of a cured fraction further implies two major sources of variation in third party intervention into civil wars: variation in (1) whether outside states will intervene eventually (or not), and (2) the timing of intervention among the countries that will intervene. Thus we are interested in modeling, as functions of covariates, both the probability of eventual intervention, and the hazard of (or duration until) intervention (conditional upon the expectation that intervention will occur). Because standard survival models only model the hazard or duration, we estimate "mixture cure" survival models instead. Mixture cure models contain the usual hazard or duration analysis component—found in all survival regression models—*and* an additional binary regression component for estimating the probability of eventually experiencing the event.

B.1 Overview of Remainder of this Appendix

Our presentation of mixture cure models proceeds as follows: We first define its mixture survivor, cumulative distribution, probability density, and hazard functions. This

is followed by overviews of maximum likelihood estimation of parametric cure models, and the functional form of the lognormal model estimated in our paper. We then discuss estimation of the semiparametric Cox cure model using the expectation maximization (EM) algorithm.[1] Our presentation draws on Farewell (1982), Peng (2003), Peng and Carriere (2002), Peng and Dear (2000), Sy and Taylor (2000), Sposto (2002), and Treasure (2000).

B.2   Mixture Functions of T

Let T denote the observed duration, and U the binary indicator of whether the event will eventually occur ($U = 1$, i.e., being uncured), or not ($U = 0$). A fraction of the total population, π, will eventually experience the event ($0 < \pi < 1$), while the remainder $(1 - \pi)$ will not. We follow the current literature on mixture cure models (e.g., Peng 2003; Sposto 2002), in which the survivor function is the first of the functions of T to be introduced. The *mixture* survivor function for the total population is

$$S(t) = \pi S_u(t) + (1 - \pi)S_{-u}(t), \tag{1}$$

where $S_u(t)$ and $S_{-u}(t)$ are the (conditional) survivor functions for the uncured and cured fractions, respectively. By definition, none of the cases in the cured fraction will experience the event, even as $t \to \infty$, so $S_{-u}(t)$ remains at unity throughout (Peng and Dear 2000, 237). Thus (1) simplifies to

$$S(t) = \pi S_u(t) + (1 - \pi), \tag{2}$$

and it can be seen that (1) and (2) reduce to the usual survivor function if all cases will eventually experience the event (i.e., $S(t) = S_u(t)$ when $\pi = 1$).

---

[1]We do not estimate the Cox cure model, but we suspect readers will be interested in this particular estimator, so we also present it for the sake of completeness.

The other functions of T that are relevant to our discussion are its

mixture cumulative distribution function (CDF)

$$
\begin{aligned}
F(t) &= 1 - S(t) \\
&= \pi - \pi S_u(t) \\
&= \pi F_u(t),
\end{aligned}
$$

mixture probability density function (PDF)

$$
\begin{aligned}
f(t) &= \frac{\mathrm{d}F(t)}{\mathrm{d}t} \\
&= \pi \frac{-\mathrm{d}S_u(t)}{\mathrm{d}t} \\
&= \pi f_u(t),
\end{aligned}
$$

and mixture hazard function

$$
\begin{aligned}
h(t) &= \frac{f(t)}{S(t)} \\
&= \frac{\pi f_u(t)}{\pi S_u(t) + (1 - \pi)}.
\end{aligned}
$$

We can also make use of the above to express $S(t)$ and $f(t)$ as

$$
\begin{aligned}
S(t) &= 1 - F(t) \\
&= 1 - \pi F_u(t)
\end{aligned}
$$

and

$$
\begin{aligned}
f(t) &= h(t)S(t) \\
&= \pi h_u(t)S_u(t).
\end{aligned}
$$

These functions also reduce to their usual forms when $\pi = 1$. (Note also that because we assume the conditional survivor function for the cured subpopulation $S_{-u}(t)$ remains at unity even as $t \to \infty$, its derivative must be zero, which further implies that the hazard faced by the cured fraction is also zero.)

We now introduce covariates into the discussion, as we are interested in their effects on the probability of being uncured and the hazard of failure. For example, the mixture

survivor function for the total population, modeled by covariates is

$$S(t|\mathbf{x}, \mathbf{z}) = \pi(\mathbf{z})S_u(t|\mathbf{x}) + [1 - \pi(\mathbf{z})]$$

$$= 1 - \pi(\mathbf{z})F_u(t|\mathbf{x}),$$

where $S_u(t|\mathbf{x}) = P(T > t|U = 1, \mathbf{x})$ and $F_u(t|\mathbf{x}) = P(T \leq t|U = 1, \mathbf{x})$ are the survivor and cumulative distribution functions respectively for the uncured cases, given a co-variate vector $\mathbf{x}$, and $\pi(\mathbf{z}) = P(U = 1|\mathbf{z})$ is the probability of being uncured given a covariate vector $\mathbf{z}$ (which may, but need not include the same covariates as $\mathbf{x}$).

B.3   Maximum Likelihood Estimation of Parametric Mixture Cure Models

Each case contributes data in the form $(t_i, \delta_i, \mathbf{x}_i, \mathbf{z}_i)$, $i = 1, 2, \ldots, n$, where t_i denotes the observed duration for the ith case, δ_i is the binary indicator that takes value 1 if t_i is un-censored and 0 otherwise, and $(\mathbf{x}_i, \mathbf{z}_i)$ are observed values of the two covariate vectors. Also define $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ as the vectors of parameters relating to $\mathbf{x}$ and $\mathbf{z}$, respectively. The likelihood function for the parametric cure model can be constructed from its standard expressions (e.g., Box-Steffensmeier and Jones 2004, 39)

$$L = \prod_{i=1}^{n} \left[f(t_i) \right]^{\delta_i} \left[S(t_i) \right]^{1-\delta_i}$$

$$= \prod_{i=1}^{n} \left[f(t_i) \right]^{\delta_i} \left[1 - F(t_i) \right]^{1-\delta_i}$$

by substituting in the mixture formulations of the functions of T

$$L(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \prod_{i=1}^{n} \left[\pi(\mathbf{z}_i) f_u(t_i|\mathbf{x}_i) \right]^{\delta_i} \left[\pi(\mathbf{z}_i) S_u(t_i|\mathbf{x}_i) + [1 - \pi(\mathbf{z}_i)] \right]^{1-\delta_i}$$

$$= \prod_{i=1}^{n} \left[\pi(\mathbf{z}_i) f_u(t_i|\mathbf{x}_i) \right]^{\delta_i} \left[1 - \pi(\mathbf{z}_i) F_u(t_i|\mathbf{x}_i) \right]^{1-\delta_i} \tag{3}$$

(e.g., Peng and Dear 2000, 238).

6

B.4 The Lognormal Mixture Cure Model

Our analyses of our interventions dataset indicate that the lognormal distribution is more appropriate than alternatives like the exponential and Weibull in characterizing the underlying hazard of intervention. The lognormal CDF for the uncured fraction is

$$F_u(t) = \Phi\Big(\ln[(\lambda t)^\alpha]\Big),$$

where $\Phi(\cdot)$ is the standard normal distribution function, λ is the scale parameter that determines the failure rate, and α is the shape parameter (Sposto 2002, 297). The lognormal mixture CDF and PDF modeled by covariates, with λ reparameterized as $\lambda = \exp(\mathbf{x}'\boldsymbol{\beta})$, are therefore

$$F(t) = \pi(\mathbf{z})\Phi[\alpha\mathbf{x}'\boldsymbol{\beta} + \alpha\ln(t)] \tag{4}$$

and

$$f(t) = \pi(\mathbf{z})\frac{\mathrm{d}\Phi[\alpha\mathbf{x}'\boldsymbol{\beta} + \alpha\ln(t)]}{\mathrm{d}t} \tag{5}$$

respectively. The likelihood function for the lognormal mixture cure model is obtained by substituting (4) and (5) into (3). The probability of being uncured can be modeled through appropriate functions, such as the logistic distribution function

$$\pi(\mathbf{z}) = [1 + \exp(-\mathbf{z}'\boldsymbol{\gamma})]^{-1}$$

(i.e., logit regression; Sposto 2002, ibid.). The maximum likelihood estimates of α, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ can be found via numerical algorithms like Newton-Raphson (Farewell 1982, 1042).

B.5 The Semiparametric Mixture Cure Model

A semiparametric survival model has the advantage of not requiring/imposing distributional assumptions about the underlying distribution of failure times. Kuk and Chen (1992) have developed the semiparametric mixture cure model, in which the failure time portion is the ubiquitous Cox Proportional Hazards model (also see Peng and Dear 2000, and Sy and Taylor 2000).

The semiparametric hazard, survivor, and probability density functions for the uncured fraction are

$$h_u(t) = h_{u0}(t) \exp(\mathbf{x}'\boldsymbol{\beta}),$$
$$S_u(t) = S_{u0}(t)^{\exp(\mathbf{x}'\boldsymbol{\beta})},$$

and

$$f_u(t) = h_u(t) S_u(t)$$
$$= h_{u0}(t) \exp(\mathbf{x}'\boldsymbol{\beta}) S_{u0}(t)^{\exp(\mathbf{x}'\boldsymbol{\beta})}$$

respectively, where $h_{u0}(t)$ and $S_{u0}(t)$ are the unspecified baseline hazard and survivor functions, respectively, for the uncured fraction. Therefore the likelihood function [refer back to (3)] for the semiparametric mixture cure model can be expressed as

$$L(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \prod_{i=1}^{n} \left[\pi(\mathbf{z}_i) h_{u0}(t_i) \exp(\mathbf{x}_i'\boldsymbol{\beta}) S_{u0}(t_i)^{\exp(\mathbf{x}_i'\boldsymbol{\beta})} \right]^{\delta_i}$$
$$\times \left[\pi(\mathbf{z}_i) S_{u0}(t_i)^{\exp(\mathbf{x}_i'\boldsymbol{\beta})} + [1 - \pi(\mathbf{z}_i)] \right]^{(1-\delta_i)} \quad (6)$$

(e.g., Sy and Taylor 2000, 228).

It may appear relatively straightforward from hereon to obtain estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. But this is not so, for the following reasons: First, it is not possible to maximize (6) using

standard maximum likelihood methods, because $h_{u0}(t)$ and $S_{u0}(t)$ are nonparametric (Treasure 2000, 4).

Second, maximum *partial* likelihood estimation is also encumbered, because we do not know exactly what cases belong to the uncured fraction. Recall that U is the binary indicator of whether the case is uncured ($U = 1$) or cured ($U = 0$). If $\delta_i = 1$, then we know $u_i = 1$. But if $\delta_i = 0$, then u_i is not observable and can take on either values 1 or 0 with $P(u_i = 1|\mathbf{z}_i) = \pi(\mathbf{z}_i)$. In other words, one cannot know for sure whether a right-censored case is really uncured and at risk, or cured and not at risk. We are thus unable to construct the requisite k partial likelihood "risk sets" containing *only* the remaining *uncured* cases, at each of the k successive distinct failure times, $t_1 < t_2 \cdots < t_k$ (ibid.).

Third, and as a consequence of the difficulty in constructing the risk sets, $h_{u0}(t)$ cannot be canceled out of the partial likelihood. Sy and Taylor further note that "[u]nlike in the [standard Cox] PH model where little information is lost by eliminating $S_0(t)$, one cannot eliminate $[S_{u0}(t)]$ in the estimation without losing information about $[\gamma]$" (2004, 228).

B.6   EM Algorithm for the Semiparametric Mixture Cure Model

As can be seen from above, estimation of the semiparametric cure model would be much easier *if* we could discern the latent cured/uncured status of right-censored cases, which means there is full observability of U. Note, in fact, that if this were indeed the case, we could then estimate parameters using the *complete data* likelihood function given the vector $\mathbf{u} = (u_1, \ldots, u_n)'$

$$L(\boldsymbol{\beta}, \boldsymbol{\gamma}|\mathbf{u}) = \prod_{i=1}^{n} \Big[\pi(\mathbf{z}_i)\Big]^{u_i} \Big[1 - \pi(\mathbf{z}_i)\Big]^{1-u_i} \Big[h_u(t_i|\mathbf{x}_i)\Big]^{\delta_i} \Big[S_u(t_i|\mathbf{x}_i)\Big]^{u_i} \tag{7}$$

(e.g., Peng 2003, 483).

Sy and Taylor (2000) and Peng and Dear (2000) propose employing the expectation maximization (EM) algorithm, hence dealing with the problem of partial observability of $\mathbf{u}$ by substituting with its conditional expectation at each iteration of the algorithm instead. The EM algorithm starts with initial values $(\boldsymbol{\beta}^{(0)}, \boldsymbol{\gamma}^{(0)})$. The E-step in the $(r+1)$th iteration calculates the expectation of (7) with respect to $\mathbf{u}$, conditional on the observed data and $(\boldsymbol{\beta}^{(r)}, \boldsymbol{\gamma}^{(r)})$, the estimates of $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ at the rth iteration. This is equivalent to calculating the conditional expectation of u_i:

$$
\begin{aligned}
p_i^{(r)} &= E(u_i|\boldsymbol{\beta}^{(r)}, \boldsymbol{\gamma}^{(r)}) \\
&= P(u_i = 1|\boldsymbol{\beta}^{(r)}, \boldsymbol{\gamma}^{(r)}) \\
&= \delta_i + \left[(1 - \delta_i) \frac{\pi(\mathbf{z}_i) S_u(t_i|\mathbf{x}_i)}{1 - \pi(\mathbf{z}_i) + \pi(\mathbf{z}_i) S_u(t_i|\mathbf{x}_i)} \right],
\end{aligned}
\tag{8}
$$

which is also the rth estimator of the probability of the ith case being uncured. Define the vector $\mathbf{p}^{(r)} = (p_1^{(r)}, \ldots, p_n^{(r)})'$. The M-step in the $(r+1)$th iteration maximizes the conditional expected complete log-likelihood function with respect to $(\boldsymbol{\beta}, \boldsymbol{\gamma})$, to obtain $(\boldsymbol{\beta}^{(r+1)}, \boldsymbol{\gamma}^{(r+1)})$. This function is the sum of

$$
\ln L_1(\boldsymbol{\beta}|\mathbf{p}^{(r)}) = \sum_{i=1}^{n} \left[p_i^{(r)} \ln S_u(t_i|\mathbf{x}_i) + \delta_i \ln h_u(t_i|\mathbf{x}_i) \right]
\tag{9}
$$

and

$$
\ln L_2(\boldsymbol{\gamma}|\mathbf{p}^{(r)}) = \sum_{i=1}^{n} \left[p_i^{(r)} \ln \pi(\mathbf{z}_i) + (1 - p_i^{(r)}) \ln(1 - \pi(\mathbf{z}_i)) \right].
\tag{10}
$$

(9) and (10) can be maximized separately. The algorithm is iterated until convergence. See Peng (2003), Peng and Carriere (2002), Peng and Dear (2000), and Sy and Taylor (2000) for more specific details on estimating the semiparametric cure model using the EM algorithm.

References

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

Farewell, V.T. 1982. "The Use of Mixture Models for the Analysis of Survival Data with Long-Term Survivors." *Biometrics* 38: 1041–1046.

Gibler, Douglas, and Meredith Reid Sarkees. 2004. "Measuring Alliances: The Correlates of War Formal Interstate Alliance Dataset, 1816–2000." *Journal of Peace Research* 41(2): 211–222.

Gurr, Ted Robert. 1993. *Minorities at Risk: A Global View of Ethnopolitical Conflicts.* Washington, DC: U.S. Institute of Peace Press.

Hensel, Paul. 1999. *Issue Correlates of War Colonial History Data.* Available at `http://garnet.acns.fsu.edu/~phensel/icowdata.html#colonies`. Accessed November 22, 2005.

Klein, James, Gary Goertz, and Paul F. Diehl. 2006. "The New Rivalry Dataset: Procedures and Patterns." *Journal of Peace Research* forthcoming.

Kuk, Anthony, and Chen-Hsin Chen. 1992. "A Mixture Model Combining Logistic Regression with Proportional Hazards Regression." *Biometrika* 79(3): 531–541.

Lacina, Bethany, and Nils Petter Gleditsch. 2005. "Monitoring Trends in Global Combat: A New Dataset of Battle Deaths." *European Journal of Population* 21(2–3): 145–166.

Marshall, Monty G., and Keith Jaggers. 2004. "Polity IV Project: Political Regime Characteristics and Transitions, 1800–2003." *Dataset Users' Manual.* Center for International Development and Conflict Management, University of Maryland, College Park.

Moore, Will H., and Stephen M. Shellman. 2004. "Fear of Persecution: Forced Migration, 1952–95." *Journal of Conflict Resolution* 48(5): 723–745.

Peng, Yingwei. 2003. "Fitting Semiparametric Cure Models." *Computational Statistics & Data Analysis* 41: 481–490.

Peng, Yingwei, and K.C. Carriere. 2002. "An Empirical Comparison of Parametric and Semiparametric Cure Models." *Biometrical Journal* 44(8): 1002–1014.

Peng, Yingwei, and Keith B.G. Dear. 2000. "A Nonparametric Mixture Model for Cure Rate Estimation." *Biometrics* 56: 237–243.

Regan, Patrick. 2002. "Third-Party Interventions and the Duration of Intrastate Conflicts." *Journal of Conflict Resolution* 46(1): 55–73.

Singer, J. David. 1987. "Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816–1985." *International Interactions* 14: 115–32.

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

Stinnett, Douglas M., Jaroslav Tir, Philip Schafer, Paul F. Diehl, and Charles Gochman. 2002. "The Correlates of War Project Direct Contiguity Data, Version 3." *Conflict Management and Peace Science* 19(2): 58–66.

Sy, Judy P., and Jeremy M.G. Taylor. 2000. "Estimation in a Cox Proportional Hazards Cure Model." *Biometrics* 56: 227–236.

Treasure, F. Peter. 2000. "Cure." Typescript. Statistical Laboratory, Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, UK. Available at `http://www.statslab.cam.ac.uk/~lab/Treasure/Cure.pdf`. Accessed November 22, 2005.