What Rules in the 'Deep' Determinants of Comparative Development?

Alvar Kangur
International Monetary Fund
Address: 1900 Pennsylvania Ave NW, Washington, DC 20431, USA
E-mail: AKangur@imf.org

Abstract

During the previous two decades or so empirical literature on comparative development of na- tions has turned to investigation of 'deep' determinants of differences in income levels, such as institutions, trade, geography and human capital. In this paper I revisit this debate and make the following contributions. First, when run in one common framework, the previous results are generally not robust to the choice of measures of institutional quality as well as their respective instruments, and might be misspecified. Institutional measures of objective constraints should be strictly preferred. Second, a careful model selection across all the instruments for all the deep determinants reveals a clear pattern of instrumentation. Most notably this indicates that settler mortality proposed by Acemoglu et al. (2001) is not a dominant instrument for institu- tional quality for which legal or linguistical instruments should be preferred. Consequently I provide evidence that the theory of colonial origins is not institutional in its nature and rather supports human capital prevalence hypothesis. Third, human capital and geography come out as 'winners' exerting separate direct effects on income levels with openness having indirect complementary effects either via institutions or natural resource exports.

JEL classification: O11, O40, P51
Keywords: comparative development, institutions, openness, geography, human capital, model selection

1. Introduction

For the past two decades or so the empirical literature on growth and development experienced a new swing. Many researchers have turned to investigating "deeper" determinants of income levels like institutions, openness, geography, and human capital. As it has been facetiously noted, what is surprising is not so much an idea exploited, but the fact that the very same people who thus far had been busy setting up state-of-the-art endogenous models of growth are now saying that differences in income levels were to a some extent shaped centuries ago and can be explained by simple two-equation instrumental variable model. In this literature it seems that the race has been twofold: first, to determine if a particular hypothesized deep determinant of relative income levels "rules" over others, and second, in many respects that has been a battle of instruments. As we are reminded by Rodrik et al. (2002)"an instrument does not a theory make", they are crucial for testing theories and stating certain causal effects.

In this article I revisit this debate. I start by reviewing the key contributions in the literature, revealing the considerable variety of different measures and instruments put forward, especially those that concern institutions. Such diversity, however, induces a considerable degree of fragility into races over what rules. To illustrate and facilitate comparison across specifications, I return the most influential races proposed in this literature within one common framework that comprises elements, the superiority of which has been put forward by a number of well-targeted studies. The striking result is that the outcomes are very different from those reported in the original articles and, even within this unified framework, exhibit considerable variation, allowing to make no firm conclusions over what rules. To a large extent this variability depends upon which particular measure or instrument of institutions is used.

To alleviate the discovered problems with identification as well as with model uncertainty over which instrument influences which endogenous determinant, thus supporting which theory, I next let the instruments speak for themselves. This is done through applying instruments to the scrutiny of three model selection strategies: Basic specific-to-general regressions, Bayesian Averaging of Classical Estimators and an automatic general-to-specific selection routine. The results clearly reject geographical variables as well as settler mortality as instruments for institutions. Even further, no variable pertinent to colonial origins has any predictive power for institutional quality, rather there is strong evidence of colonial origins reshaping the initial conditions of human capital. Both legal as well as linguistic variables seem to have the best statistical properties as far as instrument relevance is concerned.

In light of these results I next revisit the fundamental question over model uncertainty: what is it that European settlers exactly brought with them to their colonies? The main dichotomy here lies between institutional and human capital prevalence. I do so by rerunning the exercise by Acemoglu et al. (2001) under both hypotheses and further demonstrate the weakness and fragility of settler mortality as an instrument for institutions. However, settler mortality and especially the share of European settlers in 1900 turn out as relevant instruments for human capital, thus supporting the hypothesis by Lipset (1960) and Glaeser et al. (2004) of human capital prevalence.

Following Glaeser et al. (2004) and Acemoglu et al. (2005), there have been attempts to bring this question to dynamic panel data methods. While the former provides some evidence that causality should run from education to institutions, the latter finds a negative

albeit insignificant effect of education to democracy once period-specific effects are included. In Kangur (2009), a companion article to this, I argue and provide evidence that the properties of commonly used measures of institutional quality are in fact those of policy flows pursued at a particular point in time, and not those of stock variables to which institutional thinkers like North (1981) usually refer to. This is evident as the available indices already by construction do not capture any permanent or durable characteristics that should reflect the entire history of institutions - the same feature of persistence on which Acemoglu et al. (2001) have built their argument. Reinterpreting the available measures of institutional quality as policy flows allows us to find the positive effect of higher educational attainment on future institutional changes.

Finally, the natural question arises: what rules? As Sachs (2003) has put it, the empirical specifications in the literature seem to be "worryingly oversimplified" as "there is good theoretical and empirical reason to believe that the development process reflects a complex interaction of institutions, policies, and geography." I briefly discuss possible non-linearities in that higher integration with world markets may itself encourage better protection of property rights through higher monitoring of government actions, and show that direct effects of all deep determinants can be obtained. Remarkably, openness is very weakly and in most cases insignificantly linearly related to almost any variable. Second, when taking on board human capital, institutions lose significance. The direct effects of geography and natural resources are a norm, whereas the impact of openness to income levels is complementary to natural resource endowments.

The paper is organized as follows. In section 3 I re-run the main 'races' in the literature outlined above in a unified framework. Based on this I highlight the key issues that will be addressed in subsequent sections. Section 4 touches briefly on the choice of institutional indices. In section 5 I run a race between all available instruments to determine their relevance to each of the deep determinants. On the face of these results section 6 revisits the colonization theory of Acemoglu et al. (2001) and argues for human capital prevalence. In section 7 I address the problem of model misspecification, documenting the potentially substitutable roles of institutions and integration. Further evidence of human capital dominance over institutions together with direct effects of geography and natural resource endowments as well as complementary effects of trade integrations are provided. Section 8 concludes.

2. Literature Review

A nice exposition and introduction to the debate on "What Rules?" is given by Frankel and Romer (1999). Their purpose is to identify the effect of trade on international differences in income per capita levels. As an instrument for trade they use predicted trade shares from a gravity model, which explains bilateral trade with economic as well as geographical factors such as distance between trading countries and their economic size. These predicted trade shares provide in some sense a perfect instrument since geography is as exogenous a factor as an economist can possibly hope to get. Although, as we will see later, many of the geographical variables have a direct impact on current income levels and would thus violate the exclusion restriction, the geographical notions used in classical gravity models still seem to have acceptable properties for an instrument.

Hall and Jones (1999) in identifying the causal effects of broad institutions or social capital on income levels, use linguistic and geographical instruments - the extent to which five primary Western European languages are spoken today, the share of population speaking English, and distance from the equator - all of which are assumed to be correlated with Western European influence. These authors fail to reject the overidentification tests and document strong the impact of their measure of social capital to output levels.

Acemoglu et al. (2001) came up with a rather clever though more controversial idea. Their hypothesized deep factor, analogous to Hall and Jones (1999), is the quality of broadly defined institutions. To identify the effect of the latter they use exogenous variation provided by the historical data on European settler mortality. Since Europeans at the time of colonization presumably had superior institutions, and institutions are thought to be persistent, settler mortality rates could carry information of the types of early institutions set up in colonies that in turn determine current institutions.

In a related work Acemoglu et al. (2002) take one step back in the history of colonization to further investigate the settlement patterns of Europeans. They argue that in densely settled regions that were also relatively rich (urbanization seemed to be highly positively correlated with income per capita), European settlers were more likely to create extractive institutions. However, regions with lower density and also lower per capita income were easier to settle. In these places European settlers built up stronger institutions, thereby causing institutional reversal. Hence data on urbanization and population density in 1500 can also be used as a part of the instrument set.

Rodrik et al. (2002) introduced the practice of 'running races' between candidates, in their case between exogenous geography and endogenous trade and institutions. In instrumenting for the latter they alternate between settler mortality and linguistic variables. As a result they claimed the primacy of institutions: after controlling for institutional quality, trade does not exert any significant impact on income, often entering with negative sign, whereas geography measured by the distance to the equator impacts income levels only indirectly through institutions. This was an especially strong result showing that, in the words of the authors, "geography is not destiny" and "focusing on increasing the economy's links with world markets is unlikely to yield convergence."

Both of these findings, however, have been subject to scrutiny. Many researchers, most notably Diamond (1997), Bloom and Sachs (1998), Sachs (2001), Olsson and Hibbs Jr. (2005) and others, have made claims of the direct effects of geography-based variables on income. Sachs (2003), a long time advocator of the importance of geography and trade in economic development, argues that "distance from the equator, the centrepiece of testing in Rodrik et al. (2002), is an exceedingly poor choice for a serious test of geographical variables. It is at best a proxy, and a poor one at that, for climate or possibly for distance from major markets, and should not be used as the basis of the bulk of the tests in the Rodrik et al. (2002) paper when much better alternatives are available." A particular alternative Sachs (2003) has in mind is the risk of malaria transmission that he instruments using the measure of ecological conditions constructed by Kiszewski et al. (2004). Besides settler mortality he relates institutions to the share of population in temperate zones. The results are clear: malaria transmission risk has a direct negative impact on relative income levels, even after controlling for institutional quality. Hereafter the direct effects of the health environment on incomes have been documented rather frequently; see for example McArthur and Sachs (2001) and Carstensen and Gundlach (2006).

Glaeser et al. (2004) bring out two important critiques concerning the empirics of institutions. First, the choice of a measure for institutional quality used in the empirical estimation has been rather arbitrary. In particular, they note that the institutional measures exploited thus far, such as Protection against Expropriation Risk and Rule of Law, are all outcome measures reflecting choices those in power, whereas a keyword in any discussion of institutions is constraints.[1] A natural reference here is Douglass North, who defines institutions as "a set of rules, compliance procedures, and moral and ethical behavioural norms designed to constrain the behaviour of individuals in the interest of maximizing the wealth or utility of principals." (North 1981, p. 201-202). As authors show, these measures are only weakly correlated with the available indicators of written constitutional constraints, reflecting much broader concepts than institutions. Authors first choice that comes closest to capturing the political environment rather than choices, albeit still imperfectly as it is also an outcome measure, is Executive Constraints from the Polity IV dataset.

A second, and much more fundamentaal, criticism concerns the vagueness of thr underlying theory of colonial origins, or the influence of Europeans in general. The view of the role of institutions in comparative development advocated by Acemoglu et al. (2001) and followed by many others is that of 'institutional prevalence,' which will then determine factor accumulation and the choice of the production function. Another view put forward by Lipset (1960) and more recently supported by Djankov et al. (2003), Rajan and Zingales (2006) and Glaeser et al. (2007) is that of 'human capital prevalence': it is the endowment of human capital of nations that determines constituencies and institutional opportunities, and therefore, is the fundamental cause of growth and development. In usual Barro-type growth regressions, Glaeser et al. (2004) find no relationship between growth and available measures of constitutional constraints. Therefore, it is not clear what exactly was the 'contribution' of Europeans to their colonies: was the first and fundamental impact in shaping institutions, or shifting the balance of human capital endowments? Indeed, in their 2-equation type race between executive constraints and average years of schooling in the second stage, only the latter comes out as significant. Their instrument set includes both colonial instruments - settler mortality and population density in 1500 – as well as French legal origin, popularized by Porta et al. (1997, 1998, 1999) in their extensive work on legal origins of property rights protection. The argument is that countries differ considerably in investor protection and law enforcement depending on their legal practices. Common law tends to provide the best investor protection, while German and Scandinavian civil law have the best quality of law enforcement. French civil law comes out as the worst in both categories. The first-stage regression reveals that while both colonial instruments are very significant predictors of both deep determinants, their legal instrument predicts only institutions. In sum, notwithstanding identification problems, the Glaeser et al. (2004) exposition is much more supportive of human capital prevalence.

Rajan and Zingales (2006) revisit institutional and human capital explanations of slow growth and argue that it is a bad configuration of interest between different constituencies that leads to underdevelopment. In their model the most crucial role is assigned to the initial distribution of factor endowments – such as human capital – that leads to the emergence of constituencies that through the power of setting up institutions and pursuing policies are able

[1] A natural reference here is Douglass North, who defines institutions as "a set of rules, compliance procedures, and moral and ethical behavioral norms designed to constrain the behavior of individuals in the interest of maximizing the wealth or utility of principals." (North 1981, p. 201-202).

to reproduce themselves over time. In the words of the authors: "poor education policy [or institutions supporting such policies] might well be the proximate cause of underdevelopment, but the deeper (and more long-lasting) cause is the initial conditions (like the initial distribution of education) that determine political constituencies, their power and their incentives." It is even ambiguous whether institutions are endogenous or exogenous; bad policies can persist without bad institutions. In their simple regressions it is the share of European settlers in 1900 - a measure of educated constituencies – and not initial democracy levels that has strong explanatory power for both initial levels of schooling as well as modern income levels. This view clearly supports the human capital prevalence hypothesis.

Yet another determinant emerging from the endogenous growth literature is the extent to which countries are engaged in innovative activity. Lederman and Saenz (2005) have constructed a database of innovation indicators and use the stock of patents as a measure of innovations, instrumented with an average Index of Patent Rights (IPR) constructed by Ginarte and Park (1997) and Park (2001). In a race with institutions (measured by Rule of Law) they conclude that the economic magnitude of the impact of the stock of patents on relative income levels is at least as large as for institutions, if not larger.

However, conventionally, researchers have neglected innovative activity as an equivalent candidate for a 'deep determinant.' This is because innovations reflect advances in total factor productivity that increases the efficiency of reproducible factors of production, and consequentley belongs to the production function itself. Also, efficiency promoting innovations cannot take the appropriate growth-enhancing dimensions without a certain amount of property rights being enforced and a critical mass of human capital being released from other activities. Or in other words, innovative activity is itself a function of human capital and institutions – the deeper determinants. As will become clear later, in races between innovative activity and institutions (or human capital), the identification of partial effects is seriously hampered, suggesting that only one of these variables belongs to underlying equation.

3. Motivation

3.1. A Unified Framework

These aforementioned contributions form the basis of the 'State of the Art' around which the discussion here is centred. Underlying theories presented appeal to the common concepts of (broad) institutions, trade, geography or human capital – the endogenous or 'deep' deter- minants. The main question of interest is which of these determinants has direct impact on income levels and which indirect through its impact on another? Or, in another words, what rules?

Virtually every study I consider in this section sets up the following textbook 2-equation instrumental variable (IV) identification strategy:

$$\log y = \mathrm{X}\beta + \mathrm{Q}\gamma + \epsilon \tag{1.1}$$

$$\mathrm{X} = \mathrm{Q}\delta + \mathrm{Z}\theta + \nu \tag{1.2}$$

where y denotes output per capita, **X** denotes the set of hypothesized 'deep determinants' (e.g. social capital in Hall and Jones (1999), **Z** is a set of variables influencing solely **X**, and vector **Q** includes truly exogenous covariates potentially influencing both y and **X**. In order for this 2-equation strategy to be applicable, **Z** must obviously have some explanatory power for **X**, or:

$$\theta \neq 0 \tag{1.3}$$

This condition means that a given element of **X** and a corresponding element of **Z** must be partially correlated, once the effect of other exogenous regressors is netted out. In this structural model the endogeneity of **X** must be clearly recognized. Thus one needs an identifying restriction, or an exclusion restriction, if you like. The crucial assumption here is that the variables included in **Z** affect y only through deep determinants **X** and not directly or through any other possible channel. Thus our restriction takes the simple form:

$$E(\mathbf{Z}'\epsilon) = 0 \tag{1.4}$$

Under these assumptions any variable in Z is a valid instrument for a corresponding variable in X and can be used in estimating (1.1) directly without a need of complete specification of equation (1.2). Finally, the estimated system has a unique solution only if matrix E(Z'X) has full rank. Necessary for this is the order condition: one must specify at least as many instruments as endogenous determinants. Besides solving an endogeneity problem, a classic article of this type would argue that the estimation setup has the potential to solve both omitted variable and measurement error problems.

We have already seen the multitude of candidate instruments: colonial, linquistic, legal and geographical variables are all exploited to identify the partial effects of institutions.[2] In many cases it is not even clear which instrument should be used for which endogenous determinant, and consequently, which theory should be associated with which instrument. Furthermore, as much as there are competing instruments there are also different indicators used to measure endogenous determinants, especially institutions. Many of the institutional indices like Rule of Law or average Protection Against Expropriation Risk are subjective, possibly reflecting political and economic developments instead of institutional constraints. With this, such indices are open to various channels of causation and could potentially absorb the effects of other determinants, thus leading to spurious conclusions. These indicators and instruments have been used in all possible combinations, supporting contradicting results.

To put the discussion into perspective and facilitate any meaningful comparison across the aforementioned studies, there is a clear need for a common framework. In what follows, I will set up such a framework consisting of elements the superiority of which has been put forward by targeted studies. Under this unified framework I will then rerun the main specifications put forth in the literature as was described in the introduction. This makes it possible to clearly illustrate and identify the shortcomings in the current literature.

Before turning to outlining a unified framework, a note on the IV methodology is in order that in words of Sachs (2003) is "worryingly oversimplified." The limitations are evident in the presence of such variety of instruments that can identify partial effects of

[2] The list of 'institutional' instruments goes on, including measures of ethnolinguistic diversity, trade, and natural resources. I will discuss many of the related contributions in due course.

institutions (or other deep determinants) through a multitude of channels, and the researcher is rather restricted in identifying, which one matters. Any attempt to screen out the channels by running the type of horse races described before can easily lead to instrument multicollinearity and thus failure of rank and order conditions as vividly demonstrated by Dollar and Kraay (2003a,b). Relatedly, as growth is inevitably a dynamic process, many researchers would certainly prefer using dynamic specifications.

However, throughout this paper I will follow the 2-equation strategy that, once complemented by recent model selection techniques, has also many merits. First, in many respects, this strand of the literature has evolved as a battle of theories and associated instruments that are not exploitable in a dynamic framework. Second, the model selection methods that allow screening out the data-congruent instruments could not only help to alleviate the identification problems, but can also shed light on the relevance of underlying theories. Again, such 'instrument races' are not possible under dynamic specifications. Third, a dynamic setup is not suitable for investigating the direct effects of geography on income levels. Neither can it allow the identification of any indirect or complementary channels through which geography could exert its influence. Furthermore, current applications of dynamic panels do not take into account that we only observe ordinal flow measures for institutions, in which case the exploitation of time-variation for identification purposes is not straightforward.

The unified framework used in this section is defined as follows:

Institutions. Following Glaeser et al. (2004) I measure institutional quality using an average index of Executive Constraints. Although still imperfect, this *objective* measure is intuitively closer to what is meant under institutions than other indices used in the literature. As almost all earlier studies exploit popular subjective indices, it provides a subtle point of comparison, especially since this is a choice to which final results are rather sensitive. For the second source of considerable variation in final results – the instrument for institutions - I exploit settler mortality put forward by Acemoglu et al. (2001), henceforth AJR, in parallel with Hall and Jones (1999), henceforth HJ, as linguistic instruments. As these have been the two most popular 'competitors' in the literature, reporting parallel results enormously facilitates comparisons with earlier work;

Openness. The measure of openness is the current price average trade/GDP ratio taken from Rodrik et al. (2002). To be consistent with the bulk of the literature, I use Frankel and Romer (1999) gravity model predicted trade shares as instruments for openness. As these authors point out it is important to control for countries market size to account for within-country trade: smaller countries tend to engage more in international and less in within-country trade. I consider population and area as standard proxies for market size. Where appropriate, the estimation outcomes are also checked against Frankel and Rose (2002) re-estimated gravity model. The former includes a richer variety of geographical information (like various common border interactions), whereas the latter incorporates information on common languages instead while also dropping home-country population. Therefore, to the extent that geographical or market size variables might have a direct impact on income levels Frankel and Rose (2002) instrument can provide a useful robustness check. However, as will be shown later in section 1.5, when all relevant exogenous factors are netted out, there is no strong basis to prefer one instrument over another;

Geography. Following Sachs (2003), I take malaria falciparum transmission risk as the first choice for geographical conditions, instrumente using the index of malaria ecology

derived by Kiszewski et al. (2004) that in turn is built upon climatological and vector conditions and as such claimed to be exogenous to a country's level of development. Again, for the purpose of comparability with earlier studies, the only exogenous variable added at this stage is distance to equator.

Human capital and innovative activity. These are treated as in Glaeser et al. (2004) and Lederman and Saenz (2005) respectively. In particular, the measure for human capital is the standard average years of schooling of the population aged over 25 from Barro and Lee (2000) dataset.

Income levels and sample size. The dependent variable in all regressions is the log of GDP per capita PPP in 2000 taken from Heston et al. (2006) PWT version 6.2. In all the regressions throughout the paper I utilize maximum sample supported by data. First, as pointed out by Frankel and Romer (1999), given the nature of the instrument it is important to consider as broad a sample as possible in identifying the direct income effects of trade. Second, in many cases data availability considerably reduces the sample size, in which case restricting it further to a common sample would leave us with dangerously low degrees of freedom. Third, with an exception of 'neoeuropeans' for which data exists, final conclusions in general do not hinge on a few observations.

Table 1.1. IV regressions from the Literature

Panel A: Second-stage regressions for real GDP per capita PPP in 2000

	FR (1999)		AJR (2001) HJ (1999)		RST (2002)		RST (2002)		Sachs (2003)		Glaeser et al. (2004)		LS (2005)		Education vs innovation	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Ln Openness	0.45*	0.14			0.34	0.44*	0.28	0.51***								
	(0.233)	(1.044)			(0.363)	(0.249)	(0.442)	(0.175)								
Executive Constraints			0.69***	0.67***	0.61***	0.51***	0.89*	0.26*	0.84*	0.36***	-0.50	-0.02	0.59**	0.37**		
			(0.161)	(0.060)	(0.147)	(0.093)	(0.480)	(0.135)	(0.485)	(0.087)	(0.515)	(0.298)	(0.287)	(0.159)		
Distance from Equator			0.01		0.02	0.02**			0.013	0.001	-0.02	0.005	0.02	0.02*	0.02	0.04*
			(0.014)		(0.013)	(0.007)			(0.013)	(0.005)	(0.025)	(0.011)	(0.021)	(0.009)	(0.032)	(0.019)
Malaria Falciparum Risk							0.31	-1.50***	0.45	-1.18***						
							(1.36)	(0.175)	(1.35)	(0.298)						
Average Schooling											0.87**	0.40**			2.32	0.50
											(0.419)	(0.185)			(4.02)	(0.318)
Ln Stock of Patents													0.04	0.08	-1.63	-0.30
													(0.209)	(0.101)	(3.20)	(0.331)
Ln Population		0.12														
		(0.188)														
Ln Area		-0.16**														
		(0.080)														

Panel B: Tests for weak instruments

n	146	145	80	140	78	132	77	128	79	154	65	66	61	89	80	82
F-stat	3.74	3.44	29.3	123.7	21.3	65.75	14.64	76.91	15.51	74.76	12.86	35.78	25.58	71.96	2.67	30.53
Excluded instruments			AJR	HJ	AJR	HJ	AJR	HJ	AJR	HJ	AJR	HJ	AJR	HJ	AJR	HJ
Shea partial R2																
institutions			0.249	0.383	0.258	0.201	0.038	0.102	0.056	0.130	0.048	0.054	0.075	0.085		
openness	0.554	0.065			0.502	0.527	0.484	0.523								
malaria							0.066	0.204	0.107	0.310						
education											0.049	0.091			0.004	0.071
innovation													0.109	0.160	0.004	0.055
Anderson LR statistic (p-value)	118.01	9.742	22.89	67.60	22.48	29.14	2.85	13.61	4.35	20.44	2.65	3.53	4.51	7.89	0.31	4.60
	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.092)	(0.001)	(0.037)	(0.000)	(0.104)	(0.171)	(0.034)	(0.019)	(0.579)	(0.032)
Cragg-Donald F-statistic (critical value: 10% IV si	179.15	9.80	25.50	20.95	12.36	10.46	0.92	3.45	2.13	7.05	1.27	1.12	2.19	2.60	0.15	2.25
	(16.38)	(16.38)	(16.38)	(24.58)	(7.03)	(13.43)			(7.03)	(13.43)	(7.03)	(13.43)	(7.03)	(13.43)	(7.03)	(7.03)

*** - 1% significance, ** - 5% significance, * - 10% significance. Heteroskedacity corrected standard errors in parentheses.

Table 1.1 reproduces the estimation results of the main empirical specifications from articles described in the literature review in one unified framework. Columns 1 and 2 test for the main specification of Frankel and Romer (1999); although trade instrument is always significant in the first stage, under country size controls, trade share loses its power to predict income levels. Columns 3 and 4 reproduce the 'institutional' findings of Acemoglu et al. (2001) and Hall and Jones (1999), respectively, with strong identification. From here onwards all odd-numbered columns use AJR settler mortality and even-numbered columns HJ linguistic variables as instruments.

In Columns 5 and 6 I run a Rodrik et al. (2002) type race between trade, institutional quality and exogenous geography. AJR´s instrument allows us to make similar conclusions as in the original work - institutions rule, whereas trade and geography do not have a direct impact on relative income levels. Although, unlike Rodrik et al. (2002), the first stage of the regression (not reported) reveals that geography does not affect income levels through institutions, rather it plays a role in predicting integration.[3] On the other hand, results using HJ´s linguistic instruments are a clear contradiction to the original work - all three deep determinants have a significant direct impact on incomes! Even further, and again contrary to Rodrik et al. (2002) original setting, the baseline specification in Table 1.1 is rather fragile with respect to the inclusion of fixed factors (results not reported).

In columns 7 and 8, instead of latitude I measure geography using the malaria transmission risk, instrumenting it with malaria ecology. Overall results are similar to those obtained in columns 5 and 6: again a clear difference emerges between AJR and HJ instruments. Malaria falciparum risk has strong and negative direct impact only in the larger sample under HJ´s instruments that still allow us to identify the significant effects of all three deep determinants, whereas under AJR´s instrument, institutions rule with no direct role for geography. From the first stage regressions (not reported), it can be seen that settler mortality predicts both institutional quality and modern malaria environment. This raises concerns about the AJR exclusion restriction, as is reflected in the considerably weakened Shea (1997) partial R2 coefficients as well as in Anderson (1984) LR statistic that is on the border of not rejecting the null of model underdeterminancy. Finally, in columns 6 and 8, trade would maintain its significance under the larger AJR sample, only in the smaller AJR sample would it cease to be significant, though in both occasions either latitude or malaria would still have direct income effects.

In columns 9 and 10, I replicate Sachs (2003) though without any additional excluded geographical instruments. Again, a clear difference between AJR´s and HJ´s instruments strikes the eye: the malaria transmission risk has a strong and negative direct impact only under HJ´s instruments independent of sample size, contradicting the original work. In the first stage, settler mortality predicts both institutional quality and modern malaria environment that considerably weakens its partial correlation with executive constraints. Cragg and Donald (1993) - Stock and Yogo (2005) test statistic falls significantly below its critical value, indicative of weak identification, especially under AJR´s instrument.

When running the Glaeser et al. (2004) race between human capital and institutions in columns 11 and 12, it is always human capital that ‹rules› the second stage. However, settler mortality and linguistic instruments work through both deep determinants (Acemoglu et

[3] It should also be remembered that in the original work, Rodrik et al. (2002) measured institutions using the Rule of Law. As will become clear later, this choice of institutional measure enforces the result obtained by these researchers.

al. 2002) instrument of population density in 1500 is working only through human capital) and models are clearly underidentified. Similar patterns occur if one runs a race between all four endogenous determinants that is clearly too much for the simple IV identification strategy to handle. Though still human capital and malaaria would come out as winners.

The next two columns present a Lederman and Saenz (2005) type exercise using institutional quality and innovative activity as deep endogenous determinants. These results without exogenous controls suggest the dominance of institutions under both sets of instruments and once again settler mortality is a strong predictor of both deep determinants. In the original work Lederman and Saenz (2005) control for various geographical variables; the inclusion of malaria risk in particular drastically overturns the results for both specifications, making the stock of patents instead of institutions to ‹rule› the second stage.[4]

Finally, in columns 15 and 16, I run a race between human capital and innovative activity just to demonstrate that, as far as deep determinants are concerned, these two carry similar information. Nothing rules, apart from perhaps geography, specifications are rank deficient or close to it, suggesting that one of these variables might be redundant. Since human capital and innovative activity are likely to carry similar information, for the purposes of this paper I would be inclined to think of innovative activity as 'semi-deep' since it is itself an outcome made possible by 'deeper' factors (such as human capital).

3.2. The Problems

The messages that come through from Table 1.1 are rather striking. First, most of the specifications when tested in a common framework are a direct contradiction of their counterparts in the original articles. Some of these differences are due to the institutional measure, as most original pieces concentrate on subjective indices. Second, even without subjecting the specifications to standard robustness tests it is evident that, across all the regressions in Table 1.1, no deep determinant has a robustly significant impact on income levels. In other words, nothing rules. Of the long list of instruments, the two most common ones alone – AJR settler mortality and HJ linguistic variables – seem to account for a large part of the variability in the results. Below I raise the four most intriguing issues that are investigated in subsequent sections.

Indices of institutional quality. The evidence presente in the previous section reveals that the choice of measures for institutional quality can make a considerable difference in the final conclusions. This issue is briefly discussed in section 4.

Instruments and model underdeterminancy. Similarly to institutional indices, the results in Table 1.1 do vary across the institutional instruments used. In particular, this first pass suggests caution in the relevance of settler mortality as instrument. It must be stressed that this feature is not only pertinent to Executive Constraints: in Table 1.2 settler mortality does not even come close to predicting any – including subjective – measures of institutions. Model underdeterminancy - high correlation between instruments for diferent endogenous

[4] The two equations in columns 13 and 14 are estimated without included instruments. Once the original Lederman and Saenz (2005) set of exogenous instruments are included (that in the first stage predict mainly institutions), the results are reversed: innovative activity 'rules,' whereas joint exogeneity cannot be rejected. Although weaknesses in the specification do not allow any strong conclusions, these results would suggest that the level of economic development depends on a nation´s ability to engage in increasing returns activities. In this world institutions reflect mostly differences in country fixed effects.

variables - as evident from Table 1.1 that potentially render the rank and order conditions invalid is reminiscent of though not exclusive to settler mortality. Share of population in temperate zone or population density in 1500 are other frequently exploited instruments sharing the same faith. Also, in further robustness tests (results not reported), the two HJ linguistical instruments often fail the test of overidentifying restrictions, suggesting that only one of them is relevant for institutions. Therefore, greater care must be taken in selecting the instruments prior to running such horse races. I take this issue up next in section 5 by subjecting the instruments to a battery of model selection techniques. In other words, this time I let the instruments themselves to run the race for each of the five endogenous variables.

Competing theories. Importantly, such careful selection allows not only to alleviate identification problems, but to draw conclusions on the underlying theories themselves. The prime example here is Glaeser et al. (2004) question of what the European colonizers brought with them to their colonies? Was it institutions, human capital, culture of trade or something else? Building upon results from instrument selection in section 5, I revisit this issue in section 6 and demonstrate that European influence is working entirely through the human capital channel.

What rules? Finally, building upon the results from the instrument selection, I return to the prime question of interest: which endogenous determinant is 'deeper' than others. Or what rules? Although, as will become evident, instrument selection considerably alleviates identification problems, IV races between multiple endogenous determinants are still complicated. Notwithstanding these limitations, various pieces of evidence point towards a consistent and frequently observed pattern. Section 7 summarizes these findings.

4. The Choice of Institutional Indices

To further illustrate the statistical properties of the most commonly used indices, in Table 1.2 I regress three of them on potential instruments, country heterogeneity and a selection of fixed factors; a regression that could resemble the first stage of 2SLS. The results vary enormously. Average Protection Against Expropriation Risk is almost solely determined by malaria and continent heterogeneity, or by income differences once the latter is included. By contrast, Rule of Law can be predicted by almost anything one would throw into regression. The same applies to the average of six governance indices as used in Kaufmannn et al. (2006). Executive Constraints is most closely associated with all potential instruments and the only one out of these three not so evidently influenced by current income levels. Another index with similar properties is the measure of political constraints due to Henisz (2000) that is based on veto points of independent branches of government; the correlation between this measure and the Polity IV index of executive constraints is above 0.9.

Table 1.2. Institutional Measures

Dependent variable - measure of institutions

	Rule of Law		Risk of Expropriation		Executive Constraints	
Ln GDP pc PPP	0.54***		1.07***		0.35	
	(0.109)		(0.176)		(0.233)	
Ln Settler Mortality	0.02	-0.02	0.18	0.19	-0.11	-0.09
	(0.056)	(0.064)	(0.200)	(0.179)	(0.165)	(0.169)
French legal origin	-0.33**	-0.29	-0.30	-0.10	-1.36***	-1.33***
	(0.149)	(0.203)	(0.409)	(0.436)	(0.341)	(0.350)
Engfrac	0.01	0.17	-0.23	0.22	0.22	0.35
	(0.227)	(0.307)	(0.544)	(0.545)	(0.720)	(0.689)
Eurfrac	0.62**	0.77***	0.75	1.14	2.47***	2.55***
	(0.260)	(0.255)	(0.951)	(1.047)	(0.660)	(0.685)
Malaria Falciparum Risk	-0.15	-0.52*	-0.22	-1.25**	-0.82	-1.20**
	(0.244)	(0.289)	(0.480)	(0.504)	(0.539)	(0.568)
Share of population in temperate zone	0.32	0.82**	-0.15	0.70	-1.70**	-1.40**
	(0.270)	(0.336)	(0.397)	(0.474)	(0.755)	(0.682)
Landlocked	0.19*	-0.08	-0.13	-0.43	-0.17	-0.32
	(0.100)	(0.151)	(0.497)	(0.433)	(0.289)	(0.266)
Net Oil Exporter	-0.35***	-0.20	-0.05	0.25	-0.20	-0.09
	(0.096)	(0.129)	(0.194)	(0.266)	(0.274)	(0.265)
Africa	0.08	-0.27	-0.81	-1.67*	-0.36	-0.66
	(0.375)	(0.349)	(0.929)	(0.991)	(1.10)	(1.09)
Asia	-0.34	-0.53	-0.59	-0.84	-0.37	-0.67
	(0.370)	(0.420)	(1.02)	(1.127)	(1.05)	(1.03)
Latin-America and Caribbean	-0.74**	-0.89**	-2.03***	-2.64***	-1.30	-1.51
	(0.311)	(0.365)	(0.603)	(0.607)	(0.971)	(0.953)
Observations	77	78	68	69	77	78
Adj. R-squared	0.84	0.74	0.74	0.63	0.71	0.69

*** - 1% significance, ** - 5% significance, * - 10% significance. Heteroskedacity corrected standard errors in parentheses.

With these results in mind, one should be very cautious about the choice of a measure for institutions. It is no wonder that subjective indices like Rule of Law - a centrepiece of empirical testing in Rodrik et al. (2002) as well many other applications - can produce results consistent with institutional dominance over geography with the latter affecting income levels only indirectly through institutions. Because of its correlation with standard geographical controls the index absorbs the effect of the latter at the first stage of IV estimation. Similarly, an average index of Protection Against Expropriation Risk – a centrepiece of regressions in Acemoglu et al. (2001) – is more likely to be consistent with the colonization hypothesis of institutions due to its correlation with malaria and continental dummies through which settler mortality influences institutions and income. As will become clear in section 6, these two channels account for most of the explanatory power settler mortality exerts on institutions. At the same time, as shown in the context of Table 1.1, under executive constraints no conclusion is robust. A compelling case for the latter has been made by Glaeser et al. (2004), which I will follow throughout the paper.

5. Race Between Instruments

Perhaps the most intriguing message that comes through from section 3.1 is the interplay of institutional instruments with fixed factors, as well as instruments of other determinants. Various theories have come up with multiple candidate variables for instruments, but their channels of influence are unclear at best, while their improper use can easily lead to identification problems. Therefore, a careful selection of instruments would facilitate identification and help to reveal the relevance of underlying theories. This is achieved through the following three model selection procedures.

The first strategy resembles the specific-to-general selection methodology and consists of two steps. First, I regress real GDP per capita in 2000, one by one, on each of the five deep determinants (institutions, human capital, trade openness, malaria transmission and innovative activity) instrumenting these, one by one, with all variables that could potentially instrument for any of these endogenous determinants. From this step instruments with robust partial correlation with at least one endogenous determinant are selected. The second step then consists of running the same IV regressions separately for each deep determinant, but this time using the full set of instruments selected in the first step. Both stages include exogenous controls. The selection here not only ensures that the non-zero (3) condition continues to hold in a setup with multiple endogenous variables, but also minimizes the chances for rank and order conditions to fail by screening out collinear instruments. Results from this specific-to-general procedure coincide well with the other two methods and are available in the working paper version.

The second model selection procedure used is Bayesian Averaging of Classical Estimates (BACE), which is a relatively recent approach to accommodate for model uncertainty, popularized in growth literature most prominently by Sala-i-Martin et al. (2004). It incorporates a special case of Bayes' rule to average over all possible model combinations. In statistical terms this rule defines the posterior distribution of the parameters of interest as a weighted average of all possible conditional posterior densities where weights are given by the posterior probabilities of the corresponding models. The posterior probabilities themselves (under diffuse priors) are computed from a limiting case of posterior odds ratio that allows us to approximate the Bayes factor using BIC. Throughout this article I have assigned equal prior probabilities to all models. The final posterior inclusion probabilities of interest are a sum over all posterior model probabilities for all the models where a particular variable is included.

The third strategy is known as a general-to-specific model selection, based on the theory of reduction popularized by Hendry (1995). This is carried out using the PcGets automatic IV model selection routine described and exploited by Hendry and Krolzi (2001, 2004) among others. Finally, note that only candidates selected at the first stage of the specific-to-general strategy are subjected to the general-to-specific and BACE selection. This allows us to considerably economize on degrees of freedom and gain in efficiency as it would not be appropriate to subject variables that violate the basic non-zero condition to further scrutiny.

Before turning to the estimation and in addition to the variables already described in previous sections, I amend the pool of potential instruments with further popular choices of variables from the literature. The first such candidate is average ethnolinguistic fragmentation that Levine (1997) argue seriously hampers public policy-making and through this channel also prospects for growth. This is supported by Porta et al. (1999) who find ethnolinguistic

heterogeneity to be a major factor influencing the quality of governance after netting out the impact of a range of other relevant controls.

The second potential candidate instrument comes from Bockstette et al. (2002) who have constructed an index of state antiquity. The underlying idea is that countries with longer indigenous state history are more likely to have better developed institutions. Amending Hall and Jones (1999) set of instruments with their index, these authors claim it to be an important instrument for 'social capital.'

The motivation for the next two candidates for instruments – "wheat-sugar" ratio and sugarcane – comes from Engerman and Sokol (1997, 2002, 2005) who emphasize natural resource endowments in shaping the patterns of institutional quality. The latter is a dummy variable taken from Easterly and Levine (2003) indicating whether a country produced sugarcane in 1998/99 or not. The former is the log of a ratio of the share of arable land suitable for wheat to the share of arable land suitable for sugarcane constructed by Easterly (2006). The underlying idea is that regions more suitable for cultivation of economies of scale crops like sugarcane developed large scale plantations fuelled by the workforce of slaves, whereas regions more suitable for cultivating crops like wheat developed small-scale farming. With more equal distribution of wealth, the 'wheat' regions set up institutions more supportive of private property rights and ultimately leading to early industrialization.

Lastly, I take on board a set of variables motivated by Diamond (1997) and popularized by Olsson and Hibbs Jr. (2005) – biological conditions and orientation of the major axis of the continent. Diamond argues that biogeographically better endowed continents with a larger variety of domesticable animals and plants suitable for cultivation, and a vertical orientation of its major axis more favourable for the diffusion of technologies created a natural advantage for earlier agricultural transformation from hunter-gatherers to sessile societies. This transformation in turn was responsible for the emergence of superior 'guns, germs and steal,' allowing a few of the possessors of these to conquer entire civilizations, as well as to develop writing, code of law, and as a result superior institutions. To account for these biogeographical advantages Olsson and Hibbs Jr. (2005) have gathered data on early domesticable animals and cultivable plants, giving us a measure of bioconditions. I also use their measure of geoconditions and orientation of continents together with a simple one-zero dummy for continent axis.

The results for model selection are presented in Tables 1.3 and 1.4. Table 1.3. reports the posterior inclusion probabilities from BACE for each endogenous determinant. In odd-numbered columns estimation is performed over twelve potentially important 'deep instruments,' whereas in even-numbered columns most relevant instruments are combined with 5 categories of exogenous control. Estimates significant at conventional levels are indicated in bold. Table 1.4 performs the PcGets automatic general-to-specific model selection routine under a conservative strategy both with and without exogenous controls. Odd-numbered columns report selection on the full set of potential instruments only. Even numbered columns add the twelve most common exogenous factors (not reported). A conservative strategy is exploited to screen out the most relevant instruments that would survive the robustness tests and avoid weak identification. OLS 'testimation' coefficients of the selected congruent model for each deep determinant are reported.

Table 1.3. The Race Between Instruments – Posterior Inclusion Probabilities from Bayesian Averaging of Classical Estimates

	Institutions		Integration				Geography		Human Capital		Innovation	
	1	2	3	4	5	6	7	8	9	10	11	12
1 COLONIAL HISTORY												
Settler Mortality	0.487 (-)	0.803 (-)	0.147 (-)	0.143 (-)	0.158 (-)	0.141 (-)	**0.927** (+)	0.314 (+)	0.598 (-)	0.839 (-)	0.530 (-)	0.367 (-)
European Settlers in 1900	0.134 (-)	0.153 (-)	0.176 (-)	0.282 (-)	0.276 (-)	0.277 (-)	0.180 (+)	0.185 (-)	**1.000** (+)	**0.961** (+)	0.517 (+)	0.218 (+)
Ln Population density in 1500	0.183 (-)		0.353 (-)				0.300 (-)		0.125 (-)		0.627 (-)	0.832 (-)
2 LINGUISTICAL INSTRUMENTS												
Engfrac	0.155 (+)	0.132 (+)	0.617 (+)	0.161 (+)	0.144 (+)	0.183 (+)	0.221 (+)	0.161 (-)	0.650 (+)	0.217 (+)	0.390 (+)	0.318 (+)
Eurfrac	**0.999** (+)	**0.999** (+)	0.167 (+)	0.290 (-)	0.193 (-)	0.410 (-)	**0.997** (-)	0.240 (-)	0.149 (+)	0.157 (-)	**0.956** (+)	**0.836** (+)
Ethnolinguistic fractionalisation	0.155 (+)		0.723 (+)				0.173 (+)		0.214 (+)		0.142 (-)	
3 LEGAL INSTRUMENT												
French legal origin	**1.000** (-)	**1.000** (-)	0.167 (+)	0.849 (-)	**0.907** (-)	0.716 (-)	0.194 (-)	0.720 (+)	0.254 (-)	0.620 (-)	0.178 (-)	0.170 (-)
Average IPR index	0.207 (-)		0.124 (-)				0.126 (+)		0.234 (+)		**0.965** (+)	**0.959** (+)
4 PREDICTED TRADE SHARES												
Frankel-Romer (1999)	0.130 (-)	0.180 (-)	**1.000** (+)	0.591 (+)	**0.898** (+)		0.235 (+)	0.160 (+)	0.146 (-)	**0.954** (-)	0.146 (-)	0.181 (-)
Frankel-Rose (2001)	0.128 (-)	0.161 (-)	0.165 (-)	0.562 (+)		**0.891** (+)	0.339 (-)	**0.880** (-)	0.142 (+)	0.182 (+)	0.156 (+)	
5 MALARIA Malaria Ecology	0.360 (-)	0.163 (-)	0.485 (-)	0.132 (-)	0.139 (-)	0.131 (+)	**0.812** (+)	**0.922** (+)	0.572 (-)	0.214 (-)	0.516 (-)	0.329 (-)
6 GEOGRAPHY I												
% of population in temperate zone	0.145 (-)	0.369 (-)	0.122 (-)	0.364 (-)	0.426 (-)	0.298 (-)	0.331 (-)	0.507 (-)	0.159 (-)	0.241 (+)	0.223 (+)	0.166 (+)
Frostdays		0.217 (-)		0.145 (+)	0.143 (-)	0.148 (+)		0.149 (-)		**0.999** (+)		**0.863** (+)
Distance from equator		0.157 (-)		0.330 (-)	0.255 (-)	0.375 (-)		0.624 (-)		**0.958** (-)		0.148 (+)
7 GEOGRAPHY II												
Landlock		0.132 (-)		0.176 (-)	0.198 (-)	0.208 (-)		**0.862** (-)		0.233 (-)		0.163 (+)
% of pop. within 100 km of coast		0.175 (+)		0.764 (+)	**0.894** (+)	0.596 (+)		**0.918** (-)		**0.865** (+)		0.805 (+)
8 NATURAL RESOURCES												
Net exporter of oil		0.119 (-)		0.667 (+)	0.649 (+)	0.702 (+)		0.222 (-)		0.145 (+)		0.126 (+)
Share of natural capital		0.153 (+)		0.160 (-)	0.160 (-)	0.171 (-)		0.159 (+)		0.289 (-)		
9 REGION												
Africa		0.138 (+)		0.175 (+)	0.141 (-)	0.233 (+)		0.242 (-)		0.213 (-)		0.384 (-)
Asia		0.129 (+)		0.167 (+)	0.180 (+)	0.167 (+)		0.257 (-)		0.464 (+)		0.228 (-)
Latin-America & Caribbean		0.150 (-)		0.279 (-)	0.194 (-)	0.439 (-)		**0.991** (-)		0.192 (-)		0.237 (+)
10 MARKET SIZE												
Population		0.332 (-)		**0.996** (-)	**0.994** (-)	**1.000** (-)		0.150 (+)		**0.976** (-)		0.161 (-)
Area		0.147 (+)		**0.918** (+)	**0.874** (+)	**0.851** (+)		**0.905** (-)		0.153 (+)		0.254 (+)

A clear and very similar pattern emerges from all three selection strategies. Looking first at the main candidate instrument proposed by colonial theory – settler mortality – it is striking that not a single regression presented would select it as a robust predictor of institutional quality. Even more importantly, under all three selection methods settler mortality comes out as a significant predictor of current disease environment and in most cases even dominates malaria ecology, the natural climatologic instrument for malaria transmission risk. Only BACE suggests that geographical controls can screen that information out from mortality rates. Given these findings and the fact that malaaria conditions have been shown in so many cases to have direct effect on income levels, this provides prima facie evidence against using settler mortality as an instrument for institutions.

Population density in 1500, according to Acemoglu et al. (2002), as the major factor behind European settlement decisions, bears no relation whatsoever to modern institutional quality. Rather, if at all, it shows some weak statistical relation to trade or innovations. This indicates that there is too much uncertainty in assigning the story running from early population density to early settlements to current income levels to any of the deep determinants, but if anything, it is not institutional. Similarly to settler mortality and past population density, the most complete measure of European settlement outcomes, the share of European settlers in 1900, has no explanatory power for institutional quality either. Instead, across all three model selection strategies, it comes out as the most robust and the strongest predictor of average

Table 1.4. The Race Between Instruments – PcGets Model Selection under Conservative Strategy

	Institutions		Integration		Geography		Human Capital		Innovation	
	1	2	3	4	5	6	7	8	9	10
Ln Settler Mortality					0.09**	0.15***		-0.54***		
					(0.040)	(0.030)		(0.148)		
European Settlers in 1900							0.06***	0.06***		Removed
							(0.006)	(0.007)		(0.004)
Ln Population density in 1500									-0.54***	Removed
									(0.139)	(0.004)
Engfrac			Removed				1.95***	1.59***		
			(0.010)				(0.538)	(0.539)		
Eurfrac	2.74***	2.74***			-0.40**				2.29***	4.27***
	(0.338)	(0.338)			(0.086)				(0.593)	(0.431)
French legal origin	-1.97***	-1.97***								
	(0.297)	(0.297)								
Ln Predicted trade share Frankel-Romer (1999)			0.52***							
			(0.075)							
Ln Predicted trade share Frankel-Rose (2001)				Removed	-0.09*					
				(0.010)	(0.055)					
Malaria Ecolog					0.016*		-0.06***		-0.10***	
					(0.005)		(0.018)		(0.026)	
Average IPR index									1.07***	0.87***
									(0.208)	(0.227)
Share of population in temperate zone					-0.20					
					(0.126)					
Constant	4.07***	4.07***	2.46***	8.28***	-0.20	-0.41***	2.49***	4.79***	-15.1***	-14.8***
	(0.261)	(0.261)	(0.198)	(0.477)	(0.245)	(0.124)	(0.203)	(0.763)	(0.594)	(0.737)
n	65	65	65	65	65	65	55	55	55	55
R2	0.620	0.620	0.422	0.602	0.726	0.768	0.861	0.865	0.775	0.742
Root MSE	1.131	1.131	0.382	0.317	0.236	0.217	0.936	0.922	1.262	1.353
Fixed factors	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES

*** - 1% significance, ** - 5% significance, * - 10% significance. Standard errors in parentheses. "Removed" denotes that variable has been removed despite its statistical significance with implied probability values reported in square brackets.

schooling levels one could hope for. Even settler mortality across all the possible experiments has more power to predict current levels of human capital than the quality of institutions. Taken on face value, model selection provides clear evidence that colonial history does matter, but at the deepest level not through institutional transformation, but through changing the balance of human capital endowments. In this respect, human capital rules!

A slightly more diverse story emerges for HJ´s instruments. A fraction of the population speaking one of the major five European languages does come out as a significant and strong predictor of institutional quality, though not surprisingly it does appear to have some explanatory power for malaria as well as innovative activity. However, once continent heterogeneity is accounted for, in all regressions presented, it loses its power to predict modern health conditions and pledges its strongest alliance to institutions. On the contrary, the share of population speaking English has no power to instrument for institutions directly. Rather (and perhaps not so surprisingly), according to the first-pass specific to general race (not reported) and the PcGets selection in Table 1.4, it comes out as a reasonable instrument for human capital or possibly trade. Although according to BACE it has no specific robust role at all.

Our legal instruments – French legal origin and IPR – have the best statistical properties for use as instruments of what they are expected to instrument for. Across all the possible regressions, the former is the single most important predictor of institutional quality, whereas the latter is the single most important predictor of innovative activity. In their original work (Porta et al. 1997, 1998, 1999) associated legal origins with the effectiveness of

financial system, which raises the question of whether French legal origin mediates differences in financial development. Further inspection using a selection of related variables taken from Djankov et al. (2007) - private credit to GDP ratio, logarithm of days to enforce a contract, aggregate creditors right score and average annual inflation - annuls these doubts (results not reported). In all robustness tests, legal origin would still come out as a highly significant predictor of executive constraints.

At first glance the same compliment can be paid to Frankel and Romer (1999) trade instrument that even outperforms its younger relative from Frankel and Rose (2002); once the latter is included, the model selection races are generally dominated by Frankel and Romer (1999) instrument for trade. Again, the results are not so conclusive when two market size measures as well as other fixed factors are controlled. Although PCGets, if anything, favors Frankel and Rose (2002), BACE over all variable categories in Table 1.3 column 4 cannot select either as appropriate instruments. In columns 5 and 6, I show that this is due to instrument multicollinearity; once one gravity-based variable is dropped, the inclusion probability for the other increases to 0.9. Therefore, although gravity-based variables are obviously sensitive to geographical controls, they do retain appropriate, though albeit qualities as instruments for trade.

Our 'preferred' geographical instrument and popular choice in the literature – share of population in temperate zone – has no clear relation to any of the deep determinants, not even to malaaria conditions. With this, the instrumentation pattern as well as channels of influence suggested by model selection are as follows:

Institutions – French legal origin, share of population speaking one of the five major European languages.

Openness or Integration – predicted trade share from gravity models.

Malaria – malaria ecology, possibly settler mortality.

Human Capital – share of European settlers in 1900, fraction of population speaking English.

Innovative activity – index of patent rights.

6. Institutions vs. Human Capital

The previous section has outlined a very clear pattern for instruments routinely used in the literature of comparative development. The results from model selection revealed that the finaal outcome of the colonization quasi-experiment is in determining the human capital endowments, a direct contradiction to the institutional prevalence hypothesis of Acemoglu et al. (2001). Their most popular instrument – settler mortality – is exclusively related to modern health conditions. In this section I investigate further what it was that European settlers brought with them to their colonies: institutions or human capital? This is done in two steps. First, I revisit Acemoglu et al. (2001) robustness tests in light of instrument selection as well as accounting for colonization patterns. Second, consistent with the theory of human capital prevalence, I rerun the same specifications using settler mortality and the share of European settlers in 1900 as instruments for human capital. In sum, the results are twofold: settler mortality should not be used as an instrument of any kind (only if for malaria transmission risk) and the theory of colonial origins is in better concordance with human capital than with institutional prevalence hypothesis.

6.1. Acemoglu et al. (2001) Identification

The main concerns with Acemoglu et al. (2001) identification come from three important observations. First, as was cautiously pointed out by the authors themselves, the exclusion restriction (4) would fail if mortality rates were correlated with current health conditions that in turn would have a direct impact on income levels. Modern disease environment can be argued to affect institutional quality, in which case omitting it from the OLS regression would not lead to a biased estimate of the impact of institutions on economic performance. However, omitting it from the IV regression with AJR´s instrument can be very problematic. The second concern follows Olsson (2004) who claims that the two waves of colonization integrated in the single AJR instrument are far too different to be treated as one bundled quasi-experiment. Since different waves ‹cover› different continents, Olsson notes that the disaggregation of the sample into different continental groups considerably weakens the link between mortality rates and institutions. The third concern is that the relationship between settler mortality and underlying measures of institutions may be simply driven by four ‹neoeuropeans›.[5] Furthermore, it should be noted that there is nothing substitutable in the economic reasoning behind these three arguments. In other words, the intuition that allows us to control for one set of these controls does not preclude the inclusion of the others.

Table 1.5 presents the evidence on identification issues with both Executive Constraints as well as average Protection Against Expropriation Risk used in the original Acemoglu et al. (2001) work to measure institutional quality. In the latter case, the dependent variable is GDP per capita in 1995, also taken from Acemoglu et al. (2001) original database, whereas the effect of executive constraints is estimated on GDP per capita in 2000. Column 1 provides the usual baseline with strong identification. Columns 2-4 then incorporate information on the three aforementioned arguments, controlling for malaria, regional heterogeneity as well as for four neoeuropeans, respectively. Identification still comes through in the sense that IV estimation yields a significant positive coefficient on institutions in the second stage that Acemoglu et al. (2001) interpret as giving confidence to their main result. However, what does not come through from their exposition is that in a blink of an eye their instrument loses almost all of its strength. Especially with malaria and continent dummies controlled for, Shea (1997) partial correlation between the measure of institutions and its instrument dropping 3-4 fold below 10 per cent. Coefficient values on settler mortality fall in absolute value, indicative of omitted variable bias.

[5] A look at the scatterplot between a measure of institutional quality and settler mortality that Acemoglu et al. (2001) provide on page 1384 is convincing enough.

Table 1.5. Acemoglu et al. (2001) Robustness Tests

Panle A: IV results for log GDP per capita PPP

	1		2		3		4		5		6		7		8	
Average protection against expr. Risk (1985-1995)	1.01*** (0.198)		0.66** (0.260)		1.02*** (0.354)		1.41*** (0.463)		1.20 (1.35)		1.58 (3.38)		1.56* (0.854)		-1.21 (3.332)	
Executive Constraints (1960-2000)		0.76*** (0.138)		0.53** (0.220)		1.06* (0.596)		0.87*** (0.262)		0.62 (0.444)		0.53 (0.799)		1.32 (1.22)		0.61 (1.22)
Malaria falciparum			-0.76* (0.447)	-0.72 (0.529)					0.32 (2.13)	-0.91 (0.581)	0.25 (3.94)	-0.71 (1.40)			-3.20 (4.82)	-0.92 (1.09)
Africa					-0.51* (0.305)	0.70 (1.321)			-0.58 (0.652)	0.47 (0.803)			-0.35 (0.525)	1.06 (2.30)	0.46 (1.71)	0.46 (1.54)
Asia					-0.77* (0.401)	0.04 (0.797)			-0.92 (1.32)	0.07 (0.541)			-1.34 (0.835)	-0.002 (0.930)	1.57 (3.78)	0.07 (0.519)
Other_AJR					-1.00 (0.969)	Partialled out			-2.14 (3.76)	Partialled out						
Neoeurope							-2.74* (1.532)	-0.89 (0.977)			-3.21 (9.79)	-0.02 (2.19)	-3.63 (2.76)	-1.95 (3.23)	5.24 (10.61)	-0.17 (3.13)

Panel B: First-Stage Regressions for Institutions

log Settler Mortality	-0.61*** (0.150)	-0.80*** (0.156)	-0.43* (0.218)	-0.49** (0.228)	-0.43** (0.192)	-0.40 (0.253)	-0.39** (0.150)	-0.60*** (0.170)	-0.19 (0.268)	-0.32 (0.265)	-0.09 (0.222)	-0.17 (0.224)	-0.25 (0.158)	-0.27 (0.274)	0.09 (0.226)	-0.13 (0.261)
Malaria falciparum			-0.70 (0.522)	-1.19** (0.580)					-1.20* (0.637)	-0.41 (0.788)	-1.01** (0.484)	-1.47*** (0.547)			-1.59*** (0.533)	-0.67 (0.734)
Africa					-0.27 (0.353)	-1.46*** (0.544)			0.29 (0.394)	-1.27* (0.737)			-0.28 (0.342)	-1.44** (0.565)	0.50 (0.319)	-1.14 (0.740)
Asia					0.33 (0.486)	-0.52 (0.620)			0.71 (0.590)	-0.41 (0.676)			0.64 (0.464)	-0.30 (0.643)	1.20** (0.550)	-0.09 (0.679)
Other_AJR					1.24 (0.773)	Partialled out			2.36*** (0.691)	Partialled out						
Neoeurope							2.48*** (0.396)	2.05*** (0.512)			2.75*** (0.450)	2.40*** (0.492)	2.78*** (0.387)	2.13*** (0.607)	3.35*** (0.471)	2.33*** (0.565)
n	64	61	62	61	64	61	64	61	62	61	62	61	64	61	62	61
R2	0.27	0.32	0.29	0.36	0.30	0.36	0.40	0.38	0.36	0.37	0.45	0.45	0.43	0.48	0.51	0.49
F-stat for the 2nd stage	38.15	29.55	44.72	28.30	8.68	6.45	36.66	75.59	10.48	11.58	19.51	80.82	16.35	22.54	9.91	42.46

C: Tests for weak instruments

F-test for excluded instruments	22.95	26.19	3.85	4.54	5.09	2.47	6.67	12.32	0.49	1.45	0.16	0.60	2.44	0.94	0.17	0.24
(p-value)	(0.000)	(0.000)	(0.054)	(0.037)	(0.028)	(0.122)	(0.012)	(0.001)	(0.488)	(0.233)	(0.687)	(0.441)	(0.124)	(0.337)	(0.681)	(0.626)
Shea partial R2	0.270	0.316	0.080	0.080	0.096	0.073	0.127	0.178	0.014	0.032	0.004	0.010	0.039	0.033	0.004	0.005
Anderson LR statistic	20.15	23.20	5.14	5.05	6.43	4.61	8.71	11.93	0.857	1.99	0.25	0.61	2.54	2.06	0.25	0.32
(p-value)	(0.000)	(0.000)	(0.023)	(0.025)	(0.011)	(0.032)	(0.003)	(0.001)	(0.355)	(0.158)	(0.621)	(0.435)	(0.111)	(0.151)	(0.618)	(0.573)
Cragg-Donald F-statistic	22.95	27.30	5.10	5.00	6.23	4.55	8.89	12.52	0.78	1.89	0.23	0.57	2.39	1.92	0.23	0.29
(critical value: 10% IV size)	(16.38)	(16.38)	(16.38)	(16.38)	(16.38)	(16.38)	(16.38)	(16.38)	(5.53)	(5.53)	(5.53)	(5.53)	(5.53)	(5.53)	(5.53)	(5.53)
Anderson-Rubin F-stat	66.14	60.41	6.64	5.59	14.06	13.30	40.62	33.68	2.85	2.48	1.42	0.71	11.49	8.88	0.94	0.54
(p-value)	(0.000)	(0.000)	(0.013)	(0.021)	(0.000)	(0.001)	(0.000)	(0.000)	(0.097)	(0.121)	(0.239)	(0.404)	(0.001)	(0.004)	(0.336)	(0.464)

*** - 1% significance, ** - 5% significance, * - 10% significance. Heteroskedacity corrected standard errors in parentheses.

Furthermore, it is very illustrative to note that under Executive Constraints all three controls work through institutions at the first stage regressions, whereas under Protection Against Expropriation Risk they have a direct effect on income levels at the second stage instead. This is a direct result of the properties of institutional measures discussed in section (3.2) in the context of Table 1.2. Protection Against Expropriation Risk has the strongest partial correlations with malaria and continent heterogeneity that are encapsulated in settler mortality. When controlled one-by-one as in columns 2-4, settler mortality absorbs the effect at the first stage, when other controls appear insignificant. Executive Constraints, however, have the strongest partial correlations with legal and linguistic variables that are omitted throughout. Therefore, the geographical controls that are highly correlated with the omitted true predictors, pick up their impact at the first stage.

Columns 5-8 investigate the robustness to combinations of the three main sets of factors of interest. Settler mortality never comes out as significant in the first stage and no effects of institutions can be identified. Even further, in all these specifications the partial correlation between instrument and institution turns out to be almost non-existent and the null of underidentification cannot be rejected. Crucially note that now all exogenous controls are working through the first stage due to their correlation with the true omitted instruments. However, as they are not proper instruments and, as will be evident later, have a direct impact on income levels, the exclusion restriction will be violated and the second stage becomes unidentifiable.

In short, controlling for any combination of the aforementioned factors renders the instrument invalid, and does not even have first stage predictive power for institutional quality, hence violating the non-zero condition (1.3). These findings keep holding no matter what measure of institutional quality is used, whether malaria transmissioon risk is treated as exogenous or endogenous (in which case settler mortality would be a robust predictor of malaria), or should other geographical instruments for malaria such as latitude be exploited. I have carried out the same robustness tests also for HJ´s two linguistic instruments (results not reported). Interestingly enough, malaaria risk and continent dummies have similar albeit milder implications. Nevertheless, at least one of the excluded instruments retains its significance as well as acceptable statistical properties in the first stage, and the index of executive constraints will retain positive and significant impact on levels of income, which was not the case under settler mortality.

6.2. Human Capital Prevalence

Next I shed some light on the human capital prevalence hypothesis by Lipset (1960). To avoid problems of multicollinearity arising from the IV identification with multiple endogenous variables as discussed in the context of Glaeser et al. (2004), I proceed by carrying out the Acemoglu et al. (2001) exercise using settler mortality as well as the share of European settlers in 1900 as instruments for average schooling instead of institutional quality. The results are reported in Table 1.6, which follows the structure of Table 1.5.

Table 1.6. Acemoglu et al. (2001) with Human Capital

Panel A: IV results for log GDP per capita PPP

	1		2		3		4		5		6		7		8	
Average Schooling	0.47***	0.34***	0.32***	0.19***	0.49***	0.34***	0.65***	0.47***	0.34***	0.25***	0.60**	0.06	0.67***	0.47***	0.44***	0.22***
	(0.057)	(0.030)	(0.064)	(0.034)	(0.086)	(0.037)	(0.071)	(0.057)	(0.068)	(0.024)	(0.245)	(0.172)	(0.126)	(0.080)	(0.174)	(0.087)
Malaria falciparum			-0.92***	-1.39***					-1.06***	-1.28***	-0.22	-1.73***			-0.86	-1.32***
			(0.267)	(0.233)					(0.345)	(0.232)	(0.738)	(0.571)			(0.528)	(0.278)
Africa					0.001	-0.44*			0.26	0.14			0.31	-0.18	0.33	0.12
					(0.322)	(0.227)			(0.226)	(0.207)			(0.420)	(0.289)	(0.291)	(0.225)
Asia					0.43**	0.28			0.42**	0.36			0.42*	0.31	0.41**	0.36
					(0.218)	(0.228)			(0.209)	(0.228)			(0.248)	(0.208)	(0.204)	(0.235)
Other AJR					-0.85	-0.62**			-0.72*	-0.23						
					(0.696)	(0.277)			(0.420)	(0.186)						
Neoeurope							-2.57*	-1.28***			-2.27	0.93	-2.54***	-1.28**	-1.26	0.04
							(0.541)	(0.406)			(1.48)	(0.994)	(0.778)	(0.512)	(1.04)	(0.517)

Panel B: First-Stage Regressions for Education

	1		2		3		4		5		6		7		8	
log Settler Mortality	-1.53***		-1.21***		-1.19***		-1.02***		-1.03***		-0.42**		-0.79***		-0.50**	
	(0.233)		(0.350)		(0.290)		(0.119)		(0.384)		(0.188)		(0.165)		(0.194)	
European Settlers in 1900		0.08***		0.067***		0.077***		0.065***		0.07***		0.034***		0.059***		0.044***
		(0.006)		(0.008)		(0.010)		(0.009)		(0.011)		(0.013)		(0.014)		(0.012)
Malaria falciparum			-1.36*	-1.44***					-0.63	-1.03*	-2.12***	-2.04***			-1.30*	-1.41**
			(0.701)	(0.474)					(0.944)	(0.543)	(0.543)	(0.520)			(0.688)	(0.540)
Africa					-1.44***	-0.63			-1.21*	-0.11			-1.54***	-0.96*	-0.98*	-0.40
					(0.452)	(0.481)			(0.635)	(0.574)			(0.436)	(0.548)	(0.546)	(0.588)
Asia					-1.70**	1.33*			-1.53*	1.30*			-1.10*	0.96	-0.70	0.768
					(0.671)	(0.732)			(0.760)	(0.705)			(0.576)	(0.778)	(0.586)	(0.704)
Other_AJR					1.87	0.62			3.30***	0.88						
					(1.29)	(0.799)			(1.03)	(0.848)						
Neoeurope							4.71***	1.70**			5.32***	3.50***	4.50***	2.01**	4.96***	2.90***
							(0.463)	(0.767)			(0.457)	(0.917)	(0.471)	(0.983)	(0.496)	(0.881)
n	52	51	51	51	52	51	52	51	51	51	51	51	52	51	51	51
R2	0.54	0.73	0.56	0.77	0.66	0.80	0.74	0.74	0.69	0.81	0.80	0.81	0.80	0.81	0.81	0.83
F-stat for the 2nd stage	65.07	120.37	92.52	184.10	25.84	57.99	97.33	98.38	61.80	108.61	76.78	147.62	47.89	59.62	62.72	88.64

Panel C: Tests for weak instruments

F-test for excluded instruments	43.29	194.61	12.00	71.12	16.83	54.78	74.35	53.56	7.22	39.89	4.94	7.13	22.78	18.28	6.61	12.59
(p-value)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.010)	(0.000)	(0.031)	(0.010)	(0.000)	(0.000)	(0.014)	(0.001)
Shea partial R2	0.542	0.729	0.282	0.624	0.338	0.571	0.411	0.399	0.222	0.522	0.071	0.139	0.262	0.264	0.089	0.157
Anderson LR statistic	40.59	66.51	16.87	49.84	21.41	43.15	27.54	26.00	12.80	37.64	3.76	7.66	15.77	15.62	4.68	8.69
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.053)	(0.006)	(0.000)	(0.000)	(0.030)	(0.003)
Cragg-Donald F-statistic	59.14	131.55	18.82	79.54	23.94	61.20	34.21	31.92	12.83	49.12	3.59	7.62	16.65	16.49	4.33	8.36
(critical value: 10% IV size)	(16.38)	(16.38)	(16.38)	(16.38)	(16.38)	(16.38)	(16.38)	(16.38)	(16.38)	(16.38)	(16.38)	(16.38)	(16.38)	(16.38)	(16.38)	(16.38)
Anderson-Rubin F-stat	80.29	88.41	16.95	15.88	27.97	61.88	51.77	34.34	8.49	28.50	6.09	0.10	25.13	21.17	6.25	4.21
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.006)	(0.000)	(0.017)	(0.751)	(0.000)	(0.000)	(0.016)	(0.046)

*** - 1% significance, ** - 5% significance, * - 10% significance. Heteroskedacity corrected standard errors in parentheses.

Indeed, throughout columns 1-8 no robustness check, which was so devastating for institutional theory of colonial origins, can shake instrument strength as a predictor of human capital! Although some signs of weak instrumentation in the case of settler mortality are present also here, the share of European settlers in 1900 provides strong and robust identification for human capital prevalence theory. Even settler mortality has much better properties to instrument for human capital than for institutions. Under endogenous malaria transmission risk the share of European settlers in 1900 would still provide a minimum degree of identification (results not reported), whereas HJ´s instruments have no relevance to education. Crucially, it is also evident that many of the exogenous controls – most notably malaria – now have direct income effects that were blurred in Table 1.5 due to severe identification issues.

All in all, the results indicate that once the channels of influence coming from disease environment, regional heterogeneity and neoeuropeans are netted out, colonial origins are clearly more strongly related to human capital formation than to institutional theory. Settler mortality cannot be considered as capturing any information over and above modern disease environment and fixed factors, and thus should not be used as an instrument.

6.3. Recent Evidence from Dynamic Panel Data

Some of the more recent work has shifted from cross section to dynamic panel estimates. Dollar and Kraay (2003a,b), following the dynamic growth regression of Caselli et al. (1996), regress growth of per capita GDP on lagged growth and on changes in average measures of institutions and openness. In contrast to level equations they find a substantial partial effect

of changes in trade shares, while changes in measures of institutions in many cases come out as insignificant. Glaeser et al. (2004), in a specification where change in institutions is regressed on initial schooling, claim that the latter comes out as strongly significant, whereas in a reverted regression initial levels of institutions do not have any power to predict changes in schooling. In a subsequent paper, Acemoglu et al. (2005) show that once time effects are included, education looses its power to predict changes in democracy; in fact, although insignificant, the coefficient on the initial level of average schooling always comes out as negative. Therefore these authors argue for no causal links running from education to institutional quality.

Acemoglu et al. (2005) results have been criticized mainly on two grounds. First, their use of difference GMM as well as treatment of education as strictly exogenous might not be entirely appropriate in the presence of persistent variables. Results from Castelló-Climent (2006) and Bobba and Coviello (2007) reveal that once the system GMM estimator is used and current shocks to institutions are allowed to impact future education, schooling can indeed be shown to cause democracy. The second line of criticism exploited in my companion article Kangur (2009) concerns the more general interpretation and use of measures of institutional quality in dynamic estimation. (North 1981, p. 205), in describing the evolution of institutions, states the following: «The combination [of constitutional rules with the associated moral and ethical codes] produces ingrown patterns of behaviour which, like the capital stock, tend to be changed only incrementally.»

In a similar fashion, Rodrik et al. (2002) think of the quality of institutions as a predetermined stock variable representing not only the current political choices, but also those of past political rulers. Therefore, current policies affecting the institutional quality are naturally a flow variable. One can then write down an equation of motion for institutional quality I as:

$$\Delta I = \sum \alpha_c p_c - \delta I, \tag{1.5}$$

where αc denotes the impact of policy pc on institutional quality.

In my companion paper Kangur (2009), I point out that a criticism on similar lines applies practically to all past contributions that in the dynamic specifications measure institutions with observed indices. It has been well documented that almost all available measures of institutions are highly volatile and mean-reverting outcome measures that reflect policy choices in the period under consideration. These are not properties of anything permanent and credible that should reflect the entire history of policy choices. Therefore, it is my interpretation that these are features of a policy flow or pc in equation (1.5), adding to the overall stock of institutions. This point, for example, has been clearly made by Park (2001) in the context of economic freedom and IPR indices: "Another source of confusion arises from not recognizing that indexes of economic freedom and patent rights are flows, not stocks. They reflect the value for a particular year or period, and not the entire history of their respective institutions or experiences." This distinction is important as it suggests that in dynamic regressions available indices should be used to measure first differences in institutions, while the level of institutions is actually a latent variable.[6] Such reinterpretation

[6] Strictly speaking, in the previous chapters one should then measure institutional quality as a sum over the respective index values. However, in the cross-sectional regressions, and if the depreciation rate in (1.5) is negligible, an average over the entire history is an appropriate measure for a stock.

of political indices allows us to find a positive effect of higher educational attainment on future institutional changes. In reverted specifications, no indication of reverse causality running from institutions to education is evident.

7. What Rules?

To take stock of where we stand, we have seen that colonial history first and foremost has had a large influence on the emergence of educated contingency; that the contemporary distribution of European languages as well as the roots of the legal system carry information on the quality of modern institutions; that geography, especially malaria, matters; and that there is some uncertainty over identifying the effects of trade. Once reached thus far, a natural question to ask is: what rules? To answer this questions I run the "final" IV races between all the deep determinants as well as exogenous regressors, and show that by and large the findings concur with human capital prevalence with direct income effects of geography and complementary effects of trade.

Before proceeding a note on model misspecification as one source of fragility as observed in section 3.1 is in order. In the words of Sachs (2003), the models set up in this strand of the literature are so "worryingly oversimplified" because "there is good theoretical and empirical reason to believe that the development process reflects a complex interaction of institutions, policies, and geography." As suggested for example by Carstensen and Gundlach (2006), higher integration with world markets may itself encourage better protection of property rights through higher monitoring of government actions and enforcement of contracts. Therefore, there is inevitably an element of substitutability (negative interaction) between institutions and trade that has been ignored. Furthermore, the relationship between trade share and income levels cannot be 'too linear'. In the working paper version of this article I demonstrate that accounting for such 'complex interactions' between institutions and trade allows us to find more robust and significant direct effects for both of these deep determinants that Rodrik et al. (2002) were not able to find. In addition I am able to provide evidence on the 'substitutable' roles of institutions and openness.

To run "The Final Race" now I take on board the fourth deep determinant – human capital – aiming to see that – given the instrument selection - what finally rules? I provide two pieces of evidence. First, I use BACE to identify the posterior inclusion probabilities for an already well-known selection of variables in regressions where the dependent variable is GDP per capita. Second, in the spirit of the literature, I run instrumentaal variable races to correct for endogeneity and see if the pattern obtained from BACE and in the previous section continues to hold.

Table 1.7 lists our 4 deep determinants together with the 6 categories of exogenous controls most often used in the literature, and reports their BACE posterior inclusion probabilities. The first column performs BACE only across six exogenous groups. Two categories dominate: endowment with natural resources and regional heterogeneity. Interestingly, none of the 'purely' geographical variables stand out, perhaps only with the exception of share of population living in coastal regions. This can justify their use as instruments, though as was shown in section 1.5 as well as later in Table 1.8, none of them actually qualifies as an instrument on its own.

Table 1.7. Bayesian Averaging of Classical Estimates - Posterior Inclusion Probabilities of Selected Instruments and Controls

Dependent variable is GDP PPP per capita 2000

	1	2
1 INSTITUTIONS		
Executive constraints		**0.928** (+)
2 INTEGRATION		
Trade share		0.121 (-)
3 HUMAN CAPITAL		
Average Schooling		**1.000** (+)
4 MALARIA		
Malaria Falciparum		**0.993** (-)
5 GEOGRAPHY I		
% of population in temperate zone	0.228 (+)	0.193 (+)
% of land in tropics	0.546 (-)	0.180 (-)
Frostdays	0.158 (+)	0.231 (-)
Distance from equator	0.158 (+)	0.522 (+)
6 GEOGRAPHY II		
Landlock	0.234 (+)	0.199 (-)
% of population within 100 km of coast	0.802 (+)	0.472 (+)
7 NATURAL RESOURCES		
Net exporter of oil	**0.995** (+)	**0.924** (+)
Share of natural capital	**0.999** (-)	**0.941** (-)
8 REGION		
Africa	**0.999** (-)	0.224 (-)
Asia	**0.899** (-)	0.309 (-)
Latin-America & Caribbean	0.537 (-)	0.199 (-)
9 NEOEUROPE		
Neoeurope	0.470 (+)	0.137 (-)
10 MARKET SIZE		
Population	0.513 (-)	0.131 (+)
Area	0.689 (+)	0.303 (+)

Column 2 performs BACE across model specifications that incorporate both endogenous as well as exogenous variables. These results should be regarded with some caution since posterior distributions are computed for least squares estimates that do not correct for the endogeneity bias. This first pass on the data shows the clear dominance of institutions, human capital, geography as measured by modern health conditions, and endowments with natural resources. For most of the other exogenous factors posterior inclusion probabilities are significantly lowered when compared to column 1.

Table 1.8 shows the IV races with baseline instrumentation as determined in section 1.5. I start off with institutions in column 1, adding other deep determinants one at a time. The final specification is then checked for robustness. Columns 2 and 3 demonstrate the usual problem of finding the robust effect of openness that in column 4 is achieved through interaction between institutions and openness. This is motivated by Sachs (2003) in whose words the models set up in this strand of the literature are so "worryingly oversimplified" because "there is good theoretical and empirical reason to believe that the development process reflects a complex interaction of institutions, policies, and geography." As suggested for example by Carstensen and Gundlach (2006), higher integration with world markets may itself encourage better protection of property rights through higher monitoring of government actions and enforcement of contracts. Therefore, there is inevitably an element of substitutability (negative interaction) between institutions and trade that has been ignored. Furthermore, the relationship between trade share and income levels cannot be 'too linear'. In the working paper version of this article I demonstrate that, in addition to the 'substitutable' roles of institutions and openness, accounting for such 'complex interactions' between institutions and trade allows us to find more robust and significant direct effects for both of these deep determinants that Rodrik et al. (2002) were not able to find.

Table 1.8. Final Races

Panel A: Second-stage regressions for real GDP per capita PPP in 2000

	1	2	3	4	5	6	7	8	9	10	11	12
Institutions	0.49***	0.49***	0.23***	1.97**	-1.23							
	(0.077)	(0.076)	(0.077)	(0.937)	(1.024)							
Openness		0.13	0.21	2.29**	-0.99	0.25*	0.19*	0.04	0.04	0.06	-0.19	-0.14
		(0.212)	(0.144)	(1.095)	(1.0630	(0.132)	(0.116)	(0.109)	(0.123)	(0.152)	(0.507)	(0.363)
Malaria Falciparum Risk			-1.79***	-1.97***	-1.02***	-0.84***	-1.21***	-0.63**	-0.86**	-0.79***	-0.85***	-0.72***
			(0.301)	(0.372)	(0.356)	(0.315)	(0.423)	(0.275)	(0.374)	(0.306)	(0.324)	(0.261)
Institutions * Openness				-0.46*	0.27							
				(0.240)	(0.239)							
Human Capital					0.38***	0.32***	0.19**	0.28***	0.28***	0.30***	0.31***	0.25***
					(0.104)	(0.046)	(0.088)	(0.037)	(0.072)	(0.040)	(0.045)	(0.044)
Continent dummies							[0.149]					
Natural Resources								[0.002]				[0.002]
Geography I									[0.071]			
Geography II										[0.011]		
Market size											[0.295]	[0.694]
Openness * Net exp. of oil												0.48*
												(0.247)

Panel B: Tests for weak instruments

n	94	94	94	94	94	94	94	84	91	91	93	83
F-stat	39.57	20.54	62.49	30.11	38.74	104.73	72.99	100.26	60.65	78.79	61.13	71.92
R2	0.494	0.499	0.775	0.638	0.702	0.819	0.859	0.873	0.858	0.846	0.817	0.893
Shea partial R2												
institutions	0.336	0.335	0.208	0.090	0.065							
openness		0.548	0.534	0.087	0.077	0.508	0.555	0.567	0.561	0.429	0.086	0.130
malaria			0.320	0.341	0.384	0.291	0.183	0.310	0.219	0.295	0.288	0.290
institutions*openness				0.083	0.071							
human capital					0.141	0.313	0.131	0.385	0.198	0.358	0.338	0.274
Anderson LR statistic	38.45	38.12	21.13	7.56	5.76	28.63	11.35	27.96	17.07	30.00	8.08	9.84
(p-value)	(0.000)	(0.000)	(0.000)	(0.056)	(0.124)	(0.000)	(0.003)	(0.000)	(0.000)	(0.000)	(0.018)	(0.007)
Cragg-Donald F-statistic	23.00	15.01	5.61	1.22	0.78	10.68	2.76	7.60	4.23	8.20	1.95	1.84
(critical value: 20% IV size)	8.75	6.40	-	-	-	-	-	-	-	-	-	-
Sargan OI test	3.59	4.41	0.00	1.66	0.65	-	2.27	2.01	2.05	4.94	2.30	2.25
(p-value)	(0.058)	(0.036)	(0.947)	(0.437)	(0.722)	-	(0.132)	(0.157)	(0.152)	(0.026)	(0.130)	(0.133)

*** - 1% significance, ** - 5% significance, * - 10% significance. Heteroskedacity corrected standard errors in parentheses. Probability values for a Wald test of joint significance are given in square brackets.

The inclusion of human capital in column 5 seems to account for most of the information incorporated in institutions and integration. At this stage a model selection is performed on endogenous deep determinants. As was evident from the discussion in the context of Glaeser et al. (2004) race, the models incorporating both institutions and human capital are rank deficient, suggesting that only one of these variables belongs to the true model, and therefore, is consistent with the evidence presented in previous sections, allowing us to delete institutions from the regression equations. This is supported by the PcGets general-to-specific modelling routine that always selects human capital over institutions. Openness only stays moderately significant, the coefficient on malaria transmission risk more than halves the suggestion that educational attainment accounts for a large part of health conditions as well.

Column 6 forms the baseline specification that is further tested in columns 7-11 by controlling for five already familiar exogenous categories. In all these robustness tests, I add the share of population speaking English to the instrument set to enhance identification. The results suggest that natural resource endowments as well as both geographical categories do have a direct impact on income levels. This contradicts findings using BACE in the previous table, where high correlations between variables can reduce the precision of the estimates. The coefficient on trade is now rather imprecisely measured, revealing that natural resources and

geography accounts also for a large part of the extent countries integrated into world markets. In search of the effects of openness, the final column interacts trade with a dummy for net exporter of oil, controlling for the direct effects of natural resources as well as of market size measures. The results suggest that, although openness does not have a direct impact on GDP, oil rich countries do enjoy higher income effects from being more integrated into world markets.

In sum, the evidence presented here fits well with bits and pieces given in previous sections. The broad conclusions are: (i) human capital and malaria rule! (ii) geography has a direct effect also through many other channels, most notably endowments with natural resources, but seemingly also continent heterogeneity as well as other factors matter, (iii) the most important impact of colonial history is in forming highly educated constituencies, (iv) institutions matter though not those of first-order, (v) identification of the direct effects of trade openness is complicated, though there do exist complementary channels, (vi) the choice of instruments as well as institutional measures is absolutely crucial for any of the aforementioned conclusions. The identification of trade suffers from at least three factors. First, as a broad sample as possible is needed. Second, gravity-based instruments are not perfect due to their correlation with geography, though seem to be acceptable. Third, 'complex interactions' seem to be especially relevant to trade which, however, seems to be one feature where cross-sectional IV strategy is especially limited.

8. Conclusion

When first running through the regressions for all four main variables researchers have used in determining the income levels, I was rather surprised by the amount of detail people did and did not talk about as well as the enormous sensitivity of conclusions to the particular choice of the instruments and the instrumented. This was what motivated this paper in the first place. I see the main contribution provided here as an attempt to bring some order to this strand of research.

How much guidance could policy makers take away from the outcomes documented here? Contrary to Rodrik et al. (2002), I would say quite a lot. It might be stressing to learn that geography comes close to being a destiny, trade by itself does not lead to automatic convergence, or, to some extent, differences in the level of development were shaped already decades or even centuries ago. But it is good to know the channels behind these outcomes, because there are lessons to be drawn from these.

For example, one would be tempted to argue that colonization has made some poor. There might be some truth behind this, but the same colonization experience that might have not been beneficial to some, also gives them a lesson about why they might be lagging behind. Similarly, the results here would suggest that simply copying institutions (say, in the form of a judicial system) is unlikely to yield convergence. Institutions might have some independent exogenous role, but to a larger extent their effect is of second order through human capital. The growth-enhancing effect of institutions will not materialize without a supportive educated constituency. Put differently, there is an unfortunate reason why many dictators have intentionally dismantled the educated constituencies, or intervened heavily in the educational system. Lipset (1960), whose views are supported in this paper, writes:

> Dewey has suggested that the character of the educational system will influence its effect on democracy... The purpose of German education, according to Dewey, writing in 1916, was one of "disciplinary training rather than personal development." The main aim was

to produce "absorption of the aims and meaning of existing institutions," and "thoroughgoing subordination" to them.

Geography could be destiny, but working on their channels of influence like the disease environment might help. More trade might not lead to automatic convergence, but targeting trade policies to complementary factors of comparative development could. Any conclusions or policy recommendations here might seem rather general: certainly a more pragmatic policy maker would expect suggestions on a certain kind of education, type of property rights or a specific direction for a trade policy. But it is the whole research setting of this strand of comparative development that is bounded by such a level of abstraction. How detailed policy suggestions would you expect from sometimes only 60 observations with variables averaging over decades...

A final word of caution is related to instruments and whether they are suggestive of certain theories or not. Again, Rodrik et al. (2002) warn that an instrumentation strategy should not be confused with a direct test of certain theories. As they suggest, if AJR´s fit of first-stage regressions is approximately 25 per cent, there is a lot of room for other factors than colonization. This is yet another reason for more careful instrument selection as was carried out here. Indeed, while rerunning AJR´s specifications in Table 1.5, first stage regressions that still identified institutions explained about one-third of the variation. In Table 1.6, the same setup implied that colonial history can explain two-thirds or up to 80 per cent of the variation in the first-stage regression for schooling. Obviously, the instruments do not make a theory, but at such levels of variation they can certainly imply through which channels a certain theory works.

References

Acemoglu, D., Johnson, S. and Robinson, J.A. 2001. The colonial origins of comparative development: An empirical investigation. *The American Economic Review*, Vol. 91, No. 5, pp. 1369– 1401.

Acemoglu, D., Johnson, S. and Robinson, J.A. 2002. Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *The Quarterly Journal of Economics*, Vol. 117, No. 4, pp. 1231–1294.

Acemoglu, D., Johnson, S., Robinson, J.A. and P. Yared. 2005. From education to democracy? *The American Economic Review*, Vol. 95, No. 2, pp. 44–49.

Anderson, T. 1984. *Introduction to Multivariate Statistical Analysis* (2nd ed.). New York: John Wiley & Sons.

Barro, R. J. and Lee, J-W. 2000. International data on educational attainment updates and implications. *NBER Working Paper* No. 7911, p. 36.

Bloom, D. E. and Sachs, J. D. 1998. Geography, demography, and economic growth in Africa. *Brookings Papers on Economic Activity,* Vol. 1998, No. 2, pp. 207–295.

Bobba, M. and Coviello, D. 2007. States and markets: The advantage of an early start. *Economics Letters,* Vol. 96, No. 3, pp. 301–306.

Bockstette, V., Chanda, A. and Putterman, L. 2002. States and markets: The advantage of an early start. *Journal of Economic Growth,* Vol. 7, No. 4, pp. 347–369.

Carstensen, K. and Gundlach, E. 2006. The primacy of institutions reconsidered: Direct income effects of malaria prevalence. *The World Bank Economic Review*, Vol. 20, No. 3, pp. 309–339.

Caselli, F., Esquivel, G, and Lefort, F. 1996. Reopening the convergence debate: A new look at cross-country growth empirics. *Journal of Economic Growth*, Vol. 1, No. 3, pp. 363–389.

Castelló-Climent, A. 2006. On the distribution of education and democracy. *Institute of International Economics Working Paper No. 602* , pp. 1131–1175.

Cragg, J. G. and Donald, S. G. 1993. Testing identifiability and specification in instrumental variable models. *Econometric Theory*, Vol. 9, No. 2, pp. 222–240.

Diamond, J. M. 1997. *Guns, Germs, and Steel: The Fates of Human Societies*. Norton, W.W. & Company, Inc.

Djankov, S., Glaeser, E, Porta, R. L., de Silanes, F. L., and Shleifer, A. 2003. The new comparative economics. *Journal of Comaparative Economics*, Vol. 31, No. 4, pp. 595–619.

Djankov, S., McLiesh, C. and Shleifer, A. 2007. Private credit in 129 countries. *Journal of Financial Economics*, Vol. 84, No. 2, pp. 299–329.

Dollar, D. and Kraay, A. 2003a. Institutions, trade, and growth. *Journal of Monetary Economics,* Vol. 50, No. 1, pp. 133–162.

Dollar, D. and Kraay, A. 2003b. Institutions, trade, and growth: Revisiting the evidence. *The World Bank*, 29.

Easterly, W. 2006. Inequality does cause underdevelopment: insights from a new instrument. Mimeo. *Forthcoming in Journal of Development Economics*, 40.

Easterly, W. and Levine, R. 1997. Africa's growth tragedy: Policies and ethnic divisions. *The Quarterly Journal of Economics,* Vol. 112, No. 4, pp. 1203–1250.

Easterly, W. and Levine, R. 2003. Tropics, germs, and crops: how endowments influence economic development. *Journal of Monetary Economics*, Vol. 50, No. 1, pp. 3–39.

Engerman, S. L. and Sokoloff, K. L. 1997. Factor endowments, institutions, and differential paths of growth among new world economies. In S. Haber (Ed.), *How Latin America Fell Behind: Essays on the Economic Histories of Brazil and Mexico, 1800 - 1914*. Stanford: Stanford University Press.

Engerman, S. L. and Sokoloff, K. L. 2002. Factor endowments, inequality, and paths of development among new world economies. *Economia,* Vol. 3, No. 1, pp. 41–109.

Engerman, S. L. and Sokoloff, K. L. 2005. The evolution of suffrage institutions in the new world. *The Journal of Economic History,* Vol. 65, No. 4, pp. 891–921.

Frankel, J. and Rose, A. 2002. An estimate of the effect of common currencies on trade and income. *Quarterly Journal of Economics*, Vol. 117, No. 2, pp. 437–466.

Frankel, J. A. and Romer, D. 1999. Does trade cause growth? *The American Economic Review,* Vol. 89, No. (3), 379–399.

Ginarte, J. C. and Park, W. G. 1997. Determinants of patent rights: A cross-national study. *Research Policy* 26 (3), 283–301.

Glaeser, E. L., Ponzetto, G. A. and Shleifer, A. 2007. Why does democracy need education? *Journal of Economic Growth, Vol.* 12, No. 2, pp. 77–99.

Glaeser, E. L., Porta, R. L, de Silanes, F. L. and Shleifer, A. 2004. Do institutions cause growth? *NBER Working Paper No.* 10568, 53.

Hall, R. E. and Jones, C. I. 1999. Why do some countries produce so much more output per worker than others? *The Quarterly Journal of Economics,* Vol. 114, No. 1, pp. 83–116.

Hendry, D. F. 1995. *Dynamic Econometrics*. Oxford University Press.

Hendry, D. F. and Krolzig, H.-M. 2001. *Automatic Econometric Model Selection*. London: Timberlake Consultants Press.