

## APPENDIX D – VARIABLE SELECTION

This paper's primary purpose is to identify which factors determine the difference between nighttime lights and official GDP data in the form of error terms. To be more specific, we aim to explore which hidden explanatory variables play the most important role to explain the error terms. In panel data analysis, we used the traditional method to select variables. Variables are chosen based on empirical literature and intuitive observations. However, in cross-sectional analysis, we cannot apply the same strategy. This is because we are faced with two main challenges: 1. We carry out data analysis with the problem of a small number of observations; 2. There are thousands of potential explanatory variables. Therefore, we employ three alternative modern statistical techniques to select variables. They are: 1. LASSO, 2. Spike and Slab, and 3. The Minimax Concave Penalty (MCP).

The linear model:

$$LHS (outcome\ variable) = RHS (control\ variables)$$

Where

Left hand side (LHS) is dependent variable: *mean absolute value of error terms*. (Grouped by countries)

Right hand side (RHS) is a vector of 597 independent variables.

The data are collected from multiple secondary sources, including World Development Indicator (WDI), The Quality of Government (QoG), Freedom House, The WhoGov dataset (Nyrup & Bramwell, 2020), Varieties of Democracy, Center of Systemic Peace. To run Lasso, the Minimax Concave Penalty and Spike and Slab, we need to process the data set beforehand. While the original dataset has more than 2000 variables, our respond variable has only 179 countries (observations) globally. In the context of high dimensional data, the major issue to deal with is missing values. We processed data by the following steps:

- In the first step, we delete all variables that have less than 150 observations. There are still 574 variables left. Now, the average missing rate is approximately 6% (about 10 out of 179 observations). Therefore, we fill missing values with 0. This will not affect the performance of the statistical methods. We also delete some outliers of outcome variables (7 observations)

- Next step, the data set includes both continuous and categorical variables. Therefore, we break each categorical variable into dummy variables. This action increases the data set from 574 variables to 596 variables.
- Finally, to cope with the fair penalization issue, we standardize all variables.

Processed data will be used for all three methods.

## RESULTS

### 1. LASSO

#### *Selecting a value for turning parameter $\lambda$*

At the first step, we begin our analysis by determining the most appropriate value of the tuning parameter  $\lambda$  using Leave-One-Out Cross-Validation. Figure D1 represents a plot of cross-validated mean squared error (MSE) according to various values of  $\lambda$  (by default, the glmnet package in R uses the natural log of  $\lambda$ ). The two vertical dotted lines show the min MSE (left) and min MSE + 1 standard error (right). It can be seen that for larger values of the regularization parameter  $\lambda$ , the MSE is smaller. By contrast, as values of  $\lambda$  decrease, the MSE becomes larger. In addition to this, the plot seems to flatten as  $\lambda$  begins increasingly large, while the MSE becomes relatively stable low (see figure B1 as  $\log(\lambda) > -2$  )

For the lasso estimator, the most appropriate value of  $\lambda$  is the one that gives the minimum MSE value. In this context, the minimum value of MSE is 0.9842, corresponding to the value of  $\lambda = 0.2267$  ( $\log[0.2267] = -1.4841$ ). For this value of  $\lambda$ , lasso selects 5 out of 596 predictors (see column 1 in table D1). However, as displays by figure D1, the plot shows a slight change in MSE values as  $\log(\lambda) > -2$ . Therefore, in this case,  $\lambda_{min}$  is not truly visible. In order to secure we do not omit any important determinants, we select  $\lambda = 0.1353$  ( $\log[0.1353] = -2$ ). For this value of  $\lambda$ , 21 of the 596 predictors were chosen by the model (see column 2 in table D1).

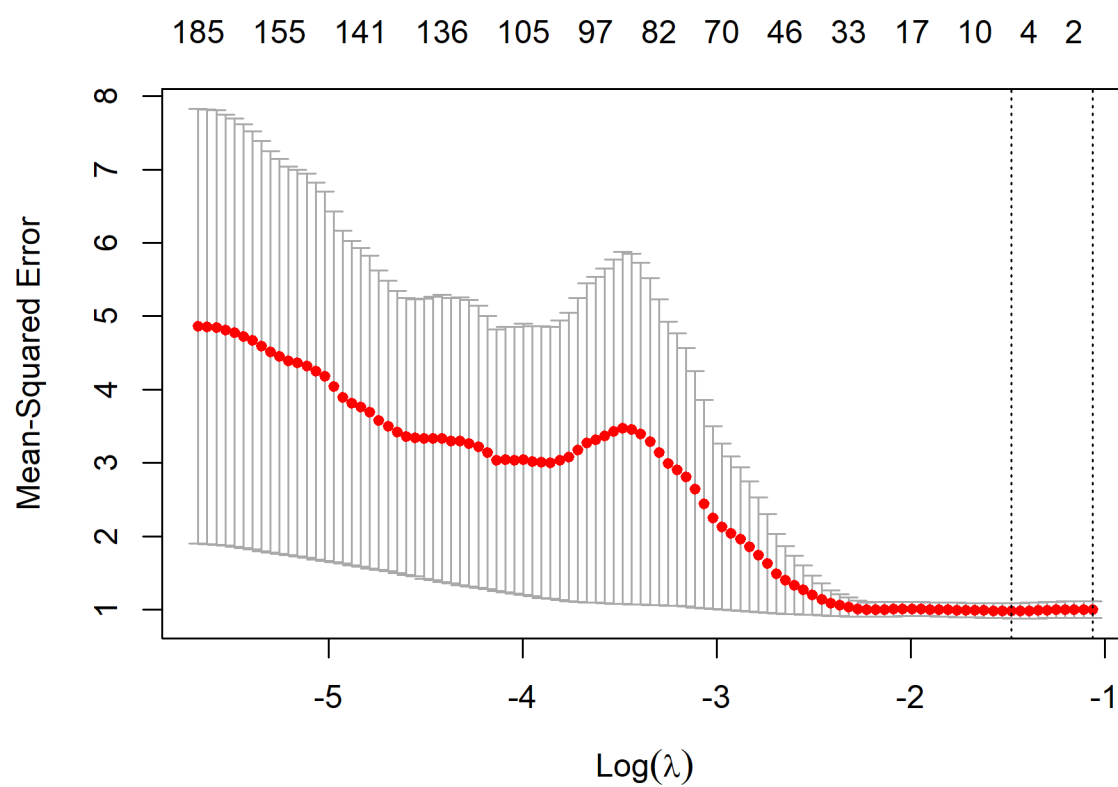


Figure D1. Plot of cross-validated mean squared error (MSE)

**Table D1.**  
**List of variables that Lasso selects**

| Variable                                                                      | $\lambda_{min} = 0.2267$<br>(1) | $\lambda = 0.1353$<br>(2) |
|-------------------------------------------------------------------------------|---------------------------------|---------------------------|
| Legal origin                                                                  |                                 | ✓                         |
| Threatened Species: Fishes (bi_fishes)                                        |                                 | ✓                         |
| State Fragility Index (cspf_sfi)                                              |                                 | ✓                         |
| Starting a business (eob_sab)                                                 | ✓                               | ✓                         |
| Arable Land (% of Agricultural land) (fao_luagrara)                           |                                 | ✓                         |
| Parliamentary Election: Compulsory Voting (ideavt_legcv)                      |                                 | ✓                         |
| Natural Resource Protection Indicator (nrmi_nrpi)                             |                                 | ✓                         |
| Average distance to nearest ice-free coast (1000 km.)                         |                                 | ✓                         |
| Estimated variance of the Property Rights Protection point estimate (prp_std) |                                 | ✓                         |
| Share of government consumption at current PPPs (pwc_gc)                      |                                 | ✓                         |
| Top 10% income share (top_top10_income_share)                                 |                                 | ✓                         |
| Official entrance age to compulsory education (years) (une_oace)              |                                 | ✓                         |
| Public sector corruption index (vdem_pubcorr)                                 |                                 | ✓                         |
| Voice and Accountability (wbgi_vae)                                           | ✓                               |                           |
| GDP growth (annual %) (wdi_gdpagr)                                            |                                 | ✓                         |
| School enrollment, preprimary, female (% gross) (wdi_gerppf)                  | ✓                               | ✓                         |
| Number of years the leader in office continuously in 2010 (wgov_leadexp)      | ✓                               | ✓                         |
| Official Religion (biu_offrel)                                                |                                 | ✓                         |
| Regime Type (ht_regtype1)                                                     |                                 | ✓                         |
| Perception Of Electoral Integrity Index (Pei_Peit)                            |                                 | ✓                         |
| Freedom status                                                                | ✓                               | ✓                         |

## 2. The Minimax Concave Penalty (MCP)

In this section, we describe how we perform the Minimax Concave Penalty regression. We used R-software package “*ncvreg*” for implementing the MCP variable selection methodology.

In the first step, we perform cross-validation to check the predictive accuracy of the model at various values of  $\lambda$ . Figure D2 represents a plot of cross-validated mean squared error (MSE) according to various values of  $\lambda$ . By using `cvfit$lambda.min` command, we can determine the value of  $\lambda$  that minimizes cross-validation errors. In this case, it is 0.2325423 ( $\log[0.2325423] = -1.4587$ ). For this value of  $\lambda$ , model selects 2 out of 596 predictors (see column 1 in table D2). However, as displays by figure D2, the plot shows a slight change in MSE values as  $\log(\lambda) > -2$ . Therefore, in this case,  $\lambda_{min}$  is not truly visible. In order to secure we do not omit any important determinants, we select  $\lambda = 0.1353$  ( $\log[0.1353] = -2$ ). For this value of  $\lambda$ , 13 of the 596 predictors were chosen by the model (see column 2 in table D2).

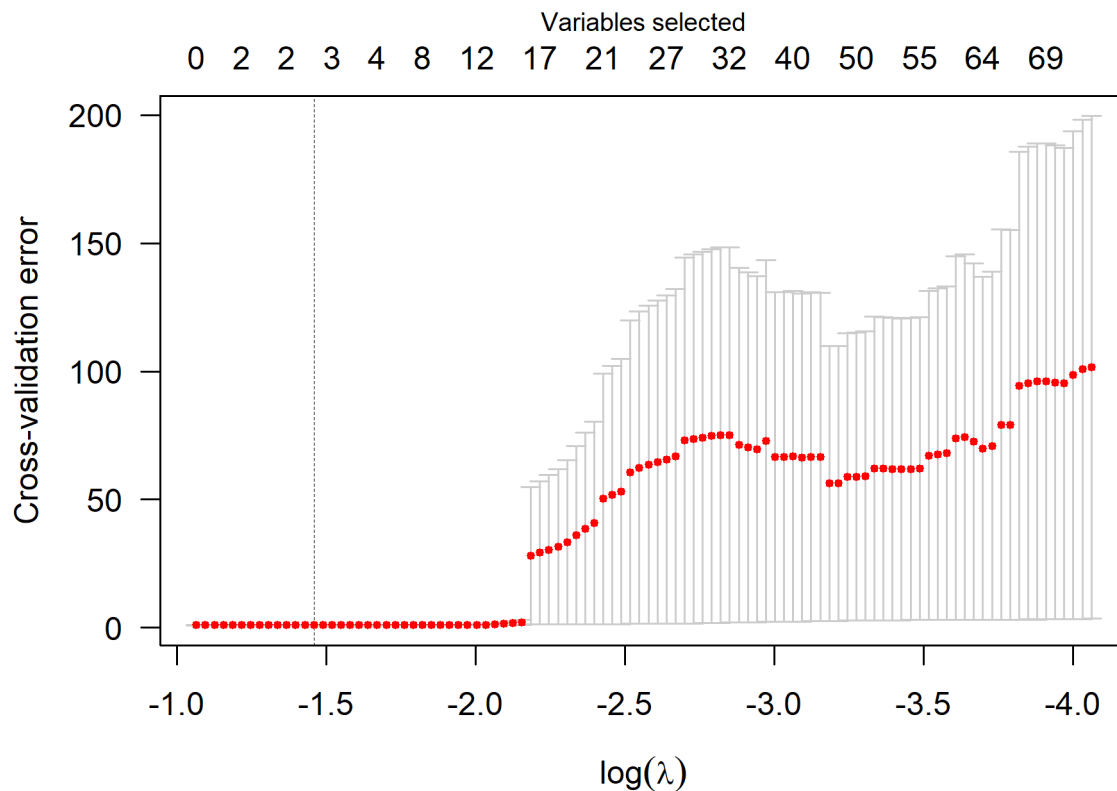


Figure D2. Plot of cross-validated mean squared error (MSE)

**Table D2.****List of variables that the Minimax Concave Penalty (MCP) selects**

| Variable                                                                 | $\lambda_{min}$<br>= 0.2325423<br>(1) | $\lambda = 0.1353$<br>(2) |
|--------------------------------------------------------------------------|---------------------------------------|---------------------------|
| Voice and Accountability, Estimate (wbgi_vae)                            | ✓                                     | ✓                         |
| Number of years the leader in office continuously in 2010 (wgov_leadexp) | ✓                                     | ✓                         |
| Starting a business (eob_sab)                                            |                                       | ✓                         |
| Public sector corruption index (vdem_pubcorr)                            |                                       | ✓                         |
| Share of government consumption at current PPPs (pwt_gc)                 |                                       | ✓                         |
| Official entrance age to compulsory education (years) (une_oeace)        |                                       | ✓                         |
| Regime Type (ht_regtype1)                                                |                                       | ✓                         |
| Legal origin                                                             |                                       | ✓                         |
| Threatened Species: Fishes (bi_fishes)                                   |                                       | ✓                         |
| Parliamentary Election: Compulsory Voting (ideavt_legcv)                 |                                       | ✓                         |
| Natural Resource Protection Indicator (nrmi_nrpi)                        |                                       | ✓                         |
| Average distance to nearest ice-free coast (1000 km.)                    |                                       | ✓                         |
| Agriculture, forestry, and _shing, value added (% of GDP) (wdi_gdpagr)   |                                       | ✓                         |
| Perception Of Electoral Integrity Index (Pei_Peit)                       |                                       | ✓                         |

The console from R studio:

```

3. MCP-penalized linear regression with n=172, p=597
4. At lambda=0.1353:
5. -----
6.   Nonzero coefficients           :   14
7.   Expected nonzero coefficients:  11.25
8.   Average mfdR (14 features)   :   0.803
9.
10.      Estimate      z      mfdR Selected
11. V560  0.1957412   3.957 0.058499      *
12. V172 -0.1657902  -3.660 0.180822      *
13. V378  0.0042798   2.039 0.688884      *
14. V304  0.0094862   2.101 0.760884      *
15. V343  0.0970584   2.979 0.770462      *
16. V578  0.0959430   2.968 0.786958      *
17. V64   0.0457330   2.471 1.000000      *
18. V95   -0.0825514  -2.835 1.000000      *
19. V279 -0.0002399  -1.962 1.000000      *
20. V291 -0.0206901  -2.222 1.000000      *
21. V293  0.0462622   2.476 1.000000      *
22. V443  0.0223186   2.239 1.000000      *
23. V576 -0.0020520  -2.038 1.000000      *
24. V584 -0.0106137  -2.121 1.000000      *

```

### 3. Spike and Slab Regression

In this section, we describe how we perform Spike and Slab regression technique. We used R-software package “*spikeslab*” for implementing spike and slab variable selection methodology. We follow the guide from Ishwaran et al. (2010) (page 69).

The spike and slab algorithm involves three key steps:

1. Filtering (dimension reduction)
2. Model Averaging (BMA)
3. Variable selection (gnet)

**Step 1** filters all but the top  $nF$  variables, where  $n$  is the sample size and  $F > 0$  is the user specified fraction. Variables are ordered on the basis of their absolute posterior mean coefficient value, where the posterior mean is calculated using Gibbs sampling applied to an approximate rescaled spike and slab posterior. Below is a toy-illustration of how filtering works [ $p$  is the total number of variables and  $(V_{(k)})_{k=1}^p$  are the ordered variables]:

$$\underbrace{V_{(1)}, \dots, V_{([nF])}}_{\text{retain these variables}} \quad \underbrace{V_{([nF]+1)}, \dots, V_{(p)}}_{\text{filter these variables}}$$

The value for  $F$  is set using the option *bigp.smalln.factor*, which by default is set to the value  $F = 1$ . The use of an approximate posterior in the filtering step is needed in

high dimensions. This yields an ultra-fast Gibbs sampling procedure, with each step of the Gibbs sampler requiring  $O(np)$  operations. Thus, computational effort for the filtering step is linear in both dimensions.

**Step 2** fits a rescaled spike and slab model using only those variables that are not filtered in Step 1. Model fitting is implemented using a Gibbs sampler. Computational times are generally rapid as the number of variables at this point are a fraction of the original size,  $p$ . A blocking technique is used to further reduce computational times. The posterior mean of the regression coefficients, which we refer to as the BMA, is calculated and returned. This (restricted) BMA is used as an estimator for the regression coefficients.

**Step 3** calculates the gnet. In the optimization, the gnet's  $\ell_2$ -regularization parameters are fixed (these being determined from the restricted BMA obtained in Step 2) and its solution path with respect to its  $\ell_1$ -regularization parameter is calculated using the **lars** R package (Hastie & Efron, 2011) [a package dependency of **spikeslab**]. The lars wrapper is called with type="lar" to produce the full LAR path solution (Efron et al., 2004). The gnet is defined as the model in this path solution minimizing the AIC criterion. Note importantly, that cross-validation is not used to optimize the  $\ell_1$ -regularization parameter. The gnet estimator is generally very stable, a property that it inherits from the BMA, and thus even simple model selection methods such as AIC work quite well in optimizing its path solution. This is different than say the elastic net (Zou & Hastie, 2005) where cross-validation is typically used to determine its regularization parameters (often this involves a double optimization over both the  $\ell_1$ - and  $\ell_2$ -regularization parameters). This is an important feature which reduces computational times in big- $p$  problems.

**Table D3** below reports all variables that the spike and slab algorithm selected.

**Table D3.**  
**List of variables that Spike and Slab algorithm selects**

| Variable                                                                 | Variable selection method |     |
|--------------------------------------------------------------------------|---------------------------|-----|
|                                                                          | Cross validation          | AIC |
| Freedom status                                                           | ✓                         | ✓   |
| Number of years the leader in office continuously in 2010 (wgov_leadexp) | ✓                         | ✓   |
| Threatened Species: Fishes (bi_fishes)                                   | ✓                         | ✓   |

|                                                                        |   |   |
|------------------------------------------------------------------------|---|---|
| Starting a business (eob_sab)                                          | ✓ | ✓ |
| State Fragility Index (cspf_sfi)                                       | ✓ | ✓ |
| Public sector corruption index (vdem_pubcorr)                          | ✓ | ✓ |
| Natural Resource Protection Indicator (nrmi_nrpi)                      | ✓ | ✓ |
| Voice and Accountability (wbgi_vae)                                    | ✓ | ✓ |
| Trading across borders (eob_tab16)                                     | ✓ | ✓ |
| Legal origin                                                           |   | ✓ |
| Agriculture, forestry, and _shing, value added (% of GDP) (wdi_gdpagr) |   | ✓ |
| School enrollment, preprimary, male (% gross) (wdi_gerppm)             |   | ✓ |
| Average distance to nearest ice-free coast (1000 km.)                  |   | ✓ |
| Regime Type (ht_regtype1)                                              |   | ✓ |
| Official entrance age to compulsory education (years) (une_oeace)      |   | ✓ |
| Share of government consumption at current PPPs (pwt_gc)               |   | ✓ |
| Parliamentary Election: Compulsory Voting (ideavt_legcv)               |   | ✓ |
| Perception Of Electoral Integrity Index (Pei_Peit)                     |   | ✓ |

The console from R studio:

```

Variable selection method      : cross-validation
Big p small n                 : TRUE
Screen variables               : TRUE
Fast processing                : TRUE
Sample size                    : 172
No. predictors                 : 597
No. burn-in values             : 2000
No. sampled values             : 2000
K-fold                         : 10
CV mean-squared error          : 0.975 +/- 0.101
Model size                     : [1,10]

```

Top variables in terms of stability:

|       | bma    | bma.cv | gnet   | gnet.cv | stability |
|-------|--------|--------|--------|---------|-----------|
| x.591 | 0.020  | 0.016  | 0.089  | 0.038   | 70        |
| x.560 | 0.010  | 0.038  | 0.047  | 0.054   | 70        |
| x.95  | -0.018 | -0.011 | -0.066 | -0.032  | 60        |
| x.172 | -0.007 | -0.022 | -0.038 | -0.057  | 60        |
| x.139 | 0.012  | 0.006  | 0.068  | 0.006   | 50        |
| x.378 | 0.001  | 0.004  | 0.000  | 0.004   | 50        |
| x.291 | -0.007 | -0.004 | -0.020 | -0.010  | 40        |
| x.395 | -0.006 | -0.004 | -0.011 | -0.005  | 40        |
| x.173 | -0.003 | -0.006 | 0.000  | -0.009  | 40        |

-----

Variable selection method : AIC

Big p small n : TRUE

Screen variables : TRUE

Fast processing : TRUE

Sample size : 172

No. predictors : 597

No. burn-in values : 2000

No. sampled values : 2000

Estimated mse : 0.814

Model size : 18

---> Top variables:

|       | bma    | gnet   | bma.scale | gnet.scale |
|-------|--------|--------|-----------|------------|
| x.560 | 0.074  | 0.149  | 0.074     | 0.150      |
| x.591 | 0.029  | 0.069  | 0.029     | 0.069      |
| x.395 | -0.021 | -0.055 | -0.021    | -0.055     |
| x.64  | 0.017  | 0.068  | 0.017     | 0.068      |
| x.95  | -0.015 | -0.064 | -0.015    | -0.064     |
| x.443 | 0.013  | 0.055  | 0.013     | 0.055      |
| x.38  | 0.009  | 0.014  | 0.009     | 0.014      |
| x.378 | 0.007  | 0.015  | 0.007     | 0.015      |

|       |        |        |        |        |
|-------|--------|--------|--------|--------|
| x.173 | -0.007 | -0.016 | -0.007 | -0.016 |
| x.172 | -0.006 | -0.032 | -0.006 | -0.032 |
| x.291 | -0.006 | -0.019 | -0.006 | -0.019 |
| x.458 | -0.005 | -0.009 | -0.005 | -0.009 |
| x.293 | 0.005  | 0.013  | 0.005  | 0.013  |
| x.578 | 0.005  | 0.015  | 0.005  | 0.015  |
| x.343 | 0.005  | 0.016  | 0.005  | 0.016  |
| x.304 | 0.005  | 0.017  | 0.005  | 0.017  |
| x.279 | -0.005 | -0.011 | -0.005 | -0.011 |
| x.584 | -0.004 | -0.018 | -0.004 | -0.018 |

---

## Reference

- Efron, B., Hastie, T., Johnstone, I., & Tibshirani, R. (2004). LEAST ANGLE REGRESSION. In *The Annals of Statistics* (Vol. 32, Issue 2).
- Hastie, T., & Efron, B. (2011). *Least Angle Regression, Lasso and Forward Stagewise*. *R package version 0.9--8*.
- Ishwaran, H., Kogalur, U. B., & Sunil Rao, J. (2010). Spikeslab: Prediction and variable selection using spike and slab regression. *R Journal*, 2(2), 68–73.  
<https://doi.org/10.32614/rj-2010-018>
- Nyrup, J., & Bramwell, S. (2020). Who Governs? A New Global Dataset on Members of Cabinets. *American Political Science Review*, 114(4), 1366–1374.  
<https://doi.org/10.1017/S0003055420000490>
- Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2), 301–320.

## APPENDIX E. CROSS-SECTIONAL ANALYSIS

**Table E1. SUMMARY STATISTICS**

| Statistic                                                                          | N   | Mean         | St. Dev.       | Min      | Max             |
|------------------------------------------------------------------------------------|-----|--------------|----------------|----------|-----------------|
| Mean absolute error                                                                | 172 | 0.5002       | 0.3328         | 0.0367   | 1.4883          |
| Land area (sq. km)                                                                 | 168 | 739,286.9000 | 1,959,270.0000 | 260.0000 | 16,389,950.0000 |
| Average precipitation in depth (mm per year) (1992-2010)                           | 164 | 1,171.9580   | 798.3112       | 18.1000  | 3,240.0000      |
| Average GDP growth (annual %) (1992-2010)                                          | 169 | 3.9886       | 2.8798         | -1.1985  | 28.8889         |
| Average GDP per capita growth (annual %) (1992-2010)                               | 169 | 2.3736       | 2.6418         | -3.2721  | 23.7464         |
| Average Gross Saving (% of GDP) (1992-2010)                                        | 153 | 21.2070      | 10.9392        | -38.1781 | 54.4391         |
| Average Agriculture, forestry, and fishing, value added (% of GDP) (1992-2010)     | 168 | 13.7796      | 12.5876        | 0.1957   | 66.5816         |
| Average Industry (including construction), value added (% of GDP) (1992-2010)      | 167 | 27.9414      | 12.4712        | 5.7734   | 79.7270         |
| Average Services, value added (% of GDP) (1992-2010)                               | 163 | 51.2838      | 11.4423        | 19.3970  | 77.2517         |
| Average Gross capital formation (% of GDP) (1992-2010)                             | 158 | 23.6171      | 6.4512         | 0.0000   | 48.4352         |
| Average General government final consumption expenditure (% of GDP) (1992-2010)    | 157 | 15.8115      | 5.7695         | 3.6291   | 38.0871         |
| Average Households and NPISHs final consumption expenditure (% of GDP) (1992-2010) | 157 | 64.9762      | 17.0667        | 19.3559  | 145.5313        |
| Average Exports of goods and services (% of GDP) (1992-2010)                       | 162 | 37.8095      | 21.1953        | 7.8515   | 139.5818        |
| Average Imports of goods and services (% of GDP) (1992-2010)                       | 162 | 43.2715      | 20.9769        | 10.4965  | 121.3219        |
| Average Population growth (annual %) (1992-2010)                                   | 170 | 1.5352       | 1.3668         | -1.2872  | 7.8143          |
| Average Urban population growth (annual %) (1992-2010)                             | 170 | 2.2909       | 1.9141         | -1.3381  | 8.1490          |
| Average Population density (people per sq. km of land area) (1992-2010)            | 169 | 119.8267     | 173.6390       | 1.5719   | 1,226.6230      |
| Average Population in the largest city (% of urban population) (1992-2010)         | 138 | 32.6612      | 15.8701        | 3.0510   | 76.3782         |
| Average Agricultural land (% of land area) (1992-2010)                             | 169 | 38.9109      | 21.5606        | 0.5263   | 84.8366         |
| Average Forest area (% of land area) (1992-2010)                                   | 169 | 34.0068      | 24.7609        | 0.0000   | 98.3075         |

|                                                                       |     |         |         |           |          |
|-----------------------------------------------------------------------|-----|---------|---------|-----------|----------|
| Average Number of years the leader in office continuously (1992-2010) | 158 | 7.1378  | 6.0729  | 1.0000    | 31.5000  |
| Average Starting a business score (1992-2010)                         | 153 | 66.0830 | 19.0286 | 2.6161    | 96.8563  |
| Average Public sector corruption index (1992-2010)                    | 159 | 0.4827  | 0.2949  | 0.0010    | 0.9614   |
| Average Voice and Accountability, Estimate (1992-2010)                | 170 | -0.0096 | 0.9479  | -1.9397   | 1.5738   |
| Average Control of Corruption: Estimate (1992-2010)                   | 170 | -0.0310 | 0.9722  | -1.4856   | 2.3757   |
| Average State Fragility Index (1995-2010)                             | 152 | 9.3567  | 6.4558  | 0.0000    | 23.3125  |
| Latitude                                                              | 167 | 18.9786 | 24.6669 | -41.8058  | 64.9899  |
| Longitude                                                             | 167 | 16.1595 | 62.9711 | -174.8472 | 171.4777 |
| % Fertile soil                                                        | 167 | 39.0006 | 25.1548 | 0.0000    | 100.0000 |
| Percentage desert in 2012                                             | 169 | 3.5625  | 11.4265 | 0.0000    | 77.2795  |
| Percentage tropical climate in 2012                                   | 169 | 41.2543 | 45.6736 | 0.0000    | 100.0000 |
| Ruggedness (Terrain Ruggedness Index, 100 m) in 2012                  | 169 | 1.3064  | 1.1914  | 0.0029    | 6.7401   |
| Average distance to nearest ice-free coast (1000 km) in 2012          | 169 | 0.2968  | 0.3746  | 0.00001   | 2.2062   |
| Percentage within 100 km. of ice-free coast in 2012                   | 169 | 45.2214 | 39.8073 | 0.0000    | 100.0000 |
| Arable Land (% of Agricultural land) (2017)                           | 169 | 51.5396 | 29.8284 | 0.1200    | 100.0000 |
| Cropland (% of Agricultural land) (2017)                              | 169 | 39.2047 | 25.7388 | 0.1200    | 98.7700  |
| % European descent                                                    | 151 | 33.6827 | 42.3845 | 0.0000    | 100.0000 |
| Ethnic Fractionalization in the year 2000                             | 167 | 0.4401  | 0.2588  | 0.0000    | 0.9302   |
| Language Fractionalization in the year 2000                           | 162 | 0.3924  | 0.2857  | 0.0021    | 0.9227   |
| Religion Fractionalization in the year 2000                           | 169 | 0.4400  | 0.2375  | 0.0023    | 0.8603   |
| Cultural Diversity                                                    | 143 | 0.3062  | 0.2072  | 0.0000    | 0.7328   |
| Ethnic Fractionalization                                              | 144 | 0.4774  | 0.2626  | 0.0040    | 1.0000   |
| Largest Minority                                                      | 136 | 0.1707  | 0.1083  | 0.0100    | 0.4400   |
| Plurality Group                                                       | 143 | 0.6456  | 0.2397  | 0.1200    | 0.9980   |
| Perception of Electoral Integrity Index Type                          | 154 | 3.0779  | 1.3554  | 1.0000    | 5.0000   |
| Legal Origin                                                          | 137 | 1.9124  | 0.9194  | 1.0000    | 5.0000   |
| Number of years the leader in office continuously in 2010             | 157 | 7.2611  | 7.9358  | 1.0000    | 41.0000  |
| Freedom status (2010)                                                 | 171 | 2.2339  | 0.7994  | 1.0000    | 3.0000   |
| The Region of the Country                                             | 172 | 3.5523  | 1.7109  | 1         | 7        |

|                                       |     |         |         |        |          |
|---------------------------------------|-----|---------|---------|--------|----------|
| Colonial Origin                       | 170 | 1.8235  | 1.5360  | 0.0000 | 4.0000   |
| Natural Resource Protection Indicator | 170 | 69.7548 | 32.7591 | 0.6133 | 100.0000 |

**Table E2. Cross-sectional analysis: Institutional quality factors**

|                                            | <i>Dependent variable:</i>                     |                      |                        |                       |                        |                        |                       |                        |
|--------------------------------------------|------------------------------------------------|----------------------|------------------------|-----------------------|------------------------|------------------------|-----------------------|------------------------|
|                                            | Mean absolute value of error terms (1992-2010) |                      |                        |                       |                        |                        |                       |                        |
|                                            | (1)                                            | (2)                  | (3)                    | (4)                   | (5)                    | (6)                    | (7)                   | (8)                    |
| Years of leader (2010)                     | 0.0136***<br>(0.0028)                          |                      |                        |                       |                        |                        |                       |                        |
| Average years of leader (1992-2010)        |                                                | 0.0112**<br>(0.0046) |                        |                       |                        |                        |                       |                        |
| Freedom status 2                           |                                                |                      | -0.2016***<br>(0.0762) |                       |                        |                        |                       |                        |
| Freedom status 3                           |                                                |                      | -0.3145***<br>(0.0677) |                       |                        |                        |                       |                        |
| Public Sector Corruption Index (1992-2010) |                                                |                      |                        | 0.2926***<br>(0.0905) |                        |                        |                       |                        |
| Voice and Accountability (1992-2010)       |                                                |                      |                        |                       | -0.1194***<br>(0.0287) |                        |                       |                        |
| Starting a business (1992-2010)            |                                                |                      |                        |                       |                        | -0.0053***<br>(0.0016) |                       |                        |
| State Fragility Index (1992-2010)          |                                                |                      |                        |                       |                        |                        | 0.0149***<br>(0.0041) |                        |
| Natural Resource Protection Indicator      |                                                |                      |                        |                       |                        |                        |                       | -0.0026***<br>(0.0008) |
| Observations                               | 157                                            | 158                  | 171                    | 159                   | 170                    | 153                    | 152                   | 170                    |
| R <sup>2</sup>                             | 0.1045                                         | 0.0412               | 0.1365                 | 0.0674                | 0.1143                 | 0.0969                 | 0.0816                | 0.0677                 |
| Adjusted R <sup>2</sup>                    | 0.0988                                         | 0.0351               | 0.1262                 | 0.0615                | 0.1090                 | 0.0909                 | 0.0755                | 0.0621                 |

*Note:* This table reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2010) on the variables shown and notes that the LASSO model selects all these variables. Authors downloaded data from various sources: The WhoGov dataset (column 1&2), Freedom House (column 3), Varieties of Democracy dataset (column 4), Center of Systemic Peace (column 7), World Development Indicator (column 5&6), and Quality of Government (column 8). Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

**Table E3. Cross-sectional analysis: Country and Geographical Data (2012)**

|                                                              | <i>Dependent variable:</i>                     |                       |                        |                     |                    |                     |                     |
|--------------------------------------------------------------|------------------------------------------------|-----------------------|------------------------|---------------------|--------------------|---------------------|---------------------|
|                                                              | Mean absolute value of error terms (1992-2010) |                       |                        |                     |                    |                     |                     |
|                                                              | (1)                                            | (2)                   | (3)                    | (4)                 | (5)                | (6)                 | (7)                 |
| Ruggedness (Terrain Ruggedness Index, 100 m) in 2012         | -0.0246<br>(0.0171)                            |                       |                        |                     |                    |                     |                     |
| Average distance to nearest ice-free coast (1000 km) in 2012 |                                                | 0.2360***<br>(0.0662) |                        |                     |                    |                     |                     |
| Percentage within 100 km. of ice-free coast in 2012          |                                                |                       | -0.0018***<br>(0.0006) |                     |                    |                     |                     |
| Percentage tropical climate in 2012                          |                                                |                       |                        | -0.0005<br>(0.0006) |                    |                     |                     |
| Percentage desert in 2012                                    |                                                |                       |                        |                     | 0.0031<br>(0.0034) |                     |                     |
| Latitude                                                     |                                                |                       |                        |                     |                    | 0.0015*<br>(0.0009) |                     |
| Longitude                                                    |                                                |                       |                        |                     |                    |                     | 0.0007*<br>(0.0004) |
| Observations                                                 | 169                                            | 169                   | 169                    | 169                 | 169                | 167                 | 167                 |
| R <sup>2</sup>                                               | 0.0076                                         | 0.0698                | 0.0458                 | 0.0041              | 0.0109             | 0.0118              | 0.0169              |
| Adjusted R <sup>2</sup>                                      | 0.0017                                         | 0.0642                | 0.0400                 | -0.0018             | 0.0050             | 0.0058              | 0.0109              |

*Note:* Each column reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2010) on the variables shown. The LASSO model only selects average distance to nearest ice-free coast (1000 km) in 2012. Data source: Ruggedness: The blessing of bad geography in Africa (Nunn & Puga, 2012). Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

**Table E4. Cross-sectional analysis: The Region of the Country**

|                         | <i>Dependent variable:</i>                     |
|-------------------------|------------------------------------------------|
|                         | Mean absolute value of error terms (1992-2010) |
| Region of country 2     | -0.1941**<br>(0.0927)                          |
| Region of country 3     | 0.0566<br>(0.1091)                             |
| Region of country 4     | 0.0230<br>(0.0920)                             |
| Region of country 5     | -0.0924<br>(0.0936)                            |
| Region of country 6     | -0.1502<br>(0.1061)                            |
| Region of country 7     | -0.1008<br>(0.1070)                            |
| Observations            | 172                                            |
| R <sup>2</sup>          | 0.0762                                         |
| Adjusted R <sup>2</sup> | 0.0426                                         |

*Note:* This table reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2010) on the variable: the region of the country (Quality of Government). There are seven different regions: 1. Eastern Europe and post-Soviet Union, 2. Latin America and The Caribbean, 3. North Africa & the Middle East, 4. Sub-Saharan Africa, 5. Western Europe and North America, 6. East Asia Pacific, 7. South Asia. Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

**Table E5. Cross-sectional analysis: Land factors**

|                                                        | <i>Dependent variable:</i>                     |                     |                     |                    |                      |
|--------------------------------------------------------|------------------------------------------------|---------------------|---------------------|--------------------|----------------------|
|                                                        | Mean absolute value of error terms (1992-2010) |                     |                     |                    |                      |
|                                                        | (1)                                            | (2)                 | (3)                 | (4)                | (5)                  |
| Land area (sq. km) (log form)                          | 0.0165*<br>(0.0095)                            |                     |                     |                    |                      |
| Average Agricultural land (% of land area) (1992-2010) |                                                | -0.0002<br>(0.0014) |                     |                    |                      |
| Average Forest area (% of land area) (1992-2010)       |                                                |                     | -0.0011<br>(0.0012) |                    |                      |
| Arable Land (% of Agricultural land) (2017)            |                                                |                     |                     | 0.0012<br>(0.0010) |                      |
| Cropland (% of Agricultural land) (2017)               |                                                |                     |                     |                    | -0.00004<br>(0.0008) |
| Observations                                           | 168                                            | 169                 | 169                 | 169                | 169                  |
| R <sup>2</sup>                                         | 0.0138                                         | 0.0001              | 0.0068              | 0.0082             | 0.00001              |
| Adjusted R <sup>2</sup>                                | 0.0079                                         | -0.0059             | 0.0008              | 0.0023             | -0.0060              |

*Note:* Each column reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2010) on the variables shown. Data source: World Development Indicator (column 1-3), Quality of Government (column 4-5). Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

**Table E6. Cross-sectional analysis: Economic Development factors (GDP sectors)**

|                                                                                | <i>Dependent variable:</i>                     |                    |                        |
|--------------------------------------------------------------------------------|------------------------------------------------|--------------------|------------------------|
|                                                                                | Mean absolute value of error terms (1992-2010) |                    |                        |
|                                                                                | (1)                                            | (2)                | (3)                    |
| Average Agriculture, forestry, and fishing, value added (% of GDP) (1992-2010) | 0.0060**<br>(0.0024)                           |                    |                        |
| Average Industry (including construction), value added (% of GDP) (1992-2010)  |                                                | 0.0042<br>(0.0025) |                        |
| Average Services, value added (% of GDP) (1992-2010)                           |                                                |                    | -0.0104***<br>(0.0023) |
| Observations                                                                   | 168                                            | 167                | 163                    |
| R <sup>2</sup>                                                                 | 0.0512                                         | 0.0237             | 0.1270                 |
| Adjusted R <sup>2</sup>                                                        | 0.0455                                         | 0.0177             | 0.1215                 |

*Note:* Each column reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2010) on the variables shown. Data source: World Development Indicator. Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

**Table E7. Cross-sectional analysis: Economic Development factors (GDP subcomponents)**

|                         | <i>Dependent variable:</i>                     |                     |                     |                    |                     |
|-------------------------|------------------------------------------------|---------------------|---------------------|--------------------|---------------------|
|                         | Mean absolute value of error terms (1992-2010) |                     |                     |                    |                     |
|                         | (1)                                            | (2)                 | (3)                 | (4)                | (5)                 |
| Consumption             | -0.0007<br>(0.0018)                            |                     |                     |                    |                     |
| Investment              |                                                | -0.0041<br>(0.0044) |                     |                    |                     |
| Government Spending     |                                                |                     | -0.0062<br>(0.0042) |                    |                     |
| Export                  |                                                |                     |                     | 0.0009<br>(0.0013) |                     |
| Import                  |                                                |                     |                     |                    | -0.0002<br>(0.0012) |
| Observations            | 157                                            | 158                 | 157                 | 162                | 162                 |
| R <sup>2</sup>          | 0.0014                                         | 0.0062              | 0.0113              | 0.0030             | 0.0002              |
| Adjusted R <sup>2</sup> | -0.0050                                        | -0.0002             | 0.0050              | -0.0033            | -0.0060             |

*Note:* Each column reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2010) on the variables shown. Consumption is Average Households and NPISHs final consumption expenditure (% of GDP) (1992-2010), Investment is Average Gross capital formation (% of GDP) (1992-2010), Government is Average General government final consumption expenditure (% of GDP) (1992-2010), Exports are Average Exports of goods and services (% of GDP) (1992-2010) and Imports are Average Imports of goods and services (% of GDP) (1992-2010). Data source: World Development Indicator. Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

**Table E8. Cross-sectional analysis: Other Economic Development factors**

|                                                      | <i>Dependent variable:</i>                     |                    |                    |
|------------------------------------------------------|------------------------------------------------|--------------------|--------------------|
|                                                      | Mean absolute value of error terms (1992-2010) |                    |                    |
|                                                      | (1)                                            | (2)                | (3)                |
| Average GDP growth (annual %) (1992-2010)            | 0.0212**<br>(0.0100)                           |                    |                    |
| Average GDP per capita growth (annual %) (1992-2010) |                                                | 0.0146<br>(0.0133) |                    |
| Average Gross Saving (% of GDP) (1992-2010)          |                                                |                    | 0.0011<br>(0.0023) |
| Observations                                         | 169                                            | 169                | 153                |
| R <sup>2</sup>                                       | 0.0331                                         | 0.0132             | 0.0014             |
| Adjusted R <sup>2</sup>                              | 0.0273                                         | 0.0073             | -0.0052            |

*Note:* Each column reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2010) on the variables shown. Data source: World Development Indicator  
Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

**Table E9. Cross-sectional analysis: Historical factors**

|                         | <i>Dependent variable:</i>                     |                       |
|-------------------------|------------------------------------------------|-----------------------|
|                         | Mean absolute value of error terms (1992-2010) |                       |
|                         | (1)                                            | (2)                   |
| Colonial origin 1       | 0.0711<br>(0.1221)                             |                       |
| Colonial origin 2       | -0.0713<br>(0.0843)                            |                       |
| Colonial origin 3       | -0.0308<br>(0.0609)                            |                       |
| Colonial origin 4       | 0.0381<br>(0.0827)                             |                       |
| Legal origin 2          |                                                | 0.1039*<br>(0.0566)   |
| Legal origin 3          |                                                | -0.0257<br>(0.1126)   |
| Legal origin 4          |                                                | 0.3167***<br>(0.1055) |
| Legal origin 5          |                                                | 0.1941*<br>(0.1028)   |
| Observations            | 170                                            | 137                   |
| R <sup>2</sup>          | 0.0132                                         | 0.0504                |
| Adjusted R <sup>2</sup> | -0.0107                                        | 0.0217                |

*Note:* Each column reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2010) on the variables shown. Column (1) is colonial origin. This is a fourfold classification of the former colonial ruler of the country. They are: 0. Never colonized by a Western overseas colonial power, 1. Others, 2. Spanish and Portuguese, 3. British, 4. French. Column (2) is legal origin. Identifies the legal origin of the Company Law or Commercial code of each country. There are five possible origins: 1. English Common Law, 2. French Commercial Code, 3. Socialist/Communist Laws, 4. German Commercial Code, 5. Scandinavian Commercial Code. Data source: Quality of Government. Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

**Table E10. Cross-sectional analysis: Ethnic and Cultural Diversity factors**

|                          | <i>Dependent variable:</i>                     |                    |                     |                     |
|--------------------------|------------------------------------------------|--------------------|---------------------|---------------------|
|                          | Mean absolute value of error terms (1992-2010) |                    |                     |                     |
|                          | (1)                                            | (2)                | (3)                 | (4)                 |
| Cultural Diversity       | 0.0617<br>(0.1363)                             |                    |                     |                     |
| Ethnic Fractionalization |                                                | 0.0170<br>(0.1016) |                     |                     |
| Largest Minority         |                                                |                    | -0.0487<br>(0.2735) |                     |
| Plurality Group          |                                                |                    |                     | -0.0509<br>(0.1133) |
| Observations             | 143                                            | 144                | 136                 | 143                 |
| R <sup>2</sup>           | 0.0014                                         | 0.0002             | 0.0002              | 0.0013              |
| Adjusted R <sup>2</sup>  | -0.0056                                        | -0.0069            | -0.0072             | -0.0058             |

*Note:* Each column reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2010) on the variables shown. Data source: Quality of Government. Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

**Table E11. Cross-sectional analysis: Demographic factors**

|                                                                            | <i>Dependent variable:</i>                     |          |          |          |
|----------------------------------------------------------------------------|------------------------------------------------|----------|----------|----------|
|                                                                            | Mean absolute value of error terms (1992-2010) |          |          |          |
|                                                                            | (1)                                            | (2)      | (3)      | (4)      |
| Average Population growth (annual %) (1992-2010)                           | 0.0382*                                        |          |          |          |
|                                                                            | (0.0203)                                       |          |          |          |
| Average Urban population growth (annual %) (1992-2010)                     |                                                | 0.0301** |          |          |
|                                                                            |                                                | (0.0144) |          |          |
| Average Population density (people per sq. km of land area) (1992-2010)    |                                                |          | -0.00002 |          |
|                                                                            |                                                |          | (0.0001) |          |
| Average Population in the largest city (% of urban population) (1992-2010) |                                                |          |          | -0.0003  |
|                                                                            |                                                |          |          | (0.0018) |
| Observations                                                               | 170                                            | 170      | 169      | 138      |
| R <sup>2</sup>                                                             | 0.0243                                         | 0.0296   | 0.0001   | 0.0002   |
| Adjusted R <sup>2</sup>                                                    | 0.0185                                         | 0.0238   | -0.0059  | -0.0072  |

*Note:* Each column reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2010) on the variables shown. Data source: World Development Indicator. Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

**Table E12. Cross-sectional analysis: Other factors**

|                                  | <i>Dependent variable:</i>                     |                     |                     |                     |
|----------------------------------|------------------------------------------------|---------------------|---------------------|---------------------|
|                                  | Mean absolute value of error terms (1992-2010) |                     |                     |                     |
|                                  | (1)                                            | (2)                 | (3)                 | (4)                 |
| European descent                 | -0.0011*<br>(0.0006)                           |                     |                     |                     |
| Ethnic_Fractionalization_in_2000 |                                                | 0.1743*<br>(0.0939) |                     |                     |
| Language Fractionalization       |                                                |                     | 0.1443*<br>(0.0861) |                     |
| Religion Fractionalization       |                                                |                     |                     | -0.0239<br>(0.0995) |
| Observations                     | 151                                            | 167                 | 162                 | 169                 |
| R <sup>2</sup>                   | 0.0198                                         | 0.0180              | 0.0158              | 0.0003              |
| Adjusted R <sup>2</sup>          | 0.0132                                         | 0.0121              | 0.0096              | -0.0057             |

*Note:* Each column reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2010) on the variables shown. Data source: Quality of Government (columns 2-4), column 1. (Nunn and Puga 2012). Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

## APPENDIX F. CROSS-SECTIONAL ANALYSIS (HARMONIZED NIGHTTIME LIGHT)

**Table F1. SUMMARY STATISTICS**

| Statistic                                                                          | N   | Mean     | St. Dev. | Min      | Max        |
|------------------------------------------------------------------------------------|-----|----------|----------|----------|------------|
| Mean absolute error                                                                | 177 | 0.5921   | 0.3374   | 0.0996   | 1.4407     |
| Average GDP growth (annual %) (1992-2018)                                          | 177 | 3.7199   | 2.2721   | -0.9375  | 19.6362    |
| Average GDP per capita growth (annual %) (1992-2018)                               | 177 | 2.2191   | 2.1411   | -1.8019  | 14.8538    |
| Average Gross Saving (% of GDP) (1992-2018)                                        | 159 | 21.7277  | 10.8582  | -36.6979 | 55.0252    |
| Average Agriculture, forestry, and fishing, value added (% of GDP) (1992-2018)     | 177 | 12.5029  | 11.8299  | 0.0176   | 54.4463    |
| Average Industry (including construction), value added (% of GDP) (1992-2018)      | 175 | 26.9081  | 12.4486  | 6.5041   | 77.0993    |
| Average Services, value added (% of GDP) (1992-2018)                               | 173 | 53.5919  | 11.7363  | 27.9936  | 86.4658    |
| Average Gross capital formation (% of GDP) (1992-2018)                             | 162 | 23.8715  | 6.1165   | 10.8323  | 49.8258    |
| Average General government final consumption expenditure (% of GDP) (1992-2018)    | 162 | 16.4714  | 7.1719   | 4.4704   | 63.3870    |
| Average Households and NPISHs final consumption expenditure (% of GDP) (1992-2018) | 162 | 64.5070  | 16.5029  | 17.7372  | 141.6347   |
| Average Exports of goods and services (% of GDP) (1992-2018)                       | 169 | 40.9335  | 25.9569  | 7.9116   | 170.7243   |
| Average Imports of goods and services (% of GDP) (1992-2018)                       | 169 | 46.9440  | 24.7189  | 10.5049  | 145.4878   |
| Average Population growth (annual %) (1992-2018)                                   | 177 | 1.4502   | 1.3180   | -1.1187  | 8.3385     |
| Average Urban population growth (annual %) (1992-2018)                             | 176 | 2.1636   | 1.7910   | -1.1649  | 8.5019     |
| Average Population density (people per sq. km of land area) (1992-2018)            | 177 | 145.6189 | 204.0805 | 1.6718   | 1,277.5920 |
| Average Population in the largest city (% of urban population) (1992-2018)         | 136 | 32.5862  | 15.7730  | 3.0621   | 75.9030    |
| Average Agricultural land (% of land area) (1992-2018)                             | 175 | 38.0233  | 21.2594  | 0.5287   | 84.2807    |
| Average Forest area (% of land area) (1992-2018)                                   | 176 | 33.5849  | 24.6292  | 0.0000   | 98.1558    |
| Average Number of years the leader in office continuously (1992-2018)              | 152 | 7.1917   | 5.9404   | 1.0000   | 34.0000    |

|                                                                       |     |         |         |           |          |
|-----------------------------------------------------------------------|-----|---------|---------|-----------|----------|
| Average Starting a business score (1992-2018)                         | 169 | 73.7831 | 14.7572 | 27.7467   | 98.5267  |
| Average Public sector corruption index (1992-2018)                    | 154 | 0.4769  | 0.2872  | 0.0010    | 0.9643   |
| Average Voice and Accountability, Estimate (1992-2018)                | 174 | 0.0247  | 0.9287  | -1.9124   | 1.6142   |
| Average Control of Corruption: Estimate (1992-2018)                   | 173 | -0.0404 | 0.9328  | -1.5570   | 2.3365   |
| Average State Fragility Index (1995-2018)                             | 150 | 8.8957  | 6.2684  | 0.0000    | 23.3333  |
| Latitude                                                              | 172 | 18.1922 | 23.9272 | -41.8060  | 64.4670  |
| Longitude                                                             | 172 | 16.4002 | 66.8299 | -174.8470 | 171.4780 |
| % Fertile soil                                                        | 171 | 40.4452 | 25.5339 | 0.0000    | 100.0000 |
| Percentage desert in 2012                                             | 168 | 3.3110  | 10.9514 | 0.0000    | 77.2795  |
| Percentage tropical climate in 2012                                   | 168 | 42.5695 | 45.6027 | 0.0000    | 100.0000 |
| Ruggedness (Terrain Ruggedness Index, 100 m) in 2012                  | 168 | 1.2867  | 1.1873  | 0.0029    | 6.7401   |
| Average distance to nearest ice-free coast (1000 km) in 2012          | 168 | 0.3027  | 0.3868  | 0.0000    | 2.2062   |
| Percentage within 100 km. of ice-free coast in 2012                   | 168 | 46.1293 | 40.7271 | 0.0000    | 100.0000 |
| Arable Land (% of Agricultural land) (2017)                           | 169 | 54.0849 | 30.0213 | 0.1200    | 100.0000 |
| Cropland (% of Agricultural land) (2017)                              | 168 | 39.9558 | 25.6953 | 0.1200    | 93.6300  |
| % European descent                                                    | 146 | 33.9309 | 42.5138 | 0.0000    | 100.0000 |
| Ethnic Fractionalization in the year 2000                             | 166 | 0.4339  | 0.2573  | 0.0000    | 0.9302   |
| Language Fractionalization in the year 2000                           | 159 | 0.3966  | 0.2835  | 0.0021    | 0.9227   |
| Religion Fractionalization in the year 2000                           | 167 | 0.4481  | 0.2377  | 0.0035    | 0.8603   |
| Threatened Species: Fishes (2019)                                     | 170 | 55.0118 | 55.4508 | 1.0000    | 345.0000 |
| State Fragility Index (2017)                                          | 149 | 7.6443  | 5.9207  | 0.0000    | 24.0000  |
| Starting a business score (2017)                                      | 169 | 82.7553 | 10.8517 | 33.7043   | 99.9635  |
| Cultural Diversity                                                    | 138 | 0.3110  | 0.2074  | 0.0000    | 0.7328   |
| Ethnic Fractionalization                                              | 139 | 0.4803  | 0.2614  | 0.0040    | 1.0000   |
| Largest Minority                                                      | 132 | 0.1694  | 0.1083  | 0.0100    | 0.4400   |
| Plurality Group                                                       | 138 | 0.6436  | 0.2392  | 0.1200    | 0.9980   |
| Parliamentary Election: Compulsory Voting                             | 164 | 0.1524  | 0.3605  | 0.0000    | 1.0000   |
| Perception of Electoral Integrity Index Type                          | 152 | 3.0526  | 1.3460  | 1.0000    | 5.0000   |
| Legal Origin                                                          | 137 | 1.8613  | 0.8505  | 1.0000    | 5.0000   |
| Public sector corruption index (2017)                                 | 154 | 0.4789  | 0.2905  | 0.0120    | 0.9680   |
| Agriculture, forestry, and _shing, value added (% of GDP) (2014-2017) | 166 | 10.8814 | 10.5342 | 0.2147    | 49.1578  |
| Number of years the leader in office continuously in 2017             | 152 | 7.3289  | 8.8091  | 1.0000    | 46.0000  |
| Regime Type (simplified) (2014)                                       | 164 | 61.0976 | 47.5048 | 1.0000    | 100.0000 |

|                                           |     |         |         |         |          |
|-------------------------------------------|-----|---------|---------|---------|----------|
| The Region of the Country                 | 177 | 3.5650  | 1.7573  | 1       | 7        |
| Voice and Accountability, Estimate (2017) | 170 | 0.0392  | 0.9311  | -1.9711 | 1.6929   |
| Colonial Origin                           | 170 | 1.8000  | 1.5292  | 0.0000  | 4.0000   |
| Natural Resource Protection Indicator     | 170 | 69.1200 | 33.3087 | 0.0000  | 100.0000 |
| Freedom status (2018)                     | 171 | 2.2398  | 0.7940  | 1.0000  | 3.0000   |

**Table F2. Cross-sectional analysis: Institutional quality factors**

|                                            | <i>Dependent variable:</i>                     |                       |                        |                       |                        |                        |                       |                        |                      |
|--------------------------------------------|------------------------------------------------|-----------------------|------------------------|-----------------------|------------------------|------------------------|-----------------------|------------------------|----------------------|
|                                            | Mean absolute value of error terms (1992-2018) |                       |                        |                       |                        |                        |                       |                        |                      |
|                                            | (1)                                            | (2)                   | (3)                    | (4)                   | (5)                    | (6)                    | (7)                   | (8)                    | (9)                  |
| Years of leader (2017)                     | 0.0093***<br>(0.0031)                          |                       |                        |                       |                        |                        |                       |                        |                      |
| Average years of leader (1992-2018)        |                                                | 0.0134***<br>(0.0048) |                        |                       |                        |                        |                       |                        |                      |
| Freedom status 2                           |                                                |                       | -0.1279<br>(0.0779)    |                       |                        |                        |                       |                        |                      |
| Freedom status 3                           |                                                |                       | -0.2536***<br>(0.0693) |                       |                        |                        |                       |                        |                      |
| Public Sector Corruption Index (1992-2018) |                                                |                       |                        | 0.3555***<br>(0.0991) |                        |                        |                       |                        |                      |
| Voice and Accountability (1992-2018)       |                                                |                       |                        |                       | -0.1063***<br>(0.0299) |                        |                       |                        |                      |
| Starting a business (1992-2018)            |                                                |                       |                        |                       |                        | -0.0079***<br>(0.0019) |                       |                        |                      |
| State Fragility Index (1992-2018)          |                                                |                       |                        |                       |                        |                        | 0.0149***<br>(0.0045) |                        |                      |
| Control of corruption (1992-2018)          |                                                |                       |                        |                       |                        |                        |                       | -0.0978***<br>(0.0301) |                      |
| Natural Resource Protection Indicator      |                                                |                       |                        |                       |                        |                        |                       |                        | -0.0015*<br>(0.0008) |
| Observations                               | 152                                            | 152                   | 171                    | 154                   | 174                    | 169                    | 150                   | 173                    | 170                  |
| R <sup>2</sup>                             | 0.0566                                         | 0.0527                | 0.0874                 | 0.0871                | 0.0850                 | 0.1176                 | 0.0726                | 0.0723                 | 0.0221               |
| Adjusted R <sup>2</sup>                    | 0.0503                                         | 0.0464                | 0.0765                 | 0.0811                | 0.0796                 | 0.1123                 | 0.0664                | 0.0669                 | 0.0162               |

*Note:* This table reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2018) on the variables shown and notes that the LASSO model selects all these variables. Authors downloaded data from various sources: The WhoGov dataset (column 1&2), Freedom House (column 3), Varieties of Democracy dataset (column 4), Center of Systemic Peace (column 7), World Development Indicator (column 5,6 &8), and Quality of Government (column 9). Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

**Table F3. Cross-sectional analysis: Country Ruggedness and Geographical factors (2012)**

|                                                              | <i>Dependent variable:</i>                     |           |          |          |          |          |           |
|--------------------------------------------------------------|------------------------------------------------|-----------|----------|----------|----------|----------|-----------|
|                                                              | Mean absolute value of error terms (1992-2018) |           |          |          |          |          |           |
|                                                              | (1)                                            | (2)       | (3)      | (4)      | (5)      | (6)      | (7)       |
| Ruggedness (Terrain Ruggedness Index, 100 m) in 2012         | -0.0332*                                       |           |          |          |          |          |           |
|                                                              | (0.0190)                                       |           |          |          |          |          |           |
| Average distance to nearest ice-free coast (1000 km) in 2012 |                                                | 0.2257*** |          |          |          |          |           |
|                                                              |                                                | (0.0666)  |          |          |          |          |           |
| Percentage within 100 km. of ice-free coast in 2012          |                                                |           | -0.0011* |          |          |          |           |
|                                                              |                                                |           | (0.0006) |          |          |          |           |
| Percentage tropical climate in 2012                          |                                                |           |          | 0.0003   |          |          |           |
|                                                              |                                                |           |          | (0.0006) |          |          |           |
| Percentage desert in 2012                                    |                                                |           |          |          | 0.0024   |          |           |
|                                                              |                                                |           |          |          | (0.0027) |          |           |
| Latitude                                                     |                                                |           |          |          |          | 0.0016   |           |
|                                                              |                                                |           |          |          |          | (0.0011) |           |
| Longitude                                                    |                                                |           |          |          |          |          | 0.0010*** |
|                                                              |                                                |           |          |          |          |          | (0.0004)  |
| Observations                                                 | 168                                            | 168       | 168      | 168      | 168      | 172      | 172       |
| R <sup>2</sup>                                               | 0.0134                                         | 0.0656    | 0.0181   | 0.0017   | 0.0059   | 0.0123   | 0.0355    |
| Adjusted R <sup>2</sup>                                      | 0.0075                                         | 0.0600    | 0.0122   | -0.0043  | -0.0001  | 0.0065   | 0.0298    |

*Note:* Each column reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2018) on the variables shown. The LASSO model only selects average distance to nearest ice-free coast (1000 km) in 2012. Data source: Ruggedness: The blessing of bad geography in Africa (Nunn & Puga, 2012). Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

**Table F4. Cross-sectional analysis: The Region of the Country**

|                         | <i>Dependent variable:</i>                     |
|-------------------------|------------------------------------------------|
|                         | Mean absolute value of error terms (1992-2018) |
| Region of country 2     | -0.3787***<br>(0.0683)                         |
| Region of country 3     | -0.1708*<br>(0.0974)                           |
| Region of country 4     | -0.0391<br>(0.0795)                            |
| Region of country 5     | -0.2869***<br>(0.0857)                         |
| Region of country 6     | -0.1388*<br>(0.0823)                           |
| Region of country 7     | -0.2279**<br>(0.1022)                          |
| Observations            | 177                                            |
| R <sup>2</sup>          | 0.1708                                         |
| Adjusted R <sup>2</sup> | 0.1415                                         |

*Note:* This table reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2018) on the variable: the region of the country (Quality of Government). There are seven different regions: 1. Eastern Europe and post-Soviet Union, 2. Latin America and The Caribbean, 3. North Africa & the Middle East, 4. Sub-Saharan Africa, 5. Western Europe and North America, 6. East Asia Pacific, 7. South Asia. Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

**Table F5. Cross-sectional analysis: Land factors**

|                                                        | <i>Dependent variable:</i>                     |                     |                    |                    |
|--------------------------------------------------------|------------------------------------------------|---------------------|--------------------|--------------------|
|                                                        | Mean absolute value of error terms (1992-2018) |                     |                    |                    |
|                                                        | (1)                                            | (2)                 | (3)                | (4)                |
| Average Agricultural land (% of land area) (1992-2018) | -0.0010<br>(0.0013)                            |                     |                    |                    |
| Average Forest area (% of land area) (1992-2018)       |                                                | -0.0006<br>(0.0011) |                    |                    |
| Arable Land (% of Agricultural land) (2017)            |                                                |                     | 0.0011<br>(0.0010) |                    |
| Cropland (% of Agricultural land) (2017)               |                                                |                     |                    | 0.0006<br>(0.0009) |
| Observations                                           | 175                                            | 176                 | 168                | 169                |
| R <sup>2</sup>                                         | 0.0039                                         | 0.0016              | 0.0074             | 0.0031             |
| Adjusted R <sup>2</sup>                                | -0.0018                                        | -0.0041             | 0.0015             | -0.0029            |

*Note:* Each column reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2018) on the variables shown. Data source: World Development Indicator (column 1-3), Quality of Government (column 4-5). Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

**Table F6. Cross-sectional analysis: Economic Development factors (GDP sectors)**

|                                                                                | <i>Dependent variable:</i>                     |                    |                        |
|--------------------------------------------------------------------------------|------------------------------------------------|--------------------|------------------------|
|                                                                                | Mean absolute value of error terms (1992-2018) |                    |                        |
|                                                                                | (1)                                            | (2)                | (3)                    |
| Average Agriculture, forestry, and fishing, value added (% of GDP) (1992-2018) | 0.0084***<br>(0.0023)                          |                    |                        |
| Average Industry (including construction), value added (% of GDP) (1992-2018)  |                                                | 0.0016<br>(0.0025) |                        |
| Average Services, value added (% of GDP) (1992-2018)                           |                                                |                    | -0.0088***<br>(0.0020) |
| Observations                                                                   | 177                                            | 175                | 173                    |
| R <sup>2</sup>                                                                 | 0.0859                                         | 0.0034             | 0.0971                 |
| Adjusted R <sup>2</sup>                                                        | 0.0807                                         | -0.0023            | 0.0918                 |

*Note:* Each column reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2018) on the variables shown. Data source: World Development Indicator. Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

**Table F7. Cross-sectional analysis: Economic Development factors (GDP subcomponents)**

|                         | <i>Dependent variable:</i>                     |                     |                    |                      |                     |
|-------------------------|------------------------------------------------|---------------------|--------------------|----------------------|---------------------|
|                         | Mean absolute value of error terms (1992-2018) |                     |                    |                      |                     |
|                         | (1)                                            | (2)                 | (3)                | (4)                  | (5)                 |
| Consumption             | 0.0020<br>(0.0018)                             |                     |                    |                      |                     |
| Investment              |                                                | -0.0043<br>(0.0039) |                    |                      |                     |
| Government Spending     |                                                |                     | 0.0007<br>(0.0046) |                      |                     |
| Export                  |                                                |                     |                    | -0.0015*<br>(0.0009) |                     |
| Import                  |                                                |                     |                    |                      | -0.0008<br>(0.0010) |
| Observations            | 162                                            | 162                 | 162                | 169                  | 169                 |
| R <sup>2</sup>          | 0.0088                                         | 0.0058              | 0.0002             | 0.0124               | 0.0034              |
| Adjusted R <sup>2</sup> | 0.0026                                         | -0.0004             | -0.0060            | 0.0065               | -0.0025             |

*Note:* Each column reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2018) on the variables shown. Consumption is Average Households and NPISHs final consumption expenditure (% of GDP) (1992-2018), Investment is Average Gross capital formation (% of GDP) (1992-2018), Government is Average General government final consumption expenditure (% of GDP) (1992-2018), Exports are Average Exports of goods and services (% of GDP) (1992-2018) and Imports are Average Imports of goods and services (% of GDP) (1992-2018). Data source: World Development Indicator. Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

**Table F8. Cross-sectional analysis: Other Economic Development factors**

|                                                      | <i>Dependent variable:</i>                     |                    |                    |
|------------------------------------------------------|------------------------------------------------|--------------------|--------------------|
|                                                      | Mean absolute value of error terms (1992-2018) |                    |                    |
|                                                      | (1)                                            | (2)                | (3)                |
| Average GDP growth (annual %) (1992-2018)            | 0.0199*<br>(0.0119)                            |                    |                    |
| Average GDP per capita growth (annual %) (1992-2018) |                                                | 0.0073<br>(0.0152) |                    |
| Average <i>Gross Saving</i> (% of GDP) (1992-2018)   |                                                |                    | 0.0005<br>(0.0025) |
| Observations                                         | 177                                            | 177                | 159                |
| R <sup>2</sup>                                       | 0.0179                                         | 0.0022             | 0.0003             |
| Adjusted R <sup>2</sup>                              | 0.0123                                         | -0.0035            | -0.0061            |

*Note:* Each column reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2018) on the variables shown. Data source: World Development Indicator Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

**Table F9. Cross-sectional analysis: Historical factors**

|                         | <i>Dependent variable:</i>                     |          |
|-------------------------|------------------------------------------------|----------|
|                         | Mean absolute value of error terms (1992-2018) |          |
|                         | (1)                                            | (2)      |
| Colonial origin 1       | 0.1750*                                        |          |
|                         | (0.1012)                                       |          |
| Colonial origin 2       | -0.1439                                        |          |
|                         | (0.0886)                                       |          |
| Colonial origin 3       | -0.1063*                                       |          |
|                         | (0.0561)                                       |          |
| Colonial origin 4       | 0.1435                                         |          |
|                         | (0.0894)                                       |          |
| Legal origin 2          |                                                | 0.1004   |
|                         |                                                | (0.0618) |
| Legal origin 3          |                                                | 0.0493   |
|                         |                                                | (0.0747) |
| Legal origin 4          |                                                | 0.2162*  |
|                         |                                                | (0.1254) |
| Legal origin 5          |                                                | 0.3645   |
|                         |                                                | (0.2219) |
| Observations            | 170                                            | 137      |
| R <sup>2</sup>          | 0.0971                                         | 0.0413   |
| Adjusted R <sup>2</sup> | 0.0752                                         | 0.0123   |

*Note:* Each column reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2018) on the variables shown. Column (1) is colonial origin. This is a fourfold classification of the former colonial ruler of the country. They are: 0. Never colonized by a Western overseas colonial power, 1. Others, 2. Spanish and Portuguese, 3. British, 4. French. Column (2) is legal origin. Identifies the legal origin of the Company Law or Commercial code of each country. There are five possible origins: 1. English Common Law, 2. French Commercial Code, 3. Socialist/Communist Laws, 4. German Commercial Code, 5. Scandinavian Commercial Code. Data source: Quality of Government. Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

**Table F10. Cross-sectional analysis: Ethnic and Cultural Diversity factors**

|                          | <i>Dependent variable:</i>                     |                    |                     |                     |
|--------------------------|------------------------------------------------|--------------------|---------------------|---------------------|
|                          | Mean absolute value of error terms (1992-2018) |                    |                     |                     |
|                          | (1)                                            | (2)                | (3)                 | (4)                 |
| Cultural Diversity       | 0.1902<br>(0.1445)                             |                    |                     |                     |
| Ethnic Fractionalization |                                                | 0.1169<br>(0.1157) |                     |                     |
| Largest Minority         |                                                |                    | -0.2494<br>(0.2788) |                     |
| Plurality Group          |                                                |                    |                     | -0.1523<br>(0.1269) |
| Observations             | 138                                            | 139                | 132                 | 138                 |
| R <sup>2</sup>           | 0.0131                                         | 0.0078             | 0.0062              | 0.0111              |
| Adjusted R <sup>2</sup>  | 0.0058                                         | 0.0006             | -0.0015             | 0.0039              |

*Note:* Each column reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2018) on the variables shown. Data source: Quality of Government. Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

**Table F11. Cross-sectional analysis: Demographic factors**

|                                                                               | <i>Dependent variable:</i>                     |                       |                    |                     |
|-------------------------------------------------------------------------------|------------------------------------------------|-----------------------|--------------------|---------------------|
|                                                                               | Mean absolute value of error terms (1992-2018) |                       |                    |                     |
|                                                                               | (1)                                            | (2)                   | (3)                | (4)                 |
| Average Population growth (annual %)<br>(1992-2018)                           | 0.0374**<br>(0.0189)                           |                       |                    |                     |
| Average Urban population growth (annual %)<br>(1992-2018)                     |                                                | 0.0415***<br>(0.0134) |                    |                     |
| Average Population density (people per sq.<br>km of land area) (1992-2018)    |                                                |                       | 0.0001<br>(0.0001) |                     |
| Average Population in the largest city (%<br>of urban population) (1992-2018) |                                                |                       |                    | -0.0003<br>(0.0019) |
| Observations                                                                  | 177                                            | 176                   | 177                | 136                 |
| R <sup>2</sup>                                                                | 0.0213                                         | 0.0484                | 0.0010             | 0.0002              |
| Adjusted R <sup>2</sup>                                                       | 0.0157                                         | 0.0429                | -0.0047            | -0.0072             |

*Note:* Each column reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2018) on the variables shown. Data source: World Development Indicator. Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

**Table F12. Cross-sectional analysis: Other factors**

|                                  | <i>Dependent variable:</i>                     |                    |                       |                    |
|----------------------------------|------------------------------------------------|--------------------|-----------------------|--------------------|
|                                  | Mean absolute value of error terms (1992-2018) |                    |                       |                    |
|                                  | (1)                                            | (2)                | (3)                   | (4)                |
| European descent                 | -0.0011<br>(0.0007)                            |                    |                       |                    |
| Ethnic_Fractionalization_in_2000 |                                                | 0.1701<br>(0.1134) |                       |                    |
| Language Fractionalization       |                                                |                    | 0.3006***<br>(0.0965) |                    |
| Religion Fractionalization       |                                                |                    |                       | 0.0302<br>(0.1134) |
| Observations                     | 146                                            | 166                | 159                   | 167                |
| R <sup>2</sup>                   | 0.0166                                         | 0.0165             | 0.0635                | 0.0004             |
| Adjusted R <sup>2</sup>          | 0.0098                                         | 0.0105             | 0.0575                | -0.0056            |

*Note:* Each column reports OLS coefficient estimates from a country-level regression of mean absolute errors (1992-2018) on the variables shown. Data source: Quality of Government (columns 2-4), column 1. (Nunn and Puga 2012). Robust standard errors are in parentheses. \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.