Sensitivity Analysis of Empirical Results on Civil War Onset

HÅVARD HEGRE
Centre for the Study of Civil War, PRIO
NICHOLAS SAMBANIS
Political Science Department
Yale University

In the literature on civil war onset, several empirical results are not robust or replicable across studies. Studies use different definitions of civil war and analyze different time periods, so readers cannot easily determine if differences in empirical results are due to those factors or if most empirical results are just not robust. The authors apply a methodology for organized specification tests to check the robustness of empirical results. They isolate causes of variation in empirical results by using the same definition of civil war and analyzing the same time period while systematically exploring the sensitivity of eighty-eight variables used to explain civil war in the literature. Several relationships with the onset of civil wars prove robust: large population and low income levels, low rates of economic growth, recent political instability and inconsistent democratic institutions, small military establishments and rough terrain, and war-prone and undemocratic neighbors. Variables representing ethnic difference in the population are robust only in relation to lower level armed conflict.

Keywords: *civil war; internal armed conflict; sensitivity analysis; ethnic fractionalization*

1. INTRODUCTION

"A fragile inference is not worth taking seriously" (Leamer 1985, 308). Most political scientists would agree with that statement. However, most empirical studies in political science typically do not subject their inferences to rigorous robustness or specification tests. Replication studies, to some extent, address this issue: in the empirical study of civil war onset, there is now consensus that the risk of war

AUTHORS' NOTE: We thank Scott Bennett, Halvard Buhaug, Scott Gates, Nils Petter Gleditsch, and Håvard Strand for helpful suggestions and Gudrun Østby and Christin Ormhaug for excellent research assistance. The research was supported by the Polarization and Conflict Project CIT-2-CT-2004-506084, funded by the European Commission-DG Research Sixth Framework Programme, and by the World Bank. It reflects only the authors' views, and the Community and the World Bank are not liable for any use made of the information contained therein. An online supplement is available at http://jcr.sagepub.com/cgi/content/full/50/4/508/DC1/ and at http://folk.uio.no/hahegre/publications.htm.

JOURNAL OF CONFLICT RESOLUTION, Vol. 50 No. 4, August 2006 508-535
DOI: 10.1177/0022002706289303

decreases as average income increases and the size of a country's population decreases. Beyond these two results, however, there is little agreement. For example, Fearon and Laitin (2003, 85) and Hegre (2003, 26) find mountainous countries to have a higher risk of war than other countries (because they provide rebels with natural sanctuaries), but Collier and Hoeffler (2004) and Sambanis (2004a) do not find this result. Collier and Hoeffler find that a high degree of dependence on primary commodity exports increases a country's civil war risk, but few other empirical studies reach this conclusion.[1] There is thus uncertainty about the inferences that we can make based on empirical results on civil war.

Even though studies such as Fearon and Laitin's (2003) and Collier and Hoeffler's (2004) are exemplary in that they conduct several robustness checks, the additional analyses presented in any single article on civil war are typically not enough to sort out which of the discrepant findings are most tenable. Part of the reason for this is that there is no single method that is clearly better than others to test the fragility of empirical results.[2] Moreover, most scholars' robustness tests are ad hoc: they identify a set of competing explanations and see if their empirical results hold once they control for some variables that might be consistent with those explanations. While testing different model specifications allows us to discriminate between rival theories and increases our knowledge about the relationship between the dependent variable and a set of independent variables, ad hoc specification tests do not allow us to clearly establish if variation in empirical results is due to differences in research design, data sources, measurement and operationalization of IVs, period coverage, or estimation methods.

To establish if small changes to the model specification influence the robustness of empirical results and the inferences that we can draw from those results, Leamer (1985, 308) argued that we need a framework for "'global sensitivity analysis' in which a neighborhood of alternative assumptions is selected and the corresponding interval of inferences is identified. Conclusions [would be] judged sturdy only if the neighborhood of assumptions is wide enough to be credible and the corresponding interval of inferences is narrow enough to be useful." We follow Leamer's suggestion and explore two possible sources of variation in empirical results in the civil war literature: first, we test the sensitivity of commonly cited, substantively important results to small changes in the set of variables included in a regression. Second, we test how fragile our substantive inferences are to small changes in the way we operationalize theoretically significant variables (which we call "concept variables"). We focus on changes in the statistical significance of parameter estimates, but we also discuss changes in substantive effects that are due to small specification changes.

The article is organized in five sections beyond the introduction. Section 2 presents the methodology for sensitivity analysis. Section 3 provides a brief overview of the

1. See, for example, the special issue of *Journal of Conflict Resolution* (Ron 2005) exploring the link between natural resources and civil war.

2. From a purely theoretical standpoint, one might reasonably argue that only a single model specification should be estimated and that this specification should follow directly from the theory. However, most theories are not comprehensive, and we often need to test rival theories to determine empirically which one fits the data better.

range of variation in empirical results in the literature on civil war onset. The large number of apparently fragile results motivates our analysis. Section 4 presents the dependent variable. Section 5 details our results and identifies the most robust and most fragile results in the literature on civil war onset. Section 6 concludes with a summary of our findings.

2. METHODOLOGY OF SENSITIVITY ANALYSIS

Leamer (1985) referred to the largest possible set of inferences that can be drawn from a given data set as "extreme bounds." The size of these "extreme bounds" depends on the number of models that can be estimated (i.e., variations in model specifications) within the limits of the data set. These variations must be theoretically consistent and aim to show how minor changes in the list of variables alter the conclusions of estimation. In Bayesian terms, the extreme bounds approach (EBA) suggests that the analyst explores the range of posterior distributions that result from specification changes to the prior distribution of a variable. To credibly identify the range of inferences that may be drawn from a model, a "global sensitivity analysis" should be applied, "large numbers of variables should be included, as should different functional forms, different distributions, different serial correlation assumptions, different measurement error processes, etcetera, etcetera" (Leamer 1985, 311). Given the severe computational burden of such an approach, a reasonable compromise is to focus on selected dimensions of the model and regression coefficients in particular.

Levine and Renelt (1992) used a variant of Leamer's (1985) EBA to check the sensitivity of cross-country regression estimates on the determinants of economic growth. They were motivated by what they perceived as frequently contradicting empirical linkages between long-run growth rates and a wide array of explanatory variables. The EBA used by Levine and Renelt specifies equations of the following form: $Y = \boldsymbol{\beta}_\iota \mathbf{I} + \beta_\mu M + \boldsymbol{\beta}_z \mathbf{Z} + u,$ where Y is the dependent variable, $\mathbf{I}$ is a set of variables always included in the regression, M is the "focus" variable (i.e., the one whose behavior we are interested in observing as we change the model specification), and $\mathbf{Z}$ is a subset of control variables selected among several potentially significant explanatory variables. They first estimate a model that includes only the $\mathbf{I}$-variables and the focus variable and then estimate regressions for "all possible combinations of up to three $\mathbf{Z}$-variables and identify the highest and lowest values for the coefficient on the variable of interest, β_μ, that cannot be rejected at the 0.05 significance level" (Levine and Renelt 1992, 944). The design tries to reduce multicollinearity problems by restricting the total number of explanatory variables to "eight or fewer," choosing a "small pool of variables from which the extreme bounds procedure selected $\mathbf{Z}$-variables," and "excluding variables that, a priori, might measure the same phenomenon" (Levine and Renelt 1992, 944–5). This specification design minimizes the risk of underspecified models while also minimizing the computer power needed to estimate the models, and it reduces problems associated with multicollinearity. The extreme bounds on the coefficient β_μ denote the confidence that we can have in the partial correlation between Y and M. The upper extreme bound is defined as the maximum value of β_μ plus two standard deviations, and the lower extreme bound is β_μ minus two standard deviations of the estimate. Coefficient β_μ is considered robust if it "remains significant and of the same sign at the extreme bounds" (Levine and Renelt 1992, 944).

Levine and Renelt's (1992, 959) analysis leads them to conclude that "very few economic variables are robustly correlated with cross-country growth rates." But this extreme result may suggest that their analysis sets too high a hurdle for robustness. According to Sala-i-Martin (1997, 179), "If the distribution of the estimators of $\beta[_\mu]$ has some positive and some negative support, then one is bound to find one regression for which the estimated coefficient changes signs if enough regressions are run. Thus, giving the label of nonrobust to all variables is all but guaranteed." This argument is reasonable, so to assess the robustness of empirical results in the literature on civil war onset, we apply Sala-i-Martin's less stringent test, which involves looking at the entire distribution of parameter estimates to determine the level of confidence in each of the explanatory variables.

Sala-i-Martin (1997) estimates M models of the following form:

$$\gamma_j = \alpha_j + \boldsymbol{\beta}_{yj}\mathbf{y} + \beta_{zj} z_j + \boldsymbol{\beta}_{xj}\mathbf{x}_j + \varepsilon, \qquad (1)$$

where γ is the dependent variable and the subscript refers to the model, $\mathbf{y}$ is a vector of three variables that always appear in the regressions, z is the variable of interest, and $\mathbf{x}$ is a vector of up to three variables taken from a pool χ of available variables. He then uses these estimates of the β_{zj} to compute the cumulative distribution function—CDF(0)—the proportion of estimates that are larger or smaller than zero, selecting the largest of the two. The distribution function is computed under two different assumptions: the first aggregation assumes that their distribution across models is normal. The average estimate is computed as $\hat{\beta}_z = \sum_{j=1}^{M} \omega_{zj}\beta_{zj}$, where the weights ω_{zj} are proportional to the integrated likelihoods $\omega_{zj} = \frac{L_{zj}}{\sum_{i=1}^{M} L_{zi}}$ (Sala-i-Martin 1997, 179). The weights ensure that models with better fit contribute more to the estimate, and the fact that the same number of variables is always included in the regression implies that we do not get "artificially" better fit by increasing the number of variables. Similarly, Sala-i-Martin (1997) computes the average variance of the estimates, $\hat{\sigma}_z^2 = \sum_{j=1}^{M} \omega_{zj}\sigma_{zj}^2$. If the assumption that the β_{zj} are normally distributed holds, the probabilities that $\beta < 0$ and $\beta > 0$ can be computed from $\hat{\beta}_z$ and as $\hat{\sigma}_z^2$ functions of the average t-ratio $\hat{\beta}_z/\hat{\sigma}_z^2$. We will refer to the smallest of these two probabilities as the "average *p*-value." If the assumption of normality does not hold, then Sala-i-Martin computes the aggregate CDF(0) from the individual CDF(0)s,

$$\mathrm{CDF}(0) = \hat{\Phi}_z(0) = \sum_{j=1}^{M} \omega_{zj}\beta_{zj}\Phi_{zj}(0|\hat{\beta}_{zj}, \hat{\sigma}_{zj}^2).$$

We make a number of adaptations to Sala-i-Martin's (1997) procedure to fit our purpose. First, an obvious difference is that, while Sala-i-Martin applied the method to cross-sectional data with a continuous dependent variable, we are using time-series cross-sectional data, and is the logit-link function of the probability of an outbreak of civil war. Nonetheless, most papers in the civil war literature treat the data as cross-sectional since they use pooled logit or pooled probit estimators. Thus, we can apply Sala-i-Martin's method directly. Yet the time series in our data may create temporal dependence, which we account for by using a function of "time at peace" as one of our core variables (Beck, Katz, and Tucker 1998; Raknerud and Hegre 1997). We include the same three **y** variables in all regressions and refer to them as our "core" variables. In a review of the quantitative literature on civil war, Sambanis (2002) identified the following three core variables that are almost always included in models of civil war onset: the natural log of population (*lnpop*), the length of peacetime until the outbreak of a war (*pt8,* which we model as a decay function of time at peace), and the natural log of per capita gross domestic product (GDP) in constant dollars (*lngdp*). From a theoretical perspective, income is perhaps the most important variable in models of civil war and other forms of political violence. Keeping these three core variables always in the model, we estimate many other models for each of the variables that we include in the pool χ of relevant variables (we explain how we selected variables in the next section).

Sala-i-Martin (1997) estimates each possible combination of three **x**-variables for each *z*-variable. We deviate somewhat from this in our treatment of the **x-**variables. Our data contain several different measures for the same concept variable (the theoretically significant variable). For instance, we can measure "the level of democracy" using data from the Polity project, from Przeworski et al. (2000), or from Freedom House. To avoid including variables that measure the same thing in the same model, we restrict the combination of **x-**variables to those that measure three different *concept* variables—not simply different operationalizations of the same concept. If a control is an alternative operationalization of the same concept as the focus variable, this is likely to diminish the effect of the focus variable. This might not be an insurmountable problem if all concepts had the same number of alternative operationalizations since any bias would be roughly equal for each focus variable. However, the number of alternative operationalizations for each concept varies (see next section). Allowing more than one operationalization for each concept then would tend to hurt variables belonging to concepts with many operationalizations.[3]

Another difference from Sala-i-Martin's (1997) procedure is that we have to allow for the fact that several variables in our data set have missing data, and the number of missing observations is different for different variables. In Sala-i-Martin's model, parameter estimates are weighted by the model's log-likelihood to ensure that models with better fit to the data are given greater weight. The log-likelihood, however, is dependent on the number of observations that are included in the estimation. Hence, we calculate weights based on the likelihood ratio index (LRI), which is analogous to the R^2. It is

3. These could be concept variables of great interest in the literature, such as ethnic heterogeneity, which explains why we have many different ways of measuring them. For example, ethnic heterogeneity is measured by Vanhanen's (1999) racial, linguistic, and religious heterogeneity index, or *ehet*; Fearon's (2003) *ef* index; and the widely used ethnolinguistic fractionalization index, *elfo*.

given as $1 - LL_m/LL_0$, where LL_0 is the log-likelihood with only the constant term, and LL_m is the log-likelihood of the model having just been estimated. This statistic has the advantage that it is less dependent on sample size, and this is important for us as some of our variables are missing observations and sample size varies across models.[4]

A few important variables in the civil war literature cannot be represented in a single *z* term; in particular, we cannot use a single variable to capture nonlinear effects, such as the hypothesized inverted U-shaped association between civil war risk and democracy level or the parabolic association between civil war risk and primary commodity export dependence. We enter each of these as individual terms (so we control, for example, for primary commodity exports and their square) and report results for the robustness of the individual terms since we cannot evaluate their joint significance or joint robustness with the method we employ.[5]

The fact that we perform extensive specification tests implies that we do not know the true model.[6] We only know that three variables should be in the model: GDP per capita, population size, and time since the previous war. While most scholars would agree that such a model is underspecified, they would disagree over which other variables to add.[7] The method that we apply here could help provide some information on what is a better specified regression equation for civil war onset by testing the fit of several theoretically relevant variables. But it certainly does not replace the need for theorizing about civil war, and the main usefulness of our approach is to give us a sense of the distribution of empirical estimates for all potentially relevant variables.

A concern with our approach as well as with all studies of civil war onset is that some of the variables included in civil war models may be endogenous. We cannot deal with this problem while trying to explore empirically the robustness of different model specifications here, and we assume exogeneity for all variables (as do almost all the studies that we surveyed in this literature). To deal with endogeneity, we would have to switch estimators for those models where we suspected endogeneity as a result of the variable combinations in the model. We could not simply use instrumental

4. Fearon and Laitin (2003) suggest that some of the differences between their results and those from other models may be due to listwise deletion of cases due to nonrandomly missing data in other studies. Our approach should be less vulnerable to this problem. The LL_0 will also change if the number of observations changes. Also, 90 percent of likelihood ratio indexes (LRIs) fall between .050 and .11, so even if differences in sample size influence weights, we do not weigh any models too heavily.

5. For related reasons, we leave out interaction terms to simplify the analysis and reduce statistical output, but the method can be applied while using interaction terms.

6. In the presence of theoretical ambiguity, others have proposed a very different approach, which prefers parsimony to empirical exploration. Achen (2002), for example, proposes using no more than three independent variables in regression analyses in the absence of a formal model that justifies the inclusion of more controls. We do not share this view and agree more with much of the literature in development and labor economics, which is now moving away from an emphasis on formal models as a motivation for empirical analysis. While we certainly see the value of mathematical models in political science, we do not think they represent the only way to theorize about politics or to identify hypotheses for empirical testing. For a related, useful methodological and applied discussion of this point, see Oneal and Russett (2005).

7. Scholars often add controls to avoid the risk of omitted variable bias, even though their theory does not call for additional variables. For a useful methodological perspective on this issue, see Clarke (2005).

variables estimation for all models since, if exogeneity cannot be rejected, this would reduce the efficiency of the estimates. And we could not hope to identify valid instrumental variables for all model specifications. Since very few of the papers in the literature on civil war deal with the issue of endogeneity,[8] we also ignore it and simply try to reduce the risk by lagging independent variables.[9]

Finally, there is some concern with how missing data influence the results of this exercise. Many of the variables in our data set have missing observations. The missing variables problem is both more and less problematic for our approach as compared to the standard regression-based analysis of civil war. It is more problematic because we include in our model variables with a large number of missing observations. However, given that there is no consensus on which variables to include in a civil war model and hence no basis for identifying or excluding the possibility of omitted variable bias, this problem applies in principle to any empirical study of civil war. A potentially more serious problem is that we aim to isolate the degree to which individual parameter estimates are sensitive to changes in the model specification. Varying the model specification, however, leads to changes in the sample when variables are missing data for different observations. This makes it difficult to identify whether changes in results are due to changes in the specification or to changes in the sample. We reduce this problem by omitting from the analysis variables that lead to large losses of cases. This problem should not bias any of the parameter estimates as long as the reason for missing observations is not correlated with the dependent variable.[10]

3. VARIATION IN EMPIRICAL RESULTS ON CIVIL WAR ONSET

A brief overview of the theory and evidence in the literature on civil war onset explains how we selected the specification of our core model. We include income per capita as a measure of the economic opportunity cost of the war (Collier and Hoeffler 2004), or of some aspect of state capacity (Fearon and Laitin 2003), and other economic factors influencing the decision to rebel. We control for population size because, for a constant per capita propensity to initiate armed resistance, the definition of a civil war, which classifies armed conflict as a civil war only if there is a high threshold of deaths, implies that civil wars are more likely to occur in

8. Exceptions include Sambanis (2000, 2001), Blomberg and Hess (2002), and Miguel, Satyanath, and Sergenti (2004).

9. Endogeneity may result in artificially good fit between models; hence, weighting by log-likelihood or the LRI could result in weighting the wrong models more. This problem is limited in magnitude, however. Ninety percent of the LRIs fall between .050 and .11, and the largest weight given is .182. In addition, some combinations of variables may present more problems than others. For example, the inclusion of both economic growth and trade may violate the exogeneity assumption since both trade and growth are affected by conflict or the expectation of conflict. Variables that measure the degree of militarization (e.g., military expenditures; size of the government army) may also be endogenous to the expectation of political violence or, certainly, to ongoing war.

10. Any systematic measurement error in the independent variables would influence parameter estimates and would therefore make it harder to identify robustly significant coefficients.

populous countries. We include time at peace since the last civil war since peace is likely to exhibit positive time dependence: the longer a country is at peace, the lower should be the risk of (another) war as conflict-specific capital remains unused and peace-specific capital is accumulated (Collier and Hoeffler 2004; Hegre et al. 2001). These three variables are always part of civil war models.

Both theoretical and empirical support for many other explanatory variables is mixed. There is no theoretical agreement on what is the "right" set of variables to include in the model, and there is also mixed empirical support for many variables. Our approach consists of creating categories of "concept variables" that correspond to different theoretical arguments found in at least one study. For each concept, we have several empirical proxies that are supposed to measure the theoretically significant variables. We list all concept categories in Table 1, along with a description of variables in each category and the sources for each variable. A full discussion of the various theories and empirical results would take us too far, so we refer readers to cited sources for more information. We only briefly address some of the key debates to motivate the empirical focus of our article and provide some context for the use of these variables.

A heavily debated question is the effect of ethnic difference (concept 1) on the risk of civil war. The most commonly used measure for ethnic difference is the now well-known index of ethnolinguistic fractionalization (*elfo*).[11] Most studies find this to be nonsignificant, although studies of economic growth and public goods provision (see, e.g., Mauro 1995; Alesina, Baqir, and Easterly 1997; Easterly and Levine 1997) find it to be significant.[12] *Elfo* ranges from 0 (ethnic homogeneity) to 100 (extreme ethnic heterogeneity) and measures the probability that two randomly selected individuals belong to different ethnolinguistic groups.[13] We also include the square term to capture arguments about nonlinear effects of fractionalization (Collier and Hoeffler 2004). We assume that *elfo* is exogenous to civil war.[14] *Elfo*, like all similar measures, is time invariant and acts much like a country-specific effect.

We also use Vanhanen's (1999) index of ethnic heterogeneity (*ehet*), which ranges from 0 (homogeneity) to 144 (heterogeneity) based on 1990s data and combining subindices of racial division, linguistic division, and religious division. Vanhanen measured the size of the largest group as a percentage of total population and

(text continues on p. 521)

11. We use an expanded version of the index that includes data from the original Soviet atlas that were not included in Taylor and Hudson (1972).

12. This suggests an obvious indirect link between elfo and civil war—ethnic difference can increase the risk of civil war through its effects on these other variables, such as economic growth and income. However, no study has yet considered these indirect effects. The same is true for inequality. Most studies look for a direct linear relationship between economic inequality and civil war. The relationship, however, may be indirect: inequality may lead to regime transition, which may lead to civil war (see Sambanis 2005).

13. We do not explore here important questions such as what constitutes a different language, how groups are identified, what level of disaggregation is used to distinguish different languages, or how bilingual and trilingual groups are coded.

14. Certainly, ethnic identity, or at least the salience of ethnic identity, is shaped by social processes, including war. However, these are long-term processes, so we maintain an assumption of weak exogeneity for ethnic and religious divisions in the short run. In addition, *elfo* was coded in 1960, the start of our data set, which further justifies the exogeneity assumption.

TABLE 1
Concept Categories and Variable Names

		Variable Name	*Description*	*Source*
Dependent variable		*warstns*	War onset	Sambanis (2004a)
Core variables		*uppac*	Internal armed conflict onset	Gleditsch et al (2002)
		ln_popns	Population, log-transformed	Fearon and Laitin (2003)
		ln_gdpen	GDP per capita, log-transformed	Authors' coding, based on Sambanis (2004a)
		pt8	2^(-years in peace/8)	
Concept Label	*Concept No.*	*Variable Name*	*Description*	*Source*
Ethnic fragmentation	1	*dlang*	Linguistic component of *ehet*	Vanhanen (1999)
	1	*drace*	Racial component of *ehet*	Vanhanen (1999)
	1	*drel*	Religious component of *ehet*	Vanhanen (1999)
	1	*ef*	Ethnic fractionalization index	Fearon and Laitin (2003)
	1	*ef2*	*Ef* squared	Fearon and Laitin (2003)
	1	*ehet*	Ethnic heterogeneity index	Vanhanen (1999)
	1	*elfo*	Ethnolinguistic diversity	Collier and Hoeffler (2004)
	1	*elfo2*	Ethnolinguistic diversity, squared	Collier and Hoeffler (2004)
	1	*numlang*	Number of languages in Ethnologue	Fearon and Laitin (2003)
	1	*plural*	Share of largest ethnic group	Fearon and Laitin (2003)
	1	*plurrel*	Size of largest confession	Fearon and Laitin (2003)
	1	*relfrac*	Religious fractionalization	Fearon and Laitin (2003)
Ethnic dominance/polarization	2	*etdo4590*	Ethnic dominance measure	Collier and Hoeffler (2004)
	2	*second*	Percent population in second largest group	Fearon and Laitin (2003)
Level of democracy	3	*auto4*	Autocracy index from Polity IV	Marshall and Jaggers (2000)
	3	*dem*	Dummy: 1 for democracies and 0 for autocracies	Marshall and Jaggers (2000)
	3	*dem4*	Democracy index from Polity IV	Marshall and Jaggers (2000)
	3	*exrec*	Executive recruitment concept variable; Polity IV	Marshall and Jaggers (2000)

	3	*mirps2*	Consistent autocracy	Gates et al. (2006)
	3	*mirps3*	Consistent democracy	Gates et al. (2006)
	3	*parcomp*	Competitiveness of participation; nonelites; Polity IV	Marshall and Jaggers (2000)
	3	*parreg*	Regulation of participation; Polity IV	Marshall and Jaggers (2000)
	3	*part*	ln(share of population voting × opposition's share of votes cast)	Vanhanen (2000)
	3	*pol4*	Polity index; Polity IV	Marshall and Jaggers (2000)
	3	*pol4m*	Polity Index; Polity IV; 77 & 88 coded = 0	Marshall and Jaggers (2000)
	3	*polcomp*	Political competition: concept variable; Polity IV	Marshall and Jaggers (2000)
	3	*reg*	Dummy: 1 for dictatorships and 0 for democracies; ACLP	Przeworski et al. (2000)
	3	*sip2*	Continuous measure of democracy	Gates et al. (2006)
	3	*xconst*	Executive constraints—operational independence of CE; Polity IV	Marshall and Jaggers (2000)
Inconsistency of political institutions	4	*anoc*	Dummy: anocracy = 1	Fearon and Laitin (2003)
	4	*mirps0*	Inconsistent polity (semi-democracy)	Gates et al. (2006)
	4	*mirps1*	Caesaristic polity	Gates et al. (2006)
	4	*partfree*	Partially free polity	Freedom House (2002)
	4	*pol4sq*	*Pol4* squared	Marshall and Jaggers (2000)
Political instability	5	*ager*	Age in years of the current regime as classified by REG; ACLP	Przeworski et al. (2000)

(continued)

TABLE 1
(continued)

Concept Label	Concept No.	Variable Name	Description	Source
		Variable Name	*Description*	*Source*
Dependent variable		*warstns*	War onset	Sambanis (2004a)
Core variables		*uppac*	Internal armed conflict onset	Gleditsch et al (2002)
		ln_popns	Population, log-transformed	Fearon and Laitin (2003)
		ln_gdpen	GDP per capita, log-transformed	Authors' coding, based on Sambanis (2004a)
		pt8	2^(-years in peace/8)	
Concept Label	*Concept No.*	*Variable Name*	*Description*	*Source*
	5	*autch98*	Autocracy annual change; Polity 98	Gurr and Jaggers (2000)
	5	*demch98*	Democracy annual change; Polity 98	Gurr and Jaggers (2000)
	5	*durable*	Years since last regime transition/since 1949; Polity IV	Marshall and Jaggers (2000)
	5	*Inst3*	Political instability; whether there was a change in Polity score in the past three years	Fearon and Laitin (2003)
	5	*nwstate*	New state	Fearon and Laitin (2003)
	5	*p4mchg*	Annual change in modified polity; Polity IV	Marshall and Jaggers (2000)
	5	*polch98*	Polity annual change; Polity 98	Gurr and Jaggers (2000)
	5	*proxregc*	2^(-durable/0.5)	Gurr and Jaggers (2000)
Political system	6	*incumb*	Consolidation of incumbent advantage; ACLP	Przeworski et al. (2000)
	6	*inst*	0-dict; 1-parliam; 2-mixed dem; 3-pres dem; ACLP	Przeworski et al. (2000)
	6	*major*	Majoritarian system	Przeworski et al. (2000)
	6	*presi*	Presidential system	Przeworski et al. (2000)
Centralization	7	*autonomy*	Country has de facto autonomous regions	Sambanis (2004a)
	7	*centpol3*	Centralized state? (Polity III data plus updates for post-1994)	Sambanis (2004a)
	7	*fedpol3*	Federal state? (Polity III data plus updates for post-1994)	Sambanis (2004a)
	7	*semipol3*	Semi-federal state? (Polity III data plus updates for post-1994)	Sambanis (2004a)

Neighborhood political economy	8	*avgnabo*	Average SIP score of neighbors	Gates et al. (2006)
	8	*nmdgdp*	Neighbors' average ln(GDP per capita)	Fearon and Laitin (2003); Sambanis (2001)
	8	*nmdp4_alt*	Neighbors' median polity (both land and water contiguity; using polity2)	Marshall and Jaggers (2000)
	8	*regd4_alt*	Median regional polity (using polity2)	Sambanis (2002)
Region	9	*geo1*	Region: Western Europe and the United States	Authors' coding
	9	*geo2*	Region: Eastern Europe and Central Asia	Authors' coding
	9	*geo34*	Region: Middle East and North Africa	Authors' coding
	9	*geo57*	Region: South and East Asia and Oceania	Authors' coding
	9	*geo69*	Region: Latin America	Authors' coding
	9	*geo8*	Region: Sub-Saharan Africa	Authors' coding
Neighborhood war	10	*nat_war*	Whether a neighbor is at war in a given year	Sambanis (2004a)
	10	*tnatwar*	Total number of neighbors at war in a given year	Sambanis (2004a)
Growth	11	*gdpgrowth*	Annual change in GDP, percent	Computed from Fearon and Laitin (2003) GDP series
Economic policy	12	*expgdp*	Exports of goods and services as percent GDP; WDI data	World Bank (2000)
	12	*trade*	Trade as percept of GDP; in 1995 constant dollars	World Bank (2000)
Social welfare	13	*illiteracy*	Percent adult population illiterate; WDI	World Bank (2000)
	13	*infant*	Infant mortality; WDI	World Bank (2000)
	13	*life*	Life Expectancy at birth; WDI	World Bank (2000)
	13	*pri*	School enrollment, primary, percent gross; WDI	World Bank (2000)
	13	*seceduc*	School enrollment, secondary, percent gross; WDI	World Bank (2000)
Resources	14	*agexp*	Agricultural raw materials exports as percentage of merchandise exports	World Bank (2000)
	14	*fuelexp*	Fuel and oil products exports as percentage of merchandise exports	World Bank (2000)
	14	*fuelexp*	Fuel and oil products exports as percentage of merchandise exports	World Bank (2000)

(continued)

TABLE 1
(continued)

		Variable Name	Description	Source
Dependent variable		*warstns*	War onset	Sambanis (2004a)
Core variables		*uppac*	Internal armed conflict onset	Gleditsch et al (2002)
		ln_popns	Population, log-transformed	Fearon and Laitin (2003)
		ln_gdpen	GDP per capita, log-transformed	Authors' coding, based on Sambanis (2004a)
		pt8	2^(-years in peace/8)	
Concept Label	*Concept No.*	*Variable Name*	*Description*	*Source*
	14	*manuexp*	Manufactures exports as percentage of merchandise exports; WDI	World Bank (2000)
	14	*oil*	Oil exports/GDP	Fearon and Laitin (2003); Sambanis (2004a)
	14	*sxpnew*	Primary commodity exports/GDP	Collier and Hoeffler (2004)
	14	*sxpsq*	Primary commodity exports/GDP, squared	Collier and Hoeffler (2004)
Terrain, geography, population distribution	15	*lmtnest*	Rough terrain	Fearon and Laitin (2003)
	15	*ncontig*	Noncontiguous state	Fearon and Laitin (2003)
	15	*popdense*	Population density: people per square kilometer; WDI	World Bank (2000)
Militarization	16	*army85*	Size of government army in 1985	World Bank (2000)
	16	*milper*	Military personnel in thousands	World Bank (2000)
Time	17	*decade1*	Dummy: 1960s	Authors' coding
	17	*decade2*	Dummy: 1970s	Authors' coding
	17	*decade3*	Dummy: 1980s	Authors' coding
	17	*decade4*	Dummy: 1990s	Authors' coding
	17	*coldwar*	Code 1 for cold war year—before 1990	Authors' coding
Colonial war	18	*warhist*	War in the country since 1945?	Authors' coding

NOTE: GDP = gross domestic product; ACLP = Alvarez, Cheibub, Limongi, and Przeworski's democracy score; CE = chief executive; REG = regime; SIP = scalar index of polities; WDI = world development indicators.

summed across the inverse of these ratios for all three categories (largest racial, linguistic, religious group).[15] We also include variables, coding the number of ethnic, linguistic, or religious groups from Fearon and Laitin (2003).

Several authors argue that polarization (a small number of large groups) or dominance (one very large group) more frequently lead to civil war than does mere fractionalization (e.g., Collier and Hoeffler 2004; Reynal-Querol 2002). This is our concept 2 variable. We include two measures of it: Fearon and Laitin's (2003) measure of the share of the population in the second largest group and Collier and Hoeffler's (2004) dichotomous "ethnic dominance" measure.

The significance of political institutions (concept 3) is also heavily debated. Some recent studies point to a significant negative association between civil war and the level of democracy (Esty et al. 1995; Esty et al. 1998; Gurr 2000), although most find the level of democracy to be nonsignificant (Collier and Hoeffler 2004; Fearon and Laitin 2003). We have several measures of democracy.

There is more evidence that concept 4, institutional inconsistency or anocracy (Hegre et al. 2001; Fearon and Laitin 2003), is significant and that it increases the risk of war, as does concept 5—political instability (Hegre et al. 2001; Fearon and Laitin 2003; Elbadawi and Sambanis 2002). We include several measures of these two concepts. Among others, we include a decay function *proxregc* of the Polity *durable* variable, which measures the number of years since an institutional change that leads to a minimum of three points' change on the Polity index.

Moving beyond democracy and considering the type of political institution (concept 6), Reynal-Querol (2002) finds that proportional representation reduces the risk of civil war. However, she only analyzes ethnic wars and focuses on war prevalence, not onset. We include her measures along with two measures from Przeworski et al. (2000).

The degree of political centralization may also influence civil war risk (concept 7). We include this concept because many civil wars are fought over self-determination claims, and several authors have argued that more decentralization could reduce the incentives for violence by peripheral groups (Lijphart 1977; Hechter 2001).

Another debated question is the regional dimension of civil wars. We group these variables into three concepts: concept 8 refers to the political economy of the neighborhood, including the average income and democracy level of neighboring countries; concept 9 includes indicator variables for the regions themselves; and concept 10 measures the extent of war in the neighborhood. Some authors find important economic spillover effects of civil war (Murdoch and Sandler 2002), and others (Sambanis 2001) find that civil war in a neighboring country in the previous year increases a country's risk of civil war onset, although Hegre et al. (2001) and Fearon and Laitin (2003) find that these neighborhood spillovers are not significant.

15. Vanhanen's (1999) computations may have an error: *ehet*'s lowest value is 0 (perfect homogeneity). But if *ehet* is the sum of the inverse ratio of the largest group over the population, then its smallest value should be 3 (a country with a single group comprising 100 percent of the population in all three cleavages should be coded as 1 + 1 + 1 = 3).

Similarly, there is disagreement over the effects of living in "bad" neighborhoods—that is, in regions with mostly poor, autocratic regimes, which some say will increase the risk of civil war in the neighborhood (Sambanis 2001; Gleditsch 2002). We explore the empirical robustness of these concept variables.

The role of economic factors has also been the focus of several studies. Several authors have shown that economic growth (concept 11) reduces civil war risk (Collier and Hoeffler 2004; Hegre 2003). Results are ambiguous, maybe since the variable used in most data sets has many missing observations. We use a variable with a more complete time series.

Some scholars have considered concept 12—the relationship between economic policy, particularly trade, and civil war (Esty et al. 1995; Hegre, Gleditsch, and Gissinger 2003; Elbadawi and Hegre 2004). So, we also control for that.

Another concept variable is social welfare (concept 13), which includes infant mortality (Esty et al. 1998) and education (Collier and Hoeffler 2004). Many scholars argue these should have a separate effect beyond that of income, but the results are again inconclusive.

The importance of natural resources (concept 14) has figured prominently in the debate (Collier and Hoeffler 2004; Collier et al. 2003; Fearon and Laitin 2003; Ross 2004; Ron 2005). We include several measures of how important natural resources are as compared to manufactures.

Other aspects of the country's demographic characteristics and geography are also potentially important (concept 15), as they affect the government's ability to stage a counterinsurgency campaign. We include measures of the terrain (from Fearon and Laitin 2003) and a measure of population density.

The degree of militarization of the country and the government's military strength (concept 16) may influence the risk of war as governments with strong militaries can more effectively deter insurgency or repress it before it rises to the level of civil war (Sambanis 2000). We include a measure of the size of military expenditures and the size of the government army.

The risk of civil war may also be a function of systemic factors (concept 17). The end of the cold war is sometimes thought to have unleashed ethnic conflict, although the evidence is inconclusive (Sambanis 2001; Fearon and Laitin 2003). We control for the cold war and also use decade dummies to control for time trends.

Finally, we include information on countries' colonial war history (concept 18) as some authors have argued that countries with different colonial histories followed different trajectories of institutional development, and this affects their civil war proneness.

The fact that we have several proxies for each concept means that we can gauge the extent to which our causal inferences about a theoretical concept are specific to a proxy variable. This is a topic of debate in the literature since scholars have identified a disjuncture between the theoretically significant variables and some of the proxies used in empirical tests (Sambanis 2004b). If results for different proxies of the same concept are inconsistent, this means that we need to define the concept better and find a more direct measure of it.

4. THE DEPENDENT VARIABLE

Differences in empirical results may also be due to the fact that there are important differences in the coding of civil war across studies (Sambanis 2004a). Depending on which data set we use, we can have twice as many civil war starts and country-years at war. These differences across civil war lists have been shown to affect some empirical estimates significantly. We do not combine our sensitivity analysis with one that checks robustness to changes in the civil war list because this would make the analysis very hard to follow.

We primarily present results based on Sambanis's (2004a) civil war definition. A civil war is defined as an armed conflict between an internationally recognized state and (mainly) domestic challengers able to mount an organized military opposition to the state. The war must have caused more than 1,000 deaths in total and in at least a three-year period. We refer to this variable as "CW" below. We also ran our analysis using the Uppsala/PRIO list of internal armed conflict (Gleditsch et al. 2002). The Uppsala/PRIO data set has an annual twenty-five battle deaths threshold and therefore is not limited to civil war. The variable is labeled IAC in what follows.[16]

Our dependent variable (civil war onset) is coded 0 for all country-years with no war, 1 for the year that a war started, and missing for periods of ongoing war. This causes us to lose some observations of wars that start while another war is ongoing in the same country. The alternative way is to code periods of ongoing war as 0s (except the year of onset), but countries with ongoing wars may have a systematically different risk of a new war, and we would need to control for that as well as to consider the effects of the ongoing war on the other explanatory variables. We therefore drop ongoing war years in the CW analyses.[17] For the IAC conflicts, where multiple conflicts going on in the same country are more frequent, we code ongoing conflict years as 0s.

5. RESULTS OF SENSITIVITY ANALYSIS OF MODELS OF CIVIL WAR ONSET

We analyze the sensitivity of the eighty-eight variables included in Table 1, organized by concept category. We cover the same set of countries over the same period and draw data from Sambanis (2001, 2004a), Hegre et al. (2001), Fearon and Laitin (2003), Reynal-Querol (2002), and Collier and Hoeffler (2004). We standardized variables to make estimates easily comparable, and we also log-transformed many variables to minimize the effect of extreme values.[18]

The combinations of the eighty-eight variables, following the method presented in the previous section, produce 4.7 million logistic regressions. Table 2 reports the results for the three core variables for the CW and IAC dependent variables.

16. See Sambanis (2004a) for more details on the coding rules and explanations for how each case was coded. See our supplement for complete results based on the Uppsala/PRIO list.

17. We did run the analysis, coding periods of ongoing war as 0s, and there were no significant differences. See our online supplement for a table with results.

18. We standardized by subtracting the mean and dividing by the standard deviation.

TABLE 2
Results for the Core Variables

Variable	*Label*	*Estimations (Sample)*	*Weighted Mean* $\hat{\beta}_z$	*Weighted Mean* SD ($\hat{\beta}_z$)	p *Normal (One-Sided Test)*	p *Nonnormal (One-Sided Test)*
Sambanis civil wars						
ln_popns	Ln(population)	4,700,394	.280	.090	.00092	.0055
ln_gdpen	Ln(GDP per capita)	4,700,394	–.504	.197	.0053	.031
pt8	2^(-years in peace/8)	4,700,394	.730	.492	.069	.105
Uppsala/PRIO internal armed conflicts						
ln_popns	Ln(population)	4,700,394	.399	.093	< .00001	.000085
ln_gdpen	Ln(GDP per capita)	4,700,394	–.438	.171	.0052	.032
pt8	2^(-years in peace/8)	4,700,394	.248	.341	.234	.175

The standardized coefficient for the population variable had a mean estimate of .280 over the CW estimations and a mean standard error of .090 (standard errors are corrected for clustering by country). On average, this corresponds to a t-value of 3.12 and an average p-value of 0.00092. Recall that this statistic is based on the assumption that the β_{zj} is normally distributed and corresponds to a one-sided test of statistical significance. The top panel in Figure 1 shows how the t-values are distributed over the 4,700,134 estimations. We have added vertical lines at $t = -1.65$ and $t = 1.65$ to represent the boundaries of a fairly lenient definition of statistical significance. The bars inside these lines indicate the frequency of nonsignificant parameter estimates. The distribution of the population variable estimates is quite close to a normal distribution. If we relax the assumption of the normal distribution of estimates, the p-value is .0055 (see the right-most column in Table 2).

The GDP per capita variable also is robustly significant, with a mean t-value of –2.5 and an average p-value of .0053 under the assumption of normal distribution. The middle panel in Figure 1 shows that the estimates have a much more skewed distribution than the population variable. The mode of the distribution of t-values is below –3, but in approximately a quarter of the models, the variable is not significant at the 5 percent level. Relaxing the distribution assumptions, the average p-value is .031. The mean parameter estimate is –.504, so income is substantively more important than population. The estimates for log-transformed variables may be interpreted as elasticities, or roughly as the relative change in the odds of conflict when the variable is increased by 1 percent. A 1 percent increase in income reduces the risk by .5 percent. The corresponding figure for population is an increase of .28 percent. This comparison is possible because we have standardized the variables, allowing us to also interpret the estimates in terms of how much the odds of war change when the variable is increased by one standard deviation of its own distribution. Decreasing income by one standard deviation increases the risk of civil war by 65 percent,

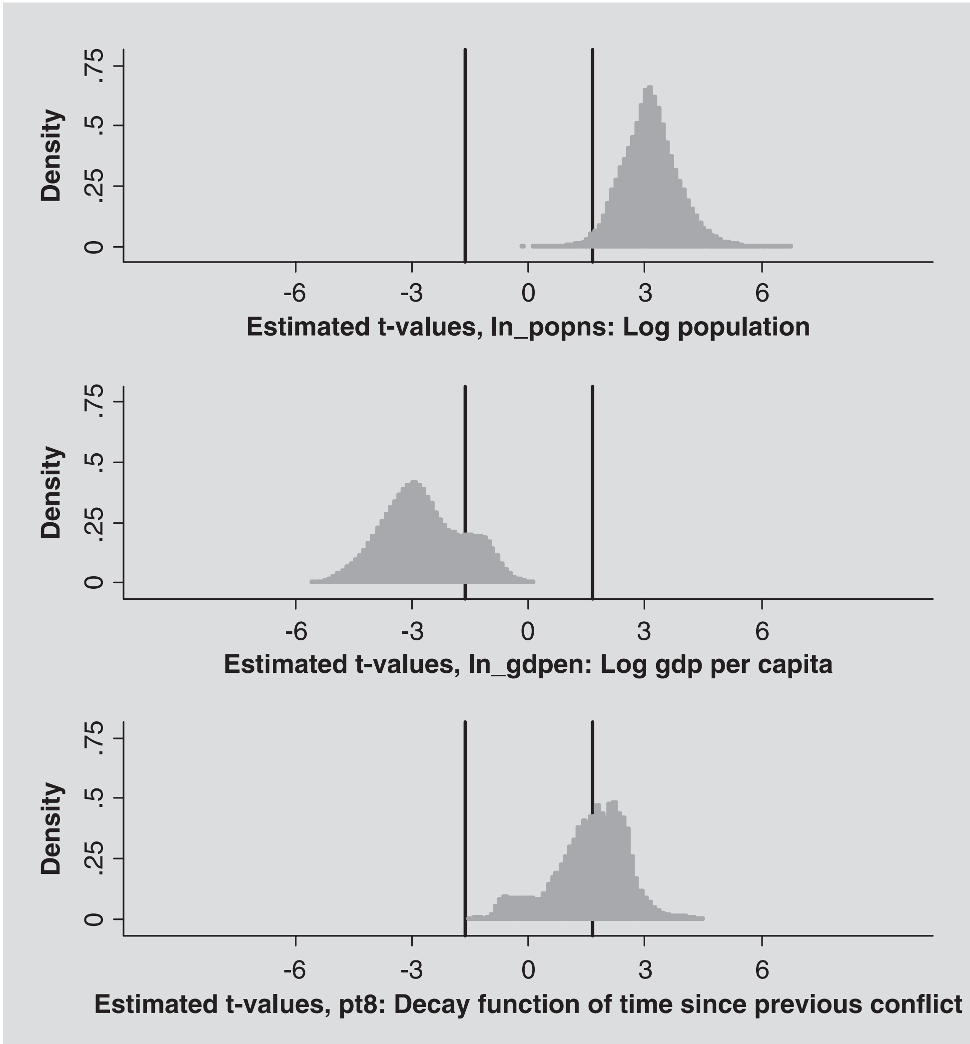

Figure 1: Distribution of Estimated *t*-Values: Core Variables

whereas *increasing* population by one standard deviation increases civil war risk by 32 percent.

Time since previous war is less robust than the other two core variables. The average *t*-statistic is 1.48, corresponding to an average *p*-value of .07 for a one-sided test. The bottom panel in Figure 1 shows the distribution of estimated *t*-statistics.

The lower half of Table 2 shows that the estimates for population size and GDP per capita are roughly comparable for the IAC dependent variable. The *pt8* variable is not robustly significant in these models.[19]

Table 3 lists the most robust variables for the CW estimations. The variables are ordered by the average *p*-value (normal distribution assumption). The third column

19. This has to do with the coding of ongoing conflict years as 0s. See Hegre (2003) for a discussion of this problem.

TABLE 3

Civil Wars: The Eighteen Most Robust Variables, Sorted by *p* Nonnormal

Variable	*Label*	*Concept*	*Weighted Mean* $\hat{\beta}_z$	*Weighted Mean* SD ($\hat{\beta}_z$)	p *Normal (One-Sided)*	p *Nonnormal (One-Sided)*	*Rank, Mean* ($\hat{\beta}_z$)
inst3	Political instability; whether Polity coded a change or 77 or 88 in previous three years	5	.448	.098	.00002	.00029	4
parreg	Regulation of participation; Polity IV	3	–.406	.140	.0018	.0064	6
geo34	Region: Middle East and North Africa	9	.289	.113	.0039	.012	13
proxregc	2^(-durable/.5)	5	.287	.115	.0049	.019	16
gdpgrowth	Annual change in GDP, percent	11	–.548	.223	.0055	.024	2
anoc	Dummy: anocracy = 1	4	.299	.122	.0060	.041	11
partfree	Partially free polity	4	.325	.131	.0062	.021	8
nat_war	Whether a neighbor is at war in a given year	10	.318	.141	.010	.025	9
lmtnest	Rough terrain	15	.294	.136	.014	.019	12
decade1	Dummy: 1960s	17	–.301	.155	.029	.041	10
pol4sq	*Pol4* squared	4	–.287	.158	.031	.072	15
nwstate	New state	5	.236	.134	.035	.065	20
regd4_alt	Median regional polity (using polity2)	8	–.416	.240	.035	.061	5
etdo4590	Ethnic dominance measure	2	.235	.142	.038	.065	21
milper	Military manpower in thousands	16	–.541	.330	.044	.015	3
geo1	Region: Western Europe and the United States	9	–.587	.351	.047	.052	1
tnatwar	Total number of neighbors at war in a given year	10	.172	.107	.049	.074	28
presi	Presidential system	6	–.373	.228	.050	.063	7

NOTE: Sambanis civil war list; ongoing war years omitted from data set. Complete results are given in the online supplement. GDP = gross domestic product.

lists the concept category corresponding to each variable. The fourth and fifth columns report the weighted average of parameter estimates and their standard errors, respectively. We report the average *p*-values in column 6 and *p*-values without the normality assumption in column 7. Table 4 presents the same information for the Uppsala/PRIO list of internal armed conflict. More detailed results for all eighty-eight variables can be found in our supplement.

We have identified robust operationalizations for up to twelve concept categories in Table 3, and eighteen variables have average *p*-values less than .05 under the least restrictive assumption (these are one-tailed tests since the hypotheses are usually directional; only nine variables are robustly significant with a two-tailed test). Since there are more onsets of Uppsala/PRIO armed conflicts (Table 4), *p*-values are generally lower. Twenty variables are robustly significant in this table. Two political instability variables are among the most robust variables in both tables. The *inst3* variable reflects whether there was a change in the country's Polity score within the three years prior to the year of observation or whether the country was coded as being in "transition" or "interregnum" by the Polity project. The inclusion of the transition cases in the definition of this variable may raise concerns about endogeneity, however. One of the Polity project's criteria for coding a country as being in a state "interregnum" is a complete collapse of central authority, which is likely to be connected with civil war outbreak. However, these transition cases form only a minority of the Polity score changes recorded in these variables and an even smaller proportion of the changes used in the coding of the *proxregc* variable, which is also significant and robust and increases the risk of civil war onset.[20] The proximity of regime change (*proxregc)* is also very robust using the Uppsala/PRIO definition (Table 4). Concept 5 variable *nwstate* (a dummy variable for the two first years of independence) is also among the robust variables (it is not robust for internal armed conflict).[21]

The second most robust variable in Table 3 is one of the "level of democracy" variables (concept 3). *Parreg* is one of the subindicators forming the Polity IV Project's Democracy index, and it reflects the extent to which political participation is regulated. However, endogeneity concerns are even more serious here than in the instability variables since Polity codes countries with extensive political violence as having nonregulated participation.[22] Other "level of democracy" variables are not robust, but several "inconsistency of political institutions" variables are quite robust (e.g., *partfree, anocracy*).

20. We reemphasize that our interest is to map the distribution of parameter estimates in the quantitative literature as it now stands; almost all of the literature assumes exogeneity of political variables, and if we change the estimator, these results will also probably change.

21. New state is a theoretically ambiguous variable. The particularities of new states are likely to be captured by several of the other variables since new states are more likely to have instability, be anocracies, and be more ethnically fractionalized. See Sambanis (2004a).

22. We reran the analysis, excluding the variables that are affected by this problem: *parreg*, *parcomp*, *anoc*, and *pol4sq*. The list of the most robust variables changes only marginally when these variables are omitted. The most notable changes are that the *p* (normal) increases from .0041 to .0082 for the *prxregc* variable and from .049 to .072 for the *tnatwar* variable. All other results change only marginally. This exercise increases our confidence that endogeneity problems do not overly distort the weights we gave to individual estimations (cf. fn. 9). Complete results are given in the online supplement.

TABLE 4

Internal Armed Conflict: The Twenty Most Robust Variables, Sorted by *p* Nonnormal

Variable	*Label*	*Concept*	*Weighted Mean* $\hat{\beta}_z$	*Weighted Mean* SD ($\hat{\beta}_z$)	p *Normal (One-Sided)*	p *Nonnormal (One-Sided)*	*Rank, Mean* β_z
ehet	Ethnic heterogeneity index	1	.388	.108	.000	.001	3
inst3	Political instability; whether Polity coded a change or 77 or 88 in previous three years	5	.344	.080	.000	.000	8
gdpgrowth	Annual change in GDP, percent	11	–.436	.093	.000	.001	2
dlang	Linguistic component of *ehet*	1	.382	.121	.001	.002	4
oil	Oil exports/GDP	14	.309	.114	.003	.007	9
decade4	Dummy: 1990s	17	.292	.106	.003	.014	11
ef	Ethnic fractionalization index	1	.364	.142	.005	.009	5
parreg	Regulation of participation; Polity IV	3	–.261	.108	.007	.020	15
elfo	Ethnolinguistic diversity	1	.348	.149	.010	.017	7
proxregc	2^(-durable/.5)	5	.198	.086	.011	.024	23
ef2	Ef squared	1	.264	.129	.019	.028	13
coldwar	Code 1 for cold war year—before 1990	17	–.215	.100	.019	.056	21
plural	Share of largest ethnic group	1	–.264	.131	.021	.031	14
autonomy	Country has de facto autonomous regions	7	.076	.039	.021	.063	60
milper	Share of population in military forces	16	–.451	.220	.021	.036	1
drel	Religious component of *ehet*	1	.232	.114	.023	.048	19
nat_war	Whether a neighbor is at war in a given year	10	.207	.101	.023	.047	22
elfo2	Ethnolinguistic diversity, squared	1	.268	.141	.029	.040	12
geo34	Region: Middle East and North Africa	9	.215	.114	.033	.066	20
sxpsq	Primary commodity exports/GDP, squared	14	.299	.175	.040	.063	10

NOTE: Uppsala/PRIO internal armed conflict list. Complete results are given in the online supplement.

The "neighborhood at war" (*natwar*) variable has robustly positive estimates for both conflict variables, lending support to hypotheses regarding the significance of diffusion and contagion effects in civil war, although we cannot distinguish among these different mechanisms through which civil war influences are transmitted across countries. The *tnatwar* version also makes the top eighteen in Table 3. Economic growth (*gdpgrowth*) is robustly negative in both tables, lending support to Collier and Hoeffler's (2004) "opportunity cost" theory of civil war. The regional democracy level (*regd4_alt*) is also robustly negative for the CW variable, suggesting that living in a bad neighborhood may be more important than a low in-country democracy level (Sambanis 2001). The rough terrain variable (*lmtnest*) is robustly positive for the CW variables. This highlights an important difference if we compare our specification tests to results obtained by using a single model while varying the measure of the dependent variable (as in Sambanis 2004a), where rough terrain is not significant.

The military personnel (*milper*) variable is negatively and robustly correlated with civil war onset. Countries with large militaries may be better able to deter insurgency or repress any opposition before it rises to the level of civil war.

The last two variables are less robust for the IAC dependent variable, however. Conversely, the square of primary commodity exports, the oil and autonomy variables, and the 1990s dummy are positive and robust for the IAC variable. There are several important differences in the robust determinants of lower armed conflict as compared to civil war, and explaining those differences should be an area of further research in the literature on political violence. The most striking difference between the two conflict variables concerns measures of ethnic fractionalization.

Only one measure of ethnic composition in the country has an average *p*-value smaller than .05 in Table 3: "ethnic dominance," as measured by Collier and Hoeffler (2004). This variable is not robust for the IAC variable. On the other hand, eight ethnic fractionalization variables are robustly associated with the risk of IAC conflict (Table 4). Apart from the "share of largest ethnic group" (*plural*), all are positive, increasing the risk of internal armed conflict. One of the fractionalization variables, *ehet*, even ranks as the most robust variable for IAC. None of these, however, are robust in the CW estimations. The weighted mean $\hat{\beta}_z$ for Fearon's (2003) ethnic fractionalization measure is many times higher when we look at all internal armed conflict using Uppsala/PRIO data relative to the civil war list (this is also true for other measures). Figure 2 shows that in the vast majority of the estimations, the ethnic fractionalization variables were not statistically significant for civil war, but they were for all internal armed conflict according to the Uppsala/PRIO definition (Figure 3). Thus, we have identified an important puzzle: why might ethnic fractionalization be associated with low-intensity armed conflict but not full-scale civil war?[23] And why does ethnic dominance seem to affect the risk of the most serious conflicts but not the lower level ones?

Finally, a result that has not received much attention is that there may be some geographic and time effects: the Middle East and North Africa region is robustly positive, and the 1960s dummy is robustly negative. Period and region effects have not been adequately explored in the civil war literature.

23. See Sambanis (2004a) for more discussion.

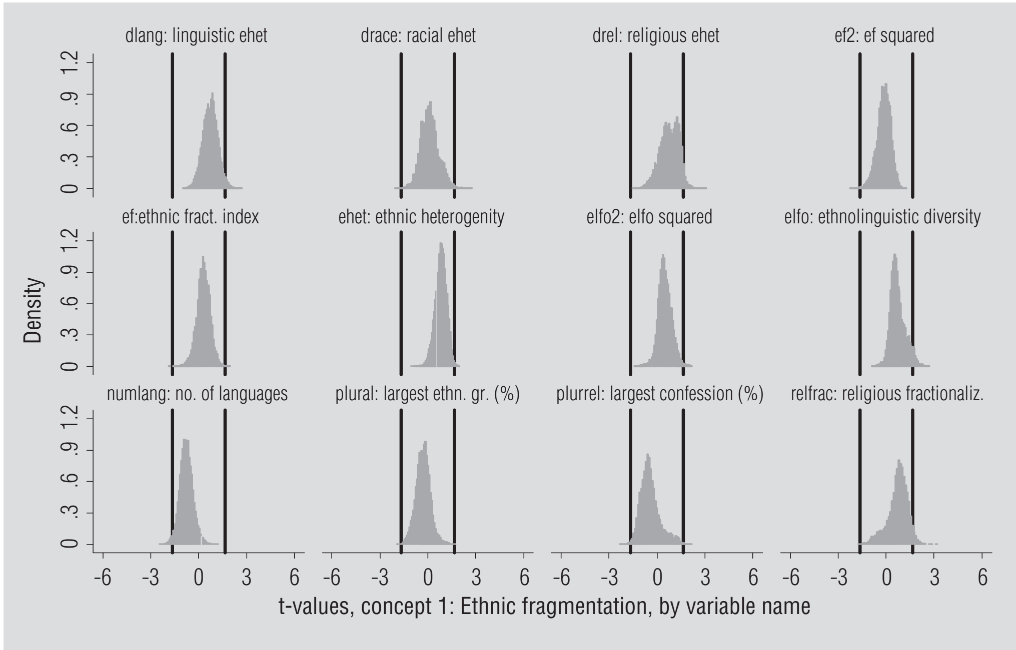

Figure 2: Distribution of Estimated *t*-Values for Ethnic Fragmentation: Civil Wars

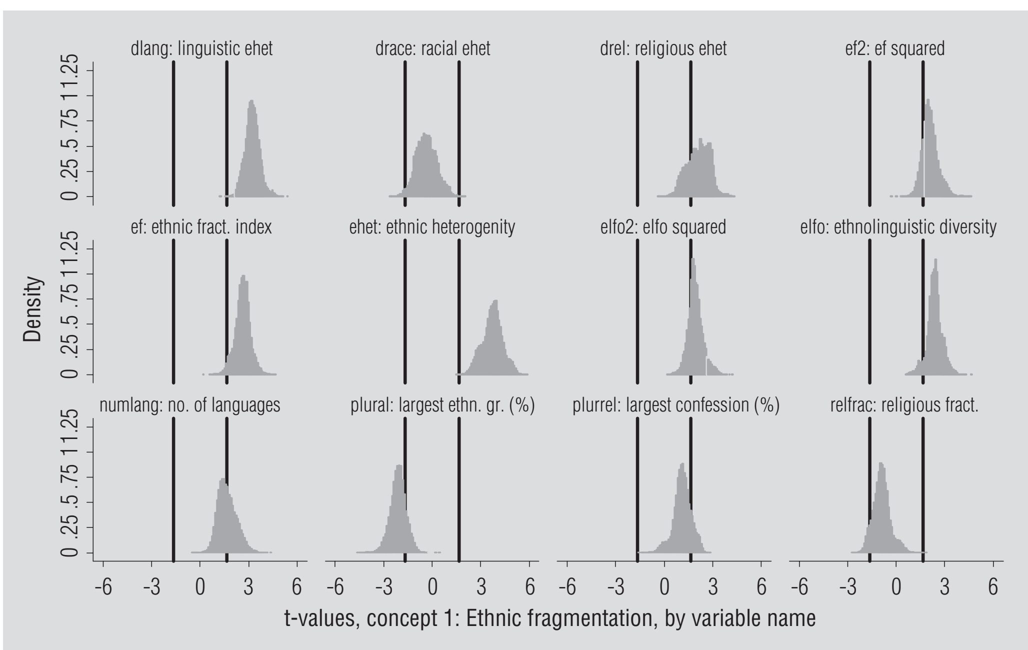

Figure 3: Distribution of Estimated *t*-Values for Ethnic Fragmentation: Internal Armed Conflict

The estimates for "weighted mean $\hat{\beta}_z$" (column 4) show that most of the variables that are significant in most models also are substantively important. The right-most column reports the variables' rank in terms of size of the mean estimate $\hat{\beta}_z$. Most of the variables in Table 3 are also among the top twenty variables with large effects (estimates larger than 0.25). There are some exceptions: military personnel (*milper*) ranks fifteenth in terms of average p-value but third in terms of average impact (mean $\hat{\beta}_z$). The variable with the largest substantive effect is a dummy for Western Europe and the United States (*geo1*). This is also statistically significant and relatively robust. Presidential political systems (*presi*) also have a large effect, but that variable is less robust. Other variables are also estimated to be substantively important but are rarely significant, as is the case with secondary school enrollment (one of Collier and Hoeffler's [2004] key variables), life expectancy at birth, and the share of manufacturing in exports.[24]

In our supplement, we present figures displaying the distribution of estimated t-values for the concept variables that are heavily debated in the literature: ethnic dominance/polarization, level of democracy, institutional inconsistency, political instability, and natural resource dependence. The graphs allow a direct comparison of the different proxies for each concept variable. The precise measure used to test the effects of democracy or instability matters, for example. Many variables that have received attention in policy debates and theoretical models of civil war are not robust. For example, nearly all measures of natural resource dependence (one of our concept variables) do not do well. Only dependence on oil exports may be considered marginally robust.

6. CONCLUSION

Using a methodology developed by Leamer (1985) and Sala-i-Martin (1997), we assess the robustness of a number of empirical results in the burgeoning civil wars literature. We find some of those results to be highly sensitive to small changes in the model specification. These specification changes amount to estimating models that include all possible seven-variable combinations of variables that conform to a set of rules set by our methodology and that might be theoretically relevant to explaining the onset of civil war. Some of the empirical results in the civil war literature are fragile, but others are not, and they are worth taking seriously. We confirm that a large population and low per capita income increase the risk of civil war, and this is consistent with many studies of civil war. We have found other robust relationships: civil wars are more likely to occur in countries with recent political instability and inconsistent democratic institutions; countries with small militaries and rough terrain; countries located in war-prone, undemocratic regions; and countries with low rates of economic growth. Most other variables fail our significance tests on average.

24. All three variables are highly correlated with gross domestic product per capita, which is a core variable in our analysis. This collinearity may lead to either inflated β_z estimates or to deflated t-values. Our analysis may therefore underestimate somewhat the robustness of these variables.

It was not our purpose in this article to sort out different mechanisms that underlie these significant correlations, and we do not claim to prove or disprove causal relationships. Rather, we only wanted to establish which correlations are more believable given the data and estimation methods used by the vast majority of researchers in this field. By establishing which empirical results are more robust, our article helps researchers focus their theory-building efforts and identifies empirical patterns that are worth more explanation. Our analysis gives readers a sense of the range of variation of parameter estimates when we try different model specifications. This sort of sensitivity analysis could be carried out in all empirical literatures and subfields in political science. While it might not be feasible for every empirical study to conduct as systematic an empirical investigation as the one that we have presented here, our results indicate that analysts should worry about the fragility of their inferences and should substitute closely related but differently measured variables in their regressions and add control variables to make sure the results on focus variables are robust before making causal inferences. There are several possible sources of uncertainty in the empirical literature on civil war that we have not considered, including uncertainty about the meaning and measuring of civil war. Different estimation methods are likely to produce different results, and adding interactions may do the same. All these sources of variability in empirical estimates should be considered. Our sensitivity analysis is therefore limited as it only explores robustness to changes in the set of control variables. But other researchers could (should) check robustness to changes in period and country coverage, sampling and estimation techniques, and so on.

Our analysis does not replace the need for careful theorizing about civil war. Our approach may be a complement to theoretically focused studies. We were motivated by the fact that no study to date has produced a clear theoretical justification for the model used in econometric tests. We do not know *the* model of civil war. Although most scholars agree that there are some variables that belong in the model, there is disagreement about others. Given the theoretical uncertainty about the causes of civil war, our analysis can help point to questions that need further scrutiny (e.g., the results on ethnic fractionalization that we mentioned earlier). It can also help by identifying empirical patterns that are robust, and we must be able to explain. The fact that all our variables have been considered by at least one other scholar as theoretically relevant for civil wars justifies their inclusion in the model and makes it hard to ignore robust empirical results. Some of our stronger results have clearly important theoretical implications that are not well understood. We found strong neighborhood effects of civil war consistent with Sambanis (2001), while other major studies (Fearon and Laitin 2003; Fearon 1998) have argued against the relevance of spillover effects. Our empirical approach, however, cannot explain why or how civil wars spread across neighboring countries. We need more theorizing and different sorts of empirical tests to uncover the mechanisms underlying this correlation. There could be different—often competing—mechanisms that explain how one country's civil war risk spreads around the region. There could be demonstration effects or diffusion effects, as rebel groups learn by observing others, or there could be contagion effects, as rebels spill over borders to incite rebellion by their coethnics in neighboring countries. Alternatively, the geographic clustering of civil wars may be due to shared attributes in countries with common borders. Our analysis points to the need to better understand the international dimensions of civil war.

Our results strongly support some, but not all, aspects of economic theories of civil war. We find oil export and primary commodities dependence to be robustly associated with lower level armed conflict only and not with civil war, contrary to arguments made by major studies of civil war. Other commonly used measures of resource dependence are not significant. The problem here may be more with the available proxies than with the underlying theory connecting resource dependence to civil war proneness. The key variables in economic theories—per capita income and, to a lesser extent, the rate of growth of income—are very robust.

But while our results on income and growth agree with Collier and Hoeffler's (2004) theory, we also find that political variables—institutional instability, incomplete democracy, and undemocratic neighborhoods—are very important and increase the risk of civil war, as argued in some recent studies (Sambanis 2001). By contrast, states with strong militaries are able to effectively deter or preempt civil war. These results are consistent with political theories of civil war (Hegre et al. 2001). The effects of neighborhood variables and political variables may have to be considered together: bad neighborhoods—those full of autocracies and civil war—might make democratic transitions in individual countries harder (cf. Gleditsch 2002). Thus, in undemocratic regions, we might be more likely to see democratization efforts resulting in prolonged periods of institutional instability, which we have argued are associated with increased risk of civil war. This conjecture is worth investigating further as we expand our theoretical focus to consider the spatial dimensions of civil war.

REFERENCES

Achen, Christopher. 2002. Toward a new political methodology: Microfoundations and ART. *Annual Review of Political Science* 5:423-50.

Alesina, Alberto, Reza Baqir, and William Easterly. 1997. Public goods and ethnic divisions. Working Paper No. 6009, NBER.

Beck, Nathaniel, Jonathan N. Katz, and Richard Tucker. 1998. Taking time seriously: Time-series-cross-section analysis with a binary dependent variable. *American Journal of Political Science* 42:1260–88.

Blomberg, S. Brock, and Gregory D. Hess. 2002. The temporal links between conflict and economic activity. *Journal of Conflict Resolution* 46:74–90.

Clarke, Kevin A. 2005. The phantom menace: Omitted variable bias in econometric research. *Conflict Management and Peace Science* 22:341-52.

Collier, Paul, Lani Elliott, Håvard Hegre, Anke Hoeffler, Marta Reynal-Querol, and Nicholas Sambanis. 2003. *Breaking the conflict trap: Civil War and development policy*. World Bank Policy Research Report. Oxford, UK: Oxford University Press/World Bank.

Collier, Paul, and Anke Hoeffler. 2004. Greed and grievance in civil war. *Oxford Economic Papers* 56:563–95.

Easterly, William, and R. Levine. 1997. Africa's growth tragedy: Policies and ethnic divisions. *Quarterly Journal of Economics* 112:1203–50.

Elbadawi, Ibrahim, and Håvard Hegre. 2004. Globalization, economic shocks, and armed conflict. Paper presented at the “Globalization, Territoriality and Conflict” conference, January 16-18, University of California, San Diego. http://folk.uio.no/hahegre/papers.

Elbadawi, Ibrahim, and Nicholas Sambanis. 2002. How much war will we see? Explaining the prevalence of civil war. *Journal of Conflict Resolution* 46:307–34.

Esty, Daniel C., Jack Goldstone, Ted Robert Gurr, Barbara Harff, Pamela T. Wurko, Alan N. Unger, and Robert S. Chen. 1998. *State failure task force report: Phase II findings*. McLean, VA: Science Applications International Corporation.

Esty, Daniel C., Jack A. Goldstone, Ted Robert Gurr, Pamela T. Surko, and Alan N. Unger. 1995. *Working papers state failure task force report*. McLean, VA: State Failure Project.

Fearon, James D. 1998. Commitment problems and the spread of ethnic conflict. In *The international spread of ethnic conflict*, edited by David A. Lake and Donald Rothchild, 107-26. Princeton, NJ: Princeton University Press.

———. 2003. Ethnic and cultural diversity by country. *Journal of Economic Growth* 8:195-222.

Fearon, James D., and David D. Laitin. 2003. Ethnicity, insurgency, and civil war. *American Political Science Review* 97:75–90.

Freedom House. 2002. *Annual surveys of freedom ratings 1972–73 to 2000–01*. http://freedomhouse.org/ratings.

Gates, Scott, Håvard Hegre, Mark P. Jones, and Håvard Strand. 2006. Institutional inconsistency and political instability: The duration of polities. *American Journal of Political Science 50 (4)*.

Gleditsch, Kristian Skrede. 2002. *All international politics is local: The diffusion of conflict, integration, and democratization*. Ann Arbor: University of Michigan Press.

Gleditsch, Nils Petter, 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.

Gurr, Ted Robert. 2000. Ethnic warfare on the wane. *Foreign Affairs* 79:52–64.

Gurr, Ted Robert, and Keith Jaggers. 2000. *Polity98 project: Regime characteristics, 1800–1998*. http://www.bsos.umd.edu/cidcm/polity.

Hechter, Michael. 2001. *Containing nationalism*. Oxford, UK: Oxford University Press.

Hegre, Håvard. 2003. Disentangling democracy and development as determinants of armed conflict. Paper presented at the 44th Annual Convention of the International Studies Association, February 26–March 1, Portland, OR. http://folk.uio.no/hahegre/Papers/DisentanglingWB.pdf.

Hegre, Håvard, Tanja Ellingsen, Scott Gates, and Nils Petter Gleditsch. 2001. Toward a democratic civil peace? Democracy, political change, and civil war, 1816–1992. *American Political Science Review* 95:33–48.

Hegre, Håvard, Nils Petter Gleditsch, and Ranveig Gissinger. 2003. Political institutions, globalization, and conflict. In *Globalization and armed conflict*, edited by Gerald Schneider, Katherine Barbieri, and Nils Petter Gleditsch, 271-5. Boulder, CO: Rowman & Littlefield.

Leamer, Edward E. 1985. Sensitivity analyses would help. *American Economic Review* 75:308–13.

Levine, Ross, and David Renelt. 1992. A sensitivity analysis of cross-country growth regressions. *American Economic Review* 82:942–63.

Lijphart, Arend. 1977. *Democracy in plural societies: A comparative exploration*. New Haven, CT: Yale University Press.

Marshall, Monty, and Keith Jaggers. 2000. *Polity IV project: Codebook and data files*. http://www.bsos.umd.edu/cidcm/inscr/polity.

Mauro, Paulo. 1995. Corruption and growth. *Quarterly Journal of Economics* 110:681–712.

Miguel, Edward, Shanker Satyanath, and Ernest Sergenti. 2004. Economic shocks and civil conflict: An instrumental variables approach. *Journal of Political Economy* 112:725–53.

Murdoch, James C., and Todd Sandler. 2002. Economic growth, civil wars, and spatial spillovers. *Journal of Conflict Resolution* 46 (1): 91–110.

Oneal, John R., and Bruce Russett. 2005. Rule of three, let it be? When more really is better. *Conflict Management and Peace Science* 21:293-310.

Przeworski, Adam, Michael E. Alvarez, José Antonio Cheibub, and Fernando Limongi. 2000. *Democracy and development: Political institutions and well-being in the world, 1950–1990*. Cambridge, UK: Cambridge University Press.

Raknerud, Arvid, and Håvard Hegre. 1997. The hazard of war: Reassessing the evidence for the democratic peace. *Journal of Peace Research* 34:385–404.

Reynal-Querol, Marta. 2002. Political systems, stability and civil wars. *Defence and Peace Economics* 13:465–83.

Ron, James. 2005. Paradigm in distress? Primary commodities and civil war [Introduction to special issue]. *Journal of Conflict Resolution* 49 (4): 443–50.

Ross, Michael. 2004. What do we know about natural resources and civil war? *Journal of Peace Research* 41:337–56.

Sala-i-Martin, Xavier X. 1997. I just ran two million regressions. *American Economic Review* 87:178–83.

Sambanis, Nicholas. 2000. Partition as a solution to ethnic war: An empirical critique of the theoretical literature. *World Politics* 52:437–83.

———. 2001. Do ethnic and non-ethnic civil wars have the same causes? A theoretical and empirical inquiry (Part 1). *Journal of Conflict Resolution* 45:259–82.

———. 2002. A review of recent advances and future directions in the quantitative literature on civil wars. *Defence and Peace Economics* 13:215–43.

———. 2004a. What is a civil war? Conceptual and empirical complexities of an operational definition. *Journal of Conflict Resolution* 48:814–58.

———. 2004b. Expanding economic models of civil war using case studies. *Perspectives on Politics* 2:259-80.

———. 2005. Poverty and the organization of political violence: A review and some conjectures. In *Brookings Trade Forum 2004*, edited by Susan M. Collins and Carol Graham, 165-222. Washington, D.C.: Brookings Institution.

Taylor, Charles Lewis, and Michael C. Hudson. 1972. *World handbook of political and social indicators*. 3rd ed., vol. 1. New Haven, CT: Yale University Press.

Vanhanen, Tatu. 1999. Domestic ethnic conflict and ethnic nepotism: A comparative analysis. *Journal of Peace Research* 36:55–73.

———. 2000. A dataset on measures of democratization. *Journal of Peace Research* 37:251–65.

World Bank. 2000. *World development indicators, 2000.* Washington, D.C.: International Bank for Reconstruction and Development/World Bank.