

Triangulating horizontal inequality: Toward improved conflict analysis

Lars-Erik Cederman

International Conflict Research, ETH Zürich

Nils B Weidmann

Department of Politics and Public Administration, University of Konstanz

Nils-Christian Bormann

International Conflict Research, ETH Zürich

Journal of Peace Research
2015, Vol. 52(6) 806–821
© The Author(s) 2015
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0022343315597969
jpr.sagepub.com



Abstract

Does economic inequality cause civil war? Deviating from individualist measures of inequality such as the Gini coefficient, recent studies have found a statistical link between group-level inequalities and conflict onset. Yet, this connection remains controversial, not least because of the difficulties associated with conceptualizing and measuring group-level differences in development. In an effort to overcome weaknesses afflicting specific methods of measurement, we introduce a new composite indicator that exploits the strengths of three sources of data. The first step of our method combines geocoded data from the G-Econ project with night lights emissions data from satellites. In a second step, we bring together the combined spatial values with survey estimates in order to arrive at an improved measure of group-level inequality that is both more accurate and more robust than any one of the component measures. We evaluate the effect of the combined indicator and its components on the onset of civil violence. As expected, the combined index yields stronger results as more information becomes available, thus confirming the initial hypothesis that horizontal economic inequality does drive conflict in the case of groups that are relatively poor compared to the country average. Furthermore, these findings appear to be considerably more robust than those relying on a single data source.

Keywords

civil war, ethnicity, geographic information systems, inequality, remote sensing

Introduction

Does economic inequality cause civil conflict? While there is a growing scholarly consensus that inequality has a number of deleterious social consequences, such as underdevelopment (Alesina & Rodrik, 1994), bad governance (Kyriacou, 2013), deficient public good provisions (Baldwin & Huber, 2010), and authoritarian rule (Acemoglu & Robinson, 2005), the empirical literature is much more inconclusive as regards internal conflict. Summing up more than two decades of intensive research, Lichbach (1989) concluded that there was at best very mixed evidence for this effect. More recently, prominent studies interpret the lack of statistical findings

as a sign that grievance-based explanations do not hold (e.g. Fearon & Laitin, 2003; Collier & Hoeffler, 2004).

However, rather than offering support for the dismissal of inequality as a source of conflict, the weak evidence may stem from the use of individual-level measures such as the Gini coefficient (Cramer, 2003). In fact, researchers who adopt a group-level approach have been able to detect a strong effect of inequality on civil war violence. Pioneering this type of research under the heading of 'horizontal' inequalities, Stewart

Corresponding author:

lcederman@ethz.ch

and her colleagues argue that inequalities between culturally defined groups are more likely to trigger conflict than 'vertical' or individual-based disparities (Stewart, 2008).

Along similar lines, other studies have been able to find more general evidence for this proposition by relying on survey data (e.g. Østby, 2008a). Yet, while surpassing case studies in generality, data limitations prevent surveys from providing truly worldwide coverage. Those studies that attempt to go beyond a regional scope are forced to patch together information from separate survey exercises conducted with widely different instruments (Huber & Mayoral, 2013).

To overcome these difficulties, Cederman, Weidmann & Gleditsch (2011) and Cederman, Gleditsch & Buhaug (2013) use geocoded information on ethnic groups' settlement areas in combination with spatial income estimates from the G-Econ dataset (Nordhaus, 2006). The spatial method applied by these studies offers the first truly global sample of economic horizontal inequality, but it is also associated with limitations. On the one hand, the G-Econ income estimates are of low quality in developing states where the official statistics are poor (Chen & Nordhaus, 2011). On the other, the estimation method leads to measurement error in cases where the ethnic groups' settlement areas overlap to a large degree.

Given these empirical complications, it does not come as a surprise that the effect of 'between-group' inequality on civil war onset remains controversial in the literature. Based on theoretical arguments, Esteban, Mayoral & Ray (2012) expect that inequalities within rather than between ethnic groups drive conflict, a claim that they attempt to validate with survey-based data at the country level. Using a more extensive, disaggregated dataset of the same type, Huber & Mayoral (2013) come to the same conclusion (see also Mitra & Ray, 2014). However, because of the inherent weaknesses of compilations based on several distinct surveys, these findings can hardly be considered conclusive.

To find a firmer basis of comparison, it therefore makes sense to look beyond surveys and the G-Econ data for additional information sources. In particular, satellite data on light emissions may offer a viable proxy for economic activities, especially in areas where the G-Econ data provide inaccurate information. As shown in the next section, researchers have already applied this method to estimating economic development and ethnic inequality, but so far, no study has systematically exploited this type of data to evaluate the link between group-level inequality and conflict.

While promising, luminosity-based estimates are no panacea. Very much as the two other measurement

methods, night light proxies for economic activities are afflicted by considerable weaknesses, such as saturation effects (Chen & Nordhaus, 2011). In addition, remote-sensing measures of ethnic inequality are as dependent on spatial separation of group settlement areas as is the aforementioned G-Econ-based spatial approach. In this respect, survey instruments offer significant advantages and may therefore serve as a useful complement to spatial measurement strategies.

Since there is no perfect way to measure group-level inequalities, this article proposes a composite indicator that draws on all three approaches in an effort to exploit their comparative advantages. We do so by first combining the G-Econ estimates with data from satellite observations. In a second step, we bring together the combined spatial values with survey estimates in order to arrive at an improved estimate of group-level inequality that is both more accurate and more robust than any one of the component measures. Evaluating its effect on civil violence, we show that the combined indicator yields stronger results as more information is added. This confirms the initial hypothesis that horizontal economic inequality does drive conflict, especially in the case of relatively poor groups. Furthermore, these findings appear to be considerably more robust than those relying on a single data source.

Research on economic inequality and civil war

The qualitative literature offers many examples of studies that associate relative poverty of ethnic groups with an increased risk of domestic conflict (Horowitz, 2000; Wood, 2003; Sambanis, 2004). Other scholars argue that wealthier ethnic groups such as the Basques in Spain or the Slovenes in Yugoslavia are as likely to rebel as deprived groups because better-off groups feel that they contribute disproportionately to the central state's budget (Gourevitch, 1979; Horowitz, 2000). We summarize these expectations in two hypotheses:

H1: Relatively poor ethnic groups are more likely to rebel than those that are closer to the country average.

H2: Relatively wealthy ethnic groups are more likely to rebel than those that are closer to the country average.

Testing these hypotheses quantitatively is challenging due to lack of data. This is why most conflict researchers have used measures of economic inequality that implicitly or explicitly assume an individualist perspective, such as the Gini coefficient (Cramer, 2003). In contrast, there have been relatively few attempts to systematically

capture economic inequality among ethnic groups. In a pioneering study, Barrows (1976) documented some influence of economic inequality among groups in sub-Saharan Africa. Until recently, most conflict researchers exploring group-level comparisons have relied on Gurr's (1993) Minorities at Risk (MAR) dataset, which provides indicators of group-level inequality, such as economic 'disadvantages'. Yet, since MAR's measures are ordinal rather than interval-based, and its sample is restricted to primarily marginalized or threatened groups, this dataset is too limited to serve as a basis for nuanced comparisons of group-level income differentials and their impact on conflict.¹

More recently, conflict researchers have used survey data to estimate economic inequality among groups. For example, relying on household surveys conducted in 39 developing countries, Østby (2008a) finds some evidence of such an effect, although it is relatively weak (see also Østby, 2008b). In a follow-up study based on geocoded conflict and survey data from sub-Saharan Africa, Østby, Nordås & Rød (2009) reach firmer conclusions, showing that both economic and social group-level differences appear to drive conflict behavior.

However, the survey-based literature is limited in a number of ways. First, data coverage is problematic, since detailed and large surveys are costly and therefore only available for certain countries. In particular, surveys are rare for states affected by civil wars.² This implies that survey-based studies of political violence could be particularly vulnerable to selection bias. Second, surveys usually measure ethnicity through self-reported categories, which raises question of standardization but also of desirability bias when dealing with marginalized and discriminated groups.

In an attempt to overcome these limitations, the aforementioned studies by Cederman, Weidmann & Gleditsch (2011) and Cederman, Gleditsch & Buhaug (2013) combine the GeoEPR data on ethnic settlement patterns (Wucherpfennig et al., 2011) with the global G-Econ dataset that provides grid-cell-level estimates of economic wealth (Nordhaus, 2006). Using geographic coincidence between the settlement areas and the G-Econ data, these studies compute the relative wealth of ethnic groups in 1990. Such calculations indicate that groups that are poorer than the country average, such as the Kosovo Albanians in the former Yugoslavia, are at

significantly higher risk of getting involved in civil wars. This effect is also present for those groups that are wealthier than the average, such as the Slovenes and Croats in the same country, although Cederman, Gleditsch & Buhaug (2013) report somewhat weaker evidence for this case.

As already mentioned, despite its generality, the spatial approach based on the G-Econ data is associated with several weaknesses. First, this measurement strategy is obviously limited to those comparisons where the relevant groups possess distinctive and mostly non-overlapping settlement areas. For this reason, it cannot estimate economic horizontal inequalities between, for example, the Tutsi and Hutu in Rwanda and Burundi. Moreover, the spatial method is unable to separate ethnic populations in larger cities, since GeoEPR specifically offers no information on purely urban settlements (Wucherpfennig et al., 2011). Furthermore, the G-Econ income maps offer relatively crude resolution, which also limits the precision of group-level measurements. Finally, the accuracy of the G-Econ data hinges on the quality of the underlying sources used to construct the dataset. In principle, this means that the data quality can be expected to be particularly poor for countries with unreliable official statistics (Chen & Nordhaus, 2011).

Against this backdrop, night lights emissions from satellite data appear as an obvious alternative that is entirely independent of possible governmental bias or quality limitations of official statistical sources. It has been shown that night lights emissions correlate strongly with economic activities (see e.g. Elvidge et al., 2009; Ghosh et al., 2010; Min et al., 2013). Relying on high resolution satellite data provided by the US Air Force Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS), social scientists have been able to improve measures of economic growth (Henderson, Storeygard & Weil, 2011) and studied the political effects of rural electrification (Min & Golden, 2014). Alesina, Michalopoulos & Papaioannou (forthcoming) are among the first to use luminosity data to compute an indicator of 'ethnic inequality'.³ Equipped with these data, they conclude that economic inequality among ethnic groups has a stronger dampening effect on development than does ethnic diversity.

¹ Obviously, this does not prevent MAR from being useful with respect to other research questions (Hug, 2013).

² See Table A-III in our online appendix.

³ Rather than relying on the GeoEPR dataset that codes overlapping settlement areas, this study uses data from the now outdated Atlas Narodov Mira, provided through shapefiles from the GREG dataset (Weidmann, Rød & Cederman, 2010).

While not specifically focusing on horizontal inequality as a cause of civil war, some conflict researchers have employed remote sensing techniques of this type to estimate the economic consequences of conflict. For example, Agnew et al. (2008) fail to find any signs of economic growth after the US ‘surge’ in Iraq in 2007. Witmer & O’Loughlin (2011) use detailed analysis of satellite imagery to study the consequences of warfare in the Caucasus. Even more recently, Shortland, Christopoulou & Makatsoris (2013) study the economic repercussions of fighting in Somalia with the help of night lights data. Finally, Kuhn & Weidmann (forthcoming) employ satellite-based data to measure within-group inequality, showing that this type of inequality is likely to drive conflict in combination with inequality between groups. Following their example, we use satellite data to measure economic inequality between groups in order to study its effect on conflict directly.

As is the case with the other two data sources, lights emissions data should be used with caution. After all, they are mere proxies of the actual underlying economic conditions, but may also pick up alternative variables. For example, night time illumination can indicate the allocation of public goods to particular regions or groups. Hodler & Raschky (2014) use luminosity data to measure how political favoritism determines the uneven provision of public goods to particular regions. While we cannot rule out that ethnic favoritism is partly driving differences in emissions across ethnic groups, we statistically take care of this problem by controlling for political exclusion. Moreover, a fine-grained validation by Weidmann & Schutte (2015) finds that night lights approximate household-level wealth with high accuracy, which means that they should be well suited for the measurement of inequality. Therefore, despite the uncertainty surrounding the interpretation of night lights, we believe that their advantages clearly outweigh the problems, in particular for research on developing countries (see Chen & Nordhaus, 2011, for a similar conclusion).

In the following, we first combine night lights (National Geophysical Data Center, 2014) and G-Econ data (Nordhaus, 2006) conditional on the quality of each country’s statistics. Second, we add survey data to the composite spatial measure to arrive at more fine-grained estimates for groups with overlapping settlement areas.

Step I: Combining spatial measures

The spatial estimation of horizontal inequality was first introduced by Cederman, Weidmann & Gleditsch (2011). The basic approach is to combine two different

spatial datasets, namely the GeoEPR dataset on group regions (Wucherpfennig et al., 2011), and the G-Econ dataset on economic activity at the subnational level (Nordhaus, 2006). The latter is essentially a map of economic performance that consists of subnational grid cells. Using the GeoEPR settlement regions as ‘cookie cutters’, for each group the relevant G-Econ cells were cut out of the G-Econ map and their values aggregated, in order to produce a total estimate of economic performance at the group level.⁴ This procedure is repeated for spatial estimates of population that allow us to compute GDP per capita estimates for each group. The final step is to compute horizontal inequality as the ratio between the per capita income of the group and that of the country as a whole.

The night lights-based measurement is similar in its approach, but obviously relies on a different data source to measure ethnic inequality. We replace the G-Econ maps with satellite imagery from the DMSP-OLS project (National Geophysical Data Center, 2014), and again apply the ‘cookie cutter’ overlay method to the luminosity layer. This step yields per capita estimates of night light emissions at the group level. Since night lights may be fluctuating due to reasons unrelated to economic activity (e.g. forest fires or cloud cover), the available night light maps filter out steady lights by aggregating multiple images over the course of a year. Still, a potential problem with the use of night lights is that the amount of light emitted may not scale linearly with income or wealth of the underlying individuals. One reason for this is the saturation of the night light measurements, where the maximum amount that can be captured by the sensor is capped at a certain level. This problem is unlikely to severely affect our analysis, as our estimates consist of aggregates across large geographic areas, whereas saturation only affects some cities mostly in developed countries. More generally, we note that the functional form of how income translates into night light emissions is still unknown. However, as reported in Chen & Nordhaus (2011: Figure 2), at the level of geographic cells, which tend to be smaller than our group polygons, the relationship is roughly linear. This is why we also use raw values of night light emissions as a linear proxy for wealth at the group level. Figure 1 shows a snapshot of the 1992 night light emission in Myanmar, overlaid with the group settlement regions of the Bamar group.⁵

⁴ For further details refer to our online appendix.

⁵ To limit measurement error, we drop observations that are more than ten times poorer or wealthier than the country average.

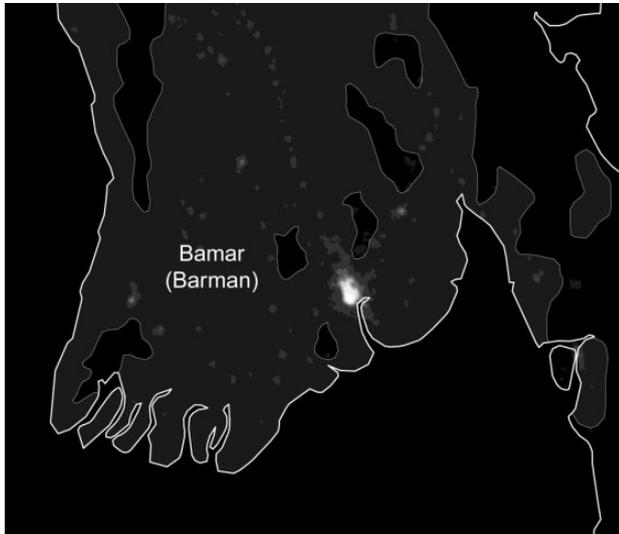
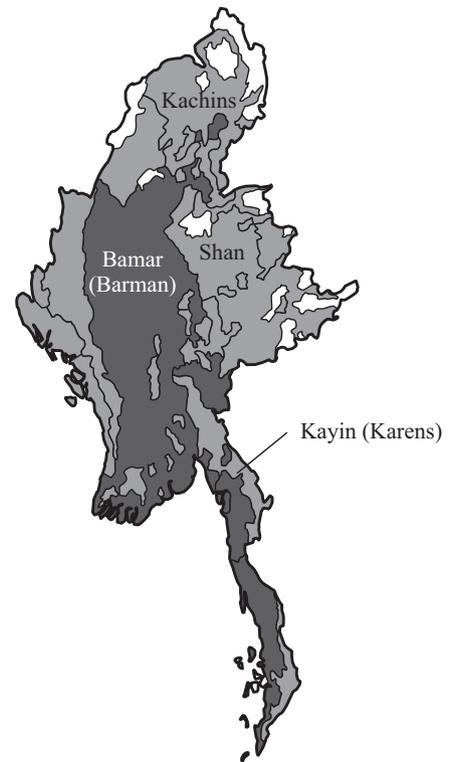


Figure 1. Example of the raw night lights data for Myanmar. The example shows nighttime illumination from the DMSP-OLS project with bright color representing more emissions, overlaid with the settlement region of the Bamar group (grey). The bright spot is Rangoon.

As discussed above, both spatial measures have strengths and weaknesses. The night lights data have global scope and can be measured with high accuracy anywhere; the obvious downside is the distance between the measured quantity (night lights) and the quantity we would like to approximate (group-level income). In fact, previous research has shown that the approximation of economic activity through night lights in general works better in poorer countries (Henderson, Storeygard & Weil, 2011). These, however, are the very same countries that the G-Econ dataset does not capture well. Chen & Nordhaus (2011) present a quality indicator for their dataset ranging from A to E (using a descending scale). Poor countries rank almost consistently in the lower categories of this classification. This is due to the fact that G-Econ draws on official statistics at the regional level; if these statistics are poor or non-existent, the ability of the G-Econ dataset to accurately reflect regional variation will be limited. This will lead G-Econ to underestimate regional variation, and scores will be more uniform across a country. Taken together, these findings suggest that G-Econ and night lights complement each other, and should be used in combination (see also Table A-II in the online appendix).

To illustrate the measurement problem in countries with low G-Econ quality scores, consider the example of Myanmar in Figure 2, which falls into the lowest quality category (E) in Chen and Nordhaus's ranking. The upper panel shows the G-Econ-based estimates and reveals very little variation: all groups cluster around the

Myanmar (G-Econ-based estimates)



Myanmar (Nightlight-based estimates)

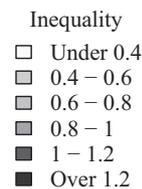
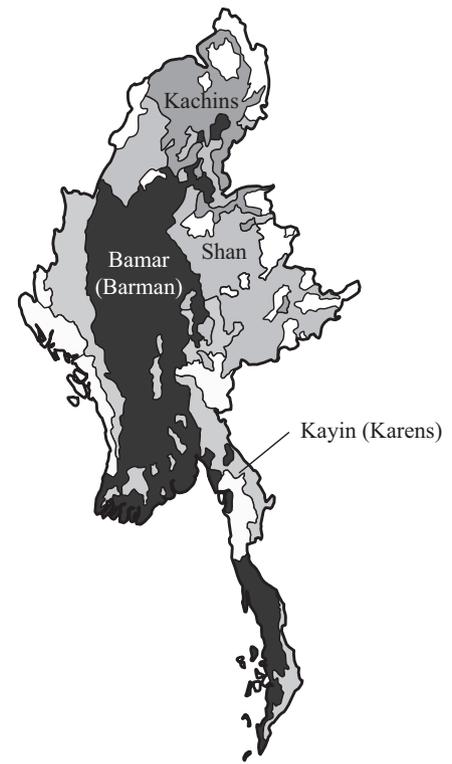


Figure 2. Comparing G-Econ and night lights-based estimates of horizontal inequality for Myanmar

Table I. Standard deviations of the G-Econ- and the night lights-based inequality estimates across the Chen & Nordhaus (2011) quality categories

Category	<i>SD(G-Econ)</i>	<i>SD(night lights)</i>
A	0.393	0.461
B	0.239	0.323
C	0.347	0.654
D	0.383	1.026
E	0.286	1.181

country average (around 1.0).⁶ In contrast, the night lights-based estimate (lower panel) provides a much more nuanced picture of horizontal inequalities between groups in Myanmar.

When we compare both spatial estimates in a cross-section of groups for 1992, we find that overall agreement is modest. The logged inequality scores correlate at 0.44, indicating that while both indicators pick up a similar direction in the data, there is a considerable degree of disagreement between them. This pattern becomes more obvious if we turn to the standard deviation of the two measures across the set of Chen and Nordhaus's quality categories (see Table I). The night lights measure records more unequal groups for lower quality categories and thus captures a well-known finding that countries in early stages of development tend to be more unequal (Williamson, 1965; UNRISD, 2010). By contrast, the G-Econ-based measure completely fails to account for this difference. Here, the spread of horizontal inequality across the categories is constant or even decreasing as we move to lower categories. This is likely due to the above-mentioned data quality problems in G-Econ, which cause the dataset to miss the significant variation in economic wealth at the subnational level.

How can we combine the G-Econ and night lights-based estimates into a measure that combines the strengths of its components? The simple approach is to weight them equally, without taking into consideration the types of countries in which they perform best. However, we propose a more nuanced combination, weighting the impact of both measures according to the above-mentioned Chen and Nordhaus categories. That is, we give the G-Econ measure a high impact for the high-quality categories in the A–E scale, and let the night lights-based measure do most of the work in the lower

categories. The G-Econ/luminosity weighting we propose is 90/10 for category A, 80/20 for category B, 50/50 for category C, 20/80 for category D, and 10/90 for category E. Since our inequality measure is a ratio, we log-transform it before computing the weighted sum, and re-transform it afterwards. For example, the category A weighting would be as follows:

$$Ineq_{spatial} = \exp\left(0.9 * \log(Ineq_{G-Econ}) + 0.1 * \log(Ineq_{nightlights})\right)$$

In this article, we use a straightforward group-level measure of economic horizontal inequality. Letting y_g denote per capita income of the ethnic group, and y_c average per capita income of all groups in the country, we measure inequality asymmetrically with two variables that correspond to groups that are poorer and wealthier than the country average, respectively:

$$Low = \begin{cases} y_c/y_g, & \text{if } y_g > y_c \\ 1 & \text{otherwise} \end{cases}$$

$$High = \begin{cases} y_g/y_c, & \text{if } y_g < y_c \\ 1 & \text{otherwise} \end{cases}$$

This operationalization guarantees that deviations from the country mean are always positive numbers above 1. For example, a group that is twice as wealthy as the average has $Low = 1$ and $High = 2$, and a group that is three times poorer, has $Low = 3$ and $High = 1$.

Our unit of analysis is the ethnic group-year and derives from the Ethnic Power Relations dataset (EPR-ETH, see Cederman, Gleditsch & Buhaug, 2013: Ch. 5). Our dependent variable measures group-level onset of civil war with at least 25 annual battle-deaths as coded by the UCDP/PRIO Armed Conflict Dataset (Gleditsch et al., 2002; Harbom & Wallensteen, 2010), which are mapped to respective ethnic groups (Wucherpfennig et al., 2012). The resulting dummy variable indicates whether a group experienced an outbreak of conflict with the government in a particular year. While onsets are coded as 1, all other years are set to 0. Observations with ongoing conflict are dropped. As in previous analyses, groups that already enjoy monopoly or dominance are not included in the analysis, since by definition they cannot rebel against themselves.

Since the main objective is to combine our indicators cross-sectionally, this article refrains from deriving time-varying measures of horizontal inequalities. The inert nature of such inequalities justifies this static assumption (Tilly, 1999; Stewart & Langer, 2008; see also the online appendix). All the same, our research design relies on

⁶ The scores measure the group's wealth relative to the average group wealth in the country; for example, 1.2 corresponds to a group that is 20% richer than the average.

panel data covering the post-Cold War period because even static calculations have to take into account major changes to the group lists and settlements areas as a consequence of geopolitical changes and mobilization processes. Furthermore, this also allows us to capture important changes in control variables. Yet, research on time-varying inequality data have to be left to future research, as indicated in the concluding section.

Relying on a similar variable list as Cederman, Gleditsch & Buhaug (2013) we proceed by describing the remaining independent variables at the group level:

- Political inequality is measured by a dichotomous variable *Excluded* that draws on EPR's distinction between excluded and included groups.
- A dummy variable labeled *Downgraded* indicates if the group suffered a loss of power in terms of EPR's status categories during the previous two years.
- *Relative group size* is measured in relation to the size of the state controlling group(s) through EPR's population size estimates.⁷
- Previous conflict history is measured with a count variable *Postwar* that records the number of previous rebellions the group has experienced since 1946 or the independence of the country.

We also introduce a number of variables to control for country-level properties:

- *GDP per capita* of the country, logged and lagged, see Penn World Table 7.0 (Heston, Summers & Aten, 2011).
- *Population* size of the country, logged and lagged, see Penn World Table 7.0 (Heston, Summers & Aten, 2011).
- *Ongoing conflict* at the country level is a dummy variable that indicates if there was an ongoing ethnic rebellion involving any other group in the country during the preceding year.
- The number of *Peace years* since the previous conflict is specified as a nonlinear function, based on natural cubic splines with three knots, as recommended by Beck, Katz & Tucker (1998).

Before assessing the impact of our data corrections, we start by replicating the original results based on G-Econ only, as reported in Cederman, Gleditsch & Buhaug (2013: Ch. 5). The results of our logistic regression

models are displayed in Table II. We limit our following discussion to the independent variables of interest, the low and high ratios. All other covariates exhibit effects similar to previous research, and across all the new models presented here. Model 1 in Table II shows that groups with a low or a high socio-economic status are more likely to experience conflict relative to groups closer to the national average. Although the coefficient for the high ratio is estimated to be positive, there seems to be no significant effect in contrast to previous research (Cederman, Weidmann & Gleditsch, 2011). Replacing the G-Econ-based indicator with the one based on night lights in Model 2, even the effect of the low ratio vanishes. On their own, the luminosity data provide no strong evidence for the effect of economic horizontal inequality on conflict. Yet, the picture changes dramatically if we use a naive 50/50 weighting of the G-Econ and night lights indicators. Here both effects re-emerge with positive signs and statistically significant. These estimates become more precise when we apply our quality-sensitive weighting of G-Econ and night lights, based on the A-E quality categories (Model 4). Compared to the initial G-Econ estimate, the new mix lends support to a roughly symmetric, bidirectional effect of inequality that includes both poorer and more affluent groups.

Step II: Factoring in survey data

The second step of our analysis aims to overcome two shortcomings of the spatial inequality estimates by adding survey-based data. First, it broadens coverage to purely urban groups for which no spatial estimates are available. Second, it helps us to distinguish between groups with overlapping settlement patterns. Our survey data come from a number of sources. We rely on data collected by Baldwin & Huber (2010) and Huber & Mayoral (2013). These authors link Fearon's (2003) list of ethnic group to individual income measures in multiple survey projects such as the Demographic and Health Surveys (DHS), the Luxemburg Income Study, the World Values Survey, and the Afrobarometer (see references in the online appendix). They then aggregate individual-level income to the level of ethnic groups and derive inequality data estimates at the country level.

Based on these data, the EPR groups were linked to the Fearon groups, and thereby to the ethnic categories identified in the relevant surveys.⁸ Equipped with estimates of group-level wealth income, we derived inequality

⁷ Denote groups' population as G and the population of all included groups as I , relative group size $g \in [0, 1) = G/(G + I)$ if the group is excluded and as G/I if the group is included.

⁸ Huber and Mayoral generously shared their group-level income data with us.

Table II. Explaining onset of ethnonationalist rebellion using spatial measures in Step I, 1946–2009

	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>
Low (G-Econ)	0.474*** (0.112)			
High (G-Econ)	0.633 (0.344)			
Low (night lights)		0.074 (0.043)		
High (night lights)		0.163 (0.099)		
Low (50/50 mix)			0.462*** (0.129)	
High (50/50 mix)			0.420* (0.166)	
Low (nuanced mix)				0.345** (0.107)
High (nuanced mix)				0.437*** (0.151)
Excluded	1.273*** (0.375)	1.255*** (0.353)	1.242*** (0.364)	1.225*** (0.360)
Downgraded	1.947*** (0.456)	2.025*** (0.436)	2.024*** (0.450)	2.082*** (0.447)
Relative group size	−0.028 (0.672)	−0.322 (0.682)	−0.149 (0.673)	−0.247 (0.702)
Postwar	0.225 (0.180)	0.157 (0.190)	0.197 (0.186)	0.184 (0.186)
Ongoing conflict	0.370 (0.408)	0.419 (0.420)	0.355 (0.419)	0.306 (0.414)
GDP per capita (log, $t-1$)	−0.324* (0.131)	−0.305* (0.129)	−0.268* (0.126)	−0.235 (0.126)
Population (log)	−0.004 (0.180)	0.003 (0.144)	−0.045 (0.156)	−0.003 (0.157)
Peace years	−0.008 (0.088)	−0.014 (0.089)	0.005 (0.091)	0.006 (0.092)
Constant	−2.946 (1.678)	−2.095 (1.501)	−2.761 (1.540)	−3.312* (1.650)
Observations	5,377	5,027	5,377	5,377

Country-clustered standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

estimates for each EPR group with a respective survey link. Dividing each group-level income estimate by the country average allows us to compare inequality across surveys and countries as the relative comparison within surveys removes scale differences between surveys.⁹

⁹ While Huber & Mayoral (2013) undertake so-called ‘intercept’ and ‘ratio’ corrections in order to guarantee better cross-country comparability and introduce a time trend in inequality, we do not implement these approaches. For the time being, our analysis is based on one survey per country even where there is more than one available. When multiple surveys per country are available, we select the one that maximizes the coverage of ethnic groups and is closest to 1992. See the online appendix for details on the specific surveys used.

In analogy to the first step, our analysis attempts to combine two data sources based on theoretically informed choices. In this case, we base our weighting on the amount of overlap a settlement area has with other areas. Recall that as per GeoEPR coding rules, groups can share parts of their region with other groups. For each group polygon, we compute the amount of overlap with other group polygons. For example, if this ratio is 0.2, 20% of the area of a group’s polygon is shared with one or more groups. Our geocoded indicator is particularly error-prone if overlap is high, since in these cases it cannot distinguish the groups that the population in the overlapping area belongs to. For that reason, we use the proportion of overlap as the weight of the

Table III. Explaining onset of ethnonationalist rebellion including surveys in Step II, 1946–2009

	<i>Model 4</i>	<i>Model 5</i>	<i>Model 6</i>	<i>Model 7</i>
Low (spatial only, nuanced mix)	0.345** (0.107)			
High (spatial only, nuanced mix)	0.437** (0.151)			
Low (surveys)		1.136*** (0.280)		
High (surveys)		-0.428 (0.719)		
Low (average spatial/surveys)			0.464*** (0.119)	
High (average spatial/surveys)			0.862*** (0.237)	
Low (overlap-weighted spatial/surveys)				0.437*** (0.116)
High (overlap-weighted spatial/surveys)				0.503** (0.154)
Excluded	1.225*** (0.360)	0.430 (0.435)	0.815* (0.331)	0.838** (0.325)
Downgraded	2.082*** (0.447)	1.114 (0.851)	1.997*** (0.433)	1.956*** (0.441)
Relative group size	-0.247 (0.702)	-0.211 (1.198)	0.098 (0.658)	0.158 (0.614)
Postwar	0.184 (0.186)	0.507*** (0.151)	0.324 (0.174)	0.344* (0.168)
Ongoing conflict	0.306 (0.414)	-0.229 (0.551)	0.254 (0.381)	0.314 (0.377)
GDP per capita (log, $t-1$)	-0.235 (0.126)	-0.419** (0.156)	-0.270* (0.128)	-0.268* (0.122)
Population (log)	-0.003 (0.157)	0.279 (0.179)	0.072 (0.127)	0.055 (0.135)
Peace years	0.006 (0.092)	-0.095 (0.080)	0.008 (0.076)	0.015 (0.078)
Constant	-3.312* (1.650)	-4.664* (2.057)	-4.667*** (1.301)	-4.158** (1.388)
Observations	5,377	3,804	6,699	6,699

Country-clustered standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

survey-based indicator compared to the GIS-based one: high-overlap cases therefore draw primarily on the survey indicator, whereas low overlap cases use the GIS-based indicator.

We are now ready to combine the spatial data from the first step of the analysis with the survey-based information in our models of civil-war onset. To facilitate comparison between the spatial mix and the survey data, Table III again displays Model 4. Model 5 shows that the survey data on their own yield a very strong and significant effect for the low variable, but the effect of richer groups is even negative and not statistically significant. Relying only on survey data yields very different results from the spatial analysis which may be a

consequence of the smaller sample for which survey data are available. We therefore proceed with combined measures that take both spatial and survey-based data into account.

Model 6 features a naive mix that puts equal weight on both spatial and survey data in cases where both are present. In addition, we rely on either geographic or survey data when the other data type is missing. Compared to the spatial mix in Model 4, the conflict-inducing effects for both poor and wealthy groups increase considerably. In particular, this estimate nearly doubles for wealthier groups and the coefficient for the poorer ones increases as well. Furthermore, the precision of both estimates increases compared to those in Model 4. However,

we note that the new estimates are considerably more sensitive to extreme values.¹⁰

As the final step in our analysis, we introduce the overlap-sensitive weighting described above. While the 'low' estimate decreases somewhat without losing its high level of significance, the 'high' effect shrinks to a value that is much closer to the corresponding spatial estimate in Model 4, without affecting the level of statistical significance. More importantly, the sophisticated mix improves the robustness of the model.¹¹

In order to further illustrate the effect, we turn to an effect plot displayed in Figure 3. Showing the predicted conflict probability as a function of the low ratio (upper panel), the graph reveals that the probability grows as groups get poorer. Due to fewer observations for extreme values, the error bands widen with increasing inequality.

We find a similar effect on conflict for wealthier groups (see Figure 3, lower panel). While the effect of richer groups on conflict is similar to that of their poorer counterparts, it is far less precise due to the low number of conflicts for extreme values of inequality. In addition to showing that the overlap-based mix of Model 7 reduces the sensitivity to extreme values, results presented in the online appendix reveal that economic horizontal inequality tends to drive territorial, as opposed to governmental, conflict. Our results also remain robust to controls for world regions or alternative spatial weighting schemes.

Compared to previous research based solely on spatial measures, the current analysis constitutes a more solid basis for inference because the survey instruments are able to detect income differences in cases where there is no spatial separation between the groups' settlement areas.¹² Yet, before drawing any further conclusions about a possible causal effect, endogeneity and reverse causation also need to be considered. Above, we have already noted that structural asymmetries such as economic horizontal inequalities tend to persist and therefore reverse causation is much less likely than in the case of political inequality (Tilly, 1999; Stewart & Langer,

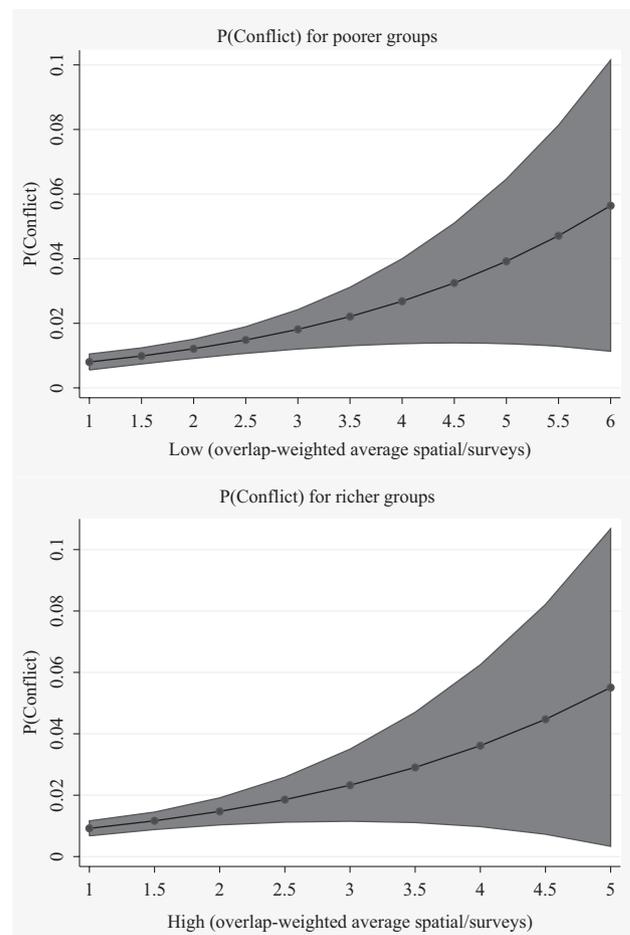


Figure 3. Predicted probability of conflict depending on values of the low (top pane) and high (bottom pane) ratios, based on Model 7 in Table III

The shaded area indicates the 95% confidence interval around the estimated effect.

2008). Although at least partly dependent on the government's provision of public goods, ethnic groups' relative economic performance typically does not exhibit much change over the decades. For example, the Kosovo Albanians were at the bottom of the Yugoslav league table in the 1950s and remained so in the 1970s (Lang, 1975; see also the online appendix).

Another potential issue is selection bias, particularly due to the limited availability of surveys in conflict cases (see Table A-III in the online appendix). We argue that selection bias is less concerning in our analysis for two reasons. First, when survey data are missing, our geographic sources provide coverage for ethnic groups with clearly defined settlement patterns. Previous research shows that these groups are more likely to rebel than dispersed groups (Toft, 2002). Thus, we include most groups that are at risk of experiencing conflict, but we

¹⁰ In particular, removing the Bakongos of Angola from the model, who are coded as almost three times wealthier than the country average, collapses the high estimate.

¹¹ For example, the sensitivity to the Bakongos' high ratio (see note 10) is considerably reduced in this version. See the online appendix for details.

¹² Additional analysis tells us that the conflict-causing impact of economic horizontal inequality is mostly limited to excluded groups, which suggests that economic grievances hinge on political horizontal inequalities, as argued by Cederman, Gleditsch & Buhaug (2013: Ch. 5) (see also the online appendix).

do not oversample groups with more extreme values on our independent variable. Second, the combination of survey data and geographic data improves our estimates specifically for groups with overlapping settlements. More specifically, it will tend to move data points away from the country average because our geography sample uses the same G-Econ and night lights data points for all groups with overlapping settlement areas, which makes groups more similar to each other, and thus to the country average. Not using survey data thus biases our results against our initial hypothesis that poorer and richer groups are more likely to rebel.

Case illustrations

So far, we have found strong statistical evidence that the combined inequality indicators lead to solid estimates of the effect of inequality on the risk of civil war. This section assesses the face validity of these general findings. We first discuss the cases that drive the main results. As regards the poorer conflict groups, there is considerable evidence that radical underdevelopment contributes to a number of groups' grievances against their respective governments. More than four times poorer than the Russian average, the Chechens are the least developed conflict group in Russia. While political disadvantage also plays a role, regional studies confirm that this economic inferiority has contributed to fueling grievances (see e.g. Derluguian, 2005). Other examples of economically marginalized groups that have rebelled include the East Timorese in Indonesia, the Naga in India, the Fur in the Sudan, the Muslim north of the Ivory Coast, and various peripheral groups in Myanmar, such as the Muslim Arakanese, the Mons, the Kayins, and the Shan. The Burmese examples are especially interesting given the difficulties of capturing income differentials based on previous datasets (see Figure 2). While persistent conflict has contributed to underdevelopment of war-torn peripheral regions of Myanmar, ethnic minority organizations have repeatedly referred to the government's economic discrimination as a major grievance motivating their secessionist bids (Walton, 2008: 897).

Whereas there is plenty of robust evidence in favor of poor groups responding violently to their underdevelopment, it is harder to find systematic support for well endowed groups rebelling. Indeed, it would be a mistake to extrapolate from the Slovenian and Croat cases to the rest of the sample. According to this interpretation of H2, wealthy minorities would seek to break away from their respective states rather than sharing their wealth with what they perceive to be ineffective and undeserving compatriot

groups. Yet, in our sample, the highest recordings of relative group wealth exceeding the values of the Yugoslav cases can be found in oil-rich regions inhabited by the Ijaw in Nigeria and the Bakongo in the Cabinda region of Angola. It can be assumed that to a large degree, these high estimates stem from gas flaring rather than from economic activities of the ethnic group in question (Chen & Nordhaus, 2011; Henderson, Storeygard & Weil, 2011). In these cases, then, potential, rather than real, wealth appears to motivate secessionist rebellions since the groups in question are anything but wealthy (Sorens, 2011; Ross, 2012). Thus, we have to conclude that while grievances are clearly at play in the link between oil and civil war, in several important cases, the specific mechanism deviates from what we originally postulated (see e.g. Asal et al., forthcoming).

Having discussed to what extent specific cases appear to confirm our hypotheses, we now rely on external data in two important cases. First, we look more closely at racial inequality in South Africa and compare it to census data on income which are likely to offer the most accurate estimate of income inequality available. Second, we discuss examples of the multiple ethnic divisions in India and how our combination of data is able to deal with such a complex case.

First, we consider South Africa, a state with a history of racial oppression and violence under the extremely exclusive Apartheid regime, as well as a lasting legacy of stark economic intergroup inequalities (Leibbrandt, Finn & Woolard, 2012: 23). Figure 4 plots inequality values based on census income data (Statistics South Africa, 2012) against the estimates based on the G-Econ dataset, the night lights data, and surveys for four different ethnic groups.¹³ If there were a perfect match between the census and any of our data sources, the points would line up on the diagonal dotted line. Three observations stand out. First, the geographic and survey data underestimate the degree of horizontal economic inequality in South Africa to a considerable degree. Second, while getting the rank order of inequality approximately right, our indicators have difficulties distinguishing between groups that are close to the country mean. Third, a combination of the three main data sources gets us closer to the more accurate

¹³ After 1990 EPR-ETH distinguishes between the various language groups that compose the umbrella category 'Blacks' but the census data do not. Therefore, we aggregate the language groups' estimates to the larger racial category. Note that both the Afrobarometer and the South African Census provide multiple data points. We average the estimates for both sources.

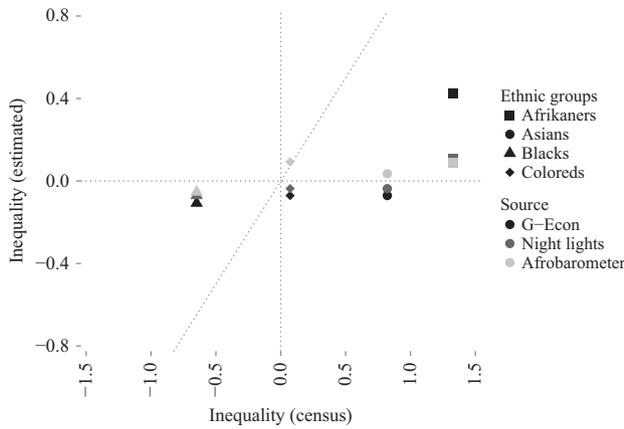


Figure 4. Horizontal inequality measures for racial ethnic groups in South Africa

Sources: census data (Statistics South Africa, 2012) and combined geographic and survey sources.

census data points. A closer look at individual groups shows why this is the case.

Beginning with Asians (circle) and Coloreds (diamond), GeoEPR codes these groups as ‘dispersed’ throughout South Africa. Our geographic estimates are therefore very close to the country average and cannot distinguish these two groups. This means our combined measures rely entirely on the information provided by surveys. In the case of the Coloreds, we almost exactly match the economic inequality estimate from the census. Our combined estimate of the Asians group gets closer to the more accurate census data and at least indicates that the group is slightly richer than the country average. Turning to the lower end of the spectrum, our three data sources show very close agreement for the inequality of the Blacks (triangle) but overestimate their relative wealth. Nevertheless combining all three data sources moves our estimate closer towards the census data relative to relying only on the night lights or the survey data. Finally, the Nordhaus data indicate a clearly higher estimate for white Afrikaners (square) and again moves the other two data sources much closer to the high-quality census estimate. If the case of South Africa is representative for other states in our sample, we can conclude that our strategy of combining three data sources improves our estimates. Nevertheless we underestimate the degree of horizontal inequality in South Africa, and thus probably obtain rather conservative estimates of the effect of horizontal inequality on civil war risk. Finally, we note that the South African census data for 2001 and 2011 reveal that between-group inequality hardly changed at all (also refer to our online appendix).

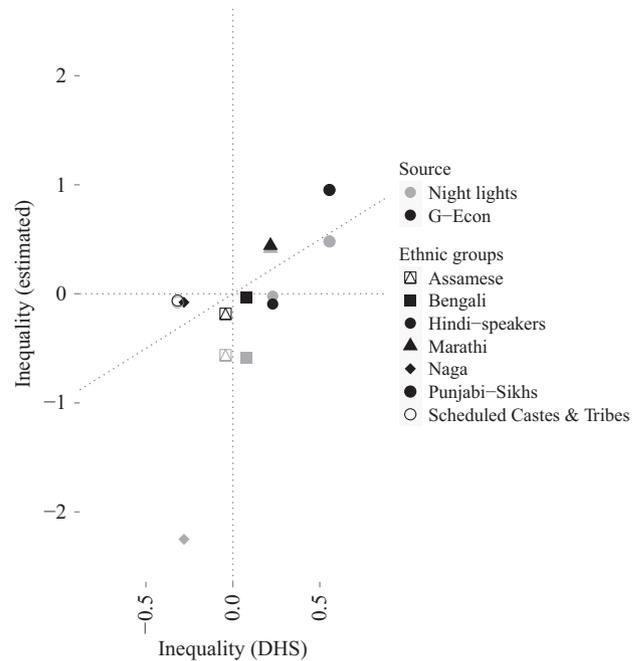


Figure 5. Horizontal inequality measures for selected Indian ethnic groups

Sources: DHS surveys and combined G-Econ/night lights sources.

We now turn to India, which is another important state in terms of both ethnic diversity and civil war occurrence as the country accounts for six civil war onsets in our sample. Unfortunately, census data on between-group inequality are not available. Nevertheless we believe that a comparison of our geographic data to DHS surveys not included in Huber and Mayoral’s database illustrates the potential of combining multiple data sources in a country that is characterized by a complex ethnic landscape including linguistic, religious, caste, and racial distinctions. Due to this complexity, Figure 5 plots the average inequality estimate from three large-scale DHS surveys against geographic data only for a subset of ethnic groups that exemplify the benefits and drawbacks of each individual data source. Note that in contrast to the South African case, the diagonally dashed line in Figure 5 indicates agreement between geographic and survey data on household items rather than proximity to very accurate census income estimates.

We start by comparing two Hindu groups whose settlement patterns overlap to a large degree but that should differ strongly in terms of relative wealth: the economically backward Scheduled Castes and Tribes (SC/ST) and non-SC/ST Hindi-speakers. Even today, the Hindu caste structure has a substantial impact on economic status, in particular in rural areas where about two-thirds of

India's population resides. Our geographic estimates do not allow us to draw a clear distinction between SCs/STs (empty circle) and Hindi-speakers (small full circle) although the former should be poorer compared to the latter (see Guha, 2007: 377). Adding survey data on these two groups with a very high overlap distinguishes the groups far more clearly. A similar improvement of the data pertains to the Assamese (square-triangle) and the Bengalis (filled square) whose settlement territories overlap to a large degree and where survey data more clearly separate the inequality estimates. Indeed, both Weiner (1978: 131) and Brass (1994: 213) indicate that the reason for the violent competition between Assamese and Bengalis is the economic backwardness of the Assamese. Our data thus underestimate the risk of civil war in this case.

Relying on survey data also has drawbacks. For example, the DHS surveys imply that Hindi-speakers and Marathis (triangle) occupy virtually the same economic rank, but the latter live in the economically advanced state of Maharashtra that includes Mumbai. As the settlement pattern of the Marathis overlaps to a far smaller degree with that of other groups, our geographic indicators would carry most of the weight and pull the Marathis above Hindi-speakers – a more reasonable assessment. To the right of the Marathis in Figure 5, the Punjabi-Sikhs (large filled circle) occupy the top-spot of the richest group – a position that is justified both through the agreement of our three indicators and qualitative case studies (Horowitz, 2000: 234). Finally, we turn to the other extreme end of the sample. The Naga (diamond), whose remote location and tribal structure make it one of the poorest ethnic groups in India, would probably be too close to the Indian average if one were to rely on Nordhaus or survey data alone. In contrast, the night lights data by themselves would probably overestimate the degree of relative inequality. The 50/50 Nordhaus–night lights mix for a category C country such as India is likely to be closer to reality (Baruah, 2005: 112). In sum, we believe that our choice of combining estimators makes sense in these cases and has a high degree of face validity.

Conclusion

Although our experiment in data merging leaves room for improvement, we feel confident enough to conclude that combining data sources on inequality can be very helpful. Using existing results published in Cederman, Weidmann & Gleditsch (2011) and Cederman, Gleditsch & Buhaug (2013) as the point of departure, we studied

how the results derived in those publications can be rendered more trustworthy and robust. The first analytical step draws on night lights emissions in order to compensate for quality limitations in Nordhaus's G-Econ data, especially in cases where official statistics are known to be poor. The second step of the analysis complements the spatial derivation with survey-based estimates in such a way that these were weighted more heavily in cases where the spatial overlap between ethnic groups is particularly high.

This article increases the confidence that the original results of Cederman, Weidmann & Gleditsch (2011) are correct. The improved measures used here indicate that both poorer and richer groups are overrepresented in the conflict statistics. Yet, as shown by the case discussion in the previous section, when it comes to causal inference, caution is still very much called for. Although clearly reflecting economic activities, the spatial approaches, and especially the night lights data, do not always offer reliable estimates of group income. Because of these complications, future efforts to measure asymmetries in ethnic groups' wealth and their link to conflict will need to take into account the role of oil and gas production, especially where these generate luminosity through flaring. Inequality and grievances are clearly involved in these cases, but the conflict effect derives from potential, rather than actual, wealth.

Beyond this specific challenge, there is plenty of room for future research on measuring group-level inequality. Such extensions could include more sophisticated ways to merge data sources, for example by exploiting information on survey size. Additional analysis will also be needed to explore the quality of the data. Most importantly, we have proceeded mostly based on theoretical reasoning. Although we offered some evidence validating our results with data that approximate an external 'gold standard' in the South African case, future research could and should go further than that. Indeed, additional in-depth analysis based on case studies and country-specific datasets will allow us to draw firmer conclusions about the extent to which the combined analysis actually represents an improvement in terms of data precision. Such detailed analysis would also be necessary to put the identification strategy on a more solid footing. Although economic inequality among ethnic groups is relatively stable over time, reverse causation cannot be excluded, and identifying such effects would clearly require a more sophisticated research design than that used in this article.

For the time being, this study has relied primarily on indirect ways of assessing the quality of our new measures. Even so, we find it reassuring that straightforward,

theoretically justified mixes of data generate stronger evidence for the link between economic horizontal inequality and internal conflict than does analysis based on any one component. Indeed, this is exactly what one would expect if our initial theoretical claim about the effect of inequality on civil war among economically marginalized groups were correct.

Replication data

The online appendix and the replication archive are available at <http://www.prio.org/jpr/datasets>. All analyses were conducted using Stata 13.

Acknowledgements

We thank Manus Midlarsky, Henrik Urdal, panel participants at the 2014 Annual Meeting of the American Political Science Association, and two anonymous reviewers for helpful comments, as well as John Huber and Laura Mayoral for generously sharing their data with us. We gratefully acknowledge support from the Swiss National Science Foundation (Grant-No. 105511-143213) for the work of Nils-Christian Bormann, and from the Alexander von Humboldt Foundation (Sofja Kovalevskaja Award) for the work of Nils B Weidmann.

References

- Acemoglu, Daron & James A Robinson (2005) *Economic Origins of Dictatorship and Democracy*. Cambridge: Cambridge University Press.
- Agnew, John; Thomas W Gillespie, Jorge Gonzales & Brian Min (2008) Baghdad nights: Evaluating the US military surge using night lights signatures. *Environment and Planning A* 40(10): 2285–2295.
- Alesina, Alberto & Dani Rodrik (1994) Distributive politics and economic growth. *Quarterly Journal of Economics* 102(2): 465–490.
- Alesina, Alberto F; Stelios Michalopoulos & Elias Papaioannou (forthcoming) Ethnic inequality. *Journal of Political Economy*.
- Asal, Victor; Michael Findley, James A Piazza & James Igoe Walsh (forthcoming) Political exclusion, oil, and ethnic armed conflict. *Journal of Conflict Resolution*. DOI: 10.1177/0022002714567948.
- Baldwin, Kate & John D Huber (2010) Economic versus cultural differences: Forms of ethnic diversity and public goods provision. *American Political Science Review* 104(4): 644–662.
- Barrows, Walter L (1976) Ethnic diversity and political stability in Black Africa. *Comparative Political Studies* 9(2): 139–170.
- Baruah, Sanjib (2005) *Durable Disorder: Understanding the Politics of Northeast India*. Oxford: Oxford University Press.
- Beck, Nathaniel; Jonathan N Katz & Richard Tucker (1998) Taking time seriously: Time-series cross-section analysis with a binary dependent variable. *American Journal of Political Science* 42(4): 1260–1288.
- Brass, Paul R (1994) *The Politics of India Since Independence*. Volume 2. New York: Cambridge University Press.
- Cederman, Lars-Erik; Kristian Skrede Gleditsch & Halvard Buhaug (2013) *Inequality, Grievances and Civil War*. New York: Cambridge University Press.
- Cederman, Lars-Erik; Nils B Weidmann & Kristian Skrede Gleditsch (2011) Horizontal inequalities and ethnonationalist civil war: A global comparison. *American Political Science Review* 105(3): 478–495.
- Chen, Xi & William D Nordhaus (2011) The value of luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences* 108(21): 8589–8594.
- Collier, Paul & Anke Hoeffler (2004) Greed and grievance in civil wars. *Oxford Economic Papers* 56(4): 563–595.
- Cramer, Christopher (2003) Does inequality cause conflict? *Journal of International Development* 15(4): 397–412.
- Derluigan, Georgi M (2005) *Bourdieu's Secret Admirer in the Caucasus: A World-System Biography*. Chicago, IL: University of Chicago Press.
- Elvidge, Christopher D; Paul C Sutton, Tilottama Ghosh, Benjamin T Tuttle, Kimberly E Baugh, Budhendra Bhaduri & Edward Bright (2009) A global poverty map derived from satellite data. *Computers and Geosciences* 35(8): 1652–1660.
- Esteban, Joan; Laura Mayoral & Debraj Ray (2012) Ethnicity and conflict: An empirical study. *American Economic Review* 102(4): 1310–1342.
- Fearon, James D (2003) Ethnic and cultural diversity by country. *Journal of Economic Growth* 8(2): 195–222.
- Fearon, James D & David D Laitin (2003) Ethnicity, insurgency, and civil war. *American Political Science Review* 97(1): 75–90.
- Ghosh, Tilottama; Rebecca L Powell, Christopher D Elvidge, Kimberly E Baugh, Paul C Sutton & Sharolyn Anderson (2010) Shedding light on the global distribution of economic activity. *Open Geography Journal* 3: 148–161.
- Gleditsch, Nils Petter; Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg & Håvard Strand (2002) Armed conflict 1946–2001: A new dataset. *Journal of Peace Research* 39(5): 615–637.
- Gourevitch, Peter A (1979) The reemergence of 'peripheral nationalisms': Some comparative speculations on the spatial distribution of political leadership and economic growth. *Comparative Studies in Society and History* 21(3): 303–322.

- Guha, Ramachandra (2007) *India after Gandhi: The History of the World's Largest Democracy*. New York: HarperCollins.
- Gurr, Ted Robert (1993) *Minorities at Risk: A Global View of Ethnopolitical Conflicts*. Washington, DC: United States Institute of Peace Press.
- Harbom, Lotta & Peter Wallensteen (2010) Armed conflicts, 1946–2009. *Journal of Peace Research* 47(4): 501–509.
- Henderson, Vernon; Adam Storeygard & David N Weil (2011) A bright idea for measuring economic growth. *American Economic Review* 101(3): 194–199.
- Heston, Alan; Robert Summers & Bettina Aten (2011) Penn World Table version 7.0 (<http://knoema.com/PWT2011Jun/penn-world-table-7-0>).
- Hodler, Roland & Paul A Raschky (2014) Regional favoritism. *Quarterly Journal of Economics* 129(2): 995–1033.
- Horowitz, Donald L (2000) *Ethnic Groups in Conflict*, 2nd edition. Berkeley, CA: University of California Press.
- Huber, John D & Laura Mayoral (2013) Inequality, ethnicity and civil conflict. Unpublished manuscript. Department of Political Science, Columbia University (http://polisci.columbia.edu/files/polisci/u86/huber%20mayoral_0.pdf).
- Hug, Simon (2013) Use and misuse of MAR. *Annual Review of Political Science* 16: 191–208.
- Kuhn, Patrick M & Nils B Weidmann (2015) Unequal we fight: Between- and within-group inequality and ethnic civil war. *Political Science Research and Methods* 3(3): 543–568.
- Kyriacou, Andreas P (2013) Ethnic group inequalities and governance: Evidence from developing countries. *Kyklos* 66(1): 77–101.
- Lang, Nicholas R (1975) The dialectics of decentralization: Economic reform and regional inequality in Yugoslavia. *World Politics* 27(3): 309–335.
- Leibbrandt, Murray; Arden Finn & Ingrid Woolard (2012) Describing and decomposing post-Apartheid income inequality in South Africa. *Development Southern Africa* 29(1): 19–34.
- Lichbach, Mark Irving (1989) An evaluation of 'Does economic inequality breed political conflict?'. *World Politics* 41(4): 431–470.
- Min, Brian & Miriam Golden (2014) Electoral cycles in electricity losses in India. *Energy Policy* 65: 619–625.
- Min, Brian; Kwawu M Gaba, Ousmane F Sarr & Alsassane Agalassou (2013) Detection of rural electrification in Africa using DMSP-OLS night lights imagery. *International Journal of Remote Sensing* 34(22): 8118–8141.
- Mitra, Anirban & Debraj Ray (2014) Implications of an economic theory of conflict: Hindu–Muslim violence in India. *Journal of Political Economy* 122(4): 719–765.
- National Geophysical Data Center (2014) DMSP-OLS nighttime lights time series, version 4 (<http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html>).
- Nordhaus, William D (2006) Geography and macroeconomics: New data and new findings. *Proceedings of the National Academy of Sciences* 103(10): 3510–3517.
- Østby, Gudrun (2008a) Inequalities, the political environment and civil conflict: Evidence from 55 developing countries. In: Frances Stewart (ed.) *Horizontal Inequalities and Conflict: Understanding Group Violence in Multiethnic Societies*. Houndmills: Palgrave Macmillan, 136–161.
- Østby, Gudrun (2008b) Polarization, horizontal inequalities and violent civil conflict. *Journal of Peace Research* 45(2): 143–162.
- Østby, Gudrun; Ragnhild Nordås & Jan Ketil Rød (2009) Regional inequalities and civil conflict in sub-Saharan Africa. *International Studies Quarterly* 53: 301–324.
- Ross, Michael (2012) *The Oil Curse: How Petroleum Wealth Shapes the Development of Nations*. Princeton, NJ: Princeton University Press.
- Sambanis, Nicholas (2004) Using case studies to expand economic models of civil war. *Perspectives on Politics* 2(2): 259–279.
- Shortland, Anja; Christopoulou Katerina & Makatsoris Charalampos (2013) War and famine, peace and light? The economic dynamics of conflict in Somalia 1993–2009. *Journal of Peace Research* 50(5): 545–561.
- Sorens, Jason (2011) Mineral production, territory, and ethnic rebellion: The role of rebel constituencies. *Journal of Peace Research* 48(5): 571–585.
- Statistics South Africa (2012) Census 2011 (<http://www.statssa.gov.za/publications/P03014/P030142011.pdf>).
- Stewart, Frances, ed. (2008) *Horizontal Inequalities and Conflict: Understanding Group Violence in Multiethnic Societies*. Houndmills: Palgrave Macmillan.
- Stewart, Frances & Arnim Langer (2008) *Horizontal Inequalities: Explaining Persistence and Change*. Houndmills: Palgrave Macmillan.
- Tilly, Charles (1999) *Durable Inequality*. Berkeley and Los Angeles, CA: University of California Press.
- Toft, Monica Duffy (2002) Indivisible territory, geographic concentration, and ethnic war. *Security Studies* 12(2): 82–119.
- UNRISD (2010) Combating poverty and inequality: Structural change, social policy and politics. United Nations Research Institute for Social Development ([http://www.unrisd.org/80256B3C005BCCF9/\(httpAuxPages\)/92B1D5057F43149CC125779600434441/](http://www.unrisd.org/80256B3C005BCCF9/(httpAuxPages)/92B1D5057F43149CC125779600434441/)).
- Walton, Matthew J (2008) Ethnicity, conflict, and history in Burma: The myths of Panglong. *Asian Survey* 48(6): 889–910.
- Weidmann, Nils B & Sebastian Schutte (2015). Using night lights for the subnational estimation of wealth. Working paper. University of Konstanz.
- Weidmann, Nils B; Jan Ketil Rød & Lars-Erik Cederman (2010) Representing ethnic groups in space: A new dataset. *Journal of Peace Research* 47(4): 87–119.
- Weiner, Myron (1978) *Sons of the Soil: Migration and Ethnic Conflict in India*. Princeton, NJ: Princeton University Press.
- Williamson, Jeffrey G (1965) Regional inequality and the process of national development: A description of the patterns. *Economic Development and Cultural Change* 13(4): 1–84.

- Witmer, Frank & John O'Loughlin (2011) Detecting the effects of wars in the Caucasus regions of Russia and Georgia using radiometrically normalized DMSP-OLS nighttime lights imagery. *GIScience and Remote Sensing* 48(4): 478–500.
- Wood, Elisabeth J (2003) *Insurgent Collective Action and Civil War in El Salvador*. Cambridge: Cambridge University Press.
- Wucherpfennig, Julian; Nils W Metternich, Lars-Erik Cederman & Kristian S Gleditsch (2012) Ethnicity, the state, and the duration of civil war. *World Politics* 64(1): 79–115.
- Wucherpfennig, Julian; Nils B Weidmann, Luc Girardin, Lars-Erik Cederman & Andreas Wimmer (2011) Politically relevant ethnic groups across space and time: Introducing the GeoEPR dataset. *Conflict Management and Peace Science* 28(5): 423–437.
- LARS-ERIK CEDERMAN, b. 1963, PhD (University of Michigan, 1994); Professor of International Conflict Research, ETH Zürich (2003–).
- NILS B WEIDMANN, b. 1976, PhD (ETH Zurich, 2009); Professor of Political Science, University of Konstanz (2012–) and head of the Communication, Networks and Contention research group.
- NILS-CHRISTIAN BORMANN, b. 1985, PhD (ETH Zürich, 2014); Post-Doc, ETH Zürich (2014–).