

Codebook for Land-Orientation

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

June 12, 2020

Contents

1	Introduction	2
1.1	Citing This Dataset	2
1.2	What is Land-Orientation?	2
2	Variable List	6
2.1	Primary Variables	6
2.2	Component Variables	6
3	Data Sources	7
3.1	Agricultural Dependence	8
3.1.1	Agriculture as a Percentage of GDP	8
3.1.2	Agricultural Employment	8
3.1.3	Energy Consumption	9
3.1.4	Urban Population	9
3.1.5	GDP per capita	9
3.2	Natural Resource Dependence	9
3.2.1	Resource Rents as a Percentage of GDP	9
3.2.2	Oil and Gas Revenues	9
3.2.3	GDP	9
3.2.4	Mining Exports as a Percentage of Merchandise Exports	10
3.2.5	Mining: Country-Specific Corrections	10
4	Measuring Land-Orientation	10

*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

1 Introduction

1.1 Citing This Dataset

When using the dataset, please cite:

Markowitz, Jonathan N., Suzie Mulesky, Benjamin A.T. Graham, and Christopher J. Fariss. “Productive Pacifists: The Rise of Production-Oriented States and Decline of Territorial Conquest.” *International Studies Quarterly*. Forthcoming.

1.2 What is Land-Orientation?

Note: Please see the online appendix for further details about how we measure land-orientation.

We develop a new measure of economic rent structure, called *land-orientation*, that quantifies states’ dependence on territory over time, capturing their dependence on rents from both agriculture and natural resources, including oil and mining. These data cover most countries in the world from 1816-2015. They begin in a period when all states were land-oriented — organized to extract rents from agrarian surplus — and extend to more recent years in which about half of states are production-oriented – organized to profit from producing goods and services.

We have a separate measurement strategy for dependence on agriculture and dependence on rents from natural resources, detailed below. We measure agricultural dependence using a combination of historic and contemporary data on agriculture as a share of GDP as well as a data imputation model. We measure natural resource dependence using data on rents from natural resources sourced from the World Bank’s World Development Indicators and the estimated value of oil and gas revenues provided by Ross and Mahdavi (2015). We combine measures of agriculture and natural resource dependence into a single measure of land-rent-dependence, or what we call land-orientation. We produce three versions of a binary land-orientation variable that accommodates multiple reasonable thresholds as well as one continuous measure of land-orientation.

We include data for the following 187 countries and territories:

Country	First Year	Last Year
Afghanistan	1816	2015
Albania	1913	2015
Algeria	1816	2013
Angola	1975	2013
Argentina	1819	2013
Armenia	1991	2015
Australia	1901	2013
Austria	1918	2013
Austria-Hungary	1816	1918
Azerbaijan	1991	2013
Baden	1816	1871
Bahrain	1993	2013
Bangladesh	1971	2015
Bavaria	1816	1871
Belarus	1991	2013
Belgium	1830	2013
Benin	1960	2015
Bhutan	1976	2015
Bolivia	1825	2013
Bosnia-Herzegovina	1992	2013
Botswana	1966	2013
Brazil	1822	2013
Bulgaria	1878	2013

Burkina Faso (Upper Volta)	1960	2015
Burundi	1962	2015
Cambodia (Kampuchea)	1953	2015
Cameroon	1960	2015
Canada	1867	2013
Central African Republic	1960	2015
Chad	1960	2015
Chile	1818	2013
China	1816	2013
Colombia	1830	2013
Comoros	1996	2015
Congo	1960	2013
Congo, Democratic Republic of (Zaire)	1960	2015
Costa Rica	1933	2013
Cote d'Ivoire	1960	2015
Croatia	1994	2013
Cuba	1902	2013
Cyprus	1960	2013
Czech Republic	1995	2013
Czechoslovakia	1918	1992
Denmark	1816	2013
Djibouti	1993	2013
Dominican Republic	1898	2013
East Timor	2002	2015
Ecuador	1830	2013
Egypt	1827	2013
El Salvador	1874	2013
Equatorial Guinea	2003	2013
Eritrea	1993	2012
Estonia	1918	2013
Ethiopia	1855	2015
Fiji	1970	2013
Finland	1917	2013
France	1816	2013
Gabon	1967	2013
Gambia	1975	2015
Georgia	1991	2013
German Democratic Republic	1949	1988
German Federal Republic	1957	2013
Germany (Prussia)	1816	1937
Ghana	1957	2015
Greece	1828	2013
Guatemala	1840	2013
Guinea	1958	2015
Guinea-Bissau	1974	2015
Guyana	1966	2015
Haiti	1816	2015
Hanover	1816	1871
Hesse-Darmstadt (Ducal)	1816	1871
Hesse-Kassel (Electoral)	1816	1871
Honduras	1905	2013
Hungary	1918	2013
India	1947	2015

Indonesia	1945	2013
Iran (Persia)	1816	2013
Iraq	1932	2013
Ireland	1921	2013
Israel	1948	2013
Italy/Sardinia	1816	2013
Jamaica	1962	2013
Japan	1816	2013
Jordan	1946	2013
Kazakhstan	1991	2013
Kenya	1963	2015
Korea	1816	1910
Korea, People's Republic of	1948	2015
Korea, Republic of	1948	2013
Kosovo	2008	2013
Kuwait	1968	2013
Kyrgyz Republic	1991	2015
Laos	1954	2015
Latvia	1918	2013
Lebanon	1944	2013
Lesotho	1966	2013
Liberia	1933	2015
Libya	1818	2012
Lithuania	1918	2013
Luxembourg	2013	2013
Macedonia (Former Yugoslav Republic of)	1993	2013
Madagascar	1816	2015
Malawi	1964	2015
Malaysia	1957	2013
Mali	1960	2015
Mauritania	1960	2015
Mauritius	1968	2013
Mecklenburg-Schwerin	1846	1871
Mexico	1821	2013
Modena	1845	1861
Moldova	1991	2013
Mongolia	1921	2013
Montenegro	2008	2013
Morocco	1816	2013
Mozambique	1975	2015
Myanmar (Burma)	1816	2015
Namibia	1992	2013
Nepal	1816	2015
Netherlands	1816	2013
New Zealand	1907	2013
Nicaragua	1907	2015
Niger	1960	2015
Nigeria	1960	2015
Norway	1905	2013
Oman	1953	2013
Pakistan	1947	2015
Panama	1932	2013
Papua New Guinea	1975	2015

Paraguay	1905	2015
Peru	1824	2013
Philippines	1946	2013
Poland	1918	2013
Portugal	1816	2013
Qatar	1996	2013
Rumania	1878	2013
Russia (Soviet Union)	1816	2013
Rwanda	1962	2015
Saudi Arabia	1932	2013
Saxony	1816	1871
Senegal	1960	2015
Serbia	1878	2013
Sierra Leone	1961	2015
Singapore	1967	2013
Slovakia	1995	2013
Slovenia	1994	2013
Solomon Islands	2010	2015
Somalia	1960	2015
South Africa	1910	2013
South Sudan	2011	2015
Spain	1816	2013
Sri Lanka (Ceylon)	1948	2013
Sudan	1956	2015
Surinam	2013	2013
Swaziland	1976	2012
Sweden	1816	2013
Switzerland	1816	2013
Syria	1946	2015
Taiwan	1949	2012
Tajikistan	1991	2015
Tanzania/Tanganyika	1961	2015
Thailand	1816	2013
Tibet	1913	1950
Togo	1960	2015
Trinidad and Tobago	1964	2013
Tunisia	1816	2013
Turkey (Ottoman Empire)	1816	2013
Turkmenistan	1991	2013
Tuscany	1816	1861
Two Sicilies	1816	1861
Uganda	1962	2015
Ukraine	1991	2013
United Arab Emirates	1977	2013
United Kingdom	1816	2013
United States of America	1816	2013
Uruguay	1883	2013
Uzbekistan	1991	2015
Venezuela	1829	2012
Vietnam (Annam/Cochin China/Tonkin)	1816	1953
Vietnam, Democratic Republic of	1954	2015
Vietnam, Republic of	1954	1975
Wurttemberg	1816	1871

Yemen (Arab Republic of Yemen)	1918	2013
Yemen, People's Republic of	1967	1990
Yugoslavia	1918	2006
Zambia	1964	2013
Zimbabwe (Rhodesia)	1965	2013

2 Variable List

The dataset contains a number of variables, including the final measures for land-orientation as well as the component variables used to create the measure. The following describes each of these variables.

2.1 Primary Variables

- `land_oriented_medium_binary` — 1 = country is land-oriented, 0 = country is production-oriented.
Details: 1 = country is either dependent on agriculture ($\text{ag/GDP} \geq 15\%$) or natural resources ($\text{rents/GDP} \geq 7.5\%$ or $\text{oil-gas-revenues/GDP} \geq 10\%$)
- `land_oriented_low_binary` — 1 = country is land-oriented, 0 = country is production-oriented.
Details: 1 = country is either dependent on agriculture ($\text{ag/GDP} \geq 10\%$) or natural resources ($\text{rents/GDP} \geq 5\%$ or $\text{oil-gas-revenues/GDP} \geq 10\%$)
- `land_oriented_high_binary` — 1 = country is land-oriented, 0 = country is production-oriented.
Details: 1 = country is either dependent on agriculture ($\text{ag/GDP} \geq 20\%$) or natural resources ($\text{rents/GDP} \geq 10\%$ or $\text{oil-gas-revenues/GDP} \geq 10\%$)
- `land_oriented_medium_continuous` — continuous measure expressed in percentage terms of how close a country is to both its agriculture and natural resources thresholds, using the medium thresholds.
Details: calculates the sum of the percent-to-threshold for both agriculture value added and natural resources rents (e.g., if a state is 50% to reaching the threshold for both agriculture and natural resources, then it will have a value of 1 — $0.5 + 0.5 = 1$).

2.2 Component Variables

- `pec_MC` — energy consumption
- `upop_MC` — urban population
- `oil_gas_value_2014_PTO` — oil and gas revenues
- `WorldBank_gdppc_2010_con_estimate` — GDP per capita
- `WorldBank_gdp_2010_con_estimate` — GDP
- `aes_AES` — share of agricultural employment (Wingender (2014))
- `ag_employ_OWID` — share of agricultural employment (Our World in Data)
- `ag_GDP_OWID` — agriculture value added as a percentage of GDP (Our World in Data)
- `ag_gdp_WDI` — agriculture value added as a percentage of GDP (the World Bank's World Development Indicators)
- `resource_rents_WDI` — natural resources rents as a percentage of GDP

- `ag_combined` — combined measure of agriculture value added as a percentage of GDP from both Our World in Data (`ag_GDP_OWID`) and the World Bank’s World Development Indicators (`ag_gdp_WDI`). It takes on the value of `ag/GDP` based on the World Bank’s figure. If that data is unavailable, then it adopts the value based on Our World in Data.
- `ag_gdp_combined_estimate` — imputed value of agriculture value added as a percentage of GDP
- `ag_gdp_combined_estimate_avg` — 5-year rolling average of `ag_gdp_combined_estimate`
- `ag_dependent_15` — Binary, 1 = country is dependent on agriculture at the 15% threshold (i.e. `ag_gdp_combined_estimate_avg` \geq 15)
- `ag_dependent_10` — Binary, 1 = country is dependent on agriculture at the 10% threshold (i.e. `ag_gdp_combined_estimate_avg` \geq 10)
- `ag_dependent_20` — Binary, 1 = country is dependent on agriculture at the 20% threshold (i.e. `ag_gdp_combined_estimate_avg` \geq 20)
- `ag_continuous_15` — continuous measure expressed in percentage terms of how close a country is to its 15% agricultural threshold as indicated by `ag_gdp_combined_estimate_avg`
- `resource_rents_WDI_avg` — 5-year rolling average of `resource_rents_WDI`
- `oil_gas_gdp` — `oil_gas_value_2014_PTO` divided by `WorldBank_gdp_2010_con_estimate`
- `oil_gas_gdp_avg` — 5-year rolling average of `oil_gas_gdp`
- `resource_dependent_7.5` — Binary, 1 = country is dependent on natural resource rents at the 7.5% threshold
- `resource_dependent_5` — Binary, 1 = country is dependent on natural resource rents at the 5% threshold
- `resource_dependent_10` — Binary, 1 = country is dependent on natural resource rents at the 10% threshold
- `petro_state` — Binary, 1 = country exceeds 10% of oil and gas revenues as a percentage of GDP (i.e. `oil_gas_gdp_avg` \geq 10)
- `rent_continuous_7.5` — continuous measure expressed in percentage terms of how close a country is to its 7.5% resource-rents threshold
- `petro_continuous` — continuous measure expressed in percentage terms of how close a country is to its 10% oil and gas revenue threshold
- `resource_continuous_7.5` — continuous measure of how close a country is to its 7.5% resource-rents threshold in percentage-terms; if data is unavailable, then this variable measures how close a country is to its 10% oil and gas revenue threshold in percentage-terms (i.e., this variable is equal to `rent_continuous_7.5` when data is available; if data is not available, then this variable is equal to `petro_continuous`)

3 Data Sources

We rely on the following data sources to develop our measure of land-orientation.

3.1 Agricultural Dependence

We impute missing values for agriculture value added (% of GDP) using a multiple imputation model developed by Honaker, King and Blackwell (2011). We utilize seven variables that are all highly correlated with agriculture value added 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 two variables that measure agriculture value added directly). After imputing missing values, we take a 5-year rolling average of agriculture value added as a share of GDP to smooth the estimates over time. We create a binary measure of agricultural dependence using 15% of GDP as the cut-off to separate agrarian from industrial economies (we also use a low threshold, 10%, and a high threshold, 20%, to accommodate multiple reasonable thresholds). 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% (10% or 20% for the other versions).

3.1.1 Agriculture as a Percentage of GDP

Our World in Data

Dataset: <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>

From 1800 to 2016, we use historic data on agriculture value added (% of GDP) for 43 countries, sourced from Our World in Data. These historic estimates are based on two sources: (1) Herrendorf, Rogerson and Valentinyi (2014) (<https://sites.google.com/site/valentinyiakos/Home/papers/handbook>) and (2) the 2015 release of the Groningen Growth and Development Centre's 10-Sector Database (<https://www.rug.nl/ggdc/productivity/10-sector/>). 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.

The World Bank's World Development Indicators

Dataset: <https://data.worldbank.org/indicator/NV.AGR.TOTL.ZS>

From 1960 to 2016, we use the World Bank's data on agriculture value added (% of GDP), downloaded on September 27, 2017. This data captures the value added from agriculture, forestry, and fishing divided by the country's GDP.

3.1.2 Agricultural Employment

Our World in Data

Dataset: <https://ourworldindata.org/grapher/gdp-vs-agriculture-employment?time=1801..>

From 1800 to 2015, we use data on the share of employment in the agricultural sector for 44 countries, sourced from Our World in Data. The estimates for agricultural employment are compiled from two data sources: (1) Herrendorf, Rogerson and Valentinyi (2014) (<https://sites.google.com/site/valentinyiakos/Home/papers/handbook>) and (2) the 2015 release of the Groningen Growth and Development Centre's 10-sector Database (<https://www.rug.nl/ggdc/productivity/10-sector/>).

Wingender (2014)

Dataset: <https://sites.google.com/site/asgerwingender/home/publications/structural-transformation-data-set>

From 1900 to 2010, we use data on the share of employment in the agricultural sector for 169 countries, sourced from Wingender (2014). These employment data are sourced from multiple organizations (e.g. the International Labor Organization, national statistical offices) and supplemented with urbanization rates when employment levels are unobserved.

3.1.3 Energy Consumption

Dataset: [https://correlatesofwar.org/data-sets/national-material-capabilities#:~:text=National%20Material%20Capabilities%20\(v5.,included%20in%20this%20data%20set](https://correlatesofwar.org/data-sets/national-material-capabilities#:~:text=National%20Material%20Capabilities%20(v5.,included%20in%20this%20data%20set).

From 1816 to 2012, we use data on energy consumption, sourced from the Correlates of War Project's National Material Capabilities dataset version 5.0 (Singer, Bremer and Stuckey, 1972; Singer, 1987).

3.1.4 Urban Population

Dataset: [https://correlatesofwar.org/data-sets/national-material-capabilities#:~:text=National%20Material%20Capabilities%20\(v5.,included%20in%20this%20data%20set](https://correlatesofwar.org/data-sets/national-material-capabilities#:~:text=National%20Material%20Capabilities%20(v5.,included%20in%20this%20data%20set).

From 1816 to 2012, we use data on urban population, sourced from the Correlates of War Project's National Material Capabilities dataset version 5.0 (Singer, Bremer and Stuckey, 1972; Singer, 1987).

3.1.5 GDP per capita

Dataset: <https://github.com/thereseanders/sdp>

We use 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. See Fariss, Anders and Markowitz (2020) for the most updated version of these data.

3.2 Natural Resource Dependence

We code countries as petrostates between 1932 and 1969 if their oil and gas revenues exceed 10% of GDP. We code countries as resource-dependent between 1970 and 2015 if their natural resource rents exceed 7.5% of GDP (we also employ low, 5%, and high thresholds, 10%, for natural resource rents to accommodate multiple reasonable thresholds). For observations after 1970 with missing values for natural resource rents, we use the 10% threshold for oil and gas revenues as a share of GDP. For both resource rents and oil & gas revenues, we take a 5-year trailing average to account for short-term fluctuations in oil prices.

3.2.1 Resource Rents as a Percentage of GDP

Dataset: <https://data.worldbank.org/indicator/NY.GDP.TOTL.RT.ZS>

From 1970 to 2015, we use the World Bank's data on total natural resources rents (% of GDP), downloaded on October 18, 2018. 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.

3.2.2 Oil and Gas Revenues

Dataset: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/ZTPW0Y>

From 1932 to 2014, 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. 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.

3.2.3 GDP

Dataset: <https://github.com/thereseanders/sdp>

We use historic latent estimates of GDP developed by Anders, Fariss and Markowitz (2020), which covers 227 countries from 1500 to 2015. We use this measure of GDP to calculate a country's dependence on oil and gas revenues. See Fariss, Anders and Markowitz (2020) for the most updated version of these data.

3.2.4 Mining Exports as a Percentage of Merchandise Exports

Dataset: <https://data.worldbank.org/indicator/TX.VAL.MMTL.ZS.UN>

From 1962 to 1970, we use the World Bank's data on ores and metals exports (as a percentage of merchandise exports) to select a list of candidate countries that could be dependent on the mining sector but not oil or natural gas.

3.2.5 Mining: Country-Specific Corrections

We code Chile as dependent on natural resources from 1948 to 1971 for the binary resource-dependence measure. 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 determine whether Chile was dependent on its mining sector from 1948 to 1972 (when World Bank data on natural resource rents becomes available for Chile), we referred to data on Chile's ores and metals exports (as a percentage of merchandise exports). From 1962 to 1973, ores and metals accounted for 82% to 89% of total merchandise exports. For data 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.

4 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. We also generate a continuous measure, which calculates the sum of the percent-to-threshold for both agriculture value added and natural resources rents.

References

- 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.
- Fariss, Christopher J., Therese Anders and Jonathan N. Markowitz. 2020. "New Estimates of Over 500 Years of Historic GDP and Population Data." *Unpublished working paper*.
- 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.
- 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>