A TWO-STAGE APPROACH TO CIVIL CONFLICT:

CONTESTED INCOMPATIBILITIES AND ARMED VIOLENCE

Henrikas Bartusevičius*
Department of Political Science, Aarhus University

* Corresponding author, email: henrikas@ps.au.dk

Kristian Skrede Gleditsch
Department of Government, University of Essex and Peace Research Institute Oslo

Abstract: We present a two-stage approach to civil conflict analysis. Unlike conventional approaches that focus only on armed conflict and treat all other cases as "at peace", we first distinguish cases with and without contested incompatibilities (Stage 1) and then whether or not contested incompatibilities escalate to armed conflict (Stage 2). This allows us to isolate factors that contribute to conflict origination (onset of incompatibilities) and factors that promote conflict militarization (onset of armed violence). Using new data on incompatibilities and armed conflict, we replicate and extend three prior studies of violent civil conflict, reformulated as a two-stage process, considering a number of different estimation procedures and potential selection problems. We find that the group-based horizontal political inequalities highlighted in research on violent civil conflict clearly influence conflict origination but have no clear effect on militarization, whereas other features emphasized as shaping the risk of civil war, such as refugee flows and soft state power, strongly influence militarization but not incompatibilities. We posit that a two-stage approach to conflict analysis can help advance theories of civil conflict, assess alternative mechanisms through which explanatory variables are thought to influence conflict, and guide new data collection efforts.

Acknowledgement: Previous versions of this paper were presented at the European Network for Conflict Research Meeting, Uppsala University, Uppsala, October 2014; Workshop on "Contemporary Conflict Research", University of Essex, Colchester, February 2015; Workshop on "Conflict, Strategies, and Tactics", University of Essex, Colchester, June 2015; 15th Jan Tinbergen European Peace Science Conference, University of Warwick, Coventry, June 2015; 5th Annual General Conference of the European Political Science Association, Vienna, June 2015; and Workshop on "Conflict and Democratization", Aarhus, November 2016. This research was supported by grants from the European Research Council (313373) and Innovation Fund Denmark.

Armed conflicts rarely erupt out of nowhere. The armed stage of a conflict is typically preceded by a formative non-armed stage, where parties articulate incompatibilities and become aware of their opposing positions.[1] Armed mobilization is not inevitable, and normally takes place much later than the emergence of the initial incompatibilities. Thus, even if some incompatibilities and armed conflicts may emerge simultaneously, the outbreak of armed conflict generally results from a two-stage process: the first involving the onset of an incompatibility (conflict origination) and the second the onset of armed violence (conflict militarization). International relations research has long acknowledged the importance of distinguishing between the onset of a "dispute" and "escalation to war". [2] However, civil conflict research has traditionally focused on armed conflicts and paid little attention to contested incompatibilities that do not see violence.

The lack of attention to the distinction between non-violent and violent incompatibilities is unfortunate on several grounds. Analyzing civil conflict onset as a dichotomous category (peace/armed conflict) potentially conceals non-monotonic effects of variables on conflict origination and conflict militarization. Recent explanations of civil conflict stress how grievances arising from group-based inequalities can motivate conflict and the use of violence.[3] However, grievances can also generate non-violent responses, and mobilization for armed action may fail. Indeed, despite grievances, the risk of militarization might be low, as privileged groups can rely on state resources to deter armed mobilization by disadvantaged groups.[4] If so, the observed effect of horizontal inequalities on armed conflict would primarily reflect their impact on conflict

[1] Boulding 1962.

[2] Fearon 1994; Huth 1996; and Reed 2000.

[3] Cederman, Gleditsch, and Buhaug 2013.

[4] Lichbach 1989, 437–8.

origination rather than violence as a specific means. The conventional approach—comparing countries with and without armed conflict—does not allow us to assess this claim, as the absence of armed conflict does not imply the absence of incompatibilities.

Furthermore, restricting civil conflict to armed violence hinders the assessment of alternative mechanisms relating specific variables to conflict. Consider the role of low GDP per capita:[5] Does it increase the risk of armed conflict because it generates incompatibilities (e.g., through poverty-related grievances) or because it facilitates rebel recruitment for armed violence (e.g., due to lower opportunity costs)? To discriminate between these claims, we must distinguish between conflict origination and conflict militarization.

This study demonstrates how the two conflict stages can be conceptually and empirically defined. In contrast to previous research that has lumped together all non-armed conflicts and the absence of manifest conflict into a compound "peace" category, we distinguish between three distinct states: "no contested incompatibility", "contested incompatibility", and "armed conflict". This allows us to isolate contested incompatibilities (that may or may not see violence) and armed conflicts over incompatibilities. This, in turn, allows us to assess factors that account for conflict origination and others that account for conflict militarization.

We first discuss existing conceptualizations of civil conflict and present a framework that disaggregates the dichotomy of peace/armed conflict into the three distinct states. We then introduce data that allow us to identify conflict origination separately from conflict militarization and present our two-stage estimation strategy. Subsequently, we demonstrate the value of the two-stage approach by replicating three studies of violent civil conflict. We start by replicating a study of horizontal inequalities and motivation for conflict, which we expect should yield clearer

[5] Hegre and Sambanis 2006.

effects on incompatibilities than armed conflicts.[6] We then replicate two other studies that illustrate factors that make armed conflicts more likely conditional on incompatibilities rather than affecting the emergence of incompatibilities.[7] Our analyses show how findings in previous research on civil war pertain to particular conflict stages and provide new theoretical insights into the likely mechanisms. Based on these findings, we conclude that a two-stage approach can help advance theories of civil conflict, help assess alternative mechanisms through which variables are theorized to influence conflicts, and guide new data collection.

Research on Civil Conflict

Previous research has shown that various demographic, economic, institutional, and geographic factors correlate cross-nationally with the onset of civil conflict.[8] However, we know less about how these factors influence different conflict stages. With some recent exceptions,[9] studies of civil conflict have traditionally focused on armed categories, ignoring non-armed incompatibilities or relegating them to a compound category of "peace", defined only by the absence of organized violence.

Many scholars of interstate conflict have long highlighted the importance of distinguishing between separate stages such as the onset of disputes or rivalries and escalation to war.[10] Some studies have found that key factors linked to international wars have distinct effects

[6] Buhaug, Cederman, and Gleditsch, 2014.

[7] Salehyan and Gleditsch, 2006; Warren 2014.

[8] Dixon 2009.

[9] E.g., Cunningham 2013; White et al. 2015.

[10] Fearon 1994; Huth 1996; and Reed 2000.

on different conflict stages. For example, Reed found that satisfaction with the status quo and joint democracy reduce the likelihood of a dispute in a dyad but do not affect the likelihood of the dispute escalating to war, while power parity has effects with opposite signs in the two stages.[11] These findings provide important insights into the mechanisms affecting crises and escalation to international wars. For example, the "democratic peace" appears to arise from a lower likelihood of disputes between democracies rather than institutional constraints that prevent the use of large-scale violence against other democratic states in disputes.[12]

Given the recognition of two-stage approaches in international relations research, it is striking that similar approaches have not been applied to the study of civil conflict. The emphasis on violence in civil conflict research stems in part from the lack of data on domestic equivalents of interstate disputes. Interstate disputes are easy to identify through overt claims or threats sanctioned by governments.[13] By contrast, domestic incompatibilities may not be publically expressed due to fears of government sanctions. Delineating parties to an incompatibility is also difficult, as potential actors such as ethnic groups often lack cohesion and collective organization prior to large-scale mobilization, and many groups opposing governments are not ethnic.

Recently, several projects have begun examining non-armed civil conflict,[14] making significant progress in conceptualizing and measuring domestic incompatibilities independently of violence. These efforts have generated new data on domestic non-violent campaigns with

[11] Reed 2000.

[12] Ibid., 90–1.

[13] Hensel et al. 2008, 128; Jones, Bremer, and Singer 1996, 169.

[14] Chenoweth and Cunningham 2013.

mass mobilization[15] or disputes and non-violent crises.[16] We use these data to identify civil conflict as a two-stage process below, first turning to our conceptual framework.

Conceptualizing Contested Incompatibilities and Armed Conflict

We define *conflict*—at the basic level of a concept[17]—as a contested incompatibility, where "incompatibility" refers to an "incompatible difference of objective…a desire on the part of both contestants to obtain what is available only to one, or only in part" [18]. "Contested" implies that the contestants pursue their objectives in a manifest and active manner (criterion *a*).

Subsequently, and in line with previous work,[19] we define *civil* conflict as an incompatibility over government and/or territory (*b*) between two or more politically organized actors (*c*), one of which is a state government (*d*), that takes place primarily within the borders of one state (*e*), and involves extra-institutional means of contention (*f*).[20] Incompatibilities can be over the government (i.e., the "type of political system, the replacement of the central government, or the change of its composition") or over territory (i.e., "demands for secession or autonomy").[21]

[15] Chenoweth and Lewis 2013.

[16] Schwank et al. 2013.

[17] Goertz 2005, 6.

[18] Dahrendorf 1959, 135; see also Boulding 1962; and Wallensteen 2015.

[19] Gleditsch et al. 2002; Small and Singer 1982.

[20] Chenoweth and Lewis 2013.

[21] Gleditsch et al. 2002, 619.

We can then define *armed* civil conflict as an incompatibility involving the systematic use of armed force. "Systematic" implies organized and sustained over an extended period. Non-armed incompatibility and armed conflict are thus both instances of civil incompatibility/conflict but distinguished by the presence/absence of the systematic use of armed force.

The "no incompatibly"/incompatibility distinction is more difficult to evaluate and requires actual data on evidence of claims as well as criteria for determining them. The criteria introduced above arguably allow for overcoming these difficulties. Criterion *a* ensures that incompatibilities are publicly displayed, with contestants actively pursuing their objectives. Hence, the mere presence of potentially contentious issues such as inequality does not by itself constitute an incompatibility unless the relevant actors mobilize and articulate specific claims. Criterion *b* establishes that incompatibilities threaten either the government or the core integrity of a state. Criteria *c* and *d* ensure that incompatibilities only include conflicts between state governments and other clearly defined political actors, thereby excluding non-state conflicts and incompatibilities between non-politically organized actors, such as criminal enterprises. Finally, criterion *f* excludes routine, legal political processes that are typically considered distinct from incompatibilities leading to armed conflict. Using these concepts, we thus redefine the outbreak of civil conflict as the end result of a two-stage process that generates the three above-mentioned states.

Measuring Contested Incompatibilities and Armed Conflict

Two-stage analysis requires data to identify the onset of incompatibilities separately from the onset of armed conflict. The Conflict Information and Analysis System (CONIAS) dataset

contains categories that approximate our conceptual definitions.[22] CONIAS seeks to catalogue all "political conflicts" between 1945 and 2008. The conceptual and operational definitions are provided in the *Conflict Barometer*, where political conflict (base category) is defined as:

> a positional difference, regarding values relevant to a society—the conflict items—between at least two decisive and directly involved actors…carried out using observable…conflict measures that lie outside established regulatory procedures and threaten core state functions. [23]

In line with our criterion *b,* "conflict items" in CONIAS include system/ideology, national power, autonomy, secession, and subnational predominance. "Decisive actors" in CONIAS are individuals, states, international organizations, or non-state actors whose "existence, actions, and communications considerably alter the practices of at least one other conflict actor pertaining to the conflict item".

Further, the CONIAS definition requires that conflicts must be carried out using measures outside established regulatory procedures that threaten core state functions. This fulfills our criteria *a* and *f* and helps exclude inactive incompatibilities or institutionalized political processes that do not threaten core state functions.

CONIAS has five conflict intensity levels: (1) disputes, (2) non-violent crises, (3) violent crises, (4) limited wars, and (5) wars. Disputes and non-violent crises constitute non-violent conflicts, whereas violent crises, limited wars, and wars constitute violent conflicts. In line with

[22] Schwank et al. 2013.

[23] Heidelberg Institute for International Conflict Research 2015, 9.

our definitions presented above, non-violent conflicts and violent conflicts are distinguished by the use of armed violence.

Disputes and non-violent crises are distinguished by the *threat* of the use of violence. Five indicators distinguish between violent crises, limited wars, and wars: (1) type of weapon and how it was used; (2) personnel involved; (3) casualties/deaths; (4) damage to infrastructure, accommodation, the economy, and culture; (5) cross-border refugees and internally displaced persons. The CONIAS conflict observers give each indicator a score, which are aggregated into the conflict intensity levels (Table 1). We provide a more elaborate discussion of the CONIAS definitions in the Online Appendix (OA).

[Table 1]

Since we focus on intrastate conflicts, we exclude all interstate events and non-state conflicts. Subsequently, to adjust the dependent variable for our analyses, we aggregate the five levels into non-armed incompatibilities (disputes and non-violent crises) and armed conflicts (violent crises, limited wars, and wars). For Stage 1, we code country-years 1 if a country experienced an incompatibility, 0 if otherwise. For Stage 2, we code country-years 1 if the incompatibilities reach the level of armed conflict and 0 if the incompatibilities remain non-armed.

The relative onset rate for incompatibilities at Stage 1 is much lower (0.041) than militarization within incompatibilities at Stage 2 (0.12). Furthermore, more than one-third of all incompatibilities that ever became armed were not armed at the outset year, and two-thirds of all incompatibilities that became limited wars did so after the onset year. Thus, although some

incompatibilities are armed during the first year,[24] many conflicts indeed appear to proceed in a two-stage fashion, often with a considerable delay between initial incompatibilities and the onset of violence.

Validating CONIAS

CONIAS is a successor of the KOSIMO dataset, initiated at the Heidelberg Institute for International Conflict Research in 1991.[25] Thus, CONIAS is an outcome of data collected over two decades. The data and coding decisions are documented in extensive case studies in *Conflict Barometer*.[26] The detail of these case studies and the transparency of coding decisions help support the face validity of the CONIAS data.[27] However, although KOSIMO has been used in many studies,[28] CONIAS has received less attention among conflict scholars. Therefore, we attempted to validate it independently.

[24] We expect that few, if any, armed conflicts break out without some history of prior incompatibilities. The proportion of incompatibilities armed at the outset would likely be lower if we could assess timing on a more fine-grained basis (e.g., by date or month).

[25] Pfetsch and Rohloff 2000.

[26] Heidelberg Institute for International Conflict Research 2015.

[27] Moreover, KOSIMO was a key source for the candidate list used to backdating the UCDP/PRIO Armed Conflict Dataset from 1990 to 1945; see Gleditsch et al. 2002. Identifying lower intensity conflicts may be more difficult due to fewer media reports; however, we see no inherent reason why such biases would apply differently to whether incompatibilities see violence or not.

[28] E.g., Chiozza 2002; Davis 2012; O'Brien 2002; Pfetsch 2006.

For conciseness, we only summarize the main points here and provide the full validation procedure in OA (2–82). We first compared the CONIAS violent categories to armed conflicts coded in the UCDP/PRIO Armed Conflict Dataset.[29] The incidence of the two categories overlaps substantially. Furthermore, as demonstrated in the replication studies below (Tables 2–4), substituting the UCDP/PRIO armed conflict or the category of civil war as defined by Sambanis[30] with the CONIAS armed conflict in conventional (dichotomous) models yields very similar results. While this does not directly validate the CONIAS non-violent categories, it demonstrates that the CONIAS coding rules capture violent phenomena largely overlapping with those identified by other well-established datasets.

Validating the CONIAS non-violent incompatibilities is more difficult, as we lack analogous alternative data. The conceptually and operationally closest categories are the NAVCO mass campaigns.[31] However, NAVCO is restricted to campaigns with at least 1,000 participants, whereas CONIAS has no such threshold. NAVCO therefore represents a more restrictive list of incompatibilities than CONIAS. Even so, there is considerable overlap between the two datasets (OA). Most importantly, we find that CONIAS and NAVCO generate very similar estimates in the two-stage analysis provided below (Table 2).

[29] Gleditsch et al. 2002; Melander, Pettersson, and Themnér 2016.

[30] Sambanis 2004.

[31] Chenoweth and Lewis 2013. We also considered comparing CONIAS to self-determination disputes or social conflict events from the SCAD dataset; see Cunningham 2013; Salehyan et al. 2012. However, the former includes only the sub-set of territorial incompatibilities, and the latter includes many events that do not fall under our definition of incompatibility (e.g., labor union strikes, conflicts between non-state actors).

Finally, we scrutinized the coding of 10 randomly selected CONIAS conflicts and evaluated the assigned values against case study evidence. While the coding of start and end dates of some conflicts could be contested, overall, the data closely matched the record in the cases.

Two-Stage Estimation Strategy

Two-stage models highlight potential selection issues, as factors that determine outcomes in the first stage may also affect conclusions about impacts in the second stage.[32] In the classical application, Heckman noted the selection problem in assessing how factors such as education affected women's wages, since many women would not enter the labor market unless they anticipated sufficiently high wages, and expectations about wages could be influenced in part by factors correlated to the feature of interest.

In a two-stage process, we may have selection on observable or non-observable characteristics. If observable, then factors affecting selection in the first stage that also influence the second stage can in principle be modeled explicitly so that the errors in the two equations will be uncorrelated. In this case, selection can easily be modeled using a so-called two-part model (2PM).[33] Monte Carlo evidence suggests that the 2PM generally outperforms a Heckman selection model when the correlation between errors is modest,[34] and the model is also more natural when the focus is on actual rather than potential outcomes.[35] Absent incompatibilities,

[32] Achen 1986; Huth 1996; and Reed 2000.

[33] See Cameron and Triverdi 2005; Vance and Ritter 2014; Wooldridge 2002.

[34] Puhani 2000.

[35] Vance and Ritter 2014.

potential militarization is arguably undefined rather than a missing value. Moreover, the 2PM has less strict identification requirements and makes it easier to test all factors, both at the incompatibility and militarization stages.

If there is selection on unobservables, so that unmeasured factors that influence selection also affect the outcome, the observed regressors that influence selection in the first stage will not suffice to remove residual correlation between the errors in the first and second stages. The Heckman selection model considers how unobserved factors in the first stage may be correlated with unobserved factors influencing second-stage outcomes by adding the "selection hazard" or inverse Mills' ratio to the second stage to address omitted variable bias.[36] Bivariate selection models are theoretically identified without restrictions on the regressors. However, if the exact same regressors are used in the two stages, then the model is identified only through assumptions about the distribution of the residuals.[37] Sample selection models tend to work better when one can introduce plausible instruments that determine the first-stage selection but not the second-stage outcome. We argue that total population size meets both of these criteria.

Validating population size as appropriate instrument

The validity of an instrument in our two-stage estimation pertains to two concerns: (i) whether the instrument is *relevant* for Stage 1 or onset of incompatibilities and (ii) whether it only influences Stage 2 or militarization through its effect on the Stage 1 but has no independent effect (*exclusion restriction*).

[36] Heckman 1979.

[37] Cameron and Trivedi 2005, 551–2; Sartori 2003.

Population size satisfies the relevance criterion with little controversy: if individuals differ in their preferences and differences in preferences or policy views at least sometimes lead to manifest incompatibilities, then the likelihood of incompatibilities will increase with the number of individuals in a population. A larger population is likely to have more heterogeneous preferences (towards ideology, the structure of political institutions, group autonomy, language policies, taxes, etc.) than a small population. This is consistent with the empirical observation that large countries like India or Indonesia have more manifest incompatibilities than small countries like Tuvalu or Liechtenstein.

Many scholars interpret the population size-conflict relationship as reflecting preference heterogeneity, arguing that "grievances…increase with size: public choices diverge more from the preferences of the average individual as heterogeneity increases"[38]. Outside conflict research, a long tradition of studies also relates increasing diversity and cleavages to population size. Dahl and Tufte argued that small countries tend to be more homogenous and that increasing size brings diversity and conflict, as "persistent and overt differences in political outlooks, interests, and demands are likely to appear" [39]. According to Alesina,

> as countries become larger, the diversity of preferences, culture, language, "identity" of their population increases…Being part of the same country implies agreeing on a set of policies: from redistributive schemes, to public goods to foreign policy; as heterogeneity

[38] Collier and Hoeffler 2004, 572.

[39] Dahl and Tufte 1973, 13–14.

> increases, more and more diverse individuals will have to agree…and individuals or regions will be less satisfied by the central government policies.[40]

In the civil war literature, population size is typically introduced as a control variable, with little elaboration of mechanisms. Many state that civil war is more likely with a higher population without specifying whether population size affects the likelihood of initial incompatibilities or prospects for violence.[41] We see no reason why population size by itself should predict to violence, over and beyond the influence on the likelihood of incompatibilities. However, given that population size plays a prominent role in civil war research, one might worry whether it has some direct influence on opportunities for violence, violating the exclusion criterion.

Studies comparing violent and non-violent tactics provide no support for a relationship between population size and violence.[42] However, since an exclusion criterion is primarily a theoretical issue, and generally cannot be evaluated empirically, we performed a comprehensive literature search to identify all studies with population size as a predictor and the suggested mechanisms linking population to civil war onset. Searching ISI Web of Science with the key words "Civil War" AND "Population" within "Political Science" and "International Relations" returned 222 studies. Of these, 32 focus on civil conflict onset and include population size (or other population characteristics such as density) among covariates. We then identified all arguments explicitly addressing the population-conflict link. We refer to OA (83–117) for an extended discussion but comment briefly here on why we see mechanisms highlighting

[40] Alesina 2003, 304–5.

[41] Sambanis 2002; Collier and Hoeffler 2004;

[42] Chenoweth and Lewis 2013.

incompatibilities as the most relevant, and why alternative arguments fail to substantiate a plausible direct link from population size to violence.

Most studies specifying a mechanism for the population-conflict link focus on population density, geographic dispersion, or growth rather than size. Arguments focusing on opportunities for violence actually highlight how *low* population density contributes to making peripheral areas inaccessible to the state and more prone to conflict.[43] Although countries with large populations could have wider population dispersion, the Collier and Hoeffler replication data show only a modest correlation between the two.

Among the many studies that examine the population-civil war link, it is possible to find arguments relating population to militarization, either in addition to the effects on heterogeneity or incompatibilities, or through a direct impact on opportunities for violence with no role for incompatibilities. Below, we briefly discuss why we find these arguments unpersuasive (see OA for details).

Collier and Hoeffler argue that larger populations can increase recruitment through lower labor costs for rebels. This seems implausible, as larger populations and more densely populated areas tend to have higher wages. It is also unclear why lower wages would not increase government recruitment more than rebel recruitment. Larger states, subsequently, would generally require a proportionally larger force, and larger firms tend to pay higher wages.[44]

In a study associated with the dismissal of grievance's role in civil conflict, Fearon and Laitin claim:

[43] Buhaug and Tollefsen 2015; Fearon and Laitin 2003; Sambanis 2002.

[44] Brown and Medoff 1989.

> a larger country population, which makes it necessary for the center to multiply layers of agents to keep tabs on who is doing what at the local level and, also, increases the number of potential recruits to an insurgency for a given level of income.[45]

The first part of the quote suggests that states face greater logistical challenges to control large populations. This could be interpreted as *by itself* increasing the likelihood of violence, irrespective of grievances or incompatibilities. In our view, this claim is questionable and difficult to defend in the face of research demonstrating a relationship between civil war and plausible measures of grievances. Group-level studies show that relative group size—a more direct measure of opportunities—is unrelated to civil war in the absence of plausible motivation or incompatibilities.[46] A more likely interpretation is that the risk of militarization increases with population because we see more incompatibilities, as argued above. Moreover, a greater need for control can be overcome by increasing military capacity, and more populous countries tend to have larger armies and higher capacity.[47]

The second part of the quote suggests that violence is simply a function of the availability of potential recruits. We also find this debatable. The historical record shows that violent rebellions can operate effectively with small numbers,[48] and median rebel troops in the UCPD

[45] Fearon and Laitin 2003, 81.

[46] Cederman et al. 2013.

[47] Whether greater military capacity can overcome challenges to control is an interesting question in its own right, but beyond the scope of this manuscript. Efforts to examine this would need to consider both plausible incompatibilities and militarization separately, in a similar two-stage approach.

[48] Cunningham et al. 2009.

data is only 4,000. Since median non-violent campaign participation is 100,000, recruitment seems even more important for non-violence. More importantly, most analyses of recruitment stress individual grievances/motivation in shaping decision to join a rebellion rather than total stock of potential recruits.[49] People without motivation are unlikely to join an armed fight risking injury and death. Sheer numbers of individuals, absent grievances, are thus unlikely to reflect potential recruitment pools.

Finally, we can easily demonstrate that if we simulate random data for a two-stage process in line with our two stage set up, then we will find a direct effect in a separate regression in Stage 2 if we introduce a non-trivial direct effect for a postulated instrument to violate the exclusion restriction (OA, 113–7). The universality of this result should not be overstated and may not extend to an unknown data generating process. However, we think that the burden of proof should rest on skeptics to explain how a non-trival direct effect of population size on militarization over and beyond the effect through incompatibilities would not be detectable in the observed data.

Exploring the two-stage approach

To assess the plausibility of our proposed instrument and examine the various aspects of the two-stage estimation, we first replicate a study by Buhaug, Cederman, and Gleditsch.[50] This study extends prior work on how horizontal inequalities increase the risk of civil war through the potential motivation of actors,[51] showing how the group-specific insights from previous work

[49] Gates 2002.

[50] Buhaug, Cederman and Gleditsch 2014.

[51] Cederman, Gleditsch, and Buhaug 2013.

can be scaled to the country level. This is an ideal example for us, given the current prominence of research on horizontal inequalities and since we can easily extend the original setup to consider how these factors influence incompatibilities and violence, respectively. The original model includes a number of horizontal inequality measures. For conciseness, however, we only focus here on their key measure of horizontal *political* inequality, given by the population share of the largest ethnic group subject to active discrimination (LDG).

Buhaug, Cederman, and Gleditsch find that horizontal inequalities are more likely to lead to armed conflict than individual inequalities not related to group cleavages. However, the mechanisms underlying the relationship remain less clear. Inequalities may generate relative deprivation among the disadvantaged, which may motivate anti-state challenges. From this perspective, horizontal inequalities should mainly relate to incompatibilities. It is less obvious that horizontal inequalities should have a positive effect on subsequent militarization over incompatibilities, although stronger grievances shared among members of identity groups may facilitate collective action, thereby contributing to violence. A two-stage analysis can assess whether horizontal inequalities contribute to the former, the latter, or both.

We first replicate the Horizontal Inequality (HI) model and find identical results (Table 2, Model 1).[52] This model is limited to post-1960. To increase the number of observations, we replicate the model using a full sample (starting with 1946). Model 2 indicates near-identical results for the expanded data. Model 3 replicates the original HI model using CONIAS-based measure of armed conflict. The results are very similar, although some coefficients are above conventional levels of statistical significance.

[52] See Model 2 in Table 1, Buhaug, Cederman, and Gleditsch 2014. We refer to the original study for a full explanation of the terms and measures.

[Table 2]

We now turn to a two-stage model, with one equation for the onset of incompatibilities and a second for whether the incompatibilities see militarization. Following the original analysis, we include a one-year lag for incidence of incompatibilities in Equation 1 and a lag for incidence of armed conflict in Equation 2. Model 4 reports estimates for a Heckman selection probit, with population size as an instrument excluded from Stage 2. The left column shows that population size is a relevant predictor of incompatibilities. The F-test in Stage 1 exceeds 35, well above the conventional threshold of 10 for weak instruments. Although exclusion restrictions generally cannot be tested empirically, we find that adding population size at Stage 2 yields a negative and not significant coefficient on militarization, conditional on incompatibilities. This is consistent with our expectation that the entire effect on violent conflict runs through the initial incompatibilities and inconsistent with most arguments that imply a direct effect on violence.

The estimate for the correlation between the residuals ρ is modest and insignificant, and the Wald test provides no evidence for rejecting the null of no correlation. One possible interpretation is that unobserved factors influencing incompatibilities are not strongly correlated with omitted variables influencing militarization after conditioning on the observed factors. Since existing Monte Carlo studies suggest that a Heckman selection model generally performs worse than a 2PM in the absence of selection,[53] we now turn to the 2PM approach for estimating the two-stage logit model.

[53] Puhani 2000.

Model 5 has a specification identical to the Heckman Model 4, save for also including population in Stage 2. The results resemble previous estimates, unsurprisingly, given the limited evidence for selection (note that the coefficient of LDG in the logit 2PM model is larger). Both models produce clear evidence of the key measure of inequality being associated with incompatibilities but not with militarization. Figure 1 illustrates the differences in the substantive effects of LDG on the two stages. An increase from 0.2 to 0.6 in the population share of the largest discriminated group increases the likelihood of conflict origination by nearly one third (from 3.65% to 4.82%; middle panel). However, an equivalent increase in the size of the discriminated group has virtually no effect on the likelihood of conflict militarization (right panel). This non-monotonic effect on the two conflict stages is concealed in the dichotomous model (left panel). One might therefore (erroneously) conclude that LDG has an independent effect on violence.

[Figure 1]

We tested the sensitivity of our estimates to alternative coding of the dependent variable. The CONIAS violent crisis includes a number of small-scale conflicts outside the conventional definition of organized armed conflict. Therefore, we reran analyses by restricting the armed conflict category to the fourth and fifth CONIAS levels. As shown in Model 6, the estimates remain almost identical to the previous results. More importantly, to assess the validity of the CONIAS data, we also replaced CONIAS with non-violent and violent campaigns from the NAVCO dataset.[54] Although the NAVCO campaigns are conceptually and operationally quite

[54] Chenoweth and Lewis 2013.

different (i.e., restricted to events with minimum 1,000 participants), we find remarkably similar estimates (Model 7).

These results demonstrate the potential of the two-stage approach. While our aim here is not to present an elaborate theoretical interpretation of this specific finding, our estimates show that horizontal political inequalities increase the likelihood of incompatibilities but do not necessarily increase the likelihood of militarization as a specific response. We see this as having a clear value for advancing the theory on the inequality-conflict nexus, showing that the mechanism via which inequality influences civil conflict primarily relates to grievances and conflict origination rather than opportunities and conflict militarization.

Demonstrating the Value of a Two-Stage Approach through Further Replications

The section above presented and assessed the validity of data on the two stages, subsequently exploring the two-stage estimation procedure with population size as an instrument. We now proceed to a more formal application of the two-stage approach, assessing its wider applicability by replicating studies selected following more impartial criteria. To identify the candidate list, we performed a comprehensive search of articles published in *International Organization* (IO) over the last decade. Our search criteria was onset of civil war as a dependent variable and a time-series cross-sectional design with country-year as a unit of analysis. This search has identified six candidate studies. Further scrutiny (see OA, 118–120) narrowed down this list to two articles with most appropriate research designs.[55]

Revisiting Refugees and Soft Power in Incompatibilities and Armed Conflict

[55] Salehyan and Gleditsch 2006; Warren 2014.

We first replicate Salehyan and Gleditsch, investigating the effects of refugees on violent civil conflict. Building on previous research, showing that violent conflict in one country influences the risk of violence in neighboring countries, Salehyan and Gleditsch argue that population movement is one likely mechanism for the transnational spread of conflict. The authors identify several mechanisms through which refugee flows may cause conflict in host countries. First, moving populations may "import" conflict-specific capital such as arms and combatants, as well as ideologies favoring violence (*Mechanism 1*). Second, large flows of refugees, particularly from a different ethnic background, may lead host populations to feel threatened, as well as increase competition over local economic resources (*Mechanism 2*). If *Mechanism 1* holds true, then we should find refugees having a stronger effect on Stage 2; conversely, if *Mechanism 2* holds true, then we should see refugees having a stronger effect on Stage 1.

Salehyan and Gleditsch find that the log of refugees from neighboring states significantly predicts outbreak of civil conflict, even when controlling for conflict in the neighborhood and trans-border ethnic kin. The fact that they only find a positive effect for refugees from neighboring countries suggests that refugees increase the risk of violence primarily in cases where incompatibilities already exist. Although couched more in terms of integration than incompatibilities, Bolfrass, Shaver, and Zhou emphasize that conditions in recipient countries (rather than characteristics of their place origin) determine whether refugees constitute a security risk.[56]

[Table 3]

[56] Bolfrass, Shaver, and Zhou 2015.

Model 8 in Table 3 replicates the main model in Salehyan and Gleditsch.[57] Changing to the CONIAS armed conflict yields similar effects for most covariates, including refugees (Model 9). Subsequently, we replicate the analysis following our two-stage approach. First, we employ the selection model (Model 10). Consistent with previous estimates, population significantly influences Stage 1 (an additional analysis indicates that population in Stage 2 yields a negative, non-significant coefficient). The key variable of interest, refugees, only significantly affects Stage 2. The estimates provide little support for selection effects, however, indicating that 2PM is more appropriate. Unsurprisingly, 2PM yields similar estimates (Model 11): population only significantly affects Stage 1 and refugees only significantly affect Stage 2. Unlike selection model, 2PM also yields a significant coefficient for trans-border ethnic kin.

This suggests that transnational shocks and ethnic ties primarily influence militarization when incompatibilities already exist. This, subsequently, supports the above-discussed *Mechanism 1,* indicating that refugees primarily affect civil conflict via the "import" of conflict-specific capital to existing incompatibilities and that refugees are unlikely to generate new incompatibilities with the host state, inconsistent with *Mechanism 2*.[58]

We now turn to the Warren study, investigating the relationship between state capacity and civil war onset. Unlike previous studies focusing on the role of "hard power" or coercive state capacity in deterring conflict, Warren argues that "soft power" and voluntary state compliance by citizens induced through political communication is a critical factor explaining

[57] Salehyan and Gleditsch 2006, Model 3 in Table 5.

[58] See also Bolfrass, Shaver, and Zhou 2015; Bove and Böhmelt 2016. We may see increasing violence against refugees, but this is distinct from organized civil violence and may not involve the government.

(the absence of) civil war. Warren hypothesizes that the expansion of mass communication technologies (TV, radio, and newspapers) gives an edge to governments over the opposition in spreading normative influence and strengthening state loyalty among citizens. This produces barriers to mobilization for violence against a state. Warren demonstrates that his Media Density Index (MDI) reduces the likelihood of civil war onset "more than a tenfold", claiming that the "mass media infrastructure represents one of the most powerful forces for peace and stability yet observed in the modern world".[59]

Does mass communication expansion reduce the likelihood of incompatibilities or does it reduce the likelihood of militarization over incompatibilities? Warren could be read as arguing that state loyalty among the citizenry undermines violent mobilization by reducing divergent preferences (*Mechanism 1*). However, mass media can also expand the opportunities for effective non-violent mobilization relative to violence,[60] and soft power might reduce the legitimacy of violent means without reducing incompatibilities (*Mechanism 2*). If so, we should expect the negative effect of MDI on civil war to work by reducing the likelihood of militarization rather than incompatibilities.

[Table 4]

Model 12 in Table 4 replicates the main model in Warren.[61] When we replicate the same model with the CONIAS armed conflict, the effects remain similar (Model 13), although MDI

[59] Warren 2014, 113.

[60] Weidmann 2015.

[61] Warren 2014, Model 3 in Table 1.

has a smaller coefficient, likely due to Warren using Sambanis' civil war measure (with a high 1,000-death threshold). Replicating the model with more similar CONIAS 4–5 increases the coefficient for MDI (Model 14).

In the two-stage estimation, using the selection model (Model 15), we again find that population size only significantly affects Stage 1. However, the key measure, MDI, does not attain significance in both stages. Yet, consistent with previous replications, the model shows no significant selection, suggesting that the 2PM is more appropriate. Model 16 shows that the 2PM yields similar results, with population size and most other covariates having significant effects in the corresponding stages. Now, however, the coefficient for MDI attains significance at the 10% level ($p = 0.077$). This supports the above-discussed *Mechanism 2*, indicating that the pacifying effect of mass media exerts its influence via reduced likelihood of militarization rather than decreasing incompatibilities (inconsistent with *Mechanism 1*).

Figure 2 illustrates differences in the substantive effects of the key variables employed in the two replications, estimated following the two-stage approach. Following the dichotomous models (left panels), one might conclude that refugees and expanding mass media influence civil conflict via mechanisms relating to incompatibilities. The middle and right panels show why this might be incorrect, vindicating mechanisms that relate the two variables to civil conflict specifically via violence.

[Figure 2]

In addition to the main results, we find notable patterns in the effects of other covariates. Throughout all replications, we find that GDP per capita clearly has stronger effects on Stage 2

than Stage 1. This can be interpreted as support for the argument that poverty aids rebel recruitment,[62] as contrasted with the arguments that interpret the low GDP per capita–civil war relationship as reflecting that poverty generates incompatibilities.[63] We also find that the prominent curvilinear relationship between democracy and civil conflict is more evidenced in Stage 1 (although this is more apparent in the replication of the Warren study).

Beyond the Domestic War and Peace Dichotomy

We have shown how a two-stage approach to civil conflict, highlighting conflict origination and militarization as separate stages, provides a useful alternative to binary conceptions of civil war. Using new data on incompatibilities and violence, we have replicated three studies of civil conflict using the two-stage approach. These results provide new insights into the existing debates. In particular, our analysis has shown that the link between horizontal inequalities and civil conflict primarily reflects the effects on conflict origination and that grievances by themselves do not facilitate violence. Refugee flows can contribute to the spread of violence between neighboring countries but primarily facilitate conflict in countries with existing incompatibilities.[64] Finally, the effects of media density on decreasing violence apparently exert themselves through militarization. This calls for new research on the relative weight of how media density relates to alternative forms of non-violent mobilization versus norms against the use of violence *per se*.

[62] Collier and Hoeffler 2013.

[63] Buhaug et al. 2011.

[64] This runs counter to some of the alarmist claims over the security implications of refugees in European states extrapolated from the original Salehyan and Gleditsch results.

Altogether, we see these results as providing ample fodder for further theorizing and refining empirical results. Here, however, our primary aim is not to advance particular debates in specific studies but rather to show the overall potential of the two-stage approach, which can be applied to study any topic in research on civil war.

We believe our study underscores a number of other, general implications for conflict research. As demonstrated above, the reasons why conflicts start and why they militarize cannot simply be assumed to be the same. A two-stage approach may help account for this, allowing us to understand the relevant causes at different stages. Similarly, researchers must consider how features can have non-uniform effects on conflict origination and militarization. Sometimes, opposing effects may even average each other out in dichotomous models, leading to misplaced conclusions that no relationship exists. Careful attention to divergent effects on particular conflict stages can thus potentially explain inconsistencies in previous studies.[65]

Conflict researchers could also use the two-stage approach to assess alternative mechanisms through which explanatory variables potentially influence civil conflict. While the two-stage approach cannot confirm or reject particular causal mechanisms, it may serve as a plausibility probe, providing clues as to more or less plausible alternatives. This, in turn, can lay the groundwork for subsequent in-depth research on actual causal mechanisms.

Finally, a two-stage framework can help inform data collection strategies, both with regards to identifying predictors of incompatibilities, and by limiting other efforts to collect data on attributes that can be observed for incompatibilities at the second-stage. Although we have found CONIAS helpful in delineating incompatibilities, this is only one approach and other alternatives can be useful. Moreover, researchers may be interested in different second-stage

[65] Dixon 2009.

outcomes, such as large-scale non-violent mobilization, and the basic framework presented here can be expanded to multiple outcomes and several sequential stages.

Although we have focused on countries as units of analysis, one might also investigate incompatibilities and violence below the country-level, including dyads of governments and self-determination groups or dissident organizations, or even between non-state actors. Actor-level analyses may also help incorporate actor-specific strategic considerations, advancing disaggregated analyses of conflict.

References

Achen, Christopher. 1986. *The Statistical Analysis of Quasi-Experiments.* Berkeley: University of California Press.

Alesina, Alberto. 2003. The Size of Countries: Does it Matter? *Journal of the European Economic Association* 1 (2-3):301–16.

Bolfrass, Alexander, Andrew Shaver, and Yang-Yang Zhou. 2015. Don't Fear Refugees: Why They Pose Little Threat to National Security. *Foreign Affairs,* available at http://www.tinyurl.com/hgjpgm2.

Boulding, Kenneth. 1962. *Conflict and Defense*. New York: Harper.

Bove, Vincenzo, and Tobias Böhmelt. 2016. Does Immigration Induce Terrorism? *Journal of Politics* 78 (2):572–88.

Brown, Charles, and James Medoff. 1989. The Employer Size Wage Effect. *Journal of Political Economy* 97 (5):1027–1059.

Buhaug, Halvard, Lars-Erik Cederman, and Kristian Skrede Gleditsch. 2014. Square Pegs in Round Holes: Inequalities, Grievances, and Civil War. *International Studies Quarterly* 58 (2):418–31.

Buhaug, Halvard, and Andreas Forø Tollefsen. 2015. Insurgency and Inaccessibility. *International Studies Review* 17 (1):6–25.

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

Cederman, Lars-Erik, Kristian S. Gleditsch, and Halvard Buhaug. 2013. *Inequality, Grievances, and Civil War*. Cambridge: Cambridge University Press.

Chenoweth, Erica, and Orion A. Lewis. 2013. Unpacking Nonviolent Campaigns: Introducing the NAVCO 2.0 Dataset. *Journal of Peace Research* 50 (3):415–23.

Chenoweth, Erica, and Kathleen Gallagher Cunningham. 2013. Understanding Nonviolent Resistance: An Introduction. *Journal of Peace Research* 50 (3):271–76.

Chiozza, Giacomo. 2002. Is There a Clash of Civilizations? Evidence from Patterns of International Conflict Involvement, 1946–97. *Journal of Peace Research* 39 (6):711–34.

Collier, Paul, and Anke Hoeffler. 2004. Greed and Grievance in Civil War. *Oxford Economic Papers* 56 (4):563–95.

Cunningham, Kathleen Gallagher. 2013. Understanding Strategic Choice: The Determinants of Civil War and Nonviolent Campaign in Self-Determination Disputes. *Journal of Peace Research* 50 (3):291–304.

Cunningham, David E., Kristian Skrede Gleditsch, and Idean Salehyan. 2009. It Takes Two: A Dyadic Analysis of Civil War Duration and Outcome. *Journal of Conflict Resolution* 53 (4): 570–97.

Dahl, Robert A., and Edward R. Tufte. 1973. *Size and Democracy.* Stanford: Stanford University Press.

Dahrendorf, Ralf. 1959. *Class and Class Conflict in Industrial Society*. Stanford: Stanford University Press.

Davis, William. 2012. Swords into Ploughshares: The Effect of Pacifist Public Opinion on Foreign Policy in Western Democracies. *Cooperation and Conflict* 47 (3):309–30.

Dixon, Jeffrey. 2009. What Causes Civil Wars? Integrating Quantitative Research Findings. *International Studies Review* 11 (4):707–35.

Fearon, James. 1994. Signaling versus the Balance of Power and Interests. *Journal of Conflict Resolution* 38 (2):236–69.

Fearon, James, and David Laitin. 2003. Ethnicity, Insurgency, and Civil War. *American Political Science Review* 97 (1):75–90.

Gates, Scott. 2002. Recruitment and Allegiance: The Microfoundations of Rebellion. *Journal of Conflict Resolution* 46 (1):111–30.

Gleditsch, Nils Peter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Håvard Strand. 2002. Armed Conflict 1946–2001: A New Dataset. *Journal of Peace Research* 39 (5):615–37.

Goertz, Gary. 2005. *Social Science Concepts: A User's Guide*. Princeton: Princeton University Press.

Heckman, James J. 1979. Sample Selection Bias as a Specification Error. *Econometrica* 47 (1):153–61.

Hegre, Håvard, and Nicholas Sambanis. 2006. Sensitivity Analysis of Empirical Results on Civil War Onset. *Journal of Conflict Resolution* 50 (4):508–35.

Heidelberg Institute for International Conflict Research. 2015. *Conflict Barometer 2014*. Heidelberg, Germany: HIIK.

Hensel, Paul R., Sara McLaughlin Mitchell, Thomas E. Sowers II, and Clayton L. Thyne. 2008. Bones of Contention: Comparing Territorial, Maritime, and River Issues. *Journal of Conflict Resolution* 52 (1):117–43.

Huth, Paul K. 1996. *Standing Your Ground: Territorial Disputes and International Conflict*. Ann Arbor: University of Michigan Press.

Jones, Daniel M., Stuart A. Bremer, and J. David Singer. 1996. Militarized Interstate Disputes, 1816-1992: Rationale, Coding Rules, and Empirical Patterns. *Conflict Management and Peace Science* 15 (2):163–213.

Lichbach, Mark. 1989. An Evaluation of "Does Economic Inequality Breed Political Conflict?" Studies. *World Politics* 41 (4):431–72.

Melander, Erik, Therése Pettersson, and Lotta Themnér. 2016. Organized Violence, 1989–2015. *Journal of Peace Research* 53 (5):727–42.

O'Brien, Sean P. 2002. Anticipating the Good, the Bad, and the Ugly: An Early Warning Approach to Conflict and Instability Analysis. *Journal of Conflict Resolution* 46 (6):791–811.

Pfetsch, Frank R. 2006. Old Wine in New Bottles: Democratic Peace as Empowerment of States in Conflict Resolution. *European Journal of Political Research* 45 (5):811–50.

Pfetsch, Frank R., and Christoph Rohloff. 2000. KOSIMO: A Databank on Political Conflict. *Journal of Peace Research* 37 (3):379–89.

Puhani, Patrick. 2000. The Heckman Correction for Sample Selection and its Critique. *Journal of Economic Surveys* 14 (1):53–68.

Reed, William. 2000. A Unified Statistical Model of Conflict Onset and Escalation. *American Journal of Political Science* 44 (1):84–93.

Salehyan, Idean, and Kristian Skrede Gleditsch. 2006. Refugee Flows and the Spread of Civil War. *International Organization* 60 (2):335–66.

Salehyan, Idean, Cullen S. Hendrix, Jesse Hamner, Christina Case, Christopher Linebarger, Emily Stull, and Jennifer Williams. 2012. Social Conflict in Africa: A New Database. *International Interactions* 38 (4): 503–11.

Sambanis, Nicholas. 2002. A Review of Recent Advances and Future Directions in the Literature on Civil War. *Defence and Peace Economics* 13 (3):215–43.

———. 2004. What Is Civil War? Conceptual and Empirical Complexities of an Operational Definition. *Journal of Conflict Resolution* 48 (6):814–58.

Sartori, Anne E. 2003. An Estimator for Some Binary-Outcome Selection Models without Exclusion Restrictions. *Political Analysis* 11 (2):111–38.

Schwank, Nicolas, Christian Baukhage, Peer Böhrnsen, Daniel Braner, Janine Conrad, Jan Deuter, David Epp, Yolanda Fernandez, Alan Götz, Natalie Hoffmann, Jens Hofmann, Pamela Jawad, Steffen Kassner, Stefanie Lerche, Anne Lübbers, Lotta Mayer, Holger Oswald, André Rapp, Friedemann Schirrmeister, Gerrit Schlomach, Rüdiger Schwarz, Suat Selcuk, Christoph Trinn, Thomas Wencker, Katja Wezel, Florian Winkler, and Silke Zaulig. 2013. CONIAS Dataset 1945–2008, Year Intensity V1.0. Heidelberg. Available at <http://test.conis-group.org/en/data_sets/available_data.html>. Accessed 28 June 2016.

Small, Melvin, and David J. Singer. 1982. *Resort to Arms: International and Civil War, 1816–1980*. Beverly Hills, CA: Sage.

Vance, Colin, and Nolan Ritter. 2014. Is Peace a Missing Value or a Zero? On Selection Models in Political Science. *Journal of Peace Research* 51 (4):528–40.

Wallensteen, Peter. 2015. *Understanding Conflict Resolution.* London: Sage.

Warren, T. Camber. 2014. Not by the Sword Alone: Soft Power, Mass Media, and the Production of State Sovereignty. *International Organization* 59 (1):39–75.

Weidmann, Nils B. 2015. Communication, Technology, and Political Conflict: Introduction to the Special Issue. *Journal of Peace Research* 52 (3):263–68.

White, Peter B., Dragana Vidovic, Belén González, Kristian Skrede Gleditsch, and David E. Cunningham. 2015. Nonviolence as a Weapon of the Resourceful: From Claims to Tactics in Mobilization. *Mobilization: An International Quarterly* 20 (4):471–91.

Wooldridge, Jeffrey. 2002. *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: MIT Press.

Table 1. The CONIAS Incompatibility Intensity Levels with Examples

Level	Name	Examples
1	Dispute	Gov. Spain vs. Catalan Nationalists (1979–2008)
2	Non-violent crisis	Gov. Moldova vs. Pridnesestrovian Moldavian Rep. (1993–2008)
3	Violent crisis	Gov. Spain versus Euskadi Ta Askatasuna/ETA (1968–2008)
4	Limited war	Gov. United Kingdom versus Irish Nationalists (1968–1998)
5	War	Afghan Civil War (1978–1994)

Table 2. Replication of Buhaug, Cederman, and Gleditsch (2014) using the two-stage approach

	(1) Original BCG	*(2) Expanded BCG*	*(3) Expanded BCG CONIAS*	*(4) Heckman CONIAS 3-5*		*(5) 2PM CONIAS 3-5*		*(6) 2PM CONIAS 4-5*		*(7) 2PM NAVCO*	
				St. 1	*St. 2*	*St. 1*	*St. 2*	*St. 1*	*St. 2*	*St. 1*	*St. 2*
ELF	**0.974***	**0.838***	**0.659***	0.207	-0.406	0.556	0.012	0.556	-0.288	0.090	-0.186
	(0.428)	**(0.412)**	**(0.335)**	(0.155)	(0.414)	(0.352)	(0.424)	(0.352)	(0.479)	(0.340)	(0.756)
Gini	-0.004	-0.000	0.007	-0.000	0.017	-0.002	0.006	-0.002	0.002	-0.011	-0.001
	(0.010)	(0.011)	(0.007)	(0.004)	(0.011)	(0.008)	(0.010)	(0.008)	(0.012)	(0.008)	(0.017)
LDG	**1.288*****	**0.823***	**0.818***	**0.352**$^{+}$	0.760	**0.722**$^{+}$	-0.147	**0.722**$^{+}$	-0.170	**0.924***	1.279
	(0.346)	**(0.359)**	**(0.404)**	**(0.190)**	(0.521)	**(0.412)**	(0.462)	**(0.412)**	(0.613)	**(0.376)**	(0.814)
PHI	-0.045	-0.101	0.137	0.046	0.219	0.043	0.049	0.043	-0.053	0.130	-0.072
	(0.175)	(0.139)	(0.091)	(0.076)	(0.221)	(0.143)	(0.137)	(0.143)	(0.175)	(0.135)	(0.272)
NHI	**0.321****	**0.333****	0.027	0.059	-0.081	0.128	0.117	0.128	0.064	0.067	0.153
	(0.119)	**(0.121)**	(0.071)	(0.043)	(0.115)	(0.084)	(0.127)	(0.084)	(0.141)	(0.139)	(0.281)
Downgrade	**0.860*****	**0.858*****	0.367	0.214	0.033	0.460	0.314	0.460	0.438	**1.116*****	0.546
	(0.255)	**(0.250)**	(0.285)	(0.155)	(0.404)	(0.326)	(0.301)	(0.326)	(0.268)	**(0.289)**	(0.551)
Power-sharing	-0.029	0.005	-0.145	-0.050	0.295	-0.098	0.068	-0.098	0.091	0.062	-0.001
	(0.221)	(0.221)	(0.173)	(0.078)	(0.257)	(0.169)	(0.209)	(0.169)	(0.231)	(0.206)	(0.379)
Democracy	0.350	0.285	0.079	-0.066	-0.043	-0.124	-0.032	-0.124	0.216	**-1.024****	0.799
	(0.345)	(0.334)	(0.184)	(0.122)	(0.325)	(0.279)	(0.245)	(0.279)	(0.275)	**(0.327)**	(0.557)
Population	**0.234****	**0.231*****	**0.393*****	**0.144*****		**0.321*****	0.149	**0.321*****	0.108	**0.322*****	-0.201
	(0.079)	**(0.067)**	**(0.055)**	**(0.033)**		**(0.070)**	(0.098)	**(0.070)**	(0.076)	**(0.062)**	(0.162)
GDP per	**-0.432****	**-0.467*****	**-0.211***	-0.062	-0.059	-0.128	**-0.214**$^{+}$	-0.128	**-0.501*****	-0.136	**-0.920*****
capita	**(0.147)**	**(0.126)**	**(0.101)**	(0.048)	(0.115)	(0.108)	**(0.114)**	(0.108)	**(0.132)**	(0.100)	**(0.189)**
Constant	**-6.311*****	**-6.275*****	**-7.442*****	**-3.247*****	-1.030	**-6.472*****	**-3.367****	**-6.472*****	**-3.389****	**-6.042*****	1.707
	(0.850)	**(0.810)**	**(0.745)**	**(0.344)**	(1.173)	**(0.754)**	**(1.135)**	**(0.754)**	**(1.068)**	**(0.656)**	(1.677)
ρ				0.182							
				(0.557)							
N	5219	6111	6058	6058		6058		6058		6058	
AIC	1275.352	1476.008	2209.011	2264.755							

Clustered standard errors in parentheses

Variables accounting for time-dependence (i.e., one year dependent variable lags) not reported

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3. Replication of Salehyan and Gleditsch (2006) using the two-stage approach

	(8)	*(9)*	*(10)*		*(11)*	
	Original SG	*SG CONIAS 3-5*	*Heckman CONIAS 3-5*		*2PM CONIAS 3-5*	
			St. 1	*St. 2*	*St. 1*	*St. 2*
Refugees	**0.033***	**0.029**$^{+}$	0.012	**0.074***	0.026	**0.033**$^{+}$
	(0.014)	**(0.017)**	(0.008)	**(0.033)**	(0.018)	**(0.018)**
Civil war in neighbor	**0.435***	0.114	-0.049	-0.253	-0.104	0.042
	(0.176)	(0.180)	(0.081)	(0.247)	(0.181)	(0.194)
Ethnic kin	**0.531****	0.256	0.111	0.171	0.246	**0.350**$^{+}$
	(0.191)	(0.182)	(0.085)	(0.243)	(0.193)	**(0.206)**
Polity	0.012	0.000	0.006	0.005	0.016	-0.017
	(0.013)	(0.014)	(0.007)	(0.019)	(0.016)	(0.015)
Polity squared	**-0.011*****	**-0.005**$^{+}$	**-0.005*****	-0.006	**-0.010*****	-0.002
	(0.003)	**(0.003)**	**(0.001)**	(0.005)	**(0.003)**	(0.003)
GDP per capita	**-0.170**$^{+}$	-0.123	-0.048	-0.154	-0.119	**-0.252****
	(0.100)	(0.104)	(0.052)	(0.137)	(0.116)	**(0.097)**
Population	**0.222*****	**0.252*****	**0.071***		**0.165****	0.076
	(0.045)	**(0.054)**	**(0.029)**		**(0.060)**	(0.051)
Ethnic heterogeneity	**0.016****	**0.009**$^{+}$	0.003	-0.010	0.007	**-0.013***
	(0.005)	**(0.005)**	(0.003)	(0.012)	(0.006)	**(0.006)**
Constant	**-2.719****	**-4.692*****	**-1.922*****	0.601	**-3.551****	-1.301
	(0.915)	**(0.991)**	**(0.522)**	(3.452)	**(1.137)**	(1.078)
ρ			0.447			
			(1.721)			
N	3591	3591	3591		3591	
AIC	1341.454	1322.930	1457.067			

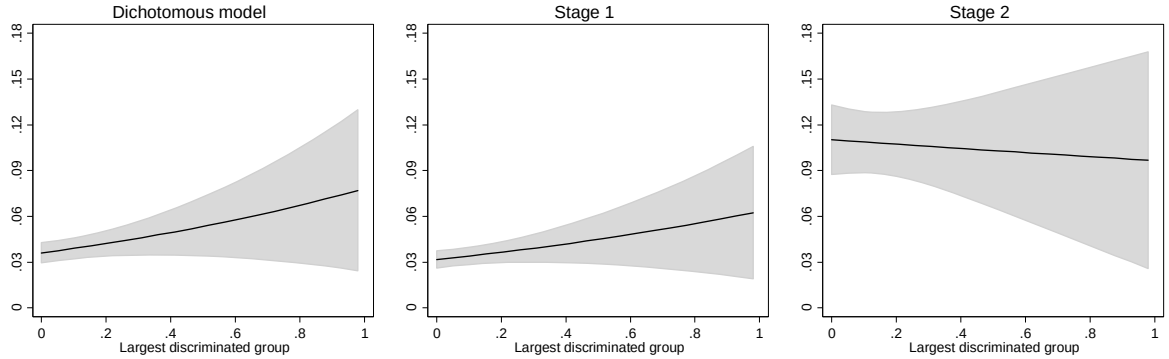
Robust standard errors in parentheses
Peace years and splines not reported
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4. Replication of Warren (2014) using the two-stage approach

	(12) *Original Warren*	*(13)* *Warren CONIAS 3-5*	*(14)* *Warren CONIAS 4-5*	*(15)* *Heckman CONIAS 4-5*		*(16)* *2PM CONIAS 4-5*	
				St. 1	*St. 2*	*St. 1*	*St. 2*
Media Density Index	**-0.026*****	**-0.006**$^{+}$	**-0.008**$^{+}$	-0.001	0.002	-0.002	**-0.007**$^{+}$
	(0.008)	**(0.003)**	**(0.004)**	(0.001)	(0.005)	(0.003)	**(0.004)**
GDP per capita	-0.038	-0.146	**-0.398***	-0.063	**-0.459****	-0.168	**-0.508****
	(0.188)	(0.182)	**(0.198)**	(0.061)	**(0.160)**	(0.141)	**(0.190)**
Area	-0.089	-0.070	-0.007	-0.026	-0.008	-0.053	0.027
	(0.089)	(0.067)	(0.101)	(0.030)	(0.087)	(0.068)	(0.101)
Mountanous terrain	**0.109***	**0.136*****	**0.135****	**0.037***	-0.002	**0.082***	0.043
	(0.049)	**(0.039)**	**(0.047)**	**(0.018)**	(0.077)	**(0.040)**	(0.069)
Population	**0.274*****	**0.412*****	**0.300*****	**0.161*****		**0.360*****	0.078
	(0.071)	**(0.070)**	**(0.071)**	**(0.040)**		**(0.087)**	(0.090)
Oil exporter	**0.756*****	0.193	**0.522***	0.088	0.175	0.199	**0.478**$^{+}$
	(0.225)	(0.227)	**(0.245)**	(0.104)	(0.339)	(0.235)	**(0.247)**
Democracy	**0.180***	**0.137***	0.083	**0.099*****	-0.034	**0.227*****	0.003
	(0.079)	**(0.063)**	(0.073)	**(0.027)**	(0.091)	**(0.064)**	(0.077)
Democracy squared	**-0.007***	**-0.005***	-0.003	**-0.004*****	0.001	**-0.010****	0.001
	(0.003)	**(0.003)**	(0.003)	**(0.001)**	(0.004)	**(0.003)**	(0.003)
Ethnic fractionalization	0.216	0.405	0.024	0.202	-0.078	0.484	-0.631
	(0.351)	(0.301)	(0.364)	(0.146)	(0.414)	(0.315)	(0.416)
Religious fractionalization	**1.381***	-0.237	0.216	0.040	**1.104***	0.057	0.419
	(0.540)	(0.450)	(0.572)	(0.168)	**(0.536)**	(0.385)	(0.597)
Constant	**-8.435*****	**-10.001*****	**-8.236*****	**-4.694*****	-0.319	**-9.901*****	**-3.821*****
	(1.083)	**(1.124)**	**(0.995)**	**(0.547)**	(2.073)	**(1.210)**	**(1.128)**
ρ				-0.045			
				(0.542)			
N	5899	5888	5888	5888		5888	
AIC	1085.098	1924.729	1359.175	1952.646			

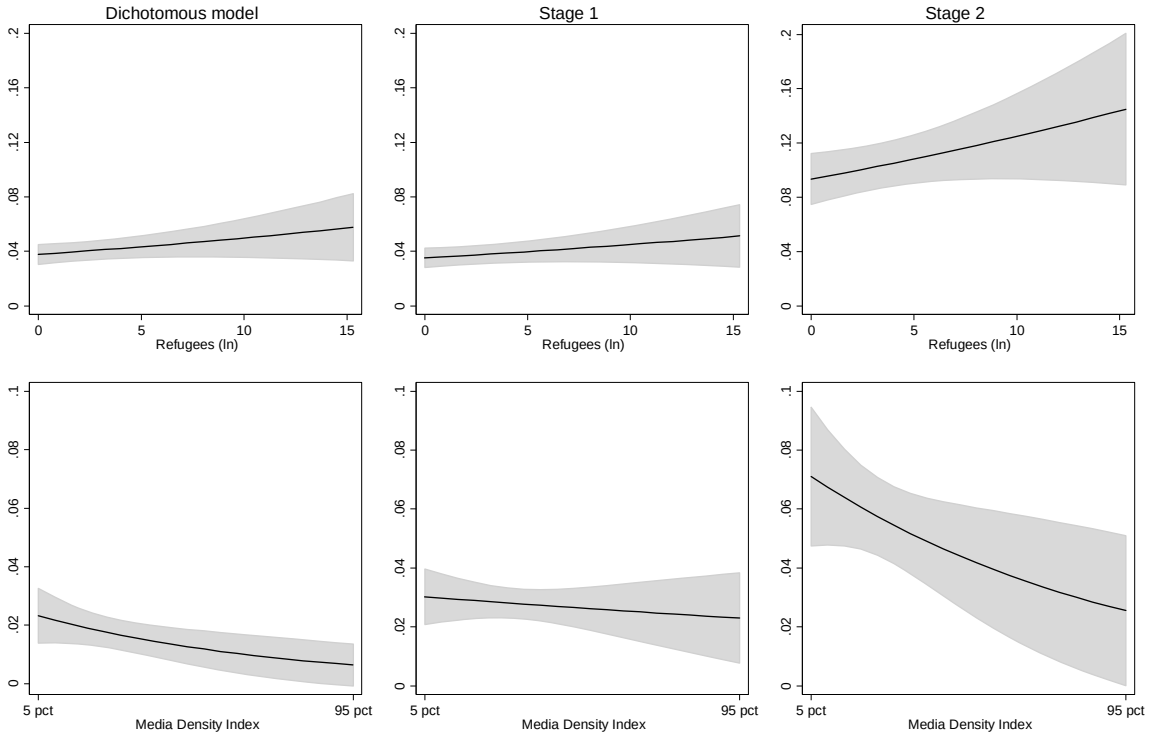
Clustered standard errors in parentheses
Peace years and splines not reported
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figures



Notes: The figure shows change in the probabilities (with 95th confidence intervals) of civil conflict onset, conflict origination, and conflict militarization as a function of *largest discriminated group* (LDG). The estimates are based on Models 3 and 5 (Table 2) when other variables are held at their mean values.

Figure 1. *Predicted probabilities of civil conflict onset, conflict origination, and conflict militarization*



Notes: The figure shows change in the probabilities (with 95th confidence intervals) of civil conflict onset, conflict origination, and conflict militarization as a function of *refugees* and *Media Density Index*. The estimates are based on Models 9 and 11 (Table 3), and 14 and 16 (Table 4), respectively, when other variables are held at their mean values.

Figure 2. *Predicted probabilities of civil conflict onset, conflict origination, and conflict militarization*