

Sunlight, culture, and state capacity

–Supporting Information–

Roberto Ezcurra

Department of Economics and INARBE

Public University of Navarre

roberto.ezcurra@unavarra.es

A Variable definitions and sources

I now present the definitions and sources of the different variables used throughout the paper.

Measures of state capacity

- *State Capacity Index (HSI)*: Aggregate measure of state capacity based on the three dimensions of state capacity that are “minimally necessary to carry out the functions of contemporary states, and [...] most plausibly distinct from one another” (Hanson and Sigman, 2021, p. 1498): extractive capacity, coercitive capacity, and administrative capacity. These three dimensions of state capacity are captured by 21 variables drawn from different sources (Table A2), and synthesized into a single aggregate indicator through Bayesian latent variable analysis. See section 3 and Hanson and Sigman (2021) for further details. Most of the econometric analyses in the paper use the average value of this measure over the period 2000-2015. Source: Hanson and Sigman (2021).
- *State Fragility Index (SFI)*: Aggregate measure of state capacity that measure the state’s “capacity to manage conflict, make and implement public policy, and deliver essential services” (Marshall and Elzinga-Marshall 2017, p. 51). The index scores each country on both effectiveness and legitimacy in four performance dimensions: security, political, economic, and social. To calculate the SFI, each of these indicators is rated on a four-point fragility scale: 0 ‘no fragility’, 1 ‘low fragility’, 2 ‘medium fragility’, and 3 ‘high fragility’ with the exception of the economic effectiveness indicator, which is rated on a five-point fragility scale (including 4 ‘extreme fragility’). In this research, the SFI has been rescaled, so that higher values of the index indicate lower state fragility. The analyses in the paper are based on the average value of this measure over the period 2000-2015. Source: Center for Systemic Peace.
- *Index of State Capacity (OMI)*: Comprehensive measure of state capacity based on aggregating the information provided by various variables from the Varieties of

Democracy dataset. The selected variables captures the degree of fiscal capacity, a state's control over its territory, the rule of law, and the provision of public goods used to support markets. The aggregate measure of state capacity is constructed using principal component analysis. The analyses in the paper are based on the average value of this measure over the period 2000-2015. Source: O'Reilly and Murphy (2022).

- *Fragile States Index (FSI)*: Aggregate measure of state capacity that captures “a state’s capacities and pressures which contribute to levels of fragility and resilience” (Fund for Peace, 2019, p. 33). Following the Conflict Assessment System Tool (CAST) analytical approach, three primary streams of data –quantitative, qualitative, and expert validation– are triangulated and subjected to critical review to obtain final scores for the FSI. The country scores are based on different key political, social, and economic indicators. In this research, the FSI has been rescaled, so that higher values of the index indicate lower state fragility. The analyses in the paper are based on the average value of this measure over the period 2000-2015. Source: The Fund for Peace.
- *Government effectiveness*: Measure that captures perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government’s commitment to such policies. The measure is obtained using an unobserved components methodology that aggregates the information provided by numerous underlying variables taken from different data sources. See Kaufmann *et al.* (2011) for further details. The analyses in the paper are based on the average value of this measure over the period 2000-2015. Source: Worldwide Governance Indicators (World Bank).
- *Tax revenue share in GDP*: Total tax revenue as a share of GDP (resource revenues are not included as tax revenues). The analyses in the paper are based on the average value of this variable over the period 2000-2015. Source: Hanson and Sigman (2021).
- *Income tax share*: Proportion of total tax revenue that comes from taxes on in-

come, profits, and capital gains. The analyses in the paper are based on the average value of this variable over the period 2000-2015. Source: Hanson and Sigman (2021).

- *Fiscal capacity index*: Measure that captures the state's capacity to fund itself through taxes that are of greater administrative complexity. The measure is calculated using a Bayesian item response theory measurement model based on a five-point ordinal scale. See Coppedge *et al.* (2021) for further details. The analyses in the paper are based on the average value of this variable over the period 2000-2015. Source: Varieties of Democracy dataset (version 11.1).
- *Rule of law*: Measure that captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence. The measure is obtained using an unobserved components methodology that aggregates the information provided by numerous underlying variables taken from different data sources. See Kaufmann *et al.* (2011) for further details. The analyses in the paper are based on the average value of this measure over the period 2000-2015. Source: Worldwide Governance Indicators (World Bank).
- *Regulatory quality*: Measure that captures perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development. The measure is obtained using an unobserved components methodology that aggregates the information provided by numerous underlying variables taken from different data sources. See Kaufmann *et al.* (2011) for further details. The analyses in the paper are based on the average value of this measure over the period 2000-2015. Source: Worldwide Governance Indicators (World Bank).
- *Impartial public administration*: Measure that assesses the level of compliance with the law by public officials. The measure is calculated using a Bayesian item response theory measurement model based on a five-point ordinal scale. See Coppedge *et al.* (2021) for further details. The analyses in the paper are based

on the average value of this measure over the period 2000-2015. Source: Varieties of Democracy dataset (version 11.1).

Measure of UV-R exposure

- *UV-R*: Measure that captures the degree of UV-R exposure at the country level. The measure has been calculated by Andersen *et al.* (2016), using satellite data from NASA available in the form of geographical grids and daily rasters with pixel size of 1° latitude by 1° longitude. Relying on data for daily local-noon irradiances for 1990 and 2000, Andersen *et al.* (2016) calculate the average yearly UV-R levels for each country (see Andersen *et al.* (2016) for further technical details). The analyses in the paper are based on the average UV-R index for each country for the years 1990 and 2000. Source: Andersen *et al.* (2016).

Cultural variables

- *Individualism*: Cultural dimension identified by Hofstede (2001), which captures the preference for a loosely-knit social framework in which individuals are expected to take care of only themselves and their immediate families. Low values of this measure indicate a preference for a tightly-knit framework in society in which individuals can expect their relatives or members of a particular ingroup to look after them in exchange for unquestioning loyalty. Further details can be found in section 5. Source: Hofstede (2001).
- *Uncertainty avoidance*: Cultural dimension identified by Hofstede (2001), which captures the degree to which the members of a society feel uncomfortable with uncertainty and ambiguity. Countries with high values in this measure tend to maintain rigid codes of belief and behaviour, and are intolerant of unorthodox behaviour and ideas. Source: Hofstede (2001).
- *Willingness to take risks*: Measure derived from the Global Preferences Survey (GPS), which captures the respondents willingness to take risks. It is constructed by combining the results of a lottery choice sequence using the staircase method

and self-assessment of one's general willingness to take risks. Source: Falk *et al.* (2018).

- *Trust*: Prevalence of generalized interpersonal trust in a country's population. The variable is constructed using information from the different waves of the World Values Survey (WVS) conducted over the period 1981-2009, and reflects the proportion of all respondents that opted for the answer 'Most people can be trusted' (as opposed to 'Can't be too careful') when responding to the survey question 'Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?'. Source: Arbath *et al.* (2020), who collected the data from the WVS.
- *Positive reciprocity*: Measure derived from the Global Preferences Survey (GPS), which captures the respondents' propensities to act in a positively reciprocal way. Source: Falk *et al.* (2018).
- *Negative reciprocity*: Measure of negative reciprocity derived from the Global Preferences Survey (GPS), which captures individuals' tendency to respond in kind to negative actions or behaviors directed towards them. Source: Falk *et al.* (2018).
- *Altruism*: Measure derived from the Global Preferences Survey (GPS), which captures the degree of altruism through a combination of qualitative and quantitative items related to donations. Source: Falk *et al.* (2018).
- *Long term orientation*: Cultural dimension identified by Hofstede (2001), which captures whether a society exhibits a pragmatic future-oriented perspective, emphasizing virtues oriented towards future rewards, such as perseverance and thrift. Societies with low values in this measure prefer to maintain time-honoured traditions and norms while viewing societal change with suspicion. Source: Hofstede (2001).
- *Masculinity*: Cultural dimension identified by Hofstede (2001), which captures the preference in society for achievement, heroism, assertiveness, and material rewards for success. Low values of this measure indicate a preference for cooperation, modesty, caring for the weak and quality of life. Source: Hofstede (2001).

- *Egalitarianism*: Cultural dimension identified by Schwartz (1994), which captures whether the members of society recognize one another as moral equals who share basic interests as human beings. Source: Schwartz (1994).
- *Harmony*: Cultural dimension identified by Schwartz (1994), which captures the degree to which individuals are content to accept and adapt to their natural and social environment, seeking to understand, preserve, and protect it, rather than change, direct, or exploit it. Source: Schwartz (1994).
- *Mastery*: Cultural dimension identified by Schwartz (1994), which captures the degree to which individuals seek success through personal action. Source: Schwartz (1994).
- *Hierarchy*: Cultural dimension identified by Schwartz (1994), which captures the degree to which individuals accept their position in the social hierarchy and are expected to be modest and have due self-control. Source: Schwartz (1994).
- *Religiosity*: Measure derived from the World Values Survey (WVS), which reflects the proportion of all respondents who believe in God. Source: Enke (2019), who collected the data from the WVS.

Control variables

- *Absolute latitude*: Absolute value of the latitude of a country's approximate geodesic centroid. Source: Andersen *et al.* (2016), who collected the data from Nunn and Puga (2012).
- *Elevation*: Mean elevation of a country in metres above sea level. Source: Andersen *et al.* (2016), who collected the data from the CIA World Factbook.
- *Temperature*: Area-weighted average air temperature in degrees Celsius over the period 1980-2008. Source: Andersen *et al.* (2016), who collected the data from the GECON dataset.
- *Precipitation*: Area-weighted average precipitation in thousands millimetres per

year over the period 1980-2008. Source: Andersen *et al.* (2016), who collected the data from the GECON dataset.

- *Frost days*: Area-weighted average number of frost days per year during the period 1901-2012. Source: Andersen *et al.* (2016), who collected the data from the Climatic Research Unit's (CRU) gridded dataset.
- *Tropical climate*: Proportion of a country's land area with tropical or subtropical climate according to the Köppen-Geiger classification system. Source: Andersen *et al.* (2016), who collected the data from Ashraf and Galor (2011).
- *Terrain ruggedness*: Terrain ruggedness index originally devised by Riley *et al.* (1999) using geospatial data. See Nunn and Puga (2012) for further details. Source: Nunn and Puga (2012).
- *Area*: Total area in square kilometres. Source: World Development Indicators (World Bank).
- *Distance to the nearest waterway*: Share of a country's total land area that is located within 100 kilometres of an ice-free coastline or sea-navigable river. Source: Bentzen *et al.* (2017), who collected the data from the Center for International Development (CID).
- *Land suitability for agriculture*: Geospatial index of the suitability of land for agriculture based on ecological indicators of climate suitability and soil suitability for cultivation. The index was originally calculated by Ramankutty *et al.* (2002), and aggregated at the country level by Michalopoulos (2012). Source: Ashraf and Galor (2013), who collected the data from Michalopoulos (2012).
- *Number of ethnic groups*: Number of distinct ethnic groups in a country's population, as compiled by Fearon (2003). See Fearon (2003) for additional details on primary data sources and methodological assumptions. Source: Fearon (2003).
- *ELF-D*: Index of ethnolinguistic fractionalization that captures the probability that two individuals randomly drawn from the population belong to different ethnolinguistic groups. See Desmet *et al.* (2012) for further details. Although Desmet

et al. (2012) consider different levels of aggregation of linguistic groups in a country's population (based on hierarchical linguistic trees), the specific fractionalization measure used in the paper is the one corresponding to the most disaggregated level. Source: Desmet *et al.* (2012).

- *POL-D*: Index of ethnolinguistic polarization that quantifies the extent to which the ethnolinguistic composition of a country's population resembles a perfectly polarized distribution, in which the national population is formed by two groups of equal size. See Desmet *et al.* (2012) for further details. Although Desmet *et al.* (2012) consider different levels of aggregation of linguistic groups in a country's population (based on hierarchical linguistic trees), the specific polarization measure used in the paper is the one corresponding to the most disaggregated level. Source: Desmet *et al.* (2012).
- *ELF-F*: Index of ethnolinguistic fractionalization that captures the probability that two individuals randomly drawn from the population belong to different ethnolinguistic groups. The index is based on the classification of ethnic groups in Fearon (2003). See Esteban *et al.* (2012) for further details. Source: Esteban *et al.* (2012).
- *POL-F*: Index of ethnolinguistic polarization based on the definition of polarization proposed by Duclos *et al.* (2004) and Esteban and Ray (2011). The index is based on the classification of ethnic groups in Fearon (2003). See Esteban *et al.* (2012) for further details. Source: Esteban *et al.* (2012).
- *ELF-E*: Index of ethnolinguistic fractionalization that captures the probability that two individuals randomly drawn from the population belong to different ethnolinguistic groups. The index is based on the information about linguistic groups provided by the Ethnologue project. See Esteban *et al.* (2012) for further details. Source: Esteban *et al.* (2012).
- *POL-E*: Index of ethnolinguistic polarization based on the definition of polarization proposed by Duclos *et al.* (2004) and Esteban and Ray (2011). The index is based on the information about linguistic groups provided by the Ethnologue project.

See Esteban *et al.* (2012) for further details. Source: Esteban *et al.* (2012).

- *Genetic diversity*: Measure of predicted genetic diversity (ancestry adjusted) of a country's contemporary national population, as developed by Ashraf and Galor (2013). The measure is based on migratory distances from East Africa to the year 1500 locations of the ancestral populations of a country's component ethnic groups in 2000 and on the pairwise migratory distances among these ancestral populations. The source of the ancestral populations are identified from the World Migration Matrix, 1500-2000 (Putterman and Weil, 2010). Source: Ashraf and Galor (2013).
- *Neolithic transition timing*: Number of years (up to the year 2000) since the majority of the population living within a country's modern national borders started practicing sedentary agriculture as their primary mode of subsistence. Source: Arbatli *et al.* (2020), who obtained the data from Putterman (2008).
- *State history*: Index that captures a country's cumulative experience with institutionalized statehood since antiquity. The index is based on the presence of a political entity above the tribal level within modern-day country borders, its degree of political autonomy, and territorial coverage across 110 50-year periods from 3500 BCE to 2000 CE. See Borcan *et al.* (2018) for further methodological details. Source: Borcan *et al.* (2018).
- *Population density in 1500 CE*: Population in 1500 CE, as reported by McEvedy and Jones (1978), divided by total land area, as reported by the World Bank's World Development Indicators. Source: Ashraf and Galor (2011).
- *Urbanization in 1500 CE*: Percentage of a country's total population residing in urban areas (each with a city population size of at least 5,000), as reported by Acemoglu *et al.* (2005). Source: Ashraf and Galor (2013).
- *Technology adoption in 1500 CE*: Level of technology adoption in 1500 CE across five sectors: agriculture, transportation, military, industry, and communications. See Comin *et al.* (2010) for further methodological details. Source: Comin *et al.* (2010).

- *Historical pathogen prevalence*: Index of historical pathogen prevalence derived from Murray and Schaller (2010), and based on epidemiological data. The index covers the prevalence of the following infectious diseases: malaria, typhus, dengue, leishmaniasis, schistosomiasis, trypanosomiasis, and filariasis. It reflects the pathogen environment populations had to face in the early-to-mid 20th century. Source: Fedderke *et al.* (2017).
- *Former colony*: Dummy variable that takes the value of one if the country in question was ever colonized by European powers, and zero otherwise. Source: Own elaboration using data from Nunn and Puga (2012).
- *Transition country*: Dummy variable that takes the value of one if the country in question is or has been in the process of changing from a centrally planned economy to a market economy, and zero otherwise. Source: Own elaboration using data from the International Monetary Fund.
- *GDP per capita*: GDP per capita expressed in constant 2015 US dollars. Average of the period 1960-2000. Source: World Development Indicators (World Bank).
- *Democracy*: Democracy index (polity2) ranging between -10 (full autocracy) and 10 (full democracy). The index is based on information about the degree of competitiveness and openness of elections, the nature of political participation in general, and the extent of checks on executive authority. Average of the period 1960-2000. Source: Polity5 dataset (Center for Systemic Peace).
- *Internal conflicts*: Proportion of years in civil war from 1950 (or independence if later) until 2000. The variable was constructed using data from the UCDP/PRIO Armed Conflict dataset. Source: Besley and Persson (2011).
- *External wars*: Proportion of years in external war from 1816 (or independence if later) until 2000. The variable was constructed using data from the Correlates of War (COW) dataset. Source: Besley and Persson (2011).
- *Dry climate*: Proportion of a country's land area with dry climate according to the Köppen-Geiger classification system. Source: Fedderke *et al.* (2017).

- *Temperate climate*: Proportion of a country's land area with temperate climate according to the Köppen-Geiger classification system. Source: Fedderke *et al.* (2017).
- *Continental climate*: Proportion of a country's land area with continental climate according to the Köppen-Geiger classification system. Source: Fedderke *et al.* (2017).
- *Climate variability*: Gini index that reflects the variability of climate across different regions within a country, according to the Köppen-Geiger classification system. Source: Fedderke *et al.* (2017).
- *Island*: Dummy variable that takes the value of one if the country in question does not share a land border with any other country, and zero otherwise. Source: Arbatlı *et al.* (2020).
- *Landlocked*: Dummy variable that takes the value of one if the country in question has no direct access to the sea, and zero otherwise. Source: Bentzen *et al.* (2017).
- *Inequality in land suitability*: Gini index based on the distribution of a measure of land suitability for agriculture, reported at a half-degree resolution by Ramankutty *et al.* (2002), across grid cells within a country. Source: Ashraf and Galor (2013).
- *Irrigation potential*: Share of land suitable for agriculture where irrigation can more than double agricultural production. Source: Bentzen *et al.* (2017).
- *Growing season days*: Average number of days in a year where temperature and precipitation allow for crop growth. Source: Bentzen *et al.* (2017).
- *Cereal/plantation crops*: The ratio between the amount of land suitable for cereal crops and the amount of land suitable for plantation crops. Source: Bentzen *et al.* (2017).
- *Ecological fractionalization*: Herfindahl index of ecological fractionalization based on the spatial distribution of ecological types across the land surface of the earth. Source: Arbatlı *et al.* (2020).

- *Cataracts*: The prevalence of cataracts is quantified using a measure of Disability-Adjusted Life Years (DALY), which can be interpreted as an estimate of the gap between current health status and an ideal health situation where the entire population lives to an old age, free of disease and disability (WHO, 2008). The indicator used in the paper captures the number of DALYs due to the incidence of cataract in 2004 per 100,000 people in the population. Source: Andersen *et al.* (2016), who collected the data from WHO (2008).
- *Skin cancer*: The prevalence of skin cancer is quantified using a measure of Disability-Adjusted Life Years (DALY), which can be interpreted as an estimate of the gap between current health status and an ideal health situation where the entire population lives to an old age, free of disease and disability (WHO, 2008). The prevalence of this condition is measured as the number of DALYs due to melanoma and other skin carcinomas in 2004 per 100,000 people in the population. Source: Andersen *et al.* (2016), who collected the data from WHO (2008).
- *Malaria*: The prevalence of malaria is quantified using a measure of Disability-Adjusted Life Years (DALY), which can be interpreted as an estimate of the gap between current health status and an ideal health situation where the entire population lives to an old age, free of disease and disability (WHO, 2008). The prevalence of malaria is measured as the number of DALYs due to this disease in 2004 per 100,000 people in the population. Source: Andersen *et al.* (2016), who collected the data from WHO (2008).
- *Trachoma*: The prevalence of trachoma is quantified using a measure of Disability-Adjusted Life Years (DALY), which can be interpreted as an estimate of the gap between current health status and an ideal health situation where the entire population lives to an old age, free of disease and disability (WHO, 2008). The prevalence of trachoma is measured as the number of DALYs due to this disease in 2004 per 100,000 people in the population. Source: Andersen *et al.* (2016), who collected the data from WHO (2008).
- *Hookworm disease*: The prevalence of hookworm disease is quantified using a measure of Disability-Adjusted Life Years (DALY), which can be interpreted as

an estimate of the gap between current health status and an ideal health situation where the entire population lives to an old age, free of disease and disability (WHO, 2008). The prevalence of hookworm disease is measured as the number of DALYs due to this affliction in 2004 per 100,000 people in the population. Source: Andersen *et al.* (2016), who collected the data from WHO (2008).

- *HIV*: The prevalence of HIV is quantified using a measure of Disability-Adjusted Life Years (DALY), which can be interpreted as an estimate of the gap between current health status and an ideal health situation where the entire population lives to an old age, free of disease and disability (WHO, 2008). The prevalence of HIV is measured as the number of DALYs due to this disease in 2004 per 100,000 people in the population. Source: Andersen *et al.* (2016), who collected the data from WHO (2008).
- *Median age population*: Median age of the population in 2004. Source: World Population Prospects 2022, Department of Economic and Social Affairs, United Nations.

Other variables

- *Blood distance*: Euclidian distance of frequency of blood types A and B in a given country relative to their frequency in the USA. Source: Gorodnichenko and Roland (2017).

Table A1: Summary statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
MEASURES OF STATE CAPACITY					
State Capacity Index ^a	143	0.520	0.949	-2.076	2.887
State Fragility Index ^b	143	60.141	27.994	0.000	100
State Capacity Index ^c	143	0.494	1.483	-3.889	3.502
Fragility States Index ^b	143	44.264	24.938	0.000	100
Government effectiveness	143	-0.109	0.960	-2.123	2.080
Tax revenue share in GDP	143	0.162	0.077	0.009	0.454
Income tax share	143	0.344	0.126	0.089	0.654
Fiscal capacity index	143	0.999	1.176	-2.140	2.931
Rule of law	143	-0.178	0.986	-2.313	1.986
Regulatory quality	143	-0.079	0.956	-2.217	1.825
Impartial public administration	143	0.423	1.450	-2.161	4.006
MEASURE OF UV-R EXPOSURE					
UV-R (log)	143	5.145	0.516	3.753	5.795
CULTURAL VARIABLES					
Individualism	90	39.578	22.406	6.000	91.000
Uncertainty avoidance	90	65.633	20.407	23.000	112
Willingness to take risks	59	0.011	0.291	-0.792	0.971
Trust	71	0.290	0.143	0.064	0.664
Positive reciprocity	59	-0.034	0.334	-1.038	0.570
Negative reciprocity	59	0.019	0.282	-0.489	0.739
Altruism	59	-0.039	0.366	-0.940	0.906
Long term orientation	77	42.922	23.432	4.000	100
Masculinity	90	48.033	19.182	5.000	110
Egalitarianism	60	4.715	0.283	4.130	5.270
Harmony	60	4.058	0.317	3.280	4.620
Mastery	60	3.941	0.168	3.540	4.410
Hierarchy	60	2.306	0.461	1.490	3.490
Religiosity	60	0.827	0.215	0.188	1.000
CONTROL VARIABLES					
Absolute latitude (log)	143	2.990	0.983	-0.627	4.166

(Continued...)

Table A1: Summary statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
Elevation (log)	143	0.202	0.879	-3.631	1.505
Temperature	143	17.845	8.378	-7.086	28.788
Precipitation	143	0.991	0.691	0.035	3.220
Frost days	143	9.511	10.400	0.002	29.797
Tropical climate	143	0.353	0.425	0.000	1.000
Terrain ruggedness	143	1.313	1.254	0.037	6.740
Area (log)	143	12.446	1.568	9.222	16.611
Distance to the nearest waterway	143	0.401	0.338	0.000	1.000
Land suitability for agriculture	143	0.388	0.246	0.003	0.951
Number of ethnic groups (log)	142	1.451	0.639	0.000	3.091
ELF-D	143	0.467	0.307	0.000	0.990
POL-D	143	0.446	0.242	0.000	0.958
ELF-F	129	0.412	0.244	0.000	0.842
POL-F	129	0.044	0.052	0.000	0.246
ELF-E	129	0.470	0.313	0.000	0.990
POL-E	129	0.042	0.045	0.000	0.214
Genetic diversity	143	0.727	0.027	0.628	0.774
Neolithic transition timing (log)	134	8.330	0.591	5.991	9.259
State history	138	0.242	0.172	0.017	0.743
Population density in 1500 CE (log)	142	0.900	1.501	-3.817	3.842
Urbanization in 1500 CE	80	7.439	5.161	0.000	28.000
Technology adoption in 1500 CE	107	0.480	0.316	0.000	1.000
Historical pathogen prevalence	143	0.211	0.633	-1.180	1.200
Former colony	143	0.622	0.486	0.000	1.000
Transition country	143	0.224	0.418	0.000	1.000
GDP per capita (log)	140	7.964	1.379	5.411	11.243
Democracy	143	-0.296	6.165	-10.000	10.000
Internal conflicts	139	0.128	0.223	0.000	1.000
External wars	142	0.036	0.073	0.000	0.555
Dry climate	143	24.989	35.048	0.000	100
Temperate climate	143	25.628	30.885	0.000	100
Cataracts (log)	90	4.707	1.673	2.122	6.855
Skin cancer (log)	90	2.637	1.068	0.374	4.269
Malaria (log)	90	1.548	1.977	0.000	6.022
Trachoma (log)	90	0.817	1.617	0.000	5.702
Hookworm disease (log)	90	1.406	1.568	0.000	4.207
HIV (log)	90	4.948	2.423	-0.441	10.279
Median age population	90	27.989	8.513	14.954	42.102

(Continued...)

Table A1: Summary statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
OTHER VARIABLES					
Blood distance	90	1.357	0.645	0	2.909

Notes: ^a Hanson and Sigman (2021). ^b The index has been rescaled, so that higher values indicate higher state capacity. ^c O'Reilly and Murphy (2022).

B Countries

List of countries included in the regression analyses in section 4:

Afghanistan	Denmark	Laos
Albania	Dominican Republic	Latvia
Algeria	Ecuador	Lebanon
Angola	Egypt	Lesotho
Argentina	El Salvador	Liberia
Armenia	Estonia	Libya
Australia	Ethiopia	Lithuania
Austria	Finland	Macedonia
Azerbaijan	France	Madagascar
Bangladesh	Gabon	Malawi
Belarus	Gambia	Malaysia
Belgium	Georgia	Mali
Benin	Germany	Mauritania
Bhutan	Ghana	Mexico
Bolivia	Greece	Moldova
Bosnia and Herzegovina	Guatemala	Mongolia
Botswana	Guinea	Morocco
Brazil	Guinea-Bissau	Mozambique
Bulgaria	Guyana	Namibia
Burkina Faso Faso	Haiti	Nepal
Burundi	Honduras	Netherlands
Cambodia	Hungary	New Zealand
Cameroon	India	Nicaragua
Canada	Indonesia	Niger
Central African Rep.	Iran	Nigeria
Chad	Iraq	Norway
Chile	Ireland	Oman
China	Israel	Pakistan
Colombia	Italy	Panama
Congo, Dem. Rep.	Japan	Papua New Guinea
Congo, Rep.	Jordan	Paraguay
Costa Rica	Kazakhstan	Peru
Cote d'Ivoire	Kenya	Philippines
Croatia	Korea, Rep.	Poland
Cuba	Kuwait	Portugal
Czech Republic	Kyrgyz Republic	Romania

Russia
Rwanda
Saudi Arabia
Senegal
Sierra Leone
Slovak Republic
Slovenia
Somalia
South Africa
Spain
Sri Lanka
Sudan

Swaziland
Sweden
Switzerland
Syria
Tajikistan
Tanzania
Thailand
Togo
Tunisia
Turkey
Turkmenistan
Uganda

Ukraine
United Arab Emirates
United Kingdom
United States
Uruguay
Uzbekistan
Venezuela
Vietnam
Yemen
Zambia
Zimbabwe

C Supplementary results

Figure A1: Latitude and state capacity: Partial regression plot conditional on the level of GDP per capita.

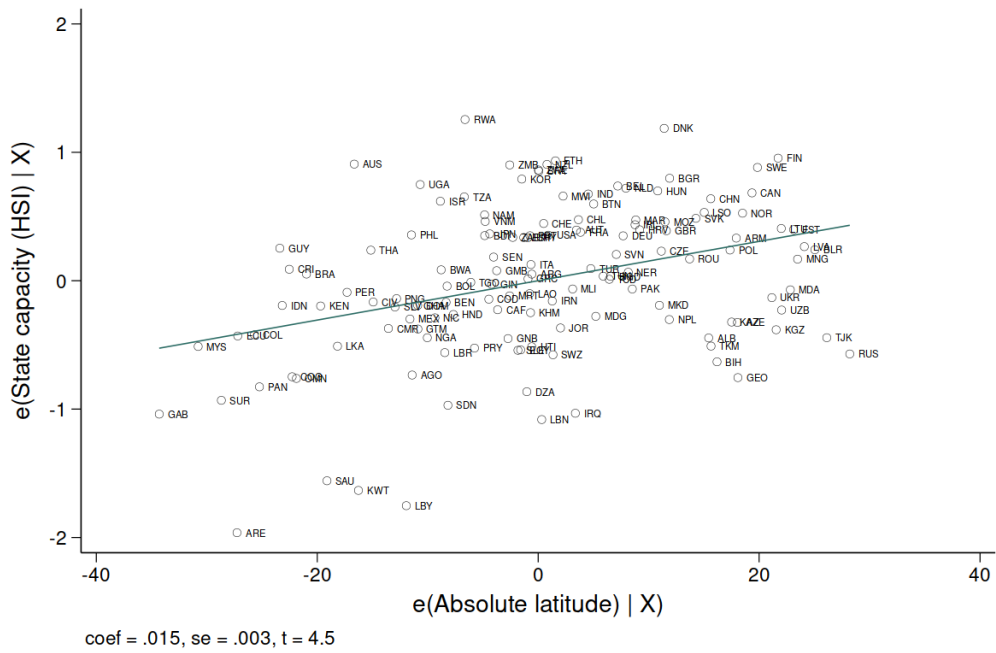


Figure A2: UV-R and state capacity: Preliminary evidence.

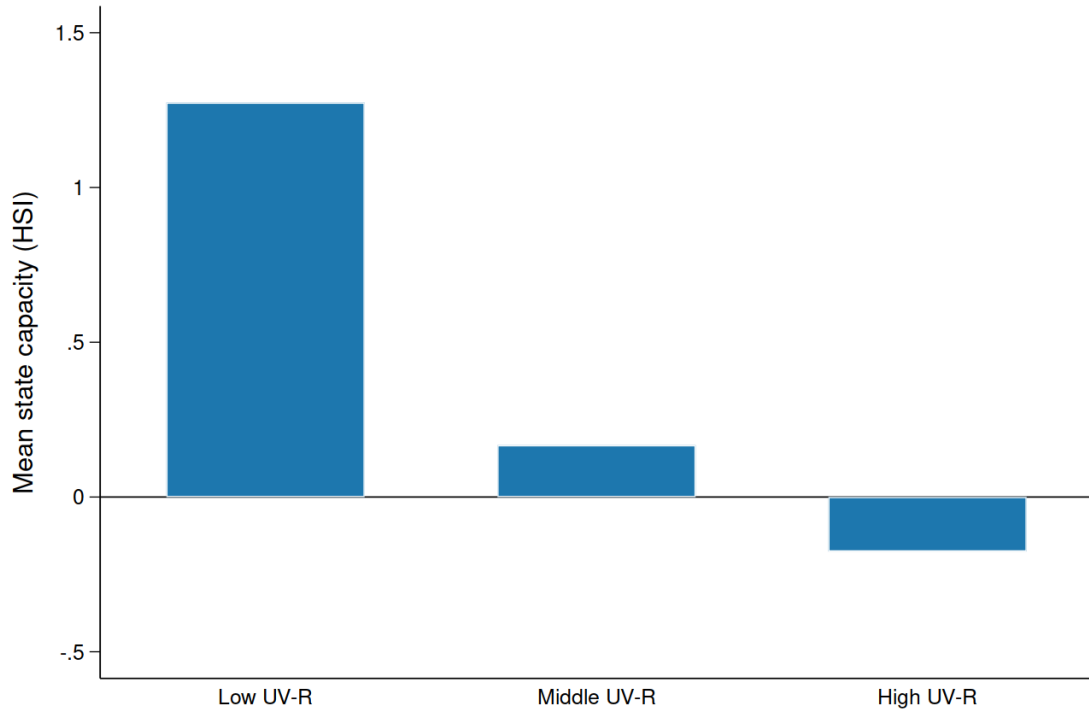
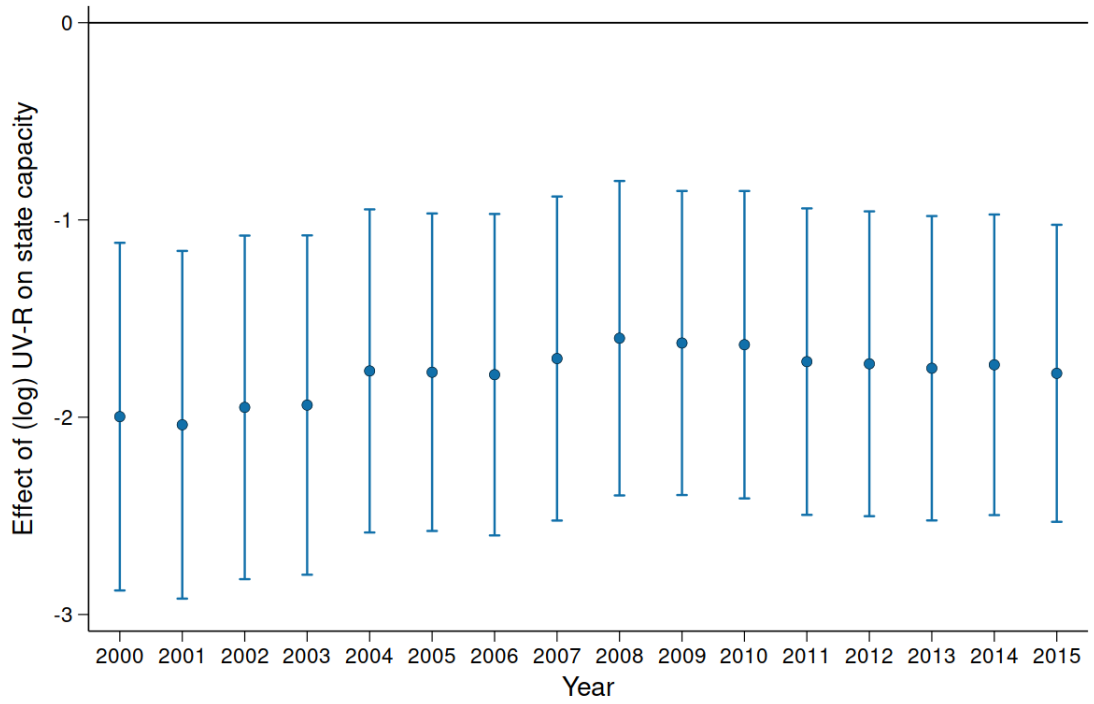


Figure A3: Effect of UV-R on state capacity over time.



Note: Point estimates with their corresponding 95% confidence intervals.

Figure A4: Spatial distribution of individualism (Hofstede's index).

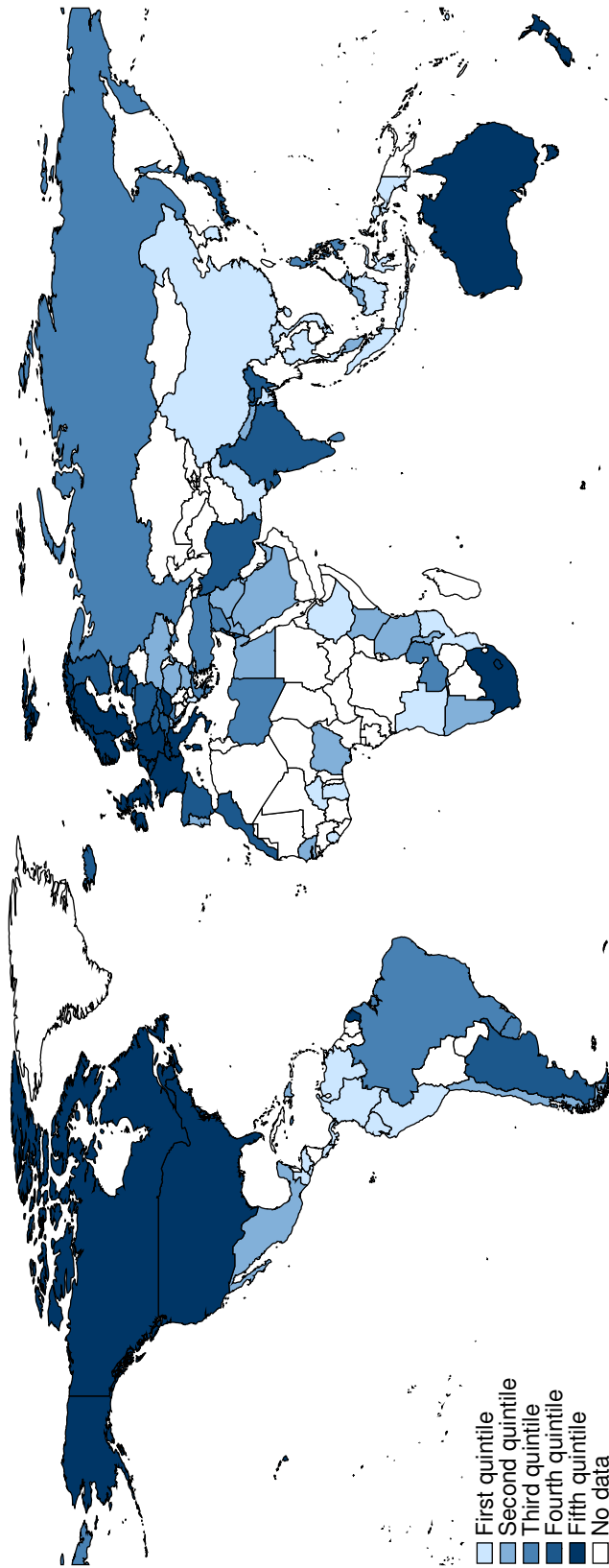
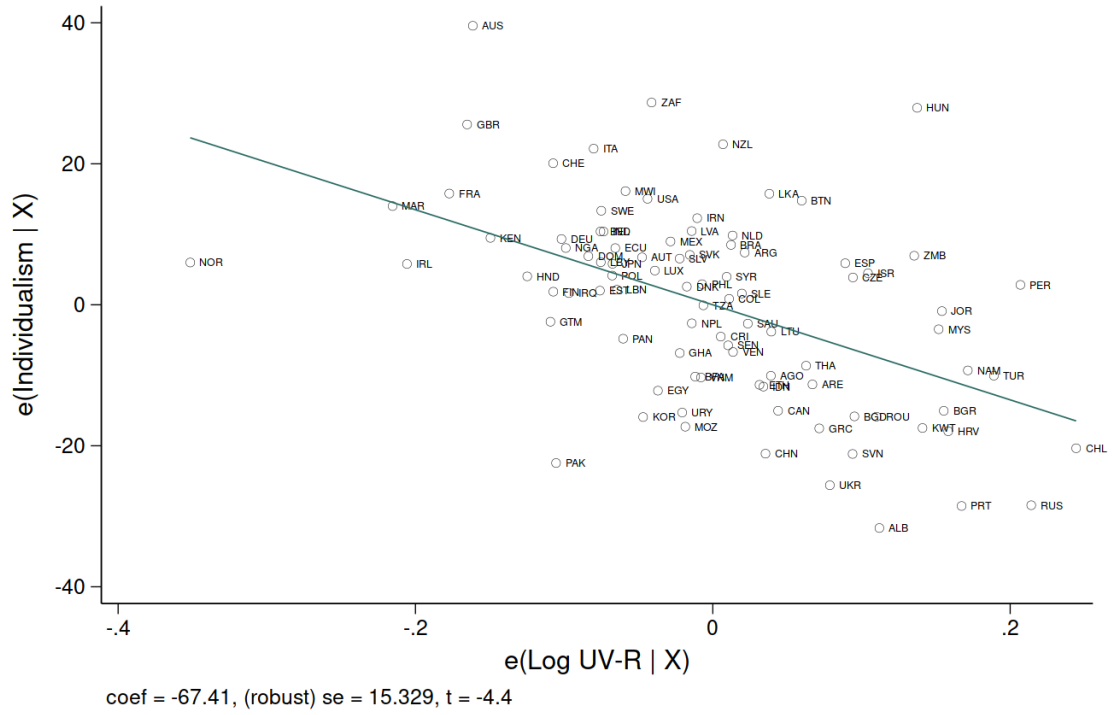
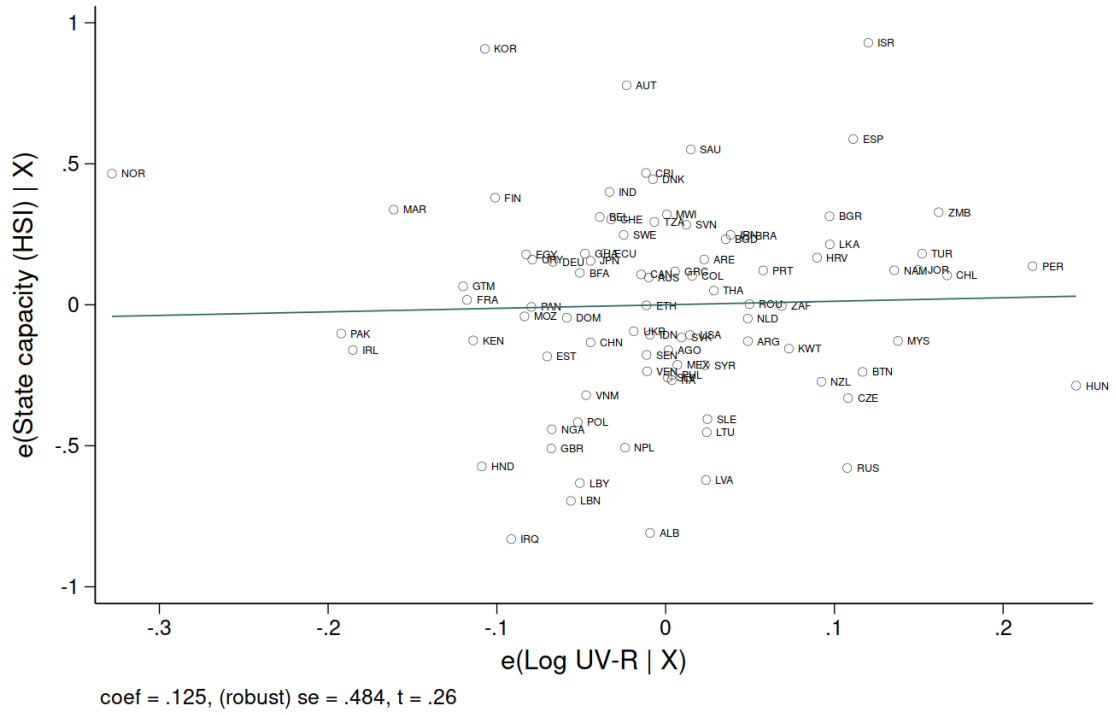


Figure A5: UV-R and individualism: Partial regression plot.



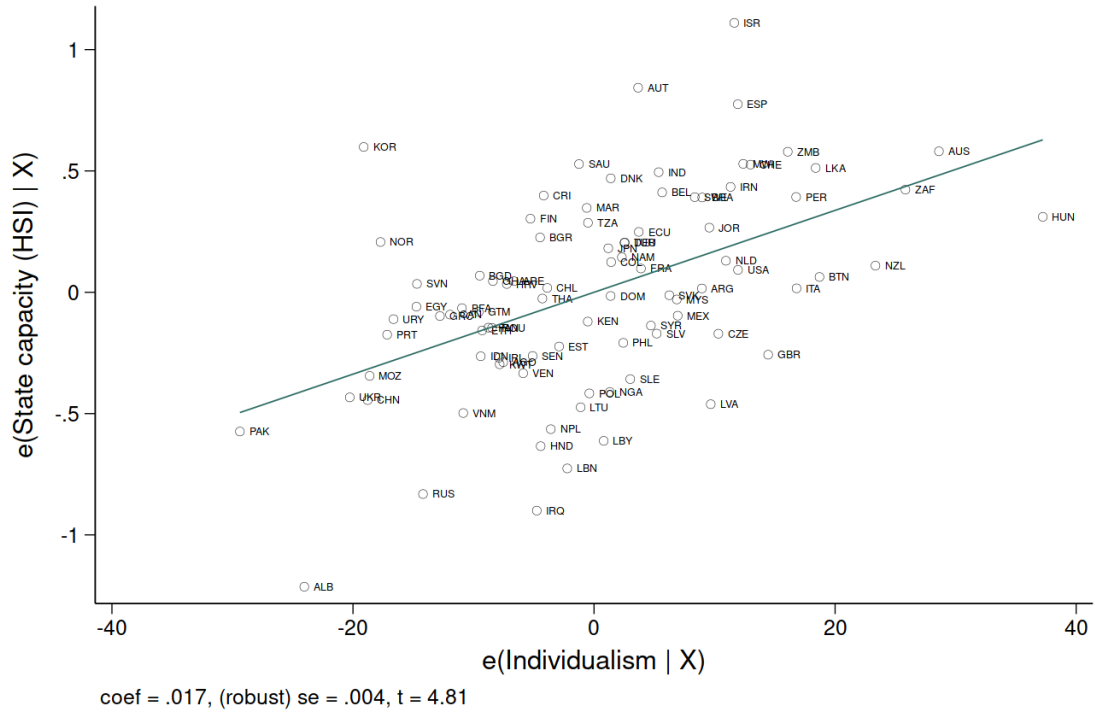
Note: Partial regression plot conditional on the full set of controls described in section 4.1 and regional fixed effects.

Figure A6: UV-R and state capacity: Partial regression plot controlling for individualism.



Note: Partial regression plot conditional on the Hofstede's index of individualism, the full set of controls described in section 4.1, and regional fixed effects.

Figure A7: Individualism and state capacity: Partial regression plot.



Note: Partial regression plot conditional on the degree of UV-R exposure, the full set of controls described in section 4.1, and regional fixed effects.

Table A2: Variables used by Hanson and Sigman (2021) to calculate their State Capacity Index.

<i>Variable</i>	<i>Source</i>
Administrative efficiency	Adelman and Morris (1967)
Bureaucratic quality	Political Risk Services
Census frequency	Hanson and Sigman (2021)
Efficiency of revenue mobilization	World Bank CPIA
Fiscal capacity	Varieties of Democracy dataset
Information capacity	Brambor <i>et al.</i> (2020)
Law and order	Political Risk Services
Military expenditure per capita	COW and SIPRI
Military personnel per capita	COW and World Development Indicators
Monopoly on use of force	Bertlesmann Transformation Index
Police officers per capita	United Nations
Quality of budgetary and financial management	World Bank CPIA
Quality of public administration	World Bank CPIA
Rigorous and impartial public administration	Varieties of Democracy dataset
State antiquity index	Bockstette <i>et al.</i> (2002)
State authority over territory	Varieties of Democracy dataset
Statistical capacity	World Bank
Taxes on income as a percentage of tax revenue	ICTD and IMF
Taxes on international trade as a percentage of tax revenue	ICTD and IMF
Total tax as a percentage of GDP	ICTD, IMF and OECD
Weberianness	Rauch and Evans (2000)

Note: See Hanson and Sigman (2021) for further details.

Table A3: UV-R and state capacity: Alternative measures of population diversity.

	State capacity					
	(1)	(2)	(3)	(4)	(5)	(6)
Log UV-R	-1.669*** (0.394)	-1.589*** (0.429)	-1.478*** (0.470)	-1.820*** (0.392)	-1.694*** (0.419)	-1.560*** (0.463)
Number of ethnic groups	0.046 (0.094)			0.056 (0.094)		
Ethnolinguistic fractionalization (Fearon)		0.243 (0.353)			0.272 (0.350)	
Ethnolinguistic polarization (Fearon)		-0.115 (1.041)			-0.800 (1.047)	
Ethnolinguistic fractionalization (Ethnologue)			-0.018 (0.257)			-0.072 (0.264)
Ethnolinguistic polarization (Ethnologue)			-0.292 (1.076)			-0.773 (1.100)
Genetic diversity				-5.408 (3.435)	-6.080 (3.735)	-4.981 (3.833)
Geographical controls	Yes	Yes	Yes	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.687	0.726	0.709	0.693	0.733	0.713
Observations	142	129	130	142	129	129

Notes: OLS estimates. The dependent variable is in all cases the state capacity index constructed by Hanson and Sigman (2021) (see section 3 for further details). All regressions include a constant term (not displayed), as well as the complete set of geographical controls described in section 4.1. Heteroskedasticity-robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A4: Determinants of UV-R exposure.

	Log UV-R					
	(1)	(2)	(3)	(4)	(5)	(6)
Absolute latitude	-0.379*** (0.047)	-0.370*** (0.046)		-0.067*** (0.018)	-0.073*** (0.019)	-0.065*** (0.018)
Elevation		0.098** (0.043)		0.186*** (0.025)	0.074** (0.031)	0.098*** (0.026)
Temperature			0.038*** (0.010)	0.044*** (0.007)	0.046*** (0.007)	0.045*** (0.005)
Precipitation			-0.093** (0.038)	-0.132*** (0.031)	-0.096*** (0.033)	-0.125*** (0.038)
Frost days			-0.008 (0.008)	-0.003 (0.006)	-0.003 (0.005)	0.004 (0.004)
Tropical climate			0.256*** (0.079)	0.207*** (0.071)	0.162** (0.066)	0.082 (0.059)
Terrain ruggedness					0.093*** (0.019)	0.071*** (0.020)
Area					0.024* (0.014)	-0.012 (0.013)
Distance to the nearest waterway					-0.178*** (0.058)	-0.113** (0.056)
Land suitability for agriculture					0.072 (0.088)	0.018 (0.089)
Regional fixed effects	No	No	No	No	No	Yes
R-squared	0.521	0.548	0.767	0.881	0.910	0.941
Observations	143	143	143	143	143	143

Notes: OLS estimates. The dependent variable is in all cases the logarithm of the measure of UV-R exposure described in section 3. All regressions include a constant term (not displayed). Heteroskedasticity-robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A5: UV-R and state capacity: Outliers and influential countries.

	State capacity						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log UV-R	-1.688*** (0.430)	-1.629*** (0.333)	-1.288*** (0.479)	-1.555*** (0.524)	-1.748*** (0.398)	-1.673*** (0.382)	-1.166*** (0.437)
Omitted countries	Norway and Somalia	$ DFBETA $ $> 2/\sqrt{n}$	None (Median reg.)	Lowest UV-R	Highest UV-R	Lowest st. capac.	Highest st. capac.
Geographical controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population diversity controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.696	0.741	0.456 ^a	0.633	0.707	0.713	0.632
Observations	141	131	143	135	135	135	135

Notes: OLS estimates with the exception of column 3 (median regression). The dependent variable is in all cases the state capacity index constructed by Hanson and Sigman (2021) (see section 3 for further details). All regressions include a constant term (not displayed), as well as the complete set of geographical and population diversity controls described in section 4.1. Heteroskedasticity-robust standard errors in parentheses. ^a Pseudo R-squared. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A6: UV-R, and fiscal and legal capacity.

	Measures of fiscal capacity			Measures of legal capacity		
	Tax reve- nue share in GDP	Income tax share	Fiscal Capacity Index	Rule of law	Regula- tory qua- lity	Impartial public administ.
	(1)	(2)	(3)	(4)	(5)	(6)
Log UV-R	-0.070*	-0.253***	-1.927***	-2.001***	-1.625***	-3.187***
	(0.040)	(0.070)	(0.608)	(0.444)	(0.414)	(0.687)
Geographical controls	Yes	Yes	Yes	Yes	Yes	Yes
Population diversity controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.536	0.359	0.544	0.634	0.632	0.572
Observations	142	142	143	143	143	143

Notes: OLS estimates. See section 4.3.2 and the Supporting Information for further details about the various measures of state capacity. All regressions include a constant term (not displayed), as well as the complete set of geographical and population diversity controls described in section 4.1. Heteroskedasticity-robust standard errors in parentheses. ^a *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A7: Measures of state capacity: Pairwise correlations.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) State Capacity Index (HSI) ^a	1.000										
(2) State Fragility Index ^b	0.854	1.000									
(3) State Capacity Index ^c	0.871	0.776	1.000								
(4) Fragile States Index ^b	0.878	0.870	0.840	1.000							
(5) Government effectiveness	0.932	0.842	0.874	0.897	1.000						
(6) Tax revenue share in GDP	0.653	0.586	0.586	0.536	0.568	1.000					
(7) Income tax share	0.460	0.288	0.380	0.415	0.474	0.228	1.000				
(8) Fiscal capacity index	0.793	0.679	0.681	0.662	0.725	0.512	0.297	1.000			
(9) Rule of law	0.897	0.823	0.897	0.905	0.967	0.553	0.425	0.689	1.000		
(10) Regulatory quality	0.890	0.830	0.854	0.871	0.943	0.528	0.372	0.731	0.938	1.000	
(11) Impartial public administration	0.865	0.751	0.943	0.822	0.864	0.575	0.373	0.677	0.878	0.860	1.000

Notes: See section 4.3.2 and Supporting Information for further details about the various measures of state capacity. ^a Hanson and Sigman (2021). ^b The index has been rescaled, so that higher values indicate higher state capacity. ^c O'Reilly and Murphy (2022).

Table A8: UV-R and state capacity: Conley (1999) standard errors.

	State capacity					
	Cut-off 500 km.	Cut-off 1,000 km.	Cut-off 1,500 km.	Cut-off 2,000 km.	Cut-off 2,500 km.	Cut-off 3,000 km.
	(1)	(2)	(3)	(4)	(5)	(6)
Log UV-R	-1.782*** (0.368)	-1.782*** (0.391)	-1.782*** (0.407)	-1.782*** (0.426)	-1.782*** (0.434)	-1.782*** (0.435)
Geographical controls	Yes	Yes	Yes	Yes	Yes	Yes
Population diversity controls	Yes	Yes	Yes	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.687	0.687	0.687	0.687	0.687	0.687
Observations	143	143	143	143	143	143

Notes: OLS estimates. The dependent variable is in all cases the state capacity index constructed by Hanson and Sigman (2021) (see section 3 for further details). All regressions include a constant term (not displayed), as well as the complete set of geographical and population diversity controls described in section 4.1. Conley (1999) standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A9: UV-R and state capacity: Additional geographical controls and continent fixed effects.

	State capacity			
	(1)	(2)	(3)	(4)
Log UV-R	-1.782*** (0.399)	-1.900*** (0.538)	-2.307*** (0.464)	-2.193*** (0.520)
Dry climate			0.007 (0.006)	0.008 (0.006)
Temperate climate			0.006 (0.005)	0.004 (0.005)
Continental climate			-0.006 (0.005)	-0.008 (0.005)
Tropical climate			-0.401 (0.338)	-0.259 (0.350)
Climate variability			0.078 (1.227)	0.168 (1.221)
Absolute latitude			-0.178 (0.246)	-0.213 (0.252)
Absolute latitude, squared			-0.045 (0.104)	-0.020 (0.107)
Island			0.637*** (0.232)	0.676** (0.267)
Landlocked			-0.054 (0.158)	-0.058 (0.158)
Inequality in land suitability			-1.790*** (0.599)	-1.747*** (0.622)
Irrigation potential			-1.502*** (0.477)	-1.404*** (0.495)
Growing season days			-1.617** (0.689)	-1.688** (0.717)
Cereal/plantation crops			-0.059 (0.268)	-0.069 (0.278)
Ecological fractionalization			0.466 (0.409)	0.525 (0.402)
Geographical controls	Yes	Yes	Yes	Yes
Population diversity controls	Yes	Yes	Yes	Yes
Regional fixed effects	Yes	No	Yes	No
Continent fixed effects	No	Yes	No	Yes
R-squared	0.687	0.700	0.752	0.760
Observations	143	143	142	142

Notes: OLS estimates. The dependent variable is in all cases the state capacity index constructed by Hanson and Sigman (2021) (see section 3 for further details). All regressions include a constant term (not displayed), as well as the complete set of geographical and population diversity controls described in section 4.1. Heteroskedasticity-robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A10: UV-R, cataracts, skin cancer, and individualism.

	Cataracts			Skin cancer			Individualism		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log UV-R	2.298*** (0.167)	1.352* (0.720)	1.866** (0.742)	-1.278*** (0.112)	-0.484 (0.529)	-0.592 (0.598)	-66.402*** (14.885)		-56.883*** (14.581)
Cataractst								-7.857*** (2.859)	-5.102*** (2.116)
Geographical controls	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Population diversity controls	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Regional fixed effects	No	No	Yes	No	No	Yes	Yes	Yes	Yes
R-squared	0.592	0.818	0.861	0.450	0.602	0.722	0.726	0.681	0.746
Observations	90	90	90	90	90	90	90	90	90

Notes: OLS estimates. In columns 1-3 the dependent variable is the prevalence of cataracts (in logs). In columns 4-6 the dependent variable is the prevalence of skin cancer (in logs). In columns 7-9 the dependent variable is the Hofstede's index of individualism (see section 5). All regressions include a constant term (not displayed). With the exception of columns 1 and 4, all regressions include the complete set of geographical and population diversity controls described in section 4.1, as well as the median age of the population. Heteroskedasticity-robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A11: Cataracts, individualism, and other diseases.

	Individualism				
	(1)	(2)	(3)	(4)	(5)
Cataracts	-7.267*** (2.341)	-7.924*** (2.360)	-6.401** (2.905)	-7.973*** (2.420)	-6.208** (3.055)
Malaria	-3.717* (2.185)				-3.617 (2.227)
Trachoma		-0.869 (1.450)			-0.628 (1.485)
Hookworm disease			-3.229 (3.781)		-2.618 (3.889)
HIV				0.291 (1.297)	0.144 (1.309)
Geographical controls	Yes	Yes	Yes	Yes	Yes
Population diversity controls	Yes	Yes	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes	Yes	Yes
R-squared	0.694	0.682	0.684	0.681	0.698
Observations	90	90	90	90	90

Notes: OLS estimates. The dependent variable is in all cases the Hofstede's index of individualism (see section 5). All regressions include a constant term (not displayed), as well as the complete set of controls described in section 4.1 and the median age of the population. Heteroskedasticity-robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A12: Cultural variables: Pairwise correlations.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) Individualism	1.000													
(2) Uncertainty avoidance	-0.336	1.000												
(3) Willingness to take risks	0.018	-0.337	1.000											
(4) Trust	0.472	-0.527	-0.227	1.000										
(5) Positive reciprocity	-0.125	-0.080	-0.185	0.206	1.000									
(6) Negative reciprocity	0.011	0.044	0.128	0.151	-0.089	1.000								
(7) Altruism	-0.211	-0.256	0.175	0.059	0.736	-0.036	1.000							
(8) Long term orientation	0.155	0.072	-0.438	0.363	-0.128	0.232	-0.219	1.000						
(9) Masculinity	0.082	0.084	-0.019	-0.205	-0.140	-0.101	-0.044	-0.049	1.000					
(10) Egalitarianism	0.314	-0.016	0.200	-0.026	0.015	-0.187	0.030	-0.292	-0.160	1.000				
(11) Harmony	0.307	0.130	-0.357	0.150	-0.249	-0.147	-0.523	0.268	-0.022	0.419	1.000			
(12) Mastery	-0.247	-0.130	0.123	0.016	0.028	0.166	0.301	0.151	0.117	-0.372	-0.572	1.000		
(13) Hierarchy	-0.478	-0.112	0.187	-0.138	0.069	0.287	0.282	0.017	0.136	-0.583	-0.635	0.556	1.000	
(14) Religiosity	-0.232	0.197	0.273	-0.548	-0.014	-0.205	0.223	-0.669	0.109	0.230	-0.261	-0.065	-0.047	1.000

Table A13: UV-R, individualism, and state capacity: Robustness to treating the individualism measure as endogenous.

	State capacity			
	OLS		2SLS	
	(1)	(2)	(3)	(4)
Individualism	0.016*** (0.003)	0.017*** (0.004)	0.014*** (0.004)	0.013** (0.005)
Log UV-R		0.125 (0.484)		-0.104 (0.528)
Geographical controls	Yes	Yes	Yes	Yes
Population diversity controls	Yes	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes	Yes
First-stage F-statistic			16.11	14.68
Overidentification test p-value			0.324	0.323
R-squared	0.840	0.840	0.838	0.838
Observations	90	90	90	90

Notes: The dependent variable is in all cases the state capacity index constructed by Hanson and Sigman (2021) (see section 3 for further details). In columns 3 and 4 the individualism index is instrumented using the degree of historical pathogen prevalence, and the Mahalanobis distance between the frequency of blood types A and B in a given country and their frequency in the United States. All regressions include a constant term (not displayed), as well as the complete set of geographical and population diversity controls described in section 4.1. Heteroskedasticity-robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

References

- Acemoglu, D., Johnson, S. and Robinson, J. A. (2005): Institutions as a fundamental cause of long-run growth. In Aghion, P. and Durlauf, S. N. (eds.): *Handbook of Economic Growth, Vol IA*, pp. 386-472. Elsevier North-Holland, Amsterdam.
- Adelman, I. and Morris, C. T. (1967): *Society, Politics, and Economic Development: A Quantitative Approach*. Johns Hopkins University Press, Baltimore.
- Andersen, T. B., Dalgaard, C. J. and Selaya, P. (2016): Climate and the emergence of global income differences, *Review of Economic Studies* 83, 1334-1363.
- Arbath, C. E., Ashraf, Q. H., Galor, O and Klemp, M. (2020): Diversity and conflict, *Econometrica* 88 (2), 727-797.
- Ashraf, Q. H. and Galor, O. (2011): Dynamics and stagnation in the Malthusian Epoch, *American Economic Review* 101 (5), 2003-2041.
- Ashraf, Q. H. and Galor, O. (2013): The “Out of Africa” hypothesis, human genetic diversity, and comparative economic development, *American Economic Review* 103 (1), 1-46.
- Bentzen, J. S., Kaarsen, N. and Wingender, A. M. (2017): Irrigation and autocracy, *Journal of the European Economic Association* 15 (1), 1-53.
- Besley, T. and Persson, T. (2011): *The Pillars of Prosperity*. Princeton University Press, Princeton (NJ).
- Bockstette, V., Chanda, A. and Putterman, L. (2002): States and markets: The advantage of an early start, *Journal of Economic Growth* 7 (4): 347–69.
- Borcan, O., Olsson, O. and Putterman, L. (2018): State history and economic development: Evidence from six millennia, *Journal of Economic Growth* 23 (1), 1-40.
- Brambor, T., Goenaga, A., Lindvall, J. and Teorell, J. (2020): The lay of the land: Information capacity and the modern state, *Comparative Political Studies* 53 (2), 175-213.
- Comin, D., Easterly, W. and Gong, E. (2010): Was the wealth of nations determined in 1000 BC?, *American Economic Journal: Macroeconomics* 2 (3), 65-97.
- Conley, T. G. (1999): GMM estimation with cross sectional dependence, *Journal of Econometrics* 92 (1), 1-45.

- Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J. *et al.* (2021): V-Dem Methodology v11.1. Varieties of Democracy (V-Dem) Project.
- Desmet, K., Ortuño-Ortín, I. and Wacziarg, R. (2012): The political economy of linguistic cleavages, *Journal of Development Economics* 97, 322-338.
- Duclos, J. Y., Esteban, J. and Ray, D. (2004): Polarization: Concepts, measurements, estimation, *Econometrica* 72 (4), 819-851.
- Enke, B. (2019): Kinship, cooperation, and the evolution of moral systems, *Quarterly Journal of Economics* 134 (2), 953-1019.
- Esteban, J. and Ray, D. (2011): Linking conflict to inequality and polarization, *American Economic Review* 101 (4), 1345-1374.
- Esteban, J., Mayoral, L. and Ray, D. (2012): Ethnicity and conflict: An empirical study, *American Economic Review* 102 (4), 1310-1342.
- Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D. and Sunde, U. (2018): Global evidence on economic preferences, *Quarterly Journal of Economics* 133 (4), 1645-1692.
- Fearon, J. D. (2003): Ethnic and cultural diversity by country, *Journal of Economic Growth* 8 (2), 195-222.
- Fedderke, J. W., Klitgaard, R. E. and Napolioni, V. (2017): Genetic adaptation to historical pathogen burdens, *Infection, Genetics and Evolution* 54, 299-307.
- Fund for Peace (2019): *Fragile States Index Annual Report 2019*. Fund for Peace, Washington DC.
- Gorodnichenko, Y. and Roland G. (2017): Culture, institutions and the wealth of nations, *Review of Economics and Statistics* 99 (3), 402-416.
- Hanson, J. K. and Sigman, R. (2021): Leviathan's latent dimensions: Measuring state capacity for comparative political research, *The Journal of Politics* 83 (4), 1495-1510.
- Hofstede, G. (2001): *Culture's Consequences: Comparing Values, Behaviors, and Organizations across Nations*. Sage, Thousand Oaks (CA).
- Kaufmann, D., Kraay, A. and Mastruzzi, M. (2011): The Worldwide Governance Indicators: Methodology and analytical issues, *Hague Journal on the Rule of Law* 3, 220-246.

- Marshall, M. G. and Elzinga-Marshall, G. (2017): *Global Report 2017: Conflict, Governance, and State Fragility*. Center for Systemic Peace, Vienna (VA).
- McEvedy, C. Jones, R. (1978): *Atlas of World Population History*. Penguin Books, New York.
- Michalopoulos, S. (2012): The origins of ethnolinguistic diversity, *American Economic Review* 102 (4), 1508-1539.
- Murray, D. R. and Schaller, M. (2010): Historical prevalence of infectious diseases within 230 geopolitical regions: A tool for investigating origins of culture, *Journal of Cross-Cultural Psychology* 41 (1), 99-108.
- Nunn, N. and Puga, D. (2012): Ruggedness: The blessing of bad geography in Africa, *Review of Economics and Statistics* 94 (1), 20-36.
- O'Reilly, C. and Murphy, R. (2022): An index measuring state capacity, 1789-2018, *Economica* 89, 713-745.
- Putterman, L. (2008): Agriculture, diffusion, and development: Ripple effects of the Neolithic Revolution, *Economica* 75 (300), 729-748.
- Putterman, L. and Weil, D. N. (2010): Post-1500 Population flows and the long-run determinants of economic growth and inequality, *Quarterly Journal of Economics* 125, 1627-1682.
- Ramankutty, N., Foley, J. A., Norman, J. and McSweeney, K. (2002): The global distribution of cultivable lands: Current patterns and sensitivity to possible climate change, *Global Ecology and Biogeography* 11 (5), 377-392.
- Rauch, J. and Evans, P. (2000): Bureaucratic structure and bureaucratic performance in less developed countries, *Journal of Public Economics* 75 (1), 49-71.
- Riley, S. J., DeGloria, S. D. and Elliot, R. (1999): A terrain ruggedness index that quantifies topographic heterogeneity, *Intermountain Journal of Sciences* 5 (1-4), 23-27.
- Schwartz, S. H. (1994): Beyond individualism/collectivism: New cultural dimensions of values. In Uichol, K., Triandis, H.C., Kagitcibasi, C., Choi, S.C. and Yoon, G. (eds.): *Individualism and Collectivism: Theory, Method, and Applications*, pp. 81-119. Sage, Thousand Oaks (CA).
- WHO (2008): *The Global Burden of Disease: 2004 Update*. World Health Organization, Geneva.