

The International Politics of Incomplete Sovereignty: How Hostile Neighbors Weaken the State

Online Appendix

Contents

A1	Data Sources and Description: Correlations with Myers Scores	1
A2	Data Sources and Description: Covariates	5
A3	Additional Tables	8
A4	Robustness and Sensitivity Checks	11
A5	Additional Analyses	14
A5.1	Military Capability and Hostile States	14
A5.2	The Moderating Effect of Distance	17
A5.3	Insurgencies and State Authority	22
A5.4	Robustness to Excluding Interventionary Rivalries	25
A5.5	Hostile States and Political Geographies	27
A5.6	Terrain Ruggedness	29
A5.7	Alternate Measure of Economic Development	30
A6	Notes on the Myers Index	33
A7	Notes on Measuring Hostility	35

A1 Data Sources and Description: Correlations with Myers Scores

This section describes the data that appears in Tables 1 and 2 in the main article. More information on the validity of the Myers Index can be found in [REDACTED].

Bertelsmann Transformation Index (BTI)

The Bertelsmann Transformation Index is a subjective evaluation of countries in terms of their progress toward achieving democracy and a market economy. The index total as well as an indicator on stateness appear in Table 1 in the main text. Experts rank countries on the stateness indicator according to their performance on the monopoly of force, the existence of administrative structures, legitimacy, and the freedom of interference from religious entities. Data are from [Bertelsmann Stiftung \(2014\)](#). Higher values for the BTI and its components indicate better state capacity or state authority. The predicted correlation with the Myers Index is therefore negative.

Birth registration

Birth registration is a form of civil registration with the state. Birth registration is generally an automatic process when births occur in a state facility or a hospital. In other cases, parents must register the birth of the child with state administrative authorities in a separate process. In some cases the state will also issue a birth certificate, but this is not yet standard practice across the world. Subnational data on birth registration are from the Demographic and Health Survey, a series of nationally-represented surveys conducted in developing countries around the world.¹ The birth registration variable that appears in Table 2 in the main text is defined as the percentage of children under the age of 5 whose births were registered with the state at the time the survey was conducted. Since the likelihood of birth registration is much higher when the degree of state administration is greater - either because birth registration is an automatic administrative process or because state privileges and rights incentivize registration with the state - higher rates of birth registration suggest greater levels of domestic sovereignty. The predicted relationship with the Myers Index is negative.

The list of countries and years for which subnational data are available is: Armenia 2001, Bangladesh 2011, Burkina Faso 2006, Cambodia 1998, Cambodia 2008, Cameroon 2005, Colombia 2005, Ethiopia 2007, Ghana 2000, Haiti 2003, Honduras 2001, India 2011, Indonesia 2000, Indonesia 2010, Kenya 1999, Kenya 2009, Maldives 2006, Mozambique 2007, Nepal 2001, Nepal 2011, Niger 2001, Nigeria 2006, Pakistan 1998, Senegal 2011, Sierra Leone 2004, Swaziland 2007, Tanzania 2002, Tanzania 2012, Uganda 2002, Zambia 2010.

Failed States Index (FSI)

The Fragile States Index, previously known as the Failed States Index, ranks countries on several components indicative of fragile statehood. The FSI aggregates data across 12 social, economic, and political indicators, such as uneven economic development, state legitimacy,

¹Available from <http://www.dhsprogram.com/>.

and group grievances. The data generation process underlying the 12 indicators is proprietary and opaque, so it is unclear how the individual measures are constructed. In addition to the overall index score, Table 1 in the main text also includes two components closely related to domestic sovereignty: the provision of public services and the security apparatus's monopoly on force. Because higher values on the FSI indicate weaker states, the expected correlation with the Myers Index is positive. Data are from [Fund for Peace \(2014\)](#).

GDP per capita

National GDP data are from the Penn World Table ([Feenstra, Inklaar, and Timmer 2013](#)) and are expressed in constant terms. GDP data reported at the subnational level are scarce, but are available for Brazil, China, Indonesia, and Turkey ([Institute of Applied Economic Research 2014](#); [China Data Center 2014](#); [Badan Pusat Statistik 2014](#); [Turkish Statistical Institute 2014](#)). Net domestic product data are available for India ([Planning Commission 2002](#)). For the subnational data, all data are reported in local currencies as constant values. As higher levels of GDP per capita should be correlated with domestic sovereignty, the expected relationship with the Myers Index is negative.

Infant Mortality

Infant mortality rates are a proxy for state authority because infant mortality is sensitive to a number of factors related to state administration and control. High infant mortality rates are indicative of state failure, akin to the canary in the coal mine ([Goldstone et al. 2000](#)). The rate of infant mortality is defined as the number of infants, per 1,000 live births, who die before their first birthday. National data come from [Institute for Health Metrics and Evaluation \(2011\)](#). Subnational data are available for China, India, and Turkey ([Bignami-Van Assche 2005](#); [Planning Commission 2002](#); [Turkish Statistical Institute 2014](#)). Higher values indicate worse levels of infant mortality. The predicted relationship with the Myers Index is therefore positive.

International Country Risk Guide (ICRG)

The International Country Risk Guide provides a set of ratings of a country's economic, political, and financial risk. The index contains several components, each with their own sub-indicators, that contribute to the overall country ranking. The components include bureaucratic quality, internal conflict, law and order, and government stability, among others. Table 1 in the main text includes the correlation between Myers and the ICRG index as a whole as well as the internal conflict component and the bureaucratic quality component. The ICRG is produced through subjective expert assessments of each country along each component. Higher values on the index indicate a state facing fewer economic, political, and financial risks, which some analysts consider to be a proxy for state capacity, so we expect to observe a negative correlation. Data are from [PRS Group \(2014\)](#). Higher values on the ICRG and its components indicate stronger state authority, so I expect to observe a negative relationship with the Myers Index.

Literacy rate

An important function of the state is to provide public goods and services. In the era

of universal mass education, primary education falls under the purview of the state, and literacy rates can therefore serve as a proxy for the presence of public schools (and more generally, the state’s administrative apparatus). Literacy appears in both the national and subnational correlation validity checks for the Myers Index. National level-data are from the World Bank’s [World Bank \(2014\)](#). The World Development Indicators define adult literacy as the percentage of people ages 15 and above who can, with understanding, read and write a short, simple statement on their everyday life. Literacy generally encompasses numeracy. Subnational data for Brazil (as illiteracy rate), India, and Turkey are from [Institute of Applied Economic Research \(2014\)](#), [Planning Commission \(2002\)](#), and [Turkish Statistical Institute \(2014\)](#), respectively. Higher literacy rates (lower illiteracy rates) suggest better state administration. The predicted relationship with the Myers Index is therefore negative.

Perceptions of security

Providing for security and public order is one of the most important functions of the state, and thus is an important component of domestic sovereignty. Subjective perceptions data on security are available at the subnational level for Indonesia ([Badan Pusat Statistik 2014](#)). Higher values indicate more positive perceptions of security, so I expect to observe a negative correlation with the Myers Index.

Post offices per 10 villages

The ratio of the number of post offices per 10 villages is a proxy for the presence of the state’s administrative apparatus in a particular territory. A higher density of post offices per 10 villages suggests a higher density of state administrative activities, offices and personnel, or functions. Data are available for Indonesia ([Badan Pusat Statistik 2014](#)). The predicted relationship with the Myers Index is negative.

Primary school enrollment and teacher-student ratio

Primary enrollment (net percentage) is the ratio of children of official school age who are enrolled in school to the population of the corresponding official school age. National-level data are from [World Bank \(2014\)](#). As no subnational data on school enrollments are available, Table 2 in the main text instead uses data on teacher-student ratio to proxy for state administration ([China Data Center 2014](#)). In both cases, higher values indicate better state administration and control, and both indicators should have negative relationships with the Myers Index.

Road density and villages connected to roads

The presence of transportation infrastructure suggests a greater degree of state administration and control, as infrastructure is often provided by the state in developing countries. To proxy for transportation infrastructure, I use subnational measures of road density from [China Data Center \(2014\)](#) (China only) and the percentage of villages connected to roads from [Planning Commission \(2002\)](#) (India only). A greater density of roads and greater levels of village-road connectivity should be associated with greater state administration and control. The expected relationship with the Myers Index is therefore negative.

Worldwide Governance Indicators (WGI)

The Worldwide Governance Indicators are one of the more commonly-used national-level indicators of state capacity. Five of six of the WGI measures appear in Table 1; these indicators in principle capture separate components of state administration and control, but in practice are highly correlated with each other. The sixth indicator from the WGI, called “voice and accountability,” is conceptually closer to a measure of the regime than a measure of the state, and I therefore do not include it in Table 1 in the main text. The source data for the WGI are primarily perceptions data from citizens, businesses, and experts. Data are from [Kaufmann, Kraay, and Mastruzzi \(2014\)](#). Higher values on the WGI indicate stronger domestic sovereignty. The expected correlations between the WGI measures and the Myers Index are negative.

A2 Data Sources and Description: Covariates

Terrain ruggedness

For each province in the sample, I calculate the average “slopedness” of the region in percent rise using world topography data from the U.S. Geological Service and modified by [Patterson \(2013\)](#). The topography data provide the average elevation for each cell in a raster image file of the world. Following a procedure similar to the one deployed in [Nunn and Puga \(2012\)](#) and using spatial software, I first calculate the maximum change in elevation between that cell and its eight neighboring cells. This process identifies the steepest downhill descent from that cell. Then, I calculate the average change in elevation of all the cells contained in each province. The resulting value is a measure of slopedness that approximates the average ruggedness of the physical terrain, where slopedness is reported as the average percent rise. For this calculation, the underlying boundary data are projected into the Cylindrical Equal Area (world) projection.

Distance from the capital

Ideally, I would like to calculate distance between the national capital and the provincial capital, or from the national capital to the largest populated area in each province. However, due to harmonization issues (requiring the merging of some provinces into superunits to obtain time-consistent units) and lack of data, it is not possible to utilize either of these methods of calculation. I therefore first determine the geographic centroid of the province and then calculate the distance between this point and the national capital. Distances are given as kilometers. For this calculation, the underlying boundary data are projected into the Azimuthal Equidistance (world) projection to minimize distortions to distance.

Population density

For each province, I use the population counts that underlie the Myers calculation to generate a total population size, and then divide population by the province’s area in square kilometers to obtain population density. For this calculation, the underlying boundary data are projected into the Cylindrical Equal Area (world) projection to minimize distortions to area.

Split ethnic groups

I adapt the procedure used in [Michalopoulos and Papaioannou \(2013\)](#) to code, by province, the presence of an ethnic group that is split across a national boundary. To identify ethnic homelands, I use geospatial data from the Georeferencing of Ethnic Groups (GREG) dataset ([Weidmann, Rød, and Cederman 2010](#)). Then, following [Michalopoulos and Papaioannou \(2013\)](#), I intersect the GREG data with a layer of national boundaries from the Global Administrative Areas database (GADM).² Because the GREG and GADM GIS data are likely drawn with some error, I identify a split group as any group in which at least 10% of their

²Available at <http://www.gadm.org/>

total surface area belongs to more than one country.

To identify split groups by province, I intersect the GREG data with the subnational boundary data used in this project. Again, due to concerns about measurement error, I drop any group whose total surface area within a given *province* represents less than 10% of the total surface area occupied by all ethnic groups (as the GREG data code some areas as occupied by no group). I then code the split groups variable as whether the predominant ethnic group in the province is split across a national border; I operationalize this as the group that has at least a plurality in the province. The ethnic groups must be contiguous with the group across the national border. For this calculation, the underlying data are projected into Cylindrical Equal Area (world) projection.

Development

Economic development is difficult to code, not only due to limitations in data, but because it is partially a consequence of governance. Measures that are available over time are often not available subnationally, presenting a tradeoff for analysts. I draw on geospatial data from the Global Land Cover project to develop a measure of human economic activity as a proxy for development ([Latham et al. 2014](#)). The Food and Agriculture Organization of the United Nations classifies land cover types into 11 classes. Two of these classes, artificial surfaces and cropland, are the direct result of human economic activity. Other surface types include mangroves, snow and glaciers, and water bodies, and are unlikely to exhibit much human activity indicative of economic development. A drawback of this approach is that land cover data are only available for a snapshot in time (approximately 2010), even though human economic activity is not time-invariant. However, given the paucity of subnational development data, the land cover approach is a suitable alternative approach even if it is measured with some error. For each province, I calculate the percentage of land covered by either artificial surfaces or cropland. For this calculation, the underlying boundary data are projected into Cylindrical Equal Area (world) projection.

Natural Resources

For each province, I calculate the percentage of territory containing known oil and gas deposits using geospatial data from PRIO's Petroleum Dataset ([Lujala, Rød, and Thieme 2007](#)). For this calculation, data are projected into the Cylindrical Equal Area (world) projection.

National Capital

For each province, I code the capital variable as 1 if the unit contains the national capital city and 0 otherwise.

Area

I calculate province area in square kilometers. For this calculation, the underlying boundary data are projected into the Cylindrical Equal Area (world) projection to minimize distortions to area.

Relative Military Capability

I code a set of categorical variables that indicate whether province i in target state j in year t borders no rival, a weaker rival, a rival at parity, or a stronger rival. To assess rival state military capability relative to target state military capability, I calculate the ratio between the rival state's capability and the target state's capability using the Composite Index of National Capability (CINC) by [Singer \(1987\)](#). CINC scores of national military capability are calculated using data on iron and steel production, primary energy consumption, the number of military personnel, expenditures on the military, and urban and total population size. I then calculate a capability ratio by dividing the CINC score of the neighboring rival state (or most powerful neighboring rival state, in the case of more than one hostile neighbor) by the target state's CINC score. Weak rivals are rivals where the CINC ratio is less than $3/5$. Rivals at "parity" are rivals whose CINC ratio is between $3/5$ and $5/3$, inclusive. Strong rivals are those with CINC ratios greater than $5/3$.

A3 Additional Tables

Table A3.1: List of Countries and Years in the Full Sample

Country	Years
Algeria	1966
Argentina	1970, 1980, 1991, 2001, 2010
Armenia	2001, 2011
Bangladesh	1981, 1991, 2001, 2011
Belarus	1999
Belize	2000
Bhutan	2005
Bolivia	1976, 1992, 2001, 2012
Brazil	1970, 1980, 1991, 2000, 2010
Burkina Faso	1985, 1996, 2006
Burma	1983
Burundi	1979, 1990
Cambodia	1962, 1998, 2008
Cameroon	1976, 1987, 2005
Chile	1960, 1970, 1982, 1992, 2002
China	1982, 1990, 2000, 2010
Colombia	1964, 1973, 1985, 1993, 2005
Congo-Brazzaville	1974, 1984, 2007
Costa Rica	1963, 1973, 1984, 2000, 2011
Croatia	2011
Ecuador	1962, 1974, 1982, 1990, 2001, 2010
Egypt	1996, 2006
El Salvador	1992
Estonia	2001, 2011
Ethiopia	1994, 2007
Gambia	1983, 1993
Ghana	1960, 1970, 1984, 2000, 2010
Guatemala	1994, 2002
Guinea	1982, 2003
Haiti	1982, 2003
Honduras	1988, 2001
India	1961, 1971, 1981, 1991, 2001, 2011
Indonesia	1971, 1980, 1990, 2000, 2010
Iran	1966, 1976, 1986, 1996, 2006
Iraq	1965, 1997
Jordan	2004
Kenya	1969, 1979, 1989, 1999, 2009

Continued on next page

Table A3.1 – *Continued from previous page*

Country	Years
Kyrgyzstan	1999, 2009
Liberia	1974
Lithuania	2001, 2011
Malawi	1987, 1998, 2008
Malaysia	1970, 1980, 1991, 2000
Mali	1976, 1987, 1998, 2009
Mauritania	1988
Mexico	1960, 1970, 1990, 1995, 2000, 2005, 2010
Mongolia	1989, 2000
Morocco	1982, 1994, 2004
Nepal	1961, 1981, 2001
Nicaragua	1963, 1971, 1995, 2005
Niger	1977, 2001
Nigeria	2006
Pakistan	1973, 1981, 1998
Panama	1960, 1970, 1980, 1990, 2000, 2010
Papua New Guinea	1980, 1990, 2000
Paraguay	2002
Peru	1993, 2007
Romania	1977, 1992, 2002
Rwanda	1978, 1991, 2002
Senegal	1976, 1988, 2002
Sierra Leone	2004
Slovenia	2002, 2011
Somalia	1975
South Africa	1970, 1980, 1991, 1996, 2001, 2007
South Korea	1960, 1970, 1980, 1985, 1990, 2000
Sudan	1993, 2008
Swaziland	1976, 1986, 1997, 2007
Syria	1960, 1970
Tanzania	1978, 1988, 2002, 2012
Thailand	1960, 1970, 1980, 1990, 2000
Togo	1981
Turkey	1975, 1985, 1990, 2000
Uganda	1991, 2002
Uruguay	1963, 1975, 1985, 1996, 2006
Venezuela	1971, 1981, 1990, 2001, 2011
Vietnam	1989, 1999, 2009
Zambia	1969, 1980, 1990, 2000, 2010

Table A3.2: Correlations for Variables

Variables	Myers	Rivalry	Terrain	Distance	Density	Ethnic	Development	Resources	Capital
Myers	1.000								
Rivalry	0.153 (0.000)	1.000							
Terrain	-0.093 (0.000)	0.133 (0.000)	1.000						
Distance	-0.001 (0.973)	0.068 (0.002)	-0.029 (0.184)	1.000					
Density	-0.122 (0.000)	-0.056 (0.010)	0.104 (0.000)	-0.206 (0.000)	1.000				
Ethnic	0.188 (0.000)	-0.051 (0.019)	0.001 (0.971)	-0.047 (0.031)	-0.029 (0.181)	1.000			
Development	0.009 (0.672)	-0.187 (0.000)	-0.152 (0.000)	-0.356 (0.000)	0.539 (0.000)	-0.011 (0.623)	1.000		
Resources	-0.112 (0.000)	0.107 (0.000)	-0.068 (0.002)	0.024 (0.277)	0.020 (0.358)	-0.057 (0.009)	0.067 (0.002)	1.000	
Capital	0.015 (0.489)	-0.065 (0.003)	-0.021 (0.327)	-0.351 (0.000)	0.104 (0.000)	0.035 (0.110)	0.032 (0.138)	-0.053 (0.015)	1.000

p-values in parentheses

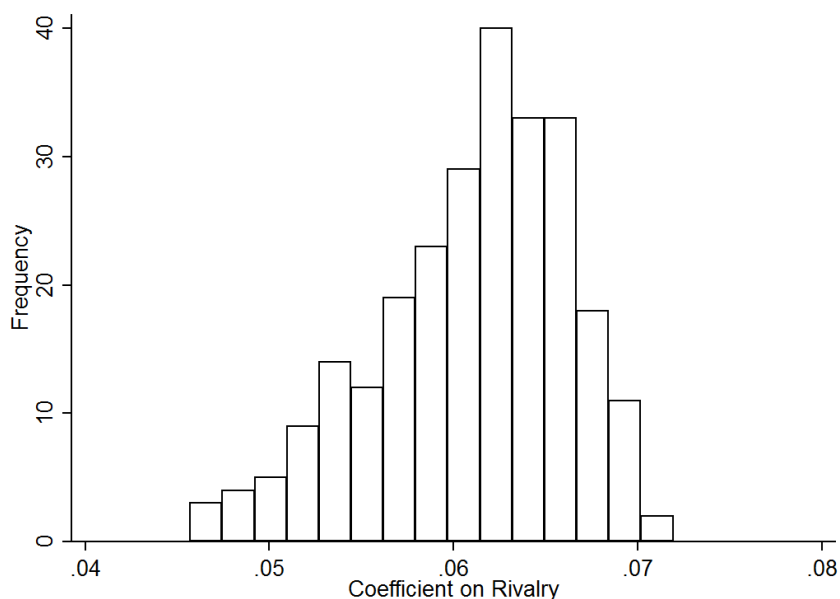
Myers, terrain, distance, and density are logged

A4 Robustness and Sensitivity Checks

First, I examine the effect of an additional covariate: area. I do not include area in the main model because the statebuilding challenge is not a problem of projecting authority over large areas, but over large areas that are sparsely populated (Herbst 2000). The results support this intuition about population density and area, as the area variable does not reach statistical significance (Column 1 of Table A4.1). The effect of rivalry is robust to the inclusion of area.

I then examine whether the results sensitive to the selection of observable covariates. To investigate this question, I iterate for each possible combination of covariates including area. A histogram of the coefficients on the main variable of interest, rivalry, is given in Figure A4.1. If the results are not sensitive to the precise specification and choice of covariates in the model, then I would expect to observe a distribution that approximates the shape of a normal distribution. Indeed, this is what the histogram shows.

Figure A4.1: Histogram of Rivalry Coefficients



The results are robust to alterations to the composition of the sample and to different specifications regarding the effects of time as well. To test sample sensitivity, I rerun the analysis and in each iteration drop a different country. Figure A4.2 shows the point estimates and 95% confidence intervals for the rivalry indicator for each iteration. To test specification sensitivity with respect to time, I repeat the main analysis using country and year fixed effects (with standard errors clustered by country) instead of country-year fixed effects, and I repeat the over-time analysis using country and year fixed effects (with standard errors clustered

Table A4.1: Robustness Checks

	(1)	(2)	(3)
	Cross-sectional	Cross-sectional	Longitudinal
Rivalry	0.0665* (0.0283)	0.0746+ (0.0399)	0.0730+ (0.0414)
Terrain ruggedness	-0.00112 (0.0154)	-0.00901 (0.0304)	
Distance from capital	0.0286 (0.0185)	0.0311 (0.0294)	
Population density	-0.133** (0.0404)	-0.0920* (0.0441)	-0.0125 (0.0333)
Split ethnic group	0.0706* (0.0295)	0.0762+ (0.0427)	
Development	0.00868 (0.0200)	-0.00945 (0.0384)	
Natural resources	-0.0206* (0.00986)	-0.0219 (0.0150)	
Capital	-0.104* (0.0480)	-0.112 (0.0696)	
Area	0.000933 (0.0361)		
Clustering	Country-year	Country	Province
Fixed Effects	Country-year	Country, year	Country, year
Observations	2115	2115	2115

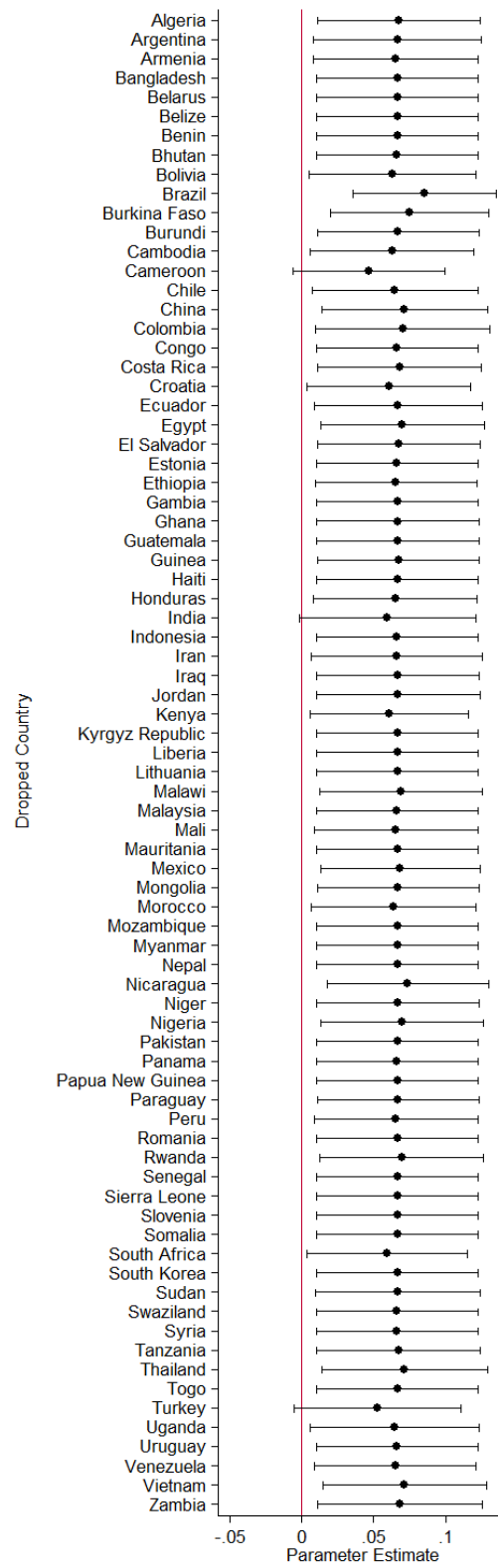
Standard errors in parentheses and clustered as indicated

Intercepts suppressed

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

by province) instead of a linear time trend (Columns 3 and 4 of Table A4.1, respectively). In both cases, the deleterious effect of rivalry remains robust.

Figure A4.2: Sample Sensitivity



A5 Additional Analyses

A5.1 Military Capability and Hostile States

The analysis in the main text investigated the effect of hostile neighbors on state authority using a binary measure of rivalry as a proxy for hostility. This binary measure ignores the heterogeneity in rival state capability, a potentially important feature of rival states that might affect their ability to influence state authority in target states.

The theoretical argument suggests that enemy states should find subversion in domestic sovereignty to be a desirable strategy, regardless of enemy state capability. Undermining state authority is attractive in part because it is cheap and often harder to detect relative to more overt policy instruments, such as conventional military force. Militarily weak rivals may especially prefer sovereignty-undermining tactics for precisely these reasons. After all, one does not need a significant level of military capacity to sponsor rebel groups or other challengers in the target. Yet sovereignty-undermining tactics are by no means a weapon of only the weak: strong states sometimes opt to degrade state authority in their neighbors. Recent Russian machinations in eastern Ukraine are only the latest example of militarily strong states subversion in the domestic sovereignty of their neighbors. Thus, it is *ex ante* not theoretically obvious that one should observe a relationship between the rival's military capability and the likelihood that it will succeed in destabilizing state authority in the target.

That said, it is still informative to examine how a rival's military capability may affect the degree to which the target state exercises effective authority. I code a new set of categorical variables that indicate whether a province in year t borders no rival, a weak rival, a rival at parity, or a stronger rival. To assess rival state military capability relative to target state military capability, I calculate the ratio between the rival state's capability and the target state's capability using the Composite Index of National Capability (CINC) by [Singer \(1987\)](#). The ratios range in principle from 0 to 1. Weak rivals are rivals where the CINC ratio is less than $3/4$. Rivals at parity are rivals whose CINC ratio is between $3/4$ and $4/3$, inclusive. Strong rivals are those with CINC ratios greater than $4/3$.

Using these new categorical variables, I rerun the cross-sectional analysis, omitting both the original binary measure of rivalry as well as the no rivalry category (the baseline). Table [A5.1](#) shows the results. As in the main text, all continuous variables have been transformed to have mean 0 and standard deviation 1 to ease interpretation, and the coefficient on the rivalry variables should be positively signed. Column 1 simply repeats the results from the main model in Table 4 of the article for purposes of comparison, while Column 2 shows the results using the categorical variables. These results indicate that both weak rivals and strong rivals degrade state authority. As one might expect, stronger rivals tend to have a more powerful deleterious effect on state authority (about a tenth of a standard deviation

shift) than weaker rivals. The estimate for weaker rivals is somewhat imprecise, falling just short of the $p < 0.10$ threshold, but is signed as theory would predict.

I do not conduct a longitudinal analysis of the effect of rival military capability on state authority for two reasons. First, as described in the article, there is limited over-time variation in rivalry for the set of countries in my sample, and splitting the rivalry variable into three categories would exacerbate this problem. Second, it would also be unclear what to make of changes in the categories: a province may border a weak rival in time t but not in time $t+1$ either because the rival's military capability has increased substantially relative to the target (a shift to the parity or strong rival categories) or because the rivalry has terminated (a shift to the no rivalry category).

Table A5.1: Effect of Hostility on State Control

	(1)	(2)
Rivalry	0.0666* (0.0285)	
Weaker rival		0.0672 (0.0445)
Rival at parity		0.0205 (0.0362)
Stronger rival		0.0926* (0.0436)
Terrain ruggedness	-0.00117 (0.0160)	0.00305 (0.0186)
Distance from capital	0.0286 (0.0183)	0.0251 (0.0201)
Population density	-0.133** (0.0308)	-0.141** (0.0335)
Split ethnic group	0.0706* (0.0291)	0.0678* (0.0310)
Development	0.00869 (0.0201)	0.0113 (0.0214)
Natural resources	-0.0207* (0.00991)	-0.0135 (0.00911)
Capital	-0.103* (0.0455)	-0.109* (0.0497)
Constant	-0.0454** (0.0155)	0.0304 ⁺ (0.0156)
Observations	2115	1857

Standard errors in parentheses and clustered by country-year

The omitted category in Column 2 is no rivalry

⁺ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

A5.2 The Moderating Effect of Distance

This study investigates how hostile neighbors undermine state authority through a comparison of *border provinces*. The focus on border provinces is both theoretically and empirically motivated. From a theoretical standpoint, although in principle enemy states could undermine state authority in the heartland of a target state, in practice very few states have the ability to project power over long distances (and the states most likely to be able to do so are developed countries like the United States, which do not enter the sample). From an empirical standpoint, the ability to draw valid inferences about the effect of neighbors derives from a comparison of borders provinces next to hostile neighbors with border provinces not next to hostile neighbors.³ Consequently, I restrict the sample only to border provinces.⁴ This sample restriction also accounts for the effect of distance from the capital, since provinces that have an international land border may be more distant from the capital than those that do not have an international border.

Still, an implication of the theory is that distance from the capital might have a moderating effect on rivalry, *even among the set of border provinces*. For example, one might think that as one moves further from the capital, the effect of rivalry might be stronger among provinces more distant from the capital, since it is less costly for the enemy state to engage in political interference and because it is harder for the target state to respond.

To examine the possible moderating effects of distance, I implement an interaction model following the state-of-the-art on analyzing interaction effects as given in [Hainmueller, Mummolo, and Xu \(2016\)](#). The authors suggest using a more flexible approach for estimating interactive effects rather than the standard multiplicative interaction model. The authors note that standard multiplicative interaction models often suffer from two problems: such models assume 1) a linear interaction effect and 2) that there are sufficient observations at the values of the treatment variable X and that these data points exhibit sufficient variation on moderator D (common support). In the context of this paper, a linear interaction effect would mean that the effect of rivalry on state authority linearly changes with distance at a constant rate. The common support assumption would mean that there are sufficient number of observations on the *no rivalry* and *rivalry* conditions and that they vary across distance from the capital.

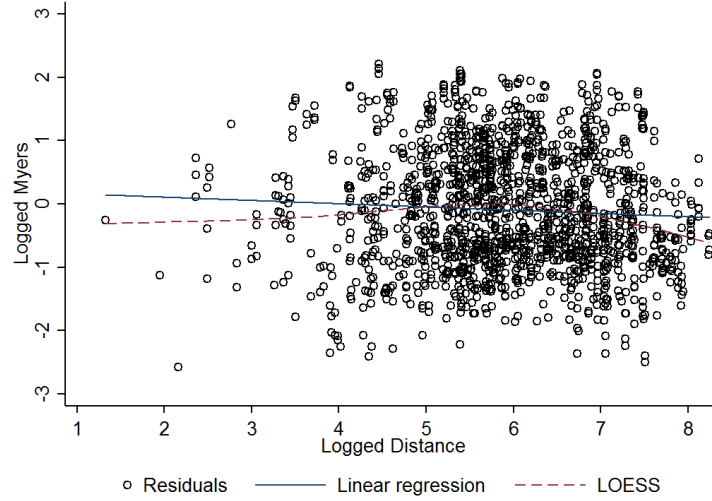
To assess these assumptions, I implement a simple diagnostic from [Hainmueller, Mummolo, and Xu \(2016\)](#) and plot residualized logged Myers scores (the dependent variable) against the moderator of logged distance (Figure A5.1). I use the residualized values of the dependent variable since the preferred regression in the main analysis includes a number of covariates. The upper panel, Figure A5.1a, shows the set of border provinces next to non-rival states,

³A secondary empirical concern is that it is impossible to code which interior provinces are at risk from neighboring states, since they by definition do not border any other states.

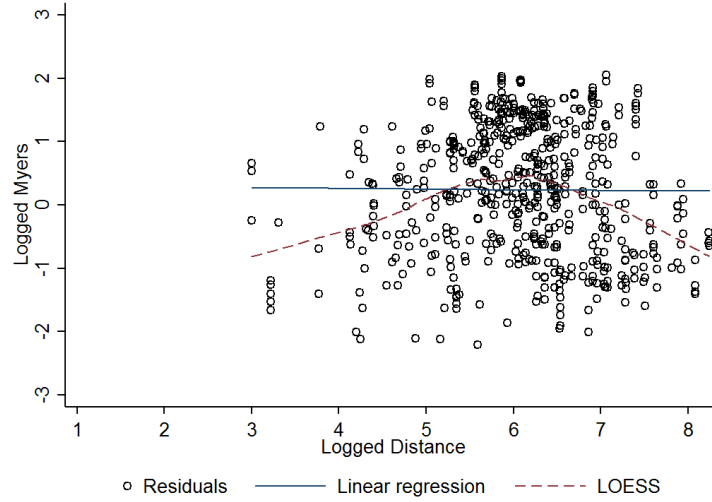
⁴[Cederman et al. \(2013\)](#) make a similar move in their paper on transnational kin groups and restrict analysis only to contiguous states.

and the middle panel, Figure A5.1b, shows the set of border provinces next to rival states. The plots also show the linear regression line (as a solid line) as well as a LOESS line (as a dashed line). Because the linear regression line and the LOESS lines diverge in both groups, the plots indicate that the relationship between state authority and distance from the capital is not linear but probably inverse-U shaped - failing the linear interactive effect assumption.

Figure A5.1: Diagnostic Scatter Plots



(a) No Rivalry

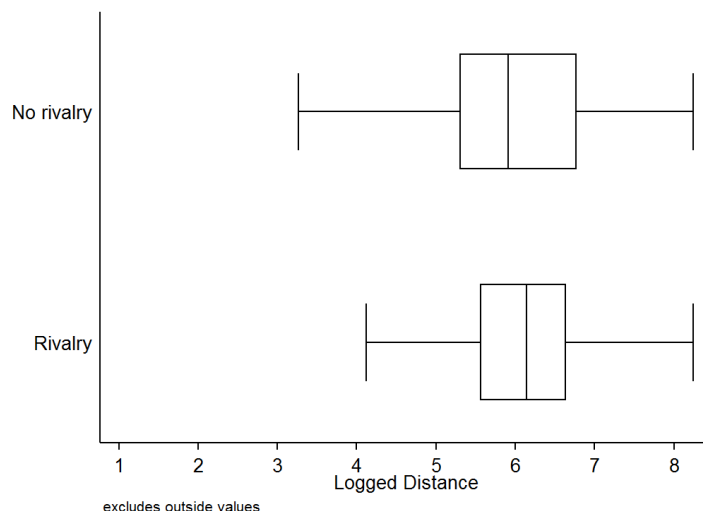


(b) Rivalry

I also check the common support assumption by plotting the distribution of logged distance (the moderator) by the two values of rivalry (rivalry or no rivalry). Figure A5.2 shows two box

plots. The box represents the 25th and 75th percentiles, while the line in the box represents the median. The whiskers of the plot indicate the lowest and highest datum within 1.5 of the interquartile range (Q3-Q1). As seen in the figure, the rivalry and non-rivalry groups share a common support of logged distance for the range of about 4 to 8.25; quite a bit of the observations in the non-rivalry group fall outside this shared range. This suggests that there may be insufficient observations in the lower part of the distribution.

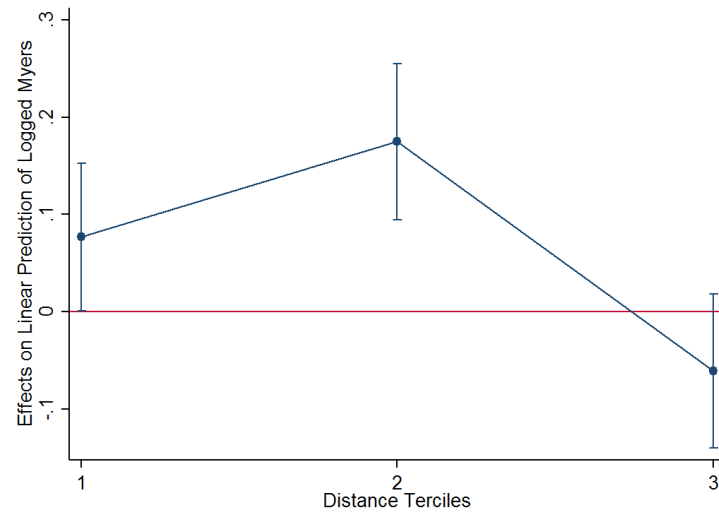
Figure A5.2: Distribution of Observations



Taken together, these diagnostics suggest that the standard multiplicative interaction model is inappropriate for assessing the effect of rivalry moderated by distance from the capital. I therefore implement a suggestion from [Hainmueller, Mummolo, and Xu \(2016\)](#), which is to discretize distance from the capital into three bins (corresponding to the terciles) and create dummy variables from these bins. These bins have a natural interpretation of low, medium, and high distance from the capital. I then estimate a model that interacts the bin dummies for distance with the rivalry dummy, and present the results in Table [A5.2](#). As before, all continuous variables including the Myers Index are standardized to have mean 0 and standard deviation 1 to aid interpretation. Because the effects are somewhat difficult to interpret from a table, Figure [A5.3](#) shows the marginal difference between rivalry and no rivalry as one moves between the terciles.

The Myers Index is scaled so that higher values indicate worse levels of state authority. As Figure [A5.3](#) indicates, the effect of rivalry is stronger among provinces in the second tercile of distance compared to those in the first tercile. As we move to the third tercile, the difference between rivalry and non-rivalry is not statistically distinguishable from zero. However, recall that the sample consists of *provinces that have an international land border* with another country. Because the sample is already “accounts” for distance by excluding the interior provinces, it is not too surprising that once one moves to the third tercile of distance, rivalry

Figure A5.3: Conditional Marginal Effects of Rivalry



does not add any additional explanatory power for the border provinces located extremely far from the capital.

Table A5.2: The Effect of Rivalry Conditional on Distance from the Capital

	(1)
No rivalry*2nd distance tercile	0.0109 (0.0213)
No rivalry*3rd distance tercile	0.109** (0.0394)
Rivalry*1st distance tercile	0.0767* (0.0388)
Rivalry*2nd distance tercile	0.186** (0.0439)
Rivalry*3rd distance tercile	0.0476 (0.0515)
Terrain ruggedness	0.00628 (0.0156)
Population density	-0.143** (0.0298)
Split ethnic group	0.0745** (0.0284)
Development	0.0131 (0.0206)
Natural resources	-0.0232* (0.00929)
Capital	-0.126** (0.0400)
Constant	-0.0875** (0.0257)
Observations	2115

Standard errors in parentheses and clustered by country-year

Country-year intercepts suppressed

The omitted category is no rivalry*1st distance tercile

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

A5.3 Insurgencies and State Authority

The main analysis does not directly test the subversion strategy with quantitative methods and instead relies on qualitative evidence from the case of the Philippines and Malaysia. The barrier to a more direct quantitative test is a paucity of data. Geo-referenced data on armed conflict exist for only the period 1989-2008 (Dittrich Hallberg 2012). Because Myers scores are not available on an annual basis, restricting the sample to 1989-2008 results in the loss of about half the observations. However, it may still be fruitful to examine whether enemy states do in fact support rebel groups (one means of subversion), and whether armed conflict reduces state authority.

To identify provinces affected by insurgency, I rely on PRIO's Conflict Site data, which provides geographic information on armed conflicts in the UCDP/PRIO Armed Conflict Dataset (ACD) (Dittrich Hallberg 2012). For each conflict in the ACD, the Conflict Site dataset provides geographic center-point coordinates and a radius that allows researchers to identify conflict zones. I restrict attention to civil wars and internationalized civil wars. Using GIS software, I calculated the percentage of territory in each province in my dataset that lies within the conflict zone polygon. Because the conflict zone polygons are very rough estimates of the spatial extent of conflict and because of the measurement error problem inherent in the analysis of spatial data, province are coded as affected by *insurgency* only if 20% or more of the province's territory lies within the conflict site polygon (i.e., the variable takes value 1 if 20% or more of the provinces territory is in the conflict site polygon and 0 otherwise). Because of data limitations with respect to the Myers Index, the unit of analysis is the province-decade, where the decade can either be 1989-1998 or 1999-2008.

I then used UCDP's External Support dataset to identify *externally-supported insurgencies*, a subset of all insurgencies that appear in the ACD (Högbladh, Pettersson, and Themnér 2011; Pettersson 2011). The indicator for externally-supported insurgencies takes value 1 if a contiguous neighbor supported the insurgency and 0 otherwise. It is coded as missing if there is no insurgency present in province-decade. External support by contiguous neighbors is conceptually and operationally distinct from rival states, though in practice states do not foment insurgency in their neighbors when they are on friendly terms. By focusing on contiguous neighbors, cases of external support from further abroad are excluded. For example, Libyan support for some sub-Saharan African rebel groups would not be coded. However, ignoring these cases of external support by non-contiguous enemy states should bias against finding an effect of hostile neighbors.

Most countries never experience civil war in the 1989-2008 period. Of the countries in my dataset, 46 observations totaling 32 unique border provinces from 17 countries experience civil war during that time period (recall that the main analysis restricts attention to border provinces for the purposes of analytic and inferential leverage).

With so few observations of insurgency, a more sophisticated mediation analysis is not possible, but one can look for evidence of simple associations between rivalry and externally-

supported insurgency and between insurgency (externally-supported or otherwise) and state authority. Given insurgency, we might ask whether rival states are more likely than other neighboring states to be involved as external sponsors. Table A5.3 diagrams this relationship and includes only province-decades affected by insurgency. Looking down the right column, we see that 32 cases involved at least one external contiguous neighbor as a sponsor, and 22 of those 32 of those sponsors were in fact rivals as coded by Thompson and Dreyer (2012). In other words, nearly two-thirds of externally-sponsored insurgencies occurred in provinces bordering a rival state.

Turning now to the bottom row of Table A5.3, it is also apparent that among provinces bordering a rival state, 22 of 23 province-decades experiencing insurgency saw the involvement of an external sponsor. The correlation between rivalry and externally-supported insurgency is positive and quite strong, with a correlation coefficient of 0.567 (statistically significant and the $p < 0.01$ level). Although the number of observations is very small due to severe data limitations, the evidence suggests that, given that an insurgency exists, border provinces next to rivals are more likely to experience externally-sponsored insurgencies compared to border provinces not next to rivals.

Table A5.3: Rivalry and Externally-Supported Insurgency, 1989-2008

		Externally-Supported Insurgency		
		No	Yes	
Rivalry	No	13	10	23
	Yes	1	22	23
		14	32	46

Notes: Correlation coefficient $\rho=0.567$, $p<0.001$. Units are border province-decades. External support refers to support by a contiguous neighbor.

Having demonstrated an association between rivalry and externally-sponsored insurgency, I now turn to the question of how insurgency affects state authority. Using logged Myers scores as the proxy for state authority, I compare the difference in means in Myers scores between conflict-affected border provinces and non-conflict affected border provinces. Since higher values of the Myers Index indicate worse state authority, the conflict-affected provinces should have a larger mean Myers scores. Table A5.4 bears out this expectation, and a t-test of the difference in means indicates that the difference is statistically significant. However, readers should be cautioned against drawing overly strong conclusions due to the small number of conflict-affected provinces.

One implication of the theory is that insurgencies that involve a neighboring rival sponsor should be more harmful than insurgencies that involve a neighbor sponsor that is not a rival. Because enemy (rival) states are highly motivated to destabilize their targets for the purpose of policy gain, their involvement in insurgencies should have a more deleterious effect. I examine this proposition by identifying all cases of external sponsorship that involved

Table A5.4: State Authority Means, 1989-2008

	Conflict	No conflict
Means, logged Myers	2.609	1.603
Observations	46	998

Notes: T-test for difference in means: $p < 0.001$. Units are border-province decades.

a contiguous rival as opposed to a contiguous non-rival, and calculate the mean logged Myers score for each group (Table A5.5). Although there are only 32 province-decades that experience externally-sponsored insurgencies, rival-sponsored conflict results in worse levels of state authority on average. A t-test indicates that the difference is weakly significant ($p < 0.10$).

Table A5.5: State Authority Means, 1989-2008

	Rival-sponsored conflict	Other externally-sponsored conflict
Means, logged Myers	2.724	2.105
Observations	25	7

Notes: T-test for difference in means: $p = 0.0971$. Units are border-province decades.

Of course, as discussed in the main text, supporting armed conflict is not the only means of subversion, nor is subversion the only strategy through which enemy states can weaken domestic sovereignty in the pursuit of their foreign policy objectives. Thus, even if more data were available, a more sophisticated analysis would not capture the entirety of enemy state efforts to degrade state authority. Yet the evidence here, while suggestive, does lend some support to one of the proposed mechanisms: rivals are associated with externally-sponsored insurgencies, insurgencies are associated with weaker state authority, and rival-sponsored insurgencies are more harmful for state authority than insurgencies involving external sponsors that are not rivals. Taken in combination with the qualitative case study on the Philippines and Malaysia in the main text of the article, this exercise more directly implicates the enemy state in the degradation of state authority in target states.

A5.4 Robustness to Excluding Interventionary Rivalries

The [Thompson and Dreyer \(2012\)](#) dataset of interstate rivalry codes a residual category of rivalry that the authors call “interventionary” rivalry. As a residual category, it is not entirely clear what counts as an interventionary rivalry. The category seems to include both rivalries that are primarily about “acquiring leverage in the other state’s decision making” ([Thompson and Dreyer 2012](#), 21), but could also include cases in which hostilities arise because of cross-border raids. Complicating the matter is the fact that Thompson rivalries can have multiple types based on the underlying issues. It is thus possible that one issue or type (say, territorial disputes) in fact prompt the intervention that earns the rivalry the interventionary label. For these two reasons, interventionary rivalries may or may not represent a reverse causality issue.

To be sure that this category of rivalries are not driving the results, I have rerun the analysis and excluded any rivalry coded as an interventionary rivalry. Table [A5.6](#) shows the results. Since rivalries can have more than one type, this coding rule excludes some rivalries of other types, resulting in an observation loss of 114 province-years. Still, the results are robust to excluding interventionary rivalries.

Table A5.6: Robustness to Excluding Interventionary Rivalries

	(1)
Non-interventionary rivalries	0.0715* (0.0327)
Terrain ruggedness	-0.00311 (0.0166)
Distance from capital	0.0299 (0.0189)
Population density	-0.140** (0.0317)
Split ethnic group	0.0694* (0.0307)
Development	0.00825 (0.0213)
Natural resources	-0.0230* (0.0102)
Capital	-0.100* (0.0468)
Constant	-0.0857** (0.0158)
Observations	2001

Standard errors in parentheses and clustered by country-year

Country-year intercepts suppressed

⁺ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

A5.5 Hostile States and Political Geographies

[Herbst \(2000\)](#) argues that political geography affects the costs of exercising authority. My article builds on Herbst's work by highlighting the role of an additional underappreciated factor that raises the costs of exercising authority: interference from hostile neighboring states. Whereas Herbst focuses on domestic factors, like distance from the capital and population density, my explanation centers on an international factor. This is not to say that Herbst is wrong. Rather, as I note in the manuscript, this international factor influences the exercise of state authority above and beyond the important domestic factors suggested by Herbst.

Herbst's argument thus suggests a set of important domestic control variables for the analysis. The main analysis reported in the article therefore accounts for Herbst's argument in three ways. First, my sampling strategy accounts for effect of distance because the sample only includes border provinces. In other words, the sample is a subset of all provinces, and this subset contains provinces already likely to be distant from the capital. Second, as an additional check, I include distance from the capital as well as the capital-containing province as control variables. Third, I also include population density as a control variable. Hostile neighbors remain a statistically and substantively important predictor of low state authority even in the presence of these control variables.

As an additional check to ensure that hostile states rather than political geographies are responsible for my results, I subset my data into Herbst's political geographies. Following [Thies \(2009\)](#), I code the African states in my sample as having *difficult*, *neutral*, *hinterland*, or *favorable* political geographies. Because subsetting my data in this manner results in each subset having too few observations for regression analysis and too little statistical power, I instead turn to t-tests. If the patterns I observe in my data are primarily the result of political geography, then I should not observe differences in the mean Myers scores between border provinces next to enemy states and border provinces not next to enemy states within the Herbst political geographies. If, however, my argument holds and enemy states exert an additional harmful effect on state authority, then the differences should be apparent in the t-tests. Indeed, as [Table A5.7](#) shows, my expectations are borne out: even within the political geography subsets, border provinces next to enemy states have less state authority than border provinces not next to enemy states.

Table A5.7: Comparison of Effect of Hostile Neighbors by Herbst Political Geographies

Political geography	Border prov. not next to enemy state	Border prov. next to enemy state	Obs.	p-value
Difficult	2.75	3.36	131	p<0.01
Neutral	2.24	2.63	160	p<0.01
Hinterland	3.11	3.25	52	p=0.27
Favorable	2.54	2.83	143	p=0.11

Notes: Higher Myers scores indicate worse levels of state authority. Therefore, the prediction is that border provinces next to enemy state should have higher Myers scores compared to border provinces not next to enemy states.

A5.6 Terrain Ruggedness

Some readers may observe that in the main analysis, there appears to be no correlation between terrain ruggedness and state authority. While this result may be surprising at first blush, it is important to recall that the sample is restricted only to border provinces. Interior provinces are thus not included. Among the sample of border provinces, it may be that mountains and economic development do not have large enough effects to be detectable in the statistical analysis.

In the case of terrain ruggedness, the fact that the estimated coefficient is close to zero also suggests a substantive explanation. On the one hand, mountains create opportunities for rebels to base and refuges for recalcitrant populations to flee the reach of the state. This would suggest that mountains might be detrimental to the exercise of state authority, and the correlation between mountains and state authority should be negative. On the other hand, many centers of political power can be found among mountainous regions. This would suggest that the apparent correlation between terrain ruggedness and state authority is the *opposite* of expectation - that mountains and high state authority are positively correlated. Santiago, Chile and Thimphu, Bhutan are both high-altitude capitals. Other capitals sit in valleys or lowland areas but are located in *administrative regions* that are mountainous relative to other provinces in the same country: Kathmandu, Nepal; Bujumbura, Burundi; and Asunción, Paraguay are examples. All three sit in regions with levels of terrain ruggedness above their country means. In my sample of border provinces, 12 out of 53 capitals are located in border provinces with above average mountainousness relative to other border provinces in the same country. 11 of these capital-containing border provinces have levels of terrain ruggedness that are a half standard deviation above their country means; 5 are more than a full standard deviation above their country means. Since the capital-containing provinces should be among the best-governed units in a country, these border provinces exert a strong effect on the estimation of terrain ruggedness. When all of these possibilities are taken together, it likely explains why the overall effect appears to be a wash in the overall sample.

A5.7 Alternate Measure of Economic Development

In the main article regression results, there appears to be no correlation between economic development and state authority. One should be cautious about the interpretation of this apparent non-result. The lack of correlation is almost certainly an artifact of data limitations related to the measurement of economic development at the subnational level. In the main analysis, I operationalize economic development as “human economic activity.” Human economic activity is coded using satellite imagery of “land cover,” or the type of surface that covers the earth. Artificial surfaces, such as the kind found in cities, and agricultural cropland are examples of types of land cover that suggest human economic activity, particularly the type associated with economic development. We can contrast these land cover types with other surfaces, such as mangrove forests or desert, that are less likely to be associated with economic development. Using this data, I operationalize economic development as the percentage of territory in a province covered either by artificial surfaces or cropland. Because the source data come from satellite imagery, I am able to construct a measure that varies subnationally (Latham et al. 2014). However, this subnational variation comes at the cost of temporal variation; there is no satellite data that would allow me to create a time-varying measure of economic development. Nor is there a suitable alternative that yields both spatial and temporal variation.

Data limitations thus force a tradeoff between cross-sectional between-province variation and over-time variation. This is not ideal. It would be more optimal to have both spatial and temporal variation in a measure of economic development. If a tradeoff has to be made, privileging between-province variation over over-time variation is the appropriate choice for my analysis. Since the outcome of interest is a property of states that varies subnationally, and because an important observable implication of the theory that I test in the empirics of this paper also pertains to spatial variation within countries, it is both theoretically and methodologically appropriate to favor between-province variation if a tradeoff has to be made. If the tradeoff was made in the other direction - using national-level data with over-time variation - there would be a mismatch between the theory and the empirics, and I would lose the important analytic leverage that the between-province comparisons afford.

Consequently, my economic development proxy exhibits subnational variation but not temporal variation. I thus caution readers against interpreting the null finding of the relationship between state authority and economic development as a true null effect. It could be that because the variable is static, over-time changes in economic development vary sufficiently such that the kinds of land cover indicative of human economic activity is not predictive with precision. Another possibility, though unlikely, is that “human economic activity” is a poor proxy for economic development independent of the issue with time invariance.

To probe whether it is the time-invariant property of the variable or whether human economic activity is not a valid measure, I created a new proxy measure for economic development that measures the suitability of land for cultivation. Agriculture is a large and important component of economic output for many developing countries around the world. For example,

in Ethiopia in 2008, agriculture accounts for 43% of the country's GDP and about 80% of its exports ([Central Statistics Agency Ethiopia 2016](#)). In India, the agricultural sector employed about 70% of all Indian workers in 1961 and 55% of all of Indian workers in 2011 ([Datenet India Private Limited 2014](#)). Like GDP, no data exist to measure agricultural productivity for subnational units over time. However, we can estimate the level of agricultural output by looking at its inputs for which data do exist. Land suitability is one such input. Where the land is more suitable for agricultural activity, there should be greater levels of agriculture, which should correlate with economic development. In Ethiopia, for example, the arid Somali and Afar regions have the highest poverty rates in the country, while Gambella, a region with high land suitability, enjoys a significantly lower poverty rate ([Oxford Poverty and Human Development Initiative 2013](#)). I measure land suitability using geospatial data drawn from an index that is the product of two components reflecting the climatic and soil suitability for cultivation ([Ramankutty et al. 2002](#)). While soil suitability might be affected by technological advances like the Green Revolution, climatic suitability would not. The land suitability index ranges from 0 (unsuitable for agriculture) to 1 (highly suitable for agriculture).

To be clear, this *land suitability* measure is also time-invariant by construction, so while it cannot address the time invariance problem, it can help us learn whether the curious results related to human economic activity are due to poor operationalization. The results using the land suitability measure do not qualitatively differ from the results for economic activity (Table [A5.8](#)). The substantive effect is slightly larger, and the estimate is less imprecise, though it still falls short of conventional statistical significance. This result would lend greater weight to the interpretation that it is the time-invariance that is the problem, rather than a lack of a true effect.

Table A5.8: Regression Results with Alternate Proxies of Development

	(1)	(2)
Rivalry	0.0666* (0.0285)	0.0684* (0.0285)
Terrain ruggedness	-0.00117 (0.0160)	-0.00303 (0.0162)
Distance from capital	0.0286 (0.0183)	0.0289 (0.0180)
Population density	-0.133** (0.0308)	-0.133** (0.0296)
Spilt ethnic group	0.0706* (0.0291)	0.0709* (0.0292)
Human economic activity	0.00869 (0.0201)	
Land suitability		0.0150 (0.0115)
Natural resources	-0.0207* (0.00991)	-0.0201* (0.00994)
Capital	-0.103* (0.0455)	-0.102* (0.0453)
Constant	-0.0454** (0.0155)	-0.0461** (0.0155)
Observations	2115	2115

Standard errors in parentheses

Country-year intercepts suppressed

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

A6 Notes on the Myers Index

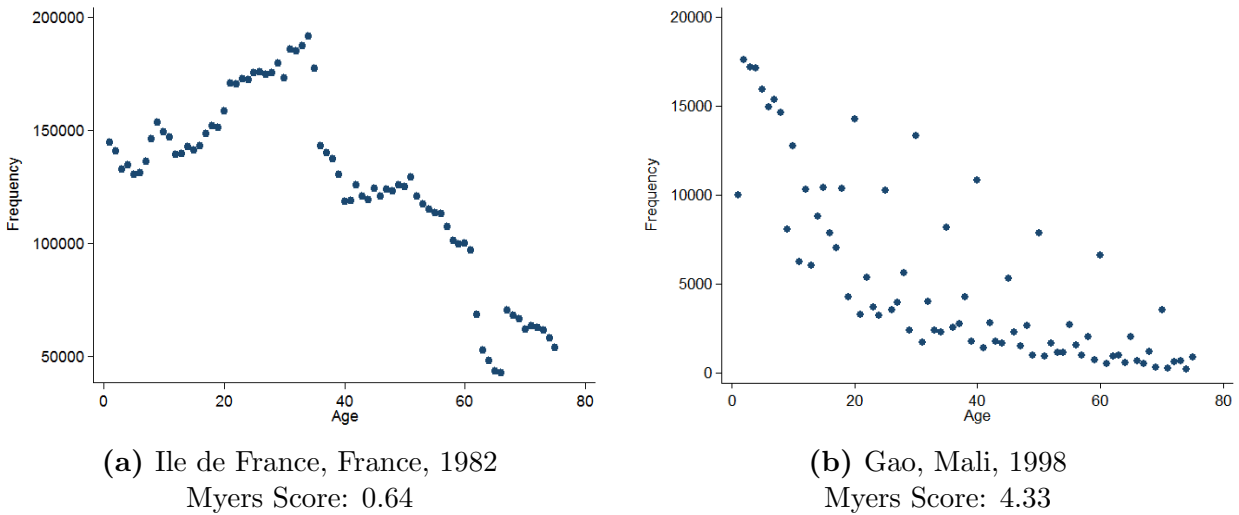
The Myers Index, which estimates the accuracy of age data reported in national population censuses, proxies for the dependent variable of spatial variation in state authority. The collection of accurate age data is a product of state authority. As described in the main text of the article and more fully in [REDACTED], there are two sources of error that are related to state authority: enumerator error and respondent error. Enumerator error is a result of *systematic* shirking by census interviewers, which typically occur in response to poor state-related conditions on the ground, such as a lack of infrastructure or physical insecurity. Respondent error typically occurs when individuals do not know their precise ages; precise age in quantitative terms is only relevant in societies where the state regulates rights, privileges, and responsibilities on the basis of age.

An important feature of the Myers Index is that it is sensitive to conflict without being biased by the long-run consequences of conflict, such as population displacement or a reduced youth population lost during fighting. Neither of these two possibilities are likely to be a source of bias; rather, the effect of conflict on the Myers Index is likely to manifest through the enumerator error effect. Consider first population displacement. Population displacement out of a unit should occur for some or all age cohorts; it is implausible that only individuals whose true ages end in particular terminal digits would leave a unit due to conflict. Instead, population displacement effectively reduces the overall size of the population in that unit. Fortunately, a particularly useful property of the Myers Index is that it is not sensitive to population size except in the case of very small populations (less than 2,000 individuals) with very small expected true error rates (less than 5%). Since these conditions never obtain in my data, even for provinces affected by conflict, population displacement is unlikely to introduce bias.

As for the youth population, while it is true that conflicts often have devastating effects on youth, a reduction in the youth population changes the shape of the age distribution curve but does not introduce significant numbers of discontinuities or “noise” into the distribution. Recall that systematic errors detectable by the Myers Index manifest as large percentage changes in the number of individuals in consecutive single ages. Graphically, these errors appear as noise around the “true” underlying distribution. It is this noise that is captured by the Myers Index; the overall shape of the underlying distribution is flexible. To see an example, consider Figure A6.1a, which shows the age distribution curve for the Île-de-France (the region containing the French capital of Paris) in 1982. The curve shows two sharp discontinuities. The discontinuity around ages 60–65 represents the loss of the generation of young men who died in World War II: men who were in their 20s during the war and who would have been in their 60s had they survived to 1982. The discontinuity around age 37 represents the post-war baby boom, as individuals born in 1945 would be in their late 30s and early 40s in 1982. Both of these discontinuities are examples of demographic shocks related to conflict; the former exemplifies the possibility raised by R2. Notice, however, that the curve does not exhibit any of the noise around the true distribution that is characteristic

of systematic error in populations. This absence of noise in the Île-de-France distribution in Figure A6.1a is especially apparent when contrasted with the distribution for Gao, a region of low state authority in Mali, in Figure A6.1b. Notice also that the Myers score for France is very small, reflecting the fact that the French data exhibit low age misreporting and thus high state authority in the Île-de-France region. Thus, the Myers Index is not likely to be biased from the decimation of the youth population from fighting.

Figure A6.1: The Effect of Demographic Shocks



A7 Notes on Measuring Hostility

The article operationalizes hostility between neighbors as an interstate rivalry based on definitions and data from [Thompson and Dreyer \(2012\)](#). The choice of the Thompson approach is deliberate. The two leading datasets on international rivalry are the Thompson (“T”) dataset and the [Klein, Goertz, and Diehl \(2006\)](#) (“KGD”) dataset. Although both datasets are commonly used, they represent significantly different approaches to operationalizing the concept. [Klein, Goertz, and Diehl \(2006\)](#) describe these disagreements as “considerable and non-trivial.” This fundamental disagreement on both theory and measurement results in the surprising lack of overlap in cases. The main difference between these approaches that is relevant for my article is that T codes rivalries based on who state leaders or their historians say are their enemies, competitors, or sources of threat, while KGD codes on certain types of revealed behavior - essentially dispute density (MIDs).

However, by coding rivalries based on only certain types of revealed behavior, KGD actually code for the kinds of behaviors substituted for by political interference in domestic sovereignty. What this means is that if subversion in sovereignty is useful because it does not engage states in direct, conventional militarized hostilities - the types captured in the Militarized Interstate Disputes dataset - then one would not expect to observe a relationship between KGD rivals and state authority in target states because MIDs are substitutes for political interference.

In contrast, the T data provide a more appropriate operationalization approach because T rivalries are coded based on leader perceptions. My theoretical argument is that states that have some relatively high degree of disagreement on issues are likely to look for ways to advance their interests on the issues under disputes. It is only this set of states with highly antagonistic relationships that should subvert domestic sovereignty, and therefore antagonism should be the starting point for operationalization. The perceptions-based approach in [Thompson and Dreyer \(2012\)](#) captures this sense of antagonism. As [Thompson and Dreyer \(2012\)](#) describe, T rivals are the states that State A’s leaders *say* are their rivals. For these reasons, I have chosen to operationalize hostility using the T definition and data of rivalry.

References

- Badan Pusat Statistik. 2014. *Badan Pusat Statistik*. Jakarta: Republik Indonesia.
URL: <http://www.bps.go.id/>
- Bertelsmann Stiftung. 2014. *The Bertelsmann Stiftung's Transformation Index*. Gutersloh: Bertelsmann Stiftung.
URL: <http://www.bti-project.org/index/>
- Bignami-Van Assche, Simona. 2005. "Province-Specific Mortality in China, 1990-2000." Paper prepared for the 2005 Annual Meeting of the Population Association of America, Philadelphia, PA.
- Cederman, Lars-Erik, Kristian Skrede Gleditsch, Idean Salehyan, and Julian Wucherpfennig. 2013. "Transborder Ethnic Kin and Civil War." *International Organization* 67 (2): 389–410.
- Central Statistics Agency Ethiopia. 2016. *CountryStat Ethiopia*. Central Statistics Agency Ethiopia.
URL: <http://www.countrystat.org/home.aspx?c=ETH>
- China Data Center. 2014. *China Data Online*. Ann Arbor: China Data Center.
URL: <http://chinadataonline.org/>
- Datanet India Private Limited. 2014. *IndiaStat*. Datanet India Private Limited.
URL: <http://www.indiastat.com>
- Dittrich Hallberg, Johan. 2012. "PRIO Conflict Site 1989-2008: A Geo-Referenced Dataset on Armed Conflict." *Conflict Management and Peace Science* 29 (2): 219–232.
- Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2013. *The Next Generation of the Penn World Table*. Groningen: University of Groningen.
URL: <http://www.ggdc.net/pwt>
- Fund for Peace. 2014. *Fragile States Index*. Washington, D.C.: The Fund for Peace.
- Goldstone, Jack A., Ted Robert Gurr, Barbara Harff, Marc A. Levy, Monty G. Marshall, Robert H. Bates, David L. Epstein, Colin H. Kahl, Pamela T. Surko, John C. Ulfelder Jr., and Alan N. Unger. 2000. *State Failure Task Force Report: Phase III Findings*. McLean: Science Applications International Corporation.
- Hainmueller, Jens, Jonathan Mummolo, and Yiqing Xu. 2016. "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." Available at SSRN.
URL: <http://papers.ssrn.com/abstract=2739221>
- Herbst, Jeffrey. 2000. *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton: Princeton University Press.

- Högbladh, Stina, Therése Pettersson, and Lotta Themnér. 2011. "External Support in Armed Conflict 1975-2009: Presenting New Data." Paper presented at the 52nd Annual International Studies Association Convention, Montreal, Canada, 16-19 March, 2011.
- Institute for Health Metrics and Evaluation. 2011. *Child Mortality Estimates and MDG4 Attainment by Country 1990-2011*. Seattle: Institute for Health Metrics and Evaluate.
- Institute of Applied Economic Research. 2014. *IPEADData*. Rio de Janeiro: Institute of Applied Economic Research.
URL: <http://www.ipeadata.gov.br/>
- Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2014. *Worldwide Governance Indicators*. Washington, D.C.: The World Bank Group.
- Klein, James P., Gary Goertz, and Paul F. Diehl. 2006. "The New Rivalry Dataset: Procedures and Patterns." *Journal of Peace Research* 43 (3): 331-348.
- Latham, John, Renato Cumani, Ilaria Rosati, and Mario Bloise. 2014. *Global Land Cover SHARE*. Rome: Food and Agriculture Organization of the United Nations.
- Lujala, Päivi, Jan Ketil Rød, and Nadia Thieme. 2007. "Fighting over Oil: Introducing a New Dataset." *Conflict Management and Peace Science* 24 (3): 239-256.
- Michalopoulos, Stelios and Elias Papaioannou. 2013. "The Long-Run Effects of the Scramble for Africa." Working Paper, Brown University and London Business School.
URL: <https://docs.google.com/file/d/0B4s-WKe-US99V2VyRjc4cksxWWs/edit>
- Nunn, Nathan and Diego Puga. 2012. "Ruggedness: The Blessing of Bad Geography in Africa." *The Review of Economics and Statistics* 94 (1): 20-36.
- Oxford Poverty and Human Development Initiative. 2013. *Multidimensional Poverty Index (MPI) Data Bank*. OPHI, University of Oxford.
URL: www.ophi.org.uk/multidimensional-poverty-index
- Patterson, Tom. 2013. "CleanTOPO2: Edited SRTM30 Plus World Elevation Data." Available from <http://www.shadedrelief.com/cleantopo2/>.
- Pettersson, Therése. 2011. "Pillars of Strength - External Support to Warring Parties." In *States in Armed Conflict 2010: Research Report 94*, ed. Therése Pettersson and Lotta Themnér. Uppsala: Universitetsstryckeriet.
- Planning Commission. 2002. *National Human Development Report 2001*. New Delhi: Government of India.
URL: <http://planningcommission.nic.in/reports/genrep/index.php?repts=nhdcont.htm>
- PRS Group. 2014. *International Country Risk Guide*. East Syracuse: PRS Group.
- Ramankutty, Navin, Jonathan A. Foley, John Norman, and Kevin McSweeney. 2002. "The Global Distribution of Cultivable Lands: Current Patterns and Sensitivity to Possible

- Climate Change.” *Global Ecology and Biogeography* 11 (5): 377–392.
URL: <http://www.sage.wisc.edu/iamdata/grid.php>
- Singer, J. David. 1987. “Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816-1985.” *International Interactions* 14 (2): 115–132.
- Thies, Cameron G. 2009. “National Design and State Building in Sub-Saharan Africa.” *World Politics* 61 (4): 623–669.
- Thompson, William R. and David R. Dreyer. 2012. *Handbook of International Rivalries: 1494-2010*. Los Angeles: CQ Press.
- Turkish Statistical Institute. 2014. “Gross Domestic Product by Expenditure Approach, 1987 Base.” Available from <http://www.turkstat.gov.tr/UstMenu>.
- Weidmann, Nils B., Jan Ketil Rød, and Lars-Erik Cederman. 2010. “Representing Ethnic Groups in Space: A New Dataset.” *Journal of Peace Research* 47 (4): 491–99.
- World Bank. 2014. *World Development Indicators*. Washington, D.C.: The World Bank.