Journal of Development Economics 118 (2016) 88–111

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Regular Article

The wealth of subnations: Geography, institutions, and within-country development☆

Todd Mitton

Department of Finance, Marriott School, Brigham Young University, United States

ARTICLE INFO

Article history:
Received 30 June 2014
Received in revised form 3 September 2015
Accepted 4 September 2015
Available online 16 September 2015

Keywords:
Geography
Institutions
Development

ABSTRACT

I study determinants of economic development in a new dataset covering 1867 subnational regions from 101 countries, focusing on within-country effects of geography and institutions. Several geographic factors have significant explanatory power for within-country differences in per-capita GDP, including terrain ruggedness, tropical climate, ocean access, temperature range, storm risk, and natural resources such as oil, diamonds, or iron. Institutions have a significant positive effect on income among subnational regions with greater autonomy, suggesting that strong subnational institutions enhance development when not dominated by national institutions.

1. Introduction

The state of Nuevo León in México has a vibrant, industrialized economy with a relatively wealthy population. With a per-capita GDP of $22,051, its level of economic development is similar to that of Portugal or South Korea.[1] Meanwhile, the state of Chiapas in México has a relatively unproductive, agricultural economy, and is much poorer than Nuevo León. With a per-capita GDP of $5330, Chiapas has a level of development similar to that of Namibia or Ukraine. Nuevo León and Chiapas have both been states of México since 1824, so why has the economy of Nuevo León prospered while the economy of Chiapas has stagnated?

In this study, I seek to better understand the impact of geography and institutions on the divergent fortunes of subnational regions within countries. An extensive literature debates the relative importance of geography and institutions for economic development. For example, Gallup et al. (1999), Sachs (2003), and Diamond (1997) argue for the importance of geography; Mauro (1995), Hall and Jones (1999), and Rodrik et al. (2004) argue that institutions are key; and Acemoglu et al. (2001, 2002), and Easterly and Levine (2003) argue that geography matters primarily to the extent that it shapes institutions.[2] However, with some exceptions (discussed below) this literature has focused on cross-country differences in development. One way to extend our understanding of the roles of geography and institutions on development is to exploit differences in these variables within countries. To this end, I assemble a comprehensive dataset of variables measuring income, geography, and institutions at the subnational level and study within-country relations between the variables.

For a within-country study to be informative, the first requirement is that the variables have enough within-country variation to capture meaningful differences. Regarding the dependent variable, my dataset shows that subnational variation in per-capita income is typically quite large. In the example above, the income of Nuevo León exceeds that of Chiapas by a factor of four, but this is by no means an extreme case of within-country disparity in development. In the United States, the District of Columbia has a per-capita GDP that exceeds that of the state of Mississippi by a factor of five. In Indonesia, the East Kalimantan province is wealthier than the North Maluku province by a factor of 21. In the Russian Federation, Nenets Oblast is wealthier than Ingushetia Republic by a factor of 69. On average, for the countries that I study, the ratio of per-capita GDP in the richest subnational region to per-capita GDP in the poorest subnational region is 6.3 (with a median of 3.3). Of course, subnational data do not capture all within-country variation in income (such as differences between cities in the same subnational region), but this approach represents an important step toward understanding the variation in development within countries.

The explanatory variables also exhibit a large amount of variation at the regional level. I find important regional differences in geography that country-level studies fail to capture, some of which are obvious (e.g., Argentina has large subnational variation in latitude), and some of which are more subtle (e.g., Tanzania has large subnational variation in malaria risk). I also find a great deal of within-country variation in

☆ I acknowledge research funding from the Ballard Center for Economic Self-reliance. For assistance with ArcGIS I thank Mark Jackson. For helpful comments I thank two anonymous referees, Nathan Nunn (editor), Gary Cornia, Simon Johnson, Grant McQueen, Catherine Moyes, Taylor Nadauld, Ryan Pratt, and Robert Schonlau. For excellent research assistance I thank Zilu Chen, Elizaveta Ezhova, Duc Ho, Zhouyang Lai, Christian Mealey, Jay Muir, Winnie Nguyen, Vincent Powell, and Cameron Thorpe.

E-mail address: tm@byu.edu.

[1] Per-capita GDP figures quoted are from 2005 and are PPP adjusted.

[2] Other contributions include Frankel and Romer (1999) who show the importance of trade (determined in large part by geography); and Dollar and Kraay (2003) who argue for the joint importance of trade and institutions. Acemoglu (2009), while emphasizing the role of institutions, offers a comprehensive review of these issues.

http://dx.doi.org/10.1016/j.jdeveco.2015.09.002
0304-3878/

subnational measures of institutions; country dummies explain only about half of the variation in the institutional variables. My dataset is comprised of an average of 18.5 subnational regions for each country, which enables me to capture much more of the complexity of the determinants of development around the world.

To assemble the dataset, I start with subnational-level measures of per-capita GDP for 2005 as the dependent variable. These data are compiled from each country separately, typically from national statistical agencies or other governmental organizations. The fact that the methods of data compilation are not standardized across countries could be problematic in a country-level study, but is of limited concern in my analysis because I use country fixed effects. Subnational per-capita GDP estimates are not available for all countries of the world, but the 101 countries in my sample constitute 96% of the world's GDP. Gennaioli et al. (2013) assemble similar subnational GDP estimates for 110 countries, although due to differing levels of aggregation they report GDP estimates for 1569 subnational regions as compared to the 1867 in my dataset.

To create subnational geographic variables I employ a number of geospatial datasets maintained by governmental and private organizations that offer detailed global mappings of a wide variety of general geographic characteristics (e.g., elevation or terrain ruggedness), climate (e.g., precipitation or temperature), and natural resources (e.g., oil or iron). The resulting set of 25 distinct geographical variables greatly expands upon the number of geographic factors tested in other subnational studies of development, which typically test between three and seven geographic factors (e.g., Dell et al., 2009; Gennaioli et al., 2013; Nordhaus, 2006; Warner, 2002). Importantly, my 25 geographic variables are generally not redundant, with an average pairwise correlation of only 0.15.

To create subnational institutional variables, I employ data from several different organizations that survey companies and individuals regarding their local institutional environment. The surveys, in total encompassing over 170,000 respondents from 86 countries, each identify the subnational region of respondents, allowing me to create subnational-level measures of institutional quality. I categorize responses from over 100 survey questions to study institutions in the areas of property rights protection, corruption control, law and order, and regulatory efficiency. The resulting set of variables constitutes a much more comprehensive dataset of subnational institutional measures than employed in prior research. For example, Gennaioli et al. (2013) use only eight survey questions to create a single institutions variable for a subsample of some 500 subnational regions.

My results can be summarized as follows. I find a significant role for geography in shaping within-country differences in development. The most robust and significant effects (both statistically and economically) come from terrain ruggedness (a negative effect), ocean access (positive), storm risk (negative), and a high temperature range (negative). Other factors that are important but that are either less robust or not as economically significant include latitude (positive), elevation (negative), and the abundance of iron (positive), alloyants (positive), oil and gas (positive), or diamonds (positive). Many of the significant factors have not been previously tested in other subnational studies of development. The importance of geographic factors persists in a series of robustness tests, including when I control for subnational measures of human capital (measured by educational attainment), or culture (measured by ethnic groups or language groups).

In contrast, I find that, in general, institutions do not play a positive role in shaping within-country differences in development. The correlation between institutions and income is usually negative, and at times the negative correlation is statistically significant. The lack of a general positive within-country effect of institutions persists across different areas of the world and for different types of institutions, i.e., property rights, corruption control, law and order, or regulatory efficiency. Additional tests suggest that the negative correlation between institutions and development is not driven by data quality or measurement issues. Two other tests suggest that the lack of a positive within-country effect of institutions may be due to the dominance of national institutions over subnational institutions. First, in regressions of income on measures of both national and subnational institutions, national institutions have much greater explanatory power, both in terms of coefficient magnitudes and r-squareds. Finally, as an additional way of assessing the dominance of national institutions over regional institutions, I exploit variation in the level of autonomy granted to regional governments. Using various proxies for autonomy or federalism, I find a significant positive relation between institutions and development in regions that have greater autonomy, suggesting that strong regional institutions may aid development when not dominated by national institutions.

In contrast to the large volume of cross-country research in this area, only a handful of other studies test the relation between geography, institutions, and development at the subnational level. Gennaioli et al. (2013) focus on the effect of human capital on regional development, but they also find a significant effect of three geographic factors and no significant effect of an institutional measure. Acemoglu et al. (2014) present within-country results while challenging the conclusions of Gennaioli et al. (2013). Acemoglu and Dell (2010), while not measuring local institutions per se, show that access to paved roads is associated with higher incomes at the municipality level for 11 countries in the Americas. Michalopoulos and Papaioannou (2013) show the effect of ethnic institutions on regional development (as proxied by satellite images of light density) in Africa. Bruhn and Gallego (2012) study the effect of colonial activities on subnational development for 17 countries in the Americas.[3] A few other papers study the effect of geography (but not institutions) on development for multiple countries at the subnational level. Dell et al. (2009) find that temperature is negatively correlated with income at the subnational level for 12 countries in the Americas. Nordhaus (2006) constructs cell-level (one degree by one degree) measures of geography and income and finds, among other results, that output per area is positively correlated with temperature or latitude. Warner (2002) studies the effect of subnational measures of geography on income in eight countries.

In comparison to these other papers, my study is intended to be a more detailed analysis of the effect of geography and institutions on subnational development, employing a more exhaustive set of geographic variables and a more comprehensive dataset of subnational institutional measures for a large set of countries. In sum, the data analyzed in this study provide the clearest picture to date of the effects of geography and institutions on world development at a disaggregated level. The results contribute to the overall debate on geography vs. institutions by showing that geography has a significant impact on development even after controlling for country-level factors, as well as by showing the conditions under which subnational institutions matter for regional development.

2. Data and summary statistics

To carry out my tests, I need subnational measures of per-capita GDP for the dependent variable and subnational measures of geography and institutions for the explanatory variables. For the geographic variables, global datasets are available to measure geographic features for all subnational regions in my dataset. But for the GDP measurement and for the institutional variables, no single data source measures these variables for all regions of the world, so I assemble the data from multiple sources. Although the sources differ across countries, they are consistent within countries, and country fixed effects control for differences in

[3] A number of related single-country studies report within-country effects of institutional variation on various economic outcomes. These include Banerjee and Iyer (2005), Iyer (2010), Dell (2010), and Acemoglu et al. (2012).

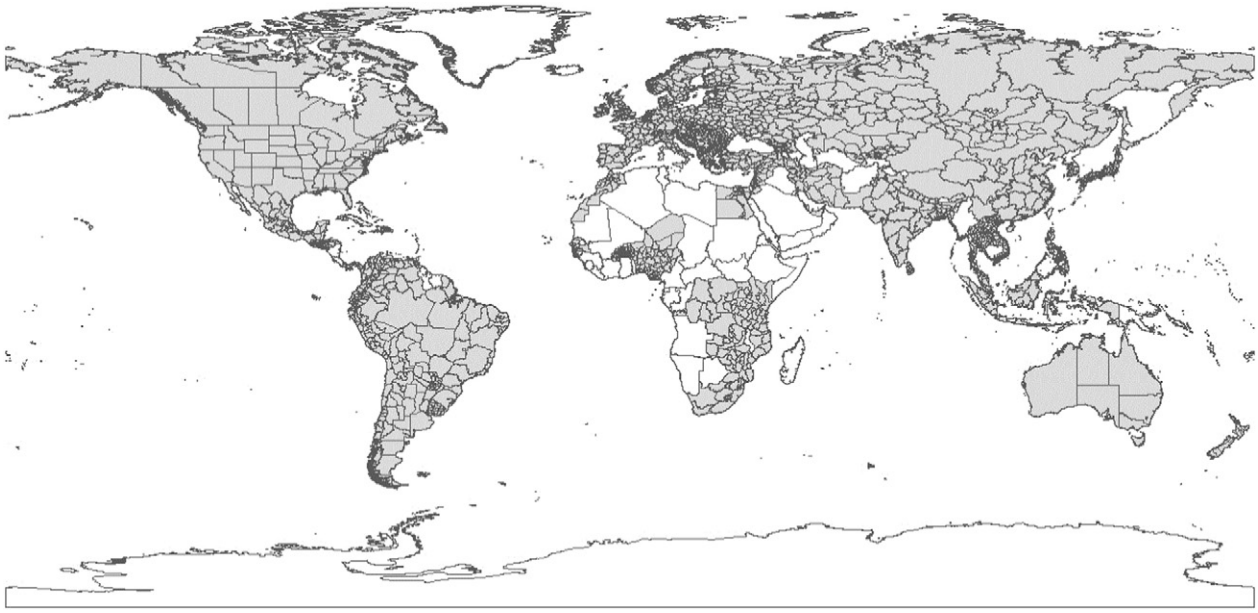

Fig. 1. The world map shows the subnational regions included in the dataset. Unshaded countries are not included in the dataset due to unavailability of subnational GDP data.

measurement across countries. In this section I discuss the compilation of the different types of data.

2.1. Economic data

Data on per-capita GDP for 2005 at the subnational level come from various sources specific to each of the countries in the sample. The most common sources are the national statistical bureaus of the individual countries, which provide data for 35 of the sample countries. Data for 28 European Union nations are provided by Eurostat. Data for 30 nations are provided by the United Nations Development Programs (UNDP) associated with individual countries. Data for the remaining eight countries come from various other sources. Appendix Table A1 provides a complete list of the GDP data sources used for each country. Although not all countries provide subnational GDP estimates, the 101 countries included in my sample represent 96% of the world's GDP and 90% of the world's population. Fig. 1 displays a world map of the subnational regions that are included in the sample. Countries that are not shaded in the figure are not included in the sample. In the figure, Africa stands out as having relatively little coverage, although 18 African nations are included in the sample.

Generally speaking, subnational GDP measures are available for the first-level administrative divisions of the respective countries, e.g., states in the U.S., provinces in China, or departments in Paraguay. In the European Union, subnational GDP is reported for NUTS 2 (Nomenclature of Units for Territorial Statistics) or NUTS 3 regions, which do not always correspond to first-level administrative divisions. Regardless of the administrative level, I use data at the level of disaggregation that is available for each country.[4] Countries that do not report subnational GDP are excluded from the sample. Table 1 reports, for each country in the sample, the type and number of subnational regions included in the dataset. Table 1 also reports the Geographic Income Disparity (GID) for each country, which I define as the ratio of per-capita GDP of the richest region in the country to the per-capita GDP of the poorest region in the country. GID thus gives a rough idea of the equality of geographic wealth distribution within the country. GID is as high as 69.2 in the Russian Federation and as low as 1.3 in Bangladesh.[5]

For most countries in the sample, the respective sources provide the data on a per-capita basis, but for 23 of the countries I scale the provided regional GDP figures using 2005 population estimates that are typically provided by the same source that provided the GDP data. When population estimates are not provided for 2005, I estimate 2005 population by interpolating between estimates given for years before and after 2005.

Most of the countries in the sample provide regional GDP data as of 2005, but 27 of the countries provide GDP data only for other years (reported in Appendix Table A1). In these cases, I convert the estimates to the 2005 estimates using inflation rates (GDP deflators) provided by the World Bank. World Bank GDP deflators are also used for conversion in cases where the estimates are provided in constant prices from a previous year (this is the case in Chile, Kenya, and Panama). All per-capita GDP estimates are adjusted for purchasing power parity using the World Bank PPP adjustment factors. Rows 1 and 2 of Table 2 report summary statistics on per-capita GDP estimates.

2.2. Geographic data

I obtain geographic data for subnational regions in my dataset from a number of different sources. Although these geographic variables are generally available for the entire world, the scope of my dataset is restricted by the availability of subnational GDP data, as discussed above. For expositional purposes, I divide the geographic variables into the categories of general geographic data, climate data, and natural resource data, which I discuss in turn.

2.2.1. General geographic data

Rows 3 through 9 of Table 2 report summary statistics for general geographic variables included in my dataset. The latitude of each

[4] Some countries in the European Union report subnational GDP for multiple NUTS levels, in which case I use the level with the greatest disaggregation.

[5] Caution should be exercised when comparing GID between countries, because GID depends not only on geographic wealth distribution in the country, but also on the number of regions into which the country is divided. The number and definition of regions in a country is not necessarily exogenous; regions could be formed endogenously due to factors such as geography, institutions, or culture (see, e.g., Alesina and Spolaore, 1997; Gries and Naude, 2008).

Table 1
Sample composition.

Country	Obs.	Region type	GID
Albania	12	Counties	4.58
Argentina	24	Provinces	8.69
Australia	8	States	1.62
Austria	9	States	2.01
Bangladesh	7	Divisions	1.26
Belarus	7	Regions	2.46
Belgium	11	Provinces	3.02
Benin	12	Departments	4.06
Bolivia	9	Departments	4.14
Bosnia and Herzegovina	11	Cantons	4.31
Brazil	27	States	9.33
Bulgaria	28	Planning regions	3.28
Burkina Faso	45	Provinces	2.74
Canada	13	Provinces	3.31
Chile	13	Regions	4.81
China	31	Provinces	10.19
Colombia	33	Departments	10.64
Congo, Dem. Rep.	11	Provinces	11.61
Croatia	21	Counties	3.18
Czech Republic	14	Regions	2.77
Denmark	5	Regions	1.67
Dominican Republic	31	Provinces	3.08
Ecuador	22	Provinces	31.38
Egypt	27	Governorates	1.56
El Salvador	14	Departments	2.33
Estonia	5	Statistical regions	2.40
Finland	19	Regions	8.65
France	26	Regions	3.27
Gambia, The	6	Divisions	3.43
Georgia	9	Regions	3.00
Germany	16	States	2.47
Greece	13	Peripheries	1.60
Guatemala	22	Departments	3.34
Honduras	18	Departments	4.21
Hungary	20	Counties	4.27
India	32	States	10.81
Indonesia	33	Provinces	21.66
Iran	30	Ostans	13.16
Ireland	8	Regional authorities	2.11
Italy	21	Regions	2.05
Japan	47	Prefectures	2.73
Jordan	12	Governorates	1.73
Kazakhstan	16	Provinces	10.68
Kenya	8	Provinces	20.42
Kosovo	30	Municipalities	2.54
Kyrgyzstan	9	Regions	4.62
Lao PDR	18	Provinces	2.83
Latvia	6	Planning regions	3.78
Lebanon	6	Governorates	2.61
Lesotho	10	Districts	2.80
Lithuania	10	Counties	3.14
Macedonia	8	Statistical regions	3.55
Malaysia	15	States	7.15
Malta	2	Statistical regions	1.39
Mexico	32	States	17.05
Micronesia	4	States	2.84
Mongolia	22	Aimags	10.71
Montenegro	4	Regions	2.75
Morocco	16	Regions	3.43
Mozambique	11	Provinces	6.47
Nepal	5	Regions	1.62
Netherlands	12	Provinces	1.66
New Zealand	15	Regions	1.95
Niger	8	Departments	3.30
Nigeria	37	States	5.79
Norway	19	Counties	2.76
Pakistan	6	Provinces	7.69
Panama	9	Provinces	7.90
Paraguay	18	Departments	3.14
Peru	24	Regions	11.15
Philippines	17	Regions	14.11
Poland	16	Provinces	2.30
Portugal	7	Statistical regions	1.79
Romania	42	Departments	4.94
Russia	83	Federal subjects	69.17
Senegal	10	Regions	1.60
Serbia	4	Statistical regions	2.84
Slovak Republic	8	Regions	4.14
Slovenia	12	Statistical regions	2.15
South Africa	9	Provinces	3.41
South Korea	16	Regions	3.41
Spain	19	Autonomous communities	1.92
Sri Lanka	9	Provinces	3.43
Swaziland	4	Regions	1.77
Sweden	21	Provinces	1.73
Switzerland	26	Cantons	3.03
Syrian Arab Republic	14	Governorates	1.44
Tanzania	21	Regions	3.36
Thailand	76	Provinces	35.35
Turkey	26	Sub-regions	4.90
Uganda	4	Administrative regions	2.49
Ukraine	27	Oblasts	6.25
United Arab Emirates	7	Emirates	12.34
United Kingdom	37	Statistical regions	4.41
United States	51	States	5.06
Uruguay	19	Departments	2.95
Uzbekistan	14	Regions	4.06
Venezuela	24	States	2.13
Vietnam	63	Provinces	37.36
Zambia	9	Provinces	2.80
Zimbabwe	10	Provinces	2.40
Total	1867	Subnational regions	

The table reports the composition of the sample by country. The number of subnational regions in each country is reported along with the type of region defined for each country. GID (Geographic Income Disparity) is defined as the ratio of per-capita GDP (PPP-adjusted) in the richest region of the country to that of the poorest region in the country. GDP data are obtained from multiple sources as listed in Appendix Table A1.

subnational region is taken from data in Google Earth, and is defined as the absolute value of the latitude of the capital city of the region, expressed in degrees.[6] *Latitude* thus measures distance from the equator, with greater distance from the equator expected to be beneficial to economic development (see, e.g., Gallup et al., 1999; Rodrik et al., 2004). *Elevation* is also taken from Google Earth and is defined as the log of one plus the elevation of the capital city of the region, expressed in thousands of feet. The region in the sample with the highest elevation, at 14,148 ft, is the Pasco region in the Peruvian Andes. *Ocean Access* is a dummy variable that is set equal to one if the subnational region has immediate access to the ocean and to zero if the subnational region is landlocked, where being landlocked would be a disadvantage to economic development to the extent that it limits trade (see, e.g., Frankel and Romer, 1999).

Terrain Ruggedness is as defined in Nunn and Puga (2012), who produce a 30 arc-second grid of terrain ruggedness across the surface of the earth. I take the average of all 30 by 30 arc-second cells in each subnational region (weighted by the actual surface area of each cell) to calculate the terrain ruggedness index for each region in my sample. The most rugged region in the sample is the canton of Uri on the north side of the Swiss Alps. A rugged terrain is generally considered a hindrance to agricultural development and trade. Nunn and Puga (2012) show that, across countries, terrain ruggedness is negatively associated with economic development, except in Africa, where terrain ruggedness also had the beneficial effect of limiting the slave trade.

The next two variables in Table 2, *Storm Risk* and *Earthquake Risk*, measure the susceptibility of subnational regions to natural disasters, which should be a hindrance to development. *Storm Risk* measures the (log) number of occurrences of hurricanes and tropical storms in each subnational region between 1842 and 2010. The data are taken from the National Oceanic Atmospheric Administration's database of over

[6] If the subnational region has no official capital then the largest city in the region (by population) is used as the reference city.

Table 2
Descriptive statistics.

	Mean	Min	25th pctile	Median	75th pctile	Max	St. dev.	N
Economic variables								
(1) GDP per capita	11,644.16	165.00	2589.81	5723.02	18,521.66	142,320.00	13,330.78	1867
(2) Log GDP per capita	8.72	5.11	7.86	8.65	9.83	11.87	1.20	1867
Basic geographic variables								
(3) *Latitude*	31.04	0.01	14.38	33.52	45.50	70.04	17.32	1867
(4) *Elevation*	0.58	−0.08	0.09	0.36	0.90	2.72	0.60	1867
(5) *Ocean Access*	0.37	0.00	0.00	0.00	1.00	1.00	0.48	1867
(6) *Terrain Ruggedness*	0.72	0.00	0.26	0.64	1.10	2.21	0.50	1866
(7) *Storm Risk*	0.68	0.00	0.00	0.00	0.69	7.59	1.34	1866
(8) *Earthquake Risk*	0.61	0.00	0.00	0.00	1.10	4.87	0.92	1866
(9) *Malaria Risk*	2.85	0.00	0.00	0.03	1.99	33.45	6.78	1651
Climate variables								
(10) *Percent Tropical*	0.38	0.00	0.00	0.00	0.99	1.00	0.46	1866
(11) *Precipitation*	90.53	0.06	50.80	76.03	122.93	420.16	56.13	1837
(12) *Rain Days*	11.03	0.02	7.98	11.33	14.24	22.22	4.28	1837
(13) *Temperature*	15.73	−15.60	9.02	14.93	24.31	29.95	8.70	1837
(14) *Temperature Range*	10.31	5.02	8.51	10.09	11.91	17.96	2.33	1837
(15) *Humidity*	71.74	28.34	67.80	74.43	79.17	90.68	10.43	1837
(16) *Sunshine*	50.82	19.59	40.73	51.18	61.04	87.30	13.13	1837
(17) *Frost Days*	5.82	0.00	0.26	4.62	9.99	25.84	5.86	1837
(18) *Wind Speed*	2.95	0.77	2.22	2.88	3.56	7.14	0.99	1837
Natural resource variables								
(19) *Oil and Gas* (number of sites)	126.81	0.00	0.00	0.00	0.00	67,796.00	1898.97	1866
(20) *Diamonds* (number of sites)	0.20	0.00	0.00	0.00	0.00	128.00	3.32	1866
(21) *Precious Metals* (number of sites)	74.09	0.00	0.00	0.00	2.00	29,261.00	961.39	1866
(22) *Base Metals* (per 1000 sq. km.)	0.80	0.00	0.00	0.01	0.36	52.34	3.38	1866
(23) *Iron* (per 1000 sq. km.)	0.15	0.00	0.00	0.00	0.01	33.16	1.15	1866
(24) *Alloyants* (per 1000 sq. km.)	0.16	0.00	0.00	0.00	0.03	14.91	0.80	1866
(25) *Water* (% of land area)	1.45	0.00	0.12	0.44	1.34	44.89	3.22	1866
(26) *Agricultural Land*	0.95	0.00	0.96	1.00	1.00	1.00	0.27	1847
(27) *Soil Quality*	−1.42	−3.80	−1.53	−1.38	−1.25	−1.00	0.27	1847
Institutions variables								
(28) Institutions in Africa	0.05	−1.58	−0.25	0.10	0.50	0.89	0.55	180
(29) Institutions in the Americas	0.11	−2.36	−0.30	0.05	0.47	1.94	0.62	362
(30) Institutions in Asia	0.01	−1.31	−0.36	0.05	0.39	1.40	0.50	147
(31) Institutions in Europe	−0.12	−2.15	−0.68	−0.10	0.66	1.67	0.88	308
(32) Institutions in Latin America	0.02	−1.61	−0.53	−0.12	0.47	2.02	0.72	248
(33) Institutions worldwide (WBES)	0.01	−2.12	−0.22	0.01	0.26	2.92	0.42	619
(34) Institutions worldwide (DB)	0.02	−1.11	−0.29	0.00	0.34	1.15	0.43	230
(35) Institutions combined	0.07	−2.16	−0.28	0.06	0.41	1.94	0.74	1343
(36) *Property Rights*	0.07	−2.50	−0.44	0.06	0.56	3.30	0.74	1130
(37) *Corruption Control*	0.06	−2.99	−0.37	0.12	0.49	5.66	0.74	1232
(38) *Law and Order*	0.07	−2.81	−0.37	0.07	0.57	1.99	0.71	1251
(39) *Regulatory Efficiency*	0.04	−3.26	−0.31	0.02	0.36	2.07	0.58	1041
(40) *Rule of Law* (country level)	0.05	−1.70	−0.78	−0.19	0.83	1.96	0.99	101
(41) *Corruption Control* (country level)	0.06	−1.43	−0.71	−0.27	0.61	2.36	1.02	101
(42) *Stability/Lack of Violence* (country level)	−0.11	−2.18	−0.79	−0.04	0.79	1.58	0.93	99
(43) *Regulatory Quality* (country level)	0.15	−2.25	−0.54	−0.12	1.04	1.69	0.94	100
(44) Institutions combined (country level)	0.04	−1.71	−0.60	−0.24	0.83	1.89	0.92	99
Other variables								
(45) Autonomy	0.06	0.00	0.00	0.00	0.00	1.00	0.24	1867
(46) Autonomy (country level)	0.17	0.00	0.00	0.00	0.00	1.00	0.36	96
(47) State elections (country level)	0.87	0.00	0.00	1.00	1.54	2.00	0.78	85
(48) Municipal elections (country level)	1.38	0.00	0.95	1.75	2.00	2.00	0.70	73
(49) State authority (country level)	0.60	0.00	0.00	1.00	1.00	1.00	0.49	43
(50) Senators elected (country level)	0.63	0.00	0.00	1.00	1.00	1.00	0.47	48
(51) Composite autonomy (country level)	4.91	0.86	4.00	5.62	6.00	7.00	1.87	19
(52) Education	7.16	0.22	4.72	7.27	9.65	13.21	3.23	1613

The table presents descriptive statistics of data used in subsequent tables. Definitions and sources for each variable are provided in Appendix Table A2. For presentation purposes, some natural resource variables are scaled differently in this table relative to subsequent tables (where noted in parentheses).

28,000 recorded storms. *Earthquake Risk* measures the (log) number of distinct fault lines present in each subnational region, and is based on a worldwide catalogue of over 3000 major faults jointly maintained by the United States Geological Survey (USGS) and Esri, the leading worldwide supplier of geodatabase software and applications.

Row 9 of Table 2 reports statistics for *Malaria Risk*, a measure of the risk of malaria transmission, which is considered a major barrier to economic development (see, e.g., Sachs, 2003). My measure of malaria risk is based on the work of Kiszewski et al. (2004), who create a global grid of malaria risk based on climatic factors (temperature and precipitation) and biologic characteristics of locally dominant vector mosquitoes. As the measure is based on predetermined factors, it is arguably exogenous to income levels (see Sachs, 2003). I define *Malaria Risk* for each subnational region as the average of the risk measures of each point estimated on the Kiszewski et al. (2004) map within the region. Table 2 shows that the maximum risk measure in the sample is 33.45, which occurs in the

Bam province of Burkina Faso. Due to the resolution of the Kiszewski et al. (2004) map, only 1651 of the regions in the dataset have a malaria risk measure.

2.2.2. Climate data

Rows 10 through 18 of Table 2 report statistics for geographic variables that examine the climate of subnational regions in depth. The first variable, *Percent Tropical*, measures the percentage, in land area, of each subnational region that falls within a "tropical and subtropical" biome as defined in the World Wildlife Fund's terrestrial ecoregions map. A tropical climate is generally thought to be a disadvantage to economic development due to agricultural productivity disadvantages or disease burden (see, e.g., Bloom and Sachs, 1998). The remaining climate variables in Table 2 examine various features of the climate of subnational regions in more detail. All of these variables come from the high-resolution global gridded dataset developed by the Climatic Research Unit at the University of East Anglia. *Precipitation* is average monthly rainfall, and *Rain Days* measures the average number of days per month with precipitation. *Temperature* is the average daily mean temperature, and *Temperature Range* is the average daily diurnal temperature range (both measured in degrees Celsius). *Humidity* is the average daily relative humidity, and *Sunshine* is the average percent of day length with sunshine. *Frost Days* is the average number of days per month with ground frost, and *Wind Speed* is the average daily windspeed (in m/s). All of these variables are averaged across all points in the gridded dataset that fall within each subnational region. As examples of regions ranking at the extremes of these measures, the Aswan Governorate in Egypt has the lowest average precipitation in the sample (the Chocó department in Colombia has the highest), and the Nunavut Territory in Canada has the lowest average temperature in the sample (the Oudalan province in Burkina Faso has the highest).

2.2.3. Natural resource data

Rows 19 through 27 of Table 2 report statistics for geographic variables that examine natural resources in more detail. I report data for three precious mineral resources, three non-precious mineral resources, and three agricultural resources. Rows 19 through 21 report statistics for the precious mineral resources: oil and gas, diamonds, and precious metals. I obtain data on oil and natural gas from the worldwide dataset of oil and gas fields compiled by the USGS and Petroconsultants International Data Corp. The dataset identifies the location of 236,625 oil and/or gas fields within the subnational regions in my dataset. Although this dataset covers the entire world, all countries do not have equal depth of coverage, so country fixed effects are particularly important for analyzing these data. *Oil and Gas* is defined as the number of oil and gas fields within each subnational region in my dataset. I obtain data on diamonds and precious metals from the Mineral Resources Data System (MRDS) of the USGS. The MRDS is an extensive database derived from a collection of reports describing metallic and non-metallic mineral resources throughout the world. Each mineral site is identified by latitude and longitude and the commodities available at the site. I extract data on the 299,335 mineral sites located within the subnational regions in my sample, among which 179 different mineral resources are identified. *Diamonds* and *Precious Metals* are defined as the number of sites for each resource located within each subnational region.[7] I do not scale precious resources by land area, under the assumption that even a small number of precious resource sites can have a large impact on wealth in the region, regardless of land area.

Rows 22 through 24 report statistics for three non-precious mineral resources included in the study: base (non-ferrous) metals, iron, and steel alloyants. Diamond (1997) emphasizes the importance of steel-making capacity for economic development. Data on these resources are also obtained from the MRDS, and *Base Metals*, *Iron*, and *Alloyants* are defined as the number of sites containing each resource located within each subnational region, scaled by total land area in the region.[8] It is important to note that whereas most geographic variables are exogenous to income levels, these mineral resource variables really measure exploitation of mineral resources rather than availability of mineral resources, and thus could be endogenous to GDP. Absent data that measure availability, the possibility that wealthier regions could have greater ability to exploit mineral resources should be considered when interpreting the results.

Rows 25 through 27 of Table 2 report statistics for variables that are important for agricultural development: water availability, agricultural land availability, and soil quality. Data on water come from the World Water Bodies database maintained by Esri, which identifies the main water features of the world using nearly three million polygon representations on the global map. *Water* is defined as the total area of the main lakes and rivers contained within the borders of each subnational region in my dataset, scaled by total area in the region. Data on agricultural land and soil quality come from the Harmonized World Soil Database (HWSD). *Agricultural Land* is defined as the percentage of land in the subnational region that meets a minimum threshold of suitability for agricultural purposes, whereas *Soil Quality* measures the level of suitability of the agricultural land. The HWSD contains estimates of soil quality for points around the world across seven dimensions: ease of tillage, nutrient availability, nutrient retention, oxygen availability, pH levels, rooting conditions, and salinity. Each dimension is rated on a scale of one to four, with four indicating the lowest soil quality. I take the average across all seven dimensions and reverse the sign to obtain a single soil quality score for each point across the globe. I define *Soil Quality* as the average soil quality for all points contained within each subnational region.

2.3. Institutions data

I compile data on institutional quality for subnational regions using survey data from several different sources. Each of these sources provides separate measurements for different subnational regions within each country. In some cases, the survey sponsor provides aggregated data at the subnational level, and in other cases, I aggregate respondent-level data into subnational measures based on the given location of each survey respondent. From each survey I extract the specific questions that relate to the quality of institutions. (In some cases, the survey sponsor aggregates the questions into sub-indexes, in which case I extract the specific sub-indexes that relate to the quality of institutions.) I exclude from my data survey questions that relate specifically to national-level institutions to ensure that the responses reflect the local situation in the subnational region as much as possible. To make the scale of questions comparable within surveys, I first take the log of responses for questions which have no maximum response (e.g., the number of days to enforce a contract), and then take the negative of responses for questions in which a higher number indicates worse institutions, so that all questions are directionally consistent. I then standardize each question to a mean of zero and a standard deviation of one so that all questions are weighted equally when aggregated. I classify all extracted questions into one of four categories of institutions: property rights protection (including contract enforcement), corruption control, law and order, and regulatory efficiency.

Rows 28 through 34 report statistics for institutions variables derived from each of the individual surveys used in the study. Data for twelve African countries come from the Afrobarometer study prepared jointly by researchers at Michigan State University, the Institute for

[7] The precious metals include gold, silver, and the platinum group (platinum, ruthenium, rhodium, palladium, osmium, and iridium).

[8] The base metals include copper, lead, nickel, zinc, aluminum, bismuth, cobalt, gallium, indium, magnesium, mercury, potassium, titanium, tin, uranium, and zirconium. The alloying elements include principal elements used in producing steel: carbon, manganese, chromium, molybdenum, vanadium, boron, cerium, niobium, and tungsten.

Democracy in South Africa, and the Ghana Centre for Democratic Development. Afrobarometer conducts public attitude surveys on democracy and governance across Africa. Detailed information on the questions used from the Afrobarometer survey is reported in Appendix Table A3. Data for 18 countries in the Americas come from the Latin American Public Opinion Project sponsored by Vanderbilt University (see Appendix Table A4). Data for eight Asian countries are provided by The Asia Foundation (TAF) in cooperation with other organizations that differ by country. Because the individual issues covered in each country in the TAF data are not exactly the same, I normalize the data for each country individually, rather than for the survey as a whole (see Appendix Table A5). Data for 17 European countries come from researchers at the Quality of Government Institute at the University of Gothenburg in Sweden (see Quality of Government Institute, 2010 and Appendix Table A6). Data for 15 Latin American nations come from Latinobarómetro, a non-profit NGO based in Santiago, Chile (see Appendix Table A7). Data for 65 countries worldwide come from the World Bank Enterprise Survey (see Appendix Table A8). Finally, data for 21 countries worldwide come from the subnational Doing Business estimates provided by the World Bank (see Appendix Table A9). Some countries are covered in multiple surveys. Combining all surveys together, I have institutions data for 1343 subnational regions in 86 different countries (see row 35 of Table 2). Because the institutions variables are based on perceptions of survey respondents, they are primarily de facto measures of institutions. The Doing Business variables are an exception; they are derived from expert surveys and are based on current laws and regulations in each region, and thus are primarily de jure measures of institutions.

Rows 36 through 39 of Table 2 report descriptive statistics for the aggregate institutional variables *Property Rights*, *Corruption Control*, *Law and Order*, and *Regulatory Efficiency*. These institutional variables are aggregated by first averaging the (standardized) data for all questions that fall within each of the four categories within each survey, and then aggregating the institutional variables across the different surveys to obtain one measure of each of the four categories of institutions. Appendix Tables A3 through A9 provide detailed classifications of the survey questions included in each category.

The variable *Property Rights* deals with issues related to the effectiveness and fairness of the legal system, protection of property rights, and enforcement of contracts. An example of within-country differences in this variable is in China, where the highest rankings for *Property Rights* are in the coastal commercial centers of Shanghai, Guangdong, and Jiangsu, and the lowest rankings are in the inland provinces of Gansu, Yunnan, and Guizhou. *Corruption Control* measures the corruption of public officials and the prevalence of bribery. As an example, the corruption problems in Russia are well known, but there is wide variation in perceptions of bribery and corruption across Russian regions. Moscow Oblast and the city of Moscow have among the worst *Corruption Control* rankings in the country, but bordering oblasts Tver and Vladimir have the best rankings, a disparity that could be due to the large concentration of bureaucrats in Moscow. *Law and Order* measures law enforcement, the prevalence of crime, and security. For example, in the U.S., the District of Columbia is often viewed as having particularly bad crime rates, and in my data it has the lowest *Law and Order* ranking in the U.S. *Regulatory Efficiency* measures regulatory obstacles to commerce and effectiveness of government services. An example of within-country variation in *Regulatory Efficiency* is in the Philippines. Many regions of the Philippines score well in *Regulatory Efficiency*, but the Autonomous Region in Muslim Mindanao is a negative outlier on this dimension of institutions and is also the poorest region in the country.

2.3.1. Measurement error in institutions data

Whereas geographic variables can be measured quite precisely, an important issue to consider in the construction of institutional variables is measurement error. While some noise is to be expected in survey data, it is important to ensure that the signal-to-noise ratio in the compiled data is not too low. A number of features of the institutions data described above are reassuring in this regard. First, the datasets used in the compilation of the institutions data come from experienced organizations (as detailed above) that have expertise in survey design. Second, each of the datasets is taken from a large sample of respondents, with sample sizes ranging from 17,000 to 40,000 individuals.[9] In total, the institutions data is based on information taken from 172,057 respondents around the world, helping to ensure that the compiled statistics are not unduly influenced by a small number of uninformative responses. Third, although cultural bias is a common concern in cross-country survey data (i.e., respondents in different countries may respond differently to questions because of variation in societal norms), my study uses country fixed effects, so survey responses are only compared within countries. In general, cultural differences within countries are far less pronounced than across countries, thereby mitigating concerns about cultural bias in my study. Fourth, my measures are aggregated from several different sources and over 100 separate survey items, and, as argued by Kaufmann and Kraay (2008), aggregated variables can increase the precision with which broad concepts of governance are measured. Finally, individual datasets used in the compilation of my institutions variables have been previously employed in other studies, including Nunn (2010), Chong et al. (2014), Ma et al. (2010), Klapper et al. (2006), Checchi et al. (2009), and Gennaioli et al. (2013). In summary, although any institutional variable (whether constructed from surveys or expert opinions) carries limitations, there are a number of factors that suggest that the variables used in this paper are as informative as possible, given currently available subnational data. I return to the issue of measurement error in Section 4 below.

3. Geography and subnational income

I empirically investigate the determinants of economic development at the subnational level using the data described in the previous section. I estimate regression equations of the following form

$$\log y_{ci} = \alpha + \mathbf{X}'_{ci}\beta + \delta_c + \varepsilon_{ci}, \quad (1)$$

where y_{ci} is GDP per capita for subnational region i in country c, $\mathbf{X}_{ci}$ is a vector containing one or more proposed determinants of development in subnational region i in country c, and δ_c represents a full set of country fixed effects. In these regressions I adjust the standard errors for clustering at the country level. To examine the impact of geographic factors on economic development I report regression results in separate tables covering general geographic variables, climate variables, and natural resource variables. To facilitate interpretation of the magnitude of the coefficients, I report standardized beta coefficients for all reported regressions.

3.1. General geographic factors

Table 3 reports results of the estimation of Eq. (1) in which $\mathbf{X}_{ci}$ contains one or more general geographic variables. Column 1 of Table 3 shows that *Latitude* is positively correlated with per-capita GDP, as expected, although the coefficient is not significant by itself. In column 2, the coefficient on *Elevation* is negative and significant at the 1% level, indicating that, within countries, subnational regions in lower elevations have higher incomes. The magnitude of the coefficient indicates that a one-standard-deviation increase in elevation (about 1900 ft) is associated with an average decrease of 8.5% of a standard deviation in log GDP per capita, which translates to roughly $658.[10] The negative

[9] The exception regarding sample size is the Doing Business dataset provided by the World Bank, which is based on an unspecified number of expert opinions.

[10] In all cases I evaluate the magnitude of the coefficient based on movements away from sample mean of log GDP per capita.

Table 3
Geographic factors and economic development within countries.

	Dependent variable is log GDP per capita								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
									Rule of Law Interaction
Latitude	0.213							0.199*	−0.194**
	(0.145)							(0.113)	(0.084)
Elevation		−0.085***						−0.036*	−0.029
		(0.023)						(0.021)	(0.021)
Ocean Access			0.089***					0.084***	−0.054***
			(0.026)					(0.026)	(0.016)
Terrain Ruggedness				−0.093***				−0.082***	0.020
				(0.021)				(0.024)	(0.018)
Storm Risk					−0.083**			−0.099**	−0.005
					(0.035)			(0.038)	(0.024)
Earthquake Risk						−0.019		0.014	0.005
						(0.033)		(0.029)	(0.017)
Malaria Risk							0.115		
							(0.098)		
R-squared (within)	0.151	0.026	0.036	0.030	0.018	0.001	0.008	0.106	0.130
Number of countries	101	101	101	101	101	101	99	101	101
Number of obs.	1867	1867	1867	1866	1866	1866	1651	1866	1866

The table reports standardized beta coefficients of regressions of log GDP per capita on variables describing general geographic factors of subnational regions around the world. *Latitude* is the absolute value of the latitude of the capital city of each subnational region; *Elevation* is the log elevation of the capital city (in thousands of feet). *Ocean Access* is a dummy variable equal to one if the subnational region has ocean access. *Terrain Ruggedness* is an index from Nunn and Puga (2012). *Storm Risk* is the log number of hurricanes and tropical storms in the subnational region and *Earthquake Risk* is the log number of fault lines in the subnational region. *Malaria Risk* comes from Kiszewski et al. (2004). Column 9 reports coefficients for an interaction with country-level rule of law (main effects not reported). Also included in the model but not reported are country fixed effects and a constant. Robust standard errors, adjusted for clustering within countries, are reported below coefficients in parentheses. Asterisks denote levels of statistical significance (*** = 1%, ** = 5%, * = 10%).

effect of elevation could be due to direct effects on agricultural productivity or indirect effects of elevation on population density (see Gallup et al., 1999; Masters and McMillan, 2001). Column 3 of Table 3 shows that subnational regions with ocean access have higher incomes, with the coefficient being significant at the 1% level. The magnitude of the coefficient indicates that subnational regions with ocean access have higher incomes than landlocked regions in the same country by $1526, on average. Column 4 shows that terrain ruggedness is associated with lower incomes within countries. The coefficient on *Terrain Ruggedness* is significant at the 1% level and indicates that a one-standard-deviation increase in ruggedness is associated with an average decrease of $723 in per-capita GDP.

Columns 5, 6, and 7 of Table 3 report the results for the risk associated with tropical storms, earthquakes, and malaria, respectively. Of the three types of risk, only *Storm Risk* has a significant coefficient. A one-standard-deviation increase in storm risk is associated with a decrease of $644 in per-capita GDP. In column 8 of Table 3, all of these geographic factors except for malaria risk (due to the smaller number of observations) are included in the regression simultaneously, and *Latitude*, *Elevation*, *Ocean Access*, *Terrain Ruggedness*, and *Storm Risk* all show significant explanatory power for incomes within countries. Of the significant variables *Latitude* has the largest standardized coefficient, indicating that, independent of the scale of units, it has the strongest within-country effect among the geographic factors.[11] The r-squared of the regression indicates that about 10.6% of the subnational variation in per-capita GDP is explained by the factors in this regression.

In the final column of Table 3 I interact the geographic variables with country-level institutions. Here institutions are measured by the rule of law index for 2005 from the World Bank's World Governance Indicators (WGI). Column 9 shows a significant negative interaction for *Latitude* and *Ocean Access*, indicating that these geographic factors have less of an impact on income in countries with stronger institutions. (For brevity, main effects are not reported in these regressions.) Although the other interaction terms do not have significant coefficients, the significant negative interactions are consistent with the view that geographic factors become less important for development when a country has effective institutions. Such could be the case, for example, if strong institutions promote within-country trade (see Rodrik et al., 2004), allowing the benefits of natural geographic endowments to be shared among subnational regions of the same country.

When assessing the economic magnitude of the variables it is important to keep in mind that the regressions include country fixed effects, so the size of the coefficients refers only to within-country variation attributable to the various factors. For example, even though malaria risk is not a significant variable in Table 3, malaria risk could still be a major reason why, for example, the per-capita GDP of South Africa (which has relatively low malaria risk) is five times the per-capita GDP of Senegal (which has relatively high malaria risk).

3.2. Climatic factors

Table 4 reports results of the estimation of Eq. (1) in which $\mathbf{X}_{ci}$ contains one or more climatic variables. In column 1, the first variable tested is the percent of the subnational region that has a tropical climate. The coefficient on *Percent Tropical* is negative and significant at the 1% level, indicating that subnational regions that are more tropical have lower incomes than less-tropical regions within the same country. The magnitude of the coefficient indicates that a non-tropical region would have higher income than a fully tropical region in the same country by $2473, on average. A potential example of this effect is in Peru, which has, among its 24 regions, 4 regions that are non-tropical (Ica, Lima, Moquegua, and Tacna) and 3 regions that are fully tropical (Loreto, Madre de Dios, and Ucayali). The per-capita GDP of the non-tropical regions averages $11,400, whereas the per-capita GDP of the tropical regions averages $5225. Although many factors may contribute to this disparity in wealth, the disadvantage of a tropical climate may be a key reason for the relative poverty of the tropical regions of Peru.

The remaining eight variables tested in Table 4, in columns 2 through 9, measure different aspects of climate that provide additional detail as to the specific climatic features that may affect development. Of the eight variables, *Temperature Range* and *Sunshine* have a statistically significant relation with per-capita GDP. These variables may be

[11] As a robustness check, I alternatively measure latitude as the geographic centroid of each region. *Latitude* is not statistically significant in this case.

Table 4
Climate and economic development within countries.

	Dependent variable is log GDP per capita										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
											Rule of Law interaction
Percent Tropical	−0.156***									−0.179**	−0.087
	(0.054)									(0.069)	(0.072)
Precipitation		0.015								0.026	−0.021
		(0.038)								(0.049)	(0.038)
Rain Days			0.023							−0.082	−0.011
			(0.052)							(0.067)	(0.041)
Temperature				−0.037						0.058	0.204*
				(0.086)						(0.125)	(0.114)
Temperature Range					−0.124***					−0.138***	−0.051
					(0.033)					(0.039)	(0.034)
Humidity						0.074				−0.012	−0.123**
						(0.049)				(0.076)	(0.061)
Sunshine							−0.102*			−0.104**	−0.035
							(0.057)			(0.052)	(0.045)
Frost Days								0.031		0.095	0.062
								(0.065)		(0.094)	(0.079)
Wind Speed									0.062	0.004	−0.030
									(0.038)	(0.036)	(0.025)
R-squared (within)	0.020	0.001	0.007	0.001	0.032	0.010	0.012	0.001	0.008	0.068	0.093
Number of countries	101	101	101	101	101	101	101	101	101	101	101
Number of obs.	1866	1837	1837	1837	1837	1837	1837	1837	1837	1837	1837

The table reports standardized beta coefficients of regressions of log GDP per capita on climatic variables for subnational regions around the world. *Percent Tropical* is the percent of the subnational region covered by a tropical or subtropical biome. The other climatic variables are monthly averages over the period 1961–1990, averaged over all estimated points within each subnational region. *Precipitation* is in mm, *Rain Days* is the number of days with rainfall >0.1 mm, *Temperature* is in average degrees Celsius, *Temperature Range* is the average diurnal temperature range, *Humidity* is the average percent, *Sunshine* is the average percent of day length, *Frost Days* is the number of days with ground frost, and *Wind Speed* is in m/s. Column 11 reports coefficients for an interaction with country-level rule of law (main effects not reported). Also included in the model but not reported are country fixed effects and a constant. Robust standard errors, adjusted for clustering within countries, are reported below coefficients in parentheses. Asterisks denote levels of statistical significance (*** = 1%, ** = 5%, * = 10%).

correlated with different aspects of having a tropical climate, but the results in column 10 show that *Temperature Range* and *Sunshine* continue to be significant when *Percent Tropical* and all other climate variables are included in the regression, indicating additional explanatory power of these variables above and beyond the tropical climate classification. In column 10, the coefficient on *Temperature Range* indicates that a one-standard-deviation increase (2.3 °C) in the diurnal temperature range is associated with a decrease in average per-capita GDP of $1103. A one-standard deviation increase (13.1 percentage points) in the average percent of day length with sunshine is associated with a decrease in income of $814. Column 10 of Table 4 shows that all of the variables together explain about 6.8% of the variation in subnational incomes.

Column 11 of Table 4 repeats the regression of column 10 but includes interaction terms of country-level rule of law with each of the climate variables. Most of the interaction effects are not significant, with the exception of those for *Temperature* and *Humidity*, indicating that, for the most part, the impact of climate on income is not moderated a great deal by the quality of a country's institutions.

3.2.1. Natural resources

Table 5 reports coefficient estimates from regressions in which measures of natural resources are the explanatory variables. I test the effect of three precious resources (oil, diamonds, precious metals), three non-precious resources (base metals, iron, alloyants) and three agricultural resources (water, agricultural land, soil quality). Of these variables, Table 5 shows that only *Diamonds* and *Iron* have a statistically significant positive relation with per-capita GDP both in regressions by themselves and in column 10, when all variables are included simultaneously. Despite the statistical significance, the economic significance of *Diamonds* and *Iron* is relatively small, with one-standard-deviation increases being associated with higher per-capita incomes of $67 and $82, respectively. One possible reason for the relatively small effects of natural resources is that the rents from natural resources may accrue more to national governments than to the regions that physically hold the resources.

Perhaps surprisingly, the coefficient on *Oil and Gas* is negative in column 1 and insignificant in column 10, although later robustness checks will show that *Oil and Gas* has a significant positive effect when other geographic factors are controlled for (see Section 5). Ecuador is an example of a country in which natural resources—and especially oil—account for much of the intra-country income differences. The three richest provinces in the country (Orellana, Pastaza, and Sucumbíos) all come from the petroleum-rich Amazonía region and have the highest oil and gas production in the country. The *Oil and Gas* variable alone explains about three-fourths of the within-country variability in per-capita GDP in Ecuador.

Perhaps reflecting evidence of a natural resource curse, *Agricultural Land* and *Soil Quality* are negatively related to income, both with statistical significance. In column 10, *Precious Metals* and *Base Metals* also have negative coefficients, though without statistical significance. Prior research on the natural resource curse argues that the curse is driven by countries with poor institutions and that abundant resources are more of a blessing in countries with strong institutions (e.g., Mehlum et al., 2006; Robinson et al., 2006).[12] Column 11 tests this hypothesis at the subnational level by including interaction terms of the natural resource variables with country-level rule of law. If the resource curse is moderated by strong institutions, then the interaction effects should be of the opposite sign, particularly on *Agricultural Land* and *Soil Quality*. Column 11 shows that the interaction effect of *Agricultural Land* and *Soil Quality* is positive, but the coefficients are not statistically significant, so it is difficult to draw strong conclusions from the results. The significant coefficient on *Base Metals* in column 11 also does not offer support for the idea that institutions moderate the resource curse, as the sign is the same as the sign on the main effect in column 10. These results do not directly contradict prior cross-country evidence, but they do suggest that strong institutions do not moderate the resource curse at the subnational level.

[12] For a more comprehensive list of studies related to this question, see the surveys by van der Ploeg (2011) and Frankel (2010).

Table 5
Natural resources and economic development within countries.

	Dependent variable is log GDP per capita										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
											Rule of Law interaction
Oil and Gas	−0.003***									0.002	−2.874***
	(0.001)									(0.002)	(0.486)
Diamonds		0.010***								0.009***	0.005
		(0.001)								(0.001)	(0.006)
Precious Metals			0.003***							−0.013	−0.079
			(0.001)							(0.013)	(0.084)
Base Metals				0.005						−0.005	−0.011***
				(0.007)						(0.009)	(0.004)
Iron					0.012***					0.011***	−0.006
					(0.003)					(0.002)	(0.009)
Alloyants						0.022				0.029	−0.009
						(0.017)				(0.020)	(0.020)
Water							0.017			0.007	−0.002
							(0.012)			(0.011)	(0.010)
Agricultural Land								−0.036***		−0.028**	0.011
								(0.013)		(0.012)	(0.008)
Soil Quality									−0.039**	−0.028*	0.004
									(0.017)	(0.016)	(0.014)
R-squared (within)	0.000	0.001	0.000	0.000	0.001	0.003	0.008	0.008	0.007	0.015	0.055
Number of countries	101	101	101	101	101	101	100	100	100	100	100
Number of obs.	1866	1866	1866	1866	1866	1866	1847	1847	1847	1847	1847

The table reports standardized beta coefficients of regressions of log GDP per capita on natural resource variables for subnational regions around the world. Mineral resources are the number of identifiable sites producing the resource in the subnational region. *Precious Metals* include gold, silver, and the platinum group; *Base Metals* include 16 valuable non-ferrous metals; alloyants include nine principal elements used in producing steel. *Water* is the area of main lakes and rivers divided by total area. All non-precious resources are scaled by land area; precious resources (oil, diamonds, precious metals) are not. *Agricultural Land* is the percentage of land suitable for agriculture, and *Soil Quality* is the average quality of the soil across seven dimensions. Column 11 reports coefficients for an interaction with country-level rule of law (main effects not reported). Also included in the model but not reported are country fixed effects and a constant. Robust standard errors, adjusted for clustering within countries, are reported below coefficients in parentheses. Asterisks denote levels of statistical significance (*** = 1%, ** = 5%, * = 10%).

4. Institutions and subnational income

In this section I turn to an analysis of the relation between institutional quality and economic development at the subnational level. A number of influential studies have demonstrated the positive effect of institutions on incomes at the national level, but it is not as clear whether the same effects operate within countries. I first test the relation between institutions and income with data aggregated to the level of each survey instrument, and then perform tests looking at all survey data aggregated into four areas of institutional quality. Lacking an appropriate instrument for institutions at the subnational level, I interpret the following regressions as demonstrating correlations for which a causal effect of institutions is but one potential interpretation.

4.1. Institutions in different world areas

Table 6 reports coefficient estimates of Eq. (1) in which $\mathbf{X}_{ci}$ contains variables measuring institutional quality in different areas of the world, with each variable corresponding to the aggregated data from one of the seven surveys from which institutional indicators are taken. Column 1 of Table 6 shows results for 180 subnational regions in 12 African countries (from the Afrobarometer survey). The coefficient on institutions shows a negative relation between institutions and per-capita income, but the coefficient is not statistically significant. Columns 2 through 5 report results for the Americas (the LAPOP survey), Asia (TAF surveys), Europe (the QOG survey), and Latin America (the Latinobarómetro survey). The coefficient on institutions is negative and significant (at the 5% level) for the Americas (column 2) and Latin America (column 5). It is negative but insignificant for Asia (column 3) and positive but insignificant for Europe (column 4). Columns 6 and 7 report results for the two worldwide surveys (WBES and DB), and the coefficient is negative and significant (at the 5% level) for WBES and positive and insignificant for DB. Finally, column 8 shows results with all surveys combined, covering 1343 subnational regions in 86 countries, and the coefficient is negative and insignificant. As a whole, Table 6 provides no statistically significant evidence of a positive relation between institutional quality and GDP per capita. To the contrary, the relation is usually negative, and sometimes negative with statistical significance.

The negative relation—or at least the lack of a positive relation—between subnational institutions and income is puzzling. The evidence for a positive relation between institutions and income at the country level is clear (e.g., Acemoglu et al., 2001; Rodrik et al., 2004), and it seems natural that this relation would carry over to the regional level, given the apparent benefits of ensuring property rights, controlling corruption, and so on. Understanding the relation between institutions and development within countries is critical to understanding how policy should be implemented to best serve underdeveloped nations. So in sections below I consider explanations for this result, ranging from a host of potential measurement issues to the possibility that subnational institutions are ineffectual because they are dominated by national institutions.

4.2. Subcategories of institutions

Table 7 reports estimates of the effect of institutions on per-capita GDP when the underlying data are aggregated across all geographic areas in different categories of institutions. Columns 1 through 4 test the effect of *Property Rights*, *Corruption Control*, *Law and Order*, and *Regulatory Efficiency* respectively. For all subcategories of institutions the coefficient is negative, and in the case of *Law and Order* the negative coefficient is significant at the 5% level. Column 5 reports results for all subcategories combined, and the coefficient is also negative.[13] Table 7 confirms the result in Table 6 of a negative relation between institutional quality and economic development. The results suggest that the positive relation between institutions and income that other studies have documented at the country level does not carry over to the regional level, at least as a general result.

[13] The combined variable in Table 7 incorporates the same underlying data as the combined variable in Table 6, but the two differ because of the method of aggregation.

Table 6
Institutional quality and economic development within countries.

	Dependent variable is log GDP per capita							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Institutions in Africa	−0.049							
	(0.053)							
Institutions in the Americas		−0.086**						
		(0.036)						
Institutions in Asia			−0.026					
			(0.092)					
Institutions in Europe				0.028				
				(0.088)				
Institutions in Latin America					−0.115**			
					(0.053)			
Institutions worldwide (WBES)						−0.077**		
						(0.038)		
Institutions worldwide (DB)							0.035	
							(0.155)	
Institutions combined								−0.056
								(0.038)
R-squared (within)	0.009	0.021	0.003	0.002	0.030	0.026	0.002	0.008
Number of countries	12	18	8	17	15	65	21	86
Number of obs.	180	362	147	308	248	619	230	1343

The table reports standardized beta coefficients of regressions of log GDP per capita on institutional quality variables for subnational regions around the world. Separate regressions for each underlying data source are reported. Data for Africa comes from Afrobarometer, data for the Americas comes from LAPOP, data for Asia comes from The Asia Foundation, data from Europe comes from the Quality of Government Institute, and data from Latin America comes from Latinobarómetro. The two worldwide sources include the World Bank Enterprise Survey (WBES) and Doing Business (DB). Also included in the model but not reported are country fixed effects and a constant. Robust standard errors, adjusted for clustering within countries, are reported below coefficients in parentheses. Asterisks denote levels of statistical significance (*** = 1%, ** = 5%, * = 10%).

As noted above, institutions are potentially endogenous to development, so employing an instrument for institutional quality at the subnational level would help clarify the relation between institutions and income within countries. However, instruments for institutions that have been employed in the literature at the country level (settler mortality, historic population density, legal origin) either do not vary at the subnational level or are unavailable at the subnational level.

4.3. Measurement issues

Why do the data not show a positive correlation between institutions and development, given that strong institutions would seem to be beneficial to economic development and have been shown to have a positive effect on institutions in country-level studies? In this section I consider whether these results can be explained by issues of data quality or measurement.

Table 7
Institutional quality and economic development within countries, subcategories.

	(1)	(2)	(3)	(4)	(5)
	Dependent variable is log GDP per capita				
Property Rights	−0.019				
	(0.029)				
Corruption Control		−0.012			
		(0.014)			
Law and Order			−0.075**		
			(0.029)		
Regulatory Efficiency				−0.032	
				(0.021)	
Institutions combined					−0.041
					(0.032)
R-squared (within)	0.001	0.001	0.017	0.005	0.005
Number of Countries	75	81	82	72	86
Number of Obs.	1130	1232	1251	1041	1343

The table reports standardized beta coefficients of regressions of log GDP per capita on institutional quality variables for subnational regions around the world. Variables are aggregated across all sources of underlying data and divided into four subcategories of institutions as noted. Also included in the model but not reported are country fixed effects and a constant. Robust standard errors, adjusted for clustering within countries, are reported below coefficients in parentheses. Asterisks denote levels of statistical significance (*** = 1%, ** = 5%, * = 10%).

One reason why the within-country regressions may not demonstrate a positive effect of institutions on income is if there is not enough variation in institutions within countries to uncover the true effect of institutional quality on development. At the extreme, if measures of institutions were identical for all subnational regions within each country, then the effect of institutions would be subsumed in the regressions by the country dummies and would have no explanatory power for income per capita. Of course, this is not the case, and in fact country dummies capture only about half of the variation in the institutional measures. For example, in a regression of the property rights measure on a set of country dummies, the r-squared is 0.55. The corresponding r-squareds are 0.51, 0.63, and 0.38 for the measures of corruption control, law and order, and regulatory efficiency, respectively. There appears to be sufficient variation in institutional quality within countries to test for subnational effects of institutions on development.

A second possible explanation for not finding a positive effect of institutions is that the institutional variables could be measured with too much noise to be informative. Section 2.3.1 above presents a number of reasons why concern about measurement error should be mitigated in the institutions data. To further address the issue of measurement error, I assess the affect on the results in Tables 6 and 7 from reducing the level of noise in the data. Responses to survey questions are noisier when the level of variation in responses is greater. That is, if the respondents to a particular question in a particular subnational region give similar responses, then we can have greater confidence in the average response for that region, but if respondents answer with high variability, then the average response for that region is less informative. The institutions data is comprised of 23,439 region/question pairs (this does not include indexes aggregated by the data provider). For each region/question pair, I calculate a coefficient of variation (CV; the standard error of the mean divided by the mean) as an indicator of the level of noise in the mean response for each question in each region. I then eliminate all region/question pairs for which the coefficient of variation is greater than one. This eliminates roughly the top 10% noisiest (2424) region/question pairs from the data. The results in Tables 6 and 7 are qualitatively unchanged when re-estimated without these noisiest responses. The primary difference is that the significant negative coefficients in Table 6 are significant at the 5% or 1%

Table 8
Institutional quality and economic development within and between countries.

	Dependent variable is log GDP per capita				
	(1)	(2)	(3)	(4)	(5)
	Institutions category				
	Property Rights	*Corruption Control*	*Law and Order*	*Regulatory Efficiency*	Institutions combined
Panel A: country level (between regression), all institutional features					
Institutions	0.093	0.476**	0.154	-0.073	0.283**
	(0.129)	(0.130)	(0.135)	(0.153)	(0.125)
R-squared (between)	0.007	0.144	0.016	0.003	0.057
Number of countries	75	81	82	72	86
Number of obs.	1130	1232	1251	1041	1343
Panel B: country level (between regression), country-relevant institutional features only					
Institutions	0.260**	0.478***	0.381***	0.400***	0.455***
	(0.116)	(0.105)	(0.135)	(0.116)	(0.113)
R-squared (between)	0.067	0.210	0.092	0.193	0.164
Number of countries	72	80	80	71	86
Number of obs.	958	1190	1128	923	1274
Panel C: subnational level (within regression), country-relevant institutional features only					
Institutions	−0.025	−0.031	−0.046**	−0.026**	−0.041
	(0.024)	(0.022)	(0.021)	(0.012)	(0.032)
R-squared (within)	0.002	0.003	0.008	0.003	0.005
Number of countries	72	80	80	71	86
Number of obs.	958	1190	1128	923	1343

The table reports standardized beta coefficients of regressions of log GDP per capita on institutional quality variables for subnational regions around the world. Panels A and B report between-country regressions and Panel C reports within-country regressions. In Panels B and C, only "country-relevant" institutional features (i.e., those which are positively correlated with income at the country level) are included in the measures of institutions. Also included in the model but not reported are country fixed effects and a constant. Robust standard errors, adjusted for clustering within countries, are reported below coefficients in parentheses. Asterisks denote levels of statistical significance (*** = 1%, ** = 5%, * = 10%).

level instead of the 10% level.[14] Although concern over mismeasurement cannot be entirely alleviated, these results, together with the factors discussed in Section 2.3.1 above, suggest that measurement error is not primarily responsible for the lack of a finding of a positive effect of institutions within countries.

A third measurement issue concerns the possibility of selection bias in the regions surveyed within countries. In some cases, surveys did not obtain data from all regions in the country, raising the possibility that richer or poorer regions were excluded from the sample. To address this question, I test the data from each country in each survey to determine if sampled regions had significantly higher or lower income than non-sampled regions. I find that for five of the surveys, there is little evidence of selection bias, as at most one country in each of the surveys has a significant difference in income between sampled and non-sampled regions (in many cases all regions are sampled). But in two of the surveys, WBES and Doing Business, roughly 28% of the countries in each survey show evidence of having targeted richer regions of the country. This is likely due to the fact that WBES and Doing Business focus on the experience of firms rather than individuals, so they naturally target more urban (richer) regions. To test whether this bias is affecting the results, I repeat the analysis of Tables 6 and 7, but exclude data from countries that display selection bias in each survey. The results (not reported) are similar to the baseline results, with coefficients that are somewhat less negative, but not different enough to alter conclusions drawn from the baseline results.[15]

A fourth reason for the difference between national and subnational results could be that my institutional measures capture different aspects of institutional quality than the measures used in existing country-level studies. In one sense, this is true by design, because in extracting survey questions to create the subnational measures I exclude any questions that deal specifically with national-level institutions. But if my subnational measures of institutions measure similar aspects of institutions as do existing country-level measures, then my measures should be positively correlated with income when tested at the national level. I study the effect of my measures of institutions on country-level incomes in Table 8. Panel A of Table 8 reports *between-country* regressions of income on my institutional measures, which allow me to test the effect of my institutional variables at the country level. Panel A shows that the between-country correlation between institutions and income is positive for four out of five of the institutional measures and statistically significant for two of the measures. Thus, consistent with existing country-level studies, the effect of my institutional measures on income at the country level is positive overall, although not as strong as has been shown in other country-level studies of institutions on development (e.g., Acemoglu et al., 2001; Rodrik et al., 2004).[16]

To further investigate the difference between country-level results and within-country results, in Panel B of Table 8 I perform another between-country regression, but in this case I intentionally create institutional measures that have a strong positive correlation with income at the country level. To do this, I extract from the data only those survey questions that have a positive correlation with income at the country level, excluding questions that have a negative correlation with income at the country level. Panel B shows that in the between-country regression using only these "country-relevant" institutional features, all measures of institutions are positive and significant at the 1% or 5% level. This is as expected since the measures are intentionally designed to have a positive correlation with income at the country level. Then in Panel C of Table 8 I take the same country-relevant institutional features back to the subnational level. That is, I use the

[14] Results are also qualitatively unchanged with more stringent CV cutoffs of 0.75 (which eliminates roughly the top 13% noisiest responses), 0.5 (which eliminates roughly the top 18% noisiest responses) and 0.25 (which eliminates roughly the top 29% noisiest responses).

[15] A related issue concerns the representativeness of the samples interviewed in each region. If the survey-conducting organizations select samples to be representative of the national population, or if samples are stratified on dimensions other than administrative regions, then the regional samples may not be representative of the regional populations. However, for most of the surveys, the samples are stratified by subnational region, ensuring that the samples are representative at the subnational level. The exceptions to this rule are some of the countries covered by Latinobarómetro and The Asia Foundation, but I find that the results (not reported) are very similar if the potentially non-representative data is excluded from the regressions.

[16] The between-country results carry the caveat that methods of data compilation are not standardized across countries.

same survey questions used in Panel B and apply them to my standard within-country regression. Panel C shows that the positive effect of institutions between countries does not carry over to the subnational level. The coefficient on institutions is negative for all four measures and significant (at the 10% level) for two of the measures. These results suggest that my subnational institutional variables are not inconsistent with existing country-level institutional measures, but rather show a reversal in the effect of institutions at the subnational level relative to the country level.

In summary, the results of this section show that the lack of a positive relation between institutions and income at the subnational level is not likely to be explained by data quality or measurement issues.

4.4. National institutions vs. subnational institutions

Another possible explanation for not finding a general positive effect of subnational institutions is that the impact of subnational institutions could be dominated by national institutions. If subnational laws and procedures are of second-order importance relative to national laws and procedures then subnational regions may be constrained in their ability to improve economic development through institutional improvement, because whatever institutions they implement are overruled by national institutions.

To test this explanation, I first run regressions to compare the explanatory power of national institutions with subnational institutions. I match each of my subnational measures of institutions with a corresponding national measure of institutions from the WGI (see rows 40 through 44 of Table 2 for descriptive statistics of the WGI variables). My *Property Rights* variable is matched with the WGI's *Rule of Law*, my *Corruption Control* is matched with the WGI's *Corruption Control*, my *Law and Order* is matched with WGI's *Stability/Lack of Violence*, and my *Regulatory Efficiency* is matched with the WGI's *Regulatory Quality*. I also match my combined measure of institutions with a combined national measure that averages the four subcategories of institutions. I run regressions of per-capita GDP on the country-level measure by itself and then on the country-level measure and its paired subnational measure (country fixed effects are excluded in these regressions).

Table 9 reports the results of these regressions. Consistent with existing literature, the country-level measures of institutions have strong positive correlations with income that are always significant at the 1% level. When added to the regression, the subnational measures of institutions are negative in all five cases, and statistically significant in three cases. But the key point in Table 9 is that the explanatory power of national institutions appears to dominate subnational institutions. For example, in column 1, the r-squared of the national-level *Rule of Law* alone is 31.9%, and the r-squared increases to only 33.0% when the subnational *Property Rights* is added to the regression (column 2). In addition, the magnitude of the standardized coefficient on *Rule of Law* in column 2 is about seven times the magnitude of the standardized coefficient on *Property Rights*, indicating that the economic significance of the country-level measure is much higher. A similar pattern is repeated for the other four pairs of variables in columns 3 through 10. In short, even if subnational institutions are negatively correlated with income, the impact of subnational institutions appears to be much smaller than that of national institutions.

4.4.1. Autonomy and subnational institutions

To further examine whether national institutions dominate subnational institutions, I study the difference in the relation between subnational institutions and income in regions with more autonomy compared to regions with less autonomy. I identify subnational regions that are designated as autonomous, meaning that they are to some degree self-governing and free from the influence of national institutions. One example of an autonomous region is the Republic of Tatarstan in the Russian Federation, where a majority of the population are ethnic Tatars and, like other autonomous Russian regions, the government has its own constitution. Another example is the island of Sicily, which, along with other autonomous regions of Italy, was granted home rule because of cultural and linguistic differences. Other examples include the Autonomous Region in Muslim Mindanao, which is a region in the Philippines that has its own government, and the Xinjiang Uyghur autonomous region in China, which, like other autonomous regions in China, enjoys more legislative rights than other Chinese provinces. If the lack of a positive relation between institutions and income at the subnational level is due to the dominance of national institutions, then the positive relation between institutions and income should reappear in the subset of regions that are autonomous.

I employ several variables to measure the degree to which subnational regions have autonomy. The first measure is a dummy variable for whether the region is on the *List of Autonomous Areas by Country* maintained on Wikipedia. This measure of autonomy, though quite simple, is advantageous for my study because it has within-country variation. The degree of autonomy of subnational regions on this list varies, but given that the degree of autonomy is difficult to quantify, I set the

Table 9
National institutions vs. regional institutions.

	Dependent variable is log GDP per capita									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Country-level measure									
	Rule of Law		*Corruption Control*		*Stability/Lack of Violence*		*Regulatory Quality*		Institutions combined	
	Subnational measure									
	Property Rights		*Corruption Control*		*Law and Order*		*Regulatory Efficiency*		Institutions combined	
Country-level institutions	0.692***	0.743***	0.780***	0.844***	0.474***	0.476***	0.691***	0.693***	0.793***	0.853***
	(0.109)	(0.125)	(0.061)	(0.064)	(0.093)	(0.096)	(0.100)	(0.098)	(0.066)	(0.077)
Subnational institutions		−0.104		−0.143***		−0.012		−0.132*		−0.141**
		(0.074)		(0.053)		(0.126)		(0.069)		(0.070)
R-squared	0.319	0.330	0.543	0.559	0.188	0.189	0.341	0.366	0.526	0.543
Number of countries	75	75	81	81	80	80	71	71	84	84
Number of obs.	1130	1130	1232	1232	1240	1240	1034	1034	1332	1332

The table reports standardized beta coefficients of regressions of log GDP per capita on institutional quality variables for subnational regions around the world. Subnational institutional measures are aggregated across all sources of underlying data and divided into four subcategories of institutions as noted. Country-level measures of institutions come from the World Bank's World Governance Indicators. Also included in the model but not reported is a constant term. Robust standard errors, adjusted for clustering within countries, are reported below coefficients in parentheses. Asterisks denote levels of statistical significance (*** = 1%, ** = 5%, * = 10%).

Table 10
Institutional quality, economic development, and subnational autonomy.

	Dependent variable is log GDP per capita							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Autonomy measure							
	Subnational region is autonomous	Country has autonomous regions	State elections held	Municipal elections held	State authority granted	Senators locally elected	Composite autonomy measure	(Autonomous regions only)
Institutions × Autonomy	0.050***	0.057*	0.071	0.098**	0.042	0.004	0.109**	
	(0.016)	(0.034)	(0.043)	(0.045)	(0.046)	(0.054)	(0.044)	
Institutions	−0.043	−0.047	−0.062	−0.055	−0.028	−0.067	−0.050	0.259**
	(0.030)	(0.033)	(0.043)	(0.042)	(0.040)	(0.043)	(0.045)	(0.094)
Autonomy	−0.020							
	(0.022)							
R-squared (within)	0.017	0.015	0.018	0.025	0.008	0.014	0.047	0.240
Number of countries	86	82	71	60	39	42	17	20
Number of obs.	1343	1322	1164	1015	663	780	377	79

The table reports standardized beta coefficients of regressions of log GDP per capita on institutional quality variables and measures of subnational autonomy. The autonomy variables measure whether the subnational region is autonomous (column 1), the presence of autonomous regions (column 2), whether state governments are locally elected (column 3), whether municipal governments are locally elected (column 4), whether state governments have authority over taxation, spending, or legislation (column 5), whether senators are locally elected (column 6), and the sum of all scores for the variables in columns 2 through 6 (column 7). Column 8 includes only autonomous regions as defined in column 1. Also included in the model but not reported are country fixed effects and a constant. Robust standard errors, adjusted for clustering within countries, are reported below coefficients in parentheses. Asterisks denote levels of statistical significance (*** = 1%, ** = 5%, * = 10%).

variable equal to one for all regions on the list.[17] In total, 115 of the 1867 subnational regions in the sample are identified as autonomous regions. My other measures of autonomy are proxies for federalism taken from the 2012 version of the Database of Political Institutions (see Beck et al., 2001), averaged over the period 1980 to 2005. These measures are available only at the country level, but they provide an opportunity to check the robustness of the results with alternative definitions of autonomy. These variables measure the presence of autonomous regions,[18] whether state governments are locally elected, whether municipal governments are locally elected, whether state governments have authority over taxation, spending, or legislation, and whether senators are locally elected. I also create a composite measure of federalism that sums these five variables (see rows 45 to 51 of Table 2 for descriptive statistics of autonomy measures).

To test for the impact of autonomy on the relation between institutions and income, I estimate regressions of the following form

$$\log y_{ci} = \alpha + \beta R_{ci} + \varphi R_{ci} * A_{ci} + \mu A_{ci} + \delta_c + \varepsilon_{ci}, \quad (2)$$

where y_{ci} is GDP per capita for subnational region i in country c, R_{ci} is the measure of combined institutions for subnational region i in country c, A_{ci} is one of the measures of autonomy, and δ_c represents a full set of country fixed effects. When the proxy for autonomy is from the Database of Political Institutions the region-level variable A_{ci} is replaced by a country-level variable A_c and the main effect of autonomy is excluded from the regression (it is taken out by the country fixed effects). The primary coefficient of interest in this regression is φ, which, as the coefficient on the interaction between institutions and autonomy, represents the differential effect of institutions on income in autonomous regions.

The results of the estimation of Eq. (2) are reported in columns 1 through 7 of Table 10. Each column reports results for one of the seven measures of autonomy. The results show that the coefficient on the interaction between institutions and autonomy (φ) is positive for all seven autonomy measures. The coefficients are statistically significant in four of the seven columns. In column 1, using the subnational measure of autonomy, the coefficient on the main effect of autonomy is insignificant, indicating that autonomous regions are not significantly richer or poorer than other regions, on average, they are just more responsive (in a positive manner) to the influence of institutions on income. The net effect of institutions, summing the coefficients on the interaction terms and the main effects, in positive in all columns except column 6, although in some cases the net positive effect is small. To further assess the net effect of institutions, in column 8, I restrict the sample to only autonomous regions, and show that the effect of institutions is significantly positive in this subsample, with an r-squared of 24%.

Table 10 shows that stronger institutions are associated with higher incomes at the subnational level, but only in a subset of regions that has greater power to determine their own governance. The results are consistent with the hypothesis that the dominance of national institutions renders local institutions less effective when regional governments lack sufficient autonomy. It should be noted, however, that these results fall short of establishing that regional autonomy *causes* a stronger relationship between institutions and development within countries. It is conceivable that some unobserved factor makes certain regions both more likely to be autonomous and to have more-effective subnational institutions. Based solely on these findings, it would be premature, for example, to pursue a policy of increasing the effectiveness of subnational institutions by granting greater autonomy to subnational governments. What the results in Table 10 do establish is that there is a set of regions in which strong subnational institutions are significantly correlated with higher incomes, so the impact of institutions on development may depend less on whether the institutions are national or local and more upon what level of government actually has the power to govern.

5. Robustness checks

Table 11 presents the results of additional regressions to assess the robustness of the main results in the paper. For each of the robustness checks two regressions are reported: odd-numbered columns include the full sample of subnational regions and even-numbered columns include the subsample of regions for which institutions data are available.

[17] Subnational regions are included on the list if only part of the region is autonomous. For example, the province of Guizhou in China is included because three of its prefectures are autonomous, and the state of Meghalaya in India is included because of its autonomous district councils.

[18] In some cases the Database of Political Institutions defines autonomy differently from Wikipedia. For example, it does not include autonomous federal districts as autonomous regions.

Table 11
Robustness checks.

	Dependent variable is log GDP per capita											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	All variables combined		Exclude largest cities		Exclude UNDP countries		Control for ethnicity		Control for language		Control for education	
Ocean Access	0.055***	0.035*	0.058***	0.038**	0.062**	0.062*	0.048**	0.036*	0.054**	0.031*	0.046**	0.029*
	(0.021)	(0.018)	(0.021)	(0.017)	(0.025)	(0.015)	(0.019)	(0.018)	(0.022)	(0.018)	(0.018)	(0.015)
Terrain Ruggedness	−0.092***	−0.093***	−0.072***	−0.069***	−0.102***	−0.099***	−0.083***	−0.085***	−0.074***	−0.080***	−0.081***	−0.083***
	(0.019)	(0.023)	(0.021)	(0.026)	(0.022)	(0.027)	(0.024)	(0.026)	(0.018)	(0.022)	(0.020)	(0.024)
Storm Risk	−0.094***	−0.126***	−0.081***	−0.105***	−0.096***	−0.131***	−0.079***	−0.117***	−0.098***	−0.125***	−0.064**	−0.080**
	(0.029)	(0.024)	(0.031)	(0.025)	(0.031)	(0.026)	(0.019)	(0.021)	(0.033)	(0.029)	(0.029)	(0.031)
Percent Tropical	−0.115**	−0.097	−0.112**	−0.096	−0.128**	−0.110	−0.091	−0.080	−0.100	−0.072	−0.039	−0.021
	(0.057)	(0.073)	(0.055)	(0.071)	(0.063)	(0.081)	(0.064)	(0.076)	(0.062)	(0.085)	(0.040)	(0.051)
Temperature Range	−0.097***	−0.109***	−0.087***	−0.098***	−0.106***	−0.115***	−0.070***	−0.090***	−0.085***	−0.096***	−0.029	−0.019
	(0.026)	(0.029)	(0.027)	(0.031)	(0.031)	(0.034)	(0.020)	(0.026)	(0.026)	(0.028)	(0.019)	(0.024)
Oil and Gas	0.011***	0.013***	0.010**	0.011***	0.012***	0.014***	0.008***	0.011***	0.011***	0.012***	0.009***	0.010***
	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)
Diamonds	0.014***	0.011***	0.015***	0.013***	0.032***	0.026***	0.014***	0.015**	0.014***	0.013***	0.007*	0.004
	(0.004)	(0.003)	(0.005)	(0.004)	(0.006)	(0.006)	(0.005)	(0.007)	(0.004)	(0.003)	(0.004)	(0.003)
Iron	0.011**	0.009***	0.012**	0.010***	0.010**	0.009***	0.011***	0.009***	0.011***	0.009***	0.007**	0.007***
	(0.004)	(0.002)	(0.005)	(0.003)	(0.004)	(0.003)	(0.004)	(0.002)	(0.004)	(0.002)	(0.003)	(0.002)
Alloyants	0.022**	0.018**	0.020**	0.017**	0.025**	0.018**	0.019**	0.015*	0.018*	0.015*	0.012**	0.008*
	(0.010)	(0.008)	(0.010)	(0.009)	(0.011)	(0.008)	(0.010)	(0.008)	(0.010)	(0.008)	(0.006)	(0.004)
Soil Quality	−0.047***	−0.051***	−0.036*	−0.034	−0.061***	−0.056***	−0.048**	−0.054*	−0.045***	−0.048***	−0.040**	−0.041**
	(0.017)	(0.018)	(0.020)	(0.021)	(0.021)	(0.020)	(0.020)	(0.020)	(0.015)	(0.018)	(0.016)	(0.017)
Institutions × Autonomy		0.060***		0.054***		0.062***		0.064***		0.063***		0.040***
		(0.015)		(0.014)		(0.015)		(0.020)		(0.014)		(0.013)
Institutions		−0.034		−0.015		−0.036		−0.035		−0.032		0.001
		(0.023)		(0.022)		(0.026)		(0.022)		(0.021)		(0.018)
Autonomy		−0.019		−0.021		−0.015		−0.008		−0.027		−0.009
		(0.017)		(0.012)		(0.017)		(0.023)		(0.019)		(0.012)
Education											0.746***	0.782***
											(0.060)	(0.065)
R-squared (within)	0.121	0.147	0.107	0.124	0.141	0.166	0.218	0.248	0.153	0.190	0.322	0.376
Number of countries	100	86	99	86	70	59	100	86	100	86	88	76
Number of obs.	1825	1329	1735	1252	1404	1031	1824	1329	1819	1329	1588	1177

The table reports standardized beta coefficients of regressions of log GDP per capita on variables representing geographic factors and institutional quality for subnational regions around the world. In columns 3 and 4, subnational regions that contain the largest city in the country are excluded. In columns 5 and 6, data is excluded from countries for which GDP data are obtained from the UNDP. Columns 7 and 8 include controls for the presence of the 77 most prominent ethnic groups in the world, and columns 9 and 10 control for the 16 most prominent language groups in the world (not reported). Columns 11 and 12 control for education (average years) as reported in Gennaioli et al. (2013). Also included in the model but not reported are country fixed effects and a constant. Asterisks denote levels of statistical significance (*** = 1%, ** = 5%, * = 10%).

Column 1 reports a baseline robustness check that includes basic geographic variables, climatic variables, and natural resource variables simultaneously. For parsimony, column 1 reports only the ten geographic variables that are significant in this combined model. Column 2 reports results including the combined institutions variable and its interaction with autonomy (measured at the subnational level). The coefficient on institutions is insignificant, and institutions interacted with autonomy is positive and significant at the 1% level. Throughout Table 11, geographic variables are occasionally not significant in the even-numbered columns due to the smaller sample size.

In columns 3 and 4 I exclude subnational regions that include the largest city within the country, in an effort to evaluate whether the results presented are driven by the effect of city agglomeration within countries. The geographic variables remain significant in column 3 and the interaction of institutions with autonomy is positive and significant in column 4, showing that city agglomeration does not appear to be driving the results. In columns 5 and 6 I exclude countries for which per-capita GDP data come from the UNDP, since the source of per-capita GDP is not as well documented for these countries, as discussed in Section 2.1 above. Again the results are robust.

In columns 7 through 10 I add other variables to control for the effects of culture on development. Tabellini (2010) shows that culture is an important within-country determinant of development in Europe. I proxy for culture using geospatial databases that define the geographic occurrence of ethnicities and language families worldwide. Columns 7 and 8 include controls for ethnicity, where ethnicity data come from the database of ethnicities developed by Weidmann et al. (2010). I create a variable for each of 1268 ethnicities which equals, for each subnational region, the proportion of the region in which people of the ethnicity live. To limit the number of variables included in the regression, I exclude variables for which this proportion, on average, is less than 0.1%, thereby limiting the analysis to the 77 most prominent ethnic groups.[19] Columns 9 and 10 include controls for languages spoken in each subnational region, where language data come from the World Language Mapping System compiled by Global Mapping International. I create a variable for each of 107 prominent language families which equals, for each subnational region, the proportion of the subnational region (by area) in which the language family is spoken. Again, I exclude from the regression language family variables for which this proportion, on average, is less than 0.1%, thereby limiting the analysis to the 16 most prominent language families.[20] Columns 7 through 10 show that the r-squared increases when controlling for either language or ethnicity, indicating that proxies for culture have explanatory power for subnational development, consistent with Tabellini (2010). At the same time, columns 7 and 9 show that all geographic variables except *Percent Tropical* retain significance when controlling for language or ethnicity, and columns 8 and 10 show that the institutions interaction also remains significant.

In columns 11 and 12 I control for human capital, defined as the average years of schooling from primary school onward for the population aged 15 years or older (see row 52 of Table 2 for descriptive statistics). The data come from Gennaioli et al. (2013), who show that education is highly correlated with per-capita GDP in subnational regions around the world.[21] Of course, education could also be an outcome of other variables in the model, but I report the results for purposes of comparison to Gennaioli et al. (2013).[22] Columns 11 and 12 confirm the strong correlation of education with development in my sample of subnational regions. Meanwhile, column 11 shows that all but two of the geographic variables retain their significance, and column 12 shows that the institutions interaction remains significant when controlling for human capital.

6. Conclusion

In this paper, I bring a comprehensive subnational dataset to bear on the ongoing debate about the relative importance of geography and institutions for economic development. By exploiting subnational differences in geography and institutions I am able to extend our understanding of the determinants of development in a direction that has previously been underemphasized in the literature.

I find a significant role for a number of geographic features in explaining within-country incomes. On the other hand, across all countries in the sample, I do not find that favorable institutions improve economic outcomes. The overall relation between institutions and income is usually negative. Given existing literature on the positive role of institutions, this result is puzzling. But I find that the lack of a positive within-country effect of institutions may be due to the fact that subnational institutions are dominated by national institutions. Consistent with this interpretation, I find a positive relation between subnational institutions and development among subnational regions that have greater autonomy.

Ultimately, sorting out the determinants of development at the national and regional levels is a complex issue. Geographic and institutional factors do not operate in exactly the same way at the subnational level as at the national level. For example, if rents from a particular natural resource tend to accrue to national governments rather than regional entities, then within-country estimates may understate the importance of that natural resource. On the flip side, if a geographic factor has substantial within-country variability, then cross-country estimates may understate the importance of that geographic factor. Regarding institutional factors, my analysis suggests that national institutions and regional institutions do not operate on a level playing field—that national institutions seem to dominate except under the right conditions.

Taking my findings back to the overall debate on the relative importance of geography and institutions, a major takeaway of this analysis is that, simply put, geography matters. In their analysis of the relative importance of geography, institutions, and trade for growth, Rodrik et al. (2004) conclude that "the quality of institutions trumps everything else"; that "geography has at best weak direct effects" on income after controlling for institutions. But my analysis shows that after completely controlling for institutions and all other country-level factors (i.e., with country fixed effects), geographic variables have a significant impact on per-capita GDP. The combined explanatory power of geographic variables is not trivial, with r-squareds of over 12% across over 1800 regions. On the other hand, after controlling for country-level factors, the importance of regional institutions is harder to find. But regional institutions do seem to play a significant role when "given a chance"; that is, when they are not dominated by national institutions.

The interplay of national institutions and local institutions in shaping economic outcomes is an issue that merits further study. From a policy perspective, it is important to know whether institutional reforms targeted at the local level can have a meaningful impact on economic development in the absence of local autonomy. If not, then either the focus for institutional reform should be on the national level, or local institutional reforms need to be implemented selectively and cautiously in order to be effective.

[19] Examples of significant correlations of ethnic groups with per-capita GDP in the regressions include a negative effect of some African ethnic groups (Basuto, Ewe, Hausa) and a positive effect of the Han Chinese and the English.

[20] Examples of significant correlations of language families with per-capita GDP in the regressions include a negative effect of native language families of Latin America (Quechuan and Oto-Manguean) and a positive effect of the primary language family of sub-Saharan Africa (Niger-Congo).

[21] Education differs from the other variables tested in the study in that it would be classified as a proximate cause of development rather than a fundamental cause. On the differences between fundamental (or "deep") causes of development and proximate causes of development, see North and Thomas (1973), Rodrik et al. (2004) or Acemoglu (2009).

[22] It is also conceivable that the ethnic and linguistic controls (in columns 7 through 10) could be outcomes of other variables in the model.

Appendix A

Appendix Table A1
Sources for GDP data.

Country	GDP data source	Year
Albania	Human Development Promotion Center	2001
Argentina	National Institute of Statistics and Census	2005
Australia	Australian Bureau of Statistics	2005
Austria	Eurostat (NUTS 2 level)	2005
Bangladesh	Canback Dangel	2005
Belarus	National Statistical Committee	2007
Belgium	Eurostat (NUTS 2 level)	2005
Benin	United Nations Development Program	2004
Bolivia	National Institute of Statistics	2005
Bosnia and Herzegovina	United Nations Development Program	2004
Brazil	Brazilian Institute of Geography and Statistics (IBGE)	2005
Bulgaria	Eurostat (NUTS 3 level)	2005
Burkina Faso	United Nations Development Program	2001
Canada	Statistics Canada	2005
Chile	Chilean Central Bank Statistical Database	2005
China	National Bureau of Statistics of China	2005
Colombia	National Administrative Department of Statistics	2005
Congo, Dem. Rep.	United Nations Development Program	2005
Croatia	Eurostat (NUTS 3 level)	2005
Czech Republic	Eurostat (NUTS 3 level)	2005
Denmark	Eurostat (NUTS 2 level)	2005
Dominican Republic	UNDP	2004
Ecuador	Central Bank of Ecuador	2005
Egypt	United Nations Development Program	2005
El Salvador	United Nations Development Program	2005
Estonia	Eurostat (NUTS 3 level)	2005
Finland	Eurostat (NUTS 3 level)	2005
France	Eurostat (NUTS 2 level)	2005
Gambia, The	Central Statistics Department/UNDP	2000
Georgia	National Statistics Office of Georgia	2006
Germany	Eurostat (NUTS 2 level)	2005
Greece	Eurostat (NUTS 2 level)	2005
Guatemala	United Nations Development Program	2006
Honduras	United Nations Development Program	2005
Hungary	Eurostat (NUTS 3 level)	2005
India	Ministry of Statistics and Programme Implementation	2005
Indonesia	BPS Statistics Indonesia	2005
Iran	Statistical Centre of Iran	2005
Ireland	Eurostat (NUTS 3 level)	2005
Italy	Eurostat (NUTS 2 level)	2005
Japan	Japan Statistics Bureau	2005
Jordan	United Nations Development Program	2002
Kazakhstan	Agency of Statistics of the Republic of Kazakhstan	2005
Kenya	UNDP Kenya	2005
Kosovo	UNDP	2004
Kyrgyzstan	National Statistical Committee	2005
Lao PDR	UNDP	2002
Latvia	Eurostat (NUTS 3 level)	2005
Lebanon	UNDP	2004
Lesotho	UNDP Lesotho	2001
Lithuania	Eurostat (NUTS 3 level)	2005
Macedonia	Eurostat (NUTS 3 level)	2005
Malaysia	Department of Statistics Malaysia	2005
Malta	Eurostat (NUTS 3 level)	2005
Mexico	National Institute of Statistics and Geography (INEGI)	2005
Micronesia	Division of Statistics	2005
Mongolia	National Statistical Office of Mongolia	2005
Montenegro	UNDP	2007
Morocco	Ministry of Economy and Finance	2005
Mozambique	Instituto Nacional de Estatística	2005
Nepal	UNDP	2001
Netherlands	Eurostat (NUTS 2 level)	2005
New Zealand	Statistics New Zealand	2003
Niger	UNDP	2001
Nigeria	Canback Dangel	2001
Norway	Statistics Norway	2005
Pakistan	Canback Dangel	2005
Panama	National Institute of Statistics and Census	2005
Paraguay	UNDP	2006
Peru	National Institute of Statistics and Information	2005
Philippines	National Statistical Coordination Board	2005
Poland	Eurostat (NUTS 2 level)	2005

Appendix Table A1 (*continued*)

Country	GDP data source	Year
Portugal	Eurostat (NUTS 2 level)	2005
Romania	Eurostat (NUTS 3 level)	2005
Russia	Russian Federation Federal State Statistics Service	2005
Senegal	UNDP	1999
Serbia	Statistical Office of the Republic of Serbia	2009
Slovak Republic	Eurostat (NUTS 3 level)	2005
Slovenia	Eurostat (NUTS 3 level)	2005
South Africa	Statistics South Africa	2005
South Korea	Statistics Korea	2005
Spain	Eurostat (NUTS 2 level)	2005
Sri Lanka	Central Bank of Sri Lanka	2005
Swaziland	UNDP Swaziland	2006
Sweden	Eurostat (NUTS 3 level)	2005
Switzerland	Swiss Federal Statistical Office	2005
Syrian Arab Republic	UNDP	2002
Tanzania	National Bureau of Statistics	2005
Thailand	Office of the National Economic and Social Development Board	2005
Turkey	Eurostat (NUTS 2 level)	2005
Uganda	United Nations Development Program	2005
Ukraine	State Statistics Committee of Ukraine	2005
United Arab Emirates	United Arab Emirates Ministry of Economy	2005
United Kingdom	Eurostat (NUTS 2 level)	2005
United States	Bureau of Economic Analysis	2005
Uruguay	United Nations Development Program	2002
Uzbekistan	United Nations Development Program	2005
Venezuela	Instituto Nacional de Estadística	2005
Vietnam	General Statistics Office of Vietnam	2005
Zambia	United Nations Development Program	2004
Zimbabwe	United Nations Development Program, collaborators	2001

The table reports, for each country in the sample, the source of the subnational GDP data used in the study and the year from which the data are taken. Data from years other than 2005 is adjusted with GDP deflators to be comparable to 2005 data.

Appendix Table A2
Description of variables.

Variable	Definition	Source
Economic variables		
GDP per capita	Log value of GDP per capita in the subnational region	Various (see Appendix Table A1)
General geographic variables		
Latitude	Absolute value of the latitude (in degrees) of the capital city of the subnational region	Google Earth
Elevation	Elevation (log thousands of feet plus one) of the capital city of the subnational region	Google Earth
Ocean Access	Dummy variable equal to 1 if the subnational region has ocean access; 0 if it is landlocked; access through bays, straits, and inlets (but not rivers) is coded as 1	Observation of maps
Terrain Ruggedness	Average terrain ruggedness (in hundreds of meters) across all 30 by 30 arc-second cells contained within the subnational region	Nunn and Puga (2012)
Storm Risk	Log number of occurrences of hurricanes and tropical storms in the region between 1842 and 2010	NOAA
Earthquake Risk	Log number of fault lines present in the subnational region	USGS and Esri
Malaria Risk	Average of all data points within the subnational region from a 0.5° global grid of malaria risk based on climatic conditions (temperature and precipitation) and the biologic characteristics of vector mosquitoes	Kiszewski et al. (2004)
Climate variables		
Percent Tropical	Percentage of subnational region (by land area) covered by tropical and subtropical biomes	World Wildlife Fund terrestrial ecoregions map
Precipitation	Average monthly rainfall (mm) over the period 1961–1990, averaged across all data points within the subnational region	Climatic Research Unit, University of East Anglia, High-resolution (10′) gridded global dataset (CRU CL 2.0)
Rain Days	Average number of days per month with rainfall > 0.1 mm over the period 1961–1990, averaged across all data points within the subnational region	Climatic Research Unit, University of East Anglia, High-resolution (10′) gridded global dataset (CRU CL 2.0)
Temperature	Monthly average of daily mean temperature (Celsius) over the period 1961–1990, averaged across all data points within the subnational region	Climatic Research Unit, University of East Anglia, High-resolution (10′) gridded global dataset (CRU CL 2.0)
Temperature Range	Monthly average of daily diurnal temperature range (Celsius) over the period 1961–1990, averaged across all data points within the subnational region	Climatic Research Unit, University of East Anglia, High-resolution (10′) gridded global dataset (CRU CL 2.0)
Humidity	Average monthly relative humidity (per day in percent) over the period 1961–1990, averaged across all data points within the subnational region	Climatic Research Unit, University of East Anglia, High-resolution (10′) gridded global dataset (CRU CL 2.0)
Sunshine	Average monthly percent of day length with sunshine over the period 1961–1990, averaged across all data points within the subnational region	Climatic Research Unit, University of East Anglia, High-resolution (10′) gridded global dataset (CRU CL 2.0)
Frost Days	Average number of days per month with ground frost over the period 1961–1990, averaged across all data points within the subnational region	Climatic Research Unit, University of East Anglia, High-resolution (10′) gridded global dataset (CRU CL 2.0)
Climate variables		
Wind Speed	Average daily wind speed per month (meters/second) over the period 1961–1990, averaged across all data points within the subnational region	Climatic Research Unit, University of East Anglia, High-resolution (10′) gridded global dataset (CRU CL 2.0)

(*continued on next page*)

Appendix Table A2 (*continued*)

Variable	Definition	Source
Natural resource variables		
Oil and Gas	The number of identifiable oil and/or natural gas sites within the subnational region (in thousands)	USGS and Petroconsultants International Data Corporation (transformed from NAD 1927 to WGS 84 6)
Diamonds	The number of identifiable mineral sites containing diamonds within the subnational region (in thousands)	Mineral Resources Data System of the United States Geological Survey
Precious Metals	The number of identifiable mineral sites containing precious metals (gold, silver, or the platinum group) within the subnational region (in thousands)	Mineral Resources Data System of the United States Geological Survey
Base Metals	The number of identifiable mineral sites containing base metals (copper, lead, nickel, zinc, aluminum, bismuth, cobalt, gallium, indium, magnesium, mercury, potassium, titanium, tin, uranium, or zirconium) within the subnational region (scaled by land area in the region in km^2)	Mineral Resources Data System of the United States Geological Survey
Iron	The number of identifiable mineral sites containing iron within the subnational region (scaled by land area in the region in km^2)	Mineral Resources Data System of the United States Geological Survey
Alloyants	The number of identifiable mineral sites containing steel alloyants (carbon, manganese, chromium, molybdenum, vanadium, boron, cerium, niobium, or tungsten) within the subnational region (scaled by land area in the region in km^2)	Mineral Resources Data System of the United States Geological Survey
Water	Area of inland lakes and rivers (in km^2) within the subnational region (scaled by land area in the region in km^2)	World Water Bodies map published by Delorme, Inc., and made available by Esri
Percent *Agricultural Land*	Percentage of land in the subnational region that is suitable for agriculture	Harmonized World Soil Database
Soil Quality	Average *Soil Quality* of points within the subnational region, averaged across 7 characteristics: ease of tillage, nutrient availability, nutrient retention, oxygen availability, pH levels, rooting conditions, and salinity. Each characteristic is rated on an integer scale of 1 to 4, with 4 indicating the worst *Soil Quality*, so the final variable is given a negative sign.	Harmonized World Soil Database
Institutions variables		
Property Rights	Aggregated score, first within surveys and then across surveys, of all survey questions that fall within the category of property rights. Before aggregation, responses with no upper bound are logged, all questions are made directionally consistent (with higher values indicating better quality), and all questions are standardized to a mean of zero and standard deviation of one.	Various (see Appendix Tables A3 through A9)
Corruption Control	Aggregated score, first within surveys and then across surveys, of all survey questions that fall within the category of corruption control. Before aggregation, responses with no upper bound are logged, all questions are made directionally consistent (with higher values indicating better quality), and all questions are standardized to a mean of zero and standard deviation of one.	Various (see Appendix Tables A3 through A9)
Law and Order	Aggregated score, first within surveys and then across surveys, of all survey questions that fall within the category of law and order. Before aggregation, responses with no upper bound are logged, all questions are made directionally consistent (with higher values indicating better quality), and all questions are standardized to a mean of zero and standard deviation of one.	Various (see Appendix Tables A3 through A9)
Regulatory Efficiency	Aggregated score, first within surveys and then across surveys, of all survey questions that fall within the category of regulatory efficiency. Before aggregation, responses with no upper bound are logged, all questions are made directionally consistent (with higher values indicating better quality), and all questions are standardized to a mean of zero and standard deviation of one.	Various (see Appendix Tables A3 through A9)
Rule of Law (country level)	Aggregated variable "capturing 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".	World Bank, World Governance Indicators
Corruption Control (country level)	Aggregated variable "capturing perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as 'capture' of the state by elites and private interests".	World Bank, World Governance Indicators
Stability/Lack of Violence (country level)	Aggregated variable "capturing perceptions of the likelihood of political instability and/or politically-motivated violence, including terrorism".	World Bank, World Governance Indicators
Regulatory Quality (country level)	Aggregated variable "capturing perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development."	World Bank, World Governance Indicators
Other variables		
Autonomy (subnational)	Dummy variable equal to 1 if the subnational region is autonomous or partly autonomous; 0 otherwise	*List of Autonomous Areas by Country*, Wikipedia
Autonomy (country level)	Dummy variable equal to 1 if the country contains autonomous regions; 0 otherwise	Database of Political Institutions (see Beck et al., 2001)
State elections	Equal to 0 if neither state executive nor state legislature are locally elected; 1 if the executive is appointed and the legislature locally elected; 2 if both are locally elected	Database of Political Institutions (see Beck et al., 2001)
Municipal elections	Equal to 0 if neither municipal executive nor municipal legislature are locally elected; 1 if the executive is appointed and the legislature locally elected; 2 if both are locally elected	Database of Political Institutions (see Beck et al., 2001)
State authority	Equal to 1 if the state has authority of taxing, spending, or legislating; 0 otherwise	Database of Political Institutions (see Beck et al., 2001)
Senators elected	Equal to 1 if senators representing the region are locally elected; 0 otherwise	Database of Political Institutions (see Beck et al., 2001)
Other variables		
Ethnicity	A set of 77 variables representing the percentage (by area) of each subnational region that is home to the given ethnicity	Weidmann et al. (2010)
Language	A set of 16 variables representing the percentage (by area) of each subnational region	World Language Mapping System, published by Global

Appendix Table A2 (*continued*)

Variable	Definition	Source
	that speaks languages in the given language family	Mapping International
Education	Average years of schooling from primary school onward for the population aged 15 years or older	Gennaioli et al. (2013)

The table presents the definition and source of each of the variables used in the empirical analysis.

Appendix Table A3
Afrobarometer survey questions used in subnational scores.

Question number	Question	Min. response	Max. response	Response Direction
Category 1: Property Rights				
55I	How much do you trust the courts of law?	0	3	+
Category 2: Corruption Control				
56C	How many elected local government councilors do you think are involved in corruption?	0	3	−
56E	How many local government officials do you think are involved in corruption?	0	3	−
56F	How many of the police do you think are involved in corruption?	0	3	−
56H	How many judges and magistrates do you think are involved in corruption?	0	3	−
57A	In the past year, how often have you had to pay a bribe to get a document or permit?	0	3	−
57B	In the past year, how often have you had to pay a bribe to get a child into school?	0	3	−
57C	In the past year, how often have you had to pay a bribe to get a household service?	0	3	−
57D	In the past year, how often have you had to pay a bribe to get medical attention?	0	3	−
57E	In the past year, how often have you had to pay a bribe to avoid a problem with police?	0	3	−
Category 3: Law and Order				
55H	How much do you trust the police?	0	3	+
70A	Could authorities enforce the law if a top government official committed a serious crime?	1	4	+
70B	Could authorities enforce the law if a person like you committed a serious crime?	1	4	+
70C	Could authorities enforce the law if a top official did not pay some income tax?	1	4	+
70D	Could authorities enforce the law if a person like you did not pay some income tax?	1	4	+
Category 4: Regulatory Efficiency				
55D	How much do you trust the elected local government council?	0	3	+
71A	Based on your experience, how easy or difficult is it to obtain an identity document?	1	4	+
71B	Based on your experience, how easy or difficult is it to obtain household services?	1	4	+
71C	Based on your experience, how easy or difficult is it to obtain help from the police?	1	4	+
71D	Based on your experience, how easy or difficult is it to obtain a place in primary school for a child?	1	4	+
71E	Based on your experience, how easy or difficult is it to obtain medical treatment at a nearby clinic?	1	4	+

The table reports the questions extracted from the Afrobarometer survey to be used in the calculation of measures of institutional quality for subnational regions in African countries. The questions from the survey (edited for brevity) and the range of possible responses are reported. "Response Direction" indicates whether a higher response indicates better institutions (+) or worse institutions (−). Data are taken from 17,950 individuals surveyed in 2005. (Data for Burkina Faso come from 2008 and include a subset of the questions listed.)

Appendix Table A4
Latin American Public Opinion Project questions used in subnational scores.

Question number	Question	Min. response	Max. response	Response Direction
Category 1: Property Rights				
AOJ12	If you were crime victim, how much faith do you have that the judicial system would punish the guilty?	1	4	−
ST2	How satisfied are you with dealings your family has had with the courts or justice tribunals?	1	4	−
ST3	How satisfied are you with dealings your family has had with the district attorney's office?	1	4	−
B1	To what extent do the courts of justice guarantee a fair trial?	1	7	+
B3	To what extent do you think citizens' basic rights are protected by the political system?	1	7	+
B10A	To what extent do you trust the system of justice?	1	7	+
B16	To what extent do you trust the state attorney general?	1	7	+
B17	To what extent do you trust the public defender's office?	1	7	+
Category 2: Corruption Control				
EXC2	Has any police official asked you for a bribe in the last year?	No	Yes	−
EXC6	During the last year has any public official asked you for a bribe?	No	Yes	−
EXC7	Based on your own experience, do you believe corruption among public officials is common?	1	4	+
EXC11	During the last year have you had to pay a bribe to process a document with the municipality?	No	Yes	−
EXC14	Have you had to give a bribe to the courts in the last year?	No	Yes	−
Category 2: Corruption Control				
EXC15	Have you had to give a bribe to obtain public health services in the last year?	No	Yes	−
Category 2: Corruption Control				
EXC16	Have you had to give a bribe at you child's school in the last year?	No	Yes	−
EXC17	Has anyone asked you for a bribe to avoid having the electricity turned off?	No	Yes	−
N9	To what extent would you say the current government combats government corruption?	1	7	+
Category 3: Law and Order				
VIC1	Have you been the victim of any type of crime in the past 12 months?	Yes	No	+
AOJ11	In your neighborhood, do you feel safe?	1	4	−
AOJ18	Do the police in your neighborhood protect people or are they involved in crime?	1	2	−
N11	To what extent would you say the current government improves citizen security?	1	7	+

(*continued on next page*)

Appendix Table A4 (*continued*)

Question number	Question	Min. response	Max. response	Response Direction
Category 4: Regulatory Efficiency				
B32	To what extent do you trust the mayor's office of your municipality?	1	7	+
ST4	How satisfied are you with dealings your family has had with the municipality?	1	4	−
SGL1	How are the services that the municipality is providing?	1	5	−

The table reports the questions extracted from the Latin American Public Opinion Project to be used in the calculation of measures of institutional quality for subnational regions in the Americas. The questions from the survey (edited for brevity) and the range of possible responses are reported. "Response Direction" indicates whether a higher response indicates better institutions (+) or worse institutions (−). Data are taken from 27,650 individuals surveyed in 2006. (Data for Argentina come from 2008 and include a subset of the questions listed. Data from the U.S. and Canada also include a subset of the questions listed.)

Appendix Table A5
Asia Foundation survey questions used in subnational scores.

Country	Sub-index	Min. score	Max. score	Score Direction
Category 1: Property Rights				
Bangladesh	Dispute resolution; access to land and security of tenure	0	20	+
Indonesia	Land access and security of tenure	0	100	+
Malaysia	Property rights and dispute resolution; access to land and security of tenure	0	20	+
Nepal	Property rights and dispute resolution	3	15	+
Philippines	Legal service quality	1	5	+
Sri Lanka	Legal institutions and conflict resolution; land access and property rights	0	18	+
Thailand	Legal institutions	3	12	−
Vietnam	Legal institutions; land access and security of tenure	0	20	+
Category 2: Corruption Control				
Bangladesh	Informal charges	0	10	+
Malaysia	Informal charges	0	10	+
Nepal	Informal charges	3	12	+
Philippines	Corruption prevention	5	20	+
Sri Lanka	Informal charges, favoritism, and discrimination	0	9	+
Vietnam	Informal charges	0	10	+
Category 3: Law and Order				
Bangladesh	Law and order	0	10	+
Indonesia	Security and conflict resolution	0	100	+
Malaysia	Crime and security	0	10	+
Nepal	Crime	2	4	+
Philippines	Crime	2	10	+
Sri Lanka	Crime and security	0	9	+
Thailand	Crime and security	5	19	−
Category 4: Regulatory Efficiency				
Bangladesh	Entry costs; time cost of regulatory compliance	0	20	+
Indonesia	Business licensing; local regulation	0	100	+
Malaysia	Entry costs; regulatory costs	0	20	+
Nepal	Time cost (in days) of regulatory compliance	1	NA	−
Philippines	Regulatory performance and trust	6	30	+
Sri Lanka	Registration, permits, and licenses; regulatory environment, compliance, and cost	0	18	+
Vietnam	Entry costs; time cost of regulatory compliance	0	20	+

The table summarizes scores taken from various surveys conducted by The Asia Foundation and their partners that are used in the calculation of measures of institutional quality for subnational regions in Asian countries. The sub-indices taken from the survey (each of which is derived from multiple questions) and the range of possible scores are reported. Sub-indices are created by the survey sponsors except for Nepal, the Philippines, and Thailand, where I aggregate and name the sub-indices from available survey questions. "Score Direction" indicates whether a higher score indicates better institutions (+) or worse institutions (−). Data are taken from 31,903 firms and individuals surveyed between 2006 and 2011.

Appendix Table A6
Quality of Government Institute survey questions used in subnational scores.

Question	Min. response	Max. response	Response Direction
Category 1: Property Rights			
NA			
Category 2: Corruption Control			
How likely is it the corruption by a public employee or politician would be exposed by the local mass media?	0	10	+
Do you agree that elections in your area are honest and clean from corruption?	0	10	+
In the past 12 months has anyone in your household paid a bribe to health or medical services?	Yes	No	+
Do you agree that corruption is prevalent in the police force in your area?	0	10	+
Do you agree that corruption is prevalent in your area's local public school system?	0	10	+
Do you agree that corruption is prevalent in the public health care system in your area?	0	10	+
Category 3: Law and Order			
How would you rate the quality of the police force in your area?	0	10	+
Does the police force give special advantages to certain people in your area?	0	10	+

Appendix Table A6 (*continued*)

Question	Min. response	Max. response	Response Direction
Are all citizens treated equally be the police force in your area?	1	4	+
Category 4: Regulatory Efficiency			
NA			

The table reports the questions extracted from the Quality of Government survey to be used in the calculation of measures of institutional quality for subnational regions in European countries. The questions from the survey (edited for brevity) and the range of possible responses are reported. "Response Direction" indicates whether a higher response indicates better institutions (+) or worse institutions (−). Data are taken from 33,540 individuals surveyed between 2009 and 2010.

Appendix Table A7
Latinobarómetro survey questions used in subnational scores.

Question number	Question	Min. response	Max. response	Response Direction
Category 1: Property Rights				
P40STB	Do you agree that justice arrives late, but it arrives?	1	4	−
P42STD	How much confidence do you have in the judiciary?	1	4	−
Category 2: Corruption Control				
P82STB	Has anyone in your family known of an act of corruption in the last 12 months?	Yes	No	+
P84ST	If the total number of public employees were 100, how many would you say are corrupted?	0	100	−
Category 3: Law and Order				
P42STF	How much confidence do you have in the police?	1	4	−
P76ST	How well can the state enforce laws?	1	10	+
P82STA	Has anyone in your family been the victim of a crime in the last 12 months?	Yes	No	+
Category 4: Regulatory Efficiency				
P30ST	How good is the operation of public institutions?	1	5	−
P40STD	Do you agree that you can trust the people who run our government to do what is right?	1	4	−
P42STE	How much confidence do you have in municipalities/local government?	1	4	−

The table reports the questions extracted from the Latinobarómetro survey to be used in the calculation of measures of institutional quality for subnational regions in Latin American countries. The questions from the survey (edited for brevity) and the range of possible responses are reported. "Response Direction" indicates whether a higher response indicates better institutions (+) or worse institutions (−). Data are taken from 20,222 individuals surveyed in 2005.

Appendix Table A8
World Bank Enterprise Survey questions used in subnational scores.

Question number	Question	Min. response	Max. response	Response Direction
Category 1: Property Rights				
H3	In reference to a payment dispute, was a court judgment made?	Yes	No	−
H4	How many weeks did it take for the court to reach judgment?	0	NA	−
H5	Was the decision of the court enforced?	Yes	No	−
H6	How many weeks did it take for court enforcement to take place?	0	NA	−
H7A	Do you agree that the court system is fair, impartial and uncorrupted?	1	4	+
H7B	Do you agree that the court system is quick?	1	4	+
H7C	Do you agree that the court system is affordable?	1	4	+
H7D	Do you agree that the court system is able to enforce its decisions?	1	4	+
H30	How much of an obstacle are the courts to the operations of this establishment?	0	4	−
Category 2: Corruption Control				
C5	Was an informal gift or payment expected or requested for an electrical connection?	Yes	No	+
C14	Was an informal gift or payment expected or requested for a water connection?	Yes	No	+
C21	Was an informal gift or payment expected or requested for a telephone connection?	Yes	No	+
G4	Was an informal gift or payment expected or requested for a construction-related permit?	Yes	No	+
Category 2: Corruption Control				
J1B	Do you agree that it is common to pay informal payments or gifts to get things done?	1	4	−
J5	In meetings with tax officials was a gift or informal payment expected or requested?	Yes	No	+
J6	In dealing with government, what percent of contract value is paid in informal payments to secure the contract?	0	NA	−
J7A	What percent of annual sales would be paid in informal payments or gifts to public officials to "get things done"?	0	NA	−
J12	Was an informal gift or payment expected or requested for an import license?	Yes	No	+
J15	Was an informal gift or payment expected or requested for an operating license?	Yes	No	+
J30F	How much of an obstacle is corruption to the operations of this establishment?	0	4	−
Category 3: Law and Order				
I1	In last fiscal year, did this establishment pay for security?	Yes	No	+
I2A	What percent of total annual sales is paid for security?	0	NA	−
I4A	What percent of annual sales is lost to theft, robbery, vandalism, or arson that occurred on your premises?	0	NA	−
I30	How much of an obstacle are crime, theft, and disorder to the operations of this establishment?	0	4	−
J30E	How much of an obstacle is political instability to the operations of this establishment?	0	4	−

(*continued on next page*)

Appendix Table A8 (*continued*)

Question number	Question	Min. response	Max. response	Response Direction
Category 4: Regulatory Efficiency				
C4	Approximately what was the wait, in days, to receive an electrical connection?	0	NA	−
C13	Approximately what was the wait, in days, to receive a water connection?	0	NA	−
C20	Approximately what was the wait, in days, to receive a telephone connection?	0	NA	−
D30B	Are customs and trade regulations are an obstacle to the operations of this establishment?	0	4	−
G3	Approximately what was the wait, in days, to receive a construction permit?	0	NA	−
J2	What percentage of senior management's time is spent in dealing with government regulations?	0	100	−
J11	Approximately what was the wait, in days, to receive an import license?	0	NA	−
J14	Approximately what was the wait, in days, to receive an operating license?	0	NA	−
J30C	How much of an obstacle are business licensing and permits to the operations of this establishment?	0	4	−

The table reports the questions extracted from the World Bank Enterprise Survey to be used in the calculation of measures of institutional quality for subnational regions around the world. The questions from the survey (edited for brevity) and the range of possible responses are reported. "Response Direction" indicates whether a higher response indicates better institutions (+) or worse institutions (−). Data are taken from 40,792 firms surveyed between 2006 and 2011.

Appendix Table A9
Doing Business questions used in subnational scores.

Question	Min. Response	Max. Response	Response Direction
Category 1: Property Rights			
How many procedures are required to register property?	1	NA	−
How much time (in days) is required to register property?	1	NA	−
What is the cost (as a % of property value) to register property?	0	NA	−
How many procedures are required to enforce a contract?	1	NA	−
How much time (in days) is required to enforce a contract?	1	NA	−
What is the cost (as a % of the claim) to enforce a contract?	0	NA	−
Category 2: Corruption Control			
NA			
Category 3: Law and Order			
NA			
Category 4: Regulatory Efficiency			
How many procedures are required to start a business?	1	NA	−
How much time (in days) is required to start a business?	1	NA	−
What is the cost (as a % of income per capita) to start a business?	0	NA	−
What is the minimum paid-in capital (as a % of income per capita) required to start a business?	0	NA	−
How many procedures are required to deal with construction permits?	1	NA	−
How much time (in days) is required to deal with construction permits?	1	NA	−
What is the cost (as a % of income per capita) to deal with construction permits?	0	NA	−

The table reports the questions covered in the World Bank's Doing Business project that are used in the calculation of measures of institutional quality for subnational regions in countries around the world. The questions and the range of possible responses are reported. "Response Direction" indicates whether a higher response indicates better institutions (+) or worse institutions (−). Data are compiled from examination of laws and expert surveys between 2006 and 2011.

Appendix B. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jdeveco.2015.09.002.

References

Acemoglu, Daron, 2009. Introduction to Modern Economic Growth. Princeton University Press, Princeton, NJ.

Acemoglu, Daron, Dell, Melissa, 2010. Productivity differences between and within countries. Am. Econ. J. Macroecon. 2, 169–188.

Acemoglu, Daron, Johnson, Simon, Robinson, James A., 2001. The colonial origins of comparative development: an empirical investigation. Am. Econ. Rev. 91, 1369–1401.

Acemoglu, Daron, Johnson, Simon, Robinson, James A., 2002. Reversal of fortune: geography and institutions in the making of the modern world income distribution. Q. J. Econ. 117, 1231–1294.

Acemoglu, Daron, Garca-Jimeno, Camilo, Robinson, James A., 2012. Finding El Dorado: slavery and long-run development in Colombia. J. Comp. Econ. 40, 534–564.

Acemoglu, Daron, Gallego, Francisco, Robinson, James A., 2014. Institutions, human capital and development. Ann. Rev. Econ. 6, 875–912.

Alesina, Alberto, Spolaore, Enrico, 1997. On the number and size of nations. Q. J. Econ. 112, 1027–1056.

Banerjee, Abhijit, Iyer, Lakshmi, 2005. History, institutions, and economic performance: the legacy of colonial land tenure systems in India. Am. Econ. Rev. 95, 1190–1213.

Beck, Thorsten, Clarke, George, Groff, Alberto, Keefer, Philip, Walsh, Patrick, 2001. New tools in comparative political economy: the database of political institutions. World Bank Econ. Rev. 15, 165–176.

Bloom, David E., Sachs, Jeffrey D., 1998. Geography, demography, and economic growth in Africa. Brook. Pap. Econ. Act. 2, 207–295.

Bruhn, Miriam, Gallego, Francisco, 2012. Good, bad, and ugly colonial activities: do they matter for economic development? Rev. Econ. Stat. 94, 433–461.

Checchi, Daniele, Florio, Massimo, Carrera, Jorge, 2009. Privatisation discontent and utility reform in Latin America. J. Dev. Stud. 45, 333–350.

Chong, Alberto, La Porta, Rafael, Lopez-de-Silanes, Florencio, Shleifer, Andrei, 2014. Letter grading government efficiency. J. Eur. Econ. Assoc. 12, 277–299.

Dell, Melissa, 2010. The persistent effects of Perú's Mining Mita. Econometrica 78, 1863–1903.

Dell, Melissa, Jones, Benjamin, Olken, Benjamin, 2009. Temperature and income: reconciling new cross-sectional and panel estimates. Am. Econ. Rev. Pap. Proc. 99, 198–204.

Diamond, Jared, 1997. Guns, Germs, and Steel: The Fates of Human Societies. W.W. Norton and Co., New York.

Dollar, David, Kraay, Aart, 2003. Institutions, trade, and growth. J. Monet. Econ. 50, 133–162.

Easterly, William, Levine, Ross, 2003. Tropics, germs, and crops: how endowments influence economic development. J. Monet. Econ. 50, 3–39.

Frankel, Jeffrey, 2010. The natural resource curse: a survey. NBER Working Paper #15836.

Frankel, Jeffrey, Romer, David, 1999. Does trade cause growth? Am. Econ. Rev. 89, 379–399.

Gallup, John Luke, Sachs, Jeffrey D., Mellinger, Andrew D., 1999. Geography and economic development. Int. Reg. Sci. Rev. 22, 179–232.

Gennaioli, Nicola, La Porta, Rafael, Lopez-de-Silanes, Florencio, Shleifer, Andrei, 2013. Human capital and regional development. Q. J. Econ. 128, 105–164.

Gries, Thomas, Naude, Wim, 2008. Trade and endogenous formation of regions in a developing country. Rev. Dev. Econ. 12, 248–275.

Hall, Robert E., Jones, Charles I., 1999. Why do some countries produce so much more output per worker than others? Q. J. Econ. 114, 83–116.
Iyer, Lakshmi, 2010. Direct versus indirect colonial rule in India: long-term consequence. Rev. Econ. Stat. 92, 693–713.
Kaufmann, Daniel, Kraay, Aart, 2008. Governance indicators: where are we, where should we be going? World Bank Res. Obs. 23, 1–30.
Kiszewski, Anthony, Mellinger, Andrew, Spielman, Andrew, Malaney, Pia, Sachs, Sonia Ehrlich, Sachs, Jeffrey, 2004. A global index representing the stability of malaria transmission. Am. J. Trop. Med. Hyg. 70, 486–498.
Klapper, Leora, Laeven, Luc, Rajan, Raghuram, 2006. Entry regulation as a barrier to entrepreneurship. J. Financ. Econ. 82, 591–629.
Ma, Yue, Qu, Baozhi, Zhang, Yifan, 2010. Judicial quality, contract intensity and trade: firm-level evidence from developing and transition countries. J. Comp. Econ. 38, 146–159.
Masters, William A., McMillan, Margaret, 2001. Climate and scale in economic growth. J. Econ. Growth 6, 167–186.
Mauro, Paolo, 1995. Corruption and growth. Q. J. Econ. 110, 681–712.
Mehlum, Halvor, Moene, Karl, Torvik, Ragnar, 2006. Institutions and the resource curse. Econ. J. 116, 1–20.
Michalopoulos, Stelios, Papaioannou, Elias, 2013. Pre-colonial ethnic institutions and contemporary African development. Econometrica 81, 113–152.
Nordhaus, William D., 2006. Geography and macroeconomics: new data and findings. Proc. Natl. Acad. Sci. U.S.A. 103, 3510–3517.
North, Douglass C., Thomas, Robert P., 1973. The Rise of the Western World: A New Economic History. Cambridge University Press, Cambridge, U.K.
Nunn, Nathan, 2010. Religious conversion in colonial Africa. Am. Econ. Rev. Pap. Proc. 100, 147–152.
Nunn, Nathan, Puga, Diego, 2012. Ruggedness: the blessing of bad geography in Africa. Rev. Econ. Stat. 94, 20–36.
Quality of Government Institute, 2010. Measuring the quality of government and subnational variation. Report for the European Commission. University of Gothenburg, Sweden.
Robinson, James, Torvik, Ragnar, Verdier, Theirry, 2006. Political foundations of the resource curse. J. Dev. Econ. 79, 447–468.
Rodrik, Dani, Subramanian, Arvind, Trebbi, Francesco, 2004. Institutions rule: the primacy of institutions over geography and integration in economic development. J. Econ. Growth 9, 131–165.
Sachs, Jeffrey D., 2003. Institutions don't rule: direct effects of geography on per capita income. NBER Working Paper #9490.
Tabellini, Guido, 2010. Culture and institutions: economic development in the regions of Europe. J. Eur. Econ. Assoc. 8, 677–716.
van der Ploeg, Frederick, 2011. Natural resources: curse or blessing? J. Econ. Lit. 49, 366–420.
Warner, Andrew, 2002. Institutions, geography, regions, countries and the mobility bias. Working Paper 91. Center for International Development at Harvard University.
Weidmann, Nils B., Rød, Jan Ketil, Cederman, Lars-Erik, 2010. Representing ethnic groups in space: a new dataset. J. Peace Res. 47, 491–499.