Online Appendix
Productive Pacifists: The Rise of Production-Oriented States and Decline of Profit-Motivated Conquest

Jonathan N. Markowitz* Suzie Mulesky† Benjamin A.T. Graham‡ Christopher J. Fariss§

June 12, 2020

Contents

1 Robustness Checks

This section summarizes the results of 186 different model specifications, including those presented in the main body of the paper. By exploring a wide range of alternative measurement and estimation strategies, we are able to comprehensively show that the support we find for our theory is not driven by the idiosyncrasies in how we measure particular variables or how we specify the regression models.

We have 4 different versions of a land-oriented variable (3 versions of the binary variable and 1 version of the continuous variable), 10 different dependent variables (5 binary and count measures each), two versions of autocracy (one from Polity IV and the other from Boix, Miller and Rosato (2013)) and two sets of control variables (one with a set of primary control variables and one with a set of alternative controls). This totals to 160 unique regressions. See Figures 3 and 4 from the main body of the paper for visual summaries of the findings from these 160 models.

We estimate an additional 26 models by probing the robustness of our findings from Table 1 from the main body of the paper. First, we re-estimate Models 1 and 3 from Table 1 by splitting the sample into autocracies only and democracies only (4 models, see Table 2 in the main body of the paper). These results demonstrate that our results are not driven by democracies or autocracies alone. The effect of land-orientation on conflict obtains across regime types.

Second, we re-estimate Models 1 and 3 from Table 1 by measuring the independent effects of agricultural dependence and natural resource dependence (6 models, see Table 3 in the main body of the paper). These results demonstrate that it is land-orientation, and not just oil, that leads states to seek territory.

Third, we subset the analysis, re-estimating Models 1-4 from Table 1 on observations prior to the start of WWII using both binary and continuous versions of land-orientation (4 models, see Table A2) as well as after the end of WWII (4 models, see Table A3). We find that land-orientation is positively related to territorial competition both prior to and after WWII. However, while land-orientation is statistically significant in all models post-WWII ($p<0.5$), it is only statistically significant in one model pre-WWII (Model 4).

*School of International Relations, University of Southern California, jonathnm@usc.edu
†Political Science and International Relations, University of Southern California, sjcaldwe@usc.edu
‡School of International Relations, University of Southern California, benjamag@usc.edu
§Political Science, University of Michigan, cjf0006@gmail.com

Fourth, we control for country fixed effects (4 models, see Table A4). While land-orientation is positively associated with territorial competition in all four models, only one is statistically significant (binary land-orientation on resource-based territorial claims).

Table A1: Alternative control variables

	Resource-based territorial claim (binary)		Resource-based territorial MID (binary)	
	(1)	(2)	(3)	(4)
Land-orientation (binary)	1.851***		0.663**	
	(0.625)		(0.270)	
Land-orientation (continuous)		0.202***		0.095
		(0.056)		(0.066)
Autocracy (binary)	0.399	0.421	0.428**	0.424**
	(0.389)	(0.412)	(0.198)	(0.204)
CINC Score	5.436	1.844	−6.301*	−6.775**
	(7.111)	(7.220)	(3.391)	(3.394)
GDP per capita, log	0.183	−0.026	0.332**	0.178
	(0.227)	(0.215)	(0.136)	(0.120)
Population, log	0.294**	0.364**	0.544***	0.541***
	(0.147)	(0.155)	(0.080)	(0.081)
Island (dummy)	0.332	0.697	0.878***	0.952***
	(0.564)	(0.565)	(0.259)	(0.262)
Trade (% of GDP)	0.824**	0.812**	−0.405***	−0.350**
	(0.366)	(0.373)	(0.155)	(0.154)
Oil & gas producing neighbors	0.208	0.277*	0.079	0.115
	(0.144)	(0.142)	(0.077)	(0.075)
Time count	−1,410.970	−1,612.910	−2,152.071***	−2,557.839***
	(951.913)	(1,070.513)	(483.108)	(514.089)
Time count^2	713.250	815.506	1,086.735***	1,291.091***
	(480.502)	(540.226)	(243.738)	(259.350)
Time count^3	−12.019	−13.745	−18.292***	−21.723***
	(8.085)	(9.087)	(4.099)	(4.361)
Constant	930,416.200	1,063,352.000	1,420,563.000***	1,689,123.000***
	(628,592.100)	(707,097.000)	(319,178.900)	(339,672.000)
Observations	6,124	5,974	6,124	5,974
Significance levels				*p<0.1; **p<0.05; ***p<0.01

Table A2: Pre-WWII

	Resource-based territorial claim (binary)		Resource-based territorial MID (binary)	
	(1)	(2)	(3)	(4)
Land-orientation (binary)	0.748		1.045	
	(0.571)		(0.832)	
Land-orientation (continuous)		0.277		0.946**
		(0.429)		(0.413)
Autocracy (binary)	−0.245	−0.286	0.776***	0.634*
	(0.276)	(0.301)	(0.300)	(0.328)
Military personnel, log	0.769***	0.848***	0.244*	0.189
	(0.142)	(0.151)	(0.134)	(0.140)
Military expenditures, log	0.198	0.119	−0.150	−0.158
	(0.142)	(0.152)	(0.123)	(0.138)
GDP per capita, log	0.335	0.580	0.137	0.893***
	(0.217)	(0.440)	(0.135)	(0.344)
Population, log	−0.483***	−0.482***	0.102	0.009
	(0.175)	(0.186)	(0.152)	(0.169)
Neighbors	−0.008	−0.008	0.018	0.023*
	(0.011)	(0.010)	(0.013)	(0.012)
Island (dummy)	0.621	0.811	−0.288	−0.177
	(0.751)	(0.759)	(0.735)	(0.739)
Time count	−61.021*	−45.112	−31.926	−52.901
	(33.078)	(39.836)	(39.275)	(43.176)
Time count^2	32.361*	23.962	17.074	28.242
	(17.592)	(21.183)	(20.823)	(22.911)
Time count^3	−0.572*	−0.424	−0.304	−0.502
	(0.312)	(0.375)	(0.368)	(0.405)
Constant	38,336.360*	28,294.240	19,877.410	33,003.080
	(20,728.600)	(24,967.550)	(24,688.850)	(27,117.730)
Observations	3,192	2,946	3,192	2,946
Significance levels				*p<0.1; **p<0.05; ***p<0.01

Table A3: Post-WWII

	Resource-based territorial claim (binary)		Resource-based territorial MID (binary)	
	(1)	(2)	(3)	(4)
Land-orientation (binary)	1.309***		0.486**	
	(0.475)		(0.200)	
Land-orientation (continuous)		0.224***		0.199***
		(0.049)		(0.036)
Autocracy (binary)	0.886***	1.091***	−0.115	−0.261
	(0.315)	(0.352)	(0.163)	(0.163)
Military personnel, log	0.039	−0.080	0.393***	0.534***
	(0.106)	(0.160)	(0.080)	(0.096)
Military expenditures, log	−0.059	0.017	−0.013	−0.119**
	(0.040)	(0.107)	(0.043)	(0.047)
GDP per capita, log	0.074	0.168	−0.095	0.001
	(0.140)	(0.162)	(0.059)	(0.064)
Population, log	0.197	0.315**	0.146*	0.165**
	(0.130)	(0.153)	(0.078)	(0.081)
Neighbors	0.008	−0.005	−0.032	−0.030
	(0.031)	(0.029)	(0.020)	(0.019)
Island (dummy)	0.641	0.646	0.599***	0.831***
	(0.395)	(0.470)	(0.231)	(0.240)
Time count	−674.115**	−667.043**	−404.000**	−433.657**
	(297.105)	(320.851)	(175.504)	(190.783)
Time count^2	341.536**	337.816**	205.323**	220.341**
	(150.528)	(162.459)	(88.791)	(96.477)
Time count^3	−5.768**	−5.703**	−3.478**	−3.732**
	(2.542)	(2.742)	(1.497)	(1.626)
Constant	443,515.600**	439,047.600**	264,965.100**	284,490.000**
	(195,463.700)	(211,215.100)	(115,629.700)	(125,753.600)
Observations	7,998	7,318	7,998	7,318
Significance levels				*p<0.1; **p<0.05; ***p<0.01

Table A4: Country Fixed Effects

	Resource-based territorial claim (binary)		Resource-based territorial MID (binary)	
	(1)	(2)	(3)	(4)
Land-orientation (binary)	1.675***		0.079	
	(0.411)		(0.241)	
Land-orientation (continuous)		0.108		0.028
		(0.114)		(0.073)
Autocracy (binary)	−0.369	−0.399	−0.370*	−0.481**
	(0.267)	(0.296)	(0.191)	(0.205)
Military personnel, log	0.627***	0.657***	0.218*	0.366***
	(0.116)	(0.144)	(0.116)	(0.130)
Military expenditures, log	−0.105*	0.060	0.045	0.004
	(0.060)	(0.112)	(0.070)	(0.083)
GDP per capita, log	−0.307	−0.738***	−0.437***	−0.525***
	(0.235)	(0.269)	(0.154)	(0.167)
Population, log	−1.064***	−1.292***	−0.165	−0.220
	(0.266)	(0.292)	(0.183)	(0.192)
Neighbors	0.050***	0.032***	0.038**	0.034**
	(0.013)	(0.012)	(0.016)	(0.016)
Island (dummy)	−0.682	−0.127	1.360**	2.569**
	(0.744)	(1.034)	(0.611)	(1.064)
Observations	11,393	10,341	11,393	10,341
R^2	0.012	0.009	0.003	0.004
Max. Possible R^2	0.130	0.116	0.170	0.169
Wald Test (df = 8)	88.850***	76.100***	27.210***	34.850***
LR Test (df = 8)	132.830***	90.476***	29.232***	39.142***
Score (Logrank) Test (df = 8)	117.656***	85.897***	29.131***	37.770***
Significance levels				*p<0.1; **p<0.05; ***p<0.01

2 Independent Variable: Measuring Economic Rent Structure

We rely on a number of data sources to compile our measure of land-orientation. We measure dependence on agriculture and natural resources separately and then combine both measures into a binary and continuous variable.

2.1 Agricultural Dependence

We measure agricultural dependence directly from 1960 to 2016 using the World Bank's data on agriculture value added as a percentage of GDP. From 1800 to 2016, we use historic data on agriculture value added (% of GDP) for 43 countries, sourced from Our World in Data.[A1] These historic estimates are based on two sources: (1) Herrendorf, Rogerson and Valentinyi (2014) and (2) the 2015 release of the Groningen Growth and Development Centre's 10-Sector Database.[A2] Both data sources can be found online (see footnotes). Most of the values from Our World in Data occur every ten years for the early industrializers (e.g. United Kingdom, United States, Belgium), and the majority of the dataset covers agriculture value added starting in 1950. This dataset goes some way to help us measure when the early industrializers transitioned from agriculture to production. However, we still require a more complete measure to capture agricultural dependence in the early to mid 1900s.

To accomplish this, we impute missing values for agriculture value added (% of GDP) using a multiple imputation model. We utilize seven variables that are all highly correlated with agricultural dependence and six of which provide the benefit of historic time coverage: share of employment in agriculture (two data sources), historic GDP per capita, energy consumption, and urban population (in addition to our two variables that measure agriculture value added directly). See Figure A1 for the correlation matrix.

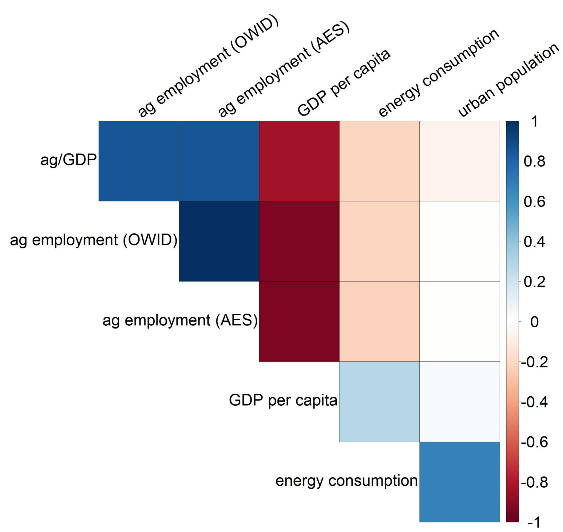

Figure A1: Pearson correlation coefficients.

We obtain historic data on agricultural employment from Our World in Data and Wingender (2014). Our World in Data measures employment share by economic sector, including agriculture, for 44 countries from 1800 to 2016.[A3] As with agriculture value added (% of GDP), the historic estimates for agricultural employment are compiled from two data sources: (1) Herrendorf, Rogerson and Valentinyi (2014) and (2) the 2015 release of the Groningen Growth and Development Centre's 10-sector Database. Wingender (2014) has compiled historic estimates of agricultural employment for 169 countries from 1900 to 2010. Wingender compiles employment data sourced from multiple organizations (e.g. the International Labor Organization, national statistical offices) and supplements the dataset with urbanization rates when employment levels are unobserved.

We use the updated historic latent estimates of GDP per capita developed by Anders, Fariss and Markowitz (2020) to measure wealth, which covers 227 countries from 1500 to 2015. We use data from the Correlates of War Project's National Material Capabilities dataset version 5.0 to measure urban population and energy consumption (Singer, Bremer and Stuckey, 1972; Singer, 1987). Both variables cover observations from 1816 to 2012. See Table A5 for a list of data sources and time coverage.

We use Amelia II software developed by Honaker, King and Blackwell (2011) to conduct the imputation.

[A1] https://ourworldindata.org/grapher/GDP-vs-agriculture-GDP?country=BOL+BWA+BRA+CHL+CHN+COL+CRI+DNK+EGY+ETH+FRA+DEU+GHA+IND+IDN+ITA+JPN+KEN+MWI+MYS+MUS+MEX+MAR+NLD+NGA+PER+PHL+SEN+SGP+ZAF+KOR+ESP+SWE+TWN+TZA+THA+GBR+USA+VEN+ZMB (accessed January 17, 2019)

[A2] https://www.rug.nl/ggdc/productivity/10-sector/ (accessed January 17, 2019)

[A3] https://ourworldindata.org/grapher/GDP-vs-agriculture-employment accessed January 17, 2019.

Table A5: Agricultural Dependence: Data Sources for Input Variables

Variable	***Data Source***	***Time coverage***
Agriculture, value added (% of GDP)	World Bank World Development Indicators	1960-2016
Agriculture, value added (% of GDP)	Our World in Data	1800-2016
Share of employment in the agricultural sector	Wingender (2014)	1900-2010
Share of employment in the agricultural sector	Our World in Data	1801-2015
GDP per capita	Anders, Fariss, and Markowitz (2019)	1500-2015
Urban population	Correlates of War, National Material Capabilities v 5.0	1816-2012
Energy consumption	Correlates of War, National Material Capabilities v 5.0	1816-2012

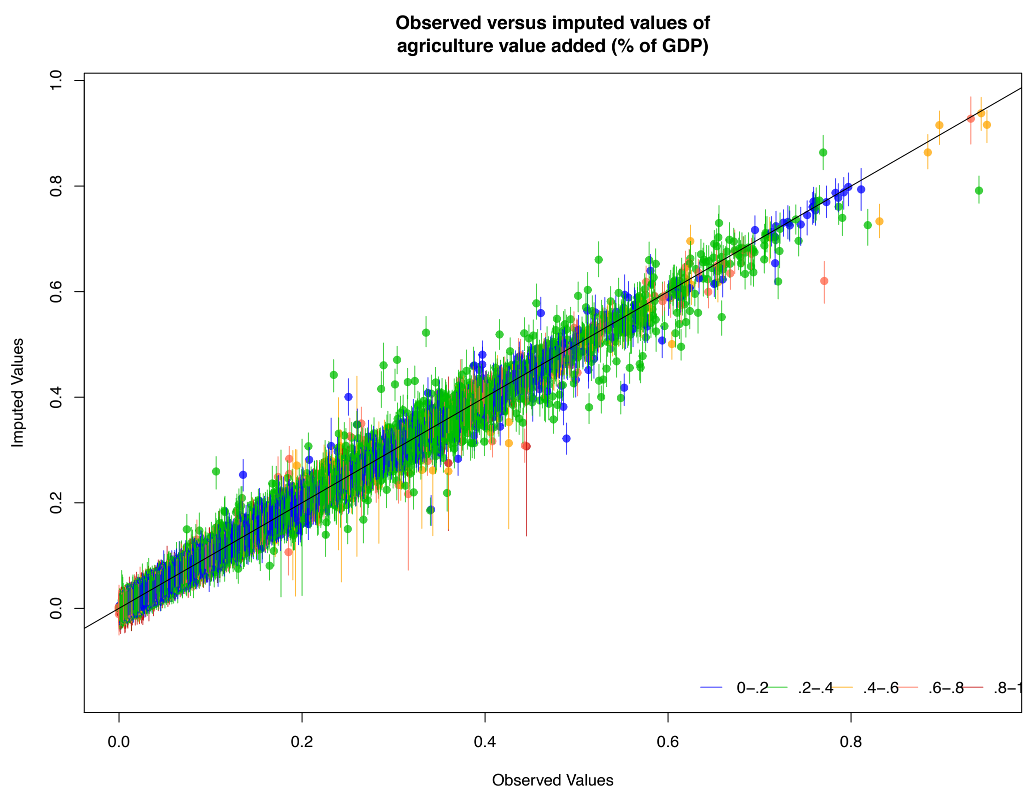

Figure A2: Observed versus imputed values of agriculture value added (share of GDP). Agriculture value added is expressed as a proportion. A $y = x$ line is fitted to demonstrate what perfect agreement would look like. Each circle is the mean of all the imputations for that value, and each circle has a corresponding 90% confidence interval. The colors of the data points reflect the percentage of covariates *not* observed, i.e. the pattern of missingness among the covariates. Blues and greens indicate there was more information available to impute the missing value. Red and orange indicate there was very little information available.

Amelia II uses expectation-maximization with bootstrapping to impute missing values using information from multiple covariates. We create five sets of estimates and average across the five to create a single estimate of agriculture value added (% of GDP). We then take a rolling 5-year average of the estimate to smooth the estimates over time. We allow the imputation algorithm to use both lagged and lead values for each country cross-section in order to improve the predictive validity of the model. In addition, we transform all non-proportion covariates using the natural logarithm to help mitigate any violation of the normality assumption.[A4] We keep agriculture value added (% of GDP) and share of agriculture in employment expressed as proportions. We bound estimates for all covariates within specific numeric ranges so that we do not generate implausible values. We bound estimates for agriculture value added and share of agriculture in employment between 0 and 1. We bound estimates for energy consumption and urban population between a number very close to zero and positive infinity. GDP per capita estimates are bounded between the natural log of 15 (near the empirical minimum) and positive infinity. Finally, we specify that the imputation model estimates missing values within country cross-sections using Gleditsch and Ward's country identification system (Gleditsch and Ward, 1999).

Overall, the model appears to perform well when estimating observations between 1960 and 2016. This is the time-range in which we can directly evaluate the accuracy of the estimates by comparing the imputed values with observed values. The developers of Amelia II have created a technique called "overimputing" to evaluate the accuracy of the imputed values. Overimputing treats each observed value as if it were missing, generating hundreds of imputed values for each observed value in order to construct 90% confidence intervals. Figure A2 plots the overimputed values against the observed values for agriculture value added (share of GDP). A $y = x$ line indicates the line of perfect agreement: if the imputation model perfectly predicted the true value, all imputed values would fall on this line. As demonstrated in Figure A2, this is nearly so.

To create a binary measure for agricultural dependence, we use 15% of GDP as the cut-off to separate agrarian from industrial economies. This threshold is somewhat arbitrary, so we test our models using two additional thresholds: 10% (our low threshold) and 20% (our high threshold). Economies in which agriculture contributes 15% or more to GDP are considered agriculturally dependent. Some countries bounce around the threshold for many years, so we consider a country as non-agriculturally dependent during the *final* year it crosses below 15%.

The choice of 15% is based on research from economic historians and development economists who consider a state to have successfully industrialized when the percentage of the population employed in agriculture falls below 25% of the working population (Ayuda, Collantes and Pinilla, 2010). We validate our 15% threshold by showing that agriculture value added as a share of GDP at the time each country crosses the 25% employment threshold is close to 15%. Figure A3 documents the distribution of agriculture value added as a share of GDP for the first and last year a country falls below 25% in agricultural employment.[A5] While there is variation in agriculture value added when employment drops below 25%, the distribution is centered around 12.8% (median), which is very close to our threshold. Further, most of the distribution sits between 10% and 20% of agriculture value added as a percentage of GDP, and these percentages represent the low and high values of our threshold.

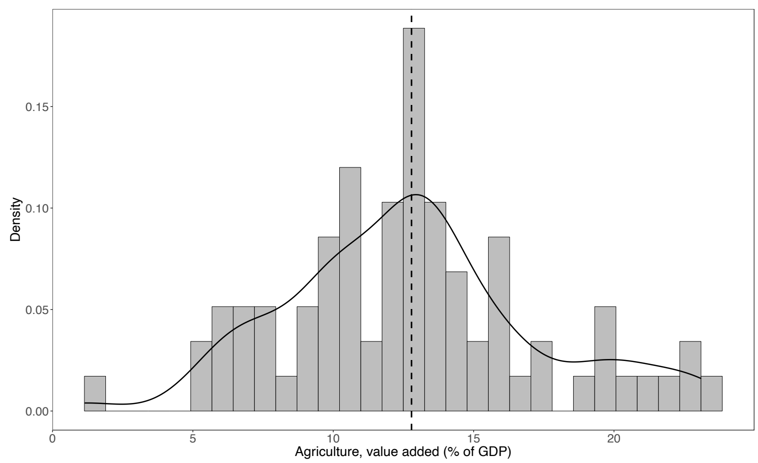

Figure A3: Distribution of agriculture value added as a share of GDP during the first and last year a country falls below 25% of agricultural employment. The dashed line represents the median value, at 12.79%.

We also compare our threshold to a secondary method of measuring industrial transition: the year the share of employment in the industrial sector supersedes the agricultural sector (Bentzen, Kaarsen and Wingender, 2013). In Table A6, we show the median difference — in terms of number of years — between the year we mark a country as industrializing and the year Bentzen, Kaarsen, and Wingender do so. For early industrializers (countries marked as having industry employment exceed agricultural employment before 1900), we tend to mark the country as industrializing much later than Bentzen, Kaarsen, and

[A4]Amelia's imputation model assumes that the complete data are multivariate normal.

[A5]Countries will often fluctuate around the threshold for several years before finally crossing over. Thus, we look at both the first and last year a country falls below 25% in agricultural employment.

Wingender do, which we believe is justifiable. For example, they mark the United Kingdom's industrial transition in 1801, while we mark it in 1898. In terms of the early industrializers, our measure is more conservative. However, the difference between our measures dissipates at the turn of the 20th century. The median difference in number of years for the 2nd wave of industrializers is 6 years and 0 years for the third wave.[A6]

Table A6: Comparing Our Measure of Agricultural Dependence with Bentzen, Kaarsen and Wingender (2013)'s Measure of Industrialization

	Difference in years between measures (median)
First wave (before 1900)	50.5 years
Second wave (1900-1959)	6 years
Third wave (1960-2004)	0 years

2.2 Natural Resource Dependence

To directly measure a country's reliance on rents from natural resources, we use data from the World Bank on total natural resources rents as a percentage of GDP (Bank, 2017). This indicator is a sum of oil, natural gas, coal, mineral, and forest rents divided by a country's GDP. The World Bank defines economic rents as the revenue above the cost of extraction. Data availability begins in 1970. To measure dependence on natural resources prior to 1970, we use data on the value of oil and natural gas production from Ross and Mahdavi (2015). This dataset covers most states in the international system from 1932 to 2014. Prior to 1932, oil and gas production were quite low globally, and we consider no state to be land-oriented due to petroleum dependence. To calculate the value of oil and gas production, Ross and Mahdavi multiply the volume of production by the world price for oil or gas. Values are in nominal dollars per million British Thermal Units of natural gas priced at the Henry Hub. For both resource rents and oil & gas production, we take a 5-year trailing average to account for short-term fluctuations in oil prices.

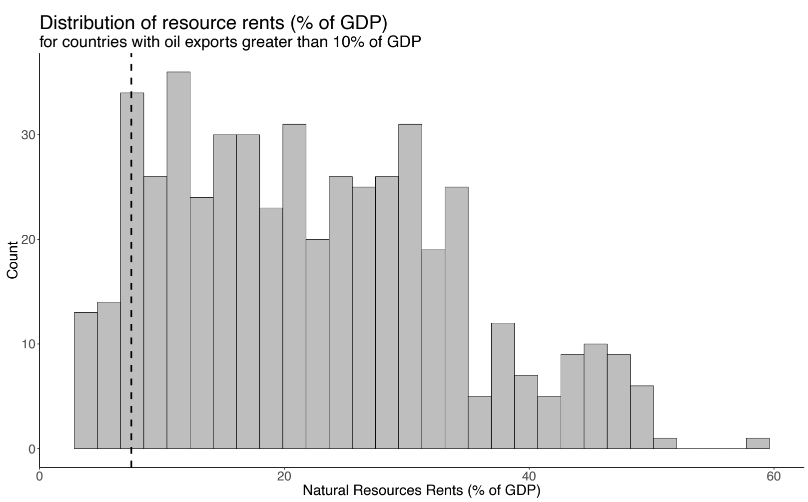

Figure A4: Distribution of natural resources rents as a share of GDP for countries with oil exports exceeding 10% of GDP. The dashed line represents our chosen threshold for resource rents, at 7.5% of GDP.

Countries may be dependent on the mining sector but *not* dependent on agriculture, oil, or natural gas prior to 1970, before the World Bank's natural resources rents variable becomes available. To code for mining dependence, we first examined which countries have mineral rents (% of GDP) greater than 5% in the 1970s, according to data from the World Bank's World Development Indicators (see rationale for threshold below). Only 12 countries make this list,[A7] but all of them stay above the 15% ag/GDP threshold before 1970, except for Chile which crosses below 15% in 1948 and stays below. To determine whether Chile was dependent on its mining sector from 1948 to 1972 (when World Bank data becomes available for Chile), we referred to data on Chile's ores and metals exports (as a percentage of merchandise exports), sourced from the World Bank's World Development Indicators. From 1962 to 1973, ores and metals accounted for 82% to 89% of total merchandise exports.[A8] For data

[A6]Their data is censored after 2005.
[A7]Botswana, Chile, DRC, Guyana, Jamaica, Liberia, Mauritania, Morocco, Papua New Guinea, Suriname, Togo, and Zambia.
[A8]https://data.worldbank.org/indicator/TX.VAL.MMTL.ZS.UN?locations=CL&view=chart (accessed January 17, 2019)

from the 1950s, we referenced multiple secondary sources on Chile's mining sector. For example, Moran (1975) notes that copper production from two companies alone accounted for 7% to nearly 20% of GDP. As a result, we code Chile as dependent on natural resources from 1948 to 1971 for the binary resource-dependence measure. However, because we do not have continuous data on Chile's mining-dependence during this interval, we keep its continuous measure missing from 1948 to 1971.

To guide our threshold for natural resource rents, we turn to the extant literature on petrostates. A state is generally considered a petrostate if its gross revenues from net oil exports exceeds 10% of its GDP in a given year (Colgan, 2013, 2, 48). We use this benchmark to set the threshold for natural resource rents as a share of GDP, which happens to be around 7.5%. Figure A4 shows that 92% of the distribution of resource rents (% of GDP) falls above 7.5% for countries with oil exports accounting for more than 10% of GDP. Data on oil exports are sourced from Emma Ashford's Oil Exports dataset contained in the IPE Data Resource (Graham and Tucker, 2017).

To create a binary measure of natural resource dependence, we use 7.5% of GDP as the cut-off to separate petrostates from non-petrostates starting in 1970 and 10% of GDP as the cut-off for observations between 1932 and 1970. We use different thresholds because one variable measures rents while the other measures output. Thus, we use the lower threshold for resource rents and the higher threshold for revenue from the oil and gas sector. As with agricultural dependence, we employ low (5%) and high thresholds (10%) for resource rents to accommodate multiple reasonable cut-offs. A country is coded as a petrostate between 1932 and 1969 if its oil and gas revenues exceed 10% of GDP, and it is coded as a petrostate between 1970 and 2015 if its resource rents exceed 7.5% of GDP. For observations after 1970 with missing values for resource rents, we use the 10% threshold for oil and gas revenues as a share of GDP.

2.3 Putting It All Together: Measuring Land-orientation

We use agricultural dependence and natural resource dependence to create a binary and continuous measure of land-orientation. We consider a state land-oriented if it is *either* dependent on agriculture or natural resources. Given that we set three different thresholds for each input variable, we generate three different versions of the binary variable, using high, medium, and low estimates for the thresholds. Figure A5 shows the decline over time in the proportion of the world's states that are land-oriented according to the three thresholds. We also generate a continuous measure, which calculates the sum of the percent-to-threshold for both agriculture value added and natural resources rents. For example, if a state is 50% to reaching the threshold for both agriculture and natural resources, then it will have a value of 1 (0.5 + 0.5 = 1). See Figure A6 for the distribution of the continuous land-orientation variable. Table A7 summarizes the thresholds for our three binary land-oriented variables. Table A8 specifies the conditions for coding an observation as land-oriented.

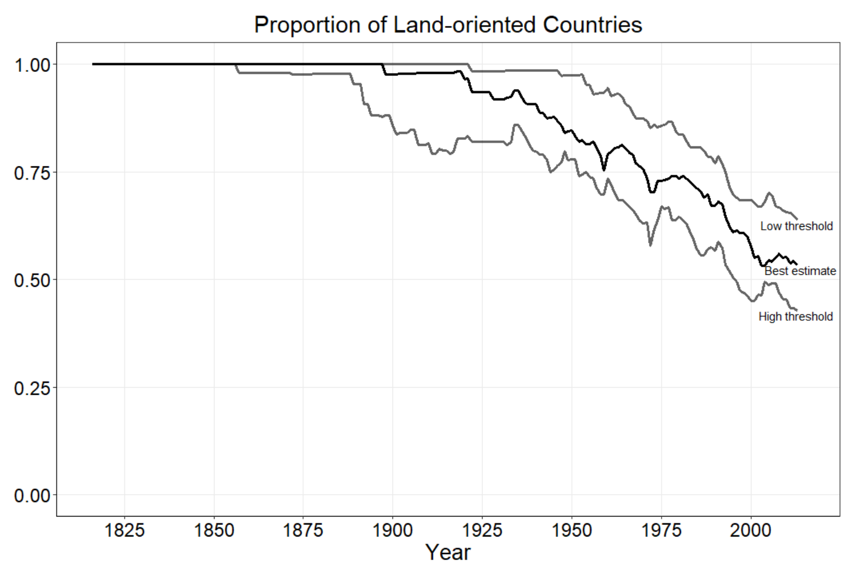

Figure A5: Decline in the proportion of land-oriented countries over time according to the high, medium, and low thresholds.

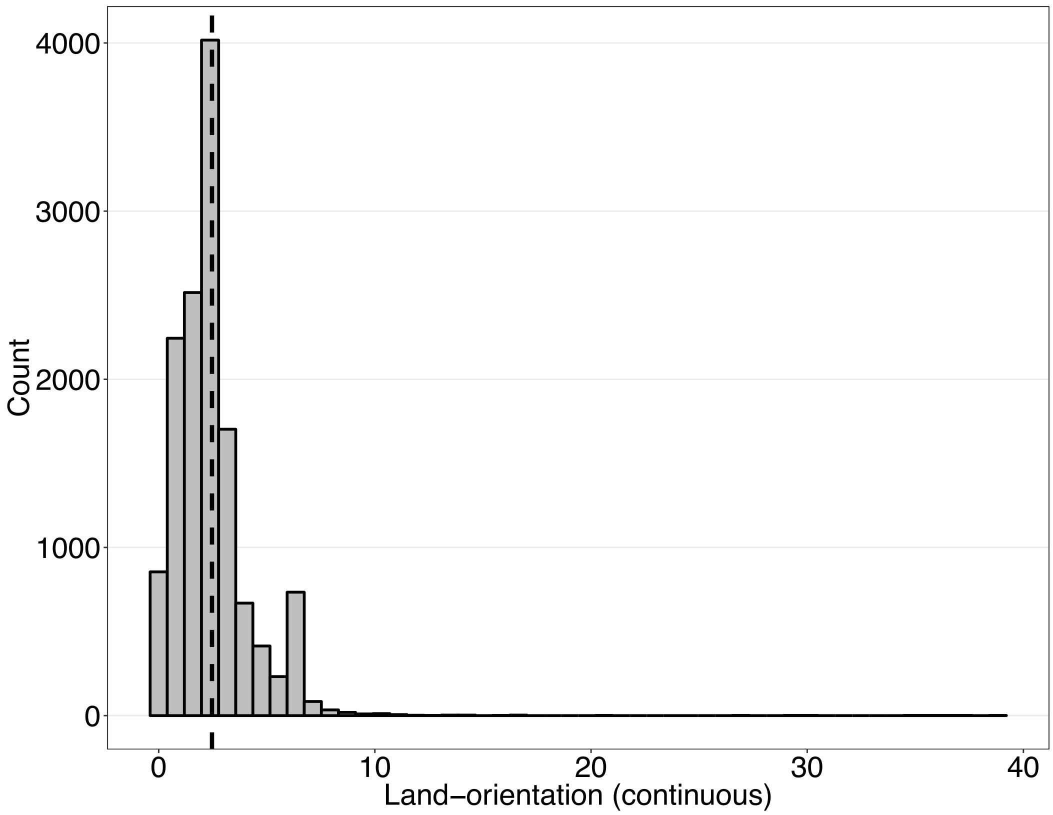

Figure A6: Distribution of the continuous land-orientation measure, using the 15% threshold for agriculture value added as a share of GDP and the 7.5% threshold for natural resources rents as a share of GDP. The dashed line represents the mean value, at 2.46.

Table A7: Thresholds for land-oriented input variables

Variable	***Low***	***Medium***	***High***
Agriculture, value added (% of GDP)	10%	15%	20%
Natural resources rents (% of GDP)	5%	7.5%	10%
Oil and gas revenues (% of GDP)		10%	

Table A8: Conditions for developing a binary measure of land-orientation

1816-2015	If agriculture value added (% of GDP) is greater than the threshold, the state is land-oriented.
1932-1969	If either agriculture value added (% of GDP) or oil and gas revenues (% of GDP) is greater than the threshold, the state is land-oriented.
1970-2015	If either agriculture value added (% of GDP) or natural resources rents (% of GDP) is greater than the threshold, the state is land-oriented.

3 Dependent Variable: Measuring Military Competition Over Territory

The outcome of interest in our study is a state's decision to compete militarily over territory. We measure military competition over territory in a number of different ways: when states initiate claims against other states' territories, when states engage in militarized interstate disputes (MIDs) over territory, and when states engage in territorial rivalries. When merging the claims and rivalry data into a full country-panel dataset, we use Gleditsch and Ward's year of independence to determine if a country-year should receive a score of zero (if it is in existence but not engaged in a claim or rivalry) or a missing value (if the state does not exist in that year). Table A9 summarizes each variable's description.

3.1 Territorial Claims

We draw territorial claims data from the Issue Correlates of War (ICOW) Project (Frederick, Hensel and Macaulay, 2017). ICOW covers territorial claims globally from 1816 to 2001. Authors define a territorial claim as an "explicit contention between two or more nation-states claiming sovereignty over a specific piece of territory" (Hensel and Mitchell, N.d.). ICOW includes information on whether the specific territory involved in the claim was known or believed by either party to contain economically valuable natural resources. When states initiate claims over this kind of territory, we call it a resource-based territorial claim. In each case, we focus on claims initiated in a given year, not the number of territorial claims a state has outstanding. We focus on initiation because a state may fail to formally relinquish a claim until long after it has ceased to actively pursue it.

We operationalize territorial claims in four ways:

1. a binary measure in which a state receives a score of one if that state initiates at least one **territorial claim** in a given year, zero otherwise,
2. a count measure capturing the number of **territorial claims** each state initiates in each year,
3. a binary measure in which a state receives a score of one if that state initiates at least one **resource-based territorial claim** in a given year, zero otherwise, and
4. a count measure capturing the number of **resource-based territorial claims** each state initiates in each year.

3.2 Militarized Disputes

ICOW uses data from The Correlates of War Project's militarized interstate dispute (MIDs) data version 3.1 to determine whether or not there was a MID related to a specific territorial claim. We operationalize militarized disputes in four ways:

1. a binary measure of whether a state is involved in at least one militarized dispute over a **territorial claim** it initiated,
2. a count measure of how many such disputes it had ongoing that year,
3. a binary measure of whether a state is involved in at least one militarized dispute over a **resource-based territorial claim** it initiated, and
4. a count measure of how many such disputes it had ongoing that year.

3.3 Territorial Rivalry

Thompson and Dreyer's (2012) *Handbook on International Rivalries* contains data on rivalries that are based on territorial contestation (Thompson and Dreyer, 2012). These data cover all states in the international system from 1816 to 2012. Each rivalry is categorized by the type of issue that appears to motivate it, including whether or not the rivalry is based on contestation over the exclusive control of territory. As above, we operationalize territorial rivalries in two ways: (1) a binary measure of whether a state is involved in at least one ongoing territorial rivalry, and (2) a count measure of how many such rivalries it had ongoing that year.

Table A9: Dependent variables and their descriptions

Territorial claim (dummy)	= 1 if a state initiates at least one territorial claim in a given year
Territorial claim (count)	the number of territorial claims each state initiates in each year
Resource-based territorial claim (dummy)	= 1 if a state initiates at least one resource-based territorial claim in a given year
Resource-based territorial claim (count)	the number of resource-based territorial claims each state initiates in each year
Territorial MID (dummy)	= 1 if a state is involved in at least one militarized dispute over a territorial claim it initiated
Territorial MID (count)	the number of militarized settlement attempts a state is involved in over a territorial claim it initiated in each year
Resource-based territorial MID (dummy)	= 1 if a state is involved in at least one militarized dispute over a resource-based territorial claim it initiated
Resource-based territorial MID (count)	the number of militarized settlement attempts a state is involved in over a resource-based territorial claim it initiated in each year
Territorial rivalry (dummy)	= 1 if a state is involved in at least one territorial-based rivalry in a given year
Territorial rivalry (count)	the number of territorial rivalries a state is involved in each year

4 Control Variables

To evaluate the conquest hypothesis, we control for a number of alternative explanations, primarily economic development, military capacity, and the opportunity to compete militarily over territory. In our primary specification, we control for GDP, population, military expenditures, military personnel, number of neighbors, and a dummy for whether or not a state is an island. In our secondary specification, we add controls for military capabilities (i.e. COW's Composite Index of National Capability (CINC) Score), trade as a share of GDP, and the number of oil and gas producing neighbors with which a state shares a contiguous land border.

We use Anders, Fariss and Markowitz (2020)'s updated latent estimates of historic GDP and population in order to achieve full sample coverage. Data for military expenditures and military personnel come from COW's National Material Capabilities Dataset Version 5.0. We use Anders, Fariss and Markowitz (2017)'s conversion of military expenditures into constant 2010 US$ to enable regression analysis using panel data. We measure the number of neighbors with which a state shares a contiguous land border using COW's Direct Contiguity (version 3.2) and Colonial/Dependency (version 3.1) datasets (Stinnett et al., 2002). To prevent creating inflated estimates of zero neighbors, we use Gleditsch and Ward's country identification system to set the year of independence for each country. We aggregate the data to the country-year level to create a count of how many neighbors with which a state shares a contiguous land border. We create a binary variable measuring whether a state is an island based on the neighbors variable. If the state has no land neighbors, it is an island.

We also include a more comprehensive measure of a state's military capacity using COW's Composite Index of National Capability (CINC) Score. CINC is an aggregation of information on military personnel, military expenditures, iron and steel production, primary energy consumption, population, and urban population. CINC measures each state's share of the system total of each element of capabilities each year, with each component weighted equally (Greig and Enterline, 2017). Data on trade is originally sourced from the World Bank's World Development Indicators (Bank, 2017). We collected the data from the IPE Data Resource (Graham and Tucker, 2017). Trade as a share of GDP is the sum of a country's exports and imports divided by its GDP. We create a variable measuring the number of oil and gas producing neighbors. Data on oil and gas production comes from Ross and Mahdavi (2015)'s oil and gas datasest, discussed in Section 2.2. We use data on land contiguities from COW's Direct Contiguity (version 3.2) and Colonial/Dependency (version 3.1) datasets, as discussed above. If a country shares a border with another state that is in the top 75% of oil and gas producing states (as measured by the value of the volume of production), we count that as an oil and gas producing neighbor. From here, we aggregate to the country-year level by counting the total number of neighbors that fall above the threshold in each year.

Table A10: Data sources for control variables

Data source	***Variable***	***Time coverage***
Anders, Fariss and Markowitz (2020)	GDP (World Bank estimate using constant 2010 US$)	1500-2015
Anders, Fariss and Markowitz (2020)	Population (CINC estimate, in thousands)	1500-2015
Correlates of War, CINC v 5.0 & Anders, Fariss and Markowitz (2017)	Military expenditures using constant 2010 US$	1816-2012
Correlates of War, CINC v 5.0	Military personnel	1816-2012
COW Direct Contiguity (version 3.2) and Colonial/Dependency (version 3.1)	Number of neighbors	1816-2012
COW Direct Contiguity (version 3.2) and Colonial/Dependency (version 3.1)	Island (dummy)	1816-2012
Correlates of War, CINC v 5.0	Composite Index of National Capability (CINC) Score	1816-2012
World Bank's World Development Indicators	Trade as a share of GDP	1960-2015
Ross and Mahdavi's oil and gas dataset; COW Direct Contiguity (version 3.2); COW Colonial/Dependency (version 3.1)	Oil & gas producing neighbors	1932-2014

References

Anders, Therese, Christopher J. Fariss and Jonathan N. Markowitz. 2017. “Revised Estimates of Military Expenditures and Military Personnel.” *Unpublished working paper* .

Anders, Therese, Christopher J. Fariss and Jonathan N. Markowitz. 2020. “Bread Before Guns or Butter: Introducing Surplus Domestic Product.” *International Studies Quarterly* 64(2):329–405.

Ayuda, María Isabel, Fernando Collantes and Vicente Pinilla. 2010. “Long-run regional population disparities in Europe during modern economic growth: a case study of Spain.” *The Annals of Regional Science; Heidelberg* 44(2):273–295.

Bank, The World. 2017. “World Development Indicators.”.
URL: *https://data.worldbank.org/data-catalog/world-development-indicators*

Bentzen, Jeanet Sinding, Nicolai Kaarsen and Asger Wingender. 2013. “The Timing of Industrialization Across Countries.”.
URL: *https://ideas.repec.org/p/kud/kuiedp/1317.html*

Boix, Carles, Michael Miller and Sebastian Rosato. 2013. “A Complete Data Set of Political Regimes, 1800–2007.” *Comparative Political Studies* 46(12):1523–1554.

Colgan, Jeff. 2013. *Petro-Aggression: When Oil Causes War*. Cambridge: Cambridge University Press.

Frederick, Bryan A, Paul R Hensel and Christopher Macaulay. 2017. “The Issue Correlates of War Territorial Claims Data, 1816–2001.” *Journal of Peace Research* 54(1):99–108.
URL: *http://journals.sagepub.com/doi/10.1177/0022343316676311*

Gleditsch, Kristian Skrede and Michael Ward. 1999. “Interstate System Membership: A Revised List of the Independent States since 1816.” *International Interactions* 25:393–413.

Graham, Benjamin A.T. and Jacob R. Tucker. 2017. “The International Political Economy Data Resource.” *Review of International Organizations* Online First.

Greig, J. Michael and Andrew J. Enterline. 2017. “Correlates of War Project: National Material Capabilities (NMC) Data Documentation Version 5.0.”.

Hensel, Paul R and Sara McLaughlin Mitchell. N.d. “Project Description.”.
URL: *http://www.paulhensel.org/icow.html*

Herrendorf, Berthold, Richard Rogerson and Akos Valentinyi. 2014. Growth and Structural Transformation. In *Handbook of Economic Growth*, ed. Philippe Aghion and Steven Durlauf. Vol. 2B Amsterdam and New York: North-Holland pp. 855–941.
URL: *https://sites.google.com/site/valentinyiakos/Home/papers/handbook*

Honaker, James, Gary King and Matthew Blackwell. 2011. “Amelia II: A Program for Missing Data.” *Journal of Statistical Software* 45(7).

Moran, Theodore H. 1975. *Multinational Corporations and the Politics of Dependence: Copper in Chile*. Princeton University Press.

Ross, Michael and Paasha Mahdavi. 2015. “Codebook for Ross-Mahdavi Oil and Gas Dataset.”.
URL: *https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/ZTPW0Y*

Singer, J. David. 1987. “Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816–1985.” *International Interactions* 14:115–132.

Singer, J. David, Stuart Bremer and John Stuckey. 1972. Capability Distribution, Uncertainty, and Major Power War. In *Peace, War, and Numbers.*, ed. Bruce M. Russett. Beverly Hills, CA: Sage pp. 19–48.

Stinnett, Douglas M., Jaroslav Tir, Philip Schafer, Paul F. Diehl and Charles S. Gochman. 2002. “The Correlates of War Project Direct Contiguity Data, Version 3.” *Conflict Management and Peace Science* 19(2):58–66.

Thompson, William R. and David Dreyer. 2012. *Handbook of International Rivalries*. CQ Press.

Wingender, Asger. 2014. Structural Transformation in the 20th Century: A New Database on Agricultural Employment Around the World. In *Discussion Paper No. 14-28*. Department of Economics, University of Copenhagen: Discussion Paper No. 14-28.
URL: *https://sites.google.com/site/asgerwingender/home/publications/structural-transformation-data-set*