#2013-056

Economic development, growth, institutions and geography
Samyukta Bhupatiraju & Bart Verspagen

AFD-MGSoG/UNU-Merit Working Paper Series on
« Institutions, Governance and Long term Growth »

In 2010, the French Development Agency (AFD) initiated a partnership with the Maastricht Graduate School of Governance (Maastricht University - UNU-Merit) with a view to exploring the conceptual and econometric relationships between institutions and long-term growth. As a development bank with a long-term lending horizon, AFD is particularly interested in better understanding the determinants of countries' long term economic, social, and political trajectory.

AFD has thus developed a programme on "Institutions, Governance, and Long-term Growth" dealing with the five following dimensions:
   (i)   Measuring institutions and discussing the meaning of such measures, notably through the Institutional Profiles Database;
   (ii)   Testing the econometric relationship between institutional measures and long term growth;
   (iii)   Exploring through a series of country case studies the historical relationship between processes of economic accumulation, forms of political organisation, and social cohesion;
   (iv)   Discussing conceptual frameworks for making sense of the interaction between political, social and economic forces in the process of development;
   (v)   Developing methodologies for political economy analyses.

The MGSoG/UNU-Merit team is involved in the five dimensions with a particular focus on the first two. Its primary objective is to explore the Institutional Profiles Database jointly developed by AFD and the French Ministry of the Economy since 2001. Institutional Profiles Database is unique by its scope (about 350 elementary questions pertaining to all institutional dimensions covering 148 countries in 2012), its entirely free access, and its ambition to incorporate the most recent theoretical advances in the field of political economy.

The present series intends to convey the results of our ongoing research, and in so doing to reflect the wealth of issues that can be fruitfully addressed from an "institutionalist" perspective. We hope that readers will find these papers stimulating and useful to develop their own understanding and research.

Nicolas Meisel (AFD)
Adam Szirmai (MGSoG/UNU-Merit)

For more information on the programme, please visit our websites:
http://www.maastrichtuniversity.nl/web/Schools/MGSoG/ProjectPages/InstitutionalProfilesDatabase.htm
http://www.afd.fr/lang/en/home/recherche/themes-recherches/institutions-gouvernance-croissance

Economic Development, Growth, Institutions and Geography

Samyukta Bhupatiraju & Bart Verspagen

bhupatiraju@merit.unu.edu

verspagen@merit.unu.edu

UNU-MERIT and Maastricht Graduate School of Governance

Abstract

In this paper, we test the Rodrik et al (2004) framework to explain differences in development levels across countries by using a broader set of definitions for institutions, geography and economic variables. We use a multi-faceted database to measure institutions in an attempt to go beyond the single-dimension measures that are often employed. We find that institutions trump other factors (geography and trade) when we use GDP per capita as an independent variable. When we expand the dependent variable to include other aspects of development, such as growth and investment, we find that institutions, growth and geography are all important variables. In this case, institutions no longer trump the other factors. In this case, we also find that the same institutions variable that was positively associated to GDP per capita is now negatively correlated with the more dynamic development variable.

Keywords: institutions, geography, openness, governance, economic development

JEL codes: O1, O16, O17

1. Introduction

Much of the empirical growth literature has focused on examining the causes of differences in growth across countries. Many theories have been put forth in this regard and in the recent decades interest in explaining and testing these differences has surged. One answer that has (re)gained popularity is that the "institutional setting for technological change"[1] determines growth in a country. In this context, the "institutional setting" (or, institutions) refers to formal rules (constitutions, laws and regulations, political systems, etc.) and informal rules (value systems, beliefs, social norms, etc.) that humans use when interacting within a wide variety of repetitive and structured situations at multiple levels of analysis (e.g., North, 2005; Ostrom, 2005). Differences in the types and quality of institutions ultimately result in different growth patterns (Acemoglu et al, 2001, 2002, 2004; Keefer and Knack, 1997; Rodrik et al, 2004; Easterly and Levine, 2003; and Bardhan, 2005).

A second influential and not so new theory, which draws on the works of Montesquieu and Weber[2] among others, states that these differences are due to an exogenously given set of geographic factors which are conducive for growth. Thus factors such as soil quality, weather conditions, and natural resource availability affect productivity and attitudes directly (Diamond, 1997; Gallup et al ,1998; Sachs, 2001; McArthur and Sachs, 2001; Masters and McMillan, 2001). A third account sees differences in economic growth arising due to a country's openness to trade, rather than as a result of its resource endowments or institutions. This stream of literature suggests that countries that follow open trade policies gain more in terms of growth performance (Frankel and Romer, 1999; Dollar and Kraay, 2001; Sachs and Warner, 1995).

These three approaches have been identified in the literature as the "deep determinants" of growth, as opposed to more "proximate determinants" such as capital resource accumulation and technological change (Rodrik et al, 2004). For most part, these three strands of literature have evolved independent of each other. More recently, studies have attempted to determine which of these three factors are most important (or "primary") in accounting for economic growth and its difference among countries (Rodrik et al, 2004; Basu, 2008; Easterly and Levine, 2003).

In this paper, we attempt to re-examine the impact of institutions, trade and geography on economic growth and development using the Rodrik et al. (2004) framework. Our contribution is to use more comprehensive measures for economic growth, for institutions, and for geography, along with the more commonly used measure of trade openness. Economic growth (or development), institutions and geography are all multi-faceted phenomena, and to capture them in a single indicator, as is often done in the empirical literature that we referred to above, is problematic. For example, economic growth is often captured by the level of GDP per capita, rather than the growth rate of that variable, and other aspects of growth, such as investment or structural change, remain poorly represented. Institutions are often captured by a single indicator,

---

[1] Darwin, J., (2008) "After Tamerlane", Bloomsbury Press. Pg 190-191
[2] Montesquieu C de S, (1748), "The Spirit of the Laws", Paris

such as a measure for "rule of law". But we know that the institutional setting has many more dimensions, and works on the factors that cause economic growth in many different subtle ways. The same holds for geography. This is often captured by a single variable such as distance from the equator, but there are obviously many more and different aspects to geography. Our aim is to capture these multi-faceted phenomena in a more comprehensive way, and see whether in this way, we can confirm the causal links between them that have been found in the existing literature.

The rest of the paper is organized as follows: Section 2 briefly reviews related literature outlining the inter-relationship between institutions, trade and geography. Section 3 describes the econometric methodology and Section 4 describes the data used. Section 5 presents the results and Section 6 concludes the paper.

2. Related Literature

2.1 "Institutions matter"

The quality of institutions as a determinant of economic growth and development is not a new theory. The revival of interest in this area of research is often attributed to Douglass North, whose definition of "institutions" is often the benchmark for what constitutes them.[3] Building on the theoretical literature, there have been numerous empirical studies using institutions as an explanatory variable in the past two decades. Among the many reasons that have boosted interest in this area of research is the development of various econometric methods that have enabled us to deal with problems such as endogeneity, reverse causality etc., which are often encountered in economic research[4], namely the 2SLS-IV methods and those that followed it.

One such influential paper that highlights the importance of institutions in studying economic growth, while correcting for endogeneity problems, is Acemoglu et al (2001). They make a case for differences in economic institutions as the "fundamental cause of different patterns of economic growth". They introduce 'settler mortality'[5] as an instrument variable to deal with the issue of endogeneity. Their results showed that institutions in fact had a large impact on economic growth, while geography did not matter (at least not in a direct way). An important critique of this paper is that the instrument they used is not universally applicable, because not all poor countries were colonized.

Hall and Jones (1999) use the percentage of population speaking English (ENGFRAC), and percentage of population speaking a European language (EURFRAC) as instruments for

---

[3] There have been many different propositions on measuring institutions. Among the more common ones include measures of corruption and democracy by various organizations.
[4] Vieira, F. V., and Damasceno, A. O. (2009).
[5] Settler mortality uses data of mortality rates of soldiers, sailors, priests/bishops (European-settlers) in their colonies per 1000 settlers.

institutions which are in turn measured by the ICRG survey indicators. They use distance from the equator to capture geographic characteristics. They conclude that differences in social infrastructure had a large impact on cross-country differences in income per capita. Other studies which use some measure of 'Western European influence' as an instrument for institutions include La Porta et al (1997, 1998, 1999) Mauro (1995). La Porta et al use legal origin as a proxy for institutions and find that common law countries perform best and civil law countries perform worst. Mauro uses ethno-linguistic divisions as a proxy for corruption and finds that corruption lowers investment and thereby growth in countries.

Easterly and Levine (2003) estimate a 2SLS-IV model using settler mortality as an instrument for institutions and distance from the equator, other geographic dummy variables such as landlockedness, crops, minerals etc to capture geography. They find that geography has an indirect effect on income differences via institutions. In their words institutions "seem to be a sufficient statistic" to account for economic growth. They do not find any direct impact of trade openness either.

Concluding, we can summarize that institutions are seen as an important determinant of economic growth, but also that they are themselves endogenous. Most of the instruments that have been proposed for institutions are somehow related to geography, which points to the fact that geography and institutions are themselves related. This means that one must indeed revert to econometric methods such as instrumental variables in order to separate the impact of these two long-run determinants of economic growth.

2.2 "Geography and Economic development"

The view that the geography of a country determines its economic prosperity dates back at least to Montesquieu.[6] There are two key versions of this hypothesis. The first states that productivity of technologies of the temperate climate zones are higher compared to technologies of the tropics.[7] The second view states that the disease environment is a burden to the country, and therefore is directly linked to the poverty that is prevailing in it. This view states that due to the highly disease prone climate in a country, productivity of workers in that economy will be low. For example, the occurrence of frost kills pests and other disease causing insects or bacteria. Many tropical countries do not have many days of frost in a year. Also, the temperatures soar high enough to not allow workers to work in agricultural fields in some seasons. All this affects growth in an economy. This outlook has become popular over the past decade and has been treated in empirical studies such as Sachs (2001) and Gallup and Sachs (2001).

While this view seems more plausible than the previous one, the causality of poverty in the wake of the disease environment and economic development is not very clear. McArthur and Sachs

---

[6] Montesquieu, C de S., "The Spirit of the Laws".
[7] Myrdal, Gunnar (1968) Asian Drama – An Inquiry into the Poverty of Nations. Vol.3 (Harmondsworth, Penguin).

(2001) conclude that institutions and geographically-related variables affect the GNP per capita of countries. However, if this were the main cause of growth or poverty in countries, then there would be no escape from being poor, and this contradicts the fact that some countries with "bad geography" have in fact developed to high income levels. Accepting that geography does have an impact through other endogenous factors could be more realistic.

Like in the case of institutions, the empirical literature has examined the importance of geography as a determinant of growth. Some of the variables used to study this hypothesis include the latitude of the country (Sala-i-Martin, 1997a,b; Bloom and Sachs, 1998; Rodrik et al, 2004), and this usually results in a positive and significant impact of the latitudinal position of a country on its growth levels. Using disease ecology as the independent variable measured by proportion of population at risk of malaria transmission, life expectancy at birth, and infant mortality rate, a negative and significant relationship between a geographically based health measure and GNP per capita is also found (Easterly and Levine, 2003; and Sachs, 2003). The number of frost days in a year, availability of arable land, the length of the coastline, rainfall, and temperature are other variables that have been used to study the effect of geography of the country on its growth rates (Bloom and Sachs, 2003; Masters and Sachs, 2001; and Bloom et al, 2003). Here, a positive and significant effect of the first three variables on growth is found, while the last two variables are found to have a negative and significant effect. Landlocked economies have been found to have lower growth (Easterly and Levine, 2001; Sachs and Gallup, 1999).

2.3 "Trade promotes growth"

International trade openness has often been cited as increasing productivity, more so in less developed countries, allowing them to catch-up or converge towards the income level of richer nations. Policy recommendations usually involve adoption of more open or liberal trade policies in order to boost growth. However, like institutions, openness may be an endogenous variable, and hence a proper test for the impact of openness on growth must account for this.

Frankel and Romer (1999) use a gravity model to propose a new instrument for overall trade share of countries as a function of geographical distance from its partners. They use this instrument in a 2SLS-IV estimation to test the relationship between trade and income level of a country. They find that overall trade of a country has a large and positive impact on income, but their result is "only moderately significant".

Sachs and Warner (1995) find that those countries that have robust trade policies converge, but add that this is not a sufficient condition for growth, as overall macroeconomic policies should accompany the trade policies. Rodriguez and Rodrik (2001) find alternative evidence, however, claiming that there is no significant relationship between open trade policies and economic growth. They further question the methodology used in Sachs and Warner (1995) and Edwards (1998), among others claiming that the measure of openness used tends to give a biased positive and significant relationship, and furthermore is open to different interpretations.

Dollar and Kraay (2001) test and find "that trade has a strong positive effect on growth". However, Easterly and Levine (2003) do not find any direct impact of trade on growth once they control for institutions. Rodrik et al (2004) come to a similar conclusion regarding the impact of trade on growth.

From the empirical evidence and counter-evidence provided, it is clear that we do not have a clear winner in the "primacy" contest. Institutions, geography and trade openness all have their own supporters as the primal factor causing growth. Most of the empirical studies however, use a uni-dimensional measure of GDP per capita or growth rates to test for the above mentioned relationships. In this paper we test the hypothesis using a different measure of growth to see if the significance of institutions, trade or geography still holds.

3. Econometric Approach

Rodrik, et al (2004) set out to estimate a regression model that encompasses all three "primal" factors behind economic growth, as we briefly surveyed them in the previous section. By including all three factors in one model, they aim to determine which of the three factors dominates over the other. Following their framework, we start with the following equation:

$$\log y_i = \alpha + \beta_1 \, \text{INS}_i + \beta_2 \text{OPENC}_i + \beta_3 \text{GEOSOIL}_i + \varepsilon_i \qquad (1)$$

where $\log y_i$ is the log of GDP per capita (in the year 2009, in international dollars) of country i, and INS_i, INT_i and GEOSOIL_i are the measures of institutional quality, trade share and soil quality conditions of the country i, respectively, and ε_i is the usual random error term.

Our brief survey of the literature has already emphasized that institutions and trade openness are both likely to be endogenous variables, and hence we need to account for this in the estimations. Cross-section growth regressions have often used the two stage least squares, or instrumental variables (2SLS-IV) estimation technique to study differences in income across countries. This method is one way of dealing with endogeneity. Rodrik et al (2004) also use the 2SLS-IV approach with instrumental variables (IV) for institutions and trade.

This approach amounts to first estimate the following regression equations:

$$\text{INS}_i = \mu + \beta_5 \text{SM}_i + \beta_6 \, \text{FR}_i + \beta_7 \, \text{GEOAREA}_i + \beta_8 \, \text{GEOSOIL}_i + \varepsilon_{\text{INS}i} \qquad (2)$$

$$\text{OPENC}_i = \gamma + \beta_9 \text{SM}_i + \beta_{10} \, \text{FR}_i + \beta_{11} \, \text{GEOAREA}_i + \beta_{12} \, \text{GEOSOIL}_i + \varepsilon_{\text{OPENC}i} \qquad (3)$$

where SM_i is settler mortality and FR_i is the constructed Frankel Romer instrument for trade share below. GEOAREA is a geographical variable that is related to the physical characteristics of countries. This variable will be presented and discussed below. Equation (2) estimates the institutional quality variable as a function of settler mortality, as well as the specific geography variables in our equation (1), and the Frankel-Romer instrument for trade. Equation (2) does the

same for the trade openness. The predicted values from equations (2) and (3) can then be used instead of the INS and INT variables (respectively) in the estimation of equation (1).[8]

As already briefly mentioned above, SM is available only for a limited group of countries (those that were colonized). Therefore, instead of using SM as an instrument, we also use two language variables, ENGFRAC (fraction of population speaking English) and EURFRAC (fraction of population speaking other European languages). This yields the following two equations for the first stage estimations:

$$INS_i = \mu + \phi_4 ENGFRAC_i + \phi_5 EURFRAC_i + \phi_6 FR_i + \phi_7 GEOAREA_i + \phi_8 GEOSOIL_i + \varepsilon_{INSi} \quad (2a)$$

$$OPENC_i = \gamma + \phi_9 ENGFRAC_i + \phi_{10} EURFRAC_i + \phi_{11} FR_i + \phi_{12} GEOAREA_i + \phi_{13} GEOSOIL_i + \varepsilon_{OPENCi} \quad (3a)$$

In order to estimate equations (1) to (3), or (1) to (3a), we first have to construct the FR instrument of trade openness as a share of GDP for the year 2006, as the data used for Frankel & Romer (1999) use data for the year 1985. We start by estimating the Frankel and Romer (1999) gravity equation where bilateral trade between countries i and j is determined by the geographical distance between the two countries, their respective areas and populations, pair wise country landlockedness, a dummy variable for whether a common border between the countries exists, as well as some interaction terms with the common border dummy. The FR instrument thus tries to obtain a measure of the "geographic component of countries' trade". The equation estimated is as follows:

$$Log(\tau_{ij}/GDP_i) = a_0 + a_1 \ln D_{ij} + a_2 \ln N_i + a_3 \ln A_i + a_4 \ln N_j + a_5 \ln A_j + a_6 \ln (L_i + L_j) + a_7 \ln B_{ij} + a_8 B_{ij} \ln D_{ij} + a_9 B_{ij} \ln N_i + a_{10} B_{ij} \ln A_i + a_{11} B_{ij} \ln N_j + a_{12} B_{ij} \ln A_j + a_{13} B_{ij}(L_i + L_j) + e_{ij}. \quad (4)$$

Where τ_{ij} is the bilateral trade between countries i and j (measured as exports plus imports), GDP_i is the nominal GDP of the country, D_{ij} is the bilateral distance between the country pairs i and j, A is the area of the country, N is the population of the country, L is the dummy for landlockedness of the country, B_{ij} is the dummy for a common border between countries i and j, and e_{ij} is the random error term.

4. Data

4.1. Construction of the variables

The International Country Risk Guide (ICRG), Corruption Perceptions Index (CPI), and the World Bank's World Governance Indicators (WGI) are a few among the many different indicators available that measure institutional quality. The WGI have the advantage that they

---

[8] We also experimented with GEOAREA as an additional exogenous variable in equation (1), i.e., not as an instrument. This yields broadly the same results as we report below, although the instruments are somewhat weaker in this case.

cover a wide range of institutional factors. They consist of six aggregated indicators, for six different aspects of governance and institutions. In this study we use the Institutional Profiles Database (IPD)[9], which offers us 367 elementary variables which are then aggregated into three-digit (133 variables) and two-digit (93 variables) levels. This database allows us to capture the multi-dimensionality of institutions, and does not require us to work with a fixed weighting scheme for the underlying indicators as in the WGI.[10] Using this multi-dimensional approach is our key contribution. Rather than relying on a single measure of institutions (and growth, and geography), as Rodrik et al (2004) do, we will test the growth – institutions relationship taking a broader perspective than just one aspect of institutions.

We start from the 2-digit level IPD data[11], but since we have 118 observations (countries) and 93 variables, this would leave us very few degrees of freedom. We thus further aggregate these 93 variables, first to 12 new variables, and then we aggregate these 12 variables to 2 variables that are, each, ultimately used in the regressions as the INS variable from equation (1). This 2-stage aggregation of the IPD variables relies on (a spatial variant of) principal components analysis. In the first stage of this procedure, we start from the four "sectors" that the IPD database covers: (i) public institutions and civil society, (ii) goods and services markets, (iii) capital markets, and (iv) labour markets and labour relations. Each of the 93 aggregated variables falls into one of these sectors. For each of the sectors, we apply our spatial principal components analysis (SPCA), and produce a number of summary measures for each sector. This yields 12 new indicators. In the second stage, these 12 indicators are again summarized by SPCA, or by (regular) PCA. In the second stage, we maintain two components each for the PCA and the SPCA version, which gives us four summary variables for the institutions variable in total.

The reason why we use this 2-stage procedure is that the sectors have an unequal number of variables in them, and without the two stages, the (S)PCA results would be influenced stronger by the sectors with more variables. The 2 stage procedure has already been applied to IPD by Verspagen (2012), but only for regular PCA (not SPCA). The methodology of the spatial version of PCA is described in Bhupatiraju et al (2013). It differs from regular PCA in terms of the objective that it sets when producing the summary variables (components). In regular PCA, the objective is to maintain a maximum fit between the summary component and the original variables. Mathematically, this is achieved by looking at the correlation matrix of the original variables. The eigenvector that is associated with the largest eigenvalue of this matrix can be used as weights to construct the "best" summary measure (component).

In the spatial version of PCA (SPCA, Bhupatiraju et al, 2012), the objective is to produce a summary measure that produces maximum spatial correlation with itself. Spatial correlation is

---

[9] Freely available at http://www.cepii.fr/anglaisgraph/bdd/institutions.htm

[10] Verspagen (2012) discusses the correlations between the WGI and IPD databases.

[11] For a detailed explanation of the database for 2009 see: Crombrugghe, D de., Farla, K., Meisel, N., Neubourg, C de., Ould Aoudia, J., Szirmai, A., (2009), "Institutional Profiles Database III: Presentation of the Institutional Profiles Database 2009 (IPD 2009)" Les Cahiers de la DGTPE – n° 2009-14.

measured by the Moran coefficient, and is essentially a measure of how spatially concentrated the underlying phenomenon is. High spatial correlation means that high values of a variable in one country tend to be surrounded by other countries with high values of the variable. SPCA also works by taking eigenvalues and eigenvectors, but in this case the matrix on which this is performed is not a simple correlation matrix, but instead some more complicated matrix product involving, among other things, the matrix of Moran coefficients between the original variables. We will discuss the differences between regular PCA and SPCA in more detail when we look at maps displaying the institutions variables below. In order to save space, we will not document the details (such as eigenvalues and eigenvectors) of the procedure, but these are available on request.

Rodrik et al (2004) use distance from the equator as their preferred measure of geography. Like in the case of institutions, this is a very partial indicator. The archipelago of Indonesia embraces the equator, and in that sense it is similar to a country like Congo, but obviously the two countries are a world of difference when it comes to geography and climatic characteristics. Therefore, we use a much wider range of variables that describe the physical geographic characteristics and the soil related characteristics of the economy.

The raw data for the geography variables are taken from Gallup et al. (1998). Their dataset contains many variables, related to physical characteristics such as access to waterways, soil quality, climate zone, etc. We use SPCA to summarize these data into two summary variables: one for the physical characteristics of a country (based on 15 underlying variables), and one for the soil quality (based on 16 variables). In the regressions, both variables are included, and denoted as GEOAREA and GEOSOIL, respectively.

It is not only on the institutions or geography side that we would like to extend the measurement horizon. Rodrik et al (2004), as many others in the literature, proxy "growth" by the level of GDP per capita. This is justified by the long-run nature of the analysis. If growth rates are different between countries for a long enough period, the countries with high growth rates will also be(come) the ones with higher levels of income per capita. In practice, however, growth rates are variable over time, and there is no strong correlation at all between growth and the level of income. Therefore, in a strict sense, the regressions of Rodrik et al (2004) have very little to say about growth, but instead tell us something about the correlation between income levels on the one hand and institutions, trade and geography on the other hand.

Therefore, we will extend our horizon with regard to the dependent variable in the regressions (equation 1) to include also a number of other variables related to economic growth and development. We include annual growth rates of GPD per capita, average of consumption, investment and government expenditure as a percentage of GDP, an indicator of openness measured as exports plus imports as a percentage of GDP and population growth rates. All these variables are taken over the period of 2000-2009. Like before, we will use SPCA to summarize

these data into two summary variables, and both of these will be used as dependent variable in the estimation of equation (1).

Since we use cross-section data for the year 2009, we have to construct the Frankel Romer instrument (FR) for 2006. In order to construct this we use the "square" gravity data set from Head et al (2008).[12] This data set contains bilateral trade flows obtained from the International Monetary Fund's *Direction of Trade Statistics* (DoTS), as well as distance and size data. Bilateral distance between countries is measured using the great circle formula, between the biggest cities of the two countries. The data on area (measured in square kilometers), population, and dummies for landlockedness as well as common borders are also taken from the CEPII gravity data set compiled by Head et al (2006).

The other instrument that we use is settler mortality (SM), as first used by Acemoglu et al (2001), from which we also obtain these data. Using the above mentioned data we perform the 2SLS-IV regressions, the results of which are presented below. In order to include non-colonized countries in the analysis, we also use the Hall and Jones (1999) suggested ENGFRAC (fraction of population speaking English) and EURFRAC (fraction of population speaking other European languages) variables as a substitute instrument for MS.

4.2. Mapping the data

The summary variables for institutions, geography (both independent variables) and economic development (dependent variable) bring out the salient differences between countries in the underlying datasets that have many more variables. The SPCA procedure that we use to construct these summary variables is designed to interpret differences and similarities between countries in these underlying variables in a spatial (geographical) way. We opted for the SPCA (instead of the regular PCA) method because geography seems to play a large role in institutions. In the standard literature (summarized in Section 2 above), this is reflected in the fact that the instruments that are used for institutions almost always have a geographical nature. Settler mortality (one of the most used instruments) is very specific to particular environments, and depends on climate, diseases, etc. Language (another popular instrument) also has a clear geographical nature. Rather than just depending on the instruments to bring out the geographical correlation of institutions, we bring in geography in the construction of the variables, by means of the SPCA procedure.

As we do not have the space to document details of the SPCA analysis that underlies almost all of the variables in our analysis, and we do want to give the reader an impression of how this method affects the variables in the analysis, we provide maps that illustrate the patterns for the main variables in the analysis. We start, in Map 1a and 1b, with the institutions variable.

---

[12] Freely available at http://www.cepii.fr/anglaisgraph/bdd/gravity.htm.

These maps clearly bring out the spatial concentration of institutions. The first component is a clearly a measure of "Western" institutions, where we use "Western" for describing Europe and its offshoots (North America, Australia and New Zealand). All these areas are darkly colored (high values), with the exception of the USA, which scores somewhat lower. This first component broadly measures the institutional side that Verspagen (2012) has dubbed as the first stylized fact of the institutions – development relation, i.e., that developed nations tend to have the type of institutions that are highly valued in the "Western World", such as democracy, civil rights, transparent justice systems and free markets. From our SPCA exercise, it appears that these kinds of institutions are strongly spatially concentrated at a global scale.
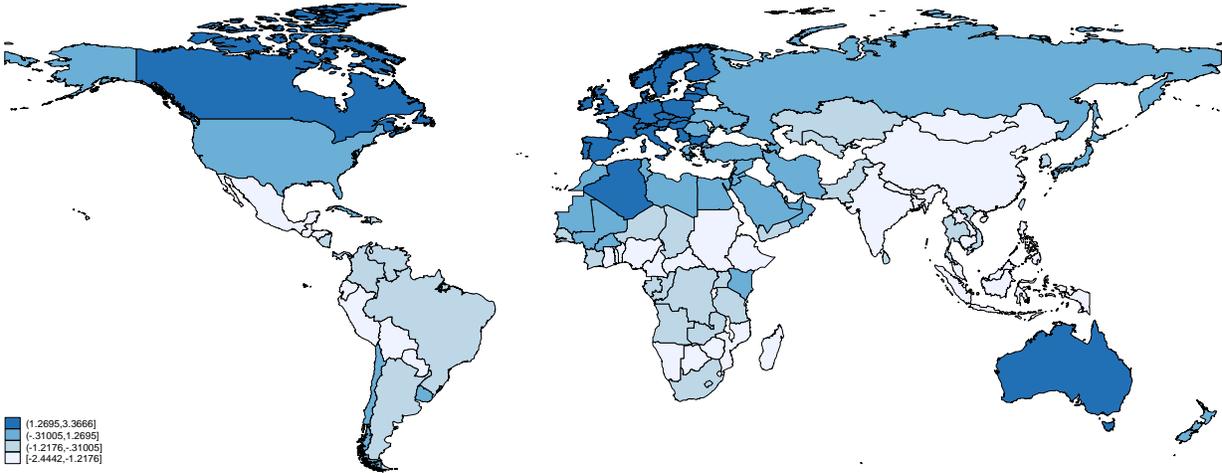
Map 1a. Institutions from the IPD database, 2009, first component
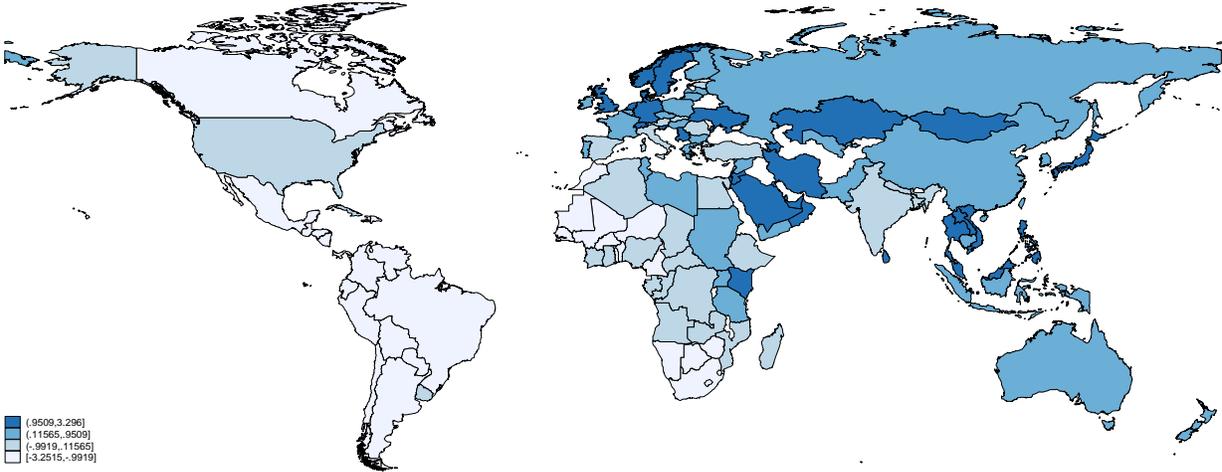
Map 1b. Institutions from the IPD database, 2009, second component

In Map 1b, we display the second component for institutions. This is again clearly spatially concentrated, but now the spatial pattern points to high values in the Eastern hemisphere, as well

as in some parts of Europe. Broadly speaking, the kind of institutional features that are stressed in this second component are those that are associated to a stronger role of the government in social and economic life. Africa is an interesting area for this indicator, as it shows a broad East-West division in itself (similar to the global pattern). Summarizing, the two maps (indicators), give us two rather distinct views on institutions at a global level, and it will be interesting to see how these are correlated to growth in the estimation of equation (1).
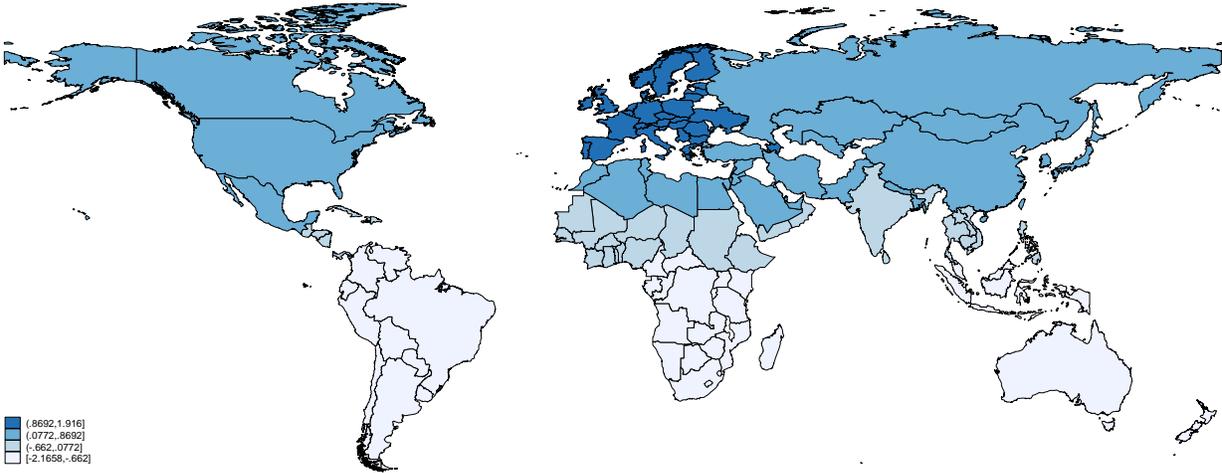
Map 2a. Physical geographical characteristics, first component
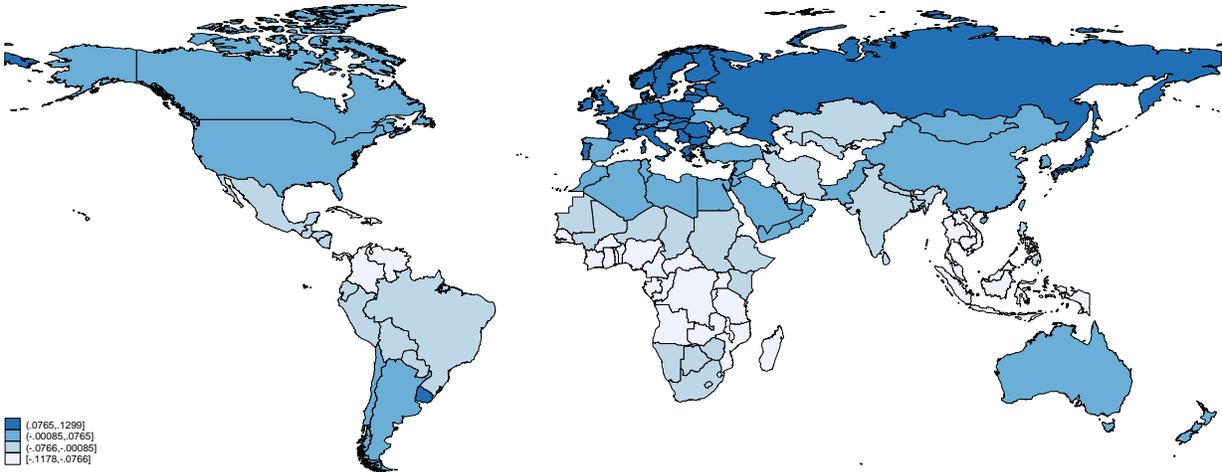
Map 2b. Physical geographical characteristics, second component

Map 2a displays the first component for the physical characteristics dataset (GEOAREA). Here, the high values are concentrated on the northern hemisphere, which is due to the fact that latitude has a high weight in the fact. On the Northern hemisphere, Europe scores particularly high, and this is mainly due to the relatively high weight of access to waterways and the small size of a country.

In Map 2b, the soil characteristics summary variable is displayed. This is the other geography related variable (GEOSOIL) that we include in the regressions as an independent variable. This variable is clearly also related to latitude, but less so than the GEOAREA variable. However, latitude does not enter in this component, and the impression that latitude affects this component actually stems from the influence of climate. The climate zones that are weighted strongly negative are savannah climates, dry climates, humid subtropical climates, and Mediterranean climates.
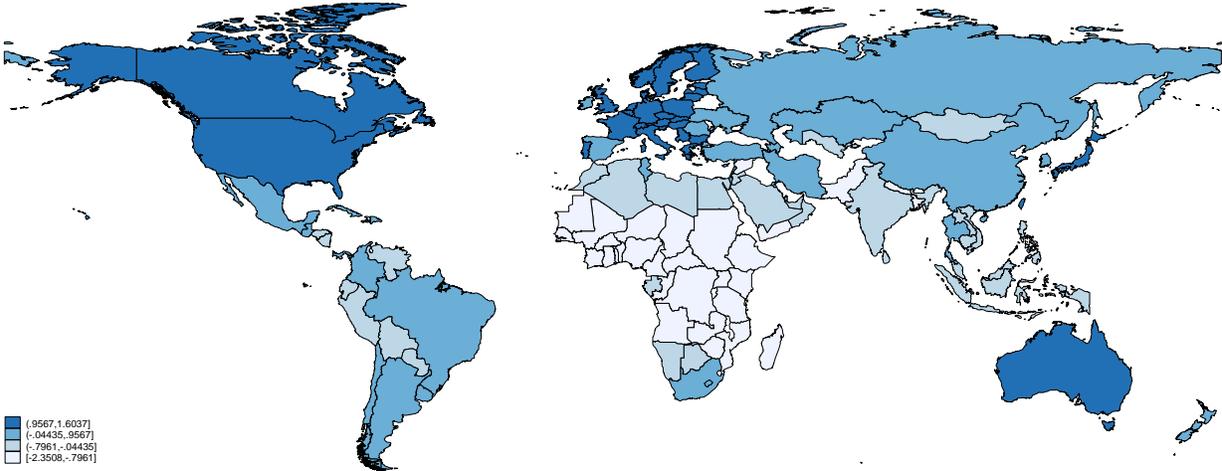
Map 3a. Economic development, first component
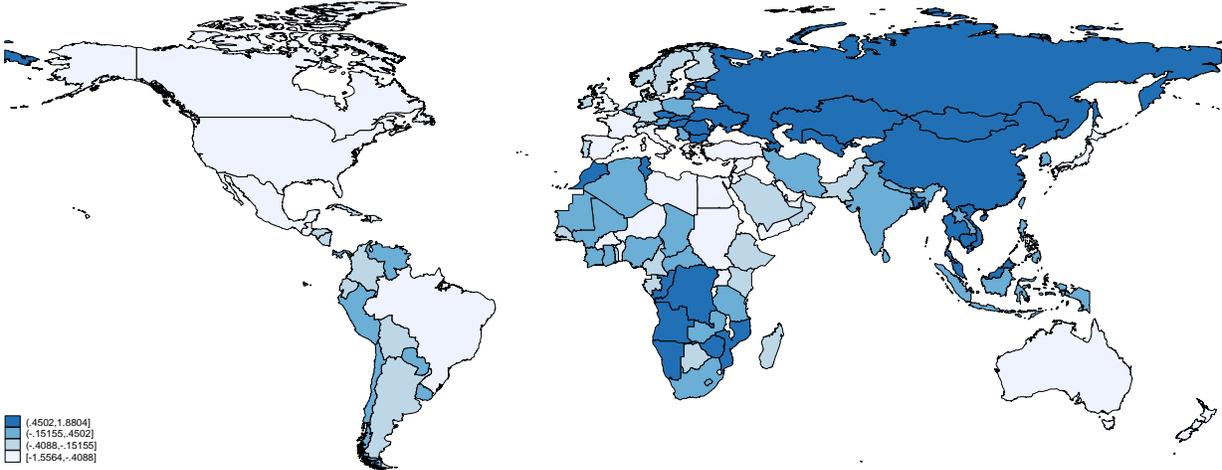
Map 3b. Economic development, second component

Finally, Maps 3a and 3b show the results of the SPCA procedure on the economic development variables. With the first factor here, in Map 3a, we capture the general development level. This

variable is mostly influenced by the initial level of GDP per capita (in 2000), but also by the share of consumption in GDP and, negatively, by population growth. All of these phenomena (low population growth, high income and high consumption) are characteristics of developed countries. Therefore, this component captures the economic side of the stylized fact that was already referred to above, i.e., that developed countries tend to have developed institutions.

Finally, Map 3b shows the second factor from the economic development dataset. This captures economic growth in a much more direct way, but in combination with some of the other indicators. We see high values in Asia, South-Eastern Europe and Central/Southern Africa. What characterizes these areas, is a relatively low initial (2000) GDP per capita level, combined with a relatively high growth rate, high openness to trade, a high investment rate and low population growth. Clearly, the indicator captures an overall growth or development strategy, rather than just the growth rate of GDP per capita.

5. Empirical Results

A. Construction of the Frankel and Romer instrument

Following Helpman et al (2008),we estimate a tobit model to take into account the missing and zero trade observations among the countries, rather than dropping those observations which have either no trade or missing trade values as in Frankel and Romer (1999). The estimated tobit equation yields broadly the same results as the OLS estimates. Table 1 shows the results of the estimated bilateral trade gravity coefficients.

The results are slightly different from those of Frankel & Romer (1999), especially for those that involve the interaction terms. We find that distance has a large, negative and significant impact on bilateral trade as expected. Bilateral trade is strongly increasing in j's population and decreasing in j's area and i's population, at a lower magnitude. Country i's area is the only non-significant variable among these. These results show that country i would trade more with countries that are closer to it, and those that have a large population but smaller areas.[13] The results also show that if both country are landlocked, trade between them is strongly negatively affected and this is significant. Sharing a border has a large effect on trade; however this estimate is not significant. This is in agreement with what Frankel and Romer (1999) find. According to them a possible explanation for this insignificance is due to the fact that "only a small fraction of country pairs share a border", and therefore are imprecisely estimated. The coefficients on A_i and L_{ij} (-0.00741 and -0.681) change significantly when we account for a shared border between countries, i.e., the coefficients on the interaction terms BA_i and BL_{ij} are -0.320 and 0.690. These interaction terms are both highly significant, showing that geography determines trade partially. Other interaction terms

---

[13] The results are interpreted as elasticities since they are log-transformed values.

however, are not significant. Out of the 12644 observations only one observation is left censored at -10.67.

Table 1: Bilateral Trade Equation (Tobit)

VARIABLES	(1) $\log T_{ij}\_GDP_i$	(2) $\log T_{ij}\_GDP_i$	(3) sigma
D_{ij}	-1.054***	-1.054***	
	(0.0251)	(0.0250)	
N_i	-0.0958***	-0.0959***	
	(0.0161)	(0.0161)	
A_i	-0.00741	-0.00736	
	(0.0146)	(0.0146)	
N_j	0.803***	0.803***	
	(0.0162)	(0.0162)	
A_j	-0.170***	-0.170***	
	(0.0144)	(0.0144)	
L_{ij}	-0.681***	-0.681***	
	(0.0318)	(0.0318)	
B	1.033	1.032	
	(1.182)	(1.181)	
BD_{ij}	0.316	0.316	
	(0.215)	(0.215)	
BN_i	0.131	0.131	
	(0.108)	(0.108)	
BA_i	-0.320***	-0.320***	
	(0.117)	(0.117)	
BN_j	0.0688	0.0687	
	(0.107)	(0.107)	
BA_j	0.0753	0.0754	
	(0.117)	(0.117)	
BL_{ij}	0.690***	0.689***	
	(0.165)	(0.165)	
Constant	6.575***	6.576***	1.922***
	(0.286)	(0.285)	(0.0121)
Observations	12,644	12,644	12,644
R-squared	0.339		

Notes: Standard errors in parentheses*** p<0.01, ** p<0.05, * p<0.1. The dependent variable is log (T_{ij}/GDP_i). Column (1) gives the results of the OLS estimation, while Columns (2) and (3) report the Tobit estimation.

Using these estimated results we obtain the fitted values of T_{ij} ($\hat{T}_{ij}$). We aggregate these values over j to get our instrument $\hat{T}_i$. This is the "geographic component" of country i's total trade. It should be noted here that correlation between the actual and the constructed trade share is 0.58 after controlling for size, which is a rather moderate correlation. This could be due to the imputed values of the geographically constructed bilateral trade data for which values are either missing or zero.[14] Since the imputation is based on the available geography data, Taiwan's trade share is not estimated due to missing GDP values.

B. Estimation of the RST model

Along with the above constructed FR instrument ($\hat{T}_i$) we also use the settler mortality instrument (SM), ENGFRAC and EURFRAC for the estimation of equation (1). As mentioned above, the IPD database has 118 countries which are used as the basis for the analysis. We perform the estimations on two sub-sets of these 118 countries. The smaller set has 52 countries for which the SM data is available, and the larger set has 104 countries for which ENGFRAC and EURFRAC data are available. We estimate and present the OLS as well as the 2SLS-IV estimations.

Table 2 below shows the results of these estimations. Columns (1) and (2) give the results of the OLS estimation of Equation 1, using the two institutional components in separate equations. We find that GDP per capita is positively and significantly related to GEOSOIL, and to both institutions variables. Columns (3) to (6) show the results of the 2SLS-IV estimation, with settler mortality as an instrument in column (3) and fraction of English and European languages as instruments in columns (4) to (6). The regression with settler mortality as an instrument has less observations (due to missing data for countries that were not colonized), and generally show less significant results. In column (3) none of the variables is significant. The 2SLS-IV estimates in columns (4) to (6) always yield a significant impact of institutions. The first component for institutions has a positive and significant impact in GDP per capita, the second component is negative, but only significant if it is entered in an equation without the first institutions component. OPENC and GEOSOIL also enter significantly in some of the equations, but these effects are not robust across the equations. OPENC has a negative and significant sign in equation (4), which is against expectations.

---

[14] Noguer and Siscart (2005).

Table 2. Regression results with GDP per capita (2009) as dependent variable, Institutions with SPCA

VARIABLES	(1) logy (OLS)	(2) logy (OLS)	(3) logy (IV-1)	(4) logy (IV-2)	(5) logy (IV-2)	(6) logy (IV-2)
GEOSOIL	6.402***	10.44***	33.14	-8.156	15.15***	-4.335
	(1.976)	(1.074)	(24.78)	(7.576)	(1.826)	(5.149)
OPENC	0.00132	0.00211	-0.0284	-0.0306**	0.0168	-0.0159
	(0.00253)	(0.00239)	(0.0411)	(0.0131)	(0.0132)	(0.0124)
INS1-SPCA	0.323***		-3.137	1.364***		1.222***
	(0.0977)		(2.869)	(0.476)		(0.309)
INS2-SPCA		0.167**			-0.508**	-0.314
		(0.0779)			(0.233)	(0.209)
Constant	8.865***	8.798***	9.452***	11.51***	7.547***	10.27***
	(0.228)	(0.225)	(2.807)	(1.080)	(1.128)	(1.056)
Observations	118	118	52	104	104	104
Underid. test (reject H_0)			N	Y	Y	N
Weak instruments			Y	Y	Y	Y
Hansen J statistic (reject H_0)			N	N	Y	N

Notes: Robust standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Columns labeled OLS are OLS regressions, columns labeled IV-1 use settler mortality as instruments, and columns labeled IV-2 use ENGFRAC and EURFRAC as instruments. The row "Underid. test (reject H_0)" reports results of Kleibergen-Paap rk LM test, a N indicates that under-identification is an issue. The row "Weak instruments" reports results of Kleibergen-Paap rk Wald F test, Y indicates that the instruments are weak. The row "Hansen J-statistic (reject H_0)" reports results of the Hansen J test for over-identification. Y indicates that over-identification is an issue.

We do not document any results for the first stage regressions (instrument equations) for any of the estimations in Table 2 or further tables. These first stage results are available on request. We did perform tests for under-identification, over-identification and for weak instruments. These results are reported at the bottom of the table. We used a 10% cut-off value for deciding on these tests. In these tests, we generally (also for the estimations below) find that the instruments are weak. This is a common problem with IV estimations, and was also reported as an issue in the original Rodrik et al. (2004) paper. Our tests do indicate that institutions and trade openness are endogenous variables, hence we cannot go on the OLS results alone. But finding strong enough instruments for these variables is a problem that the literature still needs to address.

In terms of under- or over-identification, equation (4) in Table 2 performs best. It passes both tests (but the test for weak instruments), while equations (3) and (6) do not pass the under-identification test, and equation (6) does not pass the over-identification test. Thus, equation (4) provides the best nmodel for the general development level in Table 2.

What does this table suggest in terms of the primacy of institutions in terms of explaining the level of GDP per capita? Certainly, the results in the table point to an important role of institutions in explaining the development level (GDP per capita). The two institutional variables are the only ones that are systematically significant in the 2SLS-IV regressions. The results for trade and geography are sometimes significant, which would point to an independent role of geography and/or trade, but this does not seem to be robust. We also find evidence that there are multiple sides to institutions, as both the first and the second component that we derived from the SPCA are significant. The first component seems to dominate (column 6), and the sign of the second component is negative, which leads us to conclude that what is related to the development level of a country is the degree to which institutions are of the "Western" type.

We now proceed to see whether and how these results change if we use a regular PCA method (instead of the spatial PCA) to summarize the institutions variables. These results are in Table 3. The main difference as compared to Table 2 is that the significance of the institutions variables is now much weaker. This holds for both the OLS and 2SLS-IV results. In fact, the only time that one of the institutions variables is significant, is in column (4), and this is only weakly significant (10%). Similarly, OPENC is never significant in Table 3. However, GEOSOIL is significant in all columns but equation (3).

All of the columns (3) to (6) in Table 3 suffer from under-identification. This can be interpreted as saying that the instruments are not good enough. With institutions that are not summarized in a geographical way, finding good instruments seems even more of a problem than in Table 2. In addition, equation (5) also seems to suffer from over-identification.

Table 3. Regression results with GDP per capita (2009) as dependent variable, Institutions with regular PCA

VARIABLES	(1) logy (OLS)	(2) logy (OLS)	(3) logy (IV-1)	(4) logy (IV-2)	(5) logy (IV-2)	(6) logy (IV-2)
GEOSOIL	11.54***	11.40***	8.000	13.79***	11.99***	14.48***
	(1.067)	(1.066)	(5.657)	(2.247)	(1.579)	(3.347)
OPENC	0.00316	0.00335	-0.0186	-0.0132	-0.00150	-0.0160
	(0.00247)	(0.00251)	(0.0250)	(0.0102)	(0.00838)	(0.0137)
INS1-PCA	0.330		-7.425	5.265*		6.341

	(0.346)		(9.845)	(2.824)		(4.406)
INS2-PCA		-0.124			-2.347	2.149
		(0.294)			(2.295)	(5.679)
Constant	8.709***	8.693***	10.43***	9.936***	9.065***	10.17***
	(0.228)	(0.229)	(1.548)	(0.834)	(0.702)	(1.140)
Observations	118	118	52	104	104	104
Underid. test (reject H_0)			N	N	N	N
Weak instruments			Y	Y	Y	Y
Hansen J statistic (reject H_0)			N	N	Y	N

Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Columns labeled OLS are OLS regressions, columns labeled IV-1 use settler mortality as instruments, and columns labeled IV-2 use ENGFRAC and EURFRAC as instruments. The row "Underid. test (reject H_0)" reports results of Kleibergen-Paap rk LM test, a N indicates that under-identification is an issue. The row "Weak instruments" reports results of Kleibergen-Paap rk Wald F test, Y indicates that the instruments are weak. The row "Hansen J-statistic (reject H_0)" reports results of the Hansen J test for over-identification. A Y indicates that over-identification is an issue.

The conclusion seems to be that it makes a difference whether we weight the underlying institutional variables from IPD in a spatial way, or without spatial weights. Non-spatially weighted institutions are difficult to instrument, and (therefore) do not seem to have a robust impact on the development level, while spatially weighted institutions do have such an impact. As a preliminary conclusion, it seems fair to that institutions indeed are a (the) primal factor behind economic development, but only if we take into account that institutions themselves are strongly geographically concentrated.

In order to test the effect of institutions on the broader aspects of development, including the growth rate of GDP per capita itself, we also estimate the regression models with a different dependent variable. For this purpose, we used the first and second components of the broad economic development dataset, as displayed in Maps 3a and 3b above. The results for the first factor were very similar to Tables 2 and 3, and therefore we do not document these results. This is due to the fact that the first component (Map 3a) mainly captures the development level, and hence is not too different from GDP per capita itself. Table 4 documents the results for the second factor ("growth"), using institutions with SPCA as the independent variable.

In the OLS regressions (columns 1 and 2), only the first institutions component enters significantly, but with a negative sign. OPENC also enters significantly twice, and GEOSOIL significant and negative once. Now, trade openness is significant in the 2SLS-IV results with settler mortality, but none of the other variables are significant (column 3). In the 2SLS-IV results with language instruments (columns 4 to 6), the second institutions component is never significant, but the first component remains significant and negative. GEOSOIL (twice positive and once negative) and OPENC (twice positive) are also significant. Thus, these results suggest that in terms of the more dynamic aspects of economic development, institutions are not the only

"primal" factor that contributes. Geography and trade also contribute. Moreover, "Western" institutions seem to contribute negatively to development in these regressions.

Table 4. Regression results with second economic development component (growth) as dependent variable, Institutions with SPCA

VARIABLES	(1) dev2 (OLS)	(2) dev2 (OLS)	(3) dev2 (IV-1)	(4) dev2 (IV-2)	(5) dev2 (IV-2)	(6) dev2 (IV-2)
GEOSOIL	0.125	-1.508**	-1.262	7.446*	-2.528***	7.142**
	(1.074)	(0.684)	(6.180)	(3.804)	(0.806)	(3.318)
OPENC	0.00730***	0.00661***	0.0239**	0.0193***	0.00191	0.0181**
	(0.00171)	(0.00156)	(0.0116)	(0.00709)	(0.00578)	(0.00755)
INS1-SPCA	-0.102*		0.225	-0.618**		-0.606***
	(0.0588)		(0.711)	(0.240)		(0.214)
INS2-SPCA		0.00820			0.121	0.0250
		(0.0504)			(0.0957)	(0.127)
Constant	-0.618***	-0.559***	-1.692**	-1.689***	-0.238	-1.590**
	(0.135)	(0.131)	(0.798)	(0.577)	(0.494)	(0.625)
Observations	118	118	52	104	104	104
Underid. test (reject H_0)			N	Y	Y	N
Weak instruments			Y	Y	Y	Y
Hansen J statistic (reject H_0)			N	N	Y	N

Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Columns labeled OLS are OLS regressions, columns labeled IV-1 use settler mortality as instruments, and columns labeled IV-2 use ENGFRAC and EURFRAC as instruments. The row "Underid. test (reject H_0)" reports results of Kleibergen-Paap rk LM test, a N indicates that under-identification is an issue. The row "Weak instruments" reports results of Kleibergen-Paap rk Wald F test, Y indicates that the instruments are weak. The row "Hansen J-statistic (reject H_0)" reports results of the Hansen J test for over-identification. A Y indicates that over-identification is an issue.

Like before, the instruments are weak, and all columns except (4) seem to suffer from either over- or under-identification. Thus, equation (4) is the preferred specification in Table 4, and it points to a negative impact of institutions, a positive impact of trade openness and a positive impact of GEOSOIL. Institutions no longer trump the other factors when we consider a more dynamic picture of economic development.

In Table 5, we repeat the estimations of table 4, but with using the regular PCA method to construct the institutions components. Like in Table 3, the results seem to suffer from under-identification (all columns 3 – 6). We have only one significant result for institutions in the table, in column (4), i.e., the 2SLS-IV results. Like in Table 4, the sign of institutions is negative. In

column (4), we also have a significant and positive effect of trade openness, but GEOSOIL, which was significant and positive in Table 4, is now significant and negative.

Table 5. Regression results with second economic development component (growth) as dependent variable, Institutions with regular PCA

VARIABLES	(1) dev2 (OLS)	(2) dev2 (OLS)	(3) dev2 (IV-1)	(4) dev2 (IV-2)	(5) dev2 (IV-2)	(6) dev2 (IV-2)
GEOSOIL	-1.536**	-1.434**	1.136	-2.657**	-1.615*	-3.042
	(0.613)	(0.620)	(2.335)	(1.213)	(0.853)	(1.878)
OPENC	0.00678***	0.00663***	0.0250**	0.0128**	0.00607	0.0144*
	(0.00156)	(0.00159)	(0.0115)	(0.00542)	(0.00469)	(0.00744)
INS1-PCA	-0.220		-0.712	-3.040**		-3.637
	(0.196)		(3.236)	(1.471)		(2.408)
INS2-PCA		0.131			1.385	-1.194
		(0.172)			(1.343)	(3.075)
Constant	-0.573***	-0.561***	-1.813**	-1.076**	-0.572	-1.206**
	(0.129)	(0.130)	(0.718)	(0.439)	(0.383)	(0.613)
Observations	118	118	52	104	104	104
Underid. test (reject H_0)			N	N	N	N
Weak instruments			Y	Y	Y	Y
Hansen J statistic (reject H_0)			N	N	Y	N

Notes: Robust standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Columns labeled OLS are OLS regressions, columns labeled IV-1 use settler mortality as instruments, and columns labeled IV-2 use ENGFRAC and EURFRAC as instruments. The row "Underid. test (reject H_0)" reports results of Kleibergen-Paap rk LM test, a N indicates that under-identification is an issue. The row "Weak instruments" reports results of Kleibergen-Paap rk Wald F test, Y indicates that the instruments are weak. The row "Hansen J-statistic (reject H_0)" reports results of the Hansen J test for over-identification. A Y indicates that over-identification is an issue.

Thus, for the broad development indicator that includes growth, we find some similar, and some different results as compared to just using GDP per capita. What is similar is that we have stronger results for institutions when we weight institutions in a spatial way. What is different is, first, the sign of the institutions variable, and, second, whether or not (some of) the other primal factors still matter. Our analysis of the broader development variable both stresses that institutions as they are usually measured have an adverse effect on development, and that, besides institutions, geography and trade also matter.

6. Conclusions

In this paper, we test the Rodrik et al (2004) framework to explain differences in development levels across countries by using a broader set of definitions for institutions, geography and economic variables. We use constructed indices for institutions, geography (divided into physical and soil quality) and economic variables (which take into account the variance in a set of seven variables related to the level of economic development/growth of countries).

We find from our analysis that institutions do play a primary role in explaining cross-country differences in income per capita. This part of our analysis largely confirms the Rodrik et al. (2004) results. We find, however, that it matters how a summary variable is created from our underlying multi-dimensional institutions dataset. The results are reasonably strong if we summarize institutions in a spatial way (i.e., create a summary measure that is spatially correlated). When we summarize institutions without taking into account space or geography, the results are much weaker. This suggests that institutions are indeed strongly spatially correlated, and that it matters to acknowledge this in the Rodrik et al.-type regressions. Institutions may trump the other factors (trade and geography), but they are themselves strongly geographically influenced. In other words, even if we do not see a direct impact of geography on GDP per capita, we should keep in mind that the institutions variables are spatially weighted and so 'space' does matter, thus we can conclude that spatial influence of neighbours has an impact on the economic performance of the economy.

We also estimated models in which we use a broader development variable. This is a composite index, produced by the spatial principal components analysis (SPCA) that was used to construct the institutions proxies. This broader economic development level captures growth in a more direct way, as it has high values for relatively under-developed, but fast-growing economies that are relatively open. With this variable as the dependent variable, we find that institutions, geography and openness all matter in the final equation. Institutions are no longer the only primal factor, geography and trade also matter. In addition, we find that the institutions proxy that we use has a negative influence. In other words, institutions as we usually measure them do not seem to contribute to growth of a dynamic development pattern in general.

The role of institutions in explaining the current level of income in various countries has been tested by various studies earlier. Untangling the historical and geographic reasons for the differences in growth is not an easy task. Our analysis shows that there is a need to analyze these differences using carefully constructed indices. For this line of research to go forward, we need to look at a wider choice of instruments which may prove to be better than those currently in vogue. For example the FR instrument is widely used as a proxy for trade share (as also in this paper), but the correlation between the constructed trade share and the actual trade share is only as much as 0.6. Settler mortality on the other hand tries to bring in a historical aspect in to the explanation of the quality of institutions, but is restrictive in the sense that only colonized countries can be tested and as seen by many earlier studies and by this one, is a weak instrument, much like EURFRAC and ENGFRAC.

Appendix

Variable descriptions for Table 1-5.

Variable Name	Description
D_{ij}	Distance between countries i and j
A_i, A_j	Area of country i or j
N_i, N_j	Population of country i , j
L_{ij}	Dummy for landlockedness of country i and j taken together.
B	Dummy for a common border between country i and j
BD_{ij}	Interaction term of a shared border and bilateral distance (Dij)
BN_i	Interaction term of a shared border and Population in i
BA_i	Interaction term of a shared border and Area in i
BN_j	Interaction term of a shared border and Population in j
BA_j	Interaction term of a shared border and Area in j
BL_{ij}	Interaction term of a shared border and landlockedness dummy
Logy	GDP per capita (in the year 2009, in international dollars)
INS1_SPCA	Measure of institutional quality (factor 1) using SPCA
INS2_SPCA	Measure of institutional quality (factor 2) using SPCA
INS1_PCA	Measure of institutional quality (factor 1) using PCA
INS2_PCA	Measure of institutional quality (factor 2) using PCA
dev2	Alternative measure of economic development (Factor 2)
GEOSOIL	Measure of agricultural productivity conducive for growth
GEOAREA	Measure of physical geographic factors conducive for growth
OPENC	Trade openness
$logT_{ij}\_GDP_i$	Instrument for trade openness (FR instrument)
ENGFRAC	Fraction of population speaking English
EURFRAC	Fraction of population speaking an European language

Bibliography

Acemoglu, D., S. Johnson, and J. Robinson, (2001), "The Colonial Origins of Comparative Development: An Empirical Investigation," American Economic Review, 91, 5, 1369-1401.

Acemoglu, D., S. Johnson, and J. Robinson, (2002), "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," Quarterly Journal of Economics, 117, 4, 1231-1294.

Acemoglu, D., S. Johnson, and J. Robinson, (2004), "Institutions as the Fundamental Cause of Long-run Growth," National Bureau of Economic Research Working Paper no.10481.

Bhupathiraju, S., Verspagen, B. and T. Ziesemer, (2013), *Summarizing Large Spatial Datasets: Spatial Principal Components and Spatial Canonical Correlation*, UNU-MERIT Working Paper, forthcoming.

Basu, S. R., (2008), "A New Way to Link Development to Institutions, Policies and Geography," Policy Issues in International Trade and Commodities Study Series No 37.

Bloom, D. and J. Sachs, (1998), "Geography, Demography, and Economic Growth in Africa," mimeo, Harvard Institute for International Development.

Bradhan, P., (2005), "Institutions matter, but which ones?" Economics of Transition 13.

Cameron, C. A., and Trivedi, P.K., (2009), Microeconometrics: Methods and Applications. Cambridge University Press.

Diamond, J., (1997), " Guns Germs and Steel", New York: W.W. Norton & co.

Dollar, D. and A. Kraay, (2004), "Trade, Growth and Poverty," Economic Journal, Royal Economic Society, vol. 114(493), pages F22-F49, 02.

Dollar, D. and A. Kraay, (2002), "Growth is Good for the Poor," Journal of Economic Growth, 7, 3, 195-225.

Dollar, D. and A. Kraay, (2003), "Institutions, Trade and Growth: Revisiting the Evidence," Journal of Monetary Economics, 50, 1, 133-162.

Frankel, J. and D. Romer, (1999), "Does Trade Cause Growth?," *American Economic Review*, 89, 3, 379-399.

Gallup, J.L., J. D. Sachs, A.D. Mellinger, (1998), "Geography and Economic Development" NBER Working Paper No. w6849. Available at SSRN: http://ssrn.com/abstract=145013

Hall, R., and C. Jones, (1999), "Why Do Some Countries Produce So Much More Output Per Worker Than Others?," *Quarterly Journal of Economics*, 114, 1, 83-116.

Helpman, E., Melitz, M., and Rubinstein, Y., (2008), "Estimating Trade Flows: Trading Partners and Trading Volumes," The Quarterly Journal of Economics, MIT Press, vol. 123(2), pages 441-487, 05.

Keefer, P. and S. Knack, (1997), "Why Don't Poor Countries Catch Up? A Cross-National Test of an Institutional Explanation," *Economic Inquiry*, 35, 3, 590-602.

Knack, S. and P. Keefer, (1995), "Institutions and Economic Performance: Cross Country Tests Using Alternative Institutional Measures," Economics and Politics, 7, 3, 207-27.

La Porta, R., Lopez-de-Silanes, F., Shleifer. A., and Vishny, R. W., (1997), " Legal Determinants of External Finance," Journal of Finance, American Finance Association, vol. 52(3), pages 1131-50, July

La Porta, R., Lopez-de-Silanes, F., Shleifer. A., and Vishny, R. W., (1998), "Law and Finance," Journal of Political Economy, University of Chicago Press, vol. 106(6), pages 1113-1155, December.

La Porta, R., Lopez-de-Silanes, F., Shleifer. A., and Vishny, R. W., (1999), "The Quality of Government," Journal of Law, Economics and Organization, Oxford University Press, vol. 15(1), pages 222-79, April.

Masters, W. and M. McMillan, (2001), "Climate and Scale in Economic Growth," Journal of Economic Growth, 6, 3, 167-186.

Masters, W. and J. Sachs, (2001), "Climate and Development," mimeo, Purdue University.

Mauro, P., (1995), "Corruption and Growth," Quarterly Journal of Economics, 110, 3, 681-713.

McArthur, J. and J. Sachs, (2001), "Institutions and Geography: Comment on Acemoglu, Johnson and Robinson," National Bureau of Economic Research Working Paper no. 8114.

Montesquieu C de S, (1748), "The Spirit of the Laws", Paris

Myrdal, Gunnar (1968) Asian Drama – An Inquiry into the Poverty of Nations Vol.3 (Harmondsworth, Penguin).

Nougar, M., and Siscart, M., (2005), "Trade Raises Income: A Precise and Robust Result," Journal of International Economics, 65, 447-460.

Rodriguez, F., and D. Rodrik, (2001), "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence," NBER Chapters, in: NBER Macroeconomics Annual 2000, Volume 15, pages 261-338 National Bureau of Economic Research, Inc.

Rodrik, D., A. Subramanian, and F. Trebbi, (2004), "Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development," Journal of Economic Growth, 9, 2, 131-165.

Sachs, J. and A. Warner, (1995), "Economic Reform and the Process of Global Integration (with discussion)," Brookings Papers on Economic Activity, 1, 1-118.

Sachs, J., (2003), "Institutions Don't Rule: Direct Effects of Geography on Per Capita Income," National Bureau of Economic Research Working Paper no. 9490.

Sala-i-Martin, X., (1997a), "I Just Ran 4 Million Regressions," National Bureau of Economic Research Working Paper no. 6252.

Sala-i-Martin, X., (1997b), "I Just Ran 2 Million Regressions." American Economic Review, 87, 2, 178-83.

Verspagen, B., (2012), *Stylized facts of governance, institutions and economic development. Exploring the institutional profiles database*, UNU-MERIT Working Paper 2012-036

Vieira, F. V., and Damasceno, A. O. (2009), "Is There a Primary Role of Institutions on Explaining Cross Country Differences in Income Levels and Long-Run Economic Growth?", XXXVII Encontro Nacional de Economia, 2009, Foz do Iguaçu.

The UNU-MERIT WORKING Paper Series

2013-01 *Effects of innovation on employment in Latin America* by Gustavo Crespi and Ezequiel Tacsir

2013-02 *Revisiting the porter hypothesis: An empirical analysis of green innovation for the Netherlands* George van Leeuwen and Pierre Mohnen

2013-03 *Impact of external knowledge acquisition strategies on innovation - A comparative study based on Dutch and Swiss panel data* by Spyros Arvanitis, Boris Lokshin, Pierre Mohnen and Martin Wörter

2013-04 *Interactive knowledge exchanges under complex social relations: A simulation modelRobin* by Cowan and Anant Kamath

2013-05 *Innovation systems framework: still useful in the new global context?* by Michiko Iizuka

2013-06 *The importance of intrinsic and extrinsic motivation for measuring IQ* by Lex Borghans, Huub Meijers and Bas ter Weel

2013-07 *Firms' innovation capability-building paths and the nature of changes in learning mechanisms: Multiple case-study evidence from an emerging economy* by Paulo N. Figueiredo , Marcela Cohen and Saulo Gomes

2013-08 *A set of time series data labour market stocks and flows for the Netherlands 1980 to 2010* by Manuel Müllers, Joan Muysken and Erik de Regt

2013-09 *Designing an optimal 'tech fix' path to global climate stability: R&D in a multi-phase climate policy framework* by Adriaan van Zon and Paul A. David

2013-10 *Complementarity between internal knowledge creation and external knowledge sourcing in developing countries* by Jun Hou and Pierre Mohnen

2013-11 *Summarizing large spatial datasets: Spatial principal components and spatial canonical correlation* by Samyukta Bhupathiraju, Bart Verspagen and Thomas Ziesemer

2013-12 *Regional systems of innovation in the Arab region* by Samia Satti Osman Mohamed Nour

2013-13 *Development and social justice: Education, training and health in Sudan* by Samia Satti Osman Mohamed Nour

2013-14 *The economic importance and impacts of intellectual property rights (IPRs) in Sudan* by Samia Satti Osman Mohamed Nour

2013-15 *Overview of knowledge economy in the Arab region* by Samia Satti Osman Mohamed Nour

2013-16 *The importance (impacts) of knowledge at the macro-micro levels in the Arab Gulf countries* by Samia Satti Osman Mohamed Nour

2013-17 *Political determinants and impact analysis of using a cable system as a complement to an urban transport system* by Diego Escobar-García, Francisco García-Orozco and Carlos Cadena-Gaitán

2013-18 *Women entrepreneurs in the informal economy: Is formalization the only solution for business sustainability?* By Shyama V. Ramani, Ajay Thutupalli, Tamas Medovarszki, Sutapa Chattopadhyay, Veena Ravichandran

2013-19 *Heterogeneity in innovation strategies, evolving consumer preferences and market structure: An evolutionary multi-agent based modelling approach* by Salih Çevikarslan