Proof of Proposition 1 Due to A.1, labour moves from unproductive to productive regions. Formally, Equation (11) implies that an agent with human capital h_j migrates if $w_G^{1-\theta}(h_j - \varphi)/H_G^\theta \geq w_B^{1-\theta}h_j/H_B^\theta$, where φ captures migration costs. This identifies a human capital threshold h_m such that agent j migrates if and only if $h_j \geq h_m$. By exploiting the wage equation in (6) and the equilibrium condition (9), threshold h_m can be implicitly expressed as:

$$h_m \cdot \left[1 - \left(\frac{w_B}{w_G} \right)^{1-\theta} \left(\frac{H_G}{H_B} \right)^\theta \right] = \varphi \; . \tag{Ap.1}$$

To pin down the equilibrium, note that the aggregate resource constraint is given by:

$$p \cdot H_G + (1-p) \cdot H_B = \underline{H}. \tag{Ap.2}$$

$\underline{H} = p \cdot \underline{H}_G + (1-p) \cdot \underline{H}_B$ is the country's human capital stock. Human capital employed in a productive region is equal to the initial amount $\underline{H}_G$ plus a share $1/p$ of the total human capital of migrants, namely:

$$H_G = \underline{H}_G + \frac{1-p}{p} \int_{h_m}^{+\infty} h \cdot (\mu_B \underline{h}^{\mu_B} h^{-\mu_B-1}) \cdot dh = \underline{H}_G + \underline{H}_B \frac{1-p}{p} \left(\frac{\underline{h}}{h_m} \right)^{\mu_B-1}. \tag{Ap.3}$$

By replacing (Ap.2) and (Ap.3) into (Ap.1) we find that the equilibrium is reached when:

$$\frac{\varphi}{\underline{h}} \left(\frac{p}{1-p} \right)^{\frac{1}{\mu-1}} \left(\frac{H_G - \underline{H}_G}{\underline{H}_B} \right)^{\frac{1}{\mu-1}} = 1 - \left(\frac{A_B}{A_G} \right)^{\frac{1-\theta}{1-\delta}} \left[\frac{H_G(1-p)}{\underline{H} - pH_G} \right]^{\frac{\beta(1-\theta)+\theta(1-\delta)}{1-\delta}}, \tag{Ap.4}$$

As Figure F.1 below shows, the left hand side of (Ap.4) increases in H_G, the right hand side decreases in H_G. The right hand side intersects the horizontal axis at the full mobility allocation H_G^{free}.
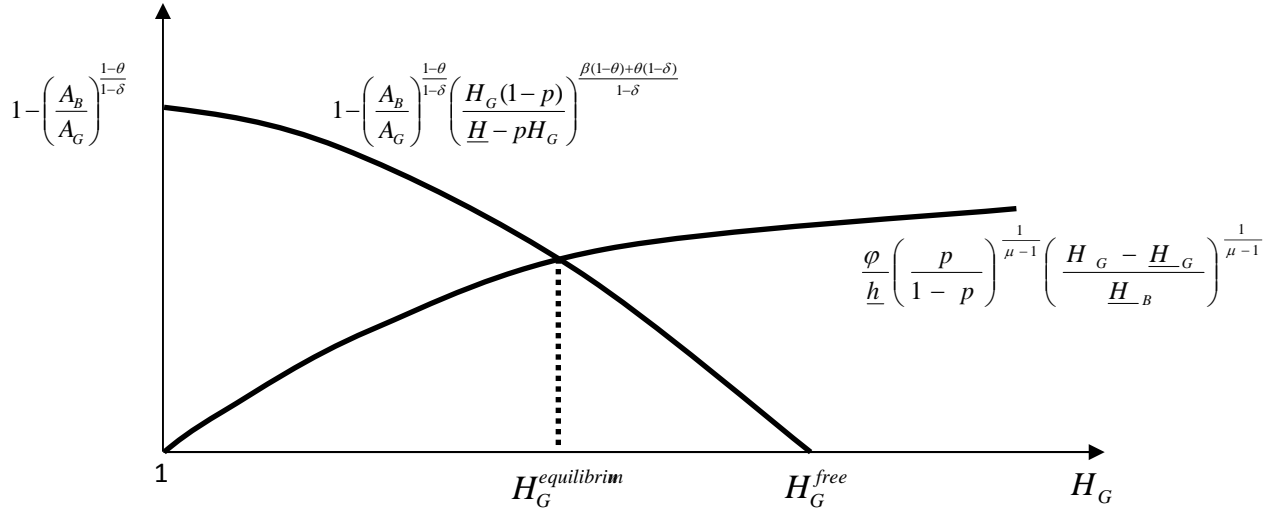
Figure F.1: *Equilibrium with mobility costs (i.e. $\varphi > 0$)*

It is immediate to see that:

$$H_G^{free} = \frac{A_G^{\frac{1-\theta}{\beta(1-\theta)+\theta(1-\delta)}}}{E\left[A^{\frac{1-\theta}{\beta(1-\theta)+\theta(1-\delta)}}\right]} \cdot \underline{H} , \qquad (Ap.5)$$

where $E\left[A^{\frac{1-\theta}{\beta(1-\theta)+\theta(1-\delta)}}\right] = pA_G^{\frac{1-\theta}{\beta(1-\theta)+\theta(1-\delta)}} + (1-p)A_B^{\frac{1-\theta}{\beta(1-\theta)+\theta(1-\delta)}}$ is the average of $A^{\frac{1-\theta}{\beta(1-\theta)+\theta(1-\delta)}}$.

The equilibrium $H_G^{equilibrium}$ is unique and is above the initial endowment $\underline{H}_G$. For given mobility costs $\varphi > 0$ and for given country level endowment $\underline{H}$, the equilibrium $\tilde{H}_G^{equilibrium}$ increases as $\underline{H}_G$ goes up. In Figure F1, such an increase in $\underline{H}_G$ (and the contextual decrease in $\underline{H}_B$), shift the upward sloping curve down, increasing the equilibrium value of H_G (the downward sloping curve stays constant).

In the full mobility equilibrium $\varphi \to 0$, we have that factor allocation fulfils:

$$H_i = \frac{A_i^{\frac{1-\theta}{\beta(1-\theta)+\theta(1-\delta)}}}{E\left[A^{\frac{1-\theta}{\beta(1-\theta)+\theta(1-\delta)}}\right]} \cdot \underline{H} \quad K_i = \frac{A_i^{\frac{\beta(1-\theta)+(1-\delta)}{(1-\delta)[\beta(1-\theta)+\theta(1-\delta)]}}}{E\left[A^{\frac{\beta(1-\theta)+(1-\delta)}{(1-\delta)[\beta(1-\theta)+\theta(1-\delta)]}}\right]} \cdot K \cdot \qquad (Ap.6)$$

55

Using Equation (Ap.6) one can find that national output is equal to:

$$Y = \hat{A}\left(H^{E}\right)^{1-\alpha-\beta-\delta}\left(H^{W}\right)^{\alpha}K^{\delta}T^{\beta}. \tag{Ap.7}$$

Here $\hat{A}$ is a function $\hat{A}(\beta,\delta,\theta,A_{G},A_{B},p)$ of exogenous parameters.

Proof of Proposition 2 In the presence of externalities and when $\psi \geq 1$, the equilibrium condition (Ap.1) can be written as:

$$h_{m}\cdot\left[1-\left(\frac{A_{B}}{A_{G}}\right)^{\frac{1-\theta}{1-\delta}}\left(\frac{L_{G}}{L_{B}}\right)^{\gamma(\psi-1)\frac{1-\theta}{1-\delta}}\left(\frac{H_{G}}{H_{B}}\right)^{\frac{(\beta-\gamma\psi)(1-\theta)+(1-\delta)\theta}{1-\delta}}\right]=\varphi. \tag{Ap.8}$$

Using Equation (Ap.3), it is immediate to express threshold h_{m} as a function of H_{G} and thus recover:

$$\frac{L_{G}}{L_{B}}=\frac{1+\dfrac{p}{1-p}\cdot\left(\dfrac{H_{G}-\underline{H}_{G}}{\underline{H}_{B}}\cdot\dfrac{p}{1-p}\right)^{\frac{\mu_{B}}{\mu_{B}-1}}}{1-\left(\dfrac{H_{G}-\underline{H}_{G}}{\underline{H}_{B}}\cdot\dfrac{p}{1-p}\right)^{\frac{\mu_{B}}{\mu_{B}-1}}}. \tag{Ap.9}$$

Under full mobility ($\varphi = 0$), using (Ap.2) one finds that the equilibrium is determined by the condition:

$$\left(\frac{A_{G}}{A_{B}}\right)^{\frac{1-\theta}{1-\delta}}\left[\frac{1-\left(\dfrac{H_{G}-\underline{H}_{G}}{\underline{H}_{B}}\cdot\dfrac{p}{1-p}\right)^{\frac{\mu_{B}}{\mu_{B}-1}}}{1+\dfrac{p}{1-p}\cdot\left(\dfrac{H_{G}-\underline{H}_{G}}{\underline{H}_{B}}\cdot\dfrac{p}{1-p}\right)^{\frac{\mu_{B}}{\mu_{B}-1}}}\right]^{\gamma(\psi-1)\frac{1-\theta}{1-\delta}}=\left[\frac{(1-p)H_{G}}{\underline{H}-pH_{G}}\right]^{\frac{(\beta-\gamma\psi)(1-\theta)+(1-\delta)\theta}{1-\delta}}. \tag{Ap.10}$$

The left hand side is decreasing in H_{G}. If $(\beta - \psi\gamma)(1- \theta) + \theta(1- \delta) > 0$, the right hand side - which captures the cost of migrating to productive regions, increases in H_{G}. As a result, when $(\beta - \psi\gamma)(1- \theta) + \theta(1- \delta) > 0$

even under full mobility in the stable equilibrium there is no universal migration to productive regions. Indeed, if all human capital moves to productive regions, then $H_G = \underline{H}/p$ and the right hand side of (Ap.10) blows to infinity. Full migration is not an equilibrium. No migration is not an equilibrium either, as in this case A.1 implies that (Ap.10) cannot hold. When $\psi = 1$ (and $\varphi = 0$) the equilibrium has:

$$H_i = \frac{A_i^{\frac{1-\theta}{(\beta-\gamma)(1-\theta)+\theta(1-\delta)}}}{E\left[A^{\frac{1-\theta}{(\beta-\gamma)(1-\theta)+\theta(1-\delta)}}\right]} \cdot \underline{H}. \tag{Ap.11}$$

With imperfect mobility $\varphi = 0$, the equilibrium fulfils the condition:

$$\frac{\varphi}{\underline{h}}\left(\frac{p}{1-p}\right)^{\frac{1}{\mu-1}}\left(\frac{H_G-\underline{H}_G}{\underline{H}_B}\right)^{\frac{1}{\mu-1}} = 1 - \left(\frac{A_B}{A_G}\right)^{\frac{1-\theta}{1-\delta}}\left[\frac{1+\dfrac{p}{1-p}\left(\dfrac{H_G-\underline{H}_G}{\underline{H}_B}\cdot\dfrac{p}{1-p}\right)^{\frac{\mu_B}{\mu_B-1}}}{1-\cdot\left(\dfrac{H_G-\underline{H}_G}{\underline{H}_B}\cdot\dfrac{p}{1-p}\right)^{\frac{\mu_B}{\mu_B-1}}}\right]^{\gamma(\psi-1)\frac{1-\theta}{1-\delta}}\left[\frac{(1-p)H_G}{\underline{H}-pH_G}\right]^{\frac{(\beta-\gamma\psi)(1-\theta)+(1-\delta)\theta}{1-\delta}}.$$

The graphical representation of the above equilibrium condition is very similar to the one displayed in Figure F.1. The only difference is that now an increase in $\underline{H}_G$ (holding $\underline{H}$ constant) does not only shift down the left hand side above, but it also shifts up the right hand side, thereby increasing H_G even more than in the no-externalities case.

Proof of Proposition 3 Given $\mu_G = \mu_B = \mu$ he migration threshold h_m can be rewritten as:

$$\int_{\underline{h}}^{h_m} \mu_B \underline{h}^\mu h^{-\mu} dh = \underline{H}\left[1-\left(\frac{h}{h_m}\right)^{\mu-1}\right] = H_B, \tag{Ap.12}$$

and H_B is identified by (Ap.11). Define $\hat{A}_i = \dfrac{A_i^{\frac{1-\theta}{(\beta-\gamma)(1-\theta)+\theta(1-\delta)}}}{E\left[A^{\frac{1-\theta}{(\beta-\gamma)(1-\theta)+\theta(1-\delta)}}\right]}$. It is then easy to find that:

$$h_m = \underline{h}\left[1 - \hat{A}_B\right]^{\frac{1}{\mu-1}}, \tag{Ap.13}$$

The higher is A_B, the more skilled must a labourer be in order for him to find migration profitable (i.e. the higher is h_m). Suppose than that:

$$\left(\frac{1-\alpha-\beta-\delta}{1-\beta-\delta}\right) < 1 - \frac{\hat{A}_B}{\hat{A}_G}, \tag{Ap.14}$$

which ensures that the least productive migrants become workers. Then the human capital thresholds above which agents become entrepreneurs in the productive and unproductive regions are equal to:

$$h_G^E = \underline{h} \cdot \left[p\left(\frac{1-\alpha-\beta-\delta}{1-\beta-\delta}\right) \cdot \hat{A}_G \right]^{-\frac{1}{\mu-1}}, h_B^E = \underline{h} \cdot \left[1 - \left(\frac{\alpha}{1-\beta-\delta}\right) \cdot \hat{A}_B \right]^{-\frac{1}{\mu-1}}, \tag{Ap.15}$$

where $h_G^E > h_B^E$, so that entrepreneurs are more skilled in productive regions. Defining f_G and f_B the number of firms in the productive and unproductive regions, respectively, and by l_G and l_B the respective workforces, one finds that in equilibrium we have that:

$$f_G = p^{\frac{1}{\mu-1}} \cdot \left[\left(\frac{1-\alpha-\beta-\delta}{1-\beta-\delta}\right)\hat{A}_G\right]^{\frac{\mu}{\mu-1}}, f_B = \left[1 - \left(\frac{\alpha}{1-\beta-\delta}\right)\hat{A}_B\right]^{\frac{\mu}{\mu-1}} - \left(1-\hat{A}_B\right)^{\frac{\mu}{\mu-1}}, \tag{Ap.16}$$

where f_i is the integral of the density of skills above h_i^E according to the distributions of Figure 2. The size of the workforce in the two regions is equal to:

$$l_G = \int_{\underline{h}}^{h_m} \mu \underline{h}^{\mu} h^{-\mu-1} dh + \frac{1}{p} \int_{h_m}^{h_G^E} \mu \underline{h}^{\mu} h^{-\mu-1} dh = 1 + \frac{1-p}{p}(1-\hat{A}_B)^{\frac{\mu}{\mu-1}} - p^{\frac{1}{\mu-1}}\left(\frac{1-\alpha-\beta-\delta}{1-\beta-\delta}\hat{A}_B\right)^{\frac{\mu}{\mu-1}}$$

$$l_B = \int_{\underline{h}}^{h_B^E} \mu \underline{h}^{\mu} h^{-\mu-1} dh = 1 - \left[1-\left(\frac{\alpha}{1-\beta-\delta}\right)\hat{A}_B\right]^{\frac{\mu}{\mu-1}}.$$

As a result, we have that $l_G/f_G > l_B/f_B$ when the following condition holds:

$$\frac{p+(1-p)(1-\hat{A}_B)^{\frac{\mu}{\mu-1}}}{\left(\frac{1-\alpha-\beta-\delta}{1-\beta-\delta}\right)^{\frac{\mu}{\mu-1}}\left[1-(1-p)(1-\hat{A}_B)^{\frac{\mu}{\mu-1}}\right]} \geq \frac{1-(1-\hat{A}_B)^{\frac{\mu}{\mu-1}}}{\left(1-\frac{\alpha}{1-\beta-\delta}\hat{A}_B\right)^{\frac{\mu}{\mu-1}}-(1-\hat{A}_B)^{\frac{\mu}{\mu-1}}}$$

The numerator/denominator of the left hand side is larger/smaller than the numerator and denominator of the right hand side if and only if:

$$\frac{1}{p(\hat{A}_G/\hat{A}_B)+(1-p)} \leq 1-\left(\frac{1-p}{2-p}\right)^{\frac{\mu-1}{\mu}}$$

$$\frac{1}{p(\hat{A}_G/\hat{A}_B)+(1-p)} \geq \frac{1-\left(\frac{\alpha}{1-\beta-\delta}\right)^{\mu-1}}{1-\left(\frac{\alpha}{1-\beta-\delta}\right)^{\mu}}$$

$\quad$ (Ap.17)

The first condition is met if $\hat{A}_G/\hat{A}_B$ is sufficiently large (i.e. larger than a certain value z_1). The second condition is met if $\hat{A}_G/\hat{A}_B$ is sufficiently low (i.e. smaller than a certain value z_2). Finally, (Ap.17) can be met (i.e. $z_1 < z_2$) when p is large.

Table 1: Descriptive Statistics

The table reports descriptive statistics for the variables in the paper. We report the total number of observations, the number of countries and medians for: (1) the number of regions with non-missing data, (2) the country average, (3) the within-country range, (4) the within-country standard deviation, and (5) the coefficient of variation for the variable in levels (rather than in logs). We also report the adjusted R squared from an univariate regression of each variable in the table on country dummies. All variables are described in Table 2.

Panel A: Regional GDP, Education, Geography, Institutions, Infrastructure, and Culture

	Number of Regions	Number of Countries	Medians for:						
			Regions per country	Mean	Minimum	Maximum	Within-country Range	Within-country std deviation	Coefficient of Variation for Variable in Levels
Ln(GDP per capita)	1,537	107	11	8.69	8.07	9.54	1.03	0.30	0.33
Years of Education	1,489	106	12	6.58	5.34	8.70	2.34	0.73	0.92
Temperature	1,568	110	12	16.84	10.23	21.13	4.47	1.45	0.09
Inverse Distance to Coast	1,569	110	12	0.90	0.80	0.99	0.13	0.05	0.05
Ln(Oil)	1,569	110	12	0.00	0.00	0.00	0.00	0.00	0.00
Informal Payments	361	76	4	1.02	0.40	1.60	0.94	0.45	0.59
ln(Tax Days)	270	58	5	1.29	1.06	1.51	0.36	0.19	0.18
Ln(Days without electricity)	222	75	2	3.03	2.73	3.37	0.54	0.36	0.32
Security costs	373	79	4	0.91	0.39	1.34	0.72	0.34	0.42
Access to land	519	81	5	0.15	0.04	0.27	0.21	0.09	0.40
Access to finance	536	82	5	0.28	0.14	0.47	0.29	0.12	0.24
Government Predictability	386	75	4	0.46	0.34	0.61	0.24	0.10	0.20
Doing Business Percentile Rank	180	19	6	0.40	0.21	0.49	0.22	0.11	0.31
Ln(Power line density)	1,569	110	12	1.34	0.00	2.53	1.87	0.61	0.61
Ln(Travel time)	1,569	110	12	5.28	4.21	6.00	1.82	0.54	0.46
Trust in others	745	69	9	0.23	0.12	0.38	0.22	0.07	0.35
Civic Values	683	75	8	2.23	1.71	3.12	1.08	0.48	0.19
Ln(Number of ethnic groups)	1,568	110	12	0.98	0.00	1.79	1.39	0.50	0.46
Probability of same language	1,545	109	12	0.67	0.28	0.79	0.26	0.09	0.21

Table 1: Descriptive Statistics (continued)

Panel B: Enterprise Survey and Census Data

	Number of Regions	Number of Countries	Regions per country	Medians for:			Within-country Range	Within-country std deviation	Coefficient of Variation for Variable in Levels
				Mean	Minimum	Maximum			
Ln(Establishments / Population)	984	65	12	-4.89	-5.45	-4.06	1.17	0.37	0.37
Ln(Employees / Establishments)	1,068	69	12	2.07	1.69	2.39	0.80	0.20	0.19
Ln(Employees / Population)	1,056	69	12	-2.66	-3.38	-1.80	1.58	0.43	0.41
Ln(Employees Big Firms / Employees)	540	31	13	-1.45	-2.17	-0.78	1.13	0.33	0.27
Ln(Sales / Employee)	549	82	5	10.21	9.79	10.59	0.79	0.35	1.22
Ln([Sales - Raw Materials]/Employee)	359	70	4	9.53	9.24	9.87	0.69	0.37	1.21
Ln(Wages / Employee)	515	77	5	8.28	8.00	8.66	0.62	0.25	1.79
Ln(Employees)	549	82	5	3.25	2.72	3.71	0.82	0.35	1.46
Ln(Expenditure on energy / Employee)	326	66	4	6.10	5.51	6.36	0.60	0.30	1.22
Ln(Property, plant and equipment / Employee)	205	41	4	8.72	8.37	9.37	0.99	0.47	1.26
Years of Education of Workers	507	74	5	9.97	8.66	10.80	2.25	0.93	3.06
Years of Education of Managers	195	38	4	14.90	14.24	15.36	1.34	0.62	0.89

Table 2 – Definitions and sources for the variables used in the paper

This table provides the names, definitions and sources of all the variables used in the tables of the paper.

Variable	Description	Sources and links
I. GDP per capita, population, employment and human capital		
Ln(GDP per capita)	The logarithm of Gross Domestic Product per capita in PPP constant 2005 international dollars in the region in 2005. Data on regional GDP is available for all countries except 20. For those 20 countries, we approximate GDP using data on income (6 countries), expenditure (8 countries), wages (3 countries), gross value added (2 countries), and consumption, investment and government expenditure (1 country). For each country, we scale regional GDP per capita values so that their population-weighted sum equals the World Development Indicators (WDI) value of Gross Domestic Product in PPP constant 2005 international dollars. Similarly, for each country, we adjust the regional population values so that their sum equals the country-level analog in WDI. For years with missing regional GDP per capita data, we interpolate using all available data for the period 1990-2008. When interpolating GDP values is not possible, we use the regional distribution of the closest year with regional GDP data. Population data for years without census data is interpolated and extrapolated from the available census data for the period 1990-2008. At the country level, we calculate this variable as the population-weighted average of regional GDP.	*Regional GDP*: See online appendix "Appendix GDP Sources". *Regional population*: Thomas Brinkhoff: City Population, http://www.citypopulation.de/ *Country-level GDP per capita and PPP exchange rates*: World Bank, (2010). Data retrieved on March 2, 2010, from World Development Indicators Online (WDI) database, http://go.worldbank.org/6HAYAHG8H0
Ln(Population)	The logarithm of the number of inhabitants in the region in 2005. Population data for years without census data is interpolated and extrapolated from the available census data for the period 1990-2008. For each country, we adjust the regional populations so that the sum of regional populations equals the country-level analog in the World Development Indicators (WDI). At the country level, we calculate this variable following the same methodology but using country boundaries.	*Regional population*: Thomas Brinkhoff: City Population, http://www.citypopulation.de/ *Regional spherical*: Collins-Bartholomew World Digital Map, http://www.bartholomewmaps.com/data.asp?pid=5.
Ln(Employment)	The number of manufacturing and service employees working in the establishments in the region. The data is for the year 2005 or the closest available. At the country level, we calculate this variable as the product of the total population and the employment ratio for the population 15 years and older.	See online appendix "Appendix on Economic Census Sources". Development Indicators Online (WDI) database, http://go.worldbank.org/6HAYAHG8H0
Years of education	The average years of schooling from primary school onwards for the population aged 15 years or older. Data for China and Georgia is for the population 6 years and older. We use the most recent information available for the period 1990-2006. To make levels of educational attainment comparable across countries, we translate educational statistics into the International Standard Classification of Education (ISCED) standard and use UNESCO data on the duration of school levels in each country for the year for which we have educational attainment data. Eurostat aggregates data for ISCED levels 0-2 and we assign such observations an ISCED level 1. Following Barro and Lee (1993): (1) we assign zero years of schooling to ISCED level 0 (i.e., pre-primary); (2) we assign zero years of *additional* schooling to (a) ISCED level 4 (i.e., vocational), and (b) ISCED level 6 (i.e. post-graduate); and (3) we assign 4 years of additional schooling to ISCED level 5 (i.e. graduate). Since regional data is not available for all countries, unlike Barro and Lee (1993), we assign zero years of *additional* schooling: (a) to all incomplete levels; and (b) to ISCED level 2 (i.e. lower secondary). Thus, the average years of schooling in a region is calculated as: (1) the product of the fraction of people whose highest attainment level is ISCED 1 or 2 and the duration of ISCED 1; plus (2) the product of the fraction of people whose highest attainment level is ISCED 3 or 4 and the cumulative duration of ISCED 3; plus (3) the product of the fraction of people whose highest attainment level is ISCED 5 or 6 and the sum of the cumulative duration of ISCED 3 plus 4 years. At the country level, we calculate this variable as the population-weighted average of the regional values.	See online appendix "Appendix on Education Sources". Links to online data: http://epdc.org/ http://epp.eurostat.ec.europa.eu/portal/page/portal/region_cities/introduction https://international.ipums.org/international/index.html http://stats.uis.unesco.org/unesco/TableViewer/document.aspx?ReportId=143&IF_Language=eng.
II. Climate, geography and natural resources		
Temperature	Average temperature during the period 1950-2000 in degrees Celsius. To produce the regional and national numbers, we create equal area projections using the Collins-Bartholomew World Digital Map and the temperature raster in ArcGIS. For each region, we sum the temperatures of all cells in that region and divide by the number of cells in that region. At the country level, we calculate this variable following the same methodology but using country boundaries.	Climate: Hijmans, R. et al. (2005), http://www.worldclim.org/ Collins-Bartholomew World Digital Map, http://www.bartholomewmaps.com/data.asp?pid=5

Variable	Description	Sources and links
Inverse distance to coast	The ratio of one over one plus the region's average distance to the nearest coastline in thousands of kilometers. To calculate each region's average distance to the nearest coastline we create an equal distance projection of the Collins-Bartholomew World Digital Map and a map of the coastlines. Using these two maps we create a raster with the distance to the nearest coastline of each cell in a given region. Finally, to get the average distance to the nearest coastline, we sum up the distance to the nearest coastline of all cells within each region and divide that sum by the number of cells in the region. At the country level, we calculate this variable following the same methodology but using country boundaries.	Collins-Bartholomew World Digital Map, http://www.bartholomewmaps.com/data.asp?pid=5
Ln(Oil)	Logarithm of one plus the estimated per capita volume of cumulative oil production and reserves by region, in millions of barrels of oil. To produce the regional measure, we load the oil map of the World Petroleum Assessment and the Collins-Bartholomew World Digital map onto ArcGIS. On-shore estimated oil in each assessment unit was allocated to the regions based on the fraction of assessment unit area covered by each region. Off-shore assessment units are not included. The World Petroleum Assessment map includes all oil fields in the world except those in the United States of America. Data for the United States is calculated using the national-level information on cumulative production and estimated reserves, available from the World Petroleum Assessment 2000 (USGS), and the United States' regional production and estimated reserves for the year 2000 from the U.S. Energy Information Administration (USEIA). The national level data for this variable is calculated following the same methodology outlined but using the data on national boundaries. The national level numbers for the U.S. are those available from the World Petroleum Assessment.	http://energy.cr.usgs.gov/oilgas/wep/products/dds60/export.htm. http://tonto.eia.doe.gov/dnav/pet/pet_crd_crpdn_adc_mbbl_a.htm. http://www.bartholomewmaps.com/data.asp?pid=5
III. Institutions		
Informal payments	The average percentage of sales spent on informal payments made to public officials to "get things done" with regard to customs, taxes, licenses, regulations, services, etc, as reported by the respondents in the region. The country-level analog of this variable is the arithmetic average of the regions in the country. Data is from the most recent year available, ranging from 2002 through 2009.	World Bank's Enterprise Surveys. https://www.enterprisesurveys.org/
Ln(Tax days)	The logarithm of one plus the average number of days spent in mandatory meetings and inspections with tax authority officials in the past year as reported by respondents in the region. The country-level analog of this variable is the arithmetic average of the regions in the country. Data is for the most recent year available, ranging from 2002 through 2009.	World Bank's Enterprise Surveys. https://www.enterprisesurveys.org/
Ln(Days without electricity)	The logarithm of one plus the average number of days without electricity in the past year as reported by the respondents in the region. The country-level analog of this variable is the arithmetic average of the regions in the country. Data is for the most recent year available, ranging from 2002 through 2009.	World Bank's Enterprise Surveys. https://www.enterprisesurveys.org/
Security costs	The average costs of security (i.e., equipment, personnel, or professional security services) as a percentage of sales as reported by the respondents in the region. The country-level analog of this variable is the arithmetic average of the regions in the country. Data is for the most recent year available, ranging from 2002 through 2009.	World Bank's Enterprise Surveys. https://www.enterprisesurveys.org/
Access to land	The percentage of respondents in the region who think that access to land is a moderate, major, or very severe obstacle to business. The country-level analog of this variable is the arithmetic average of the regions in the country. Data is for the most recent year available, ranging from 2002 through 2009.	World Bank's Enterprise Surveys. https://www.enterprisesurveys.org/
Access to finance	The percentage of respondents in the region who think that access to financing is a moderate, major, or very severe obstacle to business. The country-level analog of this variable is the arithmetic average of the regions in each respective country. Data is for the most recent year available, ranging from 2002 through 2009.	World Bank's Enterprise Surveys. https://www.enterprisesurveys.org/
Government predictability	The percentage of respondents in the region who tend to agree, agree in most cases, or fully agree that their government officials' interpretation of regulations are consistent and predictable. The country-level analog of this variable is the arithmetic average of the regions in the country. Data is for the most recent year available, ranging from 2002 through 2009.	World Bank's Enterprise Surveys. https://www.enterprisesurveys.org/
Doing Business Percentile Rank	The average of the percentile ranks in each of the following five areas: (1) starting a business; (2) dealing with construction permits; (3) registering property; (4) enforcing contracts; and (5) paying taxes. Higher values indicate more burdensome regulation. Data is for the most recent year available, ranging from 2007 through 2010.	Word Bank's Doing Business Subnational Reports. http://doingbusiness.org/Reports/Subnational-Reports/
Autocracy	This variable classifies regimes based on their degree of autocracy. Democracies are coded as 0, bureaucracies (dictatorships with a legislature) are coded as 1 and autocracies (dictatorship without a legislature) are coded as 2. Transition years are coded as the regime that emerges afterwards. This variable ranges from zero to two where higher values equal a higher degree of autocracy. This	Alvarez et al. (2000).

Variable	Description	Sources and links
	variable is measured as the average from 1960 through 1990.	
Executive Constraints	A measure of the extent of institutionalized constraints on the decision making powers of chief executives. The variable takes seven different values: (1) Unlimited authority (there are no regular limitations on the executive's actions, as distinct from irregular limitations such as the threat or actuality of coups and assassinations); (2) Intermediate category; (3) Slight to moderate limitation on executive authority (there are some real but limited restraints on the executive); (4) Intermediate category; (5) Substantial limitations on executive authority (the executive has more effective authority than any accountability group but is subject to substantial constraints by them); (6) Intermediate category; (7) Executive parity or subordination (accountability groups have effective authority equal to or greater than the executive in most areas of activity). This variable ranges from one to seven where higher values equal a greater extent of institutionalized constraints on the power of chief executives. This variable is calculated as the average from 1960 through 2000.	Jaggers and Marshall (2000).
Expropriation Risk	Risk of "outright confiscation and forced nationalization" of property. This variable ranges from zero to ten where higher values are equals a lower probability of expropriation. This variable is calculated as the average from 1982 through 1997.	International Country Risk Guide at http://www.countrydata.com/datasets/.
Proportional Representation	This variable is equal to one for each year in which candidates were elected using a proportional representation system; equals zero otherwise. Proportional representation means that candidates are elected based on the percentage of votes received by their party. This variable is measured as the average from 1975 through 2000.	Beck et al. (2001).
Corruption	The average score of the Transparency International index of corruption perception in 2005. The index provides a measure of the extent to which corruption is perceived to exist in the public and political sectors. The index focuses on corruption in the public sector and defines corruption as the abuse of public office for private gain. It is based on assessments by experts and opinion surveys. The index ranges between 0 (highly corrupt) and 10 (highly clean).	www.transparency.org

IV. Infrastructure

Variable	Description	Sources and links
Ln(Power line density)	The logarithm of one plus the length in kilometers of power lines per $10km^2$ in the year 1997. To produce the regional numbers, we load the power line map from the US Geological Survey and the Collins-Bartholomew World Digital Map onto ArcGIS. We take the ratio of total length of the power lines in the region to the spherical area of that region. At the country level, we calculate this variable following the same methodology but using country boundaries.	US Geological Survey Global GIS database, accessed through Harvard University's Geospatial Library. Collins-Bartholomew World Digital Map, http://www.bartholomewmaps.com/data.asp?pid=5
Ln(Travel time)	The logarithm of the average estimated travel time in minutes from each cell in a region to the nearest city of 50,000 or more people in the year 2000. We use the raster from the Global Environmental Monitoring Unit and the Collins-Bartholomew World Digital Map. For each region, we sum the travel time from all its cells and divide by the number of cells in that region. At the country level, we calculate this variable following the same methodology but using country boundaries.	Global Environment Monitoring Unit, http://bioval.jrc.ec.europa.eu/products/gam/index.htm Collins-Bartholomew World Digital Map, http://www.bartholomewmaps.com/data.asp?pid=5

V. Culture

Variable	Description	Sources and links
Trust in others	The percentage of respondents in the region who believe that most people can generally be trusted. The country-level analog of this variable is the arithmetic average of the regions in the country. Data is for the most recent available year, ranging from 1980 through 2005.	World Values Survey, http://www.worldvaluessurvey.org/
Civic values	The average of the value of the answers of respondents in the region about the degree of justifiability of the following four behaviors: (1) Claiming government benefits to which you are not entitled; (2) Avoiding a fare on public transport; (3) Cheating on taxes if you have a chance; and (4) Someone accepting a bribe in the course of their duties. For each question, possible answers range from 1 (never justifiable) to 10 (always justifiable). We only include observations with non-missing data for at least two of the four questions. The country-level analog of this variable is the arithmetic average of the regions in the country. Data is for the most recent available year, ranging from 1980 through 2005.	World Values Survey, http://www.worldvaluessurvey.org/
Ln(Number of ethnic groups)	The logarithm of the number of ethnic groups that inhabited the region in the year 1964. The country-level analog of this variable is constructed using country boundaries.	Weidmann et al., 2010, http://www.icr.ethz.ch/research/greg

Variable	Description	Sources and links
Probability of same language	The probability that two randomly chosen people, one from the corresponding region and one from the rest of the country, share the same mother tongue in the year 2004. Where language areas do not overlap with our regions, we compute the number of people speaking a language in a region by weighing the total number of people in a language area by the fraction of the region's surface covered by that language area. We compute the probability of same language separately for each language in a region and then calculate the surface-weighted average of the different languages in a region. The country-level analog of this variable is calculated as the population-weighted average of the regional values.	World Language Mapping System, http://www.gmi.org/wlms/

	VI. Enterprise Survey Data	
Ln(Sales / Employee)	The logarithm of the quotient of total annual revenue (in current USD) over the total number of employees in each establishment. Data is for the most recent available year, ranging from 2002 through 2009.	World Bank's Enterprise Surveys. https://www.enterprisesurveys.org/
Ln(Sales –Raw Materials / Employee)	The logarithm of the quotient of total annual revenue minus expenditure on raw materials l(in current USD) over the total number of employees in each establishment. Data is for the most recent available year, ranging from 2002 through 2009.	World Bank's Enterprise Surveys. https://www.enterprisesurveys.org/
Ln(Wages / Employee)	The logarithm of the quotient of total cost of labor (in current USD) the total number of employees in each establishment. Data is for the most recent available year, ranging from 2002 through 2009.	World Bank's Enterprise Surveys. https://www.enterprisesurveys.org/
Ln(Expenditure on energy / Employee)	The logarithm of the quotient of total energy and fuel costs over (in current USD) the total number of employees in each establishment. Data is for the most recent available year, ranging from 2002 through 2009.	World Bank's Enterprise Surveys. https://www.enterprisesurveys.org/
Years of Education of workers	The number of years of schooling from primary school onwards of the average non-management employee in each establishment. To compute this variable, we use the same assumptions and follow the same procedure as used for the previously described years of schooling variable at the regional level. Data is for the most recent available year, ranging from 2002 through 2009.	World Bank's Enterprise Surveys. https://www.enterprisesurveys.org/
Years of Education of manager	The number of years of schooling from primary school onwards of the manager of the establishment. To compute this variable, we use the same assumptions and follow the same procedure as used for the previously described years of schooling variable at the regional level. Data is for the most recent available year, ranging from 2002 through 2009.	World Bank's Enterprise Surveys. https://www.enterprisesurveys.org/
Ln(Property, plant, and equipment / Employee)	The logarithm of the quotient of the book value of property, plant and equipment (in current USD) over the total number of employees in the establishment. Data is for the most recent available year, ranging from 2002 through 2009.	World Bank's Enterprise Surveys. https://www.enterprisesurveys.org/

	VII. Economic Census Data	
Ln(Establishments / Population)	The logarithm of the quotient of the number of manufacturing and service establishments in the region and the region's population. The data is for the year 2005 or the closest available. An establishment is defined as a single physical location at which business is conducted or where services or industrial operations are performed.	See online appendix "Appendix on Economic Census Sources".
Ln(Employees / Population)	The logarithm of the quotient of the number of manufacturing and service employees working in the establishments in the region over the region's population. The data is for the year 2005 or the closest available.	See online appendix "Appendix on Economic Census Sources".
Ln(Employees / Establishments)	The logarithm of the quotient of the number of manufacturing and service employees working in the establishments in the region over the number of establishments in the region. The data is for the year 2005 or the closest available.	See online appendix "Appendix on Economic Census Sources".
Ln(Employees big firms / Employees)	The logarithm of the quotient of the number of manufacturing and service employees working in "big firms" over the total number of formal sector employees working in all the establishments in the region. We define "big firms" as establishments employing over 100 employees. However, when data for the 100 employee threshold is unavailable, we use the closest available threshold to 100 employees. The data is for the year 2005 or the closest available.	See online appendix "Appendix on Economic Census Sources".

Table 3: Univariate Fixed Effects Regressions

Fixed effects regressions of the log of GDP per capita at the regional level in the year 2005.   The independent variables are proxies for: (1) geography, (2) Institutions, and (3) Infrastructure and Culture. All variables are explained in Table 2. The table reports the number of observations, the number of countries, the R^2 within, the R^2 between, and the fraction of the variance due to countries. All variables are described in Table 2.

	Observations	Countries	R^2 Within	R^2 Between
Independent Variables:				
Years of Education	1,470	104	38%	58%
Temperature	1,536	107	1%	27%
Inverse Distance to Coast	1,537	107	4%	13%
Ln(Oil)	1,537	107	2%	4%
Informal Payments	350	74	0%	21%
ln(Tax Days)	263	56	0%	20%
Ln(Days without electricity)	219	73	2%	6%
Security costs	362	77	0%	7%
Access to land	507	79	0%	15%
Access to finance	524	80	1%	8%
Government Predictability	380	73	1%	0%
Doing Business Percentile Rank	176	18	2%	13%
Ln(Power line density)	1,537	107	5%	36%
Ln(Travel time)	1,537	107	7%	15%
Trust in others	739	68	0%	18%
Ln(Number of ethnic groups)	1,536	107	5%	17%
Probability of same language	1,518	106	1%	26%

Table 4: GDP per capita and Geography

Ordinary least squares and fixed effects regressions of the log of GDP per capita. The dependent variable is the logarithm of the 2005 level of GDP per capita at the country level in Panel A and at the logarithm of regional GDP per capita in Panel B. The independent variables are (1) temperature, (2) inverse distance to coast, (3) the logarithm of per capita oil production and reserves, (4) the average years of education, (5) the logarithm of population, and (6) the logorithm of the number of employees. Robust standard errors are shown in parentheses. All variables are described in Table 2.

Panel A: Dependent Variable is Logarithm National GDP per capita

	(1)	(2)	(3)
Temperature	-0.0914[a]	-0.0189[c]	-0.0190[c]
	(0.0100)	(0.0106)	(0.0106)
Inverse Distance to Coast	4.4768[a]	2.9647[a]	2.9499[a]
	(0.5266)	(0.5736)	(0.5782)
Ln(Oil)	1.2192[a]	0.9503[a]	0.9473[a]
	(0.1985)	(0.1371)	(0.1375)
Years of Education		0.2566[a]	0.2574[a]
		(0.0308)	(0.0311)
Ln(Population)		0.0684[c]	
		(0.0408)	
Ln(Employment)			0.0576
			(0.0398)
Constant	6.3251[a]	3.5761[a]	3.7959[a]
	(0.4598)	(0.9372)	(0.8977)
Observations	107	104	103
Adjusted R^2	50%	63%	63%

Note: a = significant at the 1% level, b = significant at the 5% level, and c = significant at the 10% level.

Table 4: GDP per capita and Geography (continued)

Panel B: Dependent Variable is Logarithm Regional GDP per capita

	(1)	(2)	(3)
Temperature	-0.0156[c]	-0.0140[c]	-0.0206[c]
	(0.0082)	(0.0084)	(0.0105)
Inverse Distance to Coast	1.0318[a]	0.4979[a]	0.5096[a]
	(0.2078)	(0.1438)	(0.1745)
Ln(Oil)	0.1651[a]	0.1752[a]	0.1941[a]
	(0.0477)	(0.0578)	(0.0440)
Years of Education		0.2755[a]	0.2751[a]
		(0.0171)	(0.0271)
Ln(Population)		0.0125	
		(0.0168)	
Ln(Employment)			0.0661[a]
			(0.0244)
Constant	8.0947[a]	6.3886[a]	5.9154[a]
	(0.2282)	(0.1944)	(0.2516)
Observations	1,545	1,478	833
Number of countries	107	104	49
R^2 Within	8%	42%	50%
R^2 Between	47%	60%	70%
R^2 Overall	34%	62%	70%
Fixed Effects	Yes	Yes	Yes

Note: a = significant at the 1% level, b = significant at the 5% level, and c = significant at the 10% level.

Table 5: National GDP per capita, Institutions, Infrastructure, and Culture

Ordinary least square regressions of the log of GDP per capita at the country level. All regressions include the years of education, logarithm of population, temperature, inverse distance to coast, and the logarithm of per capita oil production and reserves. In addition, regressions include measures of: (1) institutions (Panel A) and (2) infrastructure and culture (Panel B). Robust standard errors are shown in parenthesis. For comparison, the bottom panel shows the adjusted R^2 of two alternative specifications: (1) a regression with all regressors except the measure of institutions or culture; and (2) a regression with all regressors except education. All variables are described in Table 2.

| | Panel A: Institutions | | | | | | | | |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years of Education	0.2566a	0.2310a	0.1890a	0.2339a	0.2291a	0.2301a	0.2264a	0.2355a	0.1749b
	(0.0308)	(0.0344)	(0.0310)	(0.0316)	(0.0336)	(0.0350)	(0.0344)	(0.0332)	(0.0703)
Ln(Population)	0.0684c	-0.0022	0.0887	0.0428	0.0320	0.0067	0.0299	0.0611	-0.0782
	(0.0408)	(0.0494)	(0.0582)	(0.0488)	(0.0481)	(0.0519)	(0.0473)	(0.0457)	(0.1074)
Temperature	-0.0189c	-0.0105	-0.0276b	-0.0083	-0.0094	-0.0066	-0.0082	-0.0129	-0.0147
	(0.0106)	(0.0128)	(0.0128)	(0.0119)	(0.0114)	(0.0112)	(0.0110)	(0.0117)	(0.0306)
Inverse Distance to Coast	2.9647a	2.3086a	2.1692a	2.5170a	2.2652a	2.2826a	2.1892a	2.3979a	0.2385
	(0.5736)	(0.6321)	(0.7006)	(0.5698)	(0.5856)	(0.5406)	(0.5562)	(0.5616)	(2.1131)
Ln(Oil)	0.9503a	1.6367a	0.5257	1.1319a	1.1739a	1.1916a	1.1165a	1.2054b	0.5201
	(0.1371)	(0.5966)	(0.5050)	(0.3309)	(0.3219)	(0.3302)	(0.2950)	(0.4982)	(0.4921)
Informal Payments		-0.0121							
		(0.0499)							
ln(Tax Days)			-0.5497a						
			(0.1446)						
Ln(Days without electricity)				-0.1375					
				(0.0847)					
Security costs					-0.0332				
					(0.0250)				
Access to land						-0.7493			
						(0.5783)			
Access to finance							-0.5164		
							(0.4202)		
Government Predictability								0.3835	
								(0.4431)	
Doing Business Percentile Rank									0.6704
									(1.6413)
Constant	3.5761a	5.1927a	5.1619a	4.6815a	4.7382a	5.1545a	4.9498a	3.9328a	8.6509b
	(0.9372)	(1.1015)	(1.2918)	(0.9542)	(1.0046)	(0.9971)	(1.0246)	(0.9724)	(3.1636)
Observations	104	73	55	75	76	80	81	72	17
Adjusted R^2	63%	73%	76%	69%	69%	70%	70%	71%	34%
Adj. R^2 without institution	63%	73%	69%	69%	69%	69%	69%	71%	39%
Adj. R^2 without education	50%	53%	60%	49%	50%	52%	52%	50%	26%

Note: a = significant at the 1% level, b = significant at the 5% level, and c = significant at the 10% level.

Panel B: Infrastructure and Culture							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Years of Education	0.2566[a]	0.2379[a]	0.2642[a]	0.1935[a]	0.1818[a]	0.2534[a]	0.2394[a]
	(0.0308)	(0.0338)	(0.0325)	(0.0498)	(0.0538)	(0.0347)	(0.0377)
Ln(Population)	0.0684[c]	0.0688[c]	0.0653	0.1238	0.2169[b]	0.0999	0.0807[c]
	(0.0408)	(0.0414)	(0.0407)	(0.0788)	(0.1017)	(0.0640)	(0.0450)
Temperature	-0.0189[c]	-0.0145	-0.0191[c]	-0.0283[b]	-0.0434[a]	-0.0188[c]	-0.0163
	(0.0106)	(0.0109)	(0.0108)	(0.0135)	(0.0148)	(0.0107)	(0.0108)
Inverse Distance to Coast	2.9647[a]	2.7218[a]	3.0968[a]	3.6522[a]	4.3386[a]	2.7758[a]	2.7448[a]
	(0.5736)	(0.6025)	(0.6268)	(0.7902)	(1.0486)	(0.6473)	(0.5853)
Ln(Oil)	0.9503[a]	1.0157[a]	0.8737[a]	0.9902[a]	0.9751[a]	0.9538[a]	0.8792[a]
	(0.1371)	(0.1438)	(0.1467)	(0.3207)	(0.2895)	(0.1443)	(0.1657)
Ln(Power line density)		0.1480					
		(0.1099)					
Ln(Travel time)			0.0825				
			(0.0934)				
Trust in others				1.2472			
				(0.8796)			
Civic values					0.4180		
					(0.3105)		
Ln(Number of ethnic groups)						-0.0996	
						(0.1550)	
Probability of same language							0.4195
							(0.3391)
Constant	3.5761[a]	3.6383[a]	3.0050[b]	2.3962	-0.1572	3.4625[a]	3.3864[a]
	(0.9372)	(0.9251)	(1.2448)	(2.0122)	(3.2084)	(0.9289)	(0.9548)
Observations	104	104	104	67	57	104	103
Adjusted R^2	63%	63%	63%	49%	47%	63%	62%
Adj. R^2 without infrastructure or culture	63%	63%	63%	48%	45%	63%	62%
Adj. R^2 without education	50%	54%	50%	44%	42%	51%	52%

Note: a = significant at the 1% level, b = significant at the 5% level, and c = significant at the 10% level.

Table 6: Regional GDP per capita, Institutions, Infrastructure, and Culture

Ordinary least square regressions of the log of regional GDP per capita. All regressions include years of education, logarithm of population, temperature, inverse distance to coast, and the logarithm of per capita oil production and reserves. In addition, regressions include measures of: (1) institutions (Panel A) and (2) infrastructure and culture (Panel B). Robust standard errors are shown in parenthesis. For comparison, the bottom panel shows the adjusted R^2 of two alternative specifications: (1) a regression with all regressors except the measure of institutions or culture; and (2) a regression with all regressors except education. All variables are described in Table 2.

Panel A: Institutions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years of Education in the Region	0.2758[a]	0.3056[a]	0.3620[a]	0.3439[a]	0.3343[a]	0.3267[a]	0.3273[a]	0.3166[a]	0.4141[a]
	(0.0172)	(0.0298)	(0.0288)	(0.0481)	(0.0310)	(0.0218)	(0.0215)	(0.0207)	(0.0229)
Ln(Population in the Region)	0.0126	-0.0185	-0.0175	-0.0442	-0.0191	-0.0087	-0.0098	-0.0113	-0.0026
	(0.0168)	(0.0495)	(0.0536)	(0.0613)	(0.0432)	(0.0316)	(0.0312)	(0.0305)	(0.0229)
Temperature	-0.0140[c]	-0.0101	-0.0086	-0.0015	-0.0064	-0.0093	-0.0106	-0.0131	0.0016
	(0.0084)	(0.0096)	(0.0078)	(0.0122)	(0.0093)	(0.0086)	(0.0086)	(0.0081)	(0.0059)
Inverse Distance to Coast	0.4971[a]	0.4647	0.8290[c]	0.1810	0.2703	0.4054	0.5133[c]	0.4420	0.0913
	(0.1441)	(0.3293)	(0.4273)	(0.4312)	(0.3041)	(0.2636)	(0.2822)	(0.2788)	(0.3460)
Ln(Oil)	0.1752[a]	-0.0578	0.1555	-0.0584	-0.0473	-0.0224	-0.0040	-0.0170	0.1834
	(0.0578)	(0.1283)	(0.1319)	(0.2503)	(0.0862)	(0.1081)	(0.1113)	(0.0735)	(0.1160)
Informal Payments		-0.0089							
		(0.0353)							
ln(Tax Days)			-0.0479						
			(0.0630)						
Ln(Days without electricity)				0.0001					
				(0.0764)					
Security costs					-0.0004				
					(0.0060)				
Access to land						-0.1900			
						(0.1457)			
Access to finance							-0.0935		
							(0.1536)		
Government Predictability								-0.1251	
								(0.1426)	
Doing Business Percentile Rank									-0.6199[c]
									(0.3437)
Constant	6.3853[a]	6.5073[a]	5.7640[a]	6.8622[a]	6.4507[a]	6.3453[a]	6.2816[a]	6.4790[a]	6.3186[a]
	(0.1947)	(0.7043)	(0.8220)	(0.7867)	(0.5993)	(0.4664)	(0.4827)	(0.4629)	(0.4428)
Observations	1,469	338	255	216	352	387	381	368	172
Number of countries	104	73	55	72	76	77	76	72	17
R^2 Within	42%	58%	66%	59%	60%	62%	62%	63%	69%
R^2 Between	60%	64%	64%	53%	58%	60%	60%	63%	39%
R^2 Overall	62%	59%	60%	49%	53%	55%	55%	56%	51%
Within R^2 without institution	42%	57%	66%	59%	60%	62%	62%	62%	67%
Within R^2 without education	9%	11%	14%	10%	9%	6%	5%	7%	9%
Between R^2 without institution	60%	64%	63%	53%	58%	60%	60%	63%	41%
Between R^2 without education	42%	25%	20%	21%	26%	35%	39%	45%	50%
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: a = significant at the 1% level, b = significant at the 5% level, and c = significant at the 10% level.

Table 6: Regional GDP per capita, Institutions, Infrastructure, and Culture (Cont)

Panel B: Infrastructure and Culture							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Years of Education in the Region	0.2758[a]	0.2713[a]	0.2627[a]	0.3021[a]	0.2986[a]	0.2644[a]	0.2719[a]
	(0.0172)	(0.0187)	(0.0197)	(0.0286)	(0.0305)	(0.0181)	(0.0175)
Ln(Population in the Region)	0.0126	0.0101	0.0023	0.0091	0.0138	0.0170	0.0115
	(0.0168)	(0.0168)	(0.0184)	(0.0187)	(0.0193)	(0.0173)	(0.0157)
Temperature	-0.0140[c]	-0.0142[c]	-0.0166[c]	-0.0015	-0.0038	-0.0154[c]	-0.0140[c]
	(0.0084)	(0.0085)	(0.0085)	(0.0060)	(0.0056)	(0.0090)	(0.0080)
Inverse Distance to Coast	0.4971[a]	0.4872[a]	0.4626[a]	0.4750[c]	0.4093	0.4351[a]	0.5162[a]
	(0.1441)	(0.1427)	(0.1438)	(0.2590)	(0.2713)	(0.1358)	(0.1450)
Ln(Oil)	0.1752[a]	0.1793[a]	0.1864[a]	0.0534	0.0354	0.1922[a]	0.1772[a]
	(0.0578)	(0.0584)	(0.0582)	(0.0669)	(0.0572)	(0.0613)	(0.0591)
Ln(Power line density)		0.0199					
		(0.0198)					
Ln(Travel time)			-0.0456[c]				
			(0.0231)				
Trust in others				-0.0611			
				(0.0868)			
Civic values					-0.0040		
					(0.0231)		
Ln(Number of ethnic groups)						-0.0504[b]	
						(0.0249)	
Probability of same language							0.1723
							(0.2067)
Constant	6.3853[a]	6.4350[a]	6.9287[a]	6.0940[a]	6.0196[a]	6.5272[a]	6.2956[a]
	(0.1947)	(0.1928)	(0.3351)	(0.2863)	(0.3245)	(0.1679)	(0.2337)
Observations	1,469	1,469	1,469	699	635	1,468	1,445
Number of countries	104	104	104	65	70	104	103
R^2 Within	42%	42%	43%	49%	48%	42%	42%
R^2 Between	60%	60%	60%	50%	50%	60%	60%
R^2 Overall	62%	62%	61%	50%	47%	62%	62%
Within R^2 without institution	42%	42%	42%	49%	48%	42%	42%
Within R^2 without education	9%	13%	17%	10%	10%	14%	11%
Between R^2 without institution	60%	60%	60%	51%	50%	60%	59%
Between R^2 without education	42%	51%	47%	7%	17%	47%	50%
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: a = significant at the 1% level, b = significant at the 5% level, and c = significant at the 10% level.

Table 7: National GDP per capita and commonly used measures of institutions

Ordinary least square regressions of the log of GDP per capita at the country level. All regressions include the years of education, logarithm of population, temperature, inverse distance to coast, and the logarithm of per capita oil production and reserves. In addition, regressions include the following variables: (1) Autocracy; (2) Executive constraints; (3) Expropriation riks; (4) Proportional representation; and (5) Corruption. Robust standard errors are shown in parenthesis. For comparison, the bottom panel shows the adjusted R^2 of two alternative specifications: (1) a regression with all regressors except the measure of institutions or culture; and (2) a regression with all regressors except education.

	(1)	(2)	(3)	(4)	(5)	(6)
Years of Education	0.2567[a]	0.2200[a]	0.2069[a]	0.1626[a]	0.2448[a]	0.1850[a]
	(0.0308)	(0.0433)	(0.0438)	(0.0480)	(0.0363)	(0.0351)
Ln(Population)	0.0683[c]	0.0354	0.0559	-0.0356	0.0732	0.0504
	(0.0410)	(0.0487)	(0.0470)	(0.0482)	(0.0533)	(0.0370)
Temperature	-0.0189[c]	-0.0179	-0.0135	0.0024	-0.0181	-0.0100
	(0.0106)	(0.0118)	(0.0109)	(0.0106)	(0.0126)	(0.0104)
Inverse Distance to Coast	2.9646[a]	2.3421[a]	2.3853[a]	2.3974[a]	2.9603[a]	1.9906[a]
	(0.5742)	(0.7800)	(0.6050)	(0.5941)	(0.6208)	(0.5463)
Ln(Oil)	0.9503[a]	0.7877[c]	1.0708[a]	0.8965[a]	1.0720[b]	0.9928[a]
	(0.1373)	(0.4564)	(0.1729)	(0.1100)	(0.4094)	(0.2013)
Autocracy		-0.5994[a]				
		(0.2184)				
Executive Constraints			0.1633[b]			
			(0.0696)			
Expropriation Risk				0.3952[a]		
				(0.0986)		
Proportional Representation					0.3972[c]	
					(0.2328)	
Corruption						0.2130[a]
						(0.0479)
Constant	3.5771[a]	5.3781[a]	3.7896[a]	3.1830[b]	3.2958[a]	4.1183[a]
	(0.9416)	(1.3861)	(1.0059)	(1.3630)	(1.0503)	(0.8118)
Observations	103	80	101	81	97	103
Adjusted R^2	63%	67%	65%	70%	63%	69%
Adj. R^2 without institution	63%	64%	63%	63%	62%	63%
Adj. R^2 without education	50%	60%	59%	67%	52%	63%

Note: a = significant at the 1% level, b = significant at the 5% level, and c = significant at the 10% level.

Table 8: Firm level productivity

The table reports fixed effect regressions for for the following two dependent variables: (1) logarithm of sales per employee, (2) logarithm of sales net of raw materials per employee, and (3) logarithm of wages per employee.   All regressions include region-industry fixed effects. Errors are clustered at the regional level. The independent variables include: (1) Years of Education of manager, (2) Ln(Employees), (3) Years of Education of workers, (4) Ln(Expenditure on energy / employee), and (5) Ln(Property, Plant, Equipment / employees). All variables are described in Table 2.

| | Dependent Variable: | | | | | | | | | | | | | | |
| | Logarithm of Sales per employee | | | | | Ln[(Sales - Raw Materials)/Employee] | | | | | Logarithm of Wages per employee | | | | |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Years of Education of manager	0.0470[a]	0.0322[a]	0.0227[a]	0.0242[a]	0.0148[a]	0.0445[a]	0.0311[a]	0.0267[a]	0.0241[a]	0.0150[a]	0.0262[a]	0.0179[a]	0.0082[a]	0.0112[a]	0.0049
	(0.0033)	(0.0034)	(0.0033)	(0.0033)	(0.0047)	(0.0033)	(0.0033)	(0.0032)	(0.0033)	(0.0041)	(0.0027)	(0.0027)	(0.0028)	(0.0028)	(0.0038)
Ln(Employees)	.	0.1472[a]	.	-0.0134[b]	0.1324[a]	.	0.1388[a]	.	0.0230[a]	0.1099[a]	.	0.0822[a]	.	-0.0257[a]	0.0613[a]
	.	(0.0097)	.	(0.0062)	(0.0127)	.	(0.0101)	.	(0.0066)	(0.0121)	.	(0.0081)	.	(0.0055)	(0.0107)
Years of Education of workers	0.0318[a]	0.0243[a]	0.0434[a]	0.0443[a]	0.0071	0.0288[a]	0.0217[a]	0.0514[a]	0.0500[a]	0.0082[c]	0.0164[a]	0.0120[a]	0.0151[a]	0.0167[a]	0.0056
	(0.0033)	(0.0033)	(0.0027)	(0.0028)	(0.0044)	(0.0037)	(0.0036)	(0.0030)	(0.0030)	(0.0045)	(0.0028)	(0.0028)	(0.0025)	(0.0026)	(0.0038)
Ln(Expenditure on energy / employee)	0.3408[a]	0.3401[a]	.	.	0.2843[a]	0.3035[a]	0.3007[a]	.	.	0.2389[a]	0.2180[a]	0.2176[a]	.	.	0.1738[a]
	(0.0095)	(0.0094)	.	.	(0.0118)	(0.0102)	(0.0101)	.	.	(0.0117)	(0.0086)	(0.0086)	.	.	(0.0102)
Ln(Property, Plant, Equipment / employees)	.	.	0.3038[a]	0.3046[a]	0.1851[a]	.	.	0.2964[a]	0.2949[a]	0.1745[a]	.	.	0.1668[a]	0.1684[a]	0.1126[a]
	.	.	(0.0069)	(0.0070)	(0.0101)	.	.	(0.0072)	(0.0072)	(0.0100)	.	.	(0.0058)	(0.0058)	(0.0089)
Constant	6.9256[a]	6.6765[a]	4.2481[a]	4.2722[a]	5.9533[a]	9.3425[a]	9.1129[a]	5.1127[a]	5.0704[a]	8.9826[a]	6.2151[a]	6.0776[a]	2.7971[a]	2.8432[a]	5.5371[a]
	(0.0673)	(0.0660)	(0.0556)	(0.0564)	(0.0889)	(0.0682)	(0.0674)	(0.0564)	(0.0574)	(0.0863)	(0.0595)	(0.0604)	(0.0487)	(0.0500)	(0.0784)
Observations	13,248	13,248	19,305	19,305	7,733	10,651	10,651	17,893	17,893	6,655	12,782	12,782	19,209	19,209	7,706
Number of Regions-Industries	855	855	1,037	1,037	487	754	754	1,005	1,005	458	807	807	1,033	1,033	486
Within R^2	21%	23%	21%	21%	30%	20%	22%	20%	20%	29%	13%	14%	8%	8%	17%
Between R^2	90%	89%	67%	67%	88%	28%	24%	55%	54%	51%	89%	88%	66%	67%	86%
Overall R^2	79%	78%	58%	58%	81%	37%	33%	59%	58%	53%	76%	74%	57%	59%	78%
Regions-Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: a = significant at the 1% level, b = significant at the 5% level, and c = significant at the 10% level.

Table 9: Firm level productivity and Regional Human Capital

The table reports fixed effect regressions for for the following two dependent variables: (1) logarithm of sales per employee, (2) logarithm of sales net of raw materials per employee, and (3) logarithm of wages per employee. All regressions include country and industry fixed effects. Errors are clustered at the regional level. The independent variables include: (1) Years of Education in the Region, (2) Ln(Population in the Region), (3) Years of Education of manager, (4) Ln(Employees), (5) Years of Education of workers, (6) Ln(Expenditure on energy / employee), and (7) Ln(Property, Plant, Equipment / employees). All variables are described in Table 2.

Panel A: Basic Specification

Dependent Variable:

	Ln(Sales/Employee)					Ln[(Sales - Raw Materials)/Employee]					Logarithm of Wages per employee				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Years of Education in the Region	0.0655[a]	0.0639[a]	0.0954[a]	0.0950[a]	0.0478[b]	0.0748[a]	0.0735[a]	0.0945[a]	0.0928[a]	0.0565[a]	0.0580[a]	0.0577[a]	0.0840[a]	0.0843[a]	0.0462[a]
	(0.0202)	(0.0185)	(0.0280)	(0.0279)	(0.0185)	(0.0197)	(0.0181)	(0.0275)	(0.0270)	(0.0177)	(0.0162)	(0.0159)	(0.0233)	(0.0234)	(0.0151)
Ln(Population in the Region)	0.0920[a]	0.0803[a]	0.1437[a]	0.1409[a]	0.0917[a]	0.1332[a]	0.1178[a]	0.1046[b]	0.0938[c]	0.1188[a]	0.0682	0.0622	0.0135	0.0159	0.0787[c]
	(0.0321)	(0.0297)	(0.0501)	(0.0504)	(0.0328)	(0.0343)	(0.0327)	(0.0511)	(0.0501)	(0.0410)	(0.0425)	(0.0418)	(0.0352)	(0.0354)	(0.0413)
Years of Education of manager	0.0534[a]	0.0352[a]	0.0257[a]	0.0243[a]	0.0169[b]	0.0498[a]	0.0335[a]	0.0287[a]	0.0236[a]	0.0165[a]	0.0315[a]	0.0215[a]	0.0118[a]	0.0131[a]	0.0070
	(0.0047)	(0.0048)	(0.0062)	(0.0057)	(0.0077)	(0.0047)	(0.0042)	(0.0044)	(0.0043)	(0.0058)	(0.0038)	(0.0036)	(0.0044)	(0.0042)	(0.0052)
Ln(Employees)	.	0.1497[a]	.	0.0113	0.1468[a]	.	0.1392[a]	.	0.0401[a]	0.1177[a]	.	0.0827[a]	.	-0.0095	0.0717[a]
	.	(0.0154)	.	(0.0176)	(0.0193)	.	(0.0176)	.	(0.0141)	(0.0193)	.	(0.0150)	.	(0.0108)	(0.0187)
Years of Education of workers	0.0349[a]	0.0279[a]	0.0384[a]	0.0378[a]	0.0066	0.0332[a]	0.0264[a]	0.0490[a]	0.0468[a]	0.0112[c]	0.0195[a]	0.0151[a]	0.0146[a]	0.0152[a]	0.0049
	(0.0053)	(0.0054)	(0.0056)	(0.0058)	(0.0068)	(0.0059)	(0.0057)	(0.0059)	(0.0061)	(0.0062)	(0.0036)	(0.0036)	(0.0033)	(0.0033)	(0.0051)
Ln(Expenditure on energy / employee)	0.3577[a]	0.3554[a]	.	.	0.2902[a]	0.3183[a]	0.3133[a]	.	.	0.2440[a]	0.2248[a]	0.2232[a]	.	.	0.1721[a]
	(0.0185)	(0.0177)	.	.	(0.0220)	(0.0182)	(0.0175)	.	.	(0.0202)	(0.0173)	(0.0172)	.	.	(0.0168)
Ln(Property, Plant, Equipment / employees)	.	.	0.3258[a]	0.3250[a]	0.1946[a]	.	.	0.3150[a]	0.3118[a]	0.1818[a]	.	.	0.1787[a]	0.1794[a]	0.1230[a]
	.	.	(0.0132)	(0.0136)	(0.0162)	.	.	(0.0139)	(0.0141)	(0.0161)	.	.	(0.0086)	(0.0089)	(0.0135)
Constant	5.1202[a]	5.0055[a]	4.8529[a]	4.8850[a]	4.6033[a]	4.5073[a]	4.4682[a]	3.9685[a]	4.1094[a]	3.8699[a]	5.1007[a]	5.0322[a]	6.6732[a]	6.6461[a]	5.0701[a]
	(0.3706)	(0.3373)	(1.1885)	(1.1887)	(0.4521)	(0.4183)	(0.4008)	(0.8173)	(0.8009)	(0.6479)	(0.5225)	(0.5199)	(0.7223)	(0.7248)	(0.7073)
Observations	13,248	13,248	19,305	19,305	7,733	10,651	10,651	17,893	17,893	6,655	12,782	12,782	19,209	19,209	7,706
Number of Countries	29	29	22	22	21	25	25	21	21	20	27	27	22	22	21
Within R^2	30%	32%	31%	31%	37%	28%	30%	27%	28%	36%	20%	21%	13%	13%	22%
Between R^2	90%	90%	59%	59%	92%	87%	86%	70%	71%	89%	88%	87%	57%	57%	89%
Overall R^2	74%	74%	54%	54%	80%	74%	72%	73%	73%	81%	69%	68%	44%	44%	76%
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: a = significant at the 1% level, b = significant at the 5% level, and c = significant at the 10% level.

Table10: Firm level productivity, Regional Human Capital, and Geography

The table reports fixed effect regressions for for the following two dependent variables: (1) logarithm of sales per employee, (2) logarithm of sales net of raw materials per employee, and (3) logarithm of wages per employee. All regressions include country and industry fixed effects. Errors are clustered at the regional level. The independent variables include: (1) Temperature, (2) Inverse Distance to Coast, (3) Ln(Oil), (4) Years of Education in the Region, (5) Ln(Population in the Region), (6) Years of Education of manager, (7) Ln(Employees), (8) Years of Education of workers, (9) Ln(Expenditure on energy / employee), and (10) Ln(Property, Plant, Equipment / employees). All variables are described in Table 2.

	Dependent Variable:														
	Logarithm of Sales per employee					Ln[(Sales - Raw Materials)/Employee]					Logarithm of Wages per employee				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Temperature	0.0084	0.0105	0.0081	0.0082	0.0190	-0.0217	-0.0200	-0.0023	-0.0021	0.0278	-0.0141	-0.0128	0.0113^b	0.0112^b	-0.0046
	(0.0134)	(0.0124)	(0.0079)	(0.0079)	(0.0150)	(0.0171)	(0.0155)	(0.0079)	(0.0079)	(0.0232)	(0.0116)	(0.0112)	(0.0048)	(0.0048)	(0.0100)
Inverse Distance to Coast	0.3464	0.2275	1.1491^a	1.1458^a	0.6196	-0.2990	-0.4531	1.0464^a	1.0125^a	-0.1930	-0.0127	-0.0841	0.2830	0.2947	0.1021
	(0.3708)	(0.3652)	(0.2065)	(0.2079)	(0.4041)	(0.4233)	(0.3851)	(0.2286)	(0.2228)	(0.5291)	(0.3349)	(0.3369)	(0.1824)	(0.1824)	(0.3564)
Ln(Oil)	-0.8681^a	-0.6813^b	0.1682	0.1728	-1.2953^c	-0.7747^a	-0.5679^c	-0.0945	-0.0492	-0.6443	-0.6755^c	-0.5721	0.1596	0.1433	-1.2264^b
	(0.2856)	(0.3124)	(0.5944)	(0.5919)	(0.7382)	(0.2878)	(0.3084)	(0.4935)	(0.4942)	(0.8883)	(0.3430)	(0.3619)	(0.3833)	(0.3804)	(0.4702)
Years of Education in the Region	0.0532^a	0.0558^a	0.0236	0.0236	0.0287^c	0.0800^a	0.0839^a	0.0307	0.0310	0.0680^a	0.0560^a	0.0583^a	0.0639^b	0.0638^b	0.0467^b
	(0.0198)	(0.0184)	(0.0287)	(0.0287)	(0.0162)	(0.0205)	(0.0188)	(0.0299)	(0.0295)	(0.0180)	(0.0182)	(0.0181)	(0.0261)	(0.0261)	(0.0179)
Ln(Population in the Region)	0.0963^a	0.0817^a	0.0713^c	0.0705^c	0.1029^b	0.1235^a	0.1062^a	0.0493	0.0418	0.1195^a	0.0706^c	0.0629	-0.0257	-0.0230	0.0852^c
	(0.0321)	(0.0295)	(0.0382)	(0.0386)	(0.0416)	(0.0319)	(0.0300)	(0.0401)	(0.0395)	(0.0443)	(0.0417)	(0.0411)	(0.0291)	(0.0290)	(0.0439)
Years of Education of manager	0.0533^a	0.0353^a	0.0271^a	0.0266^a	0.0169^b	0.0495^a	0.0333^a	0.0300^a	0.0256^a	0.0168^a	0.0314^a	0.0216^a	0.0126^a	0.0142^a	0.0069
	(0.0047)	(0.0047)	(0.0060)	(0.0053)	(0.0077)	(0.0047)	(0.0042)	(0.0044)	(0.0041)	(0.0059)	(0.0037)	(0.0036)	(0.0043)	(0.0040)	(0.0052)
Ln(Employees)	.	0.1486^a	.	0.0035	0.1453^a	.	0.1393^a	.	0.0341^b	0.1178^a	.	0.0821^a	.	-0.0125	0.0719^a
	.	(0.0153)	.	(0.0175)	(0.0190)	.	(0.0174)	.	(0.0137)	(0.0191)	.	(0.0150)	.	(0.0110)	(0.0187)
Years of Education of workers	0.0344^a	0.0275^a	0.0406^a	0.0404^a	0.0058	0.0333^a	0.0266^a	0.0514^a	0.0494^a	0.0112^c	0.0194^a	0.0152^a	0.0154^a	0.0162^a	0.0047
	(0.0053)	(0.0054)	(0.0059)	(0.0061)	(0.0067)	(0.0059)	(0.0057)	(0.0061)	(0.0063)	(0.0062)	(0.0037)	(0.0037)	(0.0034)	(0.0034)	(0.0052)
Ln(Expenditure on energy / employee)	0.3571^a	0.3549^a	.	.	0.2890^a	0.3178^a	0.3128^a	.	.	0.2426^a	0.2245^a	0.2230^a	.	.	0.1717^a
	(0.0184)	(0.0176)	.	.	(0.0220)	(0.0183)	(0.0176)	.	.	(0.0204)	(0.0174)	(0.0172)	.	.	(0.0168)
Ln(Property, Plant, Equipment / employees)	.	.	0.3191^a	0.3189^a	0.1929^a	.	.	0.3108^a	0.3082^a	0.1825^a	.	.	0.1765^a	0.1773^a	0.1228^a
	.	.	(0.0122)	(0.0127)	(0.0160)	.	.	(0.0132)	(0.0134)	(0.0163)	.	.	(0.0085)	(0.0088)	(0.0135)
Constant	5.0425^a	5.1078^a	5.1076^a	5.1172^a	3.5975^a	11.2711^a	11.3696^a	4.9529^a	5.0614^a	3.8682^a	6.0203^a	5.9598^a	6.8223^a	6.7884^a	4.9977^a
	(0.4348)	(0.3958)	(0.9995)	(1.0016)	(0.5013)	(0.8349)	(0.7626)	(0.6071)	(0.6003)	(0.9306)	(0.7249)	(0.7132)	(0.6127)	(0.6112)	(0.7689)
Observations	13,248	13,248	19,305	19,305	7,733	10,651	10,651	17,893	17,893	6,655	12,782	12,782	19,209	19,209	7,706
Number of Countries	29	29	22	22	21	25	25	21	21	20	27	27	22	22	21
Within R^2	31%	32%	33%	33%	37%	29%	30%	28%	28%	36%	20%	21%	14%	14%	22%
Between R^2	93%	92%	57%	57%	89%	19%	17%	36%	36%	40%	84%	83%	58%	58%	88%
Overall R^2	74%	74%	57%	57%	78%	30%	27%	67%	66%	52%	69%	68%	54%	54%	75%
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: a = significant at the 1% level, b = significant at the 5% level, and c = significant at the 10% level.

Table 11: Regional Human Capital, the size of establishments, and participation in the economy

The table reports fixed effect regressions for for the following three dependent variables: (1) logarithm of the number of employees per establishment; (3) logarithm of the number of employees per capita; and (4) logarithm of the number of employees working in firms that employ at least 100 employees as a percent of total employment.   All regressions include the number of years of education. All variables are described in Table 2

	Dependent Variable:			
	Ln(Establishments/Population)	Ln(Employees/Establishments)	Ln(Employees/Population)	Ln(Employees Big Firms/Employees)
Years of Education in the Region	0.2967[a]	0.1233[a]	0.3418[a]	0.2445[a]
	(0.0314)	(0.0227)	(0.0273)	(0.0374)
Constant	-5.8626[a]	0.8855[a]	-4.3992[a]	-3.6568[a]
	(0.2571)	(0.2093)	(0.2119)	(0.4299)
Observations	951	983	988	501
Adjusted R^2	92%	83%	94%	95%
Country Fixed Effects	Yes	Yes	Yes	Yes

Note: a = significant at the 1% level, b = significant at the 5% level, and c = significant at the 10% level.

Figure 3: Regions in the database
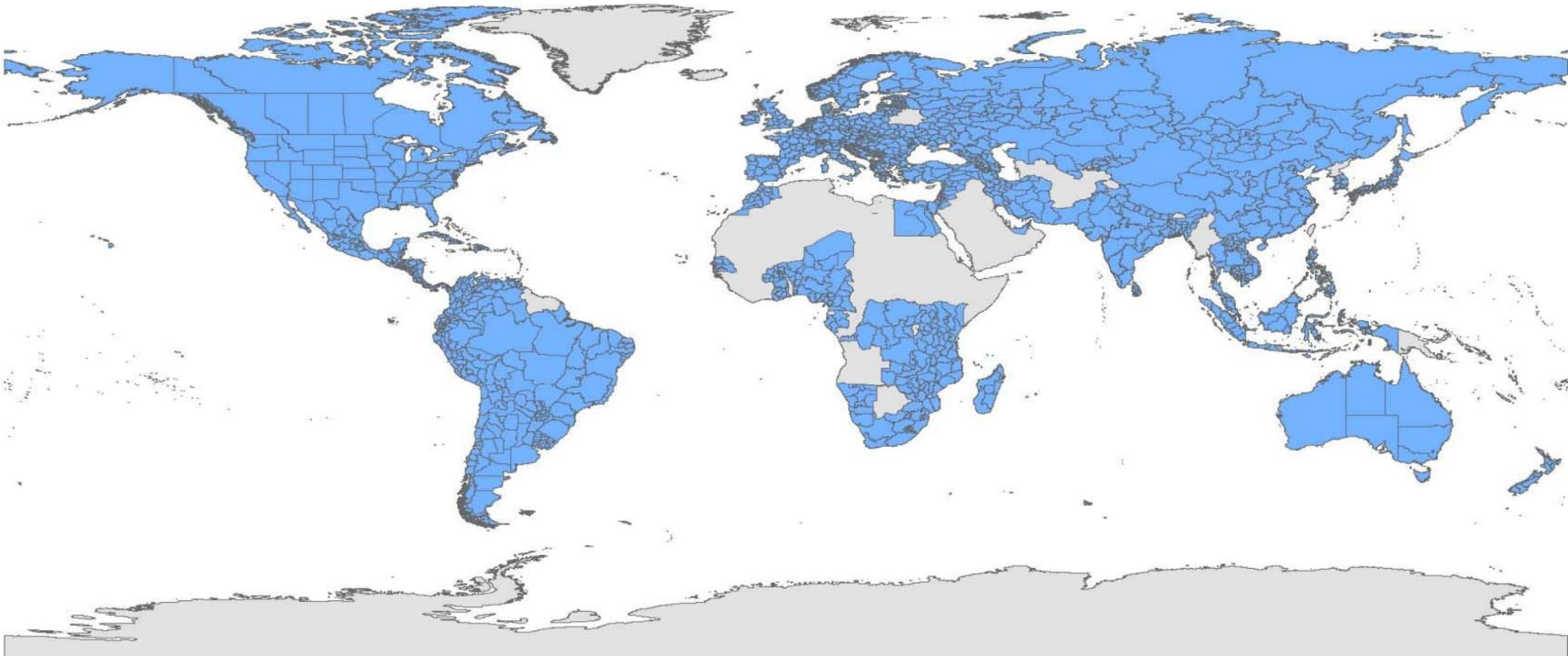
Figure 4: Within-country standard deviation of GDP per capita and development

coef = -.02489678, (robust) se = .01172037, t = -2.12

Figure 5: Within-country standard deviation of years of education and development

coef = -.12647737, (robust) se = .02655853, t = -4.76
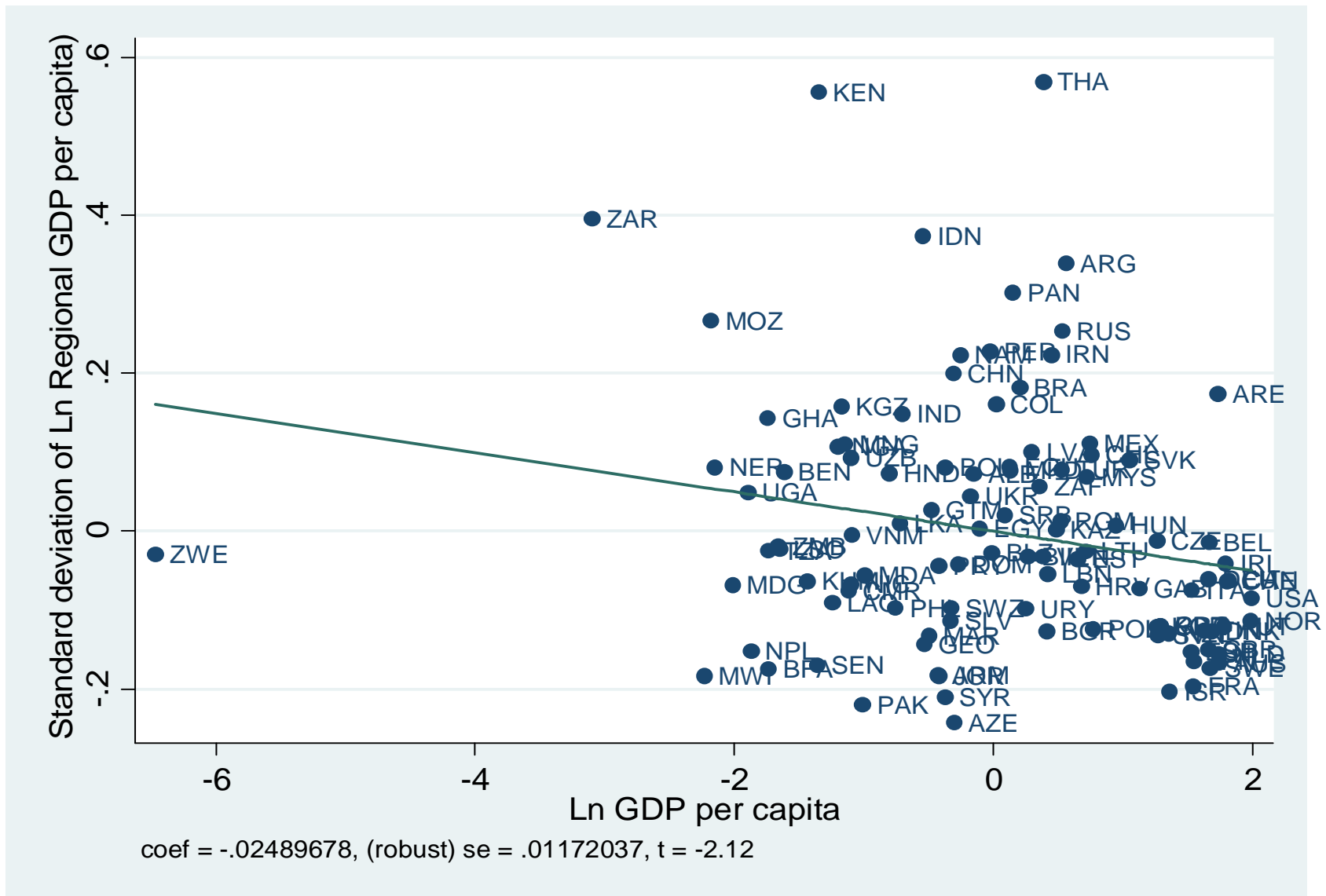
Figure 6: Partial Correlation Graph of (Log) GDP per capita and Years of Education

coef = .28573483, (robust) se = .01343762, t = 21.26
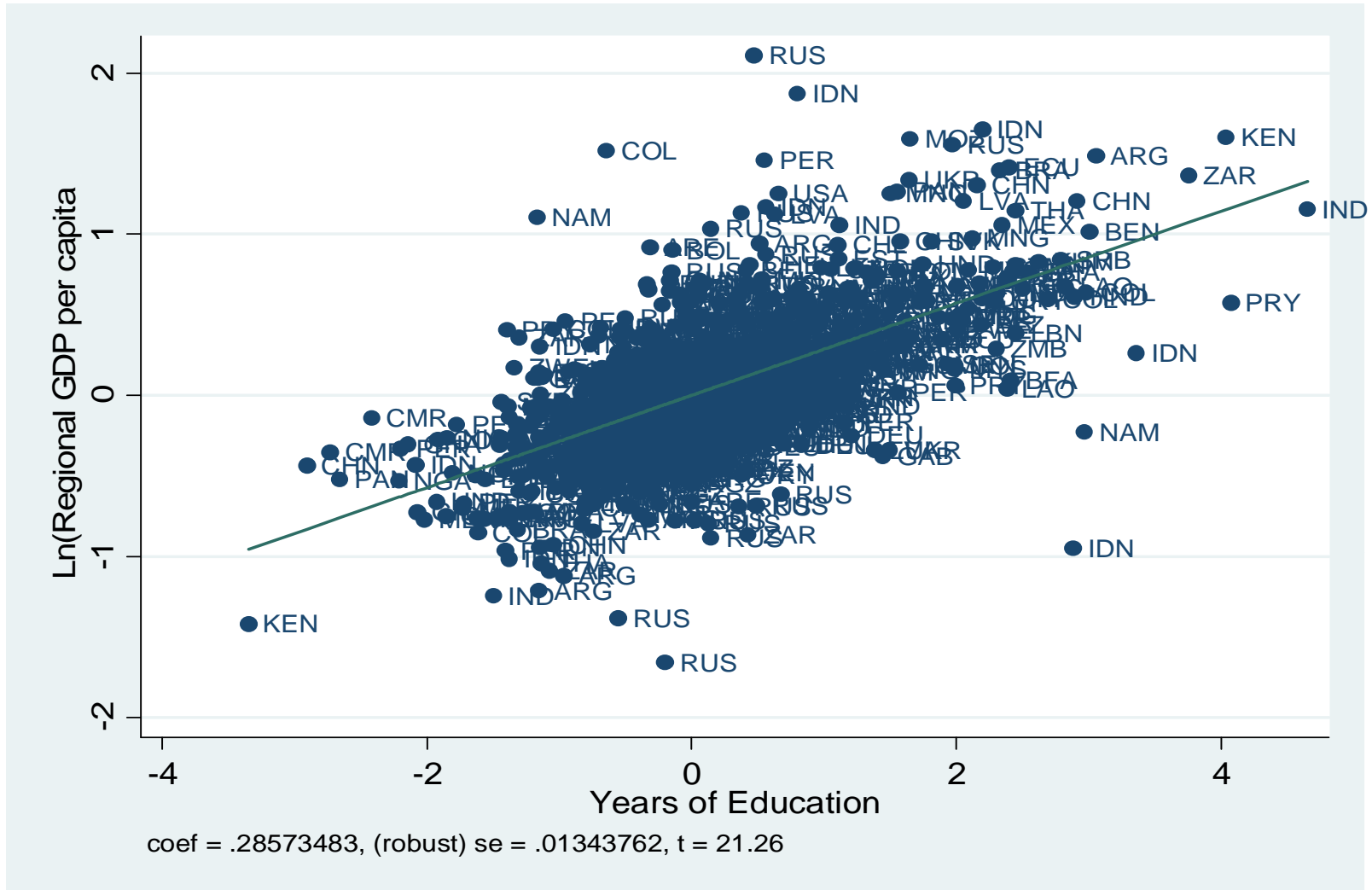
Figure 7: Years of Education and Participation in the Official Economy

coef = .29672036, (robust) se = .03135326, t = 9.46

coef = .12334688, (robust) se = .02273828, t = 5.42

coef = .34175848, (robust) se = .02725977, t = 12.54

coef = .24454543, (robust) se = .03739797, t = 6.54
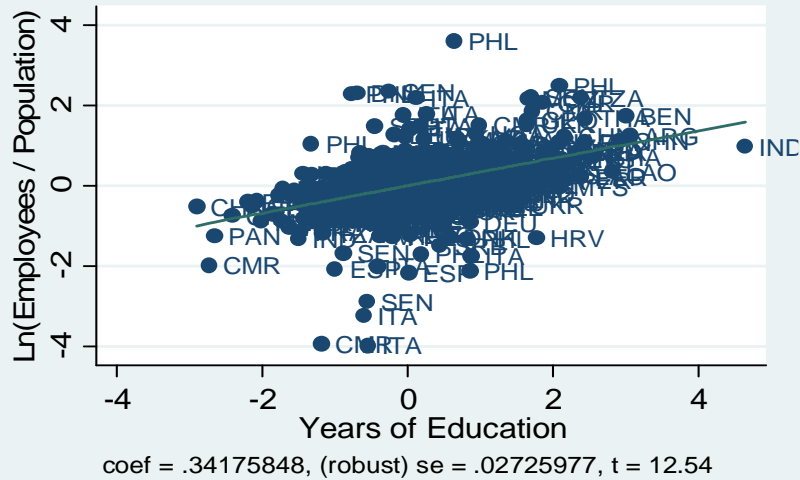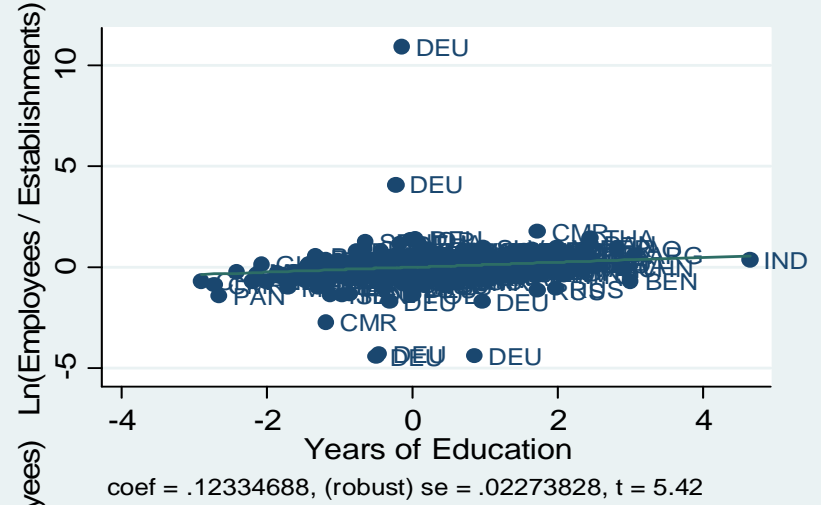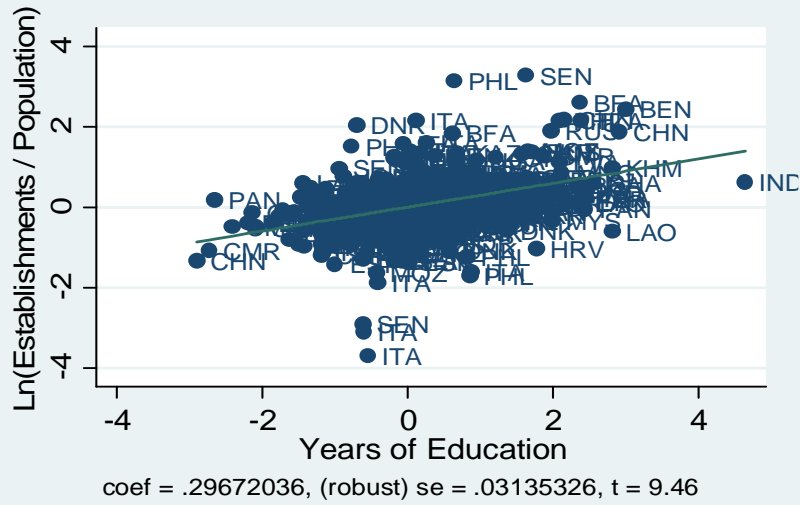
Appendix A: Reporting level for countries in our dataset

The table identifies the reporting level for the regions in our database. The table splits countries in three main groups: (1) countries where data is reported at the first-order administrative regions; (2) countries where data is reported for economic or statistical regions and where first-order administrative regions are equivalent to provinces, states or derpartments; and (3) countries where data is reported for economic or statistical regions and where first-order administrative regions are equivalent to counties, boroughs, cities, districts or municipalities. The table also subdivides countries based on the reason why the first-order administrative regions are different than the reporting regions for each of these three groups of countries.

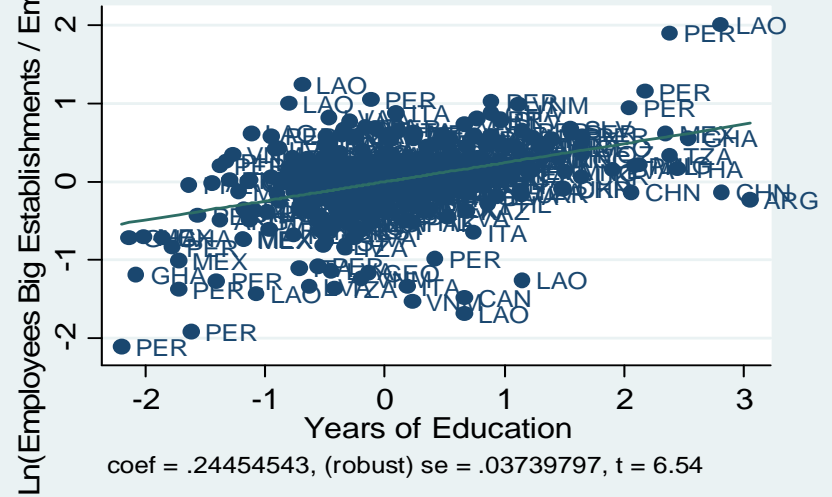|	Number of countries	First-order administrative regions	Regions in our dataset	Country names (number of first-order administrative regions lost)
1. Reporting done at the first-order administrative level:	**79**	**1,362**	**1,328**	
Our regions match first-order administrative level:	60	934	934	
Differences due to :				
Missing information for some region	7	148	130	France (4 overseas departments), Grece (1 self-governing monastic state), India (2 union territories & 1 island), Morocco (2 disputed territories), Pakistan (1 Tribal area), Tanzania (5 islands), Venezuela (2)
Aggregation of some regions	6	183	168	Croatia (1), Mozambique (1), New Zealand (3), Russia (3), Serbia (6), Switzerland (1)
Political change during sample period	6	97	96	Canada (1), Chile (2), Denmark(-10), Ecuador (2), Peru (2), Senegal (4)
2. Reporting done for economic or statistical regions. First-order administrative regions are equivalent to provinces, states or departments.	**22**	**691**	**177**	
Most data collected for statistical regions	6	78	44	Belgium(-8), Cezch Republic(6), Finland(1), Nepal(9), Portugal (13),Sweden(13)
GDP per capita collected for statistical regions	4	88	37	Dominican Republic(23), Kazakhstan(10), Cambodia(9), South Korea(9)
Education collected for statistical regions	12	525	96	Burkina Faso(32), Bulgaria(22), Egypt(22), Gabon(5), Guatemala(14), Nigeria(31), Philippines(65), Thailand(71), Turkey(69), Romania(34), Uzbekistan(9), Vietnam(55)
3. Reporting done for economic or statistical regions. First-order administrative regions are equivalent to counties, boroughs, cities, districts, or municipalities.	**9**	**782**	**64**	
Most data collected for statistical regions	7	725	52	Azerbajan (66), Great Britain (217), Ireland(32), Macedonia(76), Malawi(25), Slovenia(181), Uganda(76)
Education collected for statistical regions	2	57	12	Hungary(13), Moldova(32)
Total in the sample	**110**	**2,835**	**1,569**	

APPENDIX OF DATA SOURCES: REGIONAL GDP

Code	Country	Type of Data	Source	Available link
ALB	Albania	GDP	Data from HDR 2002	
ARE	United Arab Emirates	GDP	Data from HDR 1997 in arabic	
ARG	Argentina	GDP	1990-2001 Data from Ministry of interior	http://www.ec.gba.gov.ar/Estadistica/FTP/pbg/pbg3.html
ARM	Armenia	Expenditure	National Statistics Office	http://www.armstat.am/file/article/marz_07_e_22.pdf
AUS	Australia	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
AUT	Austria	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
AZE	Azerbaijan	Income	National Statistics Office	http://74.125.47.132/search?q=cache:http://www.azstat.org/statinfo/budget_households/en/003.shtml
BEL	Belgium	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
BEN	Benin	GDP	Data from HDR 2007/2008 and 2003	
BFA	Burkina Faso	GDP	Data from HDR for GDP per capita.	
BGD	Bangladesh	NA		
BGR	Bulgaria	GDP	Data from HDR 2003, 2002 and 2001	
BIH	Bosnia and Herzegovina	GDP	National Statistics Offices	http://www.bhas.ba/Arhiva/2007/brcko/PODACI%201-08.pdf
BLZ	Belize	Expenditure	Data from LSMS 2002	http://www.statisticsbelize.org.bz/dms20uc/dm_filedetails.asp?action=d&did=13
BOL	Bolivia	GDP	National Statistics Office	http://www.ine.gov.bo/indice/visualizador.aspx?ah=PC0104010201.HTM
BRA	Brazil	GDP	National Statistics Office	http://www.ibge.gov.br/home/estatistica/economia/contasregionais/2002_2005/contasregionais2002_2005.pdf
CAN	Canada	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
CHE	Switzerland	GDP	National Statistics Office	http://www.bfs.admin.ch/bfs/portal/fr/index/infothek/lexikon/bienvenue___login/blank/zugang_lexikon.Document.20896.xls
CHL	Chile	GDP	National Statistics Office	http://www.bcentral.cl/publicaciones/estadisticas/actividad-economica-gasto/aeg07a.htm
CHN	China	GDP	Data from National Statistics Yearbooks 2006, 2002, 1998 and 1996	http://www.stats.gov.cn/english/statisticaldata/yearlydata/YB1998e/C3-8E.htm
CMR	Cameroon	Expenditure	National Statistics Office	http://www.statistics-cameroon.org/archive/ECAM/ECAM2001/survey0/data/ECAM2001/Documentation/ECAM%20II%20-%20Rapport%20principal.pdf
COL	Colombia	GDP	National Statistics Office	http://www.dane.gov.co/index.php?option=com_content&task=category&sectionid=33&id=148&Itemid=705
CRI	Costa Rica	NA		
CUB	Cuba	Wages	Monthly wages from HDR 1996	
CZE	Czech Republic	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
DEU	Germany	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
DNK	Denmark	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
DOM	Dominican Republic	GDP	National Statistics Office	
ECU	Ecuador	GDP	National Statistics Office	
EGY	Egypt	GDP	Data from HDRs 2008, 2005, 2004, 2003, 2001; data from 2006 excluded	
ESP	Spain	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
EST	Estonia	GDP	Data from National Statistics Office	http://pub.stat.ee/px-web.2001/I_Databas/Economy_regional/23National_accounts/01Gross_Domestic_product_(GDP)/14Regional_gross_domestic_product/14Regional_gross_domestic_product.asp
FIN	Finland	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
FRA	France	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
GAB	Gabon	Expenditure	Data from HDR 2005	
GBR	United Kingdom	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
GEO	Georgia	GDP	Data from HDR 2002.	http://www.undp.org.ge/nhdr2001-02/chpt1.htm
GHA	Ghana	Income	Data from Living Standards Measurement Survey Reports for 1998/9 and 1991/2	http://siteresources.worldbank.org/INTLSMS/Resources/3358986-1181743055198/3877319-1190221709991/G3report.pdf
GRC	Greece	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
GTM	Guatemala	GDP	Data from HDR 2007/2008 annex	http://cms.fideck.com/userfiles/desarrollohumano.org/File/8012264236003654.pdf
HND	Honduras	GDP	Data from HDR 2006	
HRV	Croatia	GDP	Data from National Statistics Office	http://www.dzs.hr/default_e.htm
HUN	Hungary	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
IDN	Indonesia	GDP	Data from National Statistics Office	http://www.bps.go.id/sector/nra/grdp/table1.shtml
IND	India	GDP	National Statistics Office	http://mospi.nic.in/6_gsdp_cur_9394ser.htm
IRL	Ireland	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
IRN	Iran	GDP	Data from National Statistics Office	http://www.sci.org.ir/content/userfiles/_sci_en/sci_en/sel/year85/f21/CS_21_4.HTM
ISR	Israel	GDP	National Statistics Office	
ITA	Italy	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
JOR	Jordan	GDP	Data from HDR 2004	
JPN	Japan	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
KAZ	Kazakhstan	Income	LSMS 1996, World Bank	http://siteresources.worldbank.org/INTLSMS/Resources/3358986-1181743055198/3877319-1181930718899/finrep1.pdf
KEN	Kenya	GDP	Data from HDRs for 2006, 2005, 2003, 2001 and 1999	
KGZ	Kyrgyz Republic	GDP	Data from HDR 2005, 2001	

Code	Country	Type of Data	Source	Available link
KHM	Cambodia	Expenditure	Data from Poverty profile of Cambodia 2004; Daily consumption	http://www.mop.gov.kh/Situationandpolicyanalysis/PovertyProfile/tabid/191/Default.aspx
KOR	Korea, Rep.	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
LAO	Lao PDR	C+I+G	Data from HDR 2006; Consumption, Investment and Government Expenditure	
LBN	Lebanon	GDP	Data from HDR 2001	
LKA	Sri Lanka	GDP	Data from HDR 1998 and National Statistics Office	http://www.cbsl.gov.lk/pics_n_docs/08_statistics/_docs/xls_real_sector/table1.17.xls
LSO	Lesotho	GDP	Data from HDR 2006	
LTU	Lithuania	GDP	Data from National Statistics Office	http://db1.stat.gov.lt/statbank/SelectVarVal/Define.asp?MainTable=M2010210&PLanguage=1&PXSId=0&ShowNews=OFF
LVA	Latvia	GDP	National Statistics Office	http://data.csb.gov.lv/Dialog/varval.asp?ma=02-02a&ti=2-2.+GROSS+DOMESTIC+PRODUCT+BY+STATISTICAL+REGION,+CITY+AND+DISTRICT&path=../DATABASEEN/ekfin/Annual%20statistical%20data/02.%20Gross%20domestic%20product/&lang=1
MAR	Morocco	GDP + Expenditure	Data from HDR 1999, 2003 and Enquete Nationale sur la Consommation et les Depenses des Menages 2000/2001	
MDA	Moldova	Wages	Data from 2007 Statistical Yearbook; monthly salary	http://www.statistica.md/public/files/Yearbook/Venit_pop_1999_2006_en.doc
MDG	Madagascar	GDP	Data from HDR 2003, 2000	
MEX	Mexico	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
MKD	Macedonia, FYR	GDP	Data from National Statistics Office	http://www.stat.gov.mk/english/statistiki_eng.asp?ss=09.01&rbs=2
MNG	Mongolia	GDP	National Statistics Office	
MOZ	Mozambique	GDP	Data from HDR 2007, 2001	
MWI	Malawi	Expenditure	Data from Malawi INTEGRATED HOUSEHOLD SURVEY 2004-2005 and 1998	
MYS	Malaysia	GDP	Data from Chapter 5 of EIGHTH MALAYSIA PLAN 2001 - 2005	http://www.epu.jpm.my/new%20folder/development%20plan/RM8.htm
NAM	Namibia	Expenditure	Data from Namibia Household Income & Expenditure Survey 2003/2004; data is expenditure	http://www.npc.gov.na/publications/prenhies03_04.pdf
NER	Niger	GDP	Data from HDR 2004	
NGA	Nigeria	Income	2006 Annual Abstract of Statistics.	http://nigerianstat.gov.ng/annual_report.htm
NIC	Nicaragua	Expenditure	Data from HDR 2002	
NLD	Netherlands	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
NOR	Norway	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
NPL	Nepal	GDP	Data from HDR 2004, 2001 and 1998	
NZL	New Zealand	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
PAK	Pakistan	GDP	Data from HDR 2003	
PAN	Panama	GDP	Data from National Statistics Office	http://www.contraloria.gob.pa/dec/
PER	Peru	GDP	Cuentas Nacionales del Peru, Producto Bruto Interno por Departmentos 2001-2006	http://www1.inei.gob.pe/biblioineipub/bancopub/est/lib0763/cuadros/c037.xls
PHL	Philippines	GDP	National Statistics Office	
POL	Poland	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
PRT	Portugal	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
PRY	Paraguay	GDP	Data from Atlas de Desarrollo Humano Paraguay 2007	http://www.undp.org.py/dh/?page=atlas
ROM	Romania	GDP	Data from National Statistics Office	http://www.insse.ro/cms/files/pdf/en/cp11.pdf
RUS	Russia	GDP	National Statistics Office	http://translate.google.com/translate?ie=UTF8&oe=UTF&u=http://www.gks.ru/bgd/regl/b07_14p/IssWWW.exe/Stg/d02/10-02.htm&hl=en&ie=UTF8&sl=ru&tl=en
SEN	Senegal	GDP	Data from HDR 2001	
SLV	El Salvador	GDP	Data from HDR 2007/2008, 2005, 2003, 2001; 1996 values were in 1994 prices	
SRB	Serbia	Income	Data from National Statistics Municipal Database	http://www.statserb.sr.gov.yu/Pod/epok.asp
SVK	Slovak Republic	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
SVN	Slovenia	GDP	Data from National Statistics Office	http://www.stat.si/eng/novica_prikazi.aspx?id=1318
SWE	Sweden	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
SWZ	Swaziland	GDP	Data from HDR 2008	
SYR	Syrian Arab Republic	GDP	Data from HDR 2005	
THA	Thailand	GDP	Data from Statistical Year Book Thailand 2002	http://web.nso.go.th/eng/en/pub/pub0.htm
TUR	Turkey	GDP	National Statistics Office	http://www.turkstat.gov.tr/VeriBilgi.do?tb_id=56&ust_id=16
TZA	Tanzania	GDP	National Statistics Office	
UGA	Uganda	GDP	Data from HDR 2007	
UKR	Ukraine	GDP	Data from National Statistics Office	http://www.ukrstat.gov.ua/operativ/operativ2008/vvp/vrp/vrp2008_e.htm
URY	Uruguay	GDP	Data from HDR 2005	
USA	United States	GDP	Data from OECDStats	http://stats.oecd.org/WBOS/index.aspx
UZB	Uzbekistan	GDP	Data from HDR 2007/8, 2000 and 1998	
VEN	Venezuela	GDP	Data from HDR 2000	
VNM	Vietnam	Wages	National Statistics Office	http://www.gso.gov.vn/Modules/Doc_Download.aspx?DocID=2097
ZAF	South Africa	GDP	National Statistics Office (table 16)	http://www.statssa.gov.za/publications/statsdownload.asp?PPN=P0441&SCH=4048
ZAR	Congo, Dem. Rep.	GDP	Data from HDR 2008	
ZMB	Zambia	GDP	Data from HDR 2007 and 2003	
ZWE	Zimbabwe	GDP	Data fom HDR 2003	

APPENDIX OF DATA SOURCES: REGIONAL YEARS OF EDUCATION

Code	Country	Source	Available Link
ALB	Albania	NA	
ARE	United Arab Emirates	Ministry of Economy, 2005 Census	http://www.economy.ae/English/economicandstatisticreports/statisticreports/pages/census2005.aspx
ARG	Argentina	Education Policy and Data Center (EPDC)	http://epdc.org/
ARM	Armenia	Education Policy and Data Center (EPDC)	http://epdc.org/
AUS	Australia	National Statistics Office	http://www.abs.gov.au/
AUT	Austria	Eurostat	http://epp.eurostat.ec.europa.eu/NavTree_prod/everybody/BulkDownloadListing
AZE	Azerbaijan	Education Policy and Data Center (EPDC)	http://epdc.org/
BEL	Belgium	Eurostat	http://epp.eurostat.ec.europa.eu/NavTree_prod/everybody/BulkDownloadListing
BEN	Benin	Education Policy and Data Center (EPDC)	http://epdc.org/
BFA	Burkina Faso	Education Policy and Data Center (EPDC)	http://epdc.org/
BGD	Bangladesh	Education Policy and Data Center (EPDC)	http://epdc.org/
BGR	Bulgaria	Eurostat	http://epp.eurostat.ec.europa.eu/NavTree_prod/everybody/BulkDownloadListing
BIH	Bosnia and Herzegovina	Education Policy and Data Center (EPDC)	http://epdc.org/
BLZ	Belize	Education Policy and Data Center (EPDC)	http://epdc.org/
BOL	Bolivia	Education Policy and Data Center (EPDC)	http://epdc.org/
BRA	Brazil	Integrated Public Use Microdata Series International (IPUMS)	https://international.ipums.org/international/
CAN	Canada	National Statistics Office	http://www40.statcan.gc.ca/l01/cst01/educ43a-eng.htm
CHE	Switzerland	Swiss Labor Force Survey (SLFS) SFSO	http://www.bfs.admin.ch/bfs/portal/de/index/themen/15/04/ind4.informations.40101.401.html
CHL	Chile	National Statistics Office	http://espino.ine.cl/CuadrosCensales/apli_excel.asp
CHN	China	National Statistics Office	http://www.stats.gov.cn/ndsj/information/nj97/C091A.END
CMR	Cameroon	Education Policy and Data Center (EPDC)	http://epdc.org/
COL	Colombia	National Statistics Office	http://190.25.231.246:8080/Dane/tree.jsf
CRI	Costa Rica	Education Policy and Data Center (EPDC)	http://epdc.org/
CUB	Cuba	NA	
CZE	Czech Republic	Eurostat	http://epp.eurostat.ec.europa.eu/NavTree_prod/everybody/BulkDownloadListing
DEU	Germany	Eurostat	http://epp.eurostat.ec.europa.eu/NavTree_prod/everybody/BulkDownloadListing
DNK	Denmark	National Statistics Office	http://www.statbank.dk/statbank5a/SelectVarVal/Define.asp?Maintable=RASU1&PLanguage=1
DOM	Dominican Republic	Education Policy and Data Center (EPDC)	http://epdc.org/
ECU	Ecuador	National Statistics Office	http://190.95.171.13/cgibin/RpWebEngine.exe/PortalAction?&MODE=MAIN&BASE=ECUADOR21&MAIN=WebServerMain.inl
EGY	Egypt	Education Policy and Data Center (EPDC)	http://epdc.org/
ESP	Spain	Eurostat	http://epp.eurostat.ec.europa.eu/NavTree_prod/everybody/BulkDownloadListing
EST	Estonia	National Statistics Office	http://pub.stat.ee/px-web.2001/Dialog/varval.asp?ma=PC414&ti=ECONOMICALLY+ACTIVE+POPULATION+BY+AGE,+EDUCATIONAL+ATTAINMENT+AND+ETHNIC+NATIONALITY*&path=../I_Databas/Population_census/06Economically_active_population/(=1
FIN	Finland	Eurostat	http://epp.eurostat.ec.europa.eu/NavTree_prod/everybody/BulkDownloadListing
FRA	France	Eurostat	http://epp.eurostat.ec.europa.eu/NavTree_prod/everybody/BulkDownloadListing
GAB	Gabon	Education Policy and Data Center (EPDC)	http://epdc.org/
GBR	United Kingdom	Eurostat	http://epp.eurostat.ec.europa.eu/NavTree_prod/everybody/BulkDownloadListing
GEO	Georgia	National Statistics Office (special request of data)	
GHA	Ghana	Education Policy and Data Center (EPDC)	http://epdc.org/
GRC	Greece	Eurostat	http://epp.eurostat.ec.europa.eu/NavTree_prod/everybody/BulkDownloadListing
GTM	Guatemala	Education Policy and Data Center (EPDC)	http://epdc.org/
HND	Honduras	Education Policy and Data Center (EPDC)	http://epdc.org/
HRV	Croatia	National Statistics Office	http://www.dzs.hr/Eng/censuses/Census2001/Popis/E01_01_07/E01_01_07.html
HUN	Hungary	Eurostat	http://epp.eurostat.ec.europa.eu/NavTree_prod/everybody/BulkDownloadListing
IDN	Indonesia	Education Policy and Data Center (EPDC)	http://epdc.org/
IND	India	Education Policy and Data Center (EPDC)	http://epdc.org/
IRL	Ireland	Eurostat	http://epp.eurostat.ec.europa.eu/NavTree_prod/everybody/BulkDownloadListing
IRN	Iran	NA	
ISR	Israel	Integrated Public Use Microdata Series International (IPUMS)	https://international.ipums.org/international/
ITA	Italy	Eurostat	http://epp.eurostat.ec.europa.eu/NavTree_prod/everybody/BulkDownloadListing
JOR	Jordan	Education Policy and Data Center (EPDC)	http://epdc.org/
JPN	Japan	National Statistics Office	http://www.e-stat.go.jp/SG1/chiiki/ToukeiDataSelectDispatchAction.do
KAZ	Kazakhstan	Education Policy and Data Center (EPDC)	http://epdc.org/
KEN	Kenya	Education Policy and Data Center (EPDC)	http://epdc.org/
KGZ	Kyrgyz Republic	Education Policy and Data Center (EPDC)	http://epdc.org/
KHM	Cambodia	Education Policy and Data Center (EPDC)	http://epdc.org/
KOR	Korea, Rep.	NA	
LAO	Lao PDR	Education Policy and Data Center (EPDC)	http://epdc.org/
LBN	Lebanon	Ministry of Social Affairs	http://www.cas.gov.lb/images/PDFs/Educational%20status-2004.pdf
LKA	Sri Lanka	Education Policy and Data Center (EPDC)	http://epdc.org/

Code	Country	Source	Available Link
LSO	Lesotho	Education Policy and Data Center (EPDC)	http://epdc.org/
LTU	Lithuania	National Statistics Office	http://db.stat.gov.lt/sips/Dialog/varval.asp?ma=gs_dem17en&ti=Population+by+educational+attainment+and+age+group%A0%28aged+10+years+and+over%29&path=../Database/cen_en/p71en/demography/(=2
LVA	Latvia	National Statistics Office	http://data.csb.gov.lv/Dialog/varval.asp?ma=tsk03a&ti=EDUCATIONAL+ATTAINMENT+OF+POPULATION&path=../DATABASEEN/tautassk/Results%20of%20Population%20Census%202000in%20brief,(=1
MAR	Morocco	Education Policy and Data Center (EPDC)	http://epdc.org/
MDA	Moldova	Education Policy and Data Center (EPDC)	http://epdc.org/
MDG	Madagascar	Education Policy and Data Center (EPDC)	http://epdc.org/
MEX	Mexico	Education Policy and Data Center (EPDC)	http://epdc.org/
MKD	Macedonia, FYR	Education Policy and Data Center (EPDC)	http://epdc.org/
MNG	Mongolia	Integrated Public Use Microdata Series International (IPUMS)	https://international.ipums.org/international/
MOZ	Mozambique	Education Policy and Data Center (EPDC)	http://epdc.org/
MWI	Malawi	Education Policy and Data Center (EPDC)	http://epdc.org/
MYS	Malaysia	Integrated Public Use Microdata Series International (IPUMS)	https://international.ipums.org/international/
NAM	Namibia	Education Policy and Data Center (EPDC)	http://epdc.org/
NER	Niger	Education Policy and Data Center (EPDC)	http://epdc.org/
NGA	Nigeria	Education Policy and Data Center (EPDC)	http://epdc.org/
NIC	Nicaragua	Education Policy and Data Center (EPDC)	http://epdc.org/
NLD	Netherlands	Eurostat	http://epp.eurostat.ec.europa.eu/NavTree_prod/everybody/BulkDownloadListing
NOR	Norway	National Statistics Office	http://statbank.ssb.no/statistikkbanken/Default_FR.asp?PXSid=0&nvl=true&PLanguage=1&tilside=selecttable/hovedtabellHjem.asp&KortnavnWeb=utniv
NPL	Nepal	Education Policy and Data Center (EPDC)	http://epdc.org/
NZL	New Zealand	National Statistics Office	http://wdmzpub01.stats.govt.nz/wds/ReportFolders/reportFolders.aspx
PAK	Pakistan	Education Policy and Data Center (EPDC)	http://epdc.org/
PAN	Panama	Education Policy and Data Center (EPDC)	http://epdc.org/
PER	Peru	Education Policy and Data Center (EPDC)	http://epdc.org/
PHL	Philippines	Education Policy and Data Center (EPDC)	http://epdc.org/
POL	Poland	Eurostat	http://epp.eurostat.ec.europa.eu/NavTree_prod/everybody/BulkDownloadListing
PRT	Portugal	Eurostat	http://epp.eurostat.ec.europa.eu/NavTree_prod/everybody/BulkDownloadListing
PRY	Paraguay	National Statistics Office	http://celade.cepal.org/cgibin/RpWebEngine.exe/EasyCross?&BASE=CPVPRY2002&ITEM=INDICADO&MAIN=WebServerMain.inl
ROM	Romania	Eurostat	http://epp.eurostat.ec.europa.eu/NavTree_prod/everybody/BulkDownloadListing
RUS	Russian Federation	National Statistics Office	http://74.125.65.132/translate_c?hl=en&ie=UTF-8&sl=ru&tl=en&u=http://www.perepis2002.ru/index.html%3Fid%3D15&prev=_t&usg=ALkJrhiZr6thPp3doxH9mXdDZgf-DA1fyw
SEN	Senegal	Education Policy and Data Center (EPDC)	http://epdc.org/
SLV	El Salvador	VI Censo de la Poblacion y V de Vivienda 2007	http://www.digestyc.gob.sv/cgibin/RpWebEngine.exe/Crosstabs
SRB	Serbia	National Statistics Office	http://webrzs.statserb.sr.gov.yu/axd/en/Zip/CensusBook4.zip
SVK	Slovak Republic	National Statistics Office	http://px-web.statistics.sk/PXWebSlovak/DATABASE/En/02EmploMarket/01EconPopActiv/EA_total.px
SVN	Slovenia	National Statistics Office	http://www.stat.si/pxweb/Database/Census2002/Administrative%20units/Population/Education/Education.asp
SWE	Sweden	National Statistics Office	http://www.ssd.scb.se/databaser/makro/SubTable.asp?yp=tansss&xu=C9233001&omradekod=UF&huvudtabell=Utbildning&omradetext=Education+and+research&tabelltext=Population+16-74+years+of+age+by+highest+level+of+education,+age+and+sex.+Year&preskat=O&prodid=UF0506&starttid=1985&stopptid=2007&Fromwhere=M(=?&langdb=?
SWZ	Swaziland	Education Policy and Data Center (EPDC)	http://epdc.org/
SYR	Syrian Arab Republic	Education Policy and Data Center (EPDC)	http://epdc.org/
THA	Thailand	Education Policy and Data Center (EPDC)	http://epdc.org/
TUR	Turkey	National Statistics Office	http://www.tuik.gov.tr/isgucueng/Kurumsal.do
TZA	Tanzania	Education Policy and Data Center (EPDC)	http://epdc.org/
UGA	Uganda	Education Policy and Data Center (EPDC)	http://epdc.org/
UKR	Ukraine	National Statistics Office	http://stat6.stat.lviv.ua/PXWEB2007/Database/POPULATION/1/06/06.asp
URY	Uruguay	National Statistics Office	http://www.ine.gub.uy/microdatos/engih2006/persona.zip,
USA	United States	National Statistics Office	http://factfinder.census.gov/servlet/DatasetMainPageServlet?_program=ACS&_submenuId=&_lang=en&_ts=
UZB	Uzbekistan	Education Policy and Data Center (EPDC)	http://epdc.org/
VEN	Venezuela	Integrated Public Use Microdata Series International (IPUMS)	https://international.ipums.org/international/
VNM	Vietnam	Education Policy and Data Center (EPDC)	http://epdc.org/
ZAF	South Africa	National Statistics Office	http://www.statssa.gov.za/timeseriesdata/pxweb2006/Dialog/varval.asp?ma=Highest%20level%20of%20education%20grouped%20by%20province&ti=Table:+Census+2001+by+province,+highest+level+of+education+grouped,++population+group+and+gender.+&path=../Database/South%20Africa/Population%20Census/Census%202001%20-%20NEW%20Demarcation%20boundaries%20as%20at%209%20December%202005/Provincial%20level%20-%20Persons/(=1
ZAR	Congo, Dem. Rep.	Education Policy and Data Center (EPDC)	http://epdc.org/
ZMB	Zambia	Education Policy and Data Center (EPDC)	http://epdc.org/
ZWE	Zimbabwe	Education Policy and Data Center (EPDC)	http://epdc.org/

APPENDIX OF DATA SOURCES: ECONOMIC CENSUS DATA

Country	Code	Year	Source	Big Firms	Industry	Calculations	Links
Albania	ALB	2009	Data is from the Albanian Institute of Statistics' Economic Indicators, Business Register, table titled: Active Enterprises by Counties and Size.	50 Employees	non-agriculture	Data includes the number of establishments in four size categories. Data for the number of employees in each size category is estimated using an assumption of 2.5 employees for establishments with 1 to 4 employees, 7 for establishments with 5 to 9, 29.5 for establishments with 10 to 49 and 50 employees for establishments with over 50 employees. Data excludes establishments with no employees.	http://www.instat.gov.al/
Argentina	ARG	2003	Data is from the Instituto Nacional de Estadística y Censos, Censo Nacional Económico, Industria Manufacturera, table 2.	na	non-agriculture	Data includes the number of establishments and employees by three size categories. The number of employees are approximated here by the use of Jobs Held data.	http://www.indec.mecon.ar/
Armenia	ARM	2008	Data is from the National Statistical Service of the Republic of Armenia, Main Statistical Indicators 2005-2008	na	all sectors	No size data was found available for establishments. For employee data we use figures from the labor force survey.	http://www.armstat.am/file/RegStat/default-eng.html
Australia	AUS	2006	Data is from the Australian Bureau of Statistics, statistical catalogue number 81. Industry wide statistics, table 8165.0 Counts of Australian Businesses, including Entries and Exits, Jun 2003 to Jun 2007, Businesses by Industry Class by Main State by Employment Size Ranges - 2006-07.	200 Employees	non-agriculture (i.e. excl. 0111-0219)	Data includes the number of establishments by four size categories. Data for the number of employees in each size category is estimated using an assumption of 10 employees for establishments with 1 to 19 employees, 109.5 for establishments with 20 to 199, and 200 employees for establishments with over 200 employees.	http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/8165.0Jun%202003%20to%20Jun%202007?OpenDocument
Bangladesh	BGD	2001 & 2003	Data is from Bangladesh's Economic Census of 2001 & 2003.	na	non-agriculture and manufacturing	Data includes the total number of establishments and employees.	http://discovery.lib.harvard.edu/?itemid=|library/m/aleph|011346898
Benin	BEN	1980 & 2002	Data is from the National Institute of Statistics and Economic Analysis, Recensement des entreprises, Les Entreprises Artisanales 1980. And the Troisieme Recensement General de la Population et de L'Habitation 2002.	na	establishments data excludes the agriculture sector, employee data includes all sectors	Data includes the total number of establishments from 1980, and total number of employees from 2002. Because the regions of Benin were split in 1999, data from 1980 is disaggregated (establishments for each pre-1999 region are allocated to the two new 1999 sub regions pro rata based on formal sector employees).	http://www.insae-bj.org/
Bosnia	BIH	2009	Data is from the Federal Bureau Of Statistics, Cantons in numbers.	na	non-agriculture and manufacturing	Data includes the total number of establishments and employees.	http://www.bhas.ba/new/
Brazil	BRA	2006	Data from the Brazilian Central Registry of Businesses.	na	manufacturing [Indústrias de transformação]	Data includes the total number of establishments and employees.	http://www.ibge.gov.br/english/default.php
Burkina Faso	BFA	1998	Data is from the National Institute of Statistics and Demography, Recensement Industriel & Commercial.	na	excludes agriculture and hunting	Data includes the total number of establishments.	http://www.insd.bf/fr/
Cambodia	KHM	2009	Data is from the Preliminary Results of the Nation-wide Establishment Listing of Cambodia 2009.	na	all sectors	Data includes the total number of establishments.	http://www.nis.gov.kh/
Cameroon	CMR	1986	Data is from the Institut National de la Statistique, Recensement Industriel de 1986.	na	maufacturing	Data includes the total number of establishments and employees.	http://www.statistics-cameroon.org/
Canada	CAN	2010	Data is from Statistics Canada, CANSIM, table 281-0044 and Catalogue no. 72-002-X	100 Employees	all sectors	Data includes number of employees by seven size categories. Data for the number of establishments in each size category is estimated using an assumption of 2 employees for establishments with 0 to 4 employees, 12 for establishments with 5 to 19, 34.5 for establishments with 20 to 49 employees, 74.5 employees for establishments with 50 to 99, 199.5 for establishments with 100 to 299, 399.5 for establishments with 300 to 499 and 500 employees for establishments with over 500 employees.	http://www40.statcan.gc.ca/l01/cst01/labr83d-eng.htm
China	CHN	2004	Data is from China Online, China Economic Census Yearbook 2004, Enterprise 2, table 2-8 Number of Corporation Legal Person Units by Region and Group Interval of Empolyed Persons and 1-2 Number of Legal Person Units, Establishments and Employed Persons by Region.	100 Employees	all sectors	Data includes the number of establishments by ten size categories. The number of employees in each size category is estimated using the assumption of 3.5 employees for establishment with 0 to 7 employees, 13.5 for establishments with 8 to 19, 34.5 for establishments with 20 to 49, 74.5 for establishments with 50 to 99, 199.5 for establishments with 100 to 299, 349.5 for establishments with 300 to 499, 749.5 for establishments with 500 to 999, 1999.5 for establishments with 1000 to 2999, 3999.5 for establishments with 3000 to 4999, and 5000 employees for establishments with over 5000 employees.	http://chinadataonline.org/member/ecocensus/ybListDetail.asp?ID=1#
Colombia	COL	2001-02	Data is from the National Administrative Department of Statistics (DANE), Recensement Economique 2001-2002 Royaume du Maroc.	200 Employees	total data excludes agriculture; size data appears to include all sectors	Data includes the number of establishments and employees for ten size categories.	http://190.25.231.246:8080/Dane/tree.jsf
Costa Rica	CRI	2008	Data is from the INEC's Directory of Institutional Units and Establishments. [INEC. Directorio de Unidades Institucionales y Establecimientos.]	100 Employees	all sectors	Data includes the number of establishments and employees by five size categories.	http://www.inec.go.cr/Web/Home/pagPrincipal.aspx
Croatia	HRV	2008	Data is from the Central Bureau of Statistics, Business Entities According To Nkd 2002. By Counties at 31 December 2008, and Employed By The Activity And Counties at 31 March 2008. [Poslovni Subjekti Prema Nkd-U 2002, Po Županijama Stanje 31. Prosinca 2008. Zaposleni Po Područjima Djelatnosti Županijama Stanje 31. Ožujka 2008.]	na	non-agriculture and manufacturing	Data includes the total number of employees and establishments (excludes establishments with no employees).	http://www.dzs.hr/
Cuba	CUB	2008	Data is from the National Statistics Office of Cuba, the Provincial Statistical Yearbook 2008.	na	all sectors	Data includes the total number of business entities, coorperatives and other organisations and employees.	http://www.one.cu/
Denmark	DNK	2007	Data is from the Statistical Office of Denmark, General economic statistics, Business Demography, table GF4: General enterprise statistics by region, industry (DB07, 10-grouping) and unit	na	non-agricultural	Data includes the total number of enterprises and employees. Enterprises are used in place of establishments here.	http://www.statbank.dk/statbank5a/default.asp?w=1680
Dominican Republic	DOM	2002	Data is from the National Statistics Office of the Dominican Republic, Perfil Sociodemográfico Provincial, Censo Nacional de Población y Vivienda VIII.	na	excludes agriculture, fishing and forestry	Data includes the total number of employees.	http://www.one.gob.do/index.php?module=articles&func=view&catid=217
Ecuador	ECU	2007	Data is from the National Institute of Statistics and Censuses, the results of the Manufacturing and Mining, Table N.17.	na	manufacturing and mining	Data includes the total number of establishments and employees.	http://www.inec.gov.ec/web/guest/publicaciones/anuarios/inv_eco/man_min
Egypt	EGY	2008	Data is from the Central Agency for Public Mobilization and Statistics, Data from Egypt in Figures, table Estimates of Employed Persons (15 years old and over) By Sex, Governate.	na	all sectors	Data includes the estimated total number of employees age 15 years old and older.	http://www.msrintranet.capmas.gov.eg/ows-img2/htms/pdf/work/25.pdf
El Salvador	SLV	2005	Data is from the Ministerio de Economia Direccion General de Estadistica y Censos, VII Censos Economicos.	5 Employees	non-agriculture and manufacturing	Data includes the total number of establishments and employees.	http://www.minec.gob.sv/
Estonia	EST	2009	Data is from the National Statistical Office of Mongolia, Statistical database: Economy, Entities, Business demography, ER28: Affiliates Statistical Profile By County And Number Of Employees. [Statistika andmebaas: Majandus, Majandusüksused, Ettevõtluse demograafia, ER28: Statistilisse Profiili Kuuluvad Ettevõtted Töötajate Arvu Ja Maakonna Järgi.]	250 Employees	all sectors	Data includes the number of establishments by four size categories. Data for the number of employees in each size category is estimated using an assumption of 5 employees for establishments with less than 10 employees, 30 for establishments with 10 to 49, 150 for establishments with 50 to 249, and 250 employees for establishments with over 250 employees.	http://pub.stat.ee/px-web.2001/Dialog/varval.asp?ma=ER28&ti=STATISTILISSE+PROFIILI+KUULUVAD+ETTEV%D5TTED+T%D6%D6TAJATE+ARVU+JA+MAAKONNA+J%C4RGI+&path=../Database/Majandus_regionaalne/10Majandusuksused/02Ettevetjad/&lang=2
Georgia	GEO	2010	Data is from the National Statistics Office of Georgia, special request.	100 Employees	non-agriculture	Data includes the number of establishments and employees above and below the 100 employee size threshold.	http://www.geostat.ge/
Ghana	GHA	2003-05	Data is from the Ghana Statistical Service's 2003 National Industrial Census, Phase I and II, Table 10: Establishments by Size and Region and Table 11: Persons Engaged by Size and Region.	100 Employees	industrial sectors	Data includes the number of establishments and employees by nine size categories.	http://www.statsghana.gov.gh/Industrial_Census.html

Country	Code	Year	Source	Big Firms	Industry	Calculations	Links
India	IND	2005	Data is from the Government of India Ministry of Statistics and Programme Implementation, Economic Census 2005, tables: Tables A 3.4, A 4.4, A 3.6, A-3.22, A 4.6, and A 4.22.	na	non-agriculture and manufacturing	Data includes the total number of establishments in non-agriculture and manufacturing sectors and employees in non-agriculture employees.	http://mospi.nic.in/Mospi_New/upload/economic_census_2005/index_6june08.htm
Iran	IRN	2007	Data is from the Statistical Centre of Iran, Iran Manufacturing by Ostan 2006 and 2007, table 7.5. Manufacturing Establishments By Ostan And Type Of Ownership: 1384 and 7.8. Number Of Workers In Manufacturing Establishments By Ostan And Size: 1384.	100 Employees	manufacturing	Data includes the number of establishments and employees by three size categories.	http://www.amar.org.ir/Default.aspx
Israel	ISL	2008	Data is from the Central Bureau of Statistics, Israel in Figures 2008.	na	all sectors	Data includes the total number of establishments and employees.	http://www1.cbs.gov.il/reader/cw_usr_view_Folder?ID=141
Italy	ITA	2008	Data is from the National Statistical Institute, 8th General Census of Industry and Services October 2001, table 8.	250 Employees	non-agriculture and manufacturing	Data includes the total number of establishments and employees in the manufacturing sector and non-agriculture sectors. The data also includes non-agriculture employees by four size categories.	http://dwcis.istat.it/cis/index.htm
Japan	JPN	2006	Data is from the online Portal Site, Official Statistics of Japan, the Establishment and Enterprise Census of Japan, Results of Establishments for Japan, table 6: Establishments and Employees by Sex; by Industry (Major Groups) and Size of Employees (10 Groups); for Japan, Prefectures, 16 Major Cities, 14 Major Metropolitan Areas and Special Areas.	100 Employees	manufacturing	Data includes the number of establishments and employees by ten size categories (exludes establishments with no employees).	http://www.e-stat.go.jp/SG1/estat/ListE.do?bid=000001008300&cycode=0
Jordan	JOR	1998 & 2006	Data is from the Department of Statistics (DOS), Employment Survey 1998 and the Establishment Census 2006	na	non-agricultural	Data includes the total number of establishments and employees.	http://www.dos.gov.jo/dos_home_e/main/index.htm
Kazakhstan	KAZ	2010	Data is from the Agency of Statistics Regions of Kazakhstan Brochure.	100 Employees	all sectors	Data includes the number of establishements by three size categories. The number of employees in each size category are calculated with the assumption of 25 employees for small establishments, 75 employees for medium establishments and 100 employees for large establishments.	http://www.stat.kz/Pages/default.aspx
Kyrgyz Republic	KGZ	2008	Data is from the National Statistical Committee of Kyrgyz Republic, Key indicators (Statistical Yearbook Kyrgyz Republic 2004-2009), "Number of small and medium-sized enterprises on the territory" and "The number of workers in small and medium-sized enterprises on the territory".	na	all sectors	Data includes the total number of establishments and employees for small and medium size businesses with foreign investment.	http://www.stat.kg/rus/part/msp.htm
Lao PDR	LAO	2004	Data is from the Ministry of Industry-Handicraft	99 Employees	all sectors	Data includes the number of establishments by three size categories (establishments with no employees are excluded). The number of employees in each size category is estimated using an assumption of 5 employees for establishment with 1 to 10 employees, 55 for establishments 10 to 99, and 100 employees for establishments with more than 99 employees.	http://www.nsc.gov.la/Statistics/Selected%20Statistics/Industry.htm
Latvia	LVA	2003	Data is from the Central Statistical Bureau "Statistikas uzņēmumu reģistrs, Tabula: SR09. EKONOMISKI AKTĪVIE UZŅĒMUMI UN UZŅĒMĒJSABIEDRĪBAS, STATISTIKAJOS REĢIONOS, PILSĒTĀS UN RAJONOS, 1997. - 2003.g." [Statistical Business Register, Table SR09. Economically Active Enterprises And Business Companies, Statistical Regions, Cities And Districts, 1997. - 2003]	250 Employees	non-agriculture	Data includes the number of establishments in three size categories. The number of employees is estimated using the assumption of 25 employees for establishments with 0 to 49 employees, 149.5 for establishments with 50 to 249, and 250 employees for establishments with over 250 employees.	http://data.csb.gov.lv/DATABASE/rupnbuvn/Ikgad%C4%93jie%20statistikas%20dati/Statistikas%20vien%C4%ABbu%20re%C4%A3istrs/Statistikas%20vien%C4%ABbu%20re%C4%A3istrs.asp
Lebanon	LBN	2004	Data is from the Lebanese Census of Buildings Dwellings and Establishments 2004.	100 Employees	non-agriculture and all sectors	Data includes the number of establishments from all sectors by six size categories. The number of employees with over 100 employees is estimated with the assumption of 100 employees. The total number of establishments and employees exclude the agriculture sector.	http://www.cas.gov.lb/index.php?option=com_content&view=frontpage&Itemid=28
Lithuania	LTU	2010	Data is from the Statistics Lithuania, Database of Indicators, Business statistics, Small and medium sized enterprises in operation, Table M4010241: Number of enterprises in operation and number of employees at the beginning of the year by administrative territory, size class of enterprises	100 Employees	excluding agriculture, forestry and fishing; and all sectors	Data includes the number of establishments and employees by ten size categories for all sectors. The total number of employees and establishments exclude agriculture, forestry and fishing sectors.	http://db1.stat.gov.lt/statbank/SelectTable/omrade0.asp?SubjectCode=S4&PLanguage=1&Shownews=OFF&tree=false
Macedonia, FYR	MKD	2009	Data is from the State Statistical Office of the Republic of Macedonia, Statistical Databases, Regional Statistics, Business Entities.	50 Employees	all sectors	Data includes the number of establishments by six size categories. The number of employees in each size category is estimated using the assumption of 5 employees for establishments with 1 to 9 employees, 24.5 for establishments with 10 to 49, 149.5 for establishments with 50 to 249 and 250 employees for establishements with over 250 employees.	http://www.stat.gov.mk/PXWeb2007bazi/Database/Regional%20Statistics/databasetree.asp
Madagascar	MDG	2003	Data is from INSTAT, Recensement au niveau des Communes.	na	non-agricultural	Data includes the number of establishments.	http://www.instat.mg/pdf/rgph_6.pdf
Malaysia	MYS	2005	Data is from the Department of Statistics, Economic Census of the manufacturing sector.	na	manufacturing	Data includes the number of total establishments and Engaged Persons. Here the number of Engaged Persons is used as a proxy for employees.	http://www.statistics.gov.my/portal/index.php?lang=en
Mexico	MEX	2009	Data is from the National Institute of Statistics, Geography, and Informatics, Censos Económicos 2009.	na	excludes agriculture, livestock and forestry	Data includes the number of establishments and employees by twelve size categories.	http://www.inegi.org.mx/est/contenidos/espanol/proyectos/censos/ce2009/default.asp?s=est&c=14220
Moldova	MDA	2008	Data is from the National Bureau of Statistics, Territorial statistics, Principalii Indicatori Ai Activității Întreprinderilor Industriale, În Profil Teritorial.	na	manufacturing	Data includes the total number of employees.	http://www.statistica.md/pageview.php?l=en&idc=349&
Mongolia	MNG	2007	Data is from the Businesss Register of Mongolia table 11.3 Number of Active Legal Units, by Aimags and the Capital, Employment Size Class,	50 Employees	all sectors	Data includes the number of establishments by four size categories. The number of employees for each size category is estimated using the assumption of 5 employees for establishments with 1 to 9 employees, 14.5 for establishments with 10 to 19, 34.5 for establishments with 20 to 49, and 50 employees for establishments with over 50 employees.	http://www.statis.mn/v3/index2.php?page=free_access
Morocco	MAR	2001-02	Data is from the Department of Statistics, Economic Census, TABLEAU 1 : Répartition des établissements et de l'effectif d'emploi selon les régions et les secteurs d'activités	na	non-agriculture and manufacturing	Data includes the total number of establishments and employees in the manufacturing sector and non-agriculture sectors.	http://www.recensement-eco.hcp.ma/article.php3?id_article=11
Mozambique	MOZ	2004	Data is from the National Institute of Statistics, CEMPRE, 2004, table 5, 14 and 21. African Development Bank & African Development Fund, Mozambique Private Sector Country Profile August 2008, table 5.	100 Employees	all sectors	Data includes the total number of establishments by two size categories. The number of employees in each size category is estimated using a national total and regional averages for each size category.	http://www.ine.gov.mz/
Nepal	NPL	2006-07	Data is from the Central Bureau of Statistics, Census of Manufacturing Establishments 2006/2007.	100 Employees	maufacturing	Data includes the number of establishments and employees by six size categories. Data excludes establishments with fewer than 10 employees.	http://www.cbs.gov.np/
New Zealand	NZL	2002 [1997 update]	Data is from Statistics New Zealand, New Zealand Business Demographic Statistics, Table 2.	100 Employees	non-agriculture	Data includes the number of Enterprises and FTE Engaged Persons by five size categories. Enterprises are used as a proxy for establishments and FTE Engaged Persons are used as a proxy for employees. Data only reflects Enterprises that meet certain significance tests including minimum revenue or two or more employees. Data has been adjusted to estimate 2002, and was originally collected from a 1997 survey.	http://www.stats.govt.nz/browse_for_stats/businesses/business_characteristics/nz-business-demography-stats-std-tables.aspx
Norway	NOR	2010	Data is from Statistics Norway, table 2 Establishments, by size groups and county.	100 Employees	all sectors	Data includes the number of establishments by eight size categories. The number of employees in each size category are estimated using an assumption of 2.5 employees for establishments with 1 to 4 employees, 7 for establishments with 5 to 9, 14.5 for establishments with 10 to 19, 34.5 for establishments with 20 to 49, 74.5 for establishments with 50 to 99, 174.5 for establishments with 100 to 249, and 250 employees for establishments with over 250 employees.	http://www.ssb.no/english/subjects/10/01/bedrifter_en/tab-2010-01-29-02-en.html

Country	Code	Year	Source	Big Firms	Industry	Calculations	Links
Pakistan	PAK	2005	Data is from the Statistics Division, Federal Bureau of Statistics Division, Economic Census 2005.	100 Employees	excludes agriculture, forestry, hunting and fishing	Data includes the number of establishments by nine size categories. The number of employees in each size category is estimated using the assumption of 125.5 employees for establishments with 101 to 150 employees, 175.5 for establishments with 151 to 200, 225.5 for establishments with 201 to 250, 375.5 for establishments with 251 to 300 and 300 employees for establishments with over 300 employees.	http://www.statpak.gov.pk/depts/fbs/publications/ec_2005/ec_2005.html
Panama	PAN	2002	Data is from the Census and Statistics Directorate, V Censos Nacionales Economicos.	na	non-agriculture	Data includes the total number establishments and employees.	http://www.contraloria.gob.pa/
Paraguay	PRY	2002	Data is from the General Directorate of Statistics, Surveys and Censuses, Resultados Preliminares de la Encuesta Industrial.	na	manufacturing	Data includes total number of establishments.	http://www.dgeec.gov.py/
Peru	PER	1993-94	Data is from the National Institute of Statistics and Informatics, III Censo Nacional Económico 1993-1994 (CENEC), Peru: Numero De Establecimientos Censados E Informantes, Por Estrato De Personal Ocupado; Segun Departamento (Composicion Porcentual).	20 Employees	non-agriculture and manufacturing	Data includes number of establishments by three size categories, for non-agricultural sectors and appears to exclude establishments with no employees. The number of establishments with 20 or more employees is estimated using the assumption of 20 employees. Data also includes total establishments and employees for the non-agriculture and manufacturing sectors.	http://www.inei.gob.pe/
Philippines	PHL	2006	Data is from the 2006 Census of Philippine Business and Industry, by special request from the National Statistics Office, Republic of the Philippines.	20 Employees	non-agricultural	Data includes the number of establishments and employees by two size categories.	http://www.census.gov.ph/data/sectordata/databusind.html
Russia	RUS	2008	Data is from the Russian State Committee for Statistics, Socio-Economic Indicators 2009.	na	non-agriculture and manufacturing	Data includes establishment and employee totals for non-agriculture sectors and the manufacturing sector.	http://www.gks.ru/eng/
Senegal	SEN	2005	Data is from the National Agency of Statistics and Demography, Service Regional de la Statistique et de la Demographie.	na	non-agriculture	Data includes the total number of establishments and employees.	http://www.ansd.sn/
Serbia	SRB	2007	Data is from the Statistical Office of the Republic of Serbia, Basic Results of Business Activity of Enterprises and Entrepreneurs.	na	non-agriculture and manufacturing	Data includes the total number of establishments and employees.	http://webrzs.stat.gov.rs/axd/en/
Slovakia	SVK	2009	Data is from the Statistical Office of the Slovak Republic.	na	manufacturing	Data includes the total number of establishments and employees.	http://px-web.statistics.sk/PXWebSlovak/index_en.htm
South Africa	SFA	2007	Data is from Statistics South Africa, Labour Force Survey, table Labour market indicators (working-age population, 15-64 years) by province.	na	all sectors	Data inclues the total number of employees from a Labor Force survey.	http://www.statssa.gov.za/publications/P0210/P0210September2000,2001,2002,2003,2004,2005,2006,2007.pdf
South Korea	KOR	2004	Data is from Statistics Korea, Report of the Census on Establishments, table By province, industrial classification and type of legal organization.	na	non-agricultural	Data includes the total number of establishments and employees.	http://kostat.go.kr/nso_main/nsoMainAction.do?method=sub&catgrp=eng2009&catid1=g03&catid2=g03a&catid3=g03ac&catid4=g03acq
Sri Lanka	LKA	2003	Data is from the Department of Census and Statistics - Sri Lanka, Census of Industry 2003/2004, Table A : No.of Manufacturing Establishments and Persons Engaged by District and Type of Industry Scale.	10 Employees	manufacturing	Data includes the number of establishments and employees by two size categories.	http://www.statistics.gov.lk/industry/census%20of%20industries_2004.pdf
Switzerland	CHE	2005	Data from the The Portal Statistics Switzerland by the Federal Statistical Office (FSO), Recensement des entreprises, Structure économique, table 7. Résultats pour les grandes régions et les cantons, Etablissements and Emplois.	na	non-agricultural	Data includes the total number of establishments and employees.	http://www.bfs.admin.ch/bfs/portal/fr/index/infothek/publ.html
Syria	SYR	2007	Data is from the Central Bureau of Statistics, workers and industry statistical tables.	na	all sectors	Data includes the total number of establishments and employees.	http://www.cbssyr.org/work/2009/semi-1/TAB13.htm
Tanzania	TZA	2007	Data is from National Bureau of Statistics, Central Register of Establishments, Business Survey Tanzania Mainland Report.	100 Employees	excludes agriculture, hunting and forestry	Data includes the number of establishments and employees by eight size categories.	http://www.nbs.go.tz/index.php?option=com_phocadownload&view=category&id=72:industry&Itemid=106
Thailand	THA	2007	Data is from the National Statistical Office 2007 Industrial Census.	200 Employees	manufacturing	Data includes the number of establishments numbers by six size categories. The number of employees in each size category is estimated with the assumtion of 8 employees for establishments with 1 to 15 employees, 20.5 for establishments with 16 to 25, 28 for establishments with 26 to 30, 40.5 for establishments with 31 to 50, 125.5 for establishments with 51 to 200, and 200 employees for establishments with over 200 employees. The total number of employees are actual data and not estimated. Data excludes establishments with no employees. Establishments with 10 and fewer employees were sampled, establishments with over 11 employees were surveyed.	
Turkey	TUR	1992	Data is from Turkey's Business Statistics, General Census of Industry and Business Establishments, 2002 General Census Of Industry And Establishments, 1.1.2. Number of establishments, annual average number of persons engaged by provinces and economic activity branches.	na	non-agricultural	Data includes the total number of establishments and Persons Engaged. Persons Engaged is used as a proxy for employees.	http://www.turkstat.gov.tr/VeriBilgi.do?tb_id=28&ust_id=9
Uganda	UGA	2006	Data is from the Uganda Bureau of Statistics, Report on the Uganda Business Register 2006/7, Table 3.4.4 Distribution of Businesses by Region by Employment Size band.	100 Employees	all sectors	Data includes the number of establishments by six size categories and excludes establishments with no employees. The number of employees in each size category is estimated with the assumption of 2.5 employees for establishments with 1 to 4 employees, 7 for establishements with 5 to 9, 14.5 for establishments with 10 to 19, 34.5 for establishements with 20 to 49, 74.5 for establishments with 50 to 99, and 100 employees for establishements with over 100 employees.	
Ukraine	UKR	2008	Data is from the State Statistics Committee Of Ukraine, Activity of Enterprises: Statistical Publication 2008, Table 8.7 Number of employees the size of the enterprise, by regions, Table 8.3 Number of enterprises per 10000 of present population the size of the enterprise, by regions, and Table 8.2 Share of large, medium-sized and small enterprises, by regions.	50 Employees	all industries	Data includes the number of employees in three size categories. The number of establishments in each region and size category is implied from the number of enterprises per 10000 of present population (Table 8.3), the 2008 regional population (Statistical publication Regions of Ukraine 2009, Table 3.1) and the percentage share of firms by size and by region (Table 8.2).	http://www.ukrstat.gov.ua/
United Arab Emirates	ARE	1995	Data from the United Arab Emirates Statistical Abstract, Chapter 11 Economical Establishments.	na	all sectors, and excluding agriculture forestry and hunting	Data includes the number of establishments and employees by ten size categories in all sectors. The total number of establishments and employees excludes agriculture, forestry and hunting sectors.	http://www.economy.ae/English/EconomicAndStatisticReports/StatisticReports/StatisticAbstract/Pages/sa2007.aspx
United States	USA	2006	Data is from the U.S. Census Bureau, Statistics of U.S. Businesses.	100 Employees	all sectors and non-agriculture	Data includes the total number of establishments and employees by twelve size categories for all sectors. The total number of establishments and employees exclude agriculture.	http://www.census.gov/epcd/www/smallbus.html
Vietnam	VNM	2007	Data is from the General Statistics Office of Vietnam, Results of the 2007 Establishment Census, Table 39 Number of individual business establishments by size of employee, by size of employee.	101 Employees	non-agricultural	Data includes the number of establishments by nine size categories. Number of employees in each size category is estimated with the assumption of 3.5 employees for establishments with 2 to 5 employees, 8 for establishments with 6 to 10, 15.5 for establishments with 11 to 20, 35.5 for establishments with 21 to 50, 75.5 for establishments with 51 to 100, 150.5 for establishments with 101 to 200, 350.5 for establishments with 201 to 500, and 500 for establishments with over 500 employees.	http://www.gso.gov.vn/default_en.aspx?tabid=515&idmid=5&ItemID=9359