Modeling Structural Selection in Disaggregated Event Data

Mark David Nieman[*], Olga V. Chyzh[†], and Douglas M. Gibler[‡]

Abstract

Growing availability of disaggregated data, such as data on activity of subnational groups (e.g. protest campaigns, insurgents, terrorist groups, political parties or movements), has raised new types of theoretical and statistical challenges. In particular, rather than random, the observability and availability of disaggregated data are often a function of specific structural processes—an issue we refer to as *structural selection*. For example, domestic terrorist attacks or protester violence are conditional on the formation of domestic terrorist groups or protester movements in the first place. As a result, analytical inferences derived from subnational or other types of disaggregated data may suffer from structural selection bias, which is a type of sample selection bias. We propose a simple and elegant statistical approach to ameliorate such bias and demonstrate the advantages of this approach using a Monte Carlo example and two replications.

Dec 27, 2018

[*]Assistant Professor, Department of Political Science, Iowa State University.
[†]Assistant Professor, Department of Political Science and Department of Statistics, Iowa State University.
[‡]Professor, Institute for Social Science Research, University of Alabama.

Introduction

Growing theoretical interest in the micro-foundations of political processes, coupled with increased data availability and technological advances, have led to tremendous progress in collection and availability of disaggregated data, such as data on subnational units (e.g. protest campaigns, insurgents, terrorist groups, interest groups, political parties or movements). While creating opportunities for answering new types of research questions, these new types of data also introduce new theoretical and statistical challenges. One of the biggest challenges, and the focus of this paper, is recognizing and modeling the non-randomness of the structural processes that result in such data availability—an issue we refer to as *structural selection*.

Subnational political outcomes, such as protests, insurgencies, and domestic terrorist attacks, usually result from a two-stage non-random process. In the first stage, a group of individuals makes a decision to work together in pursuit of a common goal. In the second stage, the group makes decisions related to the promotion of their goal. The two outcomes—group formation and group activity—are interrelated, but each stage takes place at a different *level of aggregation*. In the first stage of this process, specific structural conditions, e.g. state-level factors such as a lack of government accountability or economic inequality, may lead to the formation of an insurgency group in a country. In the second stage, a set of group-level factors, such as group cohesion, values and ideology, and access to resources, affect this group's actions in pursuit of their goal. Deriving theoretical and statistical inferences regarding either of the outcomes, therefore, necessitates a two-level theoretical and statistical approach to modeling this interdependence.

Whether the outcome of interest is the formation or the activity of a subnational actor, deriving unbiased theoretical and statistical inferences of one requires an understanding of the other. Exclusion from economic resources may increase the probability of a formation of an insurgent group, yet at the group level, a lack of access to economic resources may limit the group's ability to engage in attacks. In this example, a group's exclusion from economic resources has two competing effects: a positive effect on the probability of an insurgency group formation and a negative effect on the probability or frequency of attacks. Simply controlling for horizontal inequality in statistical analysis, as is the current practice in the empirical literature, will obscure the effects of this variable and may lead to incorrect inferences regarding the outcome at the second stage—group activity. Since expectation of group's success is likely an important consideration for group formation, inferences regarding groups formation—the first stage—without an understanding of the group's probability of success, will also be biased. As a special case of the sample-selection problem, the issue of structural selection constitutes a relatively new and growing challenge for theoretical and statistical inferences.

Rather than constituting a random sample, units of observation in disaggregated datasets are observed (and enter the data) as a result of non-random national or systemic processes—what we refer to as *structural selection* processes. When not explicitly modeled, structural selection will result in the same type of bias as the infamous unit self-selection. A failure to recognize and model structural selection may result in a trivial conclusion that the very existence of insurgent or terrorist groups in a country is the best predictor of these groups' attacks, victory, or other activities of interest. Analyzing rebel groups' activity without a regard to the non-random structural conditions that led to their occurrence in the first place—the prevalent empirical practice—is akin to studying the effect of unpaid internships on starting salary. While the conditional treatment effect may reveal a positive relationships between taking one of more unpaid internships and starting salary, a failure to model the structural factors that allow some applicants to take unpaid internships in the first place (e.g., proximity to urban areas, family income and connections) may exaggerate the inferences from such a study.

Likewise, analyzing the effect of counter-insurgency policies (e.g., limitation on the freedom of movement) on a subsample of countries that experienced an insurgency will not help estimate the effect of similar policies in cases that have latent (but not active) insurgent groups. In other words, a counter-insurgency policy that is shown to fail at containing an ongoing insurgency, for example, may be very effective at preventing an onset of an insurgency. Testing counter-insurgency theories and policies only on a subsample of cases that have experienced an insurgency will obscure this very important insight.

The goal of the paper is, first, to draw attention to this important source of bias in studies that use subnational or other types of disaggregated data—a quickly growing area of research and data collection. Second, we propose an elegant and easy-to-implement statistical solution by highlighting the link between structural selection and multi-level modeling. The key to our approach is to specify a two-stage model by including the structural determinants of selecting into the sample as part of the first stage (i.e., the selection equation), and the group-level determinants of the outcome of interest (e.g., protests, attacks) as part of the second stage. As a result, the selection equation, possibly estimated at a higher level of aggregation, helps correct for the non-randomness of the sample that is used to estimate the outcome of interest. We show that our approach applies well to outcome variables drawn from common social-data distributions, including binary and count variables.

In the next section, we review the common types of sample selection bias, with a focus on structural selection. We then show how the specific type of sample selection of interest is easily corrected if the problem is recast in terms of a multi-level data structure. We discuss our approach in the context of other existing statistical techniques and highlight the advantages and scope of our approach. We support our argument with a Monte Carlo experiment and two empirical applications: first, we replicate a study on the success of non-violent protest campaigns and, second, we re-analyze a study on the relationship between civilian targeting and rebel group strength. We find that several of the inferences and conclusions drawn from these studies are determined, in part, by the underlying structural selection processes that make disaggregated events data observable.

Sample Selection Bias in Observational Data

Social scientists have long been aware of possible sample selection biases associated with observational data (Heckman 1979; Geddes 1990; Signorino 1999; Hug 2003; Cook et al. 2017; Nieman 2018). In contrast to data collected in experimental setting, observational data often yield non-random or biased samples. Uncorrected, sample selection bias leads to biased estimates in regression analysis. Using Heckman's original example, a sample of women in the workforce produces biased estimates of wages of women who chose to never enter the workforce, even controlling for levels of education and other relevant variables. Analogously, studies of political participation have long acknowledged that a sample of registered voters provides a poor estimate of turnout for unregistered voters (Erikson 1981; Squire, Wolfinger, and Glass 1987; Barreto, Segura, and Woods 2004; Nickerson 2014). Other subfields of political science have also recognized the issue: international conflict research has shown that a sample of cases of failed deterrence are not indicative of the probability of deterrence success for cases, in which the credibility of deterrence is never tested (Achen and Snidal 1989; Fearon 2002). Likewise, research on international organization has demonstrated that compliance rates of countries that enter international treaties is not indicative of those that do not (Von Stein 2005; Lupu 2013; Chyzh 2014).

In each of these examples, the bias is a result of the correlation between the outcome of interest and unit "self-selection" into the data. Collecting an unbiased (random) sample of all potential voters (rather than just registered voters) is impeded by the absence of definitive lists of unregistered voters (Barreto, Segura, and Woods 2004), just like drawing a random sample of deterrence cases requires identifying the unobservable cases of successful deterrence. In both cases, the units' probability of appearing in the sample is correlated with the outcome variable, and, even more problematically, with the probability of being observed in the first place.

Despite significant progress within certain areas of study, many types of selection remain undetected, continuing to obfuscate processes of interest. Part of the problem is that selection bias does not have a single cause, but may stem from a number of different processes related to the data-generating processes, case observability, data collection, and decisions made by the researcher. Hug (2003) identifies three general sources of selection bias. The first type—selection on the dependent variable when the whole population is observable—has so far received the most scholarly attention (Geddes 1990; King, Keohane, and Verba 1994; Dion 1998). This type of bias is easily remedied by drawing cases from the entire observable population rather than only those in which the dependent variable takes on the value of interest. This source of selection bias is perhaps the best understood and accounted for in today's literature.

The second type of selection bias may occur when cases self-sort themselves into specific outcomes, as in Heckman's canonical example where women choose to enter the workforce or stay at home. Just like women's decision to enter the workforce may be partially determined by their expected income, a country's joining of a treaty may not be independent of its subsequent compliance. In this case, correcting for possible selection bias involves specifying the two outcomes as separate equations and estimating them as part of a two-stage model, e.g. a selection model. In current research, discussions of this type of bias and the implementation of appropriate corrections are rather commonplace (Reed 2000; Signorino and Tarar 2006; Hansen, Rocca, and Ortiz 2015; Chyzh 2016; Feezell 2016; Nieman 2016).

The third type of selection bias—the focus of this paper—arises when case selection is perfectly correlated with case observability. This type of selection bias is most common in disaggregated datasets, whose cases are nested within a non-random sample of larger administrative units, e.g., insurgent groups or protesters within countries. Cases in these types of data are observed and enter the data as a result of a two-stage process. In the first stage, a subset of population decides whether to form an organization to pursue a collective goal, such as a political party, an insurgent group, or a terrorist organization. Even if such a group forms, this outcome may not be observable, as such subnational organizations are often informal or operate underground. A large number of such groups are only recorded as cases in scholarly data on the basis of a second-stage decision of whether they take specific actions towards the promotion of their goal, such as run in an election, challenge the government, or engage in an attack. As a further complication, the two decisions—to form and to take action to promote their goal—are not independent of each other. The group's expected success is likely a consideration for its formation in the first place (Nieman 2015).

Identifying the negative cases, such as the parties that never formed, or the insurgent groups that never organized, constitutes a tremendous conceptual challenge for collecting these types of data (Mahoney and Goertz 2004). Despite much effort correct for the sampling bias, such as the Minorities at Risk (MAR) Project (Minorities at Risk Project 2009) or the AMAR (A for "all") project whose goal is to collect the selection bias of the MAR data (Birnir et al. 2018), the resulting datasets are bound to suffer from various degrees of sample selection bias.

A key analytical complication for modeling structural selection is that the two outcomes—group formation and group activity—are produced by factors at different levels of aggregation or analysis. While a subnational group's decision to organize is usually driven by national or regional factors (e.g., dissatisfaction with government), the group's activity is a function of group-level factors (e.g., group's resources, ideology). Sample selection bias is introduced into the model as a result of the broader structural factors that lead to the formation of subnational groups. Correcting for such structural bias, therefore, necessitates a multi-level framework that bridges the group- and structure-levels of analysis.

Modeling Structural Selection Effects

Traditional selection estimators (e.g., Heckman 1979; Signorino 2003) are designed to model selection processes, in which both selection and outcome take place at the same unit of analysis.[1] Correcting the inferences regarding the wages of women in the workforce, for

[1]Some recent scholarship has proposed applying matching techniques to address endogeneity. Matching techniques, of course, can only match cases on *observables* and necessarily assume that data selection does not depend on potential outcomes (Ho et al. 2007). Chaudoin, Hays, and Hicks (2018) demonstrate that if

instance, is accomplished by modeling the outcome as the second stage of a process whose first stage comprised women deciding whether to enter the workforce. Importantly, the sample selection process is uncorrelated with the level of data aggregation: women exist in all countries independent of their decision to enter the workforce. Correcting for this type of sample selection, therefore, simply requires collecting additional data on women that chose not to enter the workforce and modeling this decision as the first stage of the analysis.

In contrast, structural sample selection implies not just a multi-stage, but also a multi-level, selection process—observed cases select into the second stage and level. Terrorist groups, for example, are not equally likely to form in all countries: the probability of observing a terrorist group is correlated with this group's probability of eliciting concessions from the government. While the group-level outcome (e.g., terrorist attacks) is mostly a function of group-level factors (e.g., resources, goals, member preferences), the group's existence is a function of structural factors (e.g., economic inequality, government capacity).[2] While the traditional approach would dictate that the selection bias be alleviated via collecting additional data on groups that never formed, such a task may not be productive or even practical. Instead, we propose an alternative, more elegant approach to modeling structural selection by re-conceptualizing the process of sample selection from the perspective of multi-level modeling.

The two stages of the process take place at different levels of aggregation, i.e. the first stage takes place at a higher/lower level of aggregation than the second stage. For example, subnational political actors, such as political parties, protesters, insurgents, and terrorist groups, are nested within their host-states. These groups form and act within the incentives and constraints of their host state (e.g., GDP per capita, political institutions). These groups' activity—running in an election, challenging the government, or engaging in attacks—is also

data suffer from selection on unobservables—as in the cases of structural selection we describe—matching techniques can exacerbate bias and overconfidence in estimates, as well as increase the number of falsely positive, statistically significant results.

[2]The stages of the process are, of course, rarely completely contained within levels, e.g., along with group-level factors, terrorist activity may be affected by some structural factors.

determined by the group-level factors, such as groups' resources and ideology.

More formally, denote a vector of the group-level outcome variable (e.g., number of attacks) as $\mathbf{y}$, whose i_{th} element, Y_i, equals to the number of attacks for group i. and model Y_i as a function of group-level exogenous regressors, $\boldsymbol{x}_i$, and a group-level disturbance term ϵ_i, i.e.:

$$Y_i \quad = \quad \boldsymbol{x_i}^T \boldsymbol{\beta} + \epsilon_i, \tag{1}$$

where $\boldsymbol{\beta}$ is a vector of model parameters.

According to the structural selection process, the group-level outcome y_i is observed (i.e., takes on non-missing values) under specific structural conditions, i.e. data on $\boldsymbol{x}_i$ are collected if condition α_i is met. More formally:

$$Y_i \quad = \quad \alpha_i(\boldsymbol{x_i}^T \boldsymbol{\beta} + \epsilon_i) \tag{2}$$

where α_i is an indicator variable that equals to 1 if the specific condition is met, and 0 otherwise. Note that α_i itself is a function of structural factors. That is:

$$\alpha_i^* \quad = \quad \mathbf{z}_i^T \boldsymbol{\gamma} + \eta_i \tag{3}$$

so that

$$\alpha_i = \begin{cases} 1 \text{ if } \alpha_i^* > 0 \\ 0 \text{ if } \alpha_i^* \leq 0 \end{cases}$$

where $\boldsymbol{\alpha}^*$ is a vector of the latent condition, α_i^* is an element of that vector, α_i is an element of a vector containing the latent condition's observed realization, $\boldsymbol{\gamma}$ is a vector of parameters, $\mathbf{z}_i$ is a vector of k exogenous structure-level covariates measured for observation i, and $\boldsymbol{\eta}$ is a vector of error terms at the structural level.

Importantly, Equation 3 must be estimated on a random sample drawn from the entire population of relevant units, not just the units, for which the group-level outcome and covariates are observed. In a study of domestic terrorist attacks, for example, Equation 3 would include *all* countries—not just the countries with known domestic terrorist groups—and model the outcome variable of whether a terrorist group formed within a country, α_i, as a function of the exogenous covariates, $\mathbf{z}$.[3] And the second-stage equation—Equation 2—would be specified with covariates that affect the number of attacks, Y_i, as a function of covariates $\boldsymbol{x}_i$ (e.g. group size, resources, ideology).

If $\boldsymbol{\epsilon}$ and $\boldsymbol{\eta}$ are correlated, i.e. corr $(\boldsymbol{\epsilon}, \boldsymbol{\eta}) \neq 0$, then the data availability on the group-level variables $\boldsymbol{x}_i$, as well as the values of $\boldsymbol{x}_i$, depend in part on the structural covariates $\mathbf{z}_i$. This, in turn, means that the structural covariates $\mathbf{z}_i$ affect the outcome Y_i, albeit not necessarily in a linear form. Non-zero correlation between $\boldsymbol{\epsilon}$ and $\boldsymbol{\eta}$ is likely, as this simply means that unobserved factors are correlated across the structure- and group-level. Conceptually related variables measured at the two different levels of aggregation, such as a state's military capacity and an insurgent's strength relative to the government, are likely to suffer from a some degree of measurement error which is correlated across levels. Moreover, unobservable covariates, like the degree of group and government resolve, are likely to be correlated across the two stages/equations. Non-zero correlation that is due to either measurement error or unobservable variables across the relevant levels of analysis produces selection bias in estimates of model parameters.

The proposed framework works for random variables measured on a continuous scale, as well as random variables that follow other normal or exponential family distributions (e.g., probit, logistic, poisson). Equations 2–3 may be re-written for a continuous random variable as:

$$Y_i = \begin{cases} \boldsymbol{x_i}^T \boldsymbol{\beta} + \epsilon_i & \text{if } \alpha_i = 1 \\ \text{missing} & \text{if } \alpha_i = 0. \end{cases}$$

[3]In this example, $\mathbf{Z}$ may include such covariates as government's responsiveness, GDP, geographical size and topography, or ethnic fractionalization.

For binary outcome variables, this takes on the form:

$$Y_i = \begin{cases} 1 & \text{if } \mathrm{Y}_i^* > \boldsymbol{x_i}^T\boldsymbol{\beta} + \epsilon_i \text{ and } \alpha_i = 1 \\ 0 & \text{if } \mathrm{Y}_i^* \leq \boldsymbol{x_i}^T\boldsymbol{\beta} + \epsilon_i \text{ and } \alpha_i = 1 \\ \text{missing} & \text{if } \alpha_i = 0. \end{cases}$$

The binary outcome variable case, of course, also extends to other discrete outcomes, such as ordered or nominal outcomes (see Miranda and Rabe-Hesketh 2006).

Finally, the count outcome variables differ slightly as the outcome can theoretically take on a number of discrete values (Miranda and Rabe-Hesketh 2006, 291-292). If we assume that a count random variable Y follows a Poisson distribution, so that $\Pr(Y_i|\mu) = \frac{\mu^{Y_i}e^{-\mu}}{Y_i!}$, then we can specify a log-linear model for the mean, μ. We can then write the count model as:

$$\ln(\mu) = \begin{cases} \boldsymbol{x}_i^T\boldsymbol{\beta} + \epsilon_i & \text{if } \alpha_i = 1 \\ \text{missing} & \text{if } \alpha_i = 0. \end{cases}$$

Notably, in the case of count models, the amount of overdispersion in the count is a function of the variance of $\boldsymbol{\epsilon}$ and can be identified and estimated (Miranda and Rabe-Hesketh 2006, 291). In other words, despite the assumption that the count of Y follows a Poisson distribution, the variance given a set of covariates is not equal to the conditional mean but instead permits and recovers estimates for the degree of overdispersion (see Kenkel and Terza 2001; Winkelmann 2008).

An advantage of focusing on the multi-level nature of the data—the group- (g) and structural-levels (s)—is that it provides a theoretical framework and justification to adopt a system estimator for the selection process. If the outcomes of interest are measured on a discrete scale—e.g. binary or count—they must be estimated with full information maximum likelihood (FIML), rather than a two-step estimation approach (Miranda 2004; Miranda and Rabe-Hesketh 2006; Freedman and Sekhon 2010; Greene 2010, 2018). Recovering unbiased estimates using a two-step approach, i.e. using the inverse Mills ratio—the ratio of the probability density function and the cumulative density function from the selection equation—as a regressor in the outcome equation, is predicated on two key assumptions: (1) a bivariate normal distribution of the error terms in the selection and outcome equations and (2) that the inverse Mill's ratio has a linear effect in the outcome equation. If either assumption is not met, inclusion of the inverse Mill's ratio leads to model misspecification and may induce bias (Winship and Mare 1992; Freedman and Sekhon 2010; Greene 2010, 2018). The second assumption, of course, is not met if the outcomes of interest for the group-level equation are discrete data.

An additional advantage of the structural selection approach, in contrast to typical selection models which use data on the same level of analysis, is that it lends itself to more easily overcoming concerns related to the exclusion restriction—that at least one exogenous variable is not in both equations, a problem common to Heckman-type selection models (Sartori 2003; Winship and Mare 1992).[4] This advantage stems from the different aggregation levels in the data for the structural and group-level equations—or $g \neq s$—which makes the ability to find an excluded variable much more straightforward. Most structural variables need not be included in the group-level equation, as they only affect the group-level outcome *indirectly*, by affecting the probability of the case realization (e.g., group formation) in the first-stage equation. The model, of course, does not prohibit including the relevant structural variables in both equations, as long as the restriction condition is met.

In summary, the proposed approach allows for modeling outcomes using disaggregated data—data that are only observed and collected under certain structural conditions. The scope conditions will apply to studies using subnational data (e.g., data on insurgencies, protests) or other types of disaggregated data (e.g., political candidates for various levels of administrative units, gangs operating in US states). If the outcome of interest, however,

[4] Alternatives to the exclusion restriction, such as identifying the model through the functional form (e.g., Sartori 2003), require making assumptions and inducing specific types of model dependence that may vary in appropriateness across studies.

can occur under multiple conditions, and it is only effect sizes that vary, then the data do not suffer from structural selection. Thus, a study of the effect of civil war on coups d'état would likely *not* suffer from structural selection, as coups can occur outside of a civil war context. That is, civil wars are not a necessary condition for a coup d'état; rather, coups during a civil war are a subset of the broader coups d'état category. In contrast, structural selection would likely be present in a study of temporal dependence in violent coups, as the type of coup (e.g., violent or peaceful) is conditional on observing a coup in the first place.

Monte Carlo Analysis

As an initial proof of our approach, we provide a Monte Carlo example. On this set of data, we compare several common estimation strategies used with data where the case selection is perfectly correlated with case observability. These strategies include estimating the activity of a group without accounting for its formation; estimating the activity of a group while including variables related to its structural conditions, without accounting for negative cases; and estimating the activity of groups while including fixed- or random-effects to capture omitted factors. We contrast the aforementioned estimation strategies with a censored model that treats at group activity as the second stage of a two-step process. The censored model takes information from the structural conditions associated with negative cases (i.e. where groups do not form) to inform estimates of the activity of a group.

To preview the results, all of the estimation strategies, except for the censored model, is problematic as they inflate coefficients associated with group activity. Further, estimation strategies that include structural variables, but ignore negative cases, recover parameters associated with *incorrect* signs, under some conditions. The inclusion of random effects exacerbate these problems in each set of cases. Conversely, the censored model recovers unbiased estimates for covariates associated with both group activity and structural conditions.

We start by generating $S = 100$ units at the structure level, each characterized by structure-level exogenous covariates ($\mathbf{z}$) and random disturbance term ($\boldsymbol{\eta}$). Covariates $\mathbf{z}$ are drawn from a uniform disribution, $\mathcal{U}[-2,2]$, while $\boldsymbol{\eta}$ follows a normal distribution. Next, for each of the 100 structure-level units, we generate a random variable, α^*, such that:

$$\alpha_i^* = 0.5 + 1z_i + \eta_i \tag{4}$$

For each structure-level unit $s \in \{1, 2, \ldots, \mathrm{S}\}$, such that $\alpha^* > 0$, we generate G=50 groups observations per structure-level unit to represent group-level sub-units typical to disaggragated data (e.g., 50 terrorist groups or 50 group-month observations of the same terrorist group). Each of the sub-units $g \in \{1, 2, \ldots, \mathrm{G}\}$ is characterized by fixed covariates $\boldsymbol{x}$ (e.g., group size, resources, ideology), drawn from a uniform distribution, $\mathcal{U}[-2,2]$. The group-level random variable, Y, is generated using a latent variable $\mathbf{Y}^*$, such that:

$$Y_i^* = \alpha_i(-0.5 + 1x_i + \epsilon_i). \tag{5}$$

To induce error correlation between the structural and group-level outcomes, $\boldsymbol{\epsilon}$ and $\boldsymbol{\eta}$ are drawn from a multivariate normal distribution with mean 0, variance 1, and a variance correlation such that $\mathrm{corr}(\boldsymbol{\epsilon}, \boldsymbol{\eta}) = \rho$. The group-level random variable Y_i takes on the value of 1 if $Y_i^* > 0$ and 0 otherwise. We vary the correlation between errors by setting $\rho \in \{0.7, 0.4, -0.4, -0.7\}$ and run 100 simulations at each value.

To compare the proposed approach with its alternatives, we estimate five different model specifications on the generated data. First, to mimic the most common treatments of group-level outcomes within the literature, we estimate (1) a probit model with just the group-level variables (Model 1 specified in Equation 6), and (2) a probit model that includes both the structure-level and the group-level variables in the same equation (Model 2 specified in Equation 7). These models are estimated on all cases where $\alpha = 1$ (i.e., a group-level outcome is observed) but, of course, exclude the cases for which $\alpha_i = 0$ (i.e., a group-level outcome is not observed due to structural "selection out"). We denote the structure-level variables, which are only measured if $\alpha_i = 1$, as $\mathbf{z}^*$. Model 2—a standard approach within the literature—will, of course, provide conditional estimates of the effect of structural factors, such as an estimate of the effect of government's capacity on insurgents' success, given that the government failed to deter an insurgency in the first place.

$$\mathrm{Y}_i^* = \beta_0 + \beta_1 \mathrm{x}_i + \epsilon_i, \tag{6}$$

$$\mathrm{Y}_i^* = \beta_0 + \beta_1 \mathrm{x}_i + \beta_2 \mathrm{z}_i^* + \epsilon_i \tag{7}$$

Next, we estimate two random-effects models, which allow the intercept to vary. This is a commonly used estimation technique designed to capture unobservable structure-level effects via a random intercept for each structural unit. The traditional random-effects model differs from our approach, of course, in that it excludes those structural units for which there are no group-level data, so the random intercepts, β_{0s}, are estimated only for the groups observed at the structural level. To indicate that the structural group intercepts are estimated using censored data on structural covariates (i.e. where $\alpha = 1$), we denote these estimates as β_{0s}^*. In the first of these random effects model (Model 3), we include just the group-level covariate (see Equation 8), while in the second random-effects model (Model 4) we include both the group- and the structure-level covariates (Equation 9).

$$\mathrm{Y}_i^* = \beta_{0s}^* + \beta_1 \mathrm{x}_i + \epsilon_i, \tag{8}$$

$$\mathrm{Y}_i^* = \beta_{0s}^* + \beta_1 \mathrm{x}_i + \beta_2 \mathrm{z}_i^* + \epsilon_i \tag{9}$$

Finally, as Model 5, we estimate a model that corresponds to our proposed approach—a censored probit such that the structure-level random variable is in the outcome of the selection equation and the group-level random variable is the outcome of the second equation. We expect that cases where group-level data are observed are not random but instead occur in the presence of specific structural conditions. Unobserved but related structure-level factors are also likely to be correlated with unobserved group-level characteristics. Thus, structure-level variables can be treated as a selection stage to the group-level observations.

$$\mathrm{Y}_i^* = \beta_0 + \beta_1 \mathrm{x}_i + \epsilon_i \text{ if } \alpha = 1 \text{ where } \alpha_i = 1 \text{ if } \alpha_i^* > 0, \text{ and 0 otherwise,} \quad (10)$$

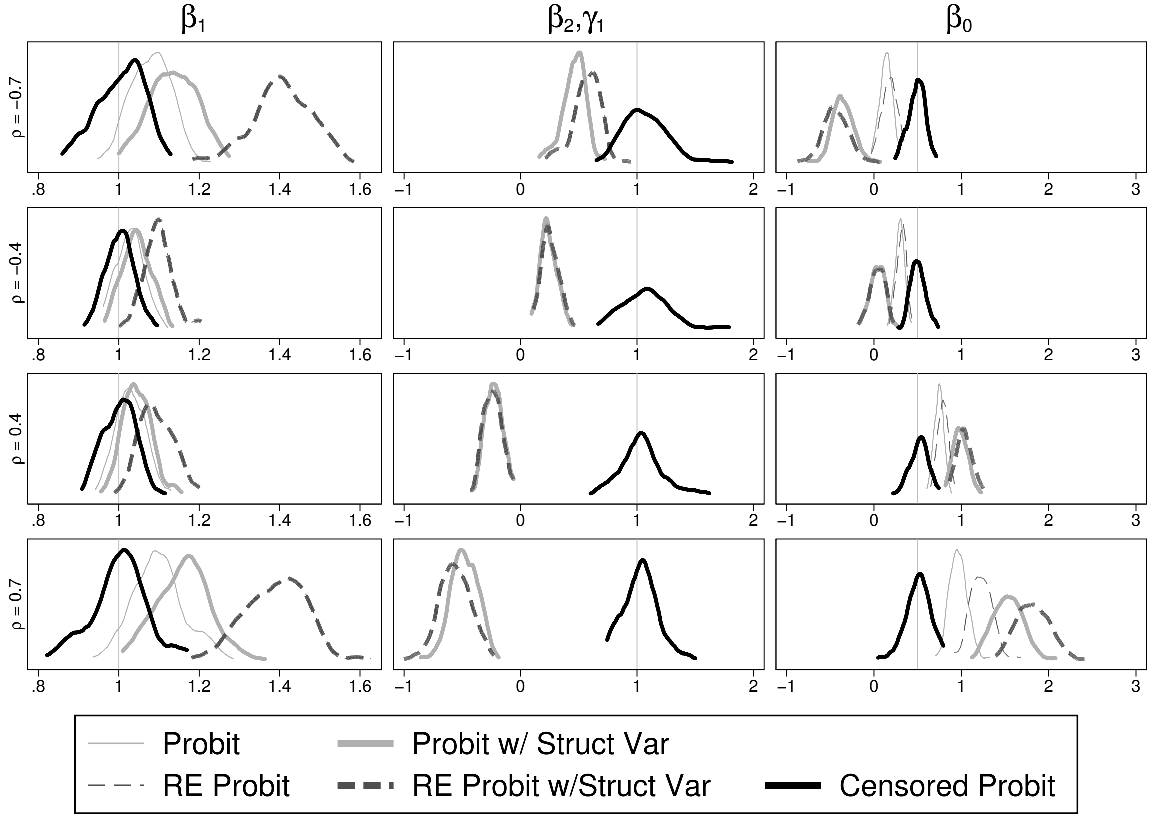
$$\alpha_s^* = \gamma_0 + \gamma_1 \mathrm{z}_s + \eta_s \quad (11)$$

We present the results of the Monte Carlo analysis in Figure 1 and Table 1.[5] As our estimates suggest, each of the probit models exaggerates the effect of x as ρ moves away from zero. Moreover, the probit models asserts a high degree of certainty in their biased estimates. Estimates of the structure-level variable in probit models that include z have the wrong sign on the parameter when ρ is positive and are biased towards zero when ρ is negative. Finally, the probit models overestimate the effect of the constant when ρ is positive and underestimate the effect (with the wrong sign) of the constant when ρ is negative. Bias on the constant is problematic given that estimators of discrete data generating processes are hyper-conditional—the estimated coefficient on one variable depends on the value of the estimates of other variables. Thus, the substantive effects of a quantity of interest, as well as the predictive power of the model, will be incorrect owing to bias in the constant.

Next, we examine parameter estimates from the probit models with random effects. Like the initial probit models, random-effects probit models exaggerate the effect of x more the farther ρ is from zero; in fact, exaggeration of the effect of x is much more pronounced here than in the traditional probit models. Random-effects probit also performs poorly when a structure-level variable is included, recovering estimates with the incorrect sign when ρ is positive and estimates that are biased towards zero when ρ is negative. The model also overestimates the constant when ρ is positive and underestimates the constant when ρ is negative. In sum, these results suggest that the estimates of random-effects probit actually exhibit the greatest degree of bias when compared to other models, which is especially problematic given how often this approach is used to address structure-level heterogeneity

[5] The coefficients reported on the *constant* in the RMSE tables are of the outcome equation for the censored probit specification.

Figure 1: Monte Carlos with Varying Degrees of Error Correlation Between Levels.



Note: Vertical line represents the true value of the coefficient. Results for 100 simulations of 100 structure-level units, with 50 group-level observations per unit, for each value of rho.

in structured data.

Finally, we turn to the probit model that accounts for potential structural selection effects. Figure 1 and Table 1 highlight, in particular, that a censored probit is the only model that recovers unbiased estimates of the true effects of the state-level covariate, z. The censored probit also performs best in terms of estimating the true value of the model's intercept, β_0. Moreover, accounting for selection, the censored probit model also recovers unbiased estimates for each parameter regardless of the value of ρ.

Our results demonstrate that ignoring underlying structural selection processes, and estimating a single-equation model (e.g., group activity) for a second-stage process (group formation and activity), produces biased estimates and possibly incorrect inferences. As the correlation between group and structure unobservables increases, estimates of parameters

Table 1: Root MSE with Varying Degrees of Error Correlation Between Levels.

		Rho = -0.7			
Var.	Probit	Probit w/ Struct.	RE	RE Struct.	Censored Probit
X	0.259	0.284	0.475	0.474	0.273
Z	—	0.591	—	0.494	0.470
Con.	0.494	0.938	0.497	1.056	0.375

		Rho = -0.4			
Var.	Probit	Probit w/ Struct.	RE	RE Struct.	Censored Probit
X	0.201	0.205	0.226	0.225	0.209
Z	—	0.788	—	0.777	0.462
Con.	0.331	0.545	0.324	0.566	0.313

		Rho = 0.4			
Var.	Probit	Probit w/ Struct.	RE	RE Struct.	Censored Probit
X	0.208	0.212	0.234	0.234	0.220
Z	—	1.256	—	1.268	0.456
Con.	0.355	0.592	0.396	0.651	0.350

		Rho = 0.7			
Var.	Probit	Probit w/ Struct.	RE	RE Struct.	Censored Probit
X	0.266	0.294	0.463	0.463	0.285
Z	—	1.498	—	1.580	0.439
Con.	0.574	1.101	0.817	1.401	0.429

associated with group activity increase, resulting in an increase likelihood of drawing false positives. Moreover, fixes common to the literature, such as the inclusion of structural-level variables or estimating random effects, do not correct for the underlying selection problem. The results are significant for many analyses taking advantage of recent data collection efforts which focus on disaggregated, group-level data, such as data on political parties, protester movements, terrorism, and insurgencies. Moreover, the approach generalizes beyond cases with binary outcome variables to other types of discrete outcome variables (e.g., Greene 2010; Miranda and Rabe-Hesketh 2006).

Empirical Applications

To further demonstrate the impact of ignoring the process of structural selection, we replicate two prominent studies of domestic political instability. First, Chenoweth and Stephan (2011) is a widely cited work that argues that non-violent protests are more likely to result in government concessions than is the use of violence. Whether the protest resulted in government concessions, however, is a second-stage outcome of a multi-stage process, in the first stage of which structural conditions result in the formation of a protest campaign. The probability of observing a violent or non-violent protest is thus conditional on the *ex ante* probability of success and the structural conditions of the state. As we demonstrate above with Monte Carlo estimates, the unmodeled non-randomness in protest data may lead to biased estimates of effects. It may be the case, for example, that nonviolent protests only occur when conditions promoting change are more likely.

Our second replication examines whether the structural conditions in the state also affect the strategies of conflict. Wood (2010) argues that rebel groups that lack the capacity to garner popular support are also the groups most likely to target civilians during civil conflicts. Of course, the observation of rebel groups is non-random and heavily conditioned by the relative strength of the government and the likelihood of a rebel group's success. The non-random nature of the data, combined with the expected correlation between structural- and group-level factors, suggests that selection processes may be at play.

Structure, Protest Occurrence, and Protest Outcomes

Chenoweth and Stephan (2011) look at how the methods utilized by protest campaigns affect how successful they are at obtaining their political goals. They expect that non-violent protest campaigns are more effective than violent protest campaigns. Chenoweth and Stephan (2011) treat the protest campaign as the unit of analysis, with data that explores government–protest campaign interactions. However, game-theoretic models hypothesize that protest campaigns are (negatively) correlated with the expectation of government repression (Pierskalla 2010; Ritter and Conrad 2016; Chyzh and Labzina 2018). Further, since the use of repression differs systematically across states (Regan and Norton 2005; Davenport 2007; Hill and Jones 2014), we expect the likelihood of both protest and success covary similarly. Complicating this relationship even more is the fact that protest strategies covary with these same structural conditions. Almost by definition, non-violent protest campaigns are impossible in extremely repressive regimes that do not tolerate dissent. To account for the possibility of a non-random sample of protest movements, we examine the probability of success by non-violent campaigns in achieving government concessions in the context of structural selection. We model state-level data as a selection equation and use those estimates to inform campaign-level data in the outcome equation.

We focus our replication on the main model (Model 1) of Table 3.1 from Chenoweth and Stephan (2011). They measure a protest movement's *success* as a binary outcome coded 1 if it achieves its stated goals, 0 otherwise. *Non-violent resistance* is measured as 1 if the movement is primarily non-violent, 0 otherwise. They also control for level of democracy, the number of participants in the movement, and the state's population.[6]

We account for structural factors using the model of civil conflict from Fearon and Laitin (2003, Table 1, Model 1).[7] We employ data from Gibler and Miller (2014), who extend and expand Fearon and Laitin's dataset following the original authors' coding rules. The structural model includes common predictors of domestic strife, such as *democracy*, political *instability*, *GDP/capita*, and whether a state has territory that is *non-contiguous*. The model also estimates conditions that favor challenges to government authority, such as the size of the *population*, amount of *mountainous terrain*, *oil exports*, and *ethnic* and *religious fractionalization*.[8]

Table 2 reports the results of our analyses using probit and censored probit selection models, where the selection equation is the structural or macro-level and the outcome equation is the micro- or group-level event data.[9] The first column displays the replication of Chenoweth and Stephan (2011) Table 3.1, Model 1 using a probit model. The second column displays the subset of the data for which the structural and campaign data overlap. This ensures that the models in Columns 2 and 3 include the same set of observations to provide a proper comparison. The third column displays the results when the multi-level selection process is also modeled.

Comparing models demonstrates that structural factors appear to influence the likelihood of whether protests occur. Once structural conditions are modeled, the type of protest

[6]See Chenoweth and Stephan (2011) for a discussion of how control variables are measured.

[7]See Regan and Norton (2005) for a similar structural/state-level approach to modeling protest behavior.

[8]See Fearon and Laitin (2003) or Gibler and Miller (2014) for a discussion of how variables are measured.

[9]Chenoweth and Stephan (2011) Table 3.3 do consider potential endogeneity in the use of violent resistance and protest-campaign success. However, they only look at the data from their protest campaign sample when constructing their instrument. By doing so, they ignore endogenity induced by selection processes.

Table 2: Probit Estimation of Protest Occurrence and Success.

Variable	Replication	Subsample	Structure-Selection
Protest Success			
Non-violent	0.548*	0.463†	0.189
	(0.290)	(0.321)	(0.168)
Democracy	0.031†	0.027†	0.022**
	(0.019)	(0.020)	(0.009)
Participants	0.229**	0.221**	0.118**
	(0.076)	(0.084)	(0.053)
Population	-0.262**	-0.295**	-0.250**
	(0.104)	(0.115)	(0.071)
Constant	-0.102	0.426	3.384**
	(0.952)	(1.052)	(0.537)
Domestic Protest			
GDP/capita			0.016
			(0.043)
Population			0.151**
			(0.021)
Mountains			0.007
			(0.026)
Non-contiguous			0.249*
			(0.135)
Oil exporter			-0.170†
			(0.113)
Democracy			-0.227**
			(0.070)
Democracy^2			-0.112**
			(0.056)
Instability			0.308**
			(0.095)
Ethnic Frac			0.079
			(0.145)
Religious Frac			0.060
			(0.132)
Constant			-3.813**
			(0.465)
Rho			-0.930**
			(0.078)
Log-likelihood	-79.88	-66.11	-616.36
Observations	141	115	7883 (115)

$^{**}p < 0.05$, $^{*}p < 0.10$ two-tailed, $^{\dagger}p < 0.10$ one-tailed. Robust standard errors in parentheses. The number under observations parentheses in the structure-selection model are uncensored cases.

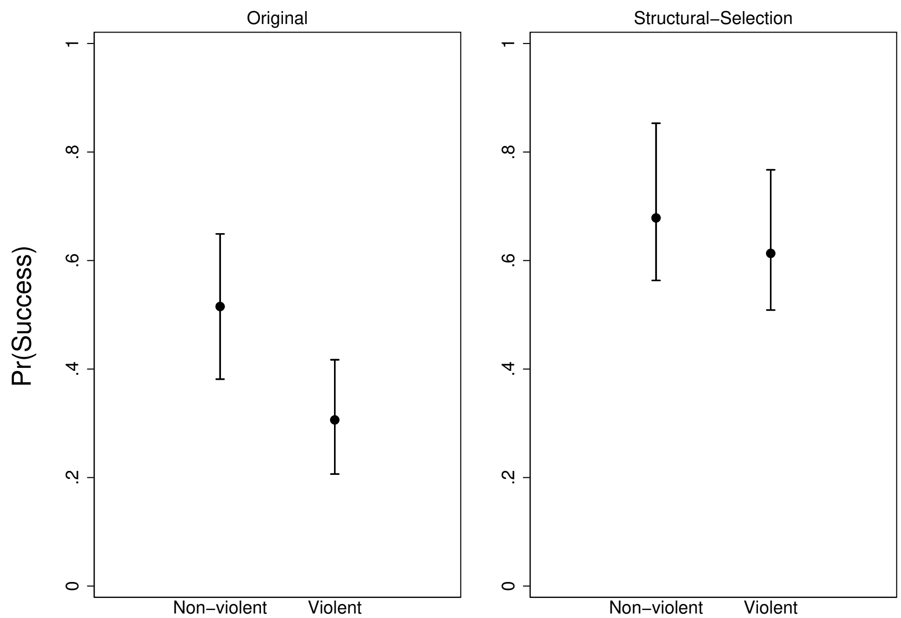
campaign does not exert a statistically significant influence in our model. The coefficient on *non-violent* protests is roughly the same as its standard error. The coefficient for *democracy*, however, is now statistically significant at conventional levels, as including estimates for the structural factors influencing protests reduces the degree of uncertainty associated with the coefficient. It is also worth noting that the constant is ten times larger in absolute value, as well as positive and statistically significant, once selection based on structure is modeled. This suggests that when protests are observed, regardless of other factors, they are more likely to succeed.

Several factors associated with government weakness—*population size*, *non-contiguous territory*, and *political instability*—increase the probability of observing protest movements, while high levels of *democracy* are associated with fewer protests. The negative constant implies the likelihood of protest at any given time is small and the negative *rho* suggests that any unobserved factors decrease the likelihood of protest at any given time. Indeed, protests are costly, and protesters often organize only after other alternatives are exhausted, but this also suggests that mean likelihood of success among the observed protest campaigns will be much higher than expected by chance. Each of these results is consistent with the formal theoretical literature which expects that protesters behave strategically and are more likely to protest (and use non-violent methods) when they expect that the government will not repress.

The findings that non-violent protests are no more successful than violent protests, and that the set of observed protests arise from specific structural conditions, also has important substantive implications. In the original analysis in Model 1, the predicted probability of success of a non-violent protest campaign, holding all other variables at the mean or modal values, is 53.8% with a 90% confidence interval of [41.1, 66.7]. The predicted probability of success from a violent protest is 32.6% [22.5, 39.5]. The first difference between these values is 21.2% [3.2, 39.5]. Once we account for the underlying structural selection processes, however, the results change dramatically. The predicted probabilities of successful non-violent and violent campaigns are now 69.7% [58.2, 86.7] and 63.4% [53.0, 78.8], respectively. The first difference between these values is now only 6.4% [-3.0, 15.3], with the 90% confidence interval now including zero. These results are summarized in Figure 2.

Accounting for the non-random process and focusing on the average treatment effect for the population, rather than a conditional effect on a subset of that population, can also have important policy implications. As illustrated by the protest example, non-violent and violent protest campaigns have similar success rates once the selection effect of whether a protest campaign is observed occurs. Efforts, then, by outside organizations that seek and

Figure 2: Substantive Effect of Non-Violent and Violent Protests on Protest Success.



Note: The figure on the left displays the predicted values and 90% confidence intervals using the replication of Chenoweth and Stephan (2011) reported in Table 2, Model 1. The figure on the right displays the predicted values and 90% confidence intervals using after accounting for structural selection, based on the estimates from Table 2, Model 3.

advocate for protest campaigns—often often citing the success rates of non-violent tactics—may inadvertently result in government repression and negative outcomes from those that participate in the protest campaign. That is, advocating policies based on conditional effects are likely to be inappropriate, and even contradictory towards one's goals, if those conditions are not present.

Overall, the results suggest that there is a selection effect in the data and estimates based solely on the group-level data will be biased. Moreover, the substantive results highlight how analysts may draw incorrect or misleading inferences if they neglect to account for the structural selection processes that make observing event data possible.

Structure, Civil Conflict, and Civilian Targeting

In the second application, we analyze the effect of structural selection on civilian targeting by rebel forces. Wood (2010) argues that rebel groups with stronger capabilities vis-à-vis the government can use a mix of selective incentives and repression to garner support and resources from the population. Weaker rebel groups, on the other hand, often lack the capacity to offer incentives to the population to garner support and instead rely to a greater degree on civilian targeting. The unit of analysis is the dyad-year, where the dyad consists of an insurgent group and the government.

We previously argued that outbreaks of civil conflict are non-random, and data on rebel groups can only be collected if civil conflicts are observed. These two points imply that observed rebel groups are likely to be more capable than the population of potential rebel groups (Nieman 2015; Chatagnier and Castelli 2016). Civil conflict, moreover, is made more likely by specific structural factors, e.g., low government capacity, institutional instability, loot-able resources (Fearon and Laitin 2003; Ross 2004; Cunningham, Gleditsch, and Salehyan 2009; Cederman, Weidmann, and Gleditsch 2011). Taken together, state factors affect the likelihood of civil conflict, which in turn likely affects the type of rebel groups that are observed and their interactions with the government. Thus, we expect that this structural selection effect influences rebel group behaviors, including the tactic of civil targeting.

Specifically, we replicate Model 1 of Table 2 from Wood (2010). Wood (2010) measures the count of *rebel–civilian one-sided killings* as the direct, intentional killings of civilians in non-combat situations by rebel forces (Eck and Hultman 2007).[10] *Rebel capability* is the ratio of troops to the scaled number of government troops (Eck and Hultman 2007).[11] He also controls for *government violence* against the population, *identity conflicts*, *territorial conflict*, the overall degree of *conflict severity*, the *age* of the conflict, *democracy*, *GDP/capita*, and whether the conflict takes place during the *Cold War*.[12] We measure structural factors related to conflict using the same model as above but also add a lagged variable of *ongoing conflict* to account for conflict duration (Fearon and Laitin 2003).

Table 3 reports the results of analyses using a negative binomial model and a count

[10]The measure does not include indirect civilian deaths resulting from sieges, disease, collateral damage, or extrajudicial executions (Wood 2010, 606).

[11]The scaling of the measure accounts for the potential presence of multiple insurgencies in one state.

[12]See Wood (2010) for a discussion of how the control variables are measured.

Table 3: Count Estimate of Rebel One-sided Civilian Killing and State Structure.

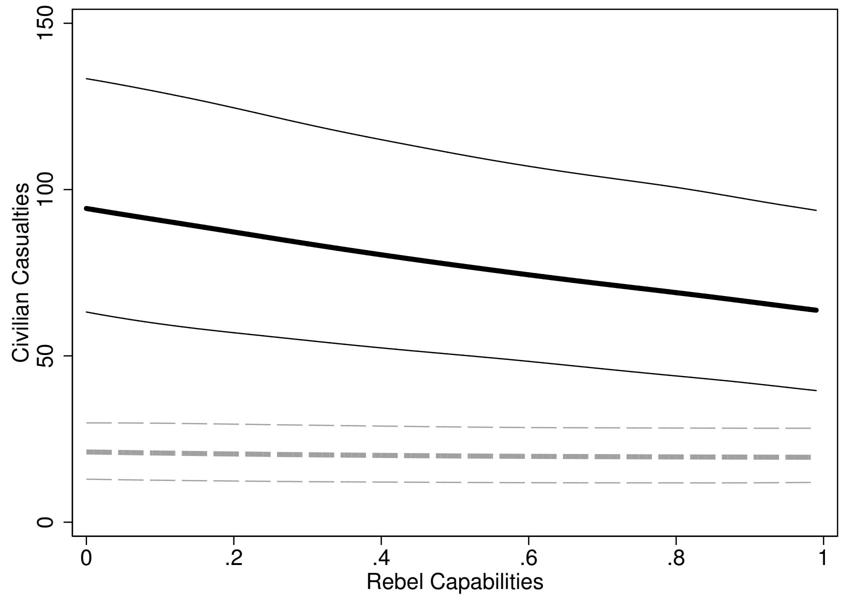
Variable	Replication	Subsample	Structure-Selection
Rebel Civilian Killing			
Rebel capacity	-0.492**	-0.403**	-0.075**
	(0.178)	(0.147)	(0.035)
Government violence	0.000*	0.000*	0.000**
	(0.000)	(0.000)	(0.000)
Identity conflict	0.892**	0.891**	-0.756**
	(0.427)	(0.293)	(0.115)
Territorial conflict	-1.008**	-1.169**	-0.565**
	(0.413)	(0.329)	(0.080)
Conflict severity	0.601**	0.536**	0.521**
	(0.087)	(0.064)	(0.022)
Age	0.224†	0.172	-0.510**
	(0.172)	(0.160)	(0.030)
Democracy	0.107**	0.104**	0.031**
	(0.037)	(0.030)	(0.008)
GDP/capita	-0.718**	-0.608**	-0.131**
	(0.229)	(0.185)	(0.036)
Cold War	-0.807**	-0.624†	-0.963**
	(0.380)	(0.405)	(0.079)
Constant	4.796**	4.494**	3.011**
	(1.725)	(1.453)	(0.337)
Log(alpha)	2.677**	2.625**	
	(0.140)	(0.083)	
Civil Conflict			
GDP/capita			-0.142**
			(0.054)
Population			0.026
			(0.029)
Mountains			0.080**
			(0.033)
Non-contiguous			-0.258†
			(0.175)
Oil exporter			0.153
			(0.149)
Democracy			-0.616**
			(0.149)
Democracy^2			-0.587**
			(0.139)
Instability			0.175†
			(0.132)
Ethnic Frac			0.374*
			(0.220)
Religious Frac			-0.421**
			(0.196)
Ongoing Conflict			2.340**
			(0.101)
Constant			-1.129*
			(0.593)
Sigma			1.117**
			(0.036)
Rho			-0.305**
			(0.037)
Log-likelihood	-1830.262	-1703.220	-6190.310
Observations	679	609	3293(609)

$^{**}p < 0.05$, $^{*}p < 0.10$ two-tailed, $^{\dagger}p < 0.10$ one-tailed. Robust standard errors in parentheses. The number under observations parentheses in the structure-selection model are uncensored cases.

model that accounts for selection (Miranda 2004; Miranda and Rabe-Hesketh 2006). The first column reports the exact replication of Wood (2010). The second column displays the subset of the data for which the structural and rebel group data overlap; this is done so that the models in Columns 2 and 3 include the same observations in order to ensure proper comparison. The third column reports the results of a count model conditioned by the structural selection process.

The estimates in Table 3 demonstrate that several structural factors influence the group-level interactions that take place within them, such as *GDP/capita* and the degree to which a

Figure 3: Substantive Effects of Changing Rebel Capacity on Civilian Targeting.



Note: The solid line displays predicted values and 90% confidence intervals using the replication of Wood (2010) reported in Table 3, Model 2. The dashed line displays predicted values and 90% confidence intervals after accounting for structural selection, based on the estimates from Table 3, Model 3.

state is *democratic*. The *rho* parameter in Column 2 is negative and statistically significant, indicating that the unobservable factors from the structural-level are negatively correlated with the unobservable group-level factors that affect one-sided rebel-civilian killing.[13] Thus, the same factors that lead an opposition to arm and to fight the government also make them less likely to engage in one-sided civilian killing.

The structural factors also affect the substantive results in the dyadic analysis of rebel tactics. The coefficient on *rebel capability*, for instance, is substantially smaller when conditioned by structural selection. To better demonstrate this change, Figure 3 compares the substantive effects of the two models using predicted values (90% confidence intervals) from Monte Carlo simulations based on estimates from Table 3. The solid line displays predicted values using Model 1 and the dashed line displays predicted values accounting for structural selection. The model that ignores selection identifies a steep, declining slope in civilian ca-

[13]The negative correlation of the structural- and group-level errors is consistent with Gibler (2017), who found that structural conditions affect reporting of crisis events in narratives compiled by the International Conflict Group.

sualties as rebel capabilities increase, while the model that accounts for structural selection factors shows almost no decline in civilian casualties at all. Moving from a *rebel capacity* of 0.2 to 0.8 in the replication without selection, for example, results in a decrease of civilian casualties of 84.8 to 69.2. Comparatively, moving from a *rebel capability* of 0.2 to 0.8 in the model where structural selection is account for results in a change in civilian casualties from 20.2 to 19.3. Substantively, this means that ignoring structure-level factors would lead one to significantly overestimate the degree to which rebel capacity reduces civilian killings by insurgent groups.

Finally, it is also worth noting that the sign on the coefficient for *identity conflict* changes from negative to positive, and is statistically significant in both models. Similarly, while *age* is positive and significant at the 0.1-level, one-tailed test in Model 1 and positive but statistically insignificant in Model 2, once structural selection processes are accounted for, the coefficient is negative and statistically significant. Ignoring structural selection may lead one to incorrectly infer that identity conflicts are more likely to result in civilian targeting than non-identity conflicts, and that older rebel groups are more likely to target civilians than younger ones.

Conclusion

We argue that structural selection impacts estimates involving disaggregated events data. We use both a Monte Carlo experiment and empirical replications to demonstrate that model estimates are improved by accounting for the the non-random processes at the structural-level that makes such groups organize in the first place. Our empirical applications demonstrate that some inferences from recent work on protest movements, and civilian targeting during civil conflicts are likely to be incorrect. Non-violent protests are not more effective once the structural environment that influences the likelihood of protest is considered. The structural factors also heavily influence the observation of rebel groups and amount of civilian targeting.

As we argue, accounting for structural selection improves estimates and associated inferences of causal variables and relationships which, in turn, enhances our theoretical understanding and increases the quality of policy prescriptions based on these theories.

Finally, though the paper focuses primarily on domestic outcomes, we believe that structural characteristics are also inherent within other types of event data. Green political parties, for example, tend to form under specific types of political and economic conditions. Likewise, international militarized disputes tend to occur in certain regions and certain times. Accounting for structural selection helps improve estimates and associated inferences, which, in turn, enriches our theoretical understanding of political processes and enhances the quality of our policy prescriptions.

References

Achen, Christopher H. and Duncan Snidal. 1989. "Rational Deterrence Theory and Comparative Case Studies". *World Politics* 41.2, 143–169.

Barreto, Matt A, Gary M. Segura, and Nathan D. Woods. 2004. "The Mobilizing Effect of Majority–Minority Districts on Latino Turnout". *American Political Science Review* 98.1, 65–75.

Birnir, Jóhanna K., David D. Laitin, Jonathan Wilkenfeld, David M. Waguespack, Agatha S. Hultquist, and Ted R. Gurr. 2018. "Introducing the AMAR (All Minorities at Risk) Data". *Journal of Conflict Resolution* 62.1, 203–226.

Cederman, Lars-Erik, Nils B. Weidmann, and Kristian Skrede Gleditsch. 2011. "Horizontal Inequalities and Ethnonationalist Civil War: A Global Comparison". *American Political Science Review* 105.03, 478–495.

Chatagnier, J. Tyson and Emanuele Castelli. 2016. "The Arc of Modernization: Economic Structure, Materialism, and the Onset of Civil Conflict". *Political Science Research and Methods.* DOI:10.1017/psrm.2016.24.

Chaudoin, Stephen, Jude Hays, and Raymond Hicks. 2018. "Do We Really Know the WTO Cures Cancer?" *British Journal of Political Science* 48.4, 903–928.

Chenoweth, Erica and Maria J. Stephan. 2011. *Why Civil Resistance Works: The Strategic Logic of Nonviolent Conflict*. Columbia University Press, 2011.

Chyzh, Olga V. 2014. "Can You Trust a Dictator: An Endogenous Model of Authoritarian Regimes' Signing and Compliance with International Treaties". *Conflict Management and Peace Science* 31.1, 3–27.

Chyzh, Olga V. 2016. "Dangerous Liaisons: An Endogenous Model of International Trade and Human Rights". *Journal of Peace Research* 52.3, 409–423.

Chyzh, Olga V. and Elena Labzina. 2018. "Bankrolling Repression? Modeling Third-Party Influence on Protests and Repression". *American Journal of Political Science* 62.2, 312–324.

Cook, Scott J., Betsabe Blas, Raymond J. Carroll, and Samiran Sinha. 2017. "Two Wrongs Make a Right: Addressing Underreporting in Binary Data from Multiple Sources". *Political Analysis* 25.2, 223–240.

Cunningham, David E., Kristian Skrede Gleditsch, and Idean Salehyan. 2009. "It Takes Two: A Dyadic Analsys of Civil War Duration and Outcome". *Journal of Conflict Resolution* 53.4, 570–597.

Davenport, Christian. 2007. "State Repression and the Tyrannical Peace". *Journal of Peace Research* 44.4, 485–504.

Dion, Douglas. 1998. "Evidence and Inference in the Comparative Case Study". *Comparative Politics* 30.2, 127–145.

Eck, Kristine and Lisa Hultman. 2007. "One-Sided Violence Against Civilians in War: Insights from New Fatality Data". *Journal of Peace Research* 44.2, 233–246.

Erikson, Robert S. 1981. "Why Do People Vote? Because They Are Registered". *American Politics Quarterly* 9.3, 259–276.

Fearon, James D. 2002. "Selection Effects and Deterrence". *International Interactions* 28.1, 5–29.

Fearon, James D. and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War". *American Political Science Review* 97.01, 75–90.

Feezell, Jessica T. 2016. "Predicting Online Political Participation: The Importance of Selection Bias and Selective Exposure in the Online Setting". *Political Research Quarterly* 69.3, 495–509.

Freedman, David A. and Jasjeet S. Sekhon. 2010. "Endogeneity in Probit Response Models". *Political Analysis* 18.2, 138–150.

Geddes, Barbara. 1990. "How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics". *Political Analysis* 2, 131–150.

Gibler, Douglas M. 2017. "Combining Behavioral and Structural Predictors of Violent Civil Conflict: Getting Scholars and Policymakers to Talk". *International Studies Quarterly* 61.1, 28–37.

Gibler, Douglas M. and Steven V. Miller. 2014. "External Territorial Threat, State Capacity, and Civil War". *Journal of Peace Research* 51.5, 634–646.

Greene, William H. 2010. "A Stochastic Frontier Model with Correction for Sample Selection". *Journal of Productivity Analysis* 34.1, 15–24.

Greene, William H. 2018. *Econometric Analysis*. New York: Peason, 2018.

Hansen, Wendy L., Michael S. Rocca, and Brittany Leigh Ortiz. 2015. "The Effects of Citizens United on Corporate Spending in the 2012 Presidential Election". *Journal of Politics* 77.2, 535–545.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error". *Econometrica* 47.1, 153–161.

Hill, Daniel W. and Zachary M. Jones. 2014. "An Empirical Evaluation of Explanations for State Repression". *American Political Science Review* 108.03, 661–687.

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2007. “Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference”. *Political Analysis* 15.3, 199–236.

Hug, Simon. 2003. “Selection Bias in Comparative Research: The Case of Incomplete Data Sets”. *Political Analysis* 11.3, 255–274.

Kenkel, Donald S. and Joseph V. Terza. 2001. “The Effect of Physician Advice on Alcohol Consumption: Count Regression with an Endogenous Treatment Effect”. *Journal of Applied Econometrics* 16.2, 165–184.

King, Gary, Robert O. Keohane, and Sydney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton University Press, 1994.

Lupu, Yonatan. 2013. “The Informative Power of Treaty Commitment: Using the Spatial Model to Address Selection Effects”. *American Journal of Political Science* 57.4, 912–925.

Mahoney, James and Gary Goertz. 2004. “The Possibility Principle: Choosing Negative Cases in Comparative Research”. *American Political Science Review* 98.4, 653–669.

Minorities at Risk Project. 2009. *Minorities at Risk Dataset*. College Park, MD: Center for International Development and Conflict Management. Retrieved from `http://www.cidcm.umd.edu/mar/` on 01/12/2012. 2009.

Miranda, Alfonso. 2004. “FIML Estimation of an Endogenous Switching Model for Count Data”. *Stata Journal* 4.1, 40–49.

Miranda, Alfonso and Sophia. Rabe-Hesketh. 2006. “Maximum Likelihood Estimation of Endogenous Switching and Sample Selection models for Binary, Ordinal, and Count Variables”. *Stata Journal* 6.3, 285–308.

Nickerson, David W. 2014. “Do Voter Registration Drives Increase Participation? For Whom and When?” *Journal of Politics* 77.1, 88–101.

Nieman, Mark David. 2015. "Statistical Analysis of Strategic Interaction with Unobserved Player Actions: Introducing a Strategic Probit with Partial Observability". *Political Analysis* 23.3, 429–448.

Nieman, Mark David. 2016. "The Return on Social Bonds: Social Hierarchy and International Conflict". *Journal of Peace Research* 53.5, 665–679.

Nieman, Mark David. 2018. "Strategic Binary Choice Models with Partial Observability". *Statistica Sinica* 28.4, 2089–2105.

Pierskalla, Jan Henryk. 2010. "Protest, Deterrence, and Escalation: The Strategic Calculus of Government Repression". *Journal of Conflict Resolution* 54.1, 117–145.

Reed, William. 2000. "A Unified Statistical Model of Conflict Onset and Escalation". *American Journal of Political Science* 44.1, 84–93.

Regan, Patrick M. and Daniel Norton. 2005. "Greed, Grievance, and Mobilization in Civil Wars". *Journal of Conflict Resolution* 49.3, 319–336.

Ritter, Emily Hencken and Courtenay R. Conrad. 2016. "Preventing and Responding to Dissent: The Observational Challenges of Explaining Strategic Repression". *American Political Science Review* 110.1, 85–99.

Ross, Michael L. 2004. "How Do Natural Resources Influence Civil War? Evidence from Thirteen Cases". *International Organization* 58.1, 35–67.

Sartori, Anne E. 2003. "An Estimator for Some Binary-Outcome Selection Models Without Exclusion Restrictions". *Political Analysis* 11.2, 111–138.

Signorino, Curtis S. 1999. "Strategic Interaction and the Statistical Analysis of International Conflict". *American Political Science Review* 93.2, 279–297.

Signorino, Curtis S. 2003. "Structure and Uncertainty in Discrete Choice Models". *Political Analysis* 11.4, 316–344.

Signorino, Curtis S. and Ahmer Tarar. 2006. "A Unified Theory and Test of Extended Immediate Deterrence". *American Journal of Political Science* 50.3, 585–605.

Squire, Peverill, Raymond E. Wolfinger, and David P. Glass. 1987. "Residential Mobility and Voter Turnout". *American Political Science Review* 81.1, 45–65.

Von Stein, Jana. 2005. "Do Treaties Constrain or Screen? Selection Bias and Treaty Compliance". *American Political Science Review* 99.4, 611–622.

Winkelmann, Rainer. 2008. *Econometric Analysis of Count Data.* Springer Science & Business Media, 2008.

Winship, Christopher and Robert D. Mare. 1992. "Models for Sample Selection Bias". *Annual Review of Sociology* 18, 327–350.

Wood, Reed M. 2010. "Rebel Capability and Strategic Violence against Civilians". *Journal of Peace Research* 47.5, 601–614.