Sensitivity Analysis of the Empirical Literature on Civil War Onset

Håvard Hegre (Centre for the Study of Civil War, PRIO)
Nicholas Sambanis (Yale University)*

Paper presented to the 46th annual meeting of the
International Studies Association, Honolulu, HI, 1-5 March 2005

February 23, 2005

Abstract

This paper applies a methodology for organized specification tests to establish the robustness of empirical findings in the quantitative literature on civil war onset. A review of the literature reveals that several empirical results are not robust or not replicable across studies, which creates uncertainty about the validity of inferences drawn from econometric studies of civil war onset. Given that most studies use different definitions of civil war and analyze datasets with different country-year coverage, readers cannot distinguish between variation in empirical results that is due to those factors and variation that is due to the measurement of independent variables or model specification when the same dependent variable and same country-years are analyzed. We reduce this uncertainty in this paper by systematically exploring how sensitive 88 variables used in the literature are to variations in the set of control variables.

* Contact: havard.hegre@stv.uio.no; Nicholas.Sambanis@yale.edu. We thank Scott Bennett, Halvard Buhaug, Scott Gates, Nils Petter Gleditsch, and Håvard Strand for helpful suggestions, and Gudrun Østby for excellent research assistance. An earlier version of this paper was presented at the Peace Science Society International Annual Meeting, Arizona, November 1–3, 2002. The research has been supported by the Polarization and Conflict Project CIT-2-CT-2004-506084 funded by the European Commission-DG Research Sixth Framework Programme. It reflects only the authors' views and the *Community* is not liable for any use that may be made of the information contained therein.

Sensitivity Analysis of the Empirical Literature on Civil War Onset

Abstract

This paper applies a methodology for organized specification tests to establish the robustness of empirical findings in the quantitative literature on civil war onset. A review of the literature reveals that several empirical results are not robust or not replicable across studies, which creates uncertainty about the validity of inferences drawn from econometric studies of civil war onset. Given that most studies use different definitions of civil war and analyze datasets with different country-year coverage, readers cannot distinguish between variation in empirical results that is due to those factors and variation that is due to the measurement of independent variables or model specification when the same dependent variable and same country-years are analyzed. We reduce this uncertainty in this paper by systematically exploring how sensitive 88 variables used in the literature are to variations in the set of control variables.

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. One reason for this may be that it is not clear which is the best way to test the fragility of empirical results.[1] Most such efforts take place in an *ad hoc* manner: researchers identify a set of competing explanations for their research questions' results and try to establish if their own results hold under different sets of conditions, controlling for different explanations. Testing different model specifications allows us to discriminate between rival theories and increases our knowledge about the relationship between the dependent variable (DV) and a set of independent variables (IVs). However, *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 clearly establish if small changes to model specification influence the robustness of inferences that are based on empirical results, Leamer (1985) 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

---

[1] 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.

only if the neighborhood of assumptions is wide enough to be credible and the corresponding interval of inferences is narrow enough to be useful" (Leamer 1985, 308). In this paper, we follow the line of inquiry suggested by Leamer (1985) and explore two possible sources of variation in empirical results in the quantitative literature on civil war. First, we test the sensitivity of important results to small changes in the model specification; and secondly, we test the robustness of key empirical findings to changes in the variable used to operationalize/measure a given concept. Our analysis allows us to establish an empirical basis on which to accept or reject theoretical propositions on the determinants of civil war onset.

The paper is organized in four sections beyond the introduction. Section two presents the methodology of sensitivity analysis. Section three provides a brief overview of the range of variation in empirical results in the literature on civil war onset. The large number of apparently fragile results motivates our analysis. Section four presents our results and identifies the most robust and most fragile inferences in the literature on civil war onset. Section five concludes with a summary of our main results.

2. Methodology of Sensitivity Analysis

Leamer referred to the largest possible set of inferences that can be drawn from a given dataset 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 dataset. 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 EBA to conduct a sensitivity analysis of cross-country regressions of economic growth, 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 form:

$Y = \beta_t \mathbf{I} + \beta_\mu M + \beta_z \mathbf{Z} + u,$ 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 of which behavior we are interested in observing as the model specification changes), 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 **I**-variables and the focus variable and then estimate regressions for "all possible combinations of up to three **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 **Z**-variables," and "excluding variables that, a priori, might measure the same phenomenon" (ibid., 944–45). This specification design minimizes the risk of underspecified models while also minimizing computer power

needed to estimate the models and reducing problems associated with multicollinearity. The extreme bounds on the coefficient β_μ denote the confidence 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. β_μ is considered robust if it "remains significant and of the same sign at the extreme bounds" (ibid., 944).

This EBA analysis leads Levine and Renelt (1992, 959) to conclude that "very few economic variables are robustly correlated with cross-country growth rates." This extreme result may suggest that their analysis sets too high a hurdle to establish 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." We agree that we need a less "extreme" test of robustness and structure our attempt to establish the robustness of empirical results on civil war onset by applying Sala-i-Martin's (1997) approach of looking at the entire distribution of parameter estimates to determine the level of confidence in each of the explanatory variables.

Sala-i-Martin estimates *M* models of the form:

$$\boldsymbol{\gamma}_j = \boldsymbol{\alpha}_j + \boldsymbol{\beta}_{yj}\mathbf{y} + \beta_{zj} z_j + \boldsymbol{\beta}_{xj}\mathbf{x}_j + \boldsymbol{\varepsilon} \quad (1)$$

where $\boldsymbol{\gamma}$ is the dependent variable and the subscript $j \in [1,2,...,M]$ 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 *x* is a vector of up to three variables taken from a pool $\aleph$ of available variables. He then uses these estimates for 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 these 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 $\omega_{zj} = \frac{L_{zi}}{\sum_{i=1}^{M} L_{zi}}$ (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 regressors. Similarly, Sala-i-Martin 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 $\hat{\sigma}_z^2$ as 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 (1997) computes the aggregate CDF(0) from the individual CDF(0)'s,

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

We make a number of adaptations to Sala-i-Martin's 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 γ_j is the logit-link function of the probability of civil war outbreak. Although we have panel data, we follow most papers in the literature in treating the data as cross-sectional, using pooled logit or pooled probit estimators. Temporal dependence is accounted for by using a function of 'time in

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 "core" variables henceforth. A recent review of the quantitative literature on civil war (Sambanis 2002) has 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 (which we model as a decay function of time at peace, (*pt8*), and the natural log of per capita GDP (in constant dollars) (*lngdp*). Keeping these three variables constant, we estimate a set of models for each of the variables in the pool $\aleph$ of relevant variables. We discuss our selection of variables to include in the pool in the next section.

Sala-i-Martin estimates each possible combination of three **x**-variables for each *z*-variable. We deviate somewhat from Sala-i-Martin in our treatment of the **x** variables. Our data contain several different measures for the same concept variable (i.e. the theoretically significant variable). For instance, we use variables from the Polity project, Przeworski et al. (2000), and Freedom House as different measures of the concept "level of democracy". We restrict the **x**-variables in the regressions so that they measure three different *concept* variables, not simply different operationalizations of the same concept. Including as a control variable an alternative operationalization of the same concept as the focus variable is likely to diminish the effect of the focus variable (because of collinearity). This would 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 section 3). Allowing more than one operationalization for each concept then would tend to hurt variables belonging to concepts with many operationalizations. (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 racial, linguistic, and religious heterogeneity index, or *ehet*, Fearon's *ef* index, as well as the more widely used ethnolinguistic fractionalization index, *elfo*). Hence, we only include combinations of three control (**x**) variables that do not belong to the same concept as the focus variable or that of the other two control variables in the model.

Another difference from Sala-i-Martin's procedure is that we have to allow for the fact that the variables in our dataset are missing data for a varying number of observations. 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 in the estimation. Hence, we calculate weights based on the likelihood ratio index (LRI), which is analogous to the R^2. It is 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.[2]

Finally, a few important variables in the civil war literature cannot be represented in a single *z* term, such as democracy–partial democracy–autocracy, or primary commodity dependence and its square term. We enter each of these as individual terms, and report results for

[2] 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 non-randomly missing data in other studies. Our approach should be less vulnerable to this problem.

their robustness on their own. It is not possible to evaluate their joint significance or joint robustness with the method employed here. [3]

That we perform extensive specification tests implies that we do not know the true model. Our core variables are robust in the theoretical and empirical literature: GDP per capita, population size, and time since last war. In that sense, our model contains the core of the true model, but is certainly under-specified. Beyond these three variables, there is little consensus on what variables to include. The method applied here can be considered as an empirical approach to developing a well-specified regression equation for civil war onset by testing the fit of several theoretically relevant variables.

Another concern is potential endogeneity of several variables in our dataset. Trying to deal with this problem in the context of our paper is not feasible as it requires a change in the estimator and use of instrumental variables for some – but not all— models estimated. Moreover, if the exogeneity assumption could not be rejected and we nevertheless used instrumental variable estimation, this would produce less efficient parameter estimates. To avoid efficiency losses, the exogeneity assumptions should have been tested before using instrumental variables regression and this further complicates the mechanics of the methodology for specification tests. Since very few of the papers in the literature on civil war deal with the issue

[3] 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.

of endogeneity,[4] we ignore it here and simply try to reduce the risk by lagging independent variables.[5]

There is also some concern with the consequences of missing observations. Many of the variables in our list have quite a few 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

---

[4] Exceptions include Sambanis (2000; 2001), Blomberg and Hess (2002), and Miguel et al (2004).

[5] 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. In addition, some combinations of variables may present more problems than others. For example, the inclusion of both economic growth and trade may be violations of the exogeneity assumption, since both trade and growth is affected by conflict or the expectation of such. Variables that measure the degree of militarization (e.g. military expenditures; size of the government army) may also be endogenous to the level of political violence at the time of war onset or the expectation of war onset.

changes in the specification or to changes in the sample. We try to reduce this problem by interpolating incomplete series and, in few cases, imputing missing values (see notes in subsequent sections). 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. Any systematic measurement error (including missing values) in the independent variables would reduce the efficiency of parameter estimates, making it harder to identify robustly significant coefficients.

3. Variation in Empirical Results on Civil War Onset

A brief overview of the theory and evidence on civil war onset explains our specification of our core model. Income per capita measures the economic opportunity cost of the war, some aspects of the state's capacity, and possibly other economic components of the decision to rebel (Collier and Hoeffler 2004). A constant per-capita propensity to initiate armed resistance implies that civil wars defined in terms of a fixed number of battle deaths are more likely in populous countries. Beyond this, the relationship between population size and the risk of civil war is relatively under-theorized (see Sambanis 2002). We include it in the core model since it is always significant in civil war models. Time at peace since the last civil war should decrease the risk of a new war, as conflict-specific capital remains unutilized and peace-specific capital is accumulated (though there is some disagreement in the literature on the path-dependence of war and peace).

Empirical support for most other explanatory variables is mixed. For example, the hypothesis that rough terrain is significant for the onset of civil war is one of the critical hypotheses in Fearon and Laitin (1999; 2003), but Sambanis (2004a) and Collier and Hoeffler (2004) do not find a significant association between mountain or forest coverage and civil war

onset. These studies utilize the same measure of rough terrain but Fearon and Laitin (2003) impute missing values using other geographic data.[6]

Our approach was to create "concept categories" that correspond to theoretically significant variables that span the range of theorizing in the literature. In each concept category, we include several empirical proxies that purport to measure the theoretical variables. Our categories are given in Table 1, along with a description of variables in each category and the sources for each variable.

[Insert Table 1 about here]

In sum, we have identified through a review of the literature on civil war several focus variables that are theoretically significant in explaining war onset. Table 1 lists the concept categories and the variables that we use under each concept and presents the main sources that first introduced the variable in statistical analysis.

A full discussion of the various theories and empirical results would take us too far, given the large number of variables considered in this paper. We refer readers to cited sources for more information, and only briefly address some of the key debates to motivate the focus of our paper and provide a background to some of our variables.

A heavily debated topic is the effect of ethnic difference on the risk of civil war (concept 1). The most commonly used measure is the now well-known index of ethno-linguistic

[6] Sambanis (2004a) however, uses the Fearon and Laitin measure and finds that it is not robust to small changes in the measurement of the *dependent* variable.

fractionalization (*elfo*).[7] Most studies of civil war find it not to be significant, though 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.[8] *Elfo* ranges from 0 (ethnic homogeneity) to 100 (extreme ethnic heterogeneity) and measures the probability that two randomly selected individuals belong to different ethno-linguistic groups.[9] We also include the square term as in Collier & Hoeffler (2004). We use *elfo* as the best available cross-national measure of ethnic distance. We assume that *elfo* is exogenous to civil war.[10] *Elfo*, like all

---

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

[8] 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 2004b).

[9] 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.

[10] Certainly, ethnic identity, or at least the salience of ethnic identity, is shaped by social processes, including war. However, these are long-term process, 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 dataset, which further justifies the exogeneity assumption.

currently available measures of social identity, is time-invariant and 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 sub-indices of racial division, linguistic division, and religious division. Vanhanen measured the size of the largest group as a percentage of total population and summed across the inverse of these ratios for all three categories (largest racial, linguistic, religious group). [11] We also include some variables coding the number of ethnic, linguistic, or religious groups from Fearon and Laitin (2003), who coded them from the *Encyclopedia Britannica*.

Several authors argue that polarization between a small number of groups or the dominance of one large group over others is more likely to lead to violent conflict than mere fragmentation (e.g., Collier & Hoeffler, 2004). This is our Concept 2 variable. We include two measures of it: Fearon & Laitin's measure of the share of the population in the second largest group and Collier & Hoeffler's dichotomous 'Ethnic dominance' measure.

The significance of political institutions (concept 3) is heavily debated. Contrary to recent studies that point to a significant negative association between civil war and the level of democracy (Esty et al. 1995; 1998; Gurr 2000), most recent studies find the level of democracy to

[11] Note that Vanhanen's computations may have an error: Ehet's lowest value is 0, indicating perfect homogeneity. But, if *ehet* is computed as the sum of the inverse ratio of the largest group over the population, then its smallest value should be 3, not 0, as a country with a single group comprising 100% of the population in all three cleavages should be coded as 1+1+1 = 3.

be non-significant (Collier and Hoeffler 2004; Fearon and Laitin 2003). We try several measures of level 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 & 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 lead 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 just onset (Collier and Sambanis 2002). 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). Many civil wars are wars over self-determination. Greater decentralization may reduce the incentives for violence, according to the logic of consociationalism (Lijphart 1977; Hechter 2001).

Another debated question is the regional dimension of civil wars. We group these variables into three concepts: The political economy of the neighborhood (concept 8), comprising variables on the average income and democracy of the neighboring countries; indicator variables for the regions themselves (concept 9), and variables denoting the extent of war in the neighborhood (concept 10). 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, though Hegre et al.

(2001) and Fearon and Laitin (2003) find that these neighborhood spillovers are not significant. Second, countries in "bad" neighborhoods, i.e. with mostly poor, autocratic regimes in the neighborhood, will be more likely to experience a civil war (Sambanis 2001; Gleditsch 2002).

The role of economic factors has been the focus of several recent studies. The rate of economic growth (concept 11) has been shown to reduce civil war risk (Collier and Hoeffler 2004; Hegre 2003). Results are ambiguous, maybe since the variable used in most datasets 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, in particular trade – and civil war (Esty et al 1995; Hegre, Gissinger, and Gleditsch 2003; Elbadawi and Hegre 2004).

Another topic is social welfare (concept 13), such as the effect of infant mortality (Esty et al 1998) and education (Collier and Hoeffler 2004) on civil war over and beyond the importance of average income levels. Again the results of these studies are inconclusive.

Finally, 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). We include several measures of the relative importance of natural resources versus manufactures.

Related to economic factors are variables describing geographical aspects of the country (concept 15), since this may affect a government's ability to maintain a monopoly on using military force. We include two of Fearon & Laitin's (2003) terrain variables and a measure of population density.

The degree of militarization of the country and the government's military strength 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 add a concept

category with variables measuring military expenditures and the size of the government army (concept 16).

The risk of civil war may have been different over time (concept 17). The end of the Cold War is sometimes thought to have unleashed ethnic conflict though the results on this variable are inconclusive (Sambanis 2001; Fearon and Laitin 2003). We control for this and also use decade dummies to control for time trends.

Finally, we include information on countries' colonial war history in concept 18.

Using the different operationalizations of each concept category we can gauge how much these differences in measurement and definition of the independent variables influence the empirical results on civil war.

4. The Dependent Variable

Differences in empirical results on civil war onset may also be due to differences in the coding of civil war, as civil war lists used in the literature differ markedly (Sambanis 2004a). The partial correlation of war onset between the recent version of the Correlates of War Project's civil war onset list (Sarkees and Singer 2001) and Collier and Hoeffler's list (2004) is 83%. Fearon and Laitin's (2003) list correlates by 58% with Singer and Small (1994) and Licklider's (1995) list correlates 57% with Singer and Small (1994) and 58% with Fearon and Laitin's list. From 1960-1993, Fearon and Laitin code 79 war starts (663 country-years at war), Singer and Small code 61 war starts (353 country-years at war) and Patrick Regan (1996) codes 116 war starts (764

country-years at war). Some of these differences are due to coding error, but most are due to different definitions of civil war and divergent coding rules.[12]

We code a civil war if: (a) the war caused more than 1,000 deaths in total and in at least a 3-year period; (b) it challenged the sovereignty of an internationally recognized state; (c) fighting took place within the territory of that state; (d) the state was a principal combatant or was represented by a communal or military group claiming the state in cases of state collapse; (e) the rebels were able to mount an organized military opposition to the state with effective resistance causing the stronger side to suffer at least 100 deaths.[13]

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 is to code periods of ongoing war as 0's (except the year of onset), but proper estimation of such a model requires that we add a lagged term denoting war in the previous period interacted with all right-hand-side variables to capture any differences in the ways these variables influence war onset while another war is ongoing. Applying our methodology to both versions of the dependent variable would have made the analysis intractable, so we chose the most commonly used coding rule.

---

[12] The State Failure Task Force, for example, uses a minimum threshold of 100 fatalities per year to code an ethnic conflict. Others require 1,000 deaths per year; or they use a cumulative (not annual) death threshold.

[13] See Sambanis (2004a) for more details on the coding of civil wars.

Two other sources of differences in empirical results in the literature are neutralized in our dataset. Different analyses often cover different time periods and countries or they use different estimation techniques. However, most of the major datasets focus on either the 1960–99 period or the 1945–99 period and employ a pooled logit or probit model to analyze the data (some cluster same-country observations and add time-dependence controls). We use the same estimation method and analyze a single dataset that includes 172 countries observed annually from 1960–99.

We now turn to a discussion of the results of our sensitivity analysis.

5. Results of Sensitivity Analysis of Models of Civil War Onset

Our sensitivity analysis analyzes the 88 variables in Table 1, organized along the 18 concept categories described above, drawing data from Sambanis (2001; 2004a), Hegre et al. (2001), Fearon and Laitin (2003), Reynal-Querol (2002), and Collier and Hoeffler (2004). All variables were standardized to make the estimates more easily comparable.[14] Several variables were log-transformed to minimize the effect of outlying observations.

We ran 6.4 million logistic regressions, but only report statistics based on a random sample of 10% of our regression results (640,182 regressions). Table 2 reports the results for the three core variables.

[Insert Table 2 here]

[14] We used an ordinary standardizing procedure: first we substracted the mean for the variable from each value, and then divided all values by the variable's standard deviation.

The standardized population variable obtained a mean parameter estimate of .279 over these estimations, and a mean estimate of .089 of the standard error. On average, this corresponds to a *t*-value of -3.13 and an 'average *p*-value' of 0.00086. Recall that this statistic is based on the assumption that the β_{zj} are normally distributed, and corresponds to a one-sided test of statistical significance. Figure 1 shows how the *t*-values are distributed over the 640,182 estimations. We have added vertical lines at t=-1.65 and t=1.65. The bars inside these lines indicate the frequency of non-significant parameter estimates. The distribution is slightly right-skewed. If we relax the assumption of normal distribution of estimates, the p-value is 0.0051 (the right-most column in Table 2). Thus, population size was significant at the .05 level (t>1.96) in 96% of the models.

[Insert Figure 1 here]

The GDP per capita variable also is robustly significant, with a mean *t*-value of –2.69 and an average *p*-value of 0.0036 under the assumption of normal distribution. Figure 2 shows that the estimates have a much more skewed distribution than the population variable. The mode of the distribution is below –3, but in as many as 24% of the models the variable is not significant at the 5% level. Relaxing the distribution assumptions, the average *p*-value is 0.029.

[Insert Figure 2 here]

The function of time since previous war is less robust than the other two core variables. The average *t*-value is 0.90, corresponding to an average *p*-value of 0.18 for a one-sided test. Figure 3 shows the distribution of estimated t-values.

[Insert Figure 3 here]

Table 3 lists the 18 most robust variables, ordered by the 'average *p*-value' under the normal distribution assumption. The third column assigns each variable to each concept category. The fourth column reports the weighted average of parameter estimates and the fifth the weighted average for the standard error of the parameter estimates. The average *p*-value is reported in the sixth column and in the seventh column we report the p-value after relaxing the normality assumption. (Note that the robustness of some of these variables may be over-estimated, since we have not taken endogeneity into account.) More detailed results for all 88 variables are given in the Appendix, Table 1.

[Insert Table 3 here]

We have identified robust operationalizations for 12 concept categories in Table 3. 14 variables for 11 concepts have average *p*-values less than 0.05 under the least restrictive assumption. Three political instability variables are among the top 4 most robust variables, and a fourth is found further down the list. The 'inst3' variable reflects whether there was a change in the country's Polity score within the three years preceding 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 raises concerns about endogeneity: 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. Finally, the concept 5 variable 'nwstate' identifying the two first years of independence of countries that became independent within the time period we analyze is also among the robust variables.[15]

The third-most robust variable is one of the 'Level of democracy' variables (concept 3). 'Parreg' is one of the sub-indicators forming the Polity IV Project's Democracy index, and codes 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 non-regulated participation. Lagging the variable by one year may be insufficient to completely remove any endogeneity bias in this case. Other 'level of democracy' variables are not robust. Several 'Inconsistency of political institutions' variables are quite robust, however, such as 'partfree', 'anoc', and 'pol4sq'.

Military personnel (milper) is negatively and robustly correlated with civil war onset, as hypothesized. 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 two 'neighborhood war' variables have robustly positive estimates, lending support to hypotheses regarding the significance of diffusion and contagion effects in civil war (though we

[15] This variable is only included in Fearon and Laitin's (2003) model. However, it is unclear what this variable measures that is also not measured by instability, anocracy, ethnic fractionalization and other variables in the model (see Sambanis 2004a for a discussion of this issue).

cannot distinguish among these different mechanisms through which civil war influences are transmitted across countries in the neighborhood.

The rough terrain variable ('lmtnest') variable is robustly positive, which 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.

Economic growth ('gdpgrowth') is robustly negative, lending support to Collier and Hoeffler's economic "opportunity cost" of civil war, and the regional democracy level 'regd4_alt' is also robustly negative, suggesting that living in a "bad" neighborhood may be more important than a low "in-country" democracy level (Sambanis 2001).

In the bottom half of the list we find 'Presidential system' with a moderately robust negative estimate, and 'oil exports/GDP' with a positive one. Only one of the ethnic variables – ethnic dominance – is robust. Finally, we found some region and period controls to be robustly associated with the risk of civil war outbreak.

Figures 4-9 display the distribution of estimated *t*-values for six concept variables that are heavily debated in the literature: (1) Ethnic fragmentation; (2) Ethnic dominance/polarization; (3) Level of democracy; (4) Institutional inconsistency; (5) Political instability; and (6) Natural Resources. Vertical lines at t=-1.65 and t=1.65 represent the boundaries of a fairly lenient definition of statistical significance.

Figure 4 shows that in the vast majority of the estimations, the ethnic fractionalization variables were not statistically significant. Figure 5 shows the distribution of estimated t-values for the ethnic dominance and the size of the second-largest group variables. The ethnic dominance

variable was positive and significant in many specifications, but not significant in a considerable fraction of them.

Figure 6 shows that most of the 'Level of democracy' variables are non-significant in most specifications. As discussed above, the 'parreg' variable is an exception. Figure 7 shows that the 'anoc', 'anocracy', and 'partfree' variables are often significant. Figure 8 demonstrates that the 'proxregc' and the three 'inst' variables differ from the other measures of political instability.

Finally, Figure 9 shows the distribution of estimated *t*-values for the natural resource variables. Only the 'oil' variable may be considered marginally robust among these.

[Insert Figures 4-9 here]

6. Conclusion

Using a methodology due to Leamer (1985) and Sala-i-Martin (1997), we have evaluated the fragility of a number of empirical results in the burgeoning civil wars literature. We have found some of those results to be highly sensitive to small changes in the model specification. These specification changes amounted to estimating systematically models including all possible (according to a set of restricting rules) seven-variable combinations of variables that might be relevant to explaining the onset of civil war. Our analysis suggests that some empirical results in the literature "are not worth taking seriously."

However, some variables perform remarkably well under a wide set of model specifications. We have confirmed that a large population and low average income increase the risk of civil war, consistent with prevailing theories of civil war. While controlling for these two variables, we also find a number of other robust relationships in the data. We find that civil wars

are more likely to occur in countries with recent political instability and inconsistent democratic institutions, countries with small militaries, countries with rough terrain, located in war-prone neighborhoods and undemocratic regions, and countries with low rates of economic growth. Other variables fail to – on average – be significant at the 0.05 level for one-sided tests.

This analysis helps sort out some of the discrepancies and disagreements between the results of different research projects on civil war. This sort of sensitivity analysis should be carried out in all empirical literatures and sub-fields in political science. While it might not be feasible for every paper that tries to test one or more hypotheses to conduct as systematic an empirical investigation as the one we have presented here, our results indicate that analysts should worry about the fragility of their inferences and (a) substitute closely related, but differently measured variables in their regressions and (b) add control variables to make sure the results on focus variables are robust before making inferences from their empirical results. In addition, other sources of variability in empirical estimates should be considered. Our sensitivity analysis only explores robustness to changes in the set of control variables. But researchers could (should) also consider robustness of empirical results to changes in period and country coverage, sampling and estimation techniques, and different operationalizations of the dependent variable.

Finally –and most importantly—our analysis also has something to say about theory. We find clear evidence of neighborhood effects of civil war. Yet, we currently do not understand those effects well and we do not know which are the main mechanisms through which civil wars in one country affect civil war risk in other countries in the region. These risks could be due to demonstration effects or they may be due to contagion, as rebels spill over borders and incite rebellion by their co-ethnics in neighboring countries. Alternatively, one might think of some civil wars as regional wars between the same set of groups in regions with disputed borders.

While our analysis does not explain how and why neighborhood variables are significant, it does point to the need for more theorizing on this question.

We also offer empirical validation of "economic" theories of civil war, though not with respect to those theories' predictions regarding the role of natural resources. We only find a weak association between civil war and oil export dependence, while other commonly used measures of resource-dependence are not significant. But the key variables in economic theories –per capita income and, to a lesser extent, the rate of growth of income—are very robust to specification changes.

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, undemocratic neighborhoods—are very important and increase the risk of civil war. 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). Our results on neighborhood variables and political variables may have to be considered together: "bad" neighborhoods (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 result 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 on the spatial dimensions of civil war.

References

Alesina, Alberto, Reza Baqir, and William Easterly. 1997. "Public Goods and Ethnic Divisions." NBER: Working Paper Number 6009

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(4):1260–1288.

Blomberg, S. Brock, and Gregory D. Hess. 2002. "The Temporal Links Between Conflict and Economic Activity." *Journal of Conflict Resolution* 46(1):74–90.

Collier, Paul, and Anke Hoeffler. 2004. "Greed and Grievance in Civil War." *Oxford Economic Papers* 56(4):563–595.

Collier, Paul, and Nicholas Sambanis. 2002. "Understanding Civil War: A New Agenda [Introduction to special issue on Understanding Civil War]." *Journal of Conflict Resolution* 46(1):3–12.

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: Oxford University Press and Washington, DC: The World Bank.

Cooley, Thomas, and Stephen LeRoy. 1981. "Identification and Estimation of Money Demand." *American Economic Review* 71(5):825–844.

Easterly, William, and R. Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics* 112(4):1203–1250.

Elbadawi, Ibrahim, and Håvard Hegre. 2004. "Globalization, Economic Shocks, and Armed Conflict." Paper presented to the 'Globalization, Territoriality and Conflict' conference, University of California, San Diego, 16–18 January (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(3):307–334.

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.

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.

Fearon, James D., and David D. Laitin. 1999. "Weak States, Rough Terrain, and Large-Scale Ethnic Violence." Paper presented at the Annual Meetings of the American Political Science Association, Atlanta, GA, 2–5 September.

Fearon, James D., and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97(1):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.2001. "Institutional Inconsistency and Political Instability: Persistence and Change in Political Systems Revisited, 1800–1998.",Paper presented to the International Studies Association Annual Meeting, Chicago, IL, 20–24 February.

Gleditsch, Kristian Skrede. 2002. *All International Politics is Local: The Diffusion of Conflict, Integration, and Democratization*. Ann Arbor, MI: University of Michigan Press.

Gurr, Ted Robert. 2000. "Ethnic Warfare on the Wane." *Foreign Affairs* 79(3):52–64.

Gurr, Ted Robert, and Keith Jaggers. 2000. *Polity98 Project: Regime Characteristics, 1800–1998*. University of Maryland.

Hechter, Michael. 2001. *Containing Nationalism*. Oxford: Oxford University Press.

Hegre, Håvard. 2003. "Disentangling Democracy and Development as Determinants of Armed Conflict." Paper presented to the 44th Annual Convention of the International Studies Association, Portland, OR, 26 February–1 March.

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(1):33–48.

Hegre, Håvard, Nils Petter Gleditsch, and Ranveig Gissinger. 2003. "Political Institutions, Globalization, and Conflict.", In *Globalization and Armed Conflict*, ed. Gerald Schneider, Katherine Barbieri, and Nils Petter Gleditsch. Boulder, CO: Rowman & Littlefield (251–275).

Leamer, Edward E. 1985. "Sensitivity Analyses Would Help." *American Economic Review* 75(3):308–313.

Levine, Ross, and David Renelt. 1992. "A Sensitivity Analysis of Cross-Country Growth Regressions." *American Economic Review* 82(4):942–963.

Licklider, Roy. 1995. "The Consequences of Negotiated Settlement in Civil Wars, 1945–1993." *American Political Science Review* 89(3):681–690.

Lijphart, Arend.1977. *Democracy in Plural Societies: A Comparative Exploration*. New Haven: Yale University Press.

Marshall, Monty, and Keith Jaggers. 2000. "Polity IV Project. Codebook and Data Files" (www.bsos.umd.edu/cidcm/inscr/polity).

Mauro, Paulo. 1995. "Corruption and Growth." *Quarterly Journal of Economics* 110(3):681–712.

Miguel, Edward; Shanker Satyanath, and Ernest Sergenti, 2004. "Economic Shocks and Civil Conflit: An Instrumental Variables Approach." *Journal of Political Economy* 112(4):725–753.

Murdoch, James C., and Todd Sandler. 2002a. "Economic Growth, Civil Wars, and Spatial Spillovers." *Journal of Conflict Resolution* 46(1):91–110.

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(4):385–404.

Regan, P. 1996. "Conditions for Successful Third Party Interventions," *Journal of Conflict Resolution* 40 (1): 336-359

Reynal-Querol, Marta. 2002. "Political Systems, Stability and Civil Wars." *Defence and Peace Economics* 13(6):465–483.

Ross, Michael. 2004. "What Do We Know About Natural Resources and Civil War?" *Journal of Peace Research* 41(3): 337–356.

Sala-i-Martin, Xavier X. 1997. "I Just Ran Two Million Regressions." *American Economic Review*, 87(2):178–183.

Sambanis, Nicholas. 2000. "Partition as a Solution to Ethnic War: An Empirical Critique of the Theoretical Literature.", *World Politics* 52(4):437– 483.

Sambanis, Nicholas. 2001. "Do Ethnic and Non-Ethnic Civil Wars Have the Same Causes? A Theoretical and Empirical Inquiry (Part 1)." *Journal of Conflict Resolution*, 45(3):259–282.

Sambanis, Nicholas. 2002. "A Review of Recent Advances and Future Directions in the Quantitative Literature on Civil Wars." *Defence and Peace Economics* 13(3):215–243.

Sambanis, Nicholas. 2004a. "What Is A Civil War? Conceptual and Empirical Complexities of an Operational Definition." *Journal of Conflict Resolution* 48(6): 814–858.

Sambanis, Nicholas. 2004b. "Poverty and the Organization of Political Violence: A Review and Some Conjectures." Brookings Trade Forum 2004: Globalization, Poverty, and Inequality. Washington, DC: Brookings Institution, pp. 165-211..

Sarkees, Meredith Reid, and J. David Singer. 2001. "The Correlates of War Dataset: The Totality of War." Paper presented at the Annual Convention of the International Studies Association, Chicago, IL, 20–24 February.

Singer, J. David, and Melwin, Small. 1994. *Correlates of War Project: International and Civil War Data, 1816–1992*. Ann Arbour, MI: ICPSR.

Taylor, Charles Lewis, and Michael C.Hudson. 1972. *World Handbook of Political and Social Indicators*. 3rd ed, vol 1. New Haven: Yale University Press.

Vanhanen, Tatu. 1999. "Domestic Ethnic Conflict and Ethnic Nepotism: A Comparative Analysis." *Journal of Peace Research* 36(1):55–73.

Vanhanen, Tatu. 2000. "A Dataset on Measures of Democratization." *Journal of Peace Research* 37(2):251–265.

World Bank.2000. *World Development Indicators, 2000.* Washington, DC: International Bank for Reconstruction and Development/The World Bank..

Table 1 – Concept Categories and Variable Names

	Variable Name	Description	Source
Dependent variable	warstns	War onset	Sambanis (2004a)
Core variables	ln_popns	Population, log-transformed	Sambanis (2004a)
	ln_gdpen	GDP per capita, log-transformed	Fearon & Laitin (2003)
	pt8	2^(-years in peace/8)	Authors' coding, based on Sambanis (2004a)

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 & Laitin (2003)
	1	ef2	Ef squared	Fearon & Laitin (2003)
	1	ehet	Ethnic heterogeneity index	Vanhanen (1999)
	1	elfo	Ethnolinguistic diversity	Collier & Hoeffler (2004)
	1	elfo2	Ethnolinguistic diversity, squared	Collier & Hoeffler (2004)
	1	numlang	Number of languages in Ethnologue	Fearon & Laitin (2003)
	1	plural	Share of largest ethnic group	Fearon & Laitin (2003)
	1	plurrel	Size of largest confession	Fearon & Laitin (2003)
	1	relfrac	Religious fractionalization	Fearon & Laitin (2003)
Ethnic dominance/ polarization	2	etdo4590	Ethnic dominance measure	Collier & Hoeffler (2004)
	2	second	Percent population in second largest group	Fearon & Laitin (2003)
Level of democracy	3	auto4	Autocracy index from Polity IV	Marshall & Jaggers (2000)
	3	dem	Dummy: 1 for democracies and 0 for autocracies	Marshall & Jaggers (2000)
	3	dem4	Democracy index from Polity IV	Marshall & Jaggers (2000)
	3	exrec	Executive recruitment concept variable; Polity IV	Marshall & Jaggers (2000)
	3	mirps2	Consistent autocracy	Gates et al.. (2001)
	3	mirps3	Consistent democracy	Gates et al.. (2001)
	3	parcomp	Competitiveness of participation; non-elites; Polity IV	Marshall & Jaggers (2000)
	3	parreg	Regulation of participation; Polity IV	Marshall & Jaggers (2000)
	3	part	ln(share of population voting x opposition's share of votes cast)	Vanhanen (2000)
	3	pol4	Polity index; Polity IV	Marshall & Jaggers (2000)
	3	pol4m	Polity Index; Polity IV; 77 & 88 coded=0	Marshall & Jaggers (2000)
	3	polcomp	Political competition: concept variable; Polity IV	Marshall & 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. (2001)
	3	xconst	Executive constraints - operational independence of CE; Polity IV	Marshall & Jaggers (2000)
Inconsistency of political institutions	4	anoc	Dummy: Anocracy=1	Fearon & Laitin (2003)
	4	mirps0	Inconsistent polity (semi-democracy)	Gates et al. (2001)
	4	mirps1	Caesaristic polity	Gates et al. (2001)
	4	partfree	Partially free polity	Freedom House (2002)
	4	pol4sq	Pol4 squared	Marshall & Jaggers (2000)
Political instability	5	ager	Age in years of the current regime as classified by REG; ACLP	Przeworski et al. (2000)
	5	autch98	Autocracy annual change; Polity 98	Gurr & Jaggers (2000)
	5	demch98	Democracy annual change; Polity 98	Gurr & Jaggers (2000)
	5	durable	Years since last regime transition/ since 1949; Polity IV	Marshall & Jaggers (2000)
	5	Inst3	Political Instability ; whether there was a change in Polity score last three years	Fearon & Laitin (2003)
	5	nwstate	New state	Fearon & Laitin (2003)
	5	p4mchg	Annual change in modified polity; Polity IV	Marshall & Jaggers (2000)
	5	polch98	Polity annual change; Polity98	Gurr & Jaggers (2000)
	5	proxregc	2^(-durable/0.5)	Gurr & 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 (2002)
	7	centpol3	Centralized state? (Polity III data plus updates for post-1994)	Sambanis (2002)
	7	fedpol3	Federal state? (Polity III data plus updates for post-1994)	Sambanis (2002)
	7	semipol3	Semi-federal state? (Polity III data plus updates for post-1994)	Sambanis (2002)
Neighborhood political economy	8	avgnabo	Average SIP score of neighbors	Gates et al. (2001)
	8	nmdgdp	Neighbors' average ln(GDP per capita)	Fearon & Laitin (2003) Sambanis (2001)
	8	nmdp4_alt	Neighbors' median polity (both land and water contiguity; using polity2)	Marshall & Jaggers (2000)
	8	regd4_alt	Median Regional polity (using polity2)	Sambanis (2002)
Region	9	geo1	Region: Western Europe and the US	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, %	Computed from Fearon & Laitin (2003) GDP series
Economic policy	12	expgdp	Exports of goods & services as % GDP; WDI data	World Bank (2000)
	12	trade	Trade as percept of GDP; in 1995 constant dollars	World Bank (2000)
Social welfare	13	illiteracy	% 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, % gross; WDI	World Bank (2000)
	13	seceduc	School enrollment,secondary, % 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	anuexp	Manufactures exports as percentage of merchandise exports; WDI	World Bank (2000)
	14	oil	Oil exports/GDP	Fearon & Laitin (2003); Sambanis (2004a)
	14	sxpnew	Primary commodity exports/GDP	Collier & Hoeffler (2001)
	14	sxpsq	Primary commodity exports/GDP, squared	Collier & Hoeffler (2001)
Terrain, geography, population distribution	15	lmtnest	Rough terrain	Fearon & Laitin (2003)
	15	ncontig	Noncontiguous state	Fearon & Laitin (2003)
	15	popdense	Population density: people per square km; WDI	World Bank (2000)
Militarization	16	army85	Size of government army in 1985	World Bank (2000)
	16	milper	Share of population in military forces, %	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

Table 2: Results for the core variables

Variable	Label	Estimations (sample)	Weighted mean betahat	Weighted mean sd(betahat)	P normal (one-sided test)	P non-normal (one-sided test)
ln_popns	Ln(population)	640,182	.279	.089	0.00086	.0051
ln_gdpen	Ln(GDP per Capita)	640,182	–.534	.199	0.0036	.029
pt8	2^(-years in peace/8)	640,182	.475	.525	0.18	.175

Table 3: The 18 most robust variables

Variable	Label	Concept	Weighted mean betahat	Weighted mean sd(betahat)	P normal (one-sided)	P non-normal (one-sided)
inst3	Instability, y=1 if poldiff> | 2 | ; Instability with 77 and 88 coded as 1 regardless of poldiff	5	.470	.097	.00000050	.00020
parreg	Regulation of participation; Polity IV	3	–.410	.141	.0019	.0067
proxregc	2^(–durable/0.5)	5	.299	.112	.0030	.014
anocracy	Anocracy = 0 if polity2>5; Anocracy = 1 if (polity2>–6 & polity2<6)	4	.295	.106	.0030	.024
geo34	Region: Middle East and North Africa	9	.311	.116	.0036	.015
nat_war	Whether a neighbor is at war in a given year.	10	.332	.142	.0085	.021
gdpgrowth	Annual change in GDP	11	–.545	.232	.0085	.028
partfree	Partially free polity	4	.314	.131	.0086	.025
milgnp92	Military expenditure as % of GNP, 1992	16	.316	.137	.010	.019
anoc	Dummy: Anocracy=1	4	.279	.121	.011	.052
regd4_alt	Median Regional polity (using polity2)	8	–.503	.233	.014	.033
instab	Political Instability	5	.234	.117	.019	.040
lmtnest	Rough terrain	15	.282	.139	.021	.027
nwstate	New state	5	.242	.137	.034	.063
etdo4590	Ethnic dominance measure	2	.247	.146	.037	.066
decade1	Dummy : 1960s	17	–.276	.157	.040	.051
tnatwar	Total number of neighbors at war in a given year.	10	.174	.107	.046	.069
milper	Military personnel as share of population	16	–.549	.346	.048	.020

Figure 1: Distribution of estimated t-values, ln_pop (population, log-transformed):

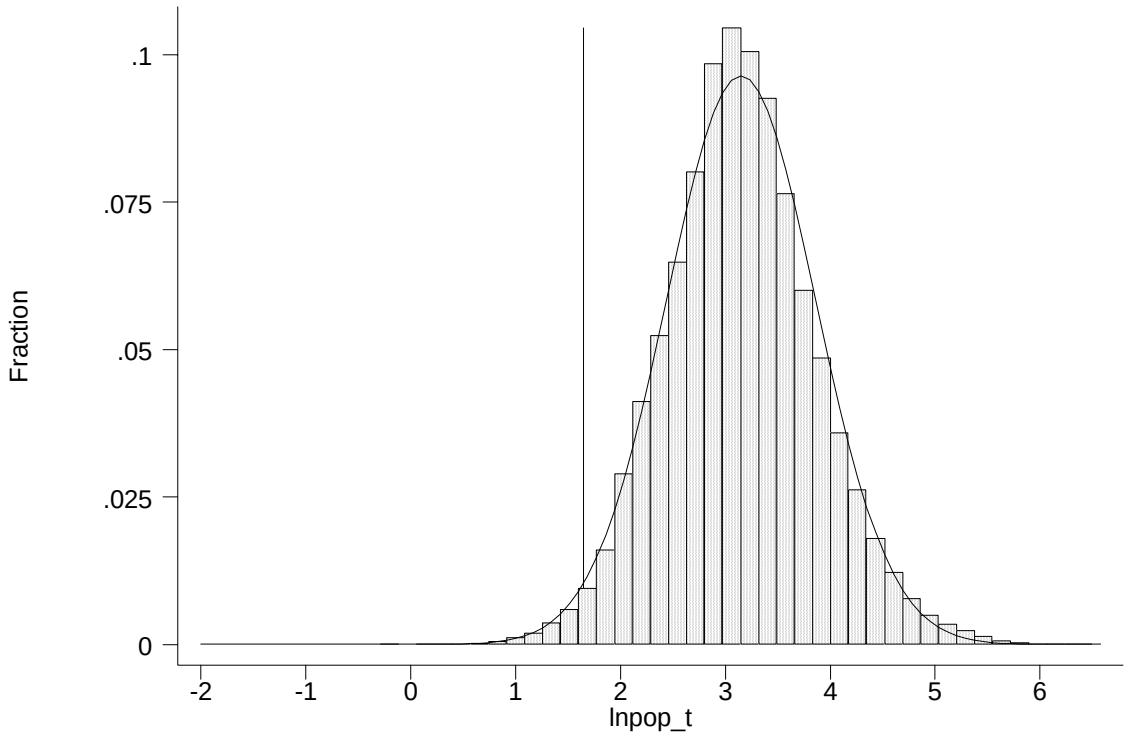


Figure 2: Distribution of estimated t-values, lngdp_t (GDP per capita, log-transformed):

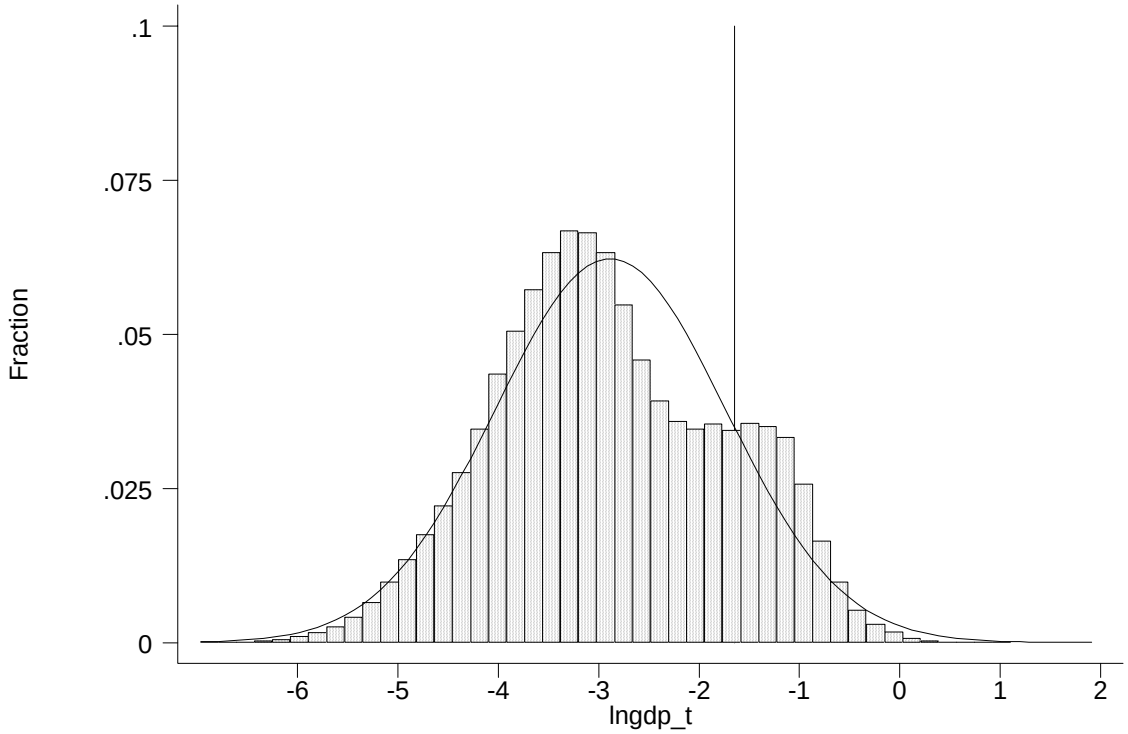


Figure 3: Distribution of estimated t-values, pt8 (2^(-years in peace/8)):

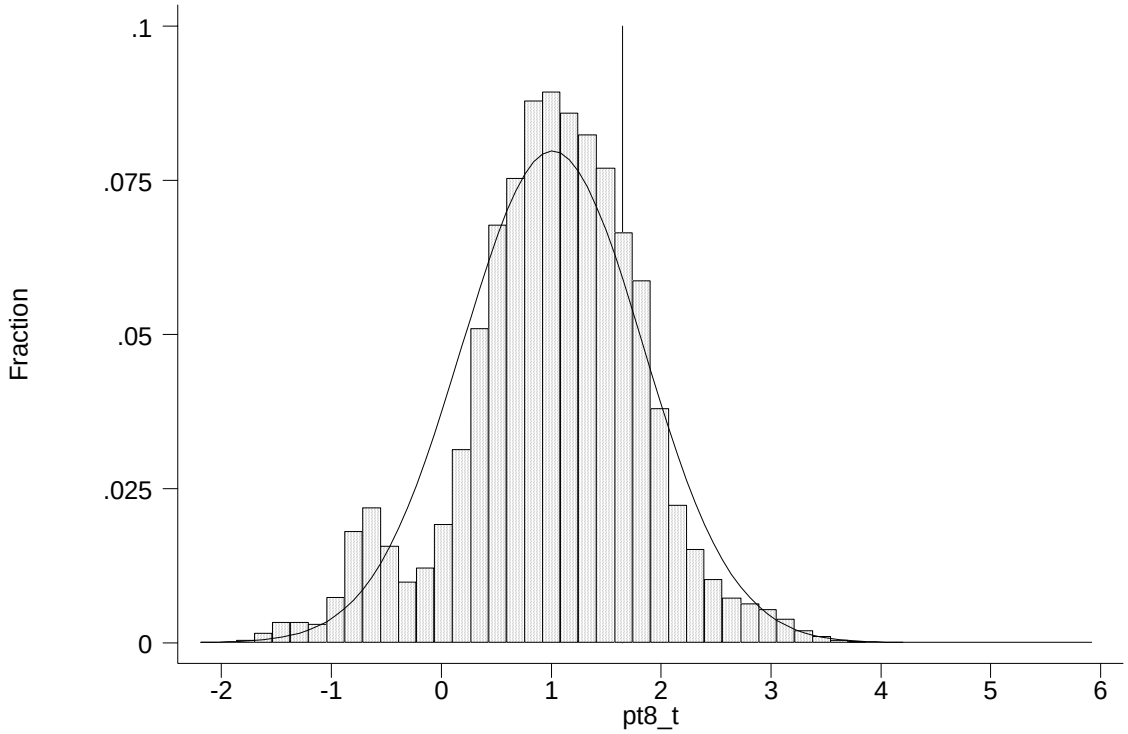


Figure 4: Distribution of estimated t-values, concept 1 (ethnic fragmentation)

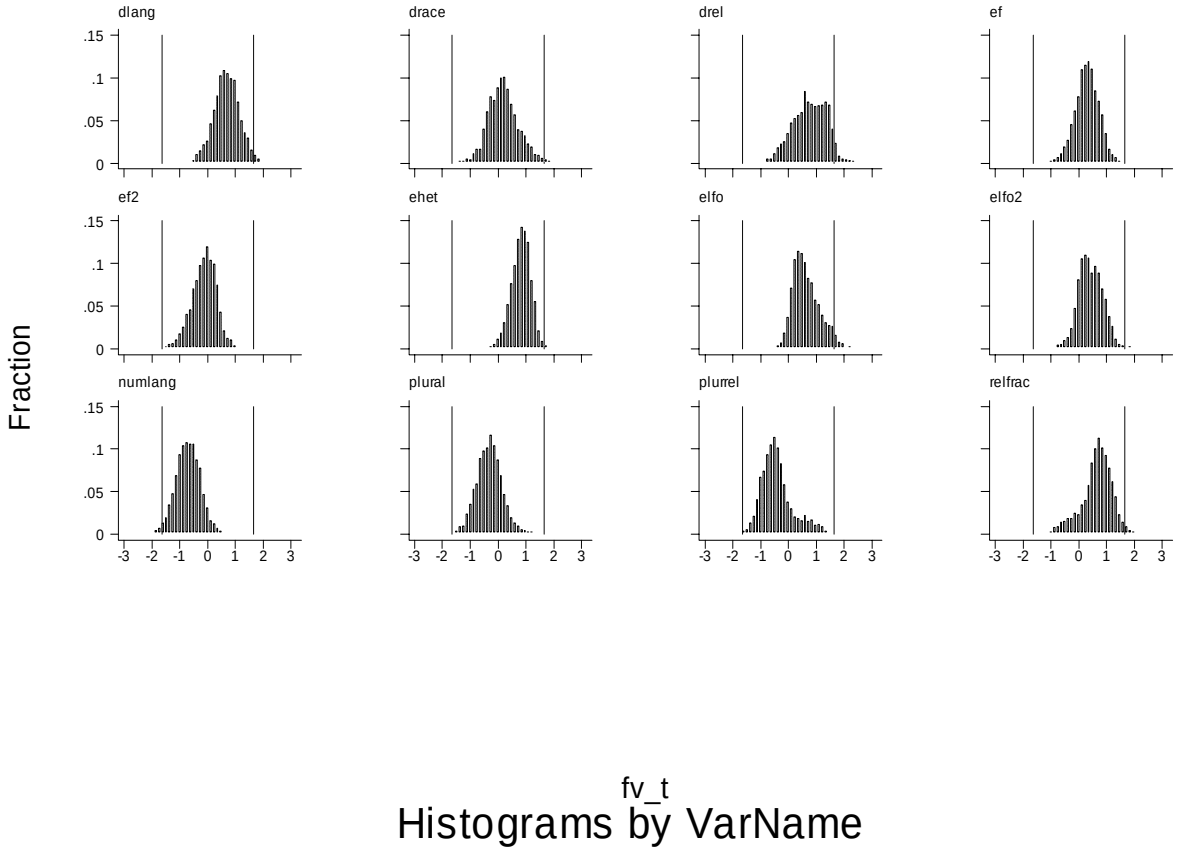


Figure 5: Distribution of estimated t-values, concept 2 (ethnic dominance/polarization)

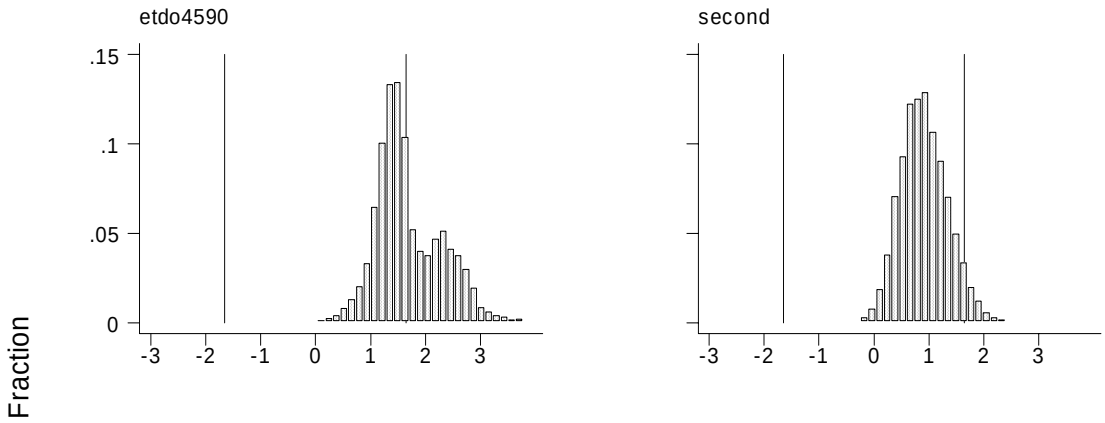


Figure 6: Distribution of estimated t-values, concept 3 (level of democracy)

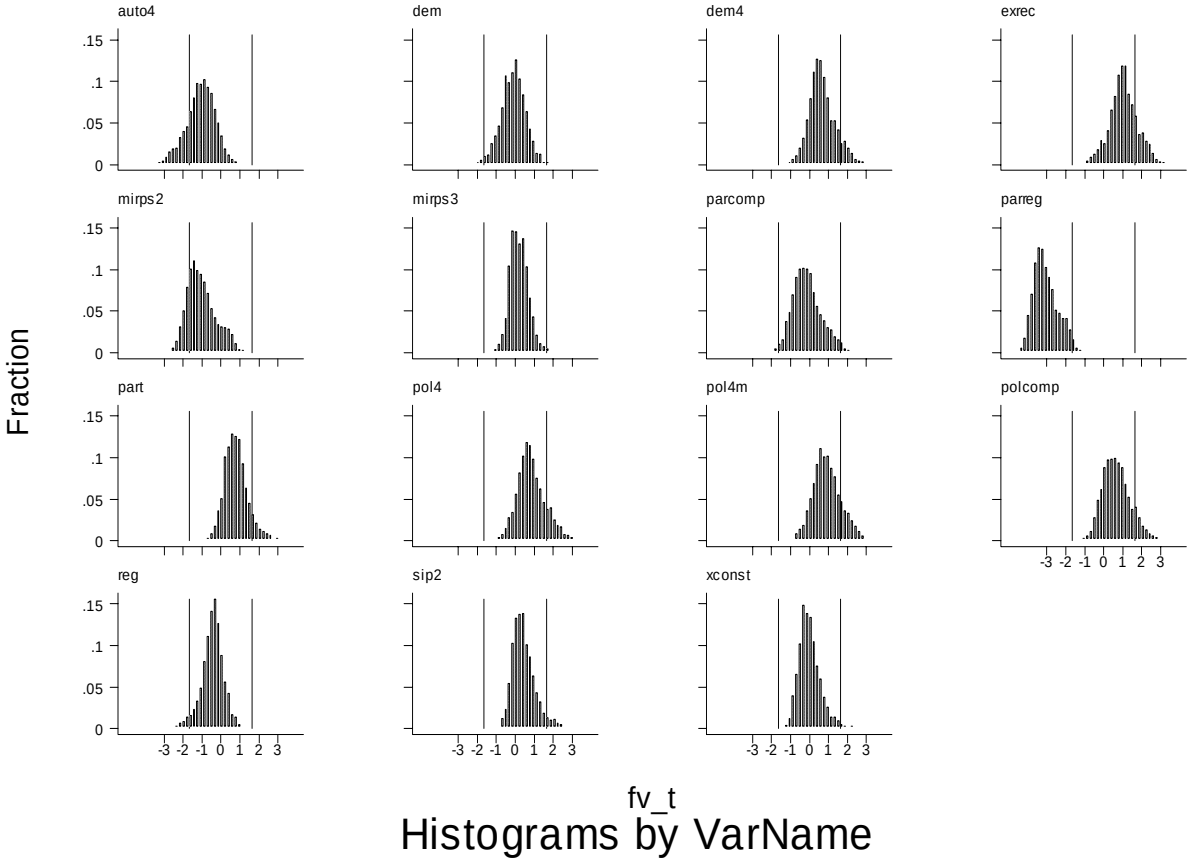


Figure 7: Distribution of estimated t-values, concept 4 (inconsistency of political institutions)

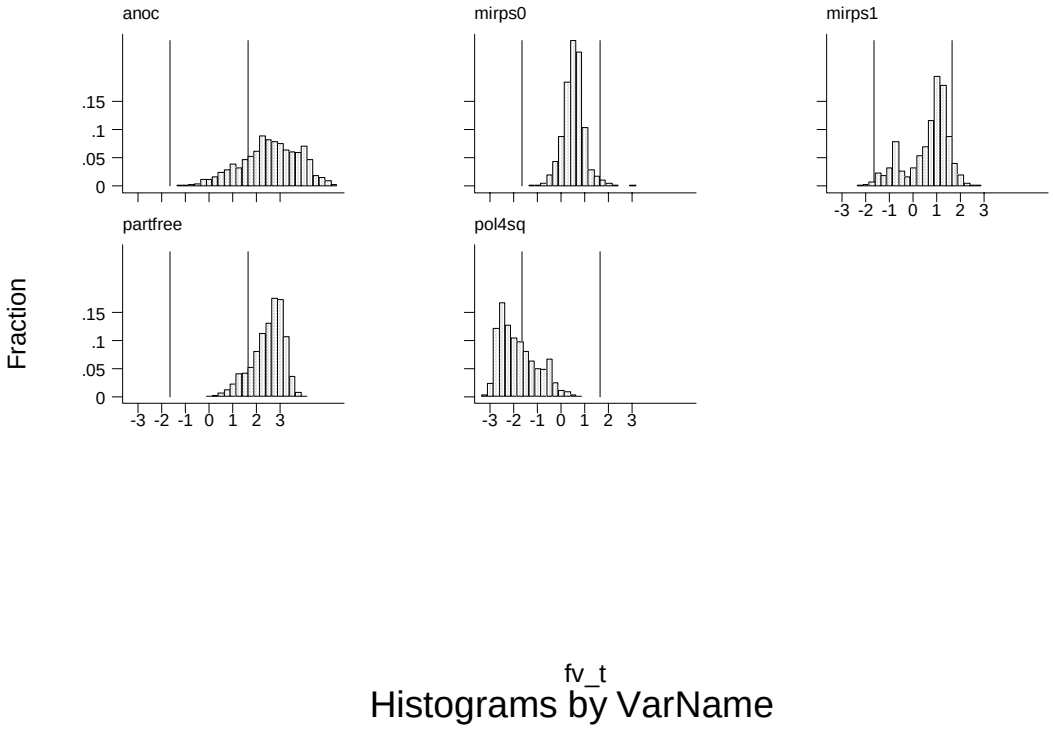


Figure 8: Distribution of estimated t-values, concept 5 (political instability)

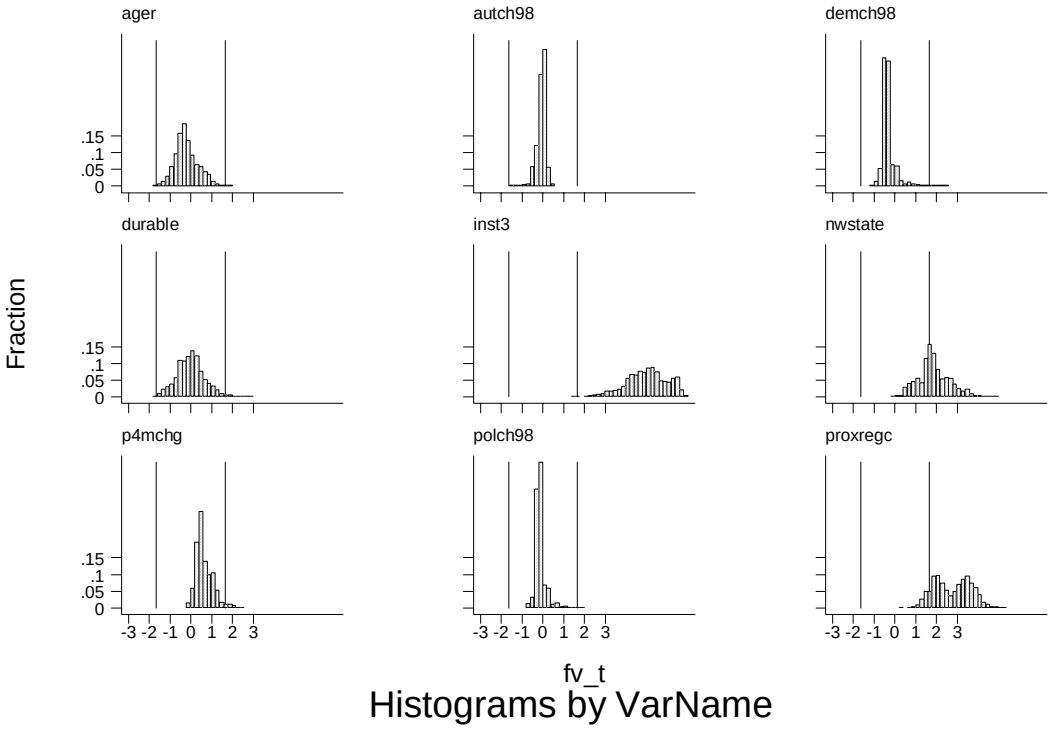
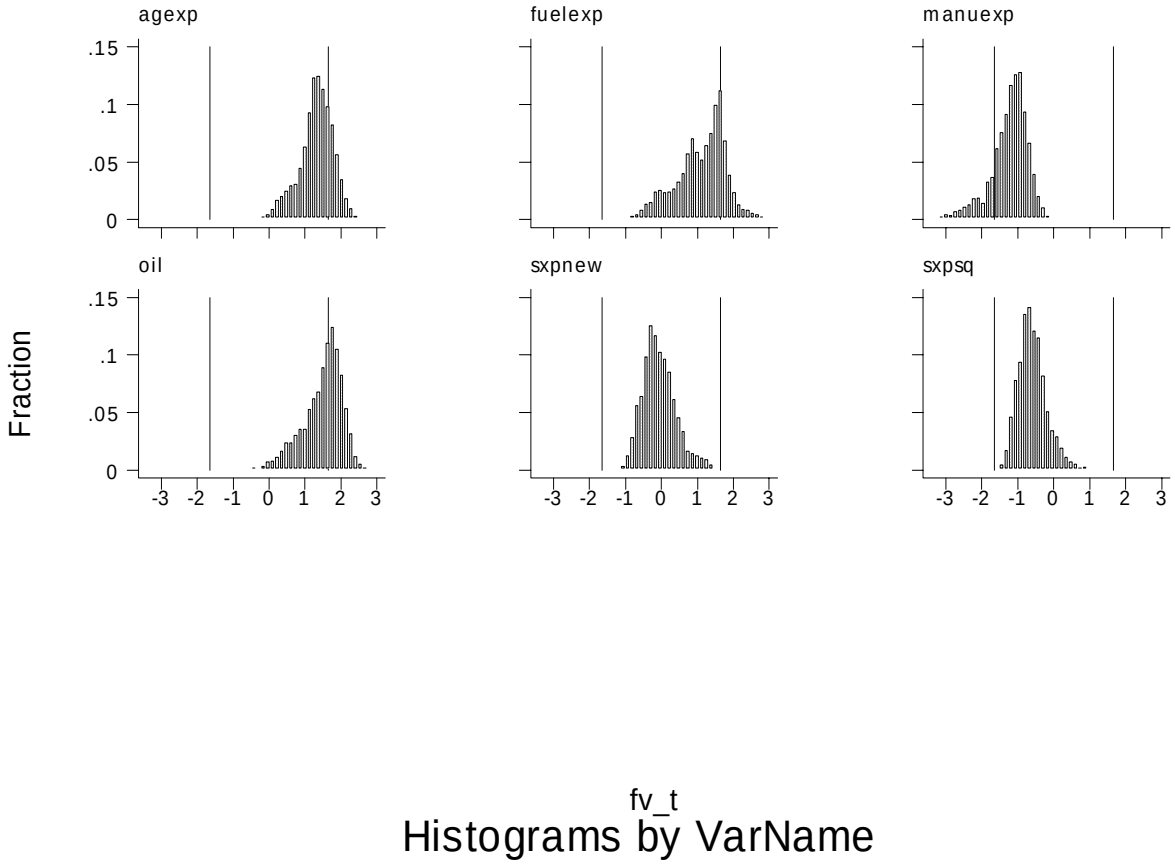


Figure 9: Distribution of estimated t-values, concept 14 (resources)



Appendix Table 1: Results for all focus variables, sorted by concept and p non-normal

Variable	Label	Concept	Estimations (sample)	Mean no. of obs.	Weighted mean betahat	Weighted mean sd(betahat)	Min (betahat)	Max (betahat)	P normal (one-sided)	P non-normal (one-sided)
ehet	Ethnic heterogeneity index	1	6599	3417	.137	.167	–.151	.547	.211	.214
drel	Religious component of Ehet	1	6424	3423	.131	.172	–.207	.583	.226	.229
relfrac	Religious fractionalization	1	6542	3429	.085	.134	–.248	.306	.272	.241
dlang	Linguistic component of Ehet	1	6524	3433	.102	.153	–.188	.431	.257	.253
elfo	Ethnolinguistic diversity	1	6360	3326	.121	.177	–.124	.613	.232	.257
numlang	Number of languages in Ethnologue	1	6536	3413	–.102	.154	–.445	.147	.252	.259
plurrel	Size of largest confession	1	6647	3426	–.046	.139	–.284	.366	.388	.290
elfo2	Ethnolinguistic diversity, squared	1	6312	3314	.079	.169	–.181	.465	.309	.306
drace	Racial component of Ehet	1	6565	3425	.032	.183	–.424	.367	.416	.330
ef	Ethnic fractionalization index	1	6538	3401	.055	.171	–.387	.360	.380	.338
plural	Share of largest ethnic group	1	6539	3429	–.046	.176	–.357	.440	.402	.341
ef2	Ef squared	1	6440	3402	–.018	.180	–.428	.248	.450	.363
etdo4590	Ethnic dominance measure	2	7611	3351	.247	.146	–.069	.670	.037	.066
second	Percent population in second largest group	2	7925	3472	.152	.166	–.069	.631	.173	.195
parreg	Regulation of participation; Polity IV	3	4938	3345	–.410	.141	–.799	–.180	.0019	.0067
exrec	Executive recruitment concept variable; Polity IV	3	4924	3334	–.190	.184	–.314	.802	.150	.168
mirps2	Consistent autocracy	3	4844	3309	–.137	.150	–.527	.271	.187	.182
auto4	Autocracy index from Polity IV	3	4906	3350	–.166	.178	–.737	.401	.174	.193
pol4m	Polity Index; Polity IV; 77 & 88 coded=0	3	4813	3377	.160	.177	–.514	.759	.182	.198
pol4	Polity index; Polity IV	3	5082	3346	.156	.188	–.410	.753	.200	.211
part	ln(share of population voting x opposition's share of votes cast)	3	4774	3352	.139	.168	–.224	.825	.206	.226
dem4	Democracy index from Polity IV	3	4707	3339	.129	.198	–.330	.755	.249	.246
polcomp	Political competition: concept variable; Polity IV	3	4956	3336	.121	.186	–.400	.666	.257	.246
parcomp	Competitiveness of participation; non-elites; Polity IV	3	4903	3351	–.018	.190	–.520	.640	.450	.290
sip2	Continuous measure of democracy	3	9936	3310	.092	.197	–.238	.897	.308	.298
reg	Dummy: 1 for dictatorships and 0 for democracies; ACLP	3	4868	3396	–.063	.183	–.550	.581	.358	.311
dem	Dummy: 1 for democracies and 0 for autocracies	3	4923	3386	–.0055	.178	–.424	.479	.497	.322
mirps3	Consistent democracy	3	4829	3307	.051	.204	–.350	.546	.389	.340
xconst	Executive constraints - operational independence of CE; Polity IV	3	9858	3363	.0092	.196	–.380	.646	.467	.343
partfree	Partially free polity	4	6208	2997	.314	.131	–.018	.516	.0086	.025
anoc	Dummy: Anocracy=1	4	6747	3446	.279	.121	–.228	.532	.011	.052
pol4sq	Pol4 squared	4	6591	3419	–.262	.158	–.644	.141	.052	.090
mirps1	Caesaristic polity	4	6668	3388	.074	.141	–.580	.352	.327	.183
mirps0	Inconsistent polity (semi-democracy)	4	6529	3368	.068	.127	–.186	.480	.295	.290

Variable	Label	Concept	Estimations (sample)	Mean no. of obs.	Weighted mean betahat	Weighted mean sd(betahat)	Min (betahat)	Max (betahat)	P normal (one-sided)	P non-normal (one-sided)
inst3	Political instability; Whether Polity coded a change or 77 or 88 in previous three years	5	5437	3393	.470	.097	.239	.724	.00000050	.00020
proxregc	2^(-durable/.5)	5	5557	3402	.299	.112	.050	.621	.0030	.014
nwstate	New state	5	9930	250	.242	.137	–.032	.902	.034	.063
p4mchg	Annual change in modified polity; Polity IV	5	11116	3385	.082	.134	–.027	.278	.270	.277
durable	Years since last regime transition/ since 1949; Polity IV	5	10980	3402	–.0053	.246	–.585	.600	.479	.311
ager	Age in years of the current regime as classified by REG; ACLP	5	10850	3409	–.031	.184	–.355	.441	.441	.325
demch98	Democracy annual change; Polity 98	5	11154	3194	–.043	.144	–.121	.204	.385	.351
polch98	Polity annual change; Polity98	5	7855	3222	–.018	.161	–.094	.201	.456	.414
autch98	Autocracy annual change; Polity 98	5	11026	3181	–.012	.173	–.216	.084	.472	.442
presi	Presidential system	6	6980	2993	–.359	.231	–.639	.067	.061	.079
inst	0-dict; 1-parliam; 2-mixed dem; 3-pres dem (Przeworski et al., 2000)	6	7176	3481	.141	.188	–.234	.707	.215	.221
major	Majoritarian system	6	6974	3000	–.0027	.158	–.396	.302	.495	.363
incumb	Consolidation of incumbent advantage(Przeworski et al., 2000)	6	7293	3482	.028	.111	–.064	.156	.399	.383
fedpol3	Federal state? (Polity III data plus updates for post-1994)	7	7159	3120	–.113	.174	–.694	.265	.24	.263
autonomy	Country has de facto autonomous regions	7	5274	3637	.034	.058	–.036	.202	.266	.275
centpol3	Centralized state? (Polity III data plus updates for post-1994)	7	7195	3123	.079	.146	–.148	.547	.276	.293
semipol3	Semi-federal state? (Polity III data plus updates for post-1994)	7	7162	3120	.0025	.104	–.291	.275	.485	.351
regd4_alt	Median Regional polity (using polity2)	8	7237	3443	–.503	.233	–1.379	.122	.014	.033
nmdp4_alt	Neighbors' median polity (both land and water contiguity; using polity2)	8	7013	3423	–.222	.198	–.686	.294	.134	.155
nmdgdp	Neighbors' average ln(GDP per capita)	8	7137	3452	–.029	.241	–.555	.560	.461	.285
avgnabo	Average SIP score of neighbors	8	7174	3408	–.009	.141	–.332	.501	.467	.306
geo34	Region: Middle East and North Africa	9	6672	3401	.311	.116	–.054	.549	.0036	.015
geo1	Region: Western Europe and the US	9	6210	3467	–.575	.351	–.820	.089	.053	.058
geo8	Region: Sub-Saharan Africa	9	6791	3418	–.130	.178	–.887	.275	.203	.240
geo69	Region: Latin America	9	6573	3422	–.067	.170	–.408	.373	.364	.280
geo2	Region: Eastern Europe and Central Asia	9	5366	3646	–.034	.257	–.593	.661	.444	.322
geo57	Region: South and East Asia and Oceania	9	6637	3408	.017	.136	–.248	.311	.434	.349

Variable	Label	Concept	Estimations (sample)	Mean no. of obs.	Weighted mean betahat	Weighted mean sd(betahat)	Min (betahat)	Max (betahat)	P normal (one-sided)	P non-normal (one-sided)
nat_war	Whether a neighbor is at war in a given year.	10	7806	3451	.332	.142	–.135	.624	.0085	.021
tnatwar	Total number of neighbors at war in a given year.	10	7746	3458	.174	.107	–.227	.458	.046	.069
gdpgrowth	Annual change in GDP, %	11	8116	3314	–.545	.232	–.961	.044	.0085	.028
expgdp	Exports of goods & services as % GDP; WDI data	12	7725	2977	–.276	.266	–.824	.324	.152	.179
trade	Trade as percept of GDP; in 1995 constant dollars	12	7857	2994	.073	.147	–.958	.341	.302	.204
pri	School enrollment, primary, % gross; WDI	13	6894	3234	–.204	.169	–.518	.169	.115	.133
life	Life Expectancy at birth; WDI	13	6933	3410	–.298	.293	–.868	.308	.154	.169
illiteracy	% adult population illiterate; WDI	13	6296	2028	.222	.253	–.278	.645	.191	.202
seceduc	School enrollment, secondary, % gross; WDI	13	6890	3145	–.305	.341	–.956	.179	.175	.202
infant	Infant mortality; WDI	13	6927	3025	.137	.231	–.273	.466	.283	.264
oil	Oil exports/GDP	14	6567	3677	.186	.128	–.160	.386	.074	.094
agexp	Agricultural raw materials exports as percentage of merchandise exports; WDI	14	6496	2630	.165	.125	–.057	.328	.094	.115
manuexp	Manufactures exports as percentage of merchandise exports; WDI	14	6620	2631	–.297	.246	–.951	.052	.111	.139
fuelexp	Fuel and oil products exports as percentage of merchandise exports; WDI	14	6473	2445	.159	.152	–.242	.429	.148	.159
sxpsq	Primary commodity exports/GDP, squared	14	6204	2111	–.251	.405	–1.066	.327	.275	.282
sxpnew	Primary commodity exports/GDP	14	6211	2106	–.014	.246	–.393	.490	.489	.359
lmtnest	Rough terrain	15	7453	3470	.282	.139	.058	.618	.021	.027
ncontig	Noncontiguous state	15	7479	3471	–.064	.154	–.455	.332	.337	.300
popdense	Population density: people per square km; WDI	15	7431	3063	–.014	.163	–.662	.450	.464	.362
milper	Share of population in military forces	16	6238	2348	–.549	.346	–2.21	–.156	.048	.020
army85	Size of government army in 1985	16	7133	3390	–.182	.162	–1.284	.029	.110	.114
decade1	Dummy : 1960s	17	6591	3436	–.276	.157	–.561	.010	.040	.051
decade2	Dummy : 1970s	17	6911	3388	.185	.120	.014	.418	.058	.071
decade4	Dummy : 1990s	17	6324	3464	.130	.150	–.264	.453	.202	.209
decade3	Dummy : 1980s	17	6882	3383	–.107	.140	–.370	.084	.218	.236
coldwar	Code 1 for Cold War year - before 1990	17	6317	3521	–.082	.154	–.471	.244	.304	.281
warhist	War in the country since 1945?	18	7755	3244	.039	.139	–.233	.306	.394	.332