Colonial Origins and Comparative Development: Institutions Matter

Chakraborty, Adrij

University of Edinburgh, School of Economics

March 2017

Colonial Origins and Comparative Development: Institutions Matter

Adrij Chakraborty
(University of Edinburgh)

Colonial Origins and Comparative Development: Institutions Matter

ABSTRACT

A large literature is present on how colonial origins through the channels of Institutions and human capital explain the income variation between countries. In most cases, it is argued that colonial European settlement outside of Europe has shaped the type of institutions present today that explain the gap between income levels of countries. Others criticize this claim on the basis that firstly, geography plays a prominent role in income variation and secondly, the studies in favor of the primacy of the role of institutions suffer from measurement error and bias. In this study, we analyze the role colonial origins and geography play in comparative income levels of countries today. In this study, we present a theoretical analysis for what determines growth in the long run and how colonial origins come into play and then make an empirical study using OLS and IV techniques, while utilizing recent data and different proxies, into the claims that colonial origins and/or geography play the prominent role. We also investigate whether the institutions primacy studies suffer from measurement error. Our findings point out that while geography may play a small role, majority of the income variation is associated with institutions and colonial origins.

JEL Codes: P51, P16, I12, N10, O57, O1, O4, O11

INTRODUCTION

In the last two decades, numerous empirical studies have taken place to investigate why some countries are poor and others are rich. One of the most prominent of these studies, Acemoglu et al[1] (from now on referred to as AJR) concentrate on the role European colonial origins play in determining income variation across countries. These studies argue that when Europeans colonized different areas of the world, they set up two types of institutions - extractive or inclusive- as well as taking with them skills, technology and culture that persisted over time. Thus, these European extorts across the globe could be the main explanation as to why we have differences in income between countries. On institutions, it is argued that where colonialists set up inclusive institutions, ones where the rights to life, liberty and property were guaranteed, growth followed and this persisted over the years creating a gap in income levels. However, there are papers that criticize this claim and find geography to be as important in determining growth and income levels. Furthermore, because the studies use Instrumental variables techniques to study the role of European colonial origins on income levels today, they are criticized on the strength and validity of their instruments. In this paper, we look into the role colonial origins play in determining income levels today. In the first section of the paper, we review the prominent literature on the topic. In the second section, we present a theoretical analysis for what determines growth in the long run and how colonial origins come into play. In the third section we take up an empirical enquiry using OLS and IV techniques, while utilizing recent data and different proxies, into the claims and criticism of whether colonial origins plays the prominent role. Lastly, we conclude our paper with a brief answer to whether the findings of AJR stands the critiques of Albouy (2004, 2006, 2012), the critique of Sachs (2003) and the test of time – given different proxies and up to date data.

[1] Acemoglu, Johnson, Robinson (2001)

REVIEW OF LITERATURE

In this section, we provide a brief review of the current empirical literature that supports the view that colonial origin through different channels play the determining role in growth. We first look at the first and most prominent channel, institutions, through which colonial origin affects growth differences. We then present the evidence against the primacy of institutions, discuss other channels such as human capital as well as some of the shortcomings and criticisms of the established papers on the topic. To this date, the most influential studies on institutions and growth address the reverse causality and the endogeneity issues by employing instrumental variable regressions. Based on the fact that institutions persist over long periods of time, the settlement decisions of colonizers and imperial powers between the 16th and 19th century can be used as econometric leverage to help unravel the real causal relationship between institutions and growth.

One of the first papers in this field is Hall and Jones (1999) who instrument the quality of current institutions, which they call "social infrastructure", with geographical and linguistic characteristics of the colonies in the 16th century: distance from the equator and the extent to which English and other European languages are spoken as a mother tongue. The authors argue that Western Europe was one of the first regions in the world to set up the social infrastructure conducive to high output growth. As such, it can be argued that regions where a larger proportion of the population who spoke these languages is correlated with Western influence, and thus early institutions. Their empirical results show that countries with the best institutions enjoy 25 to 38 times higher income per worker than countries with the worst institutions.

Acemoglu et al. (2001) criticize Hall and Jones' paper on the basis that their instruments do not represent the true causal effect of institutions on growth, as they are directly correlated with economic performance. They propose a new instrument, European settler mortality, which they argue was a crucial determinant of early colonial settlement experience which in turn led to differences in current institutions. In places where Europeans faced high mortality rates, they preferred to set up extractive institutions (Congo, Burundi, Mexico); whereas in places where they faced low mortality rates, they set up inclusive institutions which provided checks and balances against government expropriation, independent judiciary, and established the rule of law, examples being the United States, Canada and Australia. Their two-stage regression estimates show that after instrumenting institutions by the mortality

figures, institutions account for up to three quarters of the variation in output per worker across colonized countries. Their work has subsequently been revised and followed up by other authors, including Easterly and Levine (2003), Dollar and Kraay (2004), Rodrik et al (2004), and Fernandes and Kraay (2005), who all confirm that there is a significant relationship between institutions and growth thus arguing for the substantial role colonial origins play on development levels today.

In their subsequent paper, AJR (2002) argue that regions with higher population density in 1500s have some of the worst institutions today, as they were more likely to be subject to forced labour and slavery, and were ruled under extractive institutions. After instrumenting current institutions (which they measure by urbanization) with population densities, they find that a 10% increase in population density in 1500s resulted in 4% lower income per capita in 1995. Interestingly, they find that when examining countries that were never colonized by the Europeans, the relationship between population density and income per capita is positive, suggesting the relevance of population density as an instrument for institutions.

Other papers in the growth literature argue that geography matters as much as colonial origins, specifically institutions, in long-run economic development. Amongst the more notable studies in this field, include Sachs (2001, 2003) and Diamond (1997) who argue that climatic conditions have a direct effect on growth through various factors such as disease ecology, climate zone, and distance from the coast. While most of the studies indicate that geography becomes insignificant after controlling for institutions, Sachs (2003) reports that malaria transmission, which is strongly affected by malaria ecology, still has a direct effect on economic growth even after controlling for institutional quality. They find that after instrumenting malaria risk with malaria ecology and the share of country's population who live in temperate ecozones, they find that both institutions and geography are significant at the 5% level in explaining the variation in income levels, implying that geography also matters.

The most prominent criticisms against the colonial origin argument of AJR specifically institutions come from Albouy (2004, 2006) who claims that the settler mortality rates used by AJR suffer from measurement error. Albouy criticizes the data due to the fact that only 28 of the 64 countries in their original sample have actual data; whereas the other 36 countries are assigned rates based on conjectures made by the authors on the similarity of health conditions such as the disease environment. Therefore, Albouy drops these observations and

finds that the effect of settler mortality on expropriation risk (proxy for institutions) is much smaller, making it a weak instrument to study the causal effect of institutions on growth. In reply, AJR (2005, 2006, 2012) claim that their estimates of the data are supported by historical records and are therefore reliable, whereas Albouy's data suffer from selection biases based on irrational conjectures. Furthermore, Glaeser et al. (2004) argue that instrumenting institutions by European settler mortality violates the exogeneity assumption on the basis that the colonizers brought with them skills and human capital that persisted over time. They claim that it is the persistence of these skills that is causing the variation in output per worker today, and not the institutions. Moreover, they argue that the measure of institutions used by Acemoglu et al. (2005) are a measure of policy outcomes, and not the measure of "constraints" on government officials, as defined by North (1981). They run a simple OLS regression of GDP per capita on executive constraints, as well as controlling for additional variables, to show that human capital has been the main determinant of growth since the colonial period, and not institutions.

Moving on, Easterly and Levine (2015) create two new datasets, the main one of which is Euro Share. This is a dataset that captures the percentage of the total population that were Europeans during the colonial time. In line with the findings of AJR, they find that colonial origin is the prime determinant of income differences across states. There finding is a further argument in favour of the reversal of fortunes claim made by AJR. Additionally, they find that it is not only institutions that determine different income levels but also human capital and technology, so the effect of colonial origin is more significant than just through the channel of institutions. This dataset effectively merges Glaeser et al. (2004) finding with that of AJR and shows that regardless of the channel, colonial origins play arguably the most important role in GDP per capita differences today.

THEORY

Hervé Boulhol highlights[2] the importance of institutions to realize economic growth. He argues in his paper that technology in the economy is a strong component to attain productivity, which is primal to our analysis of growth and development. Boulhol's study mainly focusses on the three major effects of institutions on economic growth: efficiency in the use of technology, technology diffusion and eventual long-term TFP growth. His study

[2] Boulhol H, *Technology differences, institutions and economic growth: a conditional conditional convergence*, CEPII

wanders around the interactions between institutions and technology differences to explain cross-country growth pattern[3]. The basis of the argument presented is how Boulhol's take on technology and its approximation as a form of constructive institution impacts the cross-country growth pattern, and is extended to deduce how it helps in the economic growth of a nation. The quality of institution matters profoundly, as the degree of constructiveness decides how well the economies grow. Boulhol comments on institutions being a linear influence on growth, as well as conditional convergence being conditional on the similarity of institutions in terms of quality and efficiency of the nations concerned. AJR argue that they are consequential social decisions[4].

AJR argue that there is an intertwined relationship between the political and the economic institutions because the presence of one affects the other, which in turn influences the behaviour of the country[5]. Settlements affected early institutions; and early institutions persisted and formed the basis of current institutions. Settler mortality is emphasized owing to its ability to determine the types of institution (which affects growth – the hypothesis here) in the long term. To make matters simple, Boulhol talks about convergence in the long run, which corroborates with Acemoglu's arguments on long run stability in the economy being the prime outcome of deterministic institutions. An instrument for AJR to judge the foundation of the type of institution (extractive or inclusive) is settler mortality rate. The type of institution not only decides on the growth level, but also the quality of economic growth realised in the country concerned. With the help of the analogous model of Boulhol and the technological (hence, institutional) influences on his analysis, an explanation on how the quality and the type of institution matter for the country can be made.

The ideas are juxtaposed to explain how political institutions and the distribution of resources in turn determine the distribution of politico-economic power in society. Noted economists in the past have propounded what AJR argue is a rather elaborate extension of few of the more impactful ideas on growth. Noted theorists like Harrod-Domar, Lewis and more recently

[3] In the paper authored by Acemoglu, Johnson and Robinson, it is argued that the differences in the economic institutions are the fundamental cause of differences in economic development. They theoretically debate on the concept, as well as empirically demonstrate how economic institutions determine the incentives of and the constraints on economic actors, and shape economic outcomes.

[4] Because different groups and individuals from varied contexts typically benefit from different economic institutions, there is generally a conflict over these social choices having manifold effects, ultimately resolved in favour of groups with greater political power - AJR

[5] . Not only limiting the scope of the paper to technological and social institutions, AJR feature in the socioeconomic and political institutions that are at play by taking instances of two "quasi-natural experiments" in history, the division of Korea into two parts with very different economic institutions and the colonization of much of the world by European powers starting in the fifteenth century.

Solow, have argued that without the presence of institutions and technology economies would not have been able to grow at all. Economic institutions not only determine the aggregate economic growth potential of the economy, but also influence an array of economic outcomes, including the distribution of resources in the future (i.e., the distribution of wealth, of physical capital or human capital). The colonial past of a country and its setting determined the type of colonial presence and therefore the type of institutions the setting helped establish.

$$\begin{cases} Political\ institutions\ (t) \\ Distribution\ of\ resources\ (t) \end{cases} \rightarrow \begin{cases} de\ jure\ political\ power\ (t) \\ de\ facto\ political\ power\ (t) \end{cases}$$

$$\rightarrow \begin{cases} economic\ institutions\ (t) \rightarrow \begin{cases} economic\ performances\ (t) \\ distribution\ of\ resources\ (t+1) \end{cases} \\ political\ institutions\ (t+1) \end{cases}$$

Boulhol's argument[6]

Technological diffusion is the process by which innovations (new products, processes or management methods) spread within and across economies. According to Boulhol, technological progress is not independent from other explanatory variables, and that it directly plays a role in determining the shape of the country's future. He states that changes in output per capita isn't just explained by differences of capital per unit of labour. Boulhol sets up his argument computing the per capita variables for a Cobb-Douglas equation. This helps him identify the contribution of not only capital in the income levels, but also mark the role of exogenous variable used for (institutionalised) technology. He recognizes the productivity differences, and that both the initial productivity level and the initial output per capita are closely linked. It can be safely assumed that the initial level referred to here is the initial point for which AJR consider the settler mortality rate. This rate will go on to determine which kind of institutions crop up and how they are developed over the years. The parameter 'A' is often indistinctly designated as either the productivity or the technology level in mainstream growth models. In this analysis, Boulhol attempts to explain the technology parameter by breaking it into two explanatory parts, and calls the complement of technology in productivity "efficiency"[7].

[6] Appendix (Theory) for exposition and mathematics

[7] According to the argument setup, a pure technology level and the degree of efficiency in using this technology can explain technology, which in itself is a manifestation of institutions.

Boulhol makes no remarks about whether the institution is extractive or inclusive (depending on the instrument used by AJR– settler mortality rate). His approach is relative: he considers a benchmark country purporting to be affluent and economically developed, which serves as a comparison variable against which other countries in question are weighed. The proportions of the two components are what Boulhol believes would hint at the type of institutions for the setup in the country in question, and would give him the ideal influence of the institutions on economic growth. Boulhol considers institutional quality as an influence on the technological efficiency and possibly the technology diffusion. Here institutions can enter into growth equations through different channels like the level as well as the progress of technological efficiency, and possibly the speed of technology diffusion. Boulhol performs majority of the calculations in the proportion mode, as this form of the equation explains how exactly the institution factor figures out in the model. Taking similar origin points, the institutions (manifested in such formats) are the deterministic variables when countries decide to choose their path to economic growth.

Institutions here have an impact through technological efficiency at the inception (origin), through the long run TFP-growth deficit to the benchmark, and potentially through the long run convergence rate of technologies to the frontier which Boulhol claims to be constant across countries or institutionally related. The expression for productivity into Boulhol's argument, improves it into a growth model. This is quite similar to the macroeconomic augmented Solow Model enabled to take into account the heterogeneity of technologies and the contribution of institutions. Boulhol introduces in his argument a growth equation which is contingent mainly on the population of the economy and the capital depreciation rate. He suggests that growth of a country is driven by the conditional convergence rate and the contribution of technology diffusion.

The Boulhol Model has its similarities with that of Solow, where he makes similar assumptions on the identicalness of initial total productivity across nations and the absence of long run TFP-growth differences[8]. Boulhol considers the similar levels of technology in each country because it's difficult to scrutinize the differences that are there. The initial level of productivity (here, colonial origins) and the speed of convergence (institutionalized growth rate) are the factors, which help us, screen out the contribution of institutions in context.

[8] The model processes growth as an aggregation of forces pertaining to the likes of adjusted convergence, technological catch-up, non-convergence due to differences in long term investment rates and an additional divergence owing to long run TFP-growth disparities

Boulhol argues that income ratios can be explained by the initial productivity ratios, while the capital differences can be justified by the latter's complement. Estimates of stocks of physical and human capital to infer productivity levels can be made and it can be assessed that productivity differences account for half or more of level differences. Since productivity is born out of institutions, and enhanced by technology (resources and capital included), it can be safely assumed that institutions lead to growth and eventual economic convergence in a country[9].

Therefore, theory suggests that colonial origins have affected the countries majorly through the institutions set up. Political and economic institutions are the backbone determining the pathway for countries to grow. The type of growth rate, however, is reliant on the nature of institutions that have been positioned. Partially founded on notable work on growth economics, Boulhol, Acemoglu et al[10]. paint a similar picture arguing about how colonial nations affected the colonies in terms of growth and development.

DATA

The dependent variable - To study the effects of colonial origin on development and growth across countries we use log of GDP per capita in PPP for the year 2015. The advantages of using GDP per capita, as the independent variable is the fact that it has exponential growth thus capturing non-linear changes and there is wide availability of this measure for different countries.

The independent variables – For *Institutions* we use six proxies devised by Kraay and Kaufman (2015) which are indicators on a scale of -2.5 to 2.5 for the following qualities; rule of law, government effectiveness, political stability and nonviolence, control on corruption and regulatory quality. We then combine and average these proxies to create the *KK-index* as a proxy for institutions.

[9] The extreme simplification in the model runs a risk of misappropriation, and has the merit of highlighting the potential correlation between initial productivity and initial output per worker, as well as the comparison of how efficient an institution might be and its reflection on the corresponding level of economic growth pattern, shows how the institutions impact long term behaviour of the economy. (Klenow and Rodriguez-Clare (1997))

[10] AJR have empirically proved how the nature of institutions has been dependent on the instruments that they've considered for regression analysis. Taking settler mortality rate as an indicator, Acemoglu et al. debate that extractive institutions were considered for countries with higher numbers and inclusive for the countries with lower.

Table 1: Descriptive statistics of the major variables used in the regression analyses

	Observations	Mean	Standard deviation	Min	Max
Current Euro share	107	0.139	.251	0	0.91
Albouy's log of settler mortality	33	4.537	1.477	2.145	7.986
Log of GDP per capita	170	9.252	1.141	6.664	11.55
KK index	180	-0.019	0.882	-2.283	1.852
Euro share	118	0.074	0.168	0	0.905
Dummy for British Legal Origin	167	0.353	.0479	0	1
Executive Constraint	149	5.369	1.839	1	7
Voice and Accountability	180	-0.006	0.974	-2.112	1.656
Political Stability and non-violence	180	-0.046	0.927	-2.923	1.394
Government effectiveness	180	-0.005	0.955	-2.267	2.216
Rule of law	180	-0.032	0.965	-2.421	1.973
Control of Corruption	180	-0.022	0.976	-1.704	2.415
Regulator quality	180	-0.004	0.944	-2.367	1.920
Settler Mortality	77	4.652	1.181	2.145	7.986
Population density in 1500	151	1.015	1.612	-3.830	5.642
Latitude	180	0.290	0.190	0.002	0.713
Malaria Ecology	106	4.983	7.268	0	31.5
Indigenous mortality	117	0.307	0.463	0	1
Biogeography	78	0.021	1.334	-1.018	3.791
Education	183	72.71	33.03	0	133.5
Ethnicity	110	0.380	0.315	0	1
Independence	86	0.326	0.345	0	1
London	112	5383	2284	1381	1136
Non-European sample	118	1	0	1	1
Distance to coast	146	391.7	450.8	7.79	2374
Precious metals	119	0.302	0.461	0	1

Rule of law - is a compilation of many measures that determine rule of law some of which are property rights and their safety from expropriation, policing quality, the quality and efficiency of the judicial courts and contract enforcement, etc. (Kraay et al, 2010)

Government effectiveness – is a measure of the qualities of the government and its effectiveness on policy implementation such as government's commitments, measure of the

extent of political pressure, the speed with which policies are implemented and so on. (Kraay et al, 2010)

Political Stability and nonviolence – is a measure of the lack of politically spurred instability and violence in each state which is an amalgamation of measures such as the frequency of terrorist activities in each state, ethnic tensions that inspire and result in violence, the rate with which governments change; be it coalition break ups or military coups and so on. (Kraay et al, 2010)

Control on corruption – is an indicator of the control on the misuse of state authority by the people in power for personal gain, control and extend of bribery in each state, control and extend of nepotism, etc.

Regulatory quality - is an indicator that captures the ability of the government to formulate and implement sound policies. Some variables used to create this indicator are as follows: unfair competitive practices, price controls, discriminatory tariffs, burden of government regulations, investment freedom, ease of starting a business, etc.

Additionally, from Polity IV indicators we use *executive constraint* as a proxy for institutions. This is an indicator, scaled from 1 to 7 with 7 being maximum constraint, of the expropriation control put on the government branch of the state, a cap on their ability to abuse power. We take an average from 2006 to 2016, for a period of 11 years.

For *Geography* – we use three different proxies first of which is latitude as is used by the majority of the papers in the study of the role of geography on growth differences including AJR (2001) and Easterly Levine (2012). This proxy consists of the absolute value of latitude of each country divided by 90 to get a scale of 0 to 1. The second proxy we use is Malaria Ecology; this is used as a proxy for geography as it entails the disease environment. The last proxy used is distance from the coast.

Moving on, we use additional variables as instruments to respond to the simultaneous causality problem. The first instrument that we have at our disposal is the settler morality devised by Curtin (1998), and improved and used by Acemoglu et al. (2001; 2005)[11].

[11] AJR use *settler mortality* as an instrument based on the reasoning that since European colonialists set up extractive institutions in places with lower mortality rates, and inclusive institutions in places with low rates of mortality, it could be used a valid instrument to be used, if it only creates variation in income through the proxies of institutions or human capital. More on this will be discussed in the second section of our results.

Another instrument we use is the of the *local population density in the 1500s*. This is used by the same intuition as above. European colonialists set up extractive institutions in areas where there were more of the locals by condemning them into slavery stealing their precious metals and so on. The validity of this instrument also depends on the intuition that since population density in the 1500s cannot be affected by income today, if used as an instrument for institutions, it could solve the simultaneous causality problem; this is subject to it only affecting income through the proxies for institutions or human capital.

Additionally, as in the Easterly-Levine (2015) paper, in this study we use net gross enrolment in high schools by both genders worldwide as a proxy for human capital and label it *education15*. As EL take an average of the early 2000s for 10 years, we update the data and take a 10-year average from 2006 to 2015 that is fit for the time period we study. It should be noted that in order to see whether the main papers stand the test of time, in this study we use updated data for the period of 2005 to 2015 for our exogenous variables where it permissible to do so.

ECONOMETRIC METHODS AND ISSUES

To find the effect colonial origins in the form of institutions (or human capital and technological effects) has on incomes today we must use IV in order to deal with the problems of simultaneous causality. The problems of endogeneity and reverse causality make it difficult to estimate the true causal relationship between colonial origins in the form of institutions and growth. While there has been enough empirical evidence to suggest that causality here runs from institutions to growth (Acemoglu et al, 2001; Rodrik et al, 2002), it is still difficult to prove that improvements in institutional quality do not simply occur because wealthier countries are simply able to afford it. Thus, running a simple OLS regression that estimates the effects of institutions and geography on economic growth will give unreliable coefficients, as our estimates will be biased both in upward and downward directions.

The best way to overcome this reverse causality issue is to conduct a randomized controlled experiment whereby exogenous variations in institutions and growth can be generated. However, in the absence of such experiments, one must employ instrumental variables regressions, whereby the instrumental variable has a significant first-stage effect in explaining the variation of the endogenous variable (relevance condition) and should also be uncorrelated with the error term (validity condition). As mentioned previously, the

instrumental variable by Acemoglu et al (2001) use the log of European settler mortality and local population density in 1500 that satisfy both the above criteria.

However, the instruments might not be completely exogenous if settler mortality is correlated with present day variables that are affecting incomes such as the current disease environment. If the instruments are in fact correlated with unobservable factors that are affecting output per worker, then the coefficient estimates may be assigning the effect of diseases on income to institutions. According to AJR (2001), this is very unlikely to be the case as the vast majority of European deaths in the colonies were due to malaria and yellow fever. However, these diseases had very limited effect on the indigenous population who already developed the immunities to withstand these diseases. Their argument is supported by the mortality rates developed by Curtin (1968), who reports that among the local troops serving the British army in Bengal and Madras, only 11 to 13 out of 1000 died from local diseases, whereas the annual mortality rate of British troops was between 70 and 170 in 1000 in these regions. Therefore, according to AJR (2001), these diseases are unlikely to have contributed to the vast differences in income per capita among the colonized countries.

Similarly, for population density in 1500s (write paragraph)

Econometric Specifications

To see the relationship between institutions and income per capita, both OLS and 2SLS regressions are considered as they provide a basis of comparison with other studies. The simple OLS expression takes the form:

$$\log Y_i = \mu_i + \beta_1 GEO_i + \beta_2 INS_i + \varepsilon_i \qquad (1)$$

where i represents countries. Yi is the log of income per capita for country i, μ_i is a constant, GEOi is the proxy for geography, INSi is the proxy for institutions, and ε_i is a set of unobservable that influence income per capita. Throughout the paper, we are interested in estimating β_1 and β_2, the effects of institutions and geography on income per capita.

There are a number of reasons for why equation (1) should not be interpreted as causal. Firstly, there might be a problem of reverse causality as discussed previously. Perhaps it is higher levels of income that bring about better security, political stability and lower violence. This would result in an upward bias in our institutions coefficients. Secondly, analysists may have had a natural bias in seeing better quality institutions in richer places, which means that the data collected for institutions suffer from measurement error that biases the estimates

towards zero. Finally, there may be many other omitted determinants that affect income levels and that are also correlated with our proxies for institutions. All these issue may result in violation of the most important Gauss Markov assumption which requires that the independent variables are orthogonal to the error term, i.e. $E(u|x) = 0$ does not hold.

Thus, in order to identify the exogenous variation in institutions an instrumental variable approach must be used. In the first stage, INSi are regressed on all of the exogenous variables:

$$INS'i = \alpha + \delta 1SMi + \delta 2GEO[12]i + \eta i \qquad (2)$$

$$INS'i = \alpha + \delta 1PDi + \delta 2GEOi + \eta i \qquad (2)$$

where SMi refers to European settler mortality and GEOi refers to the geography proxy. The reason why geography is included in our first-stage regressions is because it is an exogenous determinant of income per capita, as it cannot be influenced by income or institutions. The predicted values of INS'i form the first-stage are then used in the second-stage to obtain unbiased estimates of geography and institutions on income per capita. The equation for the second-stage least squares takes the following form:

$$\log Yi = \mu i + \tau 1GEOi + \tau 2INS'i + \rho i \qquad (3)$$

where the variables are defined exactly as before. This model represents the most natural framework for estimating the effects of institutions and geography on economic growth.

RESULTS AND DISCUSSIONS: ACEMOGLU VIS-À-VIS:

SACHS

Before presenting the benchmark results, it might be useful to look at a simple bivariate relationship between income per capita and each of our proxies for institutions. As can be seen in figure 1, all of the eight scatterplots show a clear positive relationship between institutions and income per capita. This positive relationship still holds when we include latitude as a proxy for geography, as reported in Table 2. As expected, the coefficients on institutions are positive and statistically significant, which means that countries with better institutions enjoy higher levels of output. Similarly, the coefficient on latitude states that countries further from the equator are more successful in terms of GDP per capita than

[12] GEO can also be used for a matrix of other instruments X, the common intuition applies

countries close to the equator. The high R2 in all of our columns indicates that most of the variation in income levels can be explain by our independent variables.

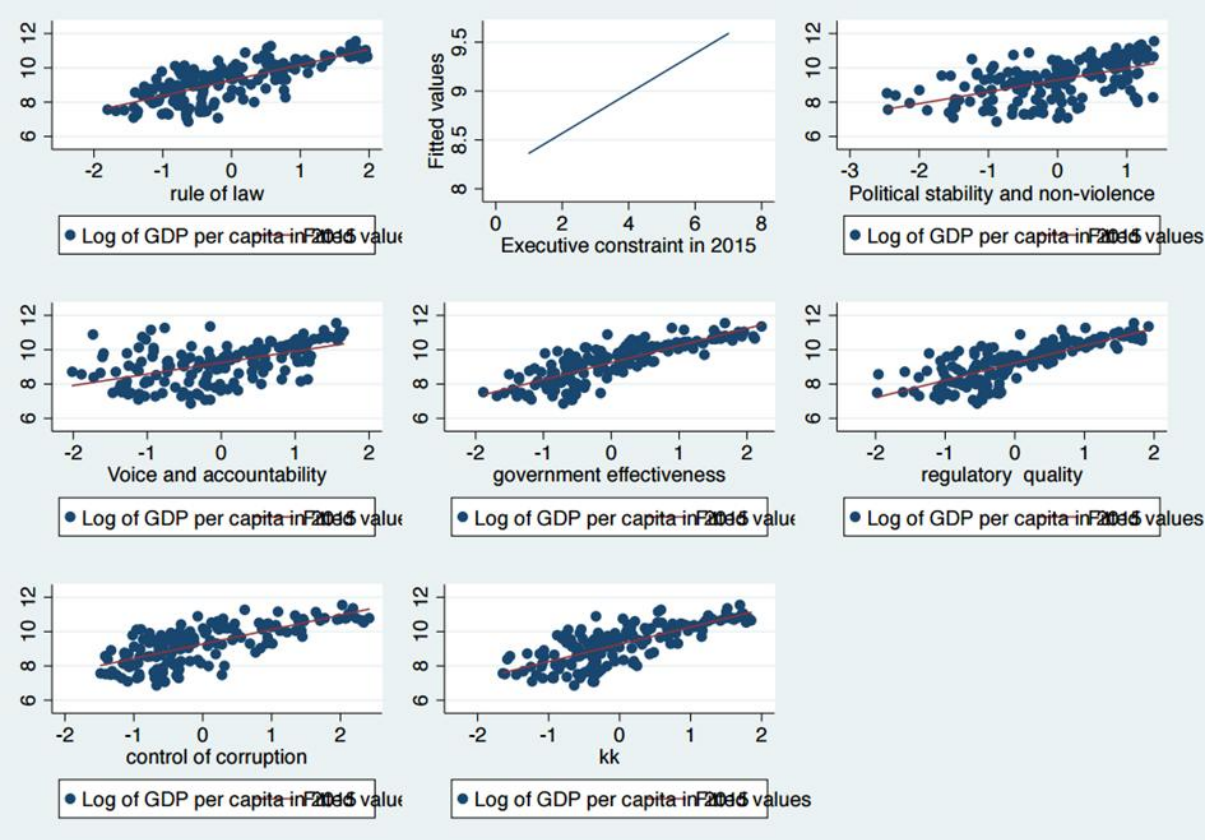

Figure 1

Given the positive and significant coefficients on all of the independent variables, it is clear that both institutions and geography have the potential to explain the cross-country variations in income per capita. However, the coefficients on institutions and geography cannot be interpreted as causal as discussed previously. Thus, we employ a two-stage-least-squares estimation procedure in Tables 3 and 4, with table 2 reporting the first-stage regressions where we instrument institutions with log of European settler mortality and log population density in 1500; and table 4 reports the second-stage regressions.

It can be seen from the first-stage regressions that there is a strong first-stage relationship between current institutions and European settler mortality rates, given the very high t-statistics in all of our settler mortality coefficients. This is consistent with the findings of Acemoglu et al (2001) that settler mortality had an important causal effect on early institutions that continued to persist to this date. Similarly, our results are consistent with AJR (2002) that population density in 1500s are negatively correlated with institutions today. The F-statistic test of being equal or greater than 10 confirms that the majority of the proxies used

for institutions pass the relevance condition, and thus are valid instruments for institutions. The R2 in the majority of

Table 2: Core specifications, ordinary least squares estimates.

	Dependent variable is log GDP per capita in 2015 (current PPP)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Rule of law	0.76***						
	(0.06)						
Executive constraint		0.11**					
		(0.05)					
Political stability and violence			0.55***				
			(0.07)				
Voice and Accountability				0.49***			
				(0.07)			
Government effectiveness					0.90***		
					(0.05)		
Regulatory Quality						0.89***	
						(0.07)	
Control of corruption							0.70***
							(0.06)
Geography (Latitude)	1.38***	3.22***	2.51***	2.30***	0.89***	0.99***	1.57***
	(0.35)	(0.40)	(0.34)	(0.36)	(0.31)	(0.34)	(0.34)
Constant	8.88***	7.71***	8.52***	8.59***	8.96***	8.92***	8.82***
	(0.13)	(0.33)	(0.14)	(0.14)	(0.11)	(0.12)	(0.13)
Observations	168	142	168	168	168	168	168
R-square	0.61	0.36	0.48	0.44	0.70	0.66	0.58

Notes: Robust standard errors are in the parentheses. All regressions are cross-sectional OLS with one observation per country. Executive constraint is on a scale of 1-7 with 7 being more control. The other proxies for institutions are on a scale of -2.5 to 2.5 where 2.5 is more control and better institutions. *Significant at 10 percent; **Significant at 5 percent ***Significant at 1 percent.

the columns also show that a large variation in institutions can be explained by settler mortality and population density during colonization.

Consistent with the findings of Glaeser et al (2004), settler mortality poorly correlates with constitutional measures of checks and balances – as measured by executive constraints taken from Polity IV data. We believe that there are both conceptual and methodologic problems with this data. First, the executive constraint scores in Polity IV is taking the multidimensional concept of institutions and reducing it to a unidimensional scale, which means that there is some serious loss of information. The Kaufman and Kraay proxies, however, reflect the views on governance of survey respondents and sector exports worldwide, and takes a continuous measure, which means there is less room for error. Second, institutions is not just about the degree of constraint the executive faces, but rather it is a broad measure of things, such as individual rights and personal freedoms. Therefore, we do not believe that this measure is an accurate representative of institutional quality, but still include it to capture just one more characteristic of institutions

Table 4 reports the second-stage least squares estimates of regressing the log of GDP per capita on institutions and geography. It can be seen that once institutions are instrumented, the coefficient on geography goes from being significant as shown in Table 2, to becoming insignificant. There are still some significant results in some columns when we instrument institutions with log of population density in 1500. One possible explanation by AJR (2002) is that the data for population density in 1500 may suffer from serious measurement error issues because the event occurred 500 years ago. However, the insignificance of the coefficients on geography in Table 4 does not mean that geography has no influence on current incomes. Rather, it indirectly influences income per capita through the channel of institutions, as can be seen in Table 3 where the geography coefficients are highly significant.

The results in Table 4 are in line with those of Acemoglu et al. (2001) and is against the papers by Sachs (2001; 2003) and McArthur and Sachs (2001) who argue that geography plays a large role in determining growth along with institutions. Table 5 reports the first-stage and second-stage results where we show that our estimation remains robust to the inclusion of two

Table 3: First-stage regressions for institutions

Dependent Variable	Rule of law		Executive Constraint		Political stability and violence		Voice and Accountability	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log of European settler mortality	-0.25***		-0.35*		-0.21***		-0.23***	
	(0.08)		(0.18)		(0.09)		(0.09)	
Population Density in 1500s		-0.14***		-0.42***		-0.22***		-0.25***
		(0.05)		(0.11)		(0.05)		(0.05)
Latitude	1.56**	2.48***	0.80	1.60	0.09	0.80	0.60	1.33**
	(0.68)	(0.61)	(1.53)	(1.27)	(0.78)	(0.64)	(0.77)	(0.61)
Constant	0.50	-0.75***	6.67***	5.21***	0.54	-0.41**	0.74	-0.29*
	(0.45)	(0.15)	(1.02)	(0.33)	(0.52)	(0.17)	(0.51)	(0.16)
Observations	75	72	69	66	75	72	75	72
R-square	0.29	0.27	0.09	0.22	0.09	0.21	0.15	0.30
F-statistic	14.95	13.06	3.31	8.77	3.67	9.26	6.28	14.58
Dependent Variable	Government effectiveness		Regulatory Quality		Control of Corruption		KK index	
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Log of European settler mortality	-0.31***		-0.23***		-0.26***		-0.25***	
	(0.08)		(0.08)		(0.08)		(0.07)	
Population Density in 1500s		-0.15***		-0.16***		-0.20***		-0.19***
		(0.05)		(0.05)		(0.05)		(0.05)
Latitude	1.23*	2.40***	0.98	1.79***	1.54**	2.48***	1.00*	1.88***
	(0.66)	(0.61)	(0.67)	(0.59)	(0.69)	(0.59)	(0.64)	(0.55)
Constant	0.96***	-0.61***	0.66	-0.44***	0.57	-0.68***	0.66	-0.53***
	(0.44)	(0.16)	(0.45)	(0.15)	(0.46)	(0.15)	(0.43)	(0.14)
Observations	75	72	75	72	75	72	75	72
R-square	0.34	0.28	0.22	0.23	0.30	0.34	0.26	0.30
F-statistic	18.6	13.12	9.89	10.51	15.12	17.86	12.41	14.91

Notes: Standard Errors are in the parentheses. All regressions are cross-sectional OLS with one observation per country. Executive constraint is on a scale of 1-7 with 7 being more control. The other proxies for institutions are on a scale of -2.5 to 2.5 where 2.5 is more control and better institutions. *Significant at 10 percent; **Significant at 5 percent ***Significant at 1 percent.

additional proxies for geography, malaria ecology and distance from the coast. Malaria ecology measures the extent to which mosquito abundance, duration of the transmission season and survival rates of mosquitos affect the stability of malaria transmission in a region, where higher value indicates that there is a greater chance for malaria transmission. As can be seen from Table 5, the coefficient on malaria ecology is negative and statistically significant at the 1% level, meaning that countries with higher malaria ecology have lower income per capita levels. The coefficient on malaria ecology, however, is very small compared to the kk index or rule of law, indicating that although malaria ecology does play a statistically significant role, it does not play a major role in explaining economic performance.

Furthermore, Mellinger, Sachs and Gallup (2000) point out those coastal regions enjoy higher incomes per capita than countries with landlocked populations. They argue that this is because coastal regions have access to a wider scope of the market than interior regions because of their ability to engage in sea-based trade, and because they enjoy much lower transport costs. Therefore, we use distance from the coast as our second measure to see whether countries

Table 4: Second-stage regressions with geography control

	Dependent variable is log GDP per capita in 2015 (current PPP)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Second-stage: Dependent variable is log GDP per capita								
Rule of law	1.85***	1.46***						
	(0.48)	(0.45)						
Executive constraint			1.30**	0.56***				
			(0.60)	(0.18)				
Political stability and violence					2.25***	0.95***		
					(0.85)	(0.28)		
Voice and Accountability							2.00***	0.84***
							(0.63)	(0.23)
Latitude	-1.49	-0.45	0.33	1.97**	1.20	2.42***	0.20	2.07***
	(1.55)	(1.40)	(2.08)	(0.93)	(1.56)	(0.77)	(1.50)	(0.74)
Constant	9.83***	9.50***	2.01	5.46***	9.54***	8.80***	9.30***	8.64***
	(0.48)	(0.44)	(2.88)	(0.87)	(0.60)	(0.25)	(0.44)	(0.20)
Observations	75	72	69	66	75	72	75	72
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Government effectiveness	1.48***	1.35***						
	(0.27)	(0.34)						
Regulatory			2.03***	1.30***				
			(0.54)	(0.36)				
Control of corruption					1.79***	1.04***		
					(0.47)	(0.29)		
KK index							1.87***	1.11***
							(0.44)	(0.28)
Latitude	-0.41	-0.07	-0.59	0.84	-1.34	0.60	-0.46	1.09
	(0.95)	(1.06)	(1.42)	(0.96)	(1.55)	(1.03)	(1.22)	(0.83)
Constant	9.34***	9.24***	9.42***	9.0***	9.74***	9.11***	9.53***	8.99***
	(0.26)	(0.30)	(0.40)	(0.27)	(0.46)	(0.30)	(0.37)	(0.25)
Observations	75	72	75	72	75	72	75	72

Notes: Standard Errors are in the parentheses. All regressions are cross-sectional OLS with one observation per country. Executive constraint is on a scale of 1-7 with 7 being more control. The other proxies for institutions are on a scale of -2.5 to 2.5 where 2.5 is more control and better institutions. *Significant at 10 percent; **Significant at 5 percent ***Significant at 1 percent.

closer to the coast have higher levels of GDP per worker. As can be seen in columns (3)-(4) and (7)-(8), distance from the coast is insignificant in determining growth after institutions are instrumented.

The reason why we only include the kk index and rule of law in Table 5 is that Sachs (2003) uses these two proxies, along with expropriation risk measure to show that geography matters. However, data on expropriation risk is no longer available but is still captured in our measure of control of corruption, thus we un-cluster the kk index in the appendix section and show that the results for geography still holds for individual proxies for institutions.

It is important to note that the R-squared in the second-stage least squares is no longer interpretable as the sum of squared residuals can exceed the total sum of squares, which would result in negative R-squares[13].

[13] (R2 = 1 – SSR/TTS).

ALBOUY

Table 5: Core specifications, instrumental variables estimates

	Dependent variable is log GDP per capita in 2015 (current PPP)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rule of law	0.93***	1.24***	1.68***	1.42***				
	(0.25)	(0.33)	(0.31)	(0.32)				
KK index					1.15***	1.02***	1.89	1.17***
					(0.31)	(0.23)	(0.35)***	(0.23)
Malaria Ecology	-0.05***	-0.04***			-0.04***	-0.04***		
	(0.01)	(0.01)			(0.01)	(0.01)		
Distance from the coast			-0.0003	-0.0004			-0.0002	-0.0002
			(0.0003)	(0.0003)			(0.0003)	(0.0003)
Constant	9.29***	9.57***	9.67***	9.57***	9.46***	9.44***	9.52***	9.34***
	(0.25)	(0.13)	(0.21)	(0.19)	(0.10)	(0.10)	(0.19)	(0.14)
Observations	72	69	70	69	72	69	70	69
R-square	0.62	0.50	0.12	0.18	0.61	0.63	0.14	0.48
Endogenous Variable	Rule of law	Rule of law	Rule of law	Rule of law	KK index	KK index	KK index	KK index
Malaria Ecology	0.005	-0.03***			-0.003	-0.03***		
	(0.01)	(0.01)			(0.01)	(0.009)		
Distance from the coast			-0.00001	-0.0002			-0.0002	-0.0004**
			(0.0002)	(0.0002)			(0.0002)	(0.0002)
Log of European settler mortality	-0.36***		-0.32***		-0.29***		-0.28***	
	(0.09)		(0.07)		(0.09)		(0.07)	
Population Density in 1500s		-0.17***		-0.20***		-0.21***		-0.24***
		(0.05)		(0.05)		(0.04)		(0.05)
Constant	1.26	-0.11	1.09	-0.18	1.05	-0.01	1.05	-0.02
	(0.40)***	(0.10)	(0.35)***	(0.14)	(0.37)***	(0.09)	(0.32)***	(0.11)
R-Square	0.25	0.23	0.23	0.18	0.25	0.36	0.23	0.32
F-statistic	11.67	9.90	10.30	7.00	11.68	18.48	10.27	15.44

Notes: Standard Errors are in the parentheses. All regressions are cross-sectional OLS with one observation per country. Executive constraint is on a scale of 1-7 with 7 being more control. The other proxies for institutions are on a scale of -2.5 to 2.5 where 2.5 is more control and better institutions. *Significant at 10 percent; **Significant at 5 percent ***Significant at 1 percent.

Albouy opposes AJR primarily on three grounds: Weak instruments, outliers in specific regions, and lack of "campaign" dummies. Albouy claims that AJR's attempt to show the impacts of property rights on economic performance isn't efficient as there is a positive correlation between the economic measures and the institutional proxies. These correlations run a risk of displaying reverse causalities of the economy on institutional development. AJR's usage of the IV is what Albouy criticizes (Albouy, 2004 a, b). Albouy takes certain specifications that have similar effects to the model[14]. Settler mortality tends to lose significance in the first stage equation determining expropriation risk[15]. On robustness checks and dropping the 36 conjectured data (that mask the collinearity between the controls and the more accurately

Table 6: Instrumental variables estimates with Albouy's settler mortality data

[14] Controls for continents, climate, geography – latitude and neo-European (control) rates, colonial nationality, and population of European origin, percentage of the population living where malaria is endemic in 1994 (used in AJR) and a combination control of Continent Indicators and Latitude

[15] To find that point estimate of the effect expropriation risk on GDP, tends to be as large as or larger than in AJR's. Albouy presents the first-stage estimates of β obtained when one applies the checks using these controls – where his mortality rates show significance in prediction of expropriation risk.

	Dependent variable is log GDP per capita in 2015 (current PPP)						
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Second-stage: Dependent variable is log GDP per capita							
Rule of law	1.394*** (0.298)						
Executive constraint		2.627 (2.736)					
Political stability and non-violence			2.088** (0.828)				
Voice and Accountability				2.523** (1.152)			
Government effectiveness					1.109*** (0.169)		
Regulatory Quality						1.425*** (0.290)	
Control of Corruption							1.389*** (0.327)
Geography (Latitude)	-0.878 (1.194)	0.399 (5.086)	0.258 (1.954)	-1.863 (2.856)	0.422 (0.741)	0.094 (1.019)	-0.974 (1.328)
Constant	9.384*** (0.312)	-5.198 (14.337)	9.73*** (0.669)	9.75*** (0.776)	8.966*** (0.191)	9.057*** (0.262)	9.438*** (0.351)
Observations	33	31	33	33	33	33	33

Panel B: First-Stage for Endogenous Variables (Institutions)							
	RL	EC	PSNV	V & A	GE	RQ	CC
Latitude	0.891 (0.954)	-0.113 (2.182)	0.051 (1.124)	0.883 (1.041)	-0.052 (0.921)	0.189 (0.979)	0.963 (1.080)
Log European settler mortality (logmal)	-0.37*** (0.101)	-0.203 (0.232)	-0.25** (0.119)	-0.205* (0.110)	-0.47*** (0.098)	-0.363 *** (0.104)	-0.37*** (0.115)
Constant	1.43** (0.61)	6.34*** (1.381)	0.792 (0.714)	0.647 (0.661)	2.175*** (0.585)	1.63*** (0.622)	1.396** (0.686)
R-square	0.432	0.032	0.160	0.196	0.493	0.357	0.378
F-statistic	11.39	0.46	2.86	3.67	14.58	8.34	9.13

Notes: Standard Errors are in the parentheses. All regressions are cross-sectional OLS with one observation per country. Executive constraint is on a scale of 1-7 with 7 being more control. The other proxies for institutions are on a scale of -2.5 to 2.5 where 2.5 is more control and better institutions. *Significant at 10 percent; **Significant at 5 percent ***Significant at 1 percent.

measured rates), Albouy notices widening standard errors and gradual increase of significance in controls with the (original) given small sample size[16].

Albouy claims that there is a weak instrument problem (aggravating on application of robustness checks), as β is not significantly different from zero[17]. AJR prove their point first focussing on the significant stage 1 relationship, and then turn to the second-stage estimates. AJR's confidence intervals (by AR) are robust in the presence of weak instruments only to conclude there is no presence of a weak instrument problem in the data series+ (Chernozukov and Hansen, 2005).

+ We have carried out the weak instrument test (see Appendix (Weak Instrument Test)) only to find a similar result to AJR, despite having the knowledge that the AJR analysis employs weak instruments for institutions. This has not, however, affected the significance of any of the variables with which we have worked on the paper.

Albouy argues that 60% of the sample, which are from Latin America and Africa, should be dropped as arbitrary[18] (Albouy, 2012) to which AJR reply with documents substantiating information on mortality of Europeans (AJR, 2012; Gutierrez, 1986). Albouy criticizes AJR's inconsistencies involving both campaigning soldiers and working labourers. He argues that in either case, the mortality rates would have been outliers. Therefrom, Albouy points out that AJR have unfairly exaggerated the mortality rates in areas with currently bad institutions[19]. AJR (2005) show that the results actually get stronger if (missing) data-points are dropped, and argue that limiting the effect of outliers has no impact on our substantive results and if anything significantly strengthens them, even making them robust[20]. The effect of institutions on income per capita is weaker when as covariates the current prevalence of malaria or the percent of population that was of European descent in 1975 (as proxies for geography) are included.

---

[16] Albouy notes that the campaign and labourer indicators become much more significant once the conjectured data are dropped.

[17] . Albouy mentions that AJR's inference based on the IV estimate using conventional asymptotic confidence regions based on the Wald statistic can be grossly incorrect (Dufour 1997)

[18] Albouy notes that there is a lack of data for 36 observations from the dataset of AJR containing 64 countries into account. AJR probably derives the value for these 36 countries from the pattern displayed by the other ones, which according to Albouy might not be plausible inferences as it noticeably reduces the significance of the results, and the mortality rates not being from actual European settlers.

[19] Albouy's paper shows the insignificance of both the first stage (regression of mortality into institutions) and the second stage (institutions on per capita income) if the usual AJR assumptions are altered, thus arguing the data series they constructed suffers from "inconsistencies, questionable judgements, and errors."

[20] AJR (2000, 2001) discussed the issue of high mortality rates at length and used logarithms to reduce the impact.

AJR believe in constructing the series by combining the mortality rates of soldiers (Curtin 1989, 1998), labourers (Curtin et al. 1995), and bishops (Gutierrez, 1986) from different time-periods, mostly prior to the twentieth century. Here, AJR first attack the two variables taken by Albouy for his argument and note issues with the same (AJR, 2005, 2012). The new series with barracks and campaign finds itself with problems including (but not restricted to) selective assignments of mortality rates (West Africa), contradiction to historical records, implausible assumptions on merging of data (Latin America), insistent usage of non-updated data after a significant political and economic overhaul, ignoring self-introduced data (Cohen, 1983) and mixing of weighted and unweighted averages in the construction of the temporal morality rates.

AJR argue that Albouy's treatment (preferred sample) on grounds of dubious decision-making[21]. Albouy is claimed to have put in values, which, if slightly altered, would yield results similar to AJR's (AJR 2005, 2012). Without offering a structured justification, Albouy's treatment selectively assigns of a few countries (Senegal, Ghana) to other countries of similar sorts (Gambia, Cote d'Ivoire)[+]. AJR notes this supposedly unfounded practice to be prevalent while the assignment of mortality rates in more than one county is concerned – wherein Albouy assigns the same rate to all the neighbours of one particular country which displays the mortality rate in question. AJR also calls out on Albouy on inconsistency of data and lack of robustness when regressions are run without the significant country (West Africa, in this case).

[+] While the logic behind Albouy's treatment can be argued to be coherence and overt similarities between countries, no proper structure is available in any of his papers (Albouy, 2004b, 2006).

AJR and build a case against Albouy's criticisms by showing that the only reason that either of Albouy's series give significantly different results from theirs is because of a sequence of odd and indefensible coding decisions.

AJR establish that after a few tiny modulations in the data Albouy's results resemble their own, and point out that Albouy selective data reporting (and misquotations and misrepresentations) as well as ignored data on regions (Austin, Feinberg, Graham et al)[22].

[21] AJR argue that Albouy's results are significantly different from theirs owing to some of the assumptions.
[22] To strengthen the case, AJR organize and discuss in detail Albouy's suggested revisions on Africa, the Americas and Asia (morality rates and data-points for consideration) by region, taking the latter's barrack and campaign series along with their original one.

They denounce the barrack and the campaign series on ground of inconsistency and not having a fixed motive or rationale for being used[23]. Albouy's accusations about the small sample size of the dataset[+] and its subsequent misinterpretations by the authors are put to rest by AJR when they deal with the established sources in reply to Albouy's country set[24]. AJR argue that Albouy's dataset lowered mortality rates making the estimates less reliable as proxies for European Settler Mortality Rate (Curtin, Tulloch).

AJR claim to extend the description of the data (64 to 84 countries)[+] such that it is more representative and robust with additional data collected. Most notably, Albouy manages to make it clear the need to state more explicitly the selection procedure and the idea behind it in the 2001 paper of AJR.[25]

[+]The issue with the AJR-Albouy debate skews heavily toward Albouy's side as he fails to come up with newer points of rebuttal after AJR modifies their argument with solid backing.

AJR show how for each segment of time (war, peace, etc.), they have considered the earliest peacetime number available, and on the non-availability of the information, the earliest mortality estimate from an "expedition" — small samples of campaigns as an alternative for data gathering purposes[+].

[+]In our paper, the logic which we have used (with arguments) for rank-based proxies for institutions when we could not avail data is similar to that of AJR. The reasons why data could not be found is standard and given, and we have pinned down the least number that was available in the rank scale. The smallest number and the largest number in the scale represent the minimum and the maximum intensities of the case at hand, respectively.

AJR uses rates campaigns more frequently in countries with greater expropriation risk and lower GDP, artificially preferring to maintain colonial origins impact countries through institutions instrumented by settler mortality rate. But Albouy claims this to be problematic for AJR[26] as it is well known that soldiers on campaign typically have higher mortality from

[23] Albouy operates with two variables for settler mortality - barracks and campaigns. He thus sets up his argument with these variables, establishing that the impact institutions supposedly have over cross-country growth pattern (Stage 2) in general are weakened or eliminated by showing the weak impact of settler mortality on institutions (Stage 1).

[24] Albouy's disregard for important information on overhauls and transient procedures skewed the data considerably, much to the dismay of AJR, who claim to have followed procedure and referenced facts to denote ranks and numbers to the appropriate cases.

[25] Albouy's success is that it has managed to address the lack of details in data construction for AJR.

[26] According to Albouy, controlling for the source of the mortality rates weakens the empirical relationship between expropriation risk and mortality rates substantially.

disease. He names AJR's result as an artefact of the data's construction (no homogenous pattern of values and data-points), because if these controls are added and the conjectured data are removed, the relationship virtually disappears. Sans environment, much of the variation assigned would have to be measurement aberrations, which, Albouy notes, is disturbing.

[+] AJR only consistently use latitude as a control variable. In our paper, we use latitude as a proxy for geography, along with malaria ecology and distance from ice-free coast.

AJR points out the specification in which the second stage estimates are sometimes insignificant is the one that includes current prevalence of malaria[+], which is a specification that biases result against finding significant effects (endogeneity problem). But even after accounting for malaria, continent dummies (as per Albouy's argument) and latitude, AJR are allowed to statistically reject the hypothesis that institutions have no effect on GDP per capita since the second stage coefficient isn't zero. They capped mortality rates in 250 (per 1000 p.a.) which typically strengthened the results (AJR 2005), so much so that it makes AJR's approach impervious to Albouy's other critiques.

[+]We have also shown the insignificance of malaria (as a proxy for geography) against institutions.

From Table 3, it's evident that almost every instrument except executive constraint is significant at the 1% level when IV regressed with log of GDP. We have controlled for geography proxied by latitude, where we see negative coefficients for three variables. We take Albouy's data of log of settler mortality rates to run regressions and add our proxies for institutions. For constructing the dataset, we take updated data and extend our analysis to show that the results hold significance till date. We see that our results are significant, where geography isn't. Given our instrument of settler mortality, the finding confirms the findings of Acemoglu et al. that colonial origins did indeed have institutions as its most profound reason for income differences across countries. In our IV regression, for example, a country with 0.1 unit more rule of law tends to have a per capita income at PPP that is higher by 1.394 log points. That is quite a significant improvement in the economic performance of a country given that the mean of GDP per capita for our sample is 9.252. The first stage shows the regressions between the settler mortality rates taken from Albouy as the instrument. Albouy's dropping the data is which is why we are also working here with 33 observations.

EASTERLY-LEVINE

Table 7: Determinants of colonial European settlement

	Dependent Variable is Euro Share								
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Population density	-0.035**	-0.035***	-0.037***	-0.040***	-0.037**	-0.0335***	-0.530***	-0.038***	-0.059***
1500	(0.014)	(0.012)	(0.012)	(0.012)	(0.014)	(0.117)	(0.0169)	(0.011)	(0.022)
Indigenous mortality	0.141***	0.144***	0.147***	0.129***	0.126***	0.158***	0.052	0.124***	0.011
	(0.033)	(0.032)	(0.032)	(0.032)	(0.034)	(0.041)	(0.043)	(0.0335)	(0.065)
Geography		0.436***	0.469***	0.525***	0.431**	0.434***	0.475***	0.406***	0.429**
(Latitude)		(0.127)	(0.132)	(0.137)	(0.162)	(0.140)	(0.149)	(0.120)	(0.176)
Precious metals			-0.283						-0.025
			(0.0272)						(0.034)
London				0.000					-6.09e-06
				(0.000)					(09.07e-06)
Biogeography					-0.006				-0.009
					(0.010)				(0.022)
Geography (Malaria						0.001			0.000
Ecology)						(0.001)			(0.003)
Settler Mortality							-0.221*		-0.052**
							(0.129)		(0.0251)
Geography (Distance								-0.000	
from coast)								(0.000)	
Constant	0.06***	-0.341*	-0.031	-0.230	-0.032	-0.050*	0.119	-0.020	0.331**
	(0.020)	(0.019)	(0.020)	(0.378)	(0.291)	(0.030)	(0.082)	(0.026)	(0.150)
Observations	107	107	107	101	75	98	71	95	56
R-square	0.35	0.46	0.47	0.49	0.52	0.49	0.56	0.49	0.66

Notes: Robust standard errors are in the parentheses. *Significant at 10 percent; **Significant at 5 percent ***Significant at 1 percent.

In the third section of the results, we look closely at Easterly and Levine's focus of colonial European share of the population. This focus builds on AJR's focus on the role of colonial origin on development. Whereas AJR focus on the role of institutions, EL create a new dataset on the share of Europeans in the colonies. They find significant results between income today and colonial origin, proxied by colonial European share of the population. They add further controls to see whether the effect of colonial European share of the population remains significant in determining development level differences across states and concur that they do. It must be noted that the use of institutions and settler mortality by AJR only focused on the type of institutions the colonialists set up. EL's use of *euro share* encompasses institutions as well as human capital and technological effects European colonization have had on the colonies and thus on the different levels of incomes today.

We follow the EL paper and run our own regressions on the role of euro share on income by using updated data on income levels for 2015, institution proxied by the KK index that we devised and executive constraint, both average of the period of 2005-2015, and human capital proxied by the education variable that is an average of the period 2006-2015. This is to see whether EL's finding stand the test of time. In addition, we use the three proxies for geography as used in the first section of the results to analyse Sachs' critique.

In table 7, we test the relationship between many independent variables and *euro share,* this is done to see where the colonialists settled. To do so, population density in the 1500s,

geography proxied by latitude, malaria ecology, biogeography, distance from London and distance from the coast as well as settler mortality of the European colonialists and the settler mortality of the local population due to European imported diseases is regressed on *euro share.* In line with the findings of EL we find that Europeans settled in areas where the local population density was higher, geography played a role on where they settled, however what we find contrasting to the finding of EL is that once we account for the European settler mortality as shown in specification (7), indigenous mortality becomes insignificant. However, EL finds indigenous mortality to be insignificant and European settler mortality significant once they include all the controls mentioned above. Once we take in all controls, we find the same result as indicated in specification (9).

Table 8: Log of GDP per capita in 2015 controlling for proportion of population of European descent

	Dependent variable is log GDP per capita in 2015 (current PPP)							
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Euro share	1.670***	1.318***	1.392***	1.926***	1.646***	-0.539	1.867***	1.853***
	(0.369)	(0.549)	(0.330)	(0.487)	(0.386)	(0.548)	(0.335)	(0.559)
Current Euro share	1.670***	1.466***	0.313	0.767	1.353***	0.730**	0.809**	0.510
	(0.369)	(0.394)	(0.323)	(0.599)	(0.408)	(0.350)	(0.358)	(0.599)
British legal origin		1.670***						0.152
		(0.369)						(0.235)
Education			0.015***					
			(0.005)					
Independence				0.610				0.464
				(0.460)				(0.433)
Executive constraint					-0.069			
					(0.081)			
kk index						1.021***		
						(0.119)		
Ethnicity							-1.106***	-0.976**
							(0.314)	(0.374)
Constant	8.640***	8.484***	7.820***	8.422***	8.981***	9.119***	9.028***	8.835**
	(0.127)	(0.157)	(0.371)	(0.160)	(0.422)	(0.120)	(0.197)	(0.293)
Observations	97	95	97	77	90	97	92	76
R-square	0.20	0.25	0.34	0.31	0.21	0.49	0.37	0.42

Notes: Robust standard errors are in the parentheses. Executive constraint is on a scale of 1-7 with 7 being more control. KK index is an average of the six institutional proxies created by Kraay and Kaufman and ranges from -2.5 to +2.5 where 2.5 is more control and better institutions. *Significant at 10 percent; **Significant at 5 percent ***Significant at 1 percent.

In order to study the role of colonial origin on development, we run OLS regressions using *Euro share, British legal origin, Education, Independence*[27]*, executive constraint, the KK index and ethnicity*[28]. To study the true effects of European colonisers, just as EL, we take out all the European countries. Additionally, we use current percentage of population of European descend –*current Euro share*- as an exogenous variable to control for the effect that may have on income levels today and the uncluster that effect from the *euro share* coefficient. From table 8 we find that there is a positive, notable and significant effect on log of GDP per

[27] Independence is the fraction of years since 1776 that a country has been independent (Easterly and Levine, 2015).

[28] Ethnicity the measure of each states degree of ethnic diversity devised by Easterly and Levine in 1997.

capita today by *euro share* except when we include the KK index (our preferred proxy for institutions) and the human capital proxy, education. *Euro share* becomes insignificant when using the latter two regressors, education and KK index, because these two variables account for the channels through which *Euro share* could affect current income. However, it should be noted that the effect of these two regressors on GDP are prone to the simultaneous causality bias when using basic OLS.

Table 9: Log of GDP per capita in 2015 controlling for proportion of population of European descent with Euro Share<12.5%

	Dependent variable: log GDP per capita in 2015 (current PPP)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Euro share	6.278**	6.813***	3.151	9.525***	7.895**	2.546	7.648***	9.542***
	(2.419)	(2.504)	(2.749)	(2.991)	(2.682)	(2.876)	(2.518)	(3.022)
Current Euro share	0.757*	1.004**	0.147	-0.111	0.908*	0.359	0.4005	-0.303
	(0.413)	(0.484)	(0.351)	(0.856)	(0.498)	(0.431)	(0.439)	(0.840)
British legal origin		0.180						0.158
		(0.230)						(0.228)
Education			0.015**					
			(0.006)					
Independence				0.792				0.675
				(0.528)				(0.496)
Executive constraint					-0.091			
					(0.081)			
kk index						1.057***		
						(0.125)		
Ethnicity							-1.07***	-0.876**
							(0.317)	(0.367)
Constant	8.57***	8.412***	7.780***	8.263***	9.01***	9.093***	8.924***	8.610***
	(0.149)	(0.169)	(0.377)	(0.188)	(0.422)	(0.138)	(0.221)	(0.303)
Observations	89	87	89	69	82	89	84	68
R-square	0.074	0.12	0.22	0.18	0.09	0.42	0.25	0.30

Notes: Robust standard errors are in the parentheses. Executive constraint is on a scale of 1-7 with 7 being more control. KK index is an average of the six institutional proxies created by Kraay and Kaufman and ranges from -2.5 to +2.5 where 2.5 is more control and better institutions. *Significant at 10 percent; **Significant at 5 percent ***Significant at 1 percent.

In table 9, we go further and run the same OLS regressions as in table 2 but we only take the sample of non-European countries that have had a colonial European population of less than 12.5%. There are two reasons in doing so, as pointed out by EL. Firstly, the results we get in table 8 could be due the case that highly developed ex-colonies today can happen to have a

large fraction of Europeans in its population. This would bias the study and show significant results in favour of *Euro share* thus colonial origin. Secondly, as argued out in the AJR and

Table 10A: Current income and colonial European settlement

Additional Control:	Dependent variable: log GDP per capita in 2015 (current PPP)						
		British Legal Origin	Education	Independence	kk index	Executive constraint	Ethnicity
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel 1: additional control variables: indigenous mortality, Latitude, Precious metals, London, Biogeography, Malaria ecology and Settler Mortality							
Table 4a sample of Non-European countries							
Euro share	0.939*	1.011	0.784	1.150**	-0.059	1.060**	0.947*
	(0.564)	(0.676)	(0.509)	(0.546)	(0.644)	(0.526)	(01.66)
Current Euro share	1.09*	1.050	-0.132	1.031	0.712	0.55	1.066
	(0.656)	(0.676)	(0.302)	(0.769)	(1.40)	(0.494)	(0.677)
Observations	56	56	56	53	56	54	56
R-squared	0.79	0.79	0.87	0.78	0.70	0.80	0.79

Table 10B: Current income and colonial European settlement

Additional Control:	Dependent variable: log GDP per capita in 2015 (current PPP)						
		British Legal Origin	Education	Independence	kk index	Executive constraint	Ethnicity
Independent Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel 1: additional control variables: indigenous mortality, Latitude, Precious metals, London, Biogeography, Malaria ecology and Settler Mortality							
Table 4b sample of Non-European countries with Euro share<0.125							
Euro share	10.52***	10.50***	2.443	11.71***	7.375***	11.71***	10.58***
	(2.765)	(2.85)	(2.433)	(3.414)	(2.149)	(3.414)	(2.74)
Current Euro share	1.010	1.026	0.051	0.863	0.667	0.862	0.978
	(0.624)	(0.639)	(0.408)	(0.729)	(0.479)	(0.729)	(0.632)
Observations	48	48	48	45	48	45	48
R-squared	0.81	0.81	0.48	0.80	0.84	0.80	0.81

Notes: Robust standard errors are in the parentheses. *Significant at 10 percent; **Significant at 5 percent ***Significant at 1 percent.

We find that once we decrease the sample to less than 12.5% *euro share* only, the coefficient on euro share increases significantly and is statistically significant except, once again, when we use education and the KK index. To sum up table 9, in a sample of countries with less than 12.5% colonial European population, the *euro share* determines income levels significantly,

Table 11: Second Stage IV.

	Dependent variable is log GDP per capita in 2015 (current PPP)					
	Non-European countries				Non-European Countries with colonial euro share <12.5%	
	(1)	(2)	(3)	(4)	(5)	(6)
Current euro share	0.122	-7.607	0.849	-7.682	0.087	-0.488
	(0.476)	(25.10)	(0.582)	(24.48)	(0.087)	(1.915)
Geography (distance from coast)	-0.000	0.001	-0.000	0.001		
	(0.000)	(0.005)	(0.00)	(0.005)		
Education	0.013***	0.0027				
	(0.004)	(0.059)				
Executive Constraint		3.127		3.208		0.502
		(9.231)		(8.080)		(0.451)
kk index	0.712**		0.800		1.777*	
	(0.310)		(0.323)		(1.016)	
Observations	88	84	88	84	89	82
Chi2	91.66	1.89	47.31	1.16	9.71	3.37

Notes: Robust standard errors are in the parentheses. Executive constraint is on a scale of 1-7 with 7 being more control. KK index is an average of the six institutional proxies created by Kraay and Kaufman and ranges from -2.5 to +2.5 where 2.5 is more control and better institutions. *Significant at 10 percent; **Significant at 5 percent ***Significant at 1 percent.

be it through human capita or/and institutions. Additionally, we find, as do EL, that even when we use a small *euro share* sample, the economic effects of European colonialism are positive. EL point out that this is the cause until the *euro share* drops to 4.7%. This is a strong claim on behalf of the colonial origins of income differences across countries.

To further examine our findings, in table 10A we add a number of control variables to see whether the role of *Euro share* is still as significant. These controls are: *indigenous mortality; Latitude, malaria ecology* and *biogeography* as a proxy for geography*; Precious metals* as a proxy for the effect and the role of natural resources; *London,* and *Settler Mortality*. And in table 10B we further use the less than 12.5% *euro share* sample. We find the same results, as in the previous tables that *euro share* is significant. Except of course, when we include human capital or institutions proxies. It must be noted that in our smaller sample of < 12.5% *euro share* even when we include proxies for institutions, the KK index and executive constraint, the effect remains significant which shows that there are other direct effects on

income levels unaccounted for by our human capital and institutions proxies. This could include, as is not limited to, technological effects.

We go a step further and see whether we could use euro share as an instrument on our instrument proxies, kk index and executive constraint. From the results mentioned above, we found that once education and institutions' proxies are accounted for, euro share effect on current income per capita becomes insignificant in the full sample of non-European countries. To test whether it's a valid instrument we run an OLS of education on euro share and current euro share as well as with and without geography and human capital proxies; distance from the coast and education. We find that colonial euro share is insignificant. Running the same regression with executive constraint and then the kk index as the dependent variables, thus the first stage of IV regression, we find euro share to be significant in the case of the kk index (and not the executive constraint). This regression together with the findings of the OLS regressions in the previous tables show that euro share only affects log of GDP per capita through the kk index. We then move on and use euro share as an instrument for kk index. The second stage IV is tabulated in table 5. We find that with and without the geography and human capital proxy, the kk index affects income levels and current share of European descendants in the population remains insignificant. However when we use the smaller sample of euro share <12.5% this results do not hold. This is because there are other channels through which euro share affects income levels making it an invalid instrument. However, overall we affirm the findings of AJR 'the reversal of fortunes'[29] claim and EL's finding that European colonial origin are the prime determinants of economic differences across countries.

[29] . 'Reversal of fortunes' in short is rich countries back in history being poorer today as the colonialists set up extractive institutions in those areas.

CONCLUSION

The paper takes Acemoglu, Johnson and Robinson's argument that colonial origins have influenced the growth pattern across nations mainly through institutions, among other factors. Given the presence of endogeneity, we tried to explain why institutions play an important role by using IV regressions, addressing three distinctive counterarguments and augmentations. We address Albouy's claim on settler mortality usage by using his own revised settler mortality and found AJR's findings stand. Our analysis on Sachs' criticism is founded by the usage of different proxies to show how Geography, though very important a determinant of cross-country growth patterns, is not as impactful as institutions. The paper addresses Sachs' arguments by taking different proxies for Geography that what he has based his arguments on. This not only helps us explore the depth of the argument further, but also provides us with an insight as to what other variables may be considered for a Geography proxy to better describe its effect. The paper finally addresses the euro share used by Easterly and Levine who typically uphold AJR's argument as well as adding that human capital and technology matters as well. Thus justifying the concentration on colonial Euro Share as it encompasses all of the above. EL run OLS regressions, but we have extended their analysis to IV regressions and tried to estimate the results. Our paper instruments current institutions with variables including, but not limited to, European settler mortality rates and population density in 1500s to prove that AJR's results categorically hold up to all of Albouy's three critiques, as well as to Sachs' arguments even when we use different proxies, varied estimation methods, extended analyses and updated data. The paper focuses on EL's analysis on European colonial settlements and human capital, and proves that the EL extension and use of Euro share is econometrically justified. The figures, taken from a plethora of sources, do conform to the dominance of institutions of other factors in influencing growth when colonial origins are concerned.

REFERENCE LIST

- Acemoglu, D., Johnson, S., and Robinson J.A., 2001. *The Colonial Origins of Comparative Development: An Empirical Investigation,* American Economic Review, December, 91, 5, 1369-1401.
- Acemoglu, D., Johnson, S. and Robinson, J.A., 2006. *Reply to the Revised (May 2006) Version of David Albouy's The Colonial Origins of Comparative Development: An Investigation of the Settler Mortality Data.* Unpublished, September, MIT and Harvard.
- Acemoglu, D., Johnson, S. and Robinson, J.A., 2005. *A response to Albouy's A Reexamination Based on Improved Settler Mortality Data*, MIT Department of Economics Working Paper.
- Acemoglu, D., Johnson, S. and Robinson, J.A., 2012. *Hither Thou Shalt Come, But No Further: Reply to "The Colonial Origins of Comparative Development: An Empirical Comment".* April, MIT and Harvard Working Paper 16966.
- Acemoglu, Daron, Simon Johnson and James A. Robinson, 2002b. "The Rise of Europe: Atlantic Trade, Institutional Change and Economic Growth" NBER Working Paper #9378.
- Acemoglu, D., Johnson S., Robinson J.A., Thaicharoen Y., 2003. "Institutional Causes, Macroeconomic Symptoms: Volatility, Crises and Growth," Journal of Monetary Economics, 50, 49-123.
- Albouy, D., 2004b. *"The Colonial Origins of Comparative Development: A Reexamination Based on Improved Settler Mortality Data",* University of California — Berkeley, December.
- Albouy, D.Y., 2008. *The colonial origins of comparative development: an investigation of the settler mortality data (No. w14130)*. National Bureau of Economic Research.
- Albouy, D.Y., 2012. *The colonial origins of comparative development: an empirical investigation: comment. The American Economic Review,* 102(6), pp.3059-3076.

- Boulhol H., 2004. *Technology differences, institutions and economic growth: a conditional conditional convergence,* Centre d'Études Prospectives et d'Informations Internationales.
- Curtin, P.D., 1998. *Disease and empire: The health of European Troops in the Conquest of Africa.* Cambridge University Press.
- Diamond, J.M., 1997. *Guns, Germs and Steel: The Fate of Human Societies,* W.W. Norton & Co., New York NY.
- Easterly W., Levine R., 2015. *The European Origins of Economic Development (Working Paper 18162).* National Bureau of Economic Research.
- Knack, Steven and Philip Keefer, 1995, *Institutions and Economic Performance: Cross-Country Tests Using Alternative Measures, Economics and Politics,* 7, pp. 207–27.
- Rodrik, D., 2000. *Institutions for high-quality growth: what they are and how to acquire them.* Studies in Comparative International Development, 35(3), pp.3-31.
- Rodrik, D., Subramanian A., and Trebbi F., (forthcoming). "*Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development",* Journal of Economic Growth.
- Hall, R.E. and Jones, C.I., 1999. *Why do some countries produce so much more output per worker than others? (No. w6564).* National bureau of economic research.
- Fernandez, A.M. and Kraay, A., 2005. *Property rights institutions, contracting institutions, and growth in South Asia: Macro and Micro Evidence.* The World Bank. Background paper prepared for the SAARC Business Leaders' Conclave: South Asia Regional Integration and Growth, New Delhi, November, pp.17-18.
- Green, A. and Moser, C., 2012. *Do Property Rights Institutions Matter at the Local Level? Evidence from Madagascar.* Journal of Development Studies, (December 2012), pp.1–15.
- Gehlbach, S. and Keefer, P., 2011. *Investment without democracy: Ruling-party institutionalization and credible commitment in autocracies.* Journal of Comparative Economics, *39(2),* pp.123-139.

- McArthur, J.W. and Sachs, J.D., 2001. *Institutions and Geography: Comment on Acemoglu, Johnson and Robinson (2000) (No. w8114).* National bureau of economic research.
- North, Douglass C. and Robert P.T., 1973. *The Rise of the Western World: A New Economic History,* Cambridge University Press, Cambridge UK
- Bloom, D.E., Sachs, J.D., Collier, P. and Udry, C., 1998. *Geography, demography, and economic growth in Africa. Brookings papers on economic activity,* 1998(2), pp.207-295.
- Haggard, S. and Tiede, L., 2011. *The rule of law and economic growth: where are we?,* World Development, 39(5), pp.673-685.
- Glaeser, E.L., La Porta, R., Lopez-de-Silanes, F. and Shleifer, A., 2004. *Do institutions cause growth?,* Journal of economic Growth, 9(3), pp.271-303.
- Vollrath D., 2014, *A Sceptic's Guide to Institution Part 1, 2, 3, 4.* The Growth Economics Blog. Available from: https://growthecon.wordpress.com/2014/11/18/the-skeptics-guide-to-institutions-part-1/, https://growthecon.wordpress.com/2014/11/20/the-skeptics-guide-to-institutions-part-2/, https://growthecon.wordpress.com/2014/11/23/the-skeptics-guide-to-institutions-part-3/, https://growthecon.wordpress.com/2014/11/26/the-skeptics-guide-to-institutions-part-4/ (Accessed 3rd March 2017).
- World Bank Group Development Indicators (2015) Historical data World Bank Group. Available at: http://www.wb.org/custom-query (Accessed: 22nd March 2017).
- World Bank Group (2015) Historical data on GDP World Bank Group. Available at: http://www.wb.org/custom-query (Accessed: 22nd March 2017).

APPNEDIX

Table 11: Variable definitions and sources of data

Variables	Sources	Description of each variable
Current Euro share	Easterly and Levine (2015)	Current European fraction of the total population
logmal	Albouy (2005)	Albouy's Log of settler mortality
LGDP	The World Bank	Log of GDP per capita in 2015 (in PPP)
KK index	Created by the Authors	Index of the six institution's proxies created by Kraay and Kaufman averaged out: va, psnv, ge, rule, cc, rq over the period of 2006-2016 (abbreviation details below)
Euro share	Easterly and Levine (2015)	Fraction of the total population in the colonies representing Europeans
Executive Constraint	Polity IV data	Constraint on the executive averaged for the period 2006-2015 with a scale of One to Seven
Voice and accountability	Kaufman and Kraay government indicators	Voice And accountability indicator averaged for the period 2006-2015. On a scale of -2.5-2.5
Political stability and non-violence	Kaufman and Kraay government indicators	Political Stability and Non-violence indicator averaged for the period 2006-2015. On a scale of -2.5-2.5
Government effectiveness	Kaufman and Kraay government indicators	The effectiveness of government indicator averaged for the period 2006-2015. On a scale of -2.5-2.5
Rule of law	- Kaufman and Kraay government indicators	Rule of Law indicator averaged for the period 2006-2015. On a scale of -2.5-2.5
Control of corruption	Kaufman and Kraay government indicators	Control on corruption indicator averaged for the period 2006-2015. On a scale of -2.5-2.5
Regulator quality	Kaufman and Kraay government indicators	Regulatory quality indicator averaged for the period 2006-2015. On a scale of -2.5-2.5
Settler Mortality	AJR (2005)	Logarithm of settler mortality in 83 ex-colonies as deaths per year per 1000 people
Population density in 1500	AJR (2001, 2005)	Log of local population each square km in the 1500s
Latitude	Easterly and Levine (2015)	Value of Latitude divided by 90 to be used as a geography proxy
Malaria Ecology	Easterly and Levine (2015)	"An index of the stability of malaria transmission based biological characteristics of mosquitos such as the proportion of blood meals taken from human

		hosts, daily survival of the mosquito, and duration of the transmission season and of extrinsic incubation." (EL, 2015)
Indigenous mortality	Easterly and Levine (2015	A dummy for the mortality rates of the local population as a result of European colonialists' diseases
Biogeography	Easterly and Levine (2015)	Log of the number of native plants species and of native animals species to exhibit domestication.
Education15	World Bank Development Indicators	Secondary school enrollment rate averaged for the period between 2005-2015
Ethnicity	Easterly and Levine (2015)	Index of ethnic diversity created by Easterly and Levine in 1997
Independence	Easterly and Levine (2015)	A country's independent as the fraction of years since 1776
London	Easterly and Levine (2015)	Distance of each state from London
Ex-colony	Easterly and Levine (2015)	A Dummy variable for ex-colonies (equaling one) created by AJR
British Legal Origin (legaluk)	AJR (2005)	A Dummy for countries with British legal origin and civil law
Distance from the coast	Easterly and Levine (2015)	Average distance of a country from the ice free coast or sea-navigable river
Precious metals (metals)	Easterly and Levine (2015)	Natural recourses dummy

THEORY (EXPOSITION)

Hervé Boulhol highlights the importance of institutions to realize economic growth. He argues in his paper[30] that technology in the economy is a strong component to attain productivity, which is primal to our analysis of growth and development. Boulhol's study mainly focusses on the three major effects of institutions on economic growth: efficiency in the use of technology, technology diffusion and eventual long-term TFP growth. In order to found a rational explanation as to how exactly do institutions form an essential part of the growth and development process for the economy, the Boulhol model provides us with an exceptional insight. His study wanders around the interactions between institutions and technology differences to explain cross-country growth pattern. In the paper[31] authored by

[30] Boulhol H, *Technology differences, institutions and economic growth : a conditional conditional convergence*, CEPII, 2004

[31] AJR

Acemoglu, Johnson and Robinson, it is argued that the differences in the economic institutions are the fundamental cause of differences in economic development. They theoretically debate on the concept, as well as empirically demonstrate how economic institutions determine the incentives of and the constraints on economic actors, and shape economic outcomes. In this paper, the basis of the argument presented is how Boulhol's take on technology and its approximation as a form of constructive institution impacts the cross-country growth pattern. The analysis is extended to deduce how technology (as a representation of institutions) help in the economic growth of a nation. The quality of institution matters profoundly, as the degree of constructiveness decides how well the economies grow. Boulhol comments on trade being a linear influence on growth, as well as conditional convergence being conditional on the similarity of institutions in terms of quality and efficiency of the nations concerned. Acemoglu, Johnson and Robinson argue that they are consequential social decisions. Because different groups and individuals typically benefit from different economic institutions, there is generally a conflict over these social choices, ultimately resolved in favour of groups with greater political power.

Not only limiting the scope of the paper to technological and social institutions, Acemoglu, Johnson and Robinson feature in the socioeconomic and political institutions that are at play by taking instances of two "quasi-natural experiments" in history, the division of Korea into two parts with very different economic institutions and the colonization of much of the world by European powers starting in the fifteenth century. Settlements affected early institutions; and early institutions persisted and formed the basis of current institutions. Settler mortality is emphasized owing to its ability to determine the types of institution (which affects growth – the hypothesis here) in the long term. To make matters simple, Boulhol talks about convergence in the long run, which corroborates with Acemoglu's arguments on long run stability in the economy being the prime outcome of deterministic institutions. An instrument for Acemoglu, Johnson and Robinson to judge the foundation of the type of institution (extractive or inclusive) is settler mortality rate. The type of institution not only decides on the growth level, but also the quality of economic growth realised in the country concerned. With the help of the analogous model of Boulhol and the technological (hence, institutional) influences on his analysis, an explanation on how the quality and the type of institution matter for the country can be made.

The arguments of AJR are considered, as their intentions can theoretically ably be explained by the three arguments forwarded by Boulhol. The ideas are juxtaposed to form a better-

looking scenario, which can substantiate the claims of how the distribution of politico-economic power in society is in turn determined, by political institutions and the distribution of resources.

Boulhol's Model for Technological Diffusion and Institution for Growth

Technological diffusion is the process by which innovations (new products, processes or management methods) spread within and across economies. Boulhol does not assume that technological profiles of the countries across the world are on a similar level, or that the differences in the level of technology in the countries are treated as residuals. Therefore, technological progress are not independent from other explanatory variables, and that it directly plays a role in determining the shape of the country's future. He states that changes in output per capita is not just explained by differences of capital per unit of labour.

Assuming a Cobb-Douglas production function and with standard notations,

1. $\mathbf{Y_i = (A_i L_i)^{1-a-b} K^a_i H_i^b}$

where Y is output, K and H are stocks of physical and human capital respectively; A is the productivity level and L the number of workers. Output per worker y_i for the country i is therefore given by

2. $\mathbf{y_i = Y_i/L_i = (A_i)^{1-a-b}(K_i/L_i)^a (H_i/L_i)^b}$

with $Z^{KH} = K(H/L)^{b/a}$ defining a capital aggregate of the physical capital stock and the human capital stock per capita. Equation (2) illustrates that initial output per worker is certainly linked to initial productivity.

If it's assumed that the productivity level A_i is independent of the country ($A_i = A, \forall\ i$), then Boulhol finds through empirics the ratio of output per capita in between two extreme cases (countries) translates into a highly unrealistic capital aggregate, Z^{KH}, per capita ratio of unrealistic proportions using a physical capital share of one-third.

Moreover, recognising the productivity differences, it is apparent from equation (2), valid at each time, that both the initial productivity level and the initial output per capita are closely linked, which renders growth equation estimates assuming identical productivity seriously biased. For the sake of exposition, it can be safely assumed that the initial level referred to here can be taken as the initial point for which Acemoglu, Johnson and Robinson consider the settler mortality rate. This settler mortality rate will go on in the later periods to determine

which kind of institutions crop up and how they are developed over the years. The parameter A is often indistinctly designated as either the productivity or the technology level. In this analysis, the parameter is broken into two parts. The distinction is made between the two notions, and the complement of technology in productivity is called "efficiency".

A_i is broken into two components: a pure technology level Bi and the degree of efficiency in using this technology X_i such that A_i equals B_iX_i. A country which is affluent and economically developed is chosen as the benchmark country, with which the comparisons of the other countries can be made (a proportion can be calculated). (Bassanini and Scarpetta (2001))

$b_i = B_i/B_{ben}$ is denoted, which is an inverse indicator of the distance to the frontier, $x_i = X_i/X_{ben}$, is the ratio of the relative efficiency to the benchmark, and $a_i = A_i / A_{ben} = b_i.x_i$, the relative productivity level. Boulhol as an influence on the technological efficiency Xi and possibly the technology diffusion that is most simply governed by considers institutional quality:

3. $\dot{b}_i(t) = v_i.(1 - b_i(t))$

where t stands for time: in the long run technologies converge to the frontier at a pace represented by v_i, which Boulhol claims to be constant across countries or institutionally related.

The pure technology component is assumed to converge (or diverge if v_i is negative), and total productivity discrepancies may persist as a result of differences in institutionally-related efficiency X_i. Here institutions enter into growth equations through three different channels: the level of technological efficiency, the progress of technological efficiency and possibly the speed of technology diffusion. Noticing that the productivity level A_i can be written as $A_{ben}.b_i.x_i$, TFP-growth can be divided into three components:

4.

$$\frac{\dot{A}i}{Ai} = \frac{\dot{A}ben}{Aben} + \frac{\dot{x}i}{xi} + \frac{\dot{b}i}{bi}$$

The first term on the right is simply the benchmark TFP-growth, denoted by g, the second, denoted $-c_i$, is the long run TFP-growth deficit to the benchmark and the third is the technological catch-up component derived from resolving the differential equation (3). The proportion mode of the equation explains how exactly the institution factor figures out in the

model. Taking similar origin points, the institutions (manifested in such formats) are the deterministic variables when countries decide to choose their path to economic growth.

5.

$$\frac{\dot{A}i}{Ai} = g - ci + \frac{vi(1 - bi(0))}{e^{vi.t} - (1 - bi(0))}$$

Institutions here have an impact through $X_i(0)$, through c_i, and potentially through v_i. Integrating equation (5) into the growth equation, the model that is developed is a growth model. This is quite similar to the macroeconomic augmented Solow Model enabled to take into account the heterogeneity of technologies and the contribution of institutions. With n_i denoting the population growth, d the physical and human capital depreciation rate, s_i^K and s_i^H the fraction of total income invested in physical and human capital respectively, Boulhol has derives extensively (the growth model derivations) and establishes the following growth equation:

6.

$$\frac{Log\ yi(t) - Log\ yi(0)}{t} = -\frac{1 - e^{-\beta i.t}}{t}.Log\ \frac{yi(0)}{ai(0).Aben(0)} + \frac{1 - e^{-\beta i.t}}{t} Log\left[\frac{(s_i^K)^a.(s_i^H)^b}{ni + d + g - ci}\right]^{\frac{1}{1-a-b}} + g - ci + f_{v,t}(bi(0))$$

where $\beta_i = (1 - a - b).(n_i + d + g - c_i)$ is the usual speed of conditional convergence and the last term

$$7.\ f_{v,t}(bi(0)) = \frac{1}{t} Log \frac{1-(1-bi(0))e^{-vi.t}}{bi(0)} \left[\approx \left[\frac{1}{bi(0)} - 1\right] vi\right]$$

is the contribution of technology diffusion to growth: it is positively related to the speed v_i and to the distance to the technological frontier. Equation (6) is to be compared to the augmented-Solow growth equation which is exactly the same as if we assume that initial total productivity level $A_i(0)$ is identical across countries, that every country is at the frontier (b_i = 1) and that there is no long term TFP-growth differences (c_i = 0). The growth process is therefore the result of four distinct forces: the "adjusted" absolute convergence – the

convergence is lessened here because of overall productivity level differences, hence the adjective “adjusted”, a second convergence component coming from technological catch-up, the usual non-convergence stemming from differences in long term paths due to different investment rates, and an additional divergence force coming from long term TFP-growth differences.

Boulhol considers the similar levels of technology in each country because it is difficult to scrutinize the differences that are there.

Equation (6) shows that there is the relative level of initial productivity. From (6) and (7), the speed of convergence can be computed and the contribution of institutionalised technology can be screened out therefrom. A part η of initial income ratios can be explained by initial productivity ratios, the complement 1 – η being explained by capital differences, and therefore:

8a.

$$\boldsymbol{ai(0) \equiv \frac{Ai(0)}{Aben(0)} = \left[\frac{yi(0)}{yben(0)}\right]^{\eta}}$$

8b.

$$\boldsymbol{\eta = \frac{Cov\,(Log(ai(0)), Log(y(0)))}{Var(Log(yi(0)))}}$$

Estimates of stocks of physical and human capital to infer productivity levels can be made and it can be assessed that productivity differences account for half or more of level differences. Since productivity is born out of institutions, and enhanced by technology (resources and capital included), it can be safely assumed that institutions lead to growth and eventual economic convergence in a country. According to equation (2), with η and $a = b = 1/3$, the output per capita ratio between the two countries (out of which one might be the benchmark country) is explained by the impact of the productivity level ratio and a capital aggregate, Z^{KH}. For sure, the extreme simplification in equation (8), which has the merit of highlighting the potential correlation between initial productivity and initial output per worker, as well as the comparison of how efficient an institution might be and its reflection

on the corresponding level of economic growth pattern, shows how countries fare in the presence of institutions (i.e. how the institutions impact long term behaviour of the economy). (Klenow and Rodriguez-Clare (1997))

WEAK INSTRUMENTS TEST

Post estimation Test: depvar indepvar (endogenous variable = instrument)

- Endogeneity Test: estat endog
  Ho: variables are exogenous
  If P value is very small, then reject null hypothesis

- Weak Instrument Test: estat firststage, all forcenonrobust
  Ho: instruments are weak.
  Partial R-squared measures the endogenous variable and the instrument after we partialling out exogenous variable, if high and if F-statistic has to be the large, then to reject the null hypothesis.
  Partial R-sq. in the table in Stata shows that post partialling correlation
- Over Identification Test: estat overid
  Ho: instrument set is valid and model is correctly specified.
  If P value for Sargan and Basman tests are large, then the null is correct.

We use the Weak Instrument Test.

<u>The Weak Instrument Test</u>

The general trend in the weak instrument test is what we expect out an AJR paper: weak instruments, but significant enough to prove a point. For the weak instrument test (see the do file), we took the three instruments that we have: logsm, lpd1500, logmal and euros.

For logsm, the instruments are weak, as we cannot reject the null hypothesis. Partial R2 is low, and so is the F-statistic, which proves the weakness in logsm as an instrument. The minimum eigenvalue statistic is equal to the F-statistic where its magnitude determines whether we accept or reject the null. The Executive Constraint proxy shows the minimum F-statistic in our analysis, while the strongest one is Government Effectiveness. lpd1500 is used for two sections: AJR and Sachs. AJR never really uses this, but we use this to explore the

concept we are dealing with. For AJR, the pattern is similar. However, the noticeable thing is the magnitude of the F-statistics. The values are extremely low in nature, far weaker than logsm, which probably explains its results. The minimum value of F-statistics is availed from the Control on Corruption proxy. For AJR, we always have used latitude (distance from the equator), as the exogenous variable and the proxy for geography. However, for Sachs' paper, things get more exciting. We run the regressions on not one, but two proxies of geography – distance from the coast and malaria ecology. Running the same test on Sachs' results would yield us the same answer, both for lpd1500 and for the lesser used logsm. It is inferred that this is a marked weak instrument. When regressed with proxies like distance from the coast and malaria ecology, Control of Corruption, Government Effectiveness, Political Stability, and Voice and Accountability, have returned decent numbers, which enabled us to reject the null hypothesis of weak instruments. Regressions on Albouy's paper returned similar results like Acemoglu indicated which again proves how he does almost the same thing AJR. Government Effectiveness and Executive Constraints have the highest and lowest values again, respectively, when logmal is used as an instrument. This proves that like logsm, logmal – derived from Albouy's revised set – is as much a weak instrument. We have extended Easterly-Levine's idea by incorporating IV regressions in our analysis (which we can see using the endogeneity test – whether or not it is at all necessary. We have not discussed about the endogeneity test in this context). By taking the euros as an instrument, we find that the F-statistic is small in every which regression, thus implying the euros is a weak instrument too. Since we do not use euros as an instrument anywhere else, we check for weak instrumentation only when we discuss Easterly-Levine's ideas extending from their analysis.

The critical values shown in the table is used to test whether or not we reject the null hypothesis. Some values are significant at 5%, some at 10%, some at 15%, and so on. However, the generic trend, with some fluctuations obviously, shows that the instruments used are weak instruments.