# ONLINE APPENDIX

David D. Laitin and Rajesh Ramachandran*

August 2015

## Organization of the online appendix

1. Section A.1 provides information on the data sources for the cross-country regressions and the micro studies.

2. Section A.2 lays out the methodological details for analyzing the extent of omitted variable bias.

3. Section A.3 provides the methodology underlying an alternative instrumental variable strategy shown in Table A.13

4. Section A.4 provides the formal exposition of the theoretical framework outlined in section 2.2 of the main text.

5. The following tables and figures are included in the online Appendix:

   (a) Table A.1 examines the robustness of the effect of average distance to alternative values of $\lambda$.

*Laitin: Department of Political Science, Stanford University, Stanford, CA 94305 (email:dlaitin@stanford.edu). Ramachandran: Department of Microeconomics and Management, Goethe University Frankfurt, Grüneburgplatz 1, 60323 Frankfurt am Main, Germany. (email:ramachandran@econ.uni-frankfurt.de).

(b) Table A.2 examines the robustness of the effect of average distance from official language to alternative measures of ethno-linguistic fractionalization.

(c) Table A.3 examines the robustness of the effect of average distance from official language to the addition of controls for temperature, rainfall and agricultural land suitability.

(d) Table A.4 examines the robustness of the effect of average distance from official language to the addition of controls for natural resources and geography.

(e) Table A.5 splits the sample into countries obtaining a share of greater than and less than 10 percent of GDP from natural resources to show that the effect of average distance from official language is more important for countries not dependent heavily on natural resources.

(f) Table A.6 examines the robustness of the effect of average distance from official language to the addition of controls for alternative measures of institutions and share of population of European descent in 1975.

(g) Table A.7 shows the regressions of average distance on Human Development Index holding constant the number of observations.

(h) Table A.8 shows the estimated lower and upper bounds of the coefficient on average distance when accounting for omitted variables. It also estimates the required strength of unobservables relative to observables for the coefficient on average distance from official language to become equal to zero.

(i) Table A.9 shows that average distance from official language is a significant predictor of life expectancy, log GDP per capita, log output per worker and zHDI when restricting the sample to only the African continent.

(j) Table A.10 examines the robustness of the effect of average distance from official language to account for the interests of the country's entrenched elites as measured

by the average duration of a leader in power.

(k) Table A.11 shows that results are robust to including a control for having a writing tradition.

(l) Table A.12 shows that the IV results in the main paper are robust to including a control for genetic diversity, genetic diversity squared and latitude.

(m) Table A.13 shows the results of our alternative instrumental variable analysis, using the share of population of partitioned ethnicities as an instrument for average distance from official language.

(n) In Figure A.1 are shown the average usage of English at home by socio-economic status and education level of parents.

(o) Figure A.2 shows the effect of exposure to English on English scores for each country in the sample.

(p) Figure A.3 shows the effect of exposure to English on Math scores for each country in the sample.

6. Data on the official language/s of countries included in the sample, the average distance from the official language, information on writing tradition, and the identity of the former colonial rulers are provided in the Excel file included in the package.

7. Data on the year of independence and the year from which the GDP data has been used are provided in the Excel file included in the package.

## A.1 Data sources

### A.1.1 Data sources for the cross-country regression

- Data on the number and size of ethnic groups comes from Fearon (2003).
- The data on Human Development Index (from 2010) is from the United Nations Development Report Programme (UNDP, 2011).
- GDP per capita (from 2005) is from the World Development Indicators (World Bank, 2014).
- Data on GDP per capita at independence comes from the Maddison Project Database (Bolt and Van Zanden, 2013) and the Penn World Tables (Heston et al., 2012).
- Data on log output per worker is from Hall and Jones (1999).
- Data on life expectancy and infant mortality rate is from the year 2010 and from the World Bank Database.
- Data on poverty headcount ratio is from the World Bank database. The data is from the latest year available from the period between 2000 and 2010.
- Data on predicted genetic diversity and diversity squared, years of schooling, institutionalized democracy score, temperature, precipitation, executive constraints, social infrastructure, log population in 1500, average land suitability for agriculture and legal origins is from Ashraf and Galor (2013).
  (Refer to www.aeaweb.org/aer/data/feb2013/20100971_app.pdf for further details.)
- Data on natural resources is from Acemoglu et al. (2001).
- Data for colonial dummies (whether country was ever a colony and if so, the former metropole) comes from Treisman (2007).

- Institutional quality data comes from Political Risk Services Group (PRS Group [Distributor] V1 [Version], 2010) averaged over the years 1995-2005.

- The data on the index of ethno-linguistic fractionalization based on list of groups from Fearon (2003) and not accounting for distance comes from the dataset of Esteban et al. (2012).

- The data on the index of polarization of Esteban, Mayoral and Ray based on list of groups from Fearon (2003) comes from the dataset of Esteban et al. (2012).

- The data on the index of ethno-linguistic fractionalization based on list of groups from Ethnologue and accounting for distance between groups comes from the dataset of Desmet et al. (2009).

- The data on the index of ethno-linguistic fractionalization based on list of groups from Ethnologue and not accounting for distance between groups comes from the dataset of Desmet et al. (2012).

- The share of population comprising partitioned ethnicities comes from the dataset of Alesina et al. (2011)

### A.1.2 Data source for the micro study on the individual distance channel

- International Institute for Population Sciences (IIPS) and Macro International. 2007. National Family Health Survey (NFHS-3), 2005-06: India: Volume II. Mumbai: IIPS.

### A.1.3 Data source for the micro evidence on the exposure channel

- The data for the evidence on the exposure channel comes from Southern and Eastern Africa Consortium for Monitoring Educational Quality. SACMEQ II Project 2000-2004

[dataset]. Version 4. Harare: SACMEQ [producer], 2004. Paris: International Institute for Educational Planning, UNESCO [distributor], 2010.

## A.2 Methodological Concerns

### A.2.1 Omitted variable bias

The documented correlation between average distance and HDI, in section 2.2 and 2.3 of the main text, could be a result of some omitted variable that affects both the measure of language distance and the HDI. Thus the observed negative correlation could be an artifact of this omitted/missing variable rather than the effect of language policy. To examine this we use the test suggested by Oster (2013), which builds upon the methodology of Altonji et al. (2005) that selection on observables can be used to assess the potential bias from unobservables. The key underlying assumption under the Altonji, Elder, and Taber (2005) test is that all of the unobservables share the same covariance properties as the observables. Oster introduces a less restrictive assumption, namely, the assumption of proportional selection. To see what this assumption implies consider the following model $Y = \beta X + W_1 + W_2$, where $W_1$ is observed, $W_2$ is unobserved and $\beta$ is the coefficient of interest. The proportional selection assumption states $\frac{Cov(X,W_2)}{Var(W_2)} = \delta \frac{Cov(X,W_1)}{Var(W_1)}$ i.e. the relationship between X and the observable index is informative about the relationship between X and the unobservable index. This link invokes a degree of proportionality, denoted $\delta$. Moreover under the Altonji, Elder, and Taber methodology the coefficient movements are used as the statistic to calculate the bias whereas Oster shows that coefficient movements alone are not a sufficient statistic to calculate bias. The omitted variable bias is proportional to coefficient movements, but only if such movements are scaled by movements in R-squared.

The regression of average distance on HDI holding number of observations constant is shown in Table A.7.

Insert Table A.7

Let $\hat{\beta}_R$ and $R_R$ be the coefficient on the variable of interest and the associated R-squared value, respectively, for the regression with no controls. Let $\hat{\beta}_F$ and $R_F$ be the coefficient on the vari-

able of interest and the associated R-squared, respectively, for the regression with all available controls. Moreover let us denote by $R_{max}$ the associated R-squared for the hypothetical regression with all controls. Now the identified set of $\beta$ can be shown to lie in the interval $\beta \in (\hat{\beta_F}, \hat{\beta_F} - \delta \frac{(\hat{\beta_R} - \hat{\beta_F})(R_{max} - R_F)}{(R_F - R_R)})$.

Insert Table A.8

The values of $\hat{\beta_F}, \hat{\beta_R}, R_F, R_R$ taken from Table A.7 are shown in column (1) and (2) of Table A.8. Assuming $\delta$ is equal to 1, which implies that the observables are at least as important as the unobservables in explaining cross-country differences in the HDI and assuming values for $R_{max}$ equal to 0.78, 0.80 and 0.85, the identified set of $\beta$ is calculated and shown in column (3). The identified set pertaining to the three values of $R_{max}$ are seen to be $[-0.185, -0.202]$, $[-0.170, -0.202]$ and $[-0.130, -0.202]$, i.e. all three exclude zero and the lower bound is reasonably close to the coefficient identified in the regression with all available controls. The final column (4) calculates what would have to be the strength of unobservables relative to observables for the coefficient on average distance from official language to become equal to zero for the three assumed values of $R_{max}$. It is seen that the explanatory power of the unobservables would have to be about 2.8 to 11 times stronger relative to the observables, which seems highly unlikely.

## A.3 An instrumental variable approach

This section provides further evidence that the documented relationship between ADOL and socio-economic development is indeed causal, by using an instrumental variable strategy distinct from the one provided in section 2.7 of the main text.

The regressions in Table IX of the main paper show that (ethno)linguistic fractionalization is an important determinant of ADOL. The link between linguistic diversity and official language choice arises as increasing diversity amplifies the problem of coordinating on the choice of an indigenous language, and increases the probability of maintaining the status quo, i.e. the colonial language remaining official. Assume first, that decision making rules of official language choice are such that the probability of a group's language being chosen as official is a non-decreasing function of their population share, and second, instituting a language as official requires unanimity or some form of a minimum winning coalition. The two assumptions will imply that the probability of a particular group's language being chosen as official decreases as population share decreases.[1] Due to this fact the expected payoff for any linguistic group participating in a game of official language choice, especially small-sized ones, reduces as linguistic diversity increases. Another channel is as the number of groups increase, implying diversity increases, it makes the commitment problem of recompensing groups whose language is not chosen harder to solve.[2] Thus higher levels of linguistic diversity tend to increase the ADOL.

One exogenous factor that has contributed to this increase in linguistic diversity in Africa has been the partitioning of Africa into spheres of influence, protectorates and colonies by the European powers at the Berlin conference of 1884-85. There is widespread agreement that the borders were arbitrarily drawn with little knowledge about ethnic homelands, and resulted in

[1] Decreasing population share, normally, would translate into increasing linguistic diversity.

[2] We develop these two points more fully in a companion paper, where the problem of choosing an official language for post-colonial multilingual states is theoretically modeled as one of coordination in a society with $n$-linguistic groups. [Citation removed for review purposes]

ethnic groups being partitioned across national borders. For instance Englebert et al. (2002) estimate that the share of partitioned groups is on an average more than 40 percent of the total population of Sub-Saharan Africa.

One mechanical consequence of partitioning ethnicities is the associated increase in linguistic fractionalization. Our theory predicts and empirical evidence (in Table IX of the main paper) shows that an increase in linguistic fractionalization increases the distance to the official language by increasing the probability of retaining the colonial language. We thus use the share of population belonging to partitioned ethnicities from the work of Alesina et al. (2011) as an instrument for ADOL. We are here assuming that the instrumental variable is statistically independent of the outcomes of interest, conditional on controlling for levels of linguistic fractionalization. Thus the key assumption is that share of partitioned ethnicities has an effect on socio-economic development only through the channel of language choice, as the Greenberg index of linguistic diversity accounts for all other effects it has through the channel of increasing fractionalization in society. The results are shown in Table A.13.

Insert Table A.13

Columns (1), (3), (5), and (7) regress life expectancy, log GDP per capita, log output per worker and zHDI, respectively, on ADOL instrumented for by the share of population comprising partitioned ethnicities, controlling for the levels of linguistic diversity using the Greenberg index.[3] In Panel (B) are shown the first stage regressions of share of partitioned ethnicities in the total population on ADOL. Although the share of partitioned ethnicities is a statistically significant predictor of ADOL, the F-statistics are seen to lie in the range of 4.63-14.4. This suggests we need to be cautious in interpreting our IV estimates, as there is the potential problem of a weak instrument. In panel A are shown the results of the second stage; we see that ADOL is a statistically significant predictor of Log GDP per capita, log output per worker and zHDI. The

[3]For the dependent variable cognitive test scores, there are only 6 African countries in the sample and hence that effect can not be estimated econometrically.

coefficient on ADOL for the dependent variable life expectancy turns statistically insignificant at the conventional level ($p - value = 0.14$), due to the small sample size, though the point estimate is negative and the beta coefficient quite large.

In columns (2), (4), (6), and (8) we add other controls outlined in section 2.4 of the main text - constraints on the executive and log GDP per capita at independence. Again the ADOL is seen to be a statistically significant predictor of log GDP per capita, log output per worker and zHDI. It is important to stress that the main objective of this exercise is to show that results using alternative approaches, here the IV methodology, are in line with the theoretically motivated cross-country regressions, and bolster our claim that the correlations we have documented indeed uncover something causal. Our intention is not to claim that the point estimate arising from the IV regressions are the actual quantitative effect of ADOL, as our sample size is small and the instrument potentially weak.

## A.4 Theoretical framework

### A.4.1 The basic framework

Consider an economy where the total output $Y$ is a function of the aggregate level of (physical and mental) human capital $H$ in the society and is given by:

$$Y = F(H) = (H)^{\alpha}, \text{where } F_1(H) > 0 \text{ and } F_{11}(H) \leq 0. \tag{1}$$

It is assumed that the markets are competitive and the wages are given by:

$$W = \alpha H^{\alpha-1} \tag{2}$$

Moreover assume that each individual $i$ has an ability given by $a_i$ and chooses $h_i$ to maximize his utility given by:

$$U(h_i) = Wh_i - (h_i)^2 C(a_i, d_{io}, e_{io}), \tag{3}$$

where the function $C$ represents the cost of obtaining human capital and is assumed to depend upon the ability $a_i$, distance $d_{ioj}$, of individual $i$ from its official language $o$ and to the amount of exposure of individual $i$ to the official language $o$ i.e. $e_{io}$. The two underlying assumptions are that greater the distance (d) of the individual $i$ to the official language $o$ the higher the cost of obtaining human capital and participating in the economy i.e.

$$\frac{dC}{dd_{io}} = \frac{df(d_{io}, e_{io})}{dd_{io}} > 0. \tag{4}$$

and greater exposure (e) to the official language, the lower the costs of obtaining human capital and participation in the economy i.e.

$$\frac{dC}{de_{io}} = \frac{df(d_{io}, e_{io})}{de_{io}} < 0. \tag{5}$$

Taking the first order condition in Equation 3 with respect to $h_i$ gives us:

$$h_i^* = \frac{W h_i'}{C(a_i, d_{io}, e_{io})} \tag{6}$$

The two underlying assumptions given by Equations 4 and 5 in turn imply:

$$\frac{dh_i}{dd_{oi}} < 0 \text{ and } \frac{dh_i}{de_{oi}} > 0 \tag{7}$$

i.e. individual outcomes (here labeled as human capital) are improving in reduced language distance from the official language and improving in increased exposure to the official language as they both reduce the costs of participating in the economy. We can now denote the output at the country level by:

$$Y = \int_{a_i} \frac{W h_i'}{C(a_i, d_{io}, e_{io})} \tag{8}$$

As $Y$ is strictly increasing in $h_i$, in light of Equation 7, this implies:

$$\frac{dY}{dd_{oi}} < 0 \text{ and } \frac{dY}{de_{oi}} > 0 \tag{9}$$

The above indicates that individual level distance and exposure will determine observed country level outcomes as seen in the cross-country framework. The calculation of the distance at the country level implies that the measure captures and subsumes the concept of both individual distance and average exposure to the official language in the same indicator.[4] It is not therefore empirically possible to disentangle and measure the separate contribution of individual distance and exposure on the dependent variable in the cross-country framework.

[4]In the cross-country analysis we attribute the distance of other ethnic groups ($i \neq j$) in the country to be a measure of exposure of the ethnic group $i$ to the official language. As the measure takes into account the distance of all ethnic groups, the concept of both individual/group distance and exposure is captured by the same measure.

# Additional Tables

# and

# Figures

Table A.1: **Using alternative values of $\lambda$ in measure of average distance from official language to check senstitivity of results to choice of $\lambda$**

| | (1) $\lambda = \mathbf{0.50}$ | (2) $\lambda = \mathbf{0.05}$ | (3) $\lambda = \mathbf{0.10}$ | (4) $\lambda = \mathbf{0.20}$ | (5) $\lambda = \mathbf{0.30}$ | (6) $\lambda = \mathbf{0.40}$ | (7) $\lambda = \mathbf{0.60}$ | (8) $\lambda = \mathbf{0.70}$ |
|---|---|---|---|---|---|---|---|---|
| Average Distance from Official Language | -1.117*** | -0.861*** | -0.909*** | -0.987*** | –1.046*** | -1.088*** | -1.137*** | -1.149*** |
| | (0.260) | (0.264) | (0.264) | (0.264) | (0.263) | (0.262) | (0.258) | (0.255) |
| | [-0.415] | [-0.326] | [-0.342] | [-0.369] | [-0.390] | [-0.404] | [-0.421] | [-0.426] |
| Linguistic fractionalization a/c for distance | -0.131 | -0.346 | -0.311 | -0.249 | -0.199 | -0.161 | -0.108 | -0.0909 |
| | (0.278) | (0.276) | (0.276) | (0.276) | (0.277) | (0.278) | (0.279) | (0.279) |
| | [-0.0271] | [-0.0717] | [-0.0643] | [-0.0515] | [-0.0413] | [-0.0333] | [-0.0224] | [-0.0188] |
| Executive constraints | 0.127*** | 0.122*** | 0.122*** | 0.124*** | 0.125*** | 0.126*** | 0.128*** | 0.129*** |
| | (0.0278) | (0.0291) | (0.0289) | (0.0285) | (0.0282) | (0.0280) | (0.0276) | (0.0275) |
| | [0.250] | [0.240] | [0.241] | [0.243] | [0.246] | [0.248] | [0.252] | [0.254] |
| Log GDP per capita at independence | 0.243*** | 0.237*** | 0.238*** | 0.240*** | 0.242*** | 0.243*** | 0.244*** | 0.244*** |
| | (0.0554) | (0.0571) | (0.0569) | (0.0566) | (0.0562) | (0.0558) | (0.0550) | (0.0547) |
| | [0.215] | [0.210] | [0.211] | [0.212] | [0.214] | [0.215] | [0.216] | [0.216] |
| Continent Dummies | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 149 | 149 | 149 | 149 | 149 | 149 | 149 | 149 |
| R-squared | 0.758 | 0.746 | 0.748 | 0.751 | 0.754 | 0.756 | 0.759 | 0.760 |

a.Linguistic fractionalization a/c for distance is the measure of ELF accounting for distance between groups from Fearon (2003).
b. Robust standard errors are shown in the parenthesis.
c. *, ** and *** significant at 10, 5 and 1 % significance level respectively.
d. In the square brackets are shown the standardized coefficients.

Table A.2: **Checking robustness of the effect of average distance from official language to using alternative measures of ELF**

| Dependent variable - zHDI in 2010 | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Average Distance from Official Language | -1.117*** | -0.852*** | -0.895*** | -1.186*** | -0.916*** |
| | (0.260) | (0.203) | (0.238) | (0.256) | (0.230) |
| | [-0.415] | [-0.311] | [-0.328] | [-0.443] | [-0.334] |
| Linguistic fractionalization a/c for distance | -0.131 | | | | |
| | (0.278) | | | | |
| | [-0.0271] | | | | |
| ELF not accounting for distance | | -0.361 | | | -0.416 |
| (list of groups from Fearon 2003) | | (0.274) | | | (0.276) |
| | | [-0.0841] | | | [-0.0970] |
| ELF not accounting for distance | | | -0.194 | | |
| (list of groups from Ethnologue) | | | (0.234) | | |
| | | | [-0.0584] | | |
| ELF accounting for distance | | | | 0.00536 | |
| (list of groups from Ethologue) | | | | (0.301) | |
| | | | | [0.000938] | |
| Polarization measure from Esteban, Mayoral and Ray | | | | | 0.940 |
| (list of groups from Fearon 2003) | | | | | (0.981) |
| | | | | | [0.0506] |
| Executive Constraints | Yes | Yes | Yes | Yes | Yes |
| Log GDP per capita at Independence | Yes | Yes | Yes | Yes | Yes |
| Continent Dummies | Yes | Yes | Yes | Yes | Yes |
| Observations | 149 | 134 | 133 | 148 | 134 |
| R-squared | 0.758 | 0.781 | 0.775 | 0.754 | 0.782 |

a. Column (1) reports the baseline specification corresponding to column (5) of Table (4) in the main text.
b. Linguistic fractionalization a/c for distance is the measure of ELF accounting for distance between groups from Fearon (2003).
c. Robust standard errors are shown in the parenthesis.
d. *, ** and *** significant at 10, 5 and 1 % significance level respectively.
e. In the square brackets are shown the standardized coefficients.
f. The two measure of ELF based on the list of groups from the Ethnologue comes from the data of Desmet et. al (2009) and Desmet et. al (2012).

f. The data on ELF measures based on ethnic groups of Fearon (2003) and the polarization measure comes from the data of Esteban, Mayoral and Ray (2012)

Table A.3: **Robustness of measure of average distance to addition of temperature, precipitation and land suitability of agriculture**

| Dependent variable - zHDI in 2010 | (1) | (2) | (3) |
|---|---|---|---|
| Average distance from official language | -1.117*** | -0.850*** | -0.883*** |
| | (0.260) | (0.282) | (0.276) |
| | [-0.415] | [-0.315] | [-0.328] |
| Linguistic fractionalization a/c for distance | -0.131 | -0.249 | -0.313 |
| | (0.278) | (0.285) | (0.276) |
| | [-0.0271] | [-0.0516] | [-0.0647] |
| Executive constraints | 0.127*** | 0.135*** | 0.146*** |
| | (0.0278) | (0.0270) | (0.0271) |
| | [0.250] | [0.265] | [0.280] |
| Log GDP per capita at independence | 0.243*** | 0.194*** | 0.210*** |
| | (0.0554) | (0.0616) | (0.0546) |
| | [0.215] | [0.171] | [0.183] |
| Log [temperature] | | -0.230 | |
| | | (0.191) | |
| | | [-0.0577] | |
| Log [precipitation] | | -0.137* | |
| | | (0.0706) | |
| | | [-0.123] | |
| Log [land suitability for agriculture] | | | -0.0810* |
| | | | (0.0419) |
| | | | [-0.103] |
| Continent Dummies | Yes | Yes | Yes |
| Observations | 149 | 149 | 143 |
| R-squared | 0.758 | 0.771 | 0.776 |

a.Linguistic fractionalization a/c for distance is the measure of ELF accounting for distance between groups from Fearon (2003).
b. Robust standard errors are shown in the parenthesis.
c. *, ** and *** significant at 10, 5 and 1 % significance level respectively.

d. In the square brackets are shown the standardized coefficients.

Table A.4: **Robustness of measure of average distance to addition of natural resources and geographical controls**

| Dependent variable - zHDI in 2010 | (1) | (2) | (3) |
|---|---|---|---|
| Average distance from official language | -1.117*** | -1.078*** | -1.065*** |
| | (0.260) | (0.254) | (0.277) |
| | [-0.415] | [-0.401] | [-0.394] |
| Linguistic fractionalization a/c for distance | -0.131 | -0.312 | -0.119 |
| | (0.278) | (0.286) | (0.284) |
| | [-0.0271] | [-0.0645] | [-0.0244] |
| Executive constraints | 0.127*** | 0.133*** | 0.122*** |
| | (0.0278) | (0.0322) | (0.0307) |
| | [0.250] | [0.255] | [0.237] |
| Log GDP per capita at independence | 0.243*** | 0.235*** | 0.247*** |
| | (0.0554) | (0.0614) | (0.0559) |
| | [0.215] | [0.188] | [0.217] |
| Percent of World Gold Reserves | | 0.00904** | |
| | | (0.00389) | |
| | | [0.0369] | |
| Percent of World Iron Reserves | | -0.0561 | |
| | | (0.0349) | |
| | | [-0.0954] | |
| Percent of World Silver Reserves | | 0.0591** | |
| | | (0.0263) | |
| | | [0.122] | |
| Percent of World Zinc Reserves | | 0.0316 | |
| | | (0.0409) | |
| | | [0.0676] | |
| Percent of World Oil Reserves | | 7.00e-08*** | |
| | | (2.14e-08) | |
| | | [0.103] | |
| Log [absolute latitude] | | | 0.00661 |
| | | | (0.0685) |
| | | | [0.00632] |
| Dummy for Landlocked | | | -0.312*** |
| | | | (0.0950) |
| | | | [-0.127] |
| Continent Dummies | Yes | Yes | Yes |
| Observations | 149 | 136 | 144 |
| R-squared | 0.758 | 0.785 | 0.774 |

a.Linguistic fractionalization a/c for distance is the measure of ELF accounting for distance between groups from Fearon (2003).
b. Robust standard errors are shown in the parenthesis.
c. *, ** and *** significant at 10, 5 and 1 % significance level respectively.

d. In the square brackets are shown the standardized coefficients.

Table A.5: **Effect of average distance on a split sample - Countries with share of GDP from natural resources with greater than and less than 10 percent**

| Dependent variable - Log GDP per capita in 2005 | (1) | (2) | (3) |
|---|---|---|---|
| Average distance from official language | -1.354*** | -0.975 | -1.514*** |
| | (0.390) | (0.813) | (0.436) |
| | [-0.383] | [-0.268] | [-0.418] |
| Linguistic fractionalization a/c for distance | 0.0519 | 0.137 | -0.307 |
| | (0.408) | (0.791) | (0.472) |
| | [0.00821] | [0.0195] | [-0.0479] |
| Executive constraints | 0.192*** | 0.0358 | 0.275*** |
| | (0.0463) | (0.149) | (0.0468) |
| | [0.289] | [0.0377] | [0.419] |
| Log GDP per capita at independence | 0.374*** | 0.792*** | 0.0139 |
| | (0.116) | (0.160) | (0.0946) |
| | [0.254] | [0.599] | [0.00906] |
| Continent Dummies | Yes | Yes | Yes |
| Observations | 149 | 40 | 109 |
| R-squared | 0.623 | 0.707 | 0.706 |

b. Column (1) considers the entire sample. Column (2) considers counties whose share of GDP from natural resources is greater than 10 percent, whereas column (3) considers those with less than 10 percen .
b. Linguistic fractionalization a/c for distance is the measure of ELF accounting for distance between groups from Fearon (2003).
c. Robust standard errors are shown in the parenthesis.
d. *, ** and *** significant at 10, 5 and 1 % significance level respectively.

e. In the square brackets are shown the standardized coefficients.

Table A.6: **Robustness of measure of average distance to alternative measure of institutions and share of population of European descent**

| Dependent variable - zHDI in 2010 | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Average distance from official language | -1.117*** | -0.930*** | -0.853*** | -1.057*** |
| | (0.260) | (0.257) | (0.200) | (0.290) |
| | [-0.415] | [-0.355] | [-0.316] | [-0.387] |
| Linguistic fractionalization a/c for distance | -0.131 | 0.0800 | -0.0944 | -0.135 |
| | (0.278) | (0.227) | (0.251) | (0.283) |
| | [-0.0271] | [0.0168] | [-0.0190] | [-0.0273] |
| Executive constraints | 0.127*** | | | 0.121*** |
| | (0.0278) | | | (0.0290) |
| | [0.250] | | | [0.230] |
| Log GDP per capita at independence | 0.243*** | 0.171*** | 0.275*** | 0.226*** |
| | (0.0554) | (0.0418) | (0.0678) | (0.0597) |
| | [0.215] | [0.156] | [0.168] | [0.195] |
| Avg. Protection against Expropriation risk | | 2.812*** | | |
| | | (0.278) | | |
| | | [0.481] | | |
| Social infrastructure | | | 1.559*** | |
| | | | (0.262) | |
| | | | [0.364] | |
| % of European descent in 1975 | | | | 0.00634* |
| | | | | (0.00340) |
| | | | | [0.268] |
| Continent Dummies | Yes | Yes | Yes | Yes |
| Observations | 149 | 127 | 112 | 137 |
| R-squared | 0.758 | 0.850 | 0.833 | 0.768 |

a.Linguistic fractionalization a/c for distance is the measure of ELF accounting for distance between groups from Fearon (2003).
b. Robust standard errors are shown in the parenthesis.
c. *, ** and *** significant at 10, 5 and 1 % significance level respectively.

d. In the square brackets are shown the standardized coefficients.

Table A.7: **Regressions of distance on HDI holding number of observations constant**

| Dependent variable - HDI in 2010 | (1) | (2) | (3) | (4) | (5) ) |
|---|---|---|---|---|---|
| Average distance from official language | -0.362*** | -0.383*** | -0.300*** | -0.266*** | -0.202*** |
| | (0.0238) | (0.0285) | (0.0302) | (0.0289) | (0.0471) |
| | [-0.743] | [-0.786] | [-0.615] | [-0.545] | [-0.415] |
| Linguistic fractionalization a/c for distance | | 0.0657 | 0.0445 | -0.00414 | -0.0237 |
| | | (0.0637) | (0.0572) | (0.0508) | (0.0504) |
| | | [0.0751] | [0.0509] | [-0.00473] | [-0.0271] |
| Executive constraints | | | 0.0360*** | 0.0310*** | 0.0230*** |
| | | | (0.00479) | (0.00429) | (0.00502) |
| | | | [0.391] | [0.337] | [0.250] |
| Log GDP per capita at independence | | | | 0.0528*** | 0.0441*** |
| | | | | (0.00933) | (0.0100) |
| | | | | [0.258] | [0.215] |
| Continent Dummies | No | No | No | No | Yes |
| Observations | 149 | 149 | 149 | 149 | 149 |
| R-squared | 0.552 | 0.556 | 0.684 | 0.742 | 0.758 |

a.Linguistic fractionalization a/c for distance is the measure of ELF accounting for distance between groups from Fearon (2003).
b. Robust standard errors are shown in the parenthesis.
c. *, ** and *** significant at 10, 5 and 1 % significance level respectively.

d. In the square brackets are shown the standardized coefficients corresponding to the equation in column (5).

Table A.8: **The Oster test: Selection on unobservables and identified set**

| Dependent variable - zHDI in 2010 | | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Treatment Variable | Baseline Effect (Std. Error) [$R^2$] | Controlled Effect (Std. Error) [$R^2$] | Identified Set | $\delta$ for $\beta = 0$ given $R_{max}$ |
| Average Distance from Official Language $R_{max} = 0.78$ | -0.362*** (0.0238), [0.552] | -0.202*** (0.0471), [0.758] | [-0.185, -0.202]$^+$ | 11.82 |
| Average Distance from Official Language $R_{max} = 0.80$ | -0.362*** (0.0238), [0.552] | -0.202*** (0.0471), [0.758] | [-0.170, -0.202]$^+$ | 6.19 |
| Average Distance from Official Language $R_{max} = 0.85$ | -0.362*** (0.0238), [0.552] | -0.202*** (0.0471), [0.758] | [-0.130, -0.202]$^+$ | 2.82 |

a. The most restricted equation controls only for average distance from official language and the fully specified equation is given by column (5) in Table A.7.
b. The standard errors are shown in the parenthesis and the R-squared in the square bracket.
c.The identified set in Column (3) is bounded below by $\hat{\beta_F}$ and above by $\beta^*$ calculated based on the denoted $R_{max}$ and $\delta = 1$.
d. The identified set excludes zero for all three assumed $R_{max}$.
e. Column (4) shows the value of $\delta$ at which $\beta$ goes to zero.

f. *** p<0.01, ** p<0.05, * p<0.1

Table A.9: **Regressions of distance on life expectancy, log GDP per capita, log output per worker and zHDI in 2010 - Only African Continent**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Life Expt. in 2010 | log GDP per capita | log Output per worker | zHDI in2010 |
| Average distance from official language | -9.481** | -1.040*** | -0.854** | -1.325*** |
| | (4.513) | (0.379) | (0.385) | (0.350) |
| | [-0.339] | [-0.296] | [-0.336] | [-0.548] |
| Linguistic fractionalization a/c for distance | -4.330 | -0.130 | -0.147 | -0.230 |
| | (4.937) | (0.553) | (0.350) | (0.390) |
| | [-0.137] | [-0.0287] | [-0.0470] | [-0.0750] |
| Executive constraints | 0.476 | 0.155 | 0.0825 | 0.0723 |
| | (0.675) | (0.194) | (0.0768) | (0.0613) |
| | [0.0847] | [0.188] | [0.150] | [0.132] |
| Log GDP per capita at independence | 2.998* | 0.799*** | 0.741*** | 0.584*** |
| | (1.573) | (0.258) | (0.138) | (0.142) |
| | [0.226] | [0.419] | [0.546] | [0.444] |
| Percent of World Gold Reserves | | 0.185*** | | |
| | | (0.0347) | | |
| | | [1.354] | | |
| Percent of World Iron Reserves | | -1.800*** | | |
| | | (0.194) | | |
| | | [-1.137] | | |
| Percent of World Zinc Reserves | | -0.360*** | | |
| | | (0.105) | | |
| | | [-0.200] | | |
| Percent of World Oil Reserves | | 4.08e-07*** | | |
| | | (1.16e-07) | | |
| | | [0.287] | | |
| HIV prevalence in 2000 | -0.440*** | | | |
| | (0.141) | | | |
| | [-0.485] | | | |
| Observations | 45 | 44 | 42 | 46 |
| R-squared | 0.420 | 0.639 | 0.460 | 0.529 |

$*p < .10; **p < .05; ***p < .01$. Robust SE's in parenthesis and standardized coefficients in square brackets.

Table A.10: **Distinguishing between general elite interests and role of language policy - showing importance of language policy independent of the constraints on development of entrenched elites**

| Dependent variable - zHDI in 2010 | (1) | (2) | (3) |
|---|---|---|---|
| Average distance from official language | -1.117*** | -1.088*** | -1.088*** |
| | (0.260) | (0.241) | (0.241) |
| | [-0.415] | [-0.406] | [-0.406] |
| Linguistic fractionalization a/c for distance | -0.131 | -0.195 | -0.195 |
| | (0.278) | (0.273) | (0.273) |
| | [-0.0271] | [-0.0405] | [-0.0405] |
| Executive constraints | 0.127*** | 0.141*** | 0.141*** |
| | (0.0278) | (0.0292) | (0.0292) |
| | [0.250] | [0.279] | [0.279] |
| Log GDP per capita at independence in 1990 US | 0.243*** | 0.212*** | 0.212*** |
| | (0.0554) | (0.0510) | (0.0510) |
| | [0.215] | [0.188] | [0.188] |
| Log duration of Leader in power (No. of days) | | 0.188** | |
| | | (0.0739) | |
| | | [0.136] | |
| Log of the squared Duration of Leader in power (No. of days) | | | 0.0938** |
| | | | (0.0370) |
| | | | [0.136] |
| Continent Dummies | Yes | Yes | Yes |
| Observations | 149 | 147 | 147 |
| R-squared | 0.758 | 0.772 | 0.772 |

a. Column (1) reports the baseline specification corresponding to column (5) of Table (4) in the main text.
b. Linguistic fractionalization a/c for distance is the measure of ELF accounting for distance between groups from Fearon (2003).
c. Robust standard errors are shown in the parenthesis.
d. *, ** and *** significant at 10, 5 and 1 % significance level respectively.
e. In the square brackets are shown the standardized coefficients.

f. The data on leader duration comes from Archigos dataset. The dataset has been accessed at

www.rochester.edu/college/faculty/hgoemans/data.htm

Table A.11: **Regressions of distance on zHDI in 2010 with additional control for having a writing tradition.**

| | (1) | (2) | (3) |
|---|---|---|---|
| Average distance from official language with delta 0.50 | -1.117*** | -0.856*** | -0.764** |
| | (0.260) | (0.275) | (0.319) |
| | [-0.415] | [-0.318] | [-0.283] |
| Linguistic fractionalization a/c for distance | -0.131 | -0.273 | -0.327 |
| | (0.278) | (0.290) | (0.314) |
| | [-0.0271] | [-0.0566] | [-0.0677] |
| Executive constraints | 0.127*** | 0.126*** | 0.151*** |
| | (0.0278) | (0.0279) | (0.0301) |
| | [0.250] | [0.248] | [0.296] |
| Log GDP per capita at independence in 1990 US | 0.243*** | 0.240*** | |
| | (0.0554) | (0.0543) | |
| | [0.215] | [0.212] | |
| Written tradition dummy | | 0.336 | 0.329 |
| | | (0.220) | (0.251) |
| | | [0.142] | [0.137] |
| State Antiquity Index | | | 0.166 |
| | | | (0.277) |
| | | | [0.0395] |
| Continent Dummies | Yes | Yes | Yes |
| Observations | 149 | 149 | 136 |
| R-squared | 0.758 | 0.762 | 0.753 |

*p < .05; **p < .01; ***p < .001. Robust SE's in parenthesis and standardized coefficients in square brackets.

Table A.12: **IV Regressions with additional controls**

| | (1) Cognitive test score | (2) Cognitive test score | (3) Life Expt. in 2010 | (4) L. Expt. in 2010 | (5) log GDP per capita | (6) log GDP per capita | (7) log Output per worker | (8) log Output per worke | (9) zHDI in2010 | (10) zHDI in2010 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Panel A: Two-Stage Least Squares** | | | | | | | | | | |
| Average distance from official language | -1.52*** | -1.45*** | -24.5*** | -27.1*** | -1.20** | -1.16* | -1.52*** | -1.53*** | -1.27*** | -1.44*** |
| | (0.49) | (0.52) | (3.37) | (3.85) | (0.53) | (0.60) | (0.41) | (0.43) | (0.31) | (0.35) |
| | [-0.67] | [-0.64] | [-0.91] | [-1.01] | [-0.34] | [-0.33] | [-0.55] | [-0.55] | [-0.47] | [-0.53] |
| Linguistic fractionalization a/c for distance | 0.28 | 0.19 | 8.64** | 10.2*** | 0.080 | 0.045 | 0.50 | 0.57 | -0.073 | 0.051 |
| | (0.35) | (0.37) | (3.50) | (3.72) | (0.55) | (0.57) | (0.42) | (0.42) | (0.32) | (0.34) |
| | [0.090] | [0.062] | [0.18] | [0.21] | [0.012] | [0.0070] | [0.099] | [0.11] | [-0.015] | [0.010] |
| Executive constraints | 0.062** | 0.078** | 0.55 | 0.58* | 0.15*** | 0.18*** | 0.095** | 0.11*** | 0.12*** | 0.13*** |
| | (0.031) | (0.030) | (0.33) | (0.35) | (0.051) | (0.052) | (0.040) | (0.041) | (0.031) | (0.032) |
| | [0.21] | [0.27] | [0.11] | [0.11] | [0.23] | [0.27] | [0.17] | [0.20] | [0.22] | [0.24] |
| Log GDP per capita at independence | 0.062 | 0.038 | 1.42** | 1.19* | 0.43*** | 0.40*** | 0.34*** | 0.31*** | 0.27*** | 0.24*** |
| | (0.064) | (0.060) | (0.61) | (0.63) | (0.093) | (0.092) | (0.097) | (0.096) | (0.057) | (0.057) |
| | [0.096] | [0.059] | [0.12] | [0.10] | [0.28] | [0.26] | [0.20] | [0.19] | [0.23] | [0.21] |
| % of European descent in 1975 | 0.00063 | 0.0017 | 0.0084 | 0.0079 | 0.0050* | 0.0042 | 0.0055** | 0.0050** | 0.0046*** | 0.0043** |
| | (0.0018) | (0.0019) | (0.018) | (0.018) | (0.0027) | (0.0027) | (0.0022) | (0.0023) | (0.0017) | (0.0017) |
| | [0.052] | [0.14] | [0.036] | [0.034] | [0.16] | [0.14] | [0.21] | [0.19] | [0.19] | [0.18] |
| America | -0.58*** | -0.60*** | -1.96 | -0.41 | -0.25 | 0.0082 | -0.27 | -0.062 | -0.23 | -0.030 |
| | (0.16) | (0.15) | (1.57) | (1.54) | (0.24) | (0.23) | (0.18) | (0.17) | (0.15) | (0.14) |
| | [-0.35] | [-0.36] | [-0.074] | [-0.015] | [-0.073] | [0.0023] | [-0.10] | [-0.023] | [-0.086] | [-0.011] |
| Predicted genetic diversity | 215* | | -677 | | 109 | | 59.8 | | -25.4 | |
| (ancestry adjusted) | (121) | | (768) | | (118) | | (90.7) | | (73.3) | |
| | [8.53] | | [-1.82] | | [2.20] | | [1.57] | | [-0.66] | |
| Predicted genetic diversity squared | -153* | | 440 | | -81.6 | | -45.8 | | 13.6 | |
| (ancestry adjusted) | (85.5) | | (542) | | (83.7) | | (64.2) | | (51.8) | |
| | [-8.61] | | [1.68] | | [-2.34] | | [-1.71] | | [0.51] | |
| Log [absolute latitude] | | -0.071 | | -0.88 | | 0.068 | | 0.051 | | -0.040 |
| | | (0.082) | | (0.75) | | (0.11) | | (0.078) | | (0.068) |
| | | [-0.097] | | [-0.085] | | [0.050] | | [0.048] | | [-0.038] |
| Observations | 66 | 66 | 139 | 139 | 135 | 135 | 110 | 110 | 137 | 137 |
| R-squared | 0.622 | 0.603 | 0.729 | 0.699 | 0.652 | 0.637 | 0.729 | 0.718 | 0.776 | 0.762 |
| **Panel B: First-Stage for ADOL** | | | | | | | | | | |
| Distance from Site of Invention of Writing | 0.000048*** | 0.000048*** | 0.000072*** | 0.000069*** | 0.000070*** | 0.000065*** | 0.000071*** | 0.000069*** | 0.000072*** | 0.000068*** |
| | (0.000011) | (0.000012) | (9.0e-06) | (0.000010) | (9.0e-06) | (0.000010) | (0.000010) | (0.000011) | (9.0e-06) | (0.000010) |
| | [0.41] | [0.41] | [0.38] | [0.37] | [0.37] | [0.35] | [0.36] | [0.35] | [0.38] | [0.36] |
| Linguistic fractionalization a/c for distance | 0.49*** | 0.52*** | 0.64*** | 0.66*** | 0.66*** | 0.68*** | 0.61*** | 0.62*** | 0.63*** | 0.65*** |
| | (0.12) | (0.13) | (0.091) | (0.099) | (0.091) | (0.099) | (0.11) | (0.11) | (0.091) | (0.10) |
| | [0.36] | [0.38] | [0.35] | [0.37] | [0.36] | [0.38] | [0.33] | [0.34] | [0.35] | [0.36] |
| Executive constraints | 0.0044 | -0.0088 | -0.011 | -0.025* | -0.013 | -0.029** | -0.016 | -0.029** | -0.0100 | -0.026* |
| | (0.014) | (0.015) | (0.012) | (0.013) | (0.012) | (0.013) | (0.014) | (0.014) | (0.012) | (0.013) |
| | [0.034] | [-0.069] | [-0.059] | [-0.13] | [-0.070] | [-0.15] | [-0.081] | [-0.15] | [-0.052] | [-0.13] |
| Log GDP per capita at independence in 1990 | -0.019 | 0.0045 | -0.00071 | 0.012 | -0.0031 | 0.0066 | -0.020 | -0.0039 | -0.00018 | 0.012 |
| | (0.028) | (0.030) | (0.023) | (0.024) | (0.023) | (0.024) | (0.034) | (0.036) | (0.023) | (0.025) |
| | [-0.067] | [0.016] | [-0.0017] | [0.028] | [-0.0072] | [0.016] | [-0.032] | [-0.0065] | [-0.00043] | [0.028] |
| % of European descent in 1975 | -0.0026*** | -0.0032*** | -0.0031*** | -0.0025*** | -0.0030*** | -0.0024*** | -0.0034*** | -0.0031*** | -0.0031*** | -0.0026*** |
| | (0.00063) | (0.00078) | (0.00057) | (0.00067) | (0.00057) | (0.00067) | (0.00067) | (0.00078) | (0.00057) | (0.00068) |
| | [-0.48] | [-0.60] | [-0.35] | [-0.29] | [-0.35] | [-0.28] | [-0.35] | [-0.33] | [-0.36] | [-0.30] |
| America | -0.065 | -0.087 | -0.070 | -0.16*** | -0.068 | -0.16*** | -0.078 | -0.16*** | -0.070 | -0.16*** |
| | (0.076) | (0.073) | (0.058) | (0.053) | (0.058) | (0.052) | (0.064) | (0.055) | (0.058) | (0.053) |
| | [-0.089] | [-0.12] | [-0.070] | [-0.16] | [-0.070] | [-0.16] | [-0.083] | [-0.17] | [-0.071] | [-0.16] |
| Predicted genetic diversity | -144*** | | -95.2*** | | -97.3*** | | -82.6*** | | -105*** | |
| (ancestry adjusted) | (51.8) | | (26.9) | | (27.3) | | (31.3) | | (27.2) | |
| | [-13.0] | | [-6.86] | | [-6.98] | | [-6.01] | | [-7.49] | |
| Predicted genetic diversity squared | 104*** | | 69.0*** | | 70.5*** | | 60.0*** | | 76.0*** | |
| (ancestry adjusted) | (36.5) | | (18.9) | | (19.2) | | (22.0) | | (19.1) | |
| | [13.3] | | [7.06] | | [7.17] | | [6.18] | | [7.70] | |
| Log [absolute latitude] | | 0.065 | | -0.047* | | -0.047* | | -0.028 | | -0.047* |
| | | (0.043) | | (0.026) | | (0.026) | | (0.028) | | (0.026) |
| | | [0.20] | | [-0.12] | | [-0.12] | | [-0.073] | | [-0.12] |
| Observations | 66 | 66 | 139 | 139 | 135 | 135 | 110 | 110 | 137 | 137 |
| R-squared | 0.599 | 0.514 | 0.730 | 0.679 | 0.740 | 0.691 | 0.731 | 0.695 | 0.736 | 0.679 |
| F-Stat | 10.6 | 8.78 | 44.0 | 39.6 | 44.8 | 40.7 | 34.3 | 33.2 | 44.6 | 39.0 |

$*p < .10; **p < .05; ***p < .01$. Robust SE's in parenthesis and standardized coefficients in square brackets.

Table A.13: **IV Regressions of distance on cognitive scores, life expectancy, log GDP per capita, log output per worker and zHDI in 2010 - Using Share of Partitioned Ethnicities**

| | (1) Life Expt. in 2010 | (2) L. Expt. in 2010 | (3) log GDP per capita | (4) log GDP per capita | (5) log Output per worker | (6) log Output per worke | (7) zHDI in2010 | (8) zHDI in2010 |
|---|---|---|---|---|---|---|---|---|
| **Panel A: Two-Stage Least Squares** | | | | | | | | |
| Average distance from official language | -16.0 | -14.3 | -5.99** | -4.72** | -3.83* | -2.91* | -4.43*** | -3.93*** |
| | (10.6) | (9.52) | (2.51) | (1.86) | (1.92) | (1.45) | (1.53) | (1.25) |
| | [-0.58] | [-0.52] | [-1.61] | [-1.27] | [-1.38] | [-1.05] | [-1.65] | [-1.47] |
| Linguistic fractionalization a/c for distance | 1.74 | 0.61 | 2.98 | 1.77 | 1.67 | 0.99 | 1.80 | 1.34 |
| | (8.67) | (7.80) | (1.99) | (1.49) | (1.31) | (0.99) | (1.19) | (0.97) |
| | [0.051] | [0.018] | [0.67] | [0.40] | [0.54] | [0.32] | [0.57] | [0.42] |
| Executive constraints | | -1.46* | | 0.19 | | 0.096 | | 0.075 |
| | | (0.76) | | (0.12) | | (0.085) | | (0.092) |
| | | [-0.24] | | [0.24] | | [0.18] | | [0.13] |
| Log GDP per capita at independence | | 5.16*** | | 0.87*** | | 0.65*** | | 0.61** |
| | | (1.87) | | (0.30) | | (0.21) | | (0.23) |
| | | [0.35] | | [0.45] | | [0.49] | | [0.44] |
| Observations | 40 | 40 | 38 | 38 | 36 | 36 | 39 | 39 |
| R-squared | 0.300 | 0.454 | | 0.239 | | 0.280 | | 0.105 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Panel B: First-Stage for ADOL** | | | | | | | | |
| Share of Partitioned Ethnicities | 0.0031*** | 0.0032*** | 0.0027** | 0.0027** | 0.0024** | 0.0024** | 0.0029** | 0.0030** |
| | (0.0011) | (0.0011) | (0.0011) | (0.0012) | (0.0011) | (0.0012) | (0.0011) | (0.0011) |
| | [0.36] | [0.37] | [0.30] | [0.31] | [0.29] | [0.30] | [0.34] | [0.35] |
| Linguistic fractionalization a/c for distance | 0.68*** | 0.67*** | 0.68*** | 0.67*** | 0.60*** | 0.60*** | 0.65*** | 0.64*** |
| | (0.15) | (0.16) | (0.15) | (0.16) | (0.15) | (0.16) | (0.15) | (0.15) |
| | [0.55] | [0.54] | [0.56] | [0.56] | [0.54] | [0.53] | [0.55] | [0.55] |
| Executive constraints | | 0.0083 | | 0.0025 | | -0.0023 | | 0.0081 |
| | | (0.028) | | (0.029) | | (0.028) | | (0.027) |
| | | [0.038] | | [0.011] | | [-0.012] | | [0.039] |
| Log GDP per capita at independence in 1990 US | | 0.021 | | 0.032 | | 0.019 | | 0.034 |
| | | (0.069) | | (0.068) | | (0.070) | | (0.068) |
| | | [0.040] | | [0.062] | | [0.038] | | [0.065] |
| Observations | 40 | 40 | 38 | 38 | 36 | 36 | 39 | 39 |
| F-Stat | 14.4 | 6.91 | 13.6 | 6.50 | 9.79 | 4.63 | 13.6 | 6.60 |

$*p < .10; **p < .05; ***p < .01$. Robust SE's in parenthesis and standardized coefficients in square brackets.

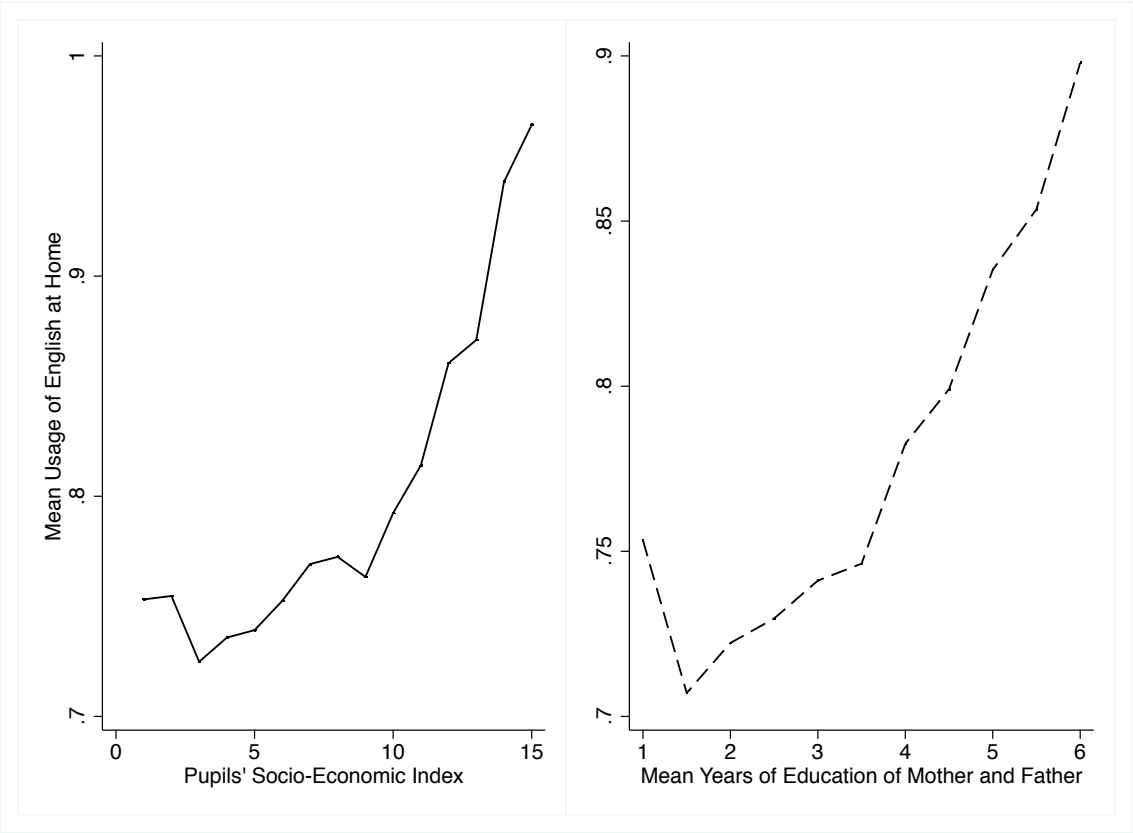


Figure A.1: **Mean of English usage by two family characteristics**

Source: SACMEQ II Dataset.

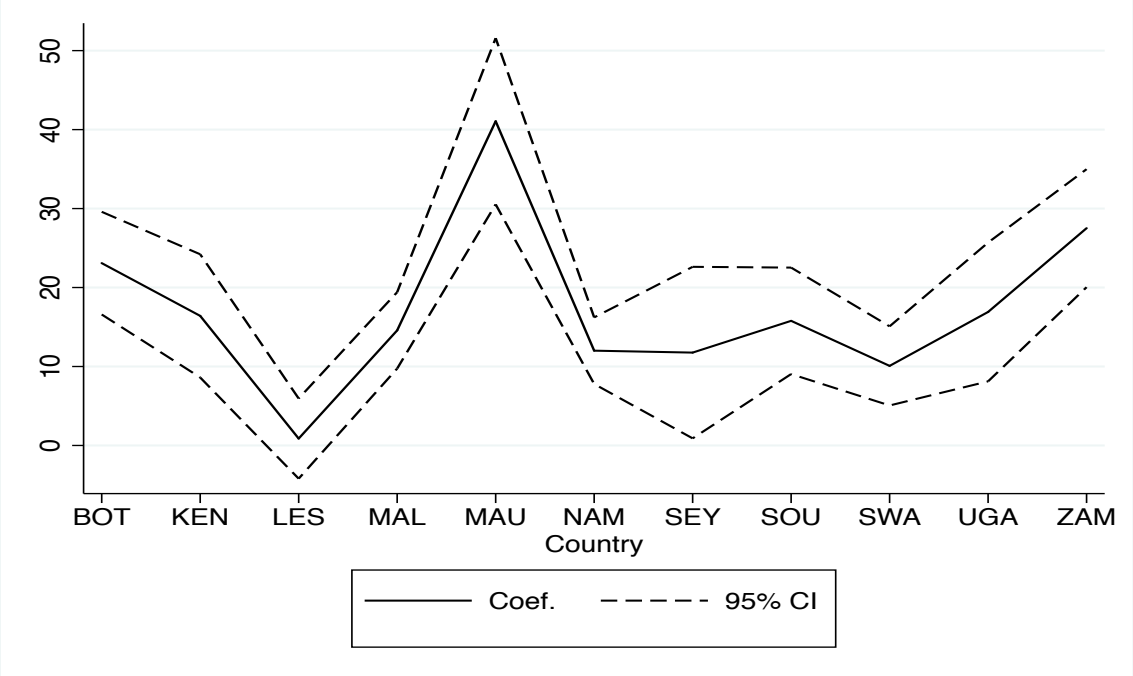


Figure A.2: **Effect of usage of English at home on English score by country**

The y-axis shows the effect on English score standardized with mean 500 and a standard deviation of 100.

Source: SACMEQ II Dataset.

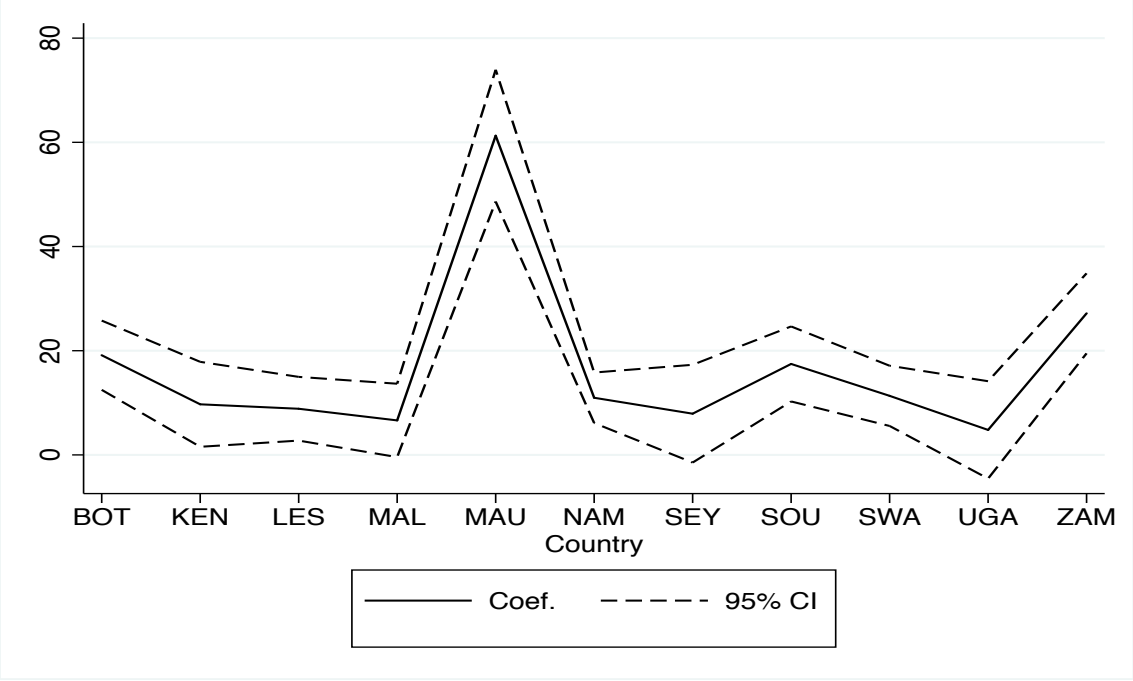


Figure A.3: **Effect of usage of English at home on Math score by country**

The y-axis shows the effect on Math score standardized with mean 500 and a standard deviation of 100.

Source: SACMEQ II Dataset.

# References


Acemoglu, D., S. Johnson, and J. A. Robinson (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review 91*(5), 1369–1401.

Alesina, A., W. Easterly, and J. Matuszeski (2011). Artificial states. *Journal of the European Economic Association 9*(2), 246–277.

Altonji, J., E. Todd, and C. Taber. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of Political Economy 113*(01), 151–184.

Ashraf, Q. and O. Galor (2013). The "Out of Africa" hypothesis, human genetic diversity, and comparative economic development. *The American Economic Review 103*(1), 1–46.

Bolt, J. and J. Van Zanden (2013). The first update of the Maddison project: Re-estimating growth before 1820. *Maddison Project Working Paper 4*.

Desmet, K., I. Ortuño-Ortín, and R. Wacziarg (2012). The political economy of linguistic cleavages. *Journal of development Economics 97*(2), 322–338.

Desmet, K., S. Weber, and I. Ortuño-Ortín (2009). Linguistic diversity and redistribution. *Journal of the European Economic Association 7*(6), 1291–1318.

Englebert, P., S. Tarango, and M. Carter (2002). Dismemberment and suffocation: A contribution to the debate on African boundaries. *Comparative Political Studies 35*(10), 1093–1118.

Esteban, J., L. Mayoral, and D. Ray (2012). Ethnicity and conflict: An empirical study. *The American Economic Review 102*(4), 1310–1342.

Fearon, J. D. (2003). Ethnic and cultural diversity by country. *Journal of Economic Growth 8*(2), 195–222.

Hall, R. E. and C. I. Jones (1999). Why do some countries produce so much more output per worker than others? *The Quarterly Journal of Economics 114*(1), 83–116.

Heston, A., R. Summers, and B. Aten (2012, Nov). Penn world table version 7.1, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Oster, E. (2013). Unobservable selection and coefficient stability: Theory and validation. No. w19054. National Bureau of Economic Research.

PRS Group [Distributor] V1 [Version] (2010). International country risk guide (ICRG) researchers dataset.

Treisman, D. (2007). What have we learned about the causes of corruption from ten years of cross-national empirical research? *Annual Review of Political Science 10*, 211–244.

UNDP (2011). UNDP (1990 through 2010). Human Development Report.

World Bank (2014). World development indicators.