Does Institutional Quality Impact Innovation? Evidence from Cross-Country Patent Grant Data

by

Edinaldo Tebaldi
Department of Economics
Bryant University
1150 Douglas Pike, Suite J-142
Smithfield, RI 02917
Phone: 401-232-6901 / Fax 401-232-6068
e-mail: etebaldi@bryant.edu

and

Bruce Elmslie
Department of Economics
University of New Hampshire
15 College Road, McConnell Hall
Durham NH, 03824
Phone: 603-862-3347; Fax: 603-862-3383
e-mail: bte@cisunix.unh.edu

ACKNOWLEDGMENT

We would like to thank the participants of the seminars at the University of New Hampshire, Union College, Bryant University, and CAEN-Graduate Program in Economics-Brazil, and anonymous referees for their helpful and insightful comments.

Does Institutional Quality Impact Innovation?
Evidence from Cross-Country Patent Grant Data

SUMMARY

This paper contributes to the literature on institutions and economic growth by conducting an empirical examination of the links between innovation and institutions. Using cross-country data and the instrumental variable method, this study finds that institutional arrangements explain much of the cross-country variations in patent production. We find evidence that control of corruption, market-friendly policies, protection of property rights and a more effective judiciary system boost an economy's rate of innovation. Our results also imply that controlling for institutional quality; geographic-related variables are not significant in explaining patent production. This paper also finds evidence to support the idea that in the long-run human capital accumulation is an important factor in shaping institutions.

Keywords: Institutions, innovation, economic growth, patent grant, cross-country

Does Institutional Quality Impact Innovation? Evidence from Cross-Country Patent Grant Data

1. INTRODUCTION

A fundamental challenge for the economics profession lies in explaining the mechanisms of economic growth. There is little doubt that significant progress has been made during the last five decades in growth modeling and economists' understanding of the mechanisms of economic growth. It also appears that economists have reached a consensus on the notion that long-run economic growth is primarily a product of technical innovation. However, there is still much to learn about the factors that ultimately determine a country's rate of innovation and economists are increasingly aware that institutional arrangements affect knowledge accumulation (e.g. Rodrik, 2000; Sala-i-Martin, 2002; Gradstein, 2004) and consequently, recognize that institutional arrangements affect the long-run growth of output. If one wants not only to diagnose the problem of growth, but also search for ways to stimulate growth, it is very important to understand how institutions and innovation are linked. However, growth economists are still struggling to understand the linkages between institutional quality and innovation and to incorporate institutions into the standard theoretical framework of economic growth (Sala-i-Martin, 2002; Huang & Xu, 1999). Furthermore, few growth models *explicitly* address this issue (e.g. Huang & Xu, 1999; Fedderke, 2001; Gradstein, 2004; Tebaldi & Elmslie, 2008) and little empirical cross-country analyses directly examine such a link.

Although the availability of datasets on institutional quality[1] greatly contributed to the opening of a new front for empirical research aimed at evaluating the influences of institutional-related variables on economic performance (e.g. Barro, 1991; Knack & Keefer, 1995; Mauro, 1995; Hall & Jones, 1999; Acemoglu *et al.* 2001 and 2005a; McArthur & Sachs, 2001; Acemoglu & Johnson, 2005; Alcala & Ciccone, 2004; Easterly & Levine, 2003), researchers in this field face a major difficulty in conducting empirical work with institutions due to inherent imprecision and limitations in the definition and measurement of institutional quality. Engerman & Sokoloff (1997) argue that institutions should be “interpreted broadly to encompass not only formal political and legal structures but culture as well” (p. 261). North (1990) proposes examining institutions in terms of formal and informal rules and enforcement of procedures.[2] The New Institutional Economic school considers institutions as the “application and extension of concepts such as transaction costs, property rights, public choice, and ideology” (Furobotn & Ricther, 2005, p. 37). These general definitions provide little aid in building a workable framework for the measurement and modeling of institutional arrangements. Sala-i-Martin (2002) suggests a pragmatic conceptualization of institutions in terms of a set of elements related to the ways that a society and its economy works in modern capitalism. He argues that *institutions* (or institutional arrangements) should account for the enforcement of contracts, protection of property rights, perceptions that the judiciary system is predictable and effective, transparency of the public administration, control of corruption, and pro-market regulations.

This study follows Sala-i-Martin (2002) and utilizes five measures of institutions (Control of Corruption, Rule of Law, Regulatory Quality, Risk of Expropriation, and an average Index of institutional quality) to examine the impacts of institutional quality on patent grant. Our empirical specification is derived from a theoretical model developed in Tebaldi & Elmslie (2008), which gives simple testable predictions regarding the mechanism by which institutions influence technical innovation.

The empirical strategy used in this article is derived from Figure 1, which presupposes that early institutions are associated with the colonization process, geography and early accumulated human capital; and the interaction of these factors determine current institutions. Furthermore, current institutions affect technical innovation (measured by patent production), which ultimately affects a country’s economic performance. The econometric analysis conducted in this paper consists of estimating OLS and instrumental variable (IV and GMM) cross-country regressions, where the dependent variable is the number of patents (a measure of innovation).

[Figure 1 about here]

2. BACKGROUND AND METHODOLOGY

It is standard procedure in the empirical growth literature to include institutions in growth regressions as an ordinary explanatory variable of steady state per capita output (e.g. Kormendi & Meguire, 1985; Barro, 1991; Mauro, 1995; Knack & Keefer, 1995 & 1997; Oliva & Rivera-Batiz, 2002; Esfahani & Ramirez, 2003; Dollar & Kraay; 2003). However, the same is not true for innovation equations. In addition, despite a large and growing literature on innovation, the *empirical* specification of the innovation equation is still subject to uncertainty with respect to the variables that should be included in the model and there are substantial differences in the empirical equations and estimation techniques usually applied in empirical innovation research. For instance, Sedgley (2006)'s specification assumes that innovation (patent grant) is fundamentally determined by human capital, population density and Research and Development (R&D) expenditure. Chen & Puttitanun (2005) estimate a system of simultaneous equations including variables such as GDP per capita, trade flows, human capital, and a proxy for protection of property rights to explain patent counts (proxy for innovation). Despite the use of alternative methodologies (e.g. GMM) aimed at addressing endogeneity, Chen & Puttitanun (2005) estimates are plagued by unaddressed endogeneity given that most of the variables included in the model are endogenous (e.g. GDP per capita, protection of property rights, and innovation). Schneider's (2005) specification includes human capital, high tech imports, R&D expenditure, foreign direct investments, GDP per capita, and a measure of protection to property rights and is also subject to the endogeneity concerns.

We choose an approach that favors parsimony (vis-à-vis to a large number of variables) and derive our innovation equation from a theoretical model developed by Tebaldi & Elmslie (2008), which shows that changes in technology are associated with institutional quality and given by:

$$\dot{A} = \delta A H_A q(T(A)) \tag{1}$$

where A measures technical knowledge, H_A is human capital engaged in R&D, and q denotes the quality of the institutional structure, which depends on both the set of institutions in place (T) and the state-of-art technology (A).[3]

This formulation is consistent with Romer (1990), Aghion & Howitt (1992), Grossman & Helpman (2001) and explicitly shows that institutions impact innovation. The logic behind this formulation is that in addition to human capital and the *stock of knowledge*, institutions are a necessary input for the production of new R&D projects. Good institutions help in the process of registering new patents, diffusion of ideas across researchers, diffusion of current knowledge, enforcement of property rights and reduce the uncertainty of new projects; all factors that stimulate R&D activities. Therefore, in this theoretical formulation, *ceteris paribus*, good institutions are expected to impact positively the rate of technical innovation. However, it is also important to recognize that the impact of institutions on innovation is assumed to be affected by the state-of-art technology because of the needed adaptation of institutions in face of changes of technology (Matthews, 1986; Engerman & Sokoloff, 2005). It is worth noticing that the model developed by Tebaldi & Elmslie (2008) also shows that the impact of institutions on innovation spillovers to income such that, at the steady state, the growth rate of output increases as the institutional quality improves. More specifically, the model implies that in the steady state the growth rate of GDP per capita is equal to the growth rate of technology. Therefore, from a theoretical standpoint, showing that institutions impact the rate of innovation implies that institutions impact the growth rate of GDP per capita.

Although equation 1 intuitively associates changes in technology to institutional arrangements, its empirical counterpart can only be derived by imposing some specific functional form. We estimate the following model:[4]

$$\ln(\Delta A_{i,t}) = \beta_0 + \beta_1 \ln(H_{Ai,t-1}) + \beta_2 \ln(A_{i,t-1}) + \beta_3 T_i + \nu_{i,t} \qquad (2)$$

where t represents time, i indexes observations, and v is a random disturbance.[5]

Despite the apparent simplicity of equation 2, its estimation is difficult due to endogeneity and the difficulties in measuring technological change (A) and institutions (*T)*. In particular, most of the measures of institutional quality are available for very short periods of time and, most importantly, the observed variations in these *perception-based*[6] variables hardly measure fundamental changes in the quality of institutions. Precisely, a fundamental idea of the institutional-related literature is that institutions changes *slowly* and smoothly over time (Matthews, 1986; Atkinson, 1998). Once institutions are built, economic, social, and political mechanisms generated as a byproduct of those institutions are expected to set constraints on future institutional changes, so that those early institutional arrangements tend to persist over time (Engerman & Sokoloff, 1997 and 2005; La Porta, *et al.*, 1999; Acemoglu *et al.*, 2001). Therefore, the short-term variations in the available measures of institutions are unrelated to the type of institutional change needed to promote fundamental changes in how a society is organized and operates and likely reflect measurement error and/or short-term changes in perception. To circumvent this problem, it is common in institutional-related studies to measure institutional quality as an *average* over the period for which the data are available (e.g. Acemoglu *et al.*, 2005a and 2001; Dollar & Kraay, 2003). This approach is appealing and consistent with the institutional theory, but its main drawback and that it generates just one observation per country and therefore, prevents the use of panel data techniques to estimate equation 2. Hence, cross-section estimation methods (OLS, 2SLS, and/or GMM) are the best feasible alternatives to estimate the innovation model.

(a) Endogeneity

Endogeneity issues plague equation 2. We address the eventual endogeneity of *H* and *A* by using *initial* values, which are predetermined and, therefore, uncorrelated with the error term. However, *T* is measured *contemporaneously* and is surely endogenous. This undermines the reliability of estimates obtained by Ordinary Least Squares (OLS). To eliminate this problem, a set of instruments for institutions, which needs to be correlated with current institutions but uncorrelated with the error term of equation 2 should be used.

The empirical literature on institutions suggests that much of the variation in current institutions can be explained by geography-related variables and historically determined factors such as colonial status and origin of the legal system (Hall & Jones, 1999; La Porta *et al.*, 1999; McArthur & Sachs, 2001; Acemoglu *et al.*, 2001; Acemoglu & Johnson, 2005). Following this approach, current institutions are modeled as follows:

$$\hat{T}_i = \delta_1 + \delta_2 H_{0,i} + \delta_3 G_i + \delta_3 R_i + \eta_i \tag{3}$$

where $\hat{T}$ denotes institutions, H_0 denotes initial human capital, G is a vector of geographical variables, R is a vector of *other* exogenous determinants of institutions (e.g., colonial status or legal origin) and η is a vector of random disturbances.

Equation 3 is very similar to the empirical specification for institutions found in La Porta *et al.* (1999), McArthur & Sachs (2001) and Acemoglu *et al.* (2001 and 2005a).[7] However, this study adds previously accumulated human capital as a determinant of current institutions. More specifically, this equation states that the initial level of human capital is an important input in the shaping of early institutional arrangements.[8]

Equations 2 and 3 form a system of equations that links innovation (proxied by patent production) to institutions. This specification implies that the origin of the legal system, geographically related variables and the initial human capital endowment determine current institutions, but are uncorrelated with the error term of equation 2 and therefore, uncorrelated with current innovation. This setup may be contentious because one could argue that these variables are directly correlated with innovation even after controlling for institutions. However, it seems to be reasonable to presuppose that the colonial legacy directly influences current institutions, but has no direct effect on current innovation, so the colonial legacy variables should not be correlated with equation 2's error term. In other words, the innovation effect from the colonial legacy is felt through the impact on current institutions rather than directly influencing current innovation. Additionally, as argued previously, the initial human capital endowment may have affected early institutions, which ultimately shaped current institutions. Because current innovation is a function of contemporary institutions this variable could have an indirect effect on current patent production via current institutions. Finally, geography-related variables may have a direct effect on current institutions as well as a direct effect on innovation. Because this is an empirical question, it is examined together with the estimation of the model. The concerns regarding the identification of the model constitute, in fact, an empirical issue that can be evaluated by testing if the instruments are correlated with the error term in equation 2. We use the Hansen's J and the Anderson Canonical correlation tests to examine whether the variables listed above satisfy the requirements for valid instruments.

(b) Measurement error

Another issue here is the potential for significant measurement error in the institutional variables. If an explanatory variable is measured with error, it is necessarily correlated with the error term. In the presence of measurement error OLS estimates will be biased and inconsistent (Davidson & MacKinnon, 1993). According to Hall & Jones (1999), this problem can be addressed together with the endogeneity issue by using the IV estimator. Consider that institutions are measured with an error, such that:

$$T_i = \overline{T}_i + \mu_i \qquad (4)$$

where T is unobserved institutions, $\overline{T}$ is measured institutions and μ is the measurement error. Substituting equation 4 into equation 2 generates:

$$\ln(\Delta A_{i,t}) = \beta_0 + \beta_1 \ln(H_{Ai,0}) + \beta_2 \ln(A_{i,0}) + \beta_3 \overline{T}_i + \beta_3 \mu_i + \nu_{i,t} \qquad (5)$$

The explanatory variables from equation 3 and 5 can be stacked in matrices $Z=[H_{A0} \quad A_0]$ and $X=[H_0 \quad G \quad R]$ respectively. If X is a valid instrument for $\overline{T}$, then $E[X'v]=0$. Assuming that μ is uncorrelated with v, Z and X; β_3 is identified by the orthogonality conditions and both the measurement error and the endogeneity concerns are addressed. Therefore, it is crucial for the reliability of estimates to select variables to instrument institutions that are uncorrelated with the error term of the second-stage regression.

3. DATA

We develop a proxy for innovation using the sum of patents granted over a specific period. This approach follows recent studies that suggest that one can use "patents to create systematic measures of intangibles [innovation/knowledge] that drive economic growth" (Romer, 2002, p. ix, Foreword). Of course, this approach is subject to major concerns including the fact that not all patents have the same quality and not all inventions are patented because some do not meet certain criteria necessary for an invention to be patentable or because economic agents may strategically "rely on secrecy or other means of appropriability" (Jaffe & Trajtenberg (2002, p. 4). "[I]n spite of all the difficulties, patents statistics remain a unique resource for the analysis of the process of technical change" (Griliches, 1990, p.1702). [9]

The patent data are from the United States Patent and Trademark Office (USPTO) and record information about the numbers of patents granted to non-residents, that is, inventions created in countries other than the U.S. whose inventors wish to patent their ideas in the U.S. market.[10] Non-resident patenting suggests that those inventions have some non-negligible economic value and may embody a valuable contribution to the stock of knowledge. Our proxy for innovation using USPTO data consists of the accumulated number of patents granted since 1970 and is defined as follows:

$$\Delta A_i^{USPTO} = A_{i,2003} - A_{i,1970} = \sum_{t=1970}^{2003} p_{it}^{USPTO} = P_i^{USPTO}$$

where p denotes the number of patents granted to country i.[11]

We follow a growing branch of the growth literature that suggests that subjective institutional measures provide relevant information about growth-promoting institutional arrangements and thus can be used as proxies for institutions (e.g. Kormendi & Meguire, 1985; Mauro, 1995; Grogan & Moers, 2001; Hall & Jones, 1999; McArthur & Sachs, 2001; Acemoglu *et al.*, 2001 & 2005a; Dollar & Kraay, 2003). The most common measures of institutions used in the growth and institutions literature include autocracy, government effectiveness, executive constraints, risk of expropriation, and (control of) corruption. However, many of these proxies for institutions are highly correlated (see La Porta *et at.*, 1999; Glaeser *et al*., 2004) and the choice of the institutional measure for conducting empirical work is, in general, *ad hoc*. This study uses five different measures of institutions. Four of these measures are from a dataset developed by Kaufmann *et al.* (2003), which covers 199 countries. Specifically, it utilizes an *average* through the time periods 1996, 1998, 2000 and 2002 of Rule of Law, Control of Corruption and Regulatory Quality. [12] We also construct an overall index of institutional quality based on the average of Regulatory quality, Rule of Law, Control of corruption, Voice and Accountability, Political Stability, and Government Effectiveness. These measures of institutions "are based on several hundred variables measuring perceptions of governance, drawn from 25 separate data sources constructed by 18 different organizations" (Kaufmann *et al.*, 2003, p. 3). The Rule of Law, Control of Corruption, Regulatory Quality, and the average index variables range between -2.5 and 2.5, with higher scores indicating better institutional arrangements.

The fifth measure of institutional quality is a proxy for market institutions and measures the risk of confiscation and forced nationalization. This variable is used with the purpose to conform to other studies in the growth and institutions literature. It is calculated as the *average* value for each country over the period 1985-1995 and ranges between 0 and 10 with higher scores representing better institutions and thus lower risk of confiscation or forced nationalization. This variable is originally from Political Risk Services, but it was taken from McArthur & Sachs (2001). [13]

We utilize two alternative variables as proxies for the initial stock of knowledge: patents granted in the initial period (as suggested by Griliches, 1990) and a relative measure of book production.[14] The patents granted in the initial period is calculated as a short-term stock, that is: $A^{p}_{0,i} = \sum_{t=1963}^{1970} p_{t,i}$, where p denotes the number of patents granted to country *i* and *t* indexes times.

Data on book production[15] are from several years of the UNESCO Statistical Yearbook and defined as the number of book titles produced (non-periodical publications) by each country.[16] In order to minimize noise from a single year's data (the initial values play a key role in determining the parameter estimates) the calculations are made using the average number of books produced over a five year period. In a few cases, the statistics are not available for all five years during the time period considered and the average is calculated using the available information. Moreover, we normalize it using the United States book production as a reference. Specifically, A_0 is calculated as follows:

$$A_{0,i} = \left(\sum_{t=1968}^{1972} B_{t,i} \Big/ n_i \right) / \left(\sum_{t=1968}^{1972} B_{t,US} \Big/ n_{US} \right)$$

where B denotes yearly book title production, i indexes countries, t indexes time and n denotes the number of years for which the data is available.

Personnel engaged in R&D is measured as the total number of scientists, engineers, technicians and supporting staff engaged in research and development. The data on personnel is expressed as a full-time equivalent. This variable is from several years of the UNESCO Statistical Yearbook and was calculated as an average from 1968 to 1972.

The idea that the development of institutions is an evolutionary process depending on previously accumulated human capital is accounted for in the empirical model by including a variable that measures human capital accumulation in the early 20th century. This variable is calculated as the number of students enrolled in primary and secondary schools in 1920 per square kilometer.

$$H_{0,i} = h_{0,i} \big/ area_i$$

where h_0 denotes the number of students in school in 1920, *area* denotes the country land area and i indexes countries. The data on students enrolled in primary and secondary schools in early 20th century are from Mitchell (2003a, 2003b, 2003c).[17] The country area is from the United Nations and based upon the current geopolitical arrangement.[18]

The geographic variables are taken from McArthur & Sachs (2001) and La Porta *et al. (1999).* We use i) mean temperature, which measures the 1987 mean annual temperature in Celsius; ii) coastal land, which quantifies the proportion of land area within 100 km of the coast and iii) latitude, which quantifies the absolute value of the latitude, is scaled to take values between 0 and 1. This variable is taken from La Porta *et al.* (1999). The colonial legacy is measured by a set of dummy variables that identify the origin of a country's legal system. Specifically, these dummies identify if the origin of the legal system is English, French, German, Scandinavian, or Socialist. These variables were taken from La Porta *et al.* (1999).

4. REGRESSION ANALYSIS

Table 1 shows that patent production is highly correlated with all measures of institutions. We use regression analysis to test if the positive relationship between institutions and patent production holds as we account for other factors that influence patent production. Several alternative specifications and different methods of estimation (OLS, 2SLS, and GMM) are utilized to evaluate the robustness of the results.

[Table 1 about here]

(a) The Determinants of institutions

Despite the fact that appraising the determinants of institutions is not the primary goal of this study; our model of patent production requires finding suitable instruments for institutions. This necessitates at least some evaluation of the determinants of institutions.[19] Our empirical strategy closely follows the literature by instrumenting institutions with legal origins (La Porta *et al.*, 1999), geography (Acemoglu *et al.,* 2001; Rodrik, 2000) and historical levels of human capital. Our estimates of the first stage regressions for the four measures of institutions considered in the paper conform to most of the results previously found in the literature.

Tables 2 and 3 provide the estimates of equation 3 using the four different institutional measures. The adjusted R-squared indicates that over 60 percent of the variation in Control of Corruption and Rule of Law can be explained by the set of explanatory variables included in the estimates. The model explains about 50 percent of the variation in Regulatory Quality and Risk of Expropriation. The overall fit of the model is in accordance with previous studies in the field (e.g. Acemoglu *et al.*, 2001 & 2005a; La Porta *et al.,* 1999). It is worth noticing that mean temperature and latitude are highly correlated (-0.85) and provide comparable information about climate conditions. Including these two variables simultaneously may cause severe multicollinearity but not improve the model's fit. The best parsimonious specification seems to include only mean temperature and the share of land in the coast as controls for geographic influences.

[Tables 2 and 3 about here]

The estimates provide evidence that historical levels of human capital, legal origins, and geographically related variables are important in explaining institutions. In addition, controlling for the variables mentioned above, all of the tested regressions suggest that historical population density has no effect on shaping current institutions.[20]

Our estimates of the first stage regressions contribute to the literature because the regression analysis is conducted using a recent dataset on institutional quality and a measure of human capital accumulation in the early 20th century. Furthermore, the results discussed above are in accordance with previous findings in this field. For instance, La Porta *et al.* (1999) find that French and Socialist legal origins have negative effects on the measures of governance, i.e., a property rights index, a business regulation index and a corruption index. Hall & Jones (1999), La Porta *et al.* (1999), Acemoglu *et al.* (2001), McArthur & Sachs (2001) and Acemoglu & Johnson (2005) also find that geographically related variables affect current institutions. While we are able to replicate the findings of previous empirical examinations, this study adds a new finding that schooling density in the early 1900's also helps to explain current institutions. This finding is consistent with the reasoning of Lipset (1960), which argues that human capital accumulation contributes to shape more benign politics, less violence and more political stability. As noticed by Glaeser *et al*. (2004, p. 282), "[t]he key human capital externality is not technological but political: courts and legislatures replace guns. These improvements in turn bring about greater security of property and economic growth".

(b) Institutions and innovation

Tables 4 through 8 report the regressions of the determinants of innovation using different measures of institutions. In these five tables, the dependent variable is the USPTO patent count between 1970 and 2003. Models 1, 2, 4 and 5 only include variables suggested by the theoretical model (Equation 1). The other models are augmented to evaluate the model's robustness. Model 1 is estimated using OLS, models 2 thru 4 are estimated using 2SLS, and models 5 thru 7 are GMM estimates. The results show that the OLS estimates underestimate the impact of institutions on innovation. In all 2SLS/GMM regressions the coefficients on institutions are larger than the corresponding coefficients generated by OLS. This clearly indicates that there exists significant measurement error in the institutional variables, which is not accounted for in the OLS regressions, thus generating significant bias in the OLS estimates.[21] Moreover, we find that the 2SLS and GMM estimates are very similar and produce identical quantitative and qualitative conclusions. This suggests that the disturbances in the model are not subject to noteworthy correlation and/or heteroskedasticity and that the coefficient estimates are efficient.[22]

[Tables 4-8 about here]

In all regressions the coefficients on the share of book production have the expected signs (positive) and are statistically significant at the 5 percent level of significance. In addition, the estimates using *patents granted in the initial period* as a proxy for initial knowledge (not reported in the paper) instead of the share of book production show no significant differences compared to the results reported in tables 5 through 8. This is not surprising because Figure 2 shows that these two proxies for the initial stock of knowledge are highly correlated, so the estimates should not differ significantly when using either of these proxies. In addition, the high correlation between the share of book production and the initial stock of patents also suggests that knowledgeable societies are more innovative. Therefore, our strategy of using the share of book production to proxy initial knowledge seems to be reasonable. [23] To save space, the results using the initial flow of patents are not reported.

[Figure 2 about here]

The coefficients on personnel engaged in R&D are positive and statistically significant in all regressions of Tables 4 through 8. The results do not support the presence of a scale effect from personnel engaged in R&D on patent production. In all of the regressions, the null hypothesis that the coefficient on personnel engaged in R&D is equal to one is rejected (at the 1% and 5% levels of significance) in favor of the alternative hypothesis that the coefficients are less than one. In fact, the point estimates are below 0.5 and suggest that innovation is subject to significant decreasing returns to scale with respect to human capital. This implies that long-run growth of technology (patent production) cannot be sustained only with the increase of the workforce engaged in R&D. This finding contradicts the predictions from Romer (1990)'s model, but is in accordance with Jones (1995).

The estimates provide evidence that, controlling for the size of an economy and the position of a country on the world knowledge frontier, institutions have a strong positive effect on innovation. Specifically, the coefficients on institutional measures are positive and statistically significant in all of the regressions reported in Tables 4 through 8. This suggests that economies that have market-friendly policies, lower perception of corruption, or whose judiciary systems are more effective and predictable will experience higher rates of innovation (patent production). For example, the estimates suggest that if Brazil had the same market regulatory system as that of the U.S. (the difference in the regulatory quality index between these two countries is about 1.2), Brazilian residents would have produced about 2.5 times more patents than what they actually produced between 1970 and 2003. The model also suggests that controlling for the differences between the Brazilian and the U.S. indexes of control of corruption or rule of law, Brazilian patent production, between 1970 and 2003, would be approximately four times larger than it actually was. The presence of measurement errors in quantifying institutions implies that interpretation of the coefficients (marginal effects) of the model may be imprecise. However, the impacts of improving institutions on patent production seem to be quite large. Institutions are fundamentally important in the determination of how innovative an economy will be.

(c) Robustness

The USPTO patent data used in this study are subject to the home bias (home advantage). To alleviate concerns regarding this issue, we run the same set of regressions reported in tables 4 thru 8 using patent data from the World Bank,[24] which records data on patents granted by each country's patent office to its residents, as the dependent variable. It is worth noticing that despite the fact that patent counts greatly differ across the USPTO and the World Bank datasets; there is a 0.81 correlation between the natural logarithms of the accumulated patent counts from the USPTO and from the World Bank for the period 1995 to 2001 and a 0.79 correlation between these measures using data from different periods

(USPTO: 1970-2003 and World Bank: 1995-2001). This finding suggests that these two measures encompass similar information about innovation. Moreover, reinforcing this idea, the regressions using the World Bank patent dataset as the dependent variable (not reported in the paper) produce results very similar to the ones reported in tables 4 thru 8.[25] This suggests that our results are robust to the use of alternative measures of patent counts.

Models 1, 2, 4, 5 and 7 of Tables 4 through 8 assume that geography affects innovation through an indirect effect from institutions. This specification is similar to that used by Hall & Jones (1999). However, Gallup *et al.* (1999), Sachs (2000) and McArthur & Sachs (2001) argue that geography may have a direct effect on production as well as an indirect effect from institutions. The argument could be made that spatial location contributes directly to innovation. For instance, mean temperature (climate) may affect the health of the personnel engaged in R&D[26]. Models 3 and 6 of Tables 4 through 8, report the estimates of the model augmented with mean temperature and the share of land in the coast. The regressions provide no support to the claim that geographically related variables have direct effects on innovation. Specifically, once we have controlled for institutions, the coefficients of mean temperature and share of land in the coast are insignificant at the standard levels of significance in all regressions of Tables 4 thru 8.[27]

The results discussed above are subject to the validity of the IV estimates, which ultimately depend on the assumption that the model is properly identified. Of course, we recognize that the instruments are not perfect and that the exclusivity assumption may be violated. For instance, although initial human capital is expected to not be correlated with current patent production (which implies that it would also be correlated with the error term of equation 2), there are channels (e.g. omitted variables, measurement error) by which initial human capital may be correlated with the error term of the second-stage regressions. To alleviate concerns about the validity of the instruments, we use the Hansen's J (Hansen, 1982) and the Anderson Canonical Correlation tests (Anderson, 1951) to evaluate the overidentifying restrictions in the 2SLS/GMM regressions. In general, the overidentification tests (see the bottom panel of Tables 4 through 8) suggest that the correlation between the instruments and the error term is not significant. The results give more confidence in the validity and robustness of the estimates. However, the test does cast some doubt that the model for risk of expropriation is correctly identified.

The functional form chosen for the empirical model is also subject to imperfections, so alternative specifications can be used to check the robustness of the results. A key concern regarding the validity of the specification is motivated by the idea that countries that are far from the technological frontier do not patent as much as the others, independently from the quality of institutions. Another concern originates from the idea that innovation in developed and developing countries may respond differently to changes in institutional quality. We address these concerns by estimating alternative models whose estimates are shown in Table 10. A quadratic specification on institutions (Model 1) shows no differential effect. Models 2 and 3 test if innovation in OECD[28] and non-OECD countries responds differently to the quality of institutions. The estimates show no differential effect for countries in the technological frontier (OECD) compared to countries far from the technological frontier (non-OECD countries).

We also test if, controlling for institutional quality, the initial GDP per capita, a measure of economic development (and a measure of the distance from the technological frontier), affects innovation. The coefficients on initial level of GDP per capita (Model 4 of Table 9) turned out insignificant at the standard levels of confidence. Following Stern *et al.* (2002) and Schneider (2005), we dropped the book production variable and included GDP per capita as a proxy for the stock of knowledge. Column 5 of Table 9 shows that the coefficient of GDP per capita is positive and significant at the 10 percent level of significance, but more importantly, this approach does not affect the results (size and significance) on the institution variable and alleviate concerns regarding the appropriateness of our strategy top proxy for the initial stock of knowledge.

Finally, we also examine if the results are sensitive to the time period used to calculate the cumulative patent count by estimating a model that considers patent grants between 1995 and 2001 (this period was chosen ad hoc). Model 6 of Table 9 provides no evidence that the time period influences the qualitative results. All regressions of Table 9 use an average index of institutional quality as a proxy for institutions. However, similar results are also obtained using individual variables (e.g. control of corruption, rule of Law, etc.) instead of the average index.

These results together suggest that the impact of institutional quality on innovation is important for countries in the technological frontier as well as for countries far from the technological frontier. This finding also refutes the idea that the link between institutions and patenting found in this study could be by showing a relationship between the level of development and a country's distance from the technological frontier.

[Table 9 about here]

5. CONCLUSION

This study contributes to the literature by conducting an empirical examination on the links between innovation and institutions. The estimates obtained show that institutional arrangements positively contribute to explain much of the cross-country variations in patents granted. Specifically, we find evidence that control of corruption, market-friendly policies, protection of property rights and a more effective judiciary system boost an economy’s rate of innovation (patent production).

Although we do not formally test for a steady state growth effect, a major prediction of the classical Solow Model (1956), Romer (1990), Jones (1995), Aghion & Howitt (1992), Grossman & Helpman (2001), and Tebaldi & Elmslie (2008), among many others, is that the steady state growth rate of output per capita is equal to the rate of technological innovation. Therefore, growth theory together with our results provides evidence of a *growth effect* through innovation, i.e., institutions have a growth effect on income because institutional quality affects an economy’s rate of innovation, the engine of economic growth.

The econometric estimates also provide evidence that the impact of institutional quality on innovation is important for countries in the technological frontier as well as for countries far from the technological frontier. This result is important because it refutes the idea that countries that are far from the technological frontier do not patent as much as the others, independently from the quality of institutions. This research also finds that geography, *per se*, cannot explain differences in innovation across countries. Geography affects innovation, but only through institutions. This paper also finds empirical evidence to support Lipset's (1960) theoretical arguments that in the long-run human capital accumulation is a key variable in shaping institutions.

ENDNOTES

[1] For example, datasets from Gastil (1979), the International Country Risk Guide (ICRG), the Transparency International (TI), Business Environmental Risk Intelligence (BERI) and Kaufmann, Kraay & Mastruzzi (2003 and 2008) are now available.

[2] See also North & Thomas (1973).

[3] This specification overlooks the influences of physical capital on knowledge production. The rationale supporting this formulation is the idea that knowledge production is more human capital intensive. See Tebaldi & Elmslie (2008) for more detail.

[4] It is worth noticing that equation 2 differs from Griliches (1990) and Jaffe & Lerner (2002), which specify an equation for patents as a function of R&D expenditure and previous production of patents (proxy for existing stock of knowledge). The use of personnel engaged in R&D instead of R&D expenditure represents a change of the focus (from input costs to the input itself), but as shown by Griliches (1994), these variables are highly correlated and should contain comparable information about the production process.

[5] This empirical equation is specified in terms of T rather than $ln(T)$ because i) this study uses standardized measures of institutions, which have zero mean and assume negative values, preventing the use of logarithmic transformations and ii) the true measure of institutions is not known, so T is used as a proxy for $ln(T)$. This procedure has become standard in the literature (e.g. Dollar & Kraay, 2003; Acemoglu *et al*., 2001; La Porta *et al*., 1999; Hall & Jones, 1999).

[6] The measures of institutions are mainly assembled by private companies (e.g. Transparency International, Political Risk Services) and based on an assessment of *perception*. These companies conduct perception surveys of "economic agents who make growth-relevant decisions" (Grogan & Moers, 2001, p. 326) about factors such as corruption, contract enforcement, protection of property rights, political instability, etc.

[7] For a detailed discussion regarding the impacts of colonial status and legal origin on current institutions see La Porta *et al.* (1999) and Acemoglu *et al.* (2001).

[8] This proposition is motivated by the work of Bernard Mandeville (early 1700), who argues that the development of institutions is an evolutionary process depending on generations of accumulated knowledge (Rosenberg, 1963). In

addition, a recent article by Glaeser *at al*. (2004) also shows that human capital positively impacts institutions, "even over a relatively short horizon of 5 years" (p. 296).

[9] Schmookler (1966), Griliches (1979) and Griliches (1984) are the pioneers in using patent count as a measure of innovation.

[10] It could be argued that even if two countries have the same number of innovations, the country with *weak institutions* would have fewer number of patents registered. However, it is important to notice that the USPTO patent data produce a valid measure of innovation in the context of weak institutions because the "US patent system can be viewed as exogenous to any developing country"(Chen & Puttitanun, 2005, p. 483).

[11] This proxy for innovation may be biased against countries whose inventors do not target the U.S. market or due to high transaction costs due to distance for example, which may affect an inventor's decision to patent a new invention in the U.S. Therefore, for the sake of robustness, we also estimate the model using patent data from the World Bank, which records data on patents granted by each country's patent office to its residents. The results are briefly discussed in the next section.

[12]Regulatory Quality "includes measures of the incidence of market-unfriendly policies such as price controls or inadequate bank supervision, as well as perceptions of the burdens imposed by excessive regulation in areas such as foreign trade and business development" (Kaufmann *et al.*, 2003, p. 3). Rule of Law includes "several indicators which measure the extent to which agents have confidence in and abide by the rules of society. These include perceptions of the incidence of crime, the effectiveness and predictability of the judiciary and the enforceability of contracts. Together, these indicators measure the success of a society in developing an environment in which fair and predictable rules form the basis for economic and social interactions and importantly, the extent to which property rights are protected" (Kaufmann *et al.*, 2003, p. 3). Control of Corruption "measures perceptions of corruption, conventionally defined as the exercise of public power for private gain…. The presence of corruption is often a manifestation of a lack of respect of both the corrupter (typically a private citizen or firm) and the corrupted (typically a public official or politician) for the rules which govern their interactions and hence represents a failure of governance according to our definition" (Kaufmann *et al.*, 2003, p. 4).

[13] Glaeser *at al.* (2004) argue that these measures of institutions (Risk of Expropriation, Control of Corruption, Rule of Law and Regulatory Quality (the last three are labeled as "government effectiveness") are actually *outcome* measures rather than ´deep´ measures of institutions. While this argument is correct, that is, the institutional

measures used in this paper are outcome measures, if this were not so; we would not be concerned about finding instruments for these measures (see Acemoglu *et al.* 2005b & 2005c for a detailed discussion on this issue).

[14] Sedgley (2006) proposes to estimate the initial stock of knowledge by utilizing the ratio of total patents issued during a specific period over the growth of the number of volumes in the collection of the national library during the same period. However, data limitation prevents the use of this procedure here.

[15] The logic supporting the idea that the stock of knowledge can be proxied by the country's share of world book production (new titles) is that a significant share of a society's knowledge is transmitted across individuals and from generation to generation through books, so that a highly knowledgeable society is expected to produce a large number of book titles and consequently should have a larger share of world book production. So, a country's share of world book production is used as a proxy for the starting stock of knowledge (concrete measure of the current knowledge stock) in the regression analysis. For robustness purposes, we use two measures: the share of applied and pure sciences book production, and the share of all new titles.

[16] The statistics include book production in applied sciences, pure sciences, social sciences, philosophy, arts, geography, history and generalities.

[17] Mitchell provides these statistics back to the eighteenth century only for a few countries. A representative cross-country sample can be only collected around 1920. Mitchell reports the number of children enrolled in primary and secondary schools for 68 countries in 1920 and statistics for 52 countries around the 1930s. Therefore, combining the actual 1920 data with estimates of the number of students enrolled in 1920 based upon the 1930 numbers allows one to get a sample comprised of 120 countries. We use the geometric growth rates in the estimations. For instance, if a country has data on enrollment between 1930 and 1940, the geometric growth rate between these periods is utilized to estimate enrollment back to 1920.

[18] Countries that experienced changes in their boundaries, such as the former USSR republics, Paraguay, Peru, Bolivia, India, Ivory Coast, Mali, Mauritania, Algeria and Zaire were not included in the regression analysis. In fact, these countries were not included in our analysis because of missing data or simply because they did not exist back to the beginning of the 20th century.

[19] A detailed discussion of the determinants of institutions can be found in Acemoglu *et al.* (2001 & 2005a), Acemoglu & Johnson (2005) and La Porta *et al.* (1999).

[20] This was tested using both a linear and a quadratic specification for population density. In addition, there is a strong correlation between schooling density and population density, which suggests that a potential impact of agglomeration on institutions may have already been accounted for by including the schooling density variable.

[21] This finding is consistent with Acemoglu *et al. (*2001).

[22] If the disturbances of the model are uncorrelated and homoscedastic, then the GMM estimation reduces simply to 2SLS (for more detail see Davidson & MacKinnon, 1993).

[23] In a set of regressions not reported in the paper, we tested an alternative measure of book production. More precisely, we use the share of applied and pure science book production instead of the share of all new titles produced. There are no significant quantitative differences compared to those results reported in Tables 5-8, which corroborates the model's specification. However, the coefficient on applied and pure science book production is found to be positive but insignificant at the 5 % significance levels in the regressions for the Risk of Expropriation. The significance and sign of the other coefficients are not affected.

[24] Due to data availability we use the accumulate number of patents granted from 1995 to 2001.

[25] The authors will gladly provide the estimates using the World Bank patent count upon request.

[26] McArthur & Sachs (2001) refer to this effect as "disease ecology".

[27] In a set of regressions not reported, mean temperature was substituted for absolute latitude. The results are similar, that is, the coefficients of absolute latitude turned out to be insignificant in all regressions.

[28] OECD members as of 1990 including Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, United Kingdom, and United States. Iceland and Luxembourg were not included due to missing data.

REFERENCES

Acemoglu, D. & Johnson, S. (2005). Unbundling institutions. *Journal of Political Economy*, *113*(5), 949-95.

Acemoglu, D., Johnson, S. & Robinson, J. A. (2001). The colonial origins of comparative development: an empirical investigation. *American Economic Review*, *91*(5), 1367-1401.

Acemoglu, D., Johnson, S. & Robinson, J. A. (2005a). Institutions as the fundamental cause of long-run growth. In: P. Aghion & S. Durlauf (eds.), *Handbook of Economic Growth (pp. 385-472).* North Holland: Elsevier.

Acemoglu, D., Johnson, S., Robinson, J. A. & Yared, P. (2005b). From education to democracy? *NBER Working Paper 11204*. http://www.nber.org/papers/w11204.

Acemoglu, D., Johnson, S., Robinson, J. A. & Yared, P. (2005c). Income and democracy. *NBER Working Paper 11205*. http://www.nber.org/papers/w11205.

Aghion, P. & Howitt, P. (1992). A Model of growth through creative destruction. *Econometrica*, *60*(2), 323-351.

Alcala, F. & Ciccone, A. (2004). Trade and productivity. *Quarterly Journal of Economics*, *119*(2), 613-46.

Anderson, T. W. (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions. *Annals of Mathematical Statistics*, *22*, 327-51.

Atkinson, G. (1998). The evolutionary theory of the development of property and the state. In: Samuels, W. *The Foundation of Institutional Economics*, New York: Routledge.

Barro, R. (1991). Economic growth in a cross section of countries. *Quarterly Journal of Economics*, *106*(2), 407-43.

Chen, Y. & and Puttitanun, T. (2005). Intellectual property rights and innovation in developing countries. *Journal of Development Economics, 78*(2), 474-93.

Davidson, R. & MacKinnon, J. G. (1993). *Estimation and inference in econometrics*. New York: Oxford University Press.

Dollar, D. & Kraay, A. (2003). Institutions, trade and growth. *Journal of Monetary Economics*, *50*, 133-162.

Easterly, W. & Levine, R. (2003). Tropics, germs, and crops: how endowments influence economic development. *Journal of Monetary Economics, 50,* 3-39.

Engerman, S. L. & Sokoloff, K. L. (1997). Factor endowments, institutions and differential paths of growth among new world economies: a view from economic historians of the United States. In: S. Haber (Ed.), *How Latin America Fell Behind* (pp. 260-292). Stanford.

Engerman, S. L. & Sokoloff, K. L. (2005). Institutions and non-institutional explanations of economic differences. In: Ménard, C. & Shirley, M. M. (Eds.), *Handbook of New Institutional Economics*. Dordrecht and New York: Springer.

Esfahani, H. S. & Ramírez, M. T. (2003). Institutions, infrastructure and economic growth. *Journal of Economic Development, 70,* 443-447.

Fedderke, J. (2001). Growth and institutions. *Journal of International Development, 13,* 645-670.

Furubotn, E. G. & Richter, R. (2005). *Institutions and economic growth***.** Ann Arbor: The University of Michigan Press (2nd ed.).

Gallup, J. L., Sachs, J. D. & Mellinger, A. D. (1999). Geography and economic development. *CID Working Paper No. 1.*

Gastil, R. D. (1979). *Freedom in the world: political rights and civil liberties.* New York: Freedom House.

Glaeser, E. L., La Porta, R. Lopes-de-Silanes, F. & Shleifer, A. (2004). Do institutions cause growth*? Journal of Economic Growth*, *9*(1), 271-303.

Gradstein, M. (2004). Governance and growth, *Journal of Development Economics*, *73,* 505– 518.

Griliches, Z. (1979). Issues in assessing the contribution of R&D to productivity growth. *Bell Journal of Economics*, *10*(1), 92-116.

Griliches, Z. (1990). Patent statistics as economic indicators: a survey. *Journal of Economic Literature*, *28*(4), 1661-1770.

Griliches, Z. (1994). Productivity, R&D and the data constraint. *American Economic Review*, *84*(1), 1-23.

Griliches, Z. (ed) (1984). *R&D, Patents and productivity*. Chicago: University of Chicago Press.

Grogan,L. & Moers, L. (2001). Growth empirics with institutional measures for transition countries. *Economic Systems*, *25*(4), 323-344.

Grossman, G. M. & Helpman, E. (2001). *Innovation and growth in the global economy*. Cambridge, MA: The MIT Press. Seventh printing.

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

Hansen, L. (1982). Large Sample properties of generalized method of moments estimators. *Econometrica*, *50*, 1029-1054.

Huang, H. & Xu, C. (1999). Institutions, innovations and growth. *American Economic Review*, 89(2), 438-443.

Jaffe, A. B. & Lerner, J. (2002). Reinventing public R&D: patent citations on the impact of nasa and other federal las on commercial innovation. In: A. B. Jaffe & M. Trajtenberg (Eds.) *Patents, Citations & Innovations* (pp. 287-335). Cambridge, MA: The MIT Press.

Jaffe, A. B. & Trajtenberg, M. (2002). *Patents, citations & innovations***.** Cambridge, MA: The MIT Press.

Jones, C. (1995). R&D-based models of economic growth. *Journal of Political Economy*, *103*:4: 759-784.

Kaufmann, D., Kraay. A. & Mastruzzi, M. (2003). Governance matters iii: governance indicators for 1996–2002. *Policy Research Working Paper No. 3106.*

Kaufmann, D., Kraay. A. & Mastruzzi, M. (2008). Governance matters vii: aggregate and individual governance indicators, 1996-2007. *World Bank Policy Research Working Paper No. 4654.*

Knack, S. & Keefer, P. (1995). Institutions and economic performance: cross-country tests using alternative institutional measures. *Economics and Politics*, *7*(3), 207-227.

Knack, S. & Keefer, P. (1997). Does social capital have an economic payoff? a cross-country investigation. *Quarterly Journal of Economics*, *112*(4), 1251-*1288.*

Kormendi, R. C. & Meguire, P. G. (1985). Macroeconomic determinants of growth. *Journal of Monetary Economics*, *16*, 141-163.

La Porta, R., FLopes-de-Silanes, F., Shleifer, A. & Vishny, R. (1999). The quality of government. *Journal of Law, Economics, & Organization, 15*(1), 222-279.

Lipset, S. M. (1960). *Political man: the social basis of modern politics*. New York: Doubleday.

Matthews, R.C.O. (1986). The Economics of Institutions and the Sources of Growth. *Economic Journal*, *96*, 903-918.

Mauro, P. (1995). Corruption and growth. *Quarterly Journal of Economics,* 110(3), 681-712.

McArthur, J. W. & Sachs, J. D. (2001). Institutions and geography: comment on Acemoglu, Johnson and Robinson (2000). *NBER*, Working Paper 8114.

Mitchell, B. R. (2003a). *International historical statistics: Africa, Asia & Oceania, 1750-2001***.** New York: Palgrave Macmillan (fourth ed).

Mitchell, B. R. (2003b). *International historical statistics Europe 1754-2000: Europe, 1750-2000.* New York: Palgrave Macmillan (fifth ed.).

Mitchell, B. R. (2003c). *International historical statistics: The Americas 1750-2000.* New York: Palgrave Macmillan (fifth ed.).

North, D. C. & Thomas, R. P. (1973). *The Rise of the western world: a new economic history.* Cambridge University Press.

North, D.C. (1990). *Institutions, institutional change and economic performance*. New York, Cambridge: University Press.

Oliva, M. A. & Rivera-Batiz, L. A. (2002). Political institutions, capital flows and developing country growth: an empirical investigation. R*eview of Development Economics*, *6*(2), 248-262.

Rodrik, D. (2000). Institutions for high-quality growth: what they are and how to acquire them. *NBER Working Paper 7540*.

Romer, P. (1990). Endogenous technological change. *Journal of Political Economy, 98*, S71-S102.

Romer, P. (2002). *Preface*. In: A. B. Jaffe & M. Trajtenberg (Eds.) *Patents, citations & innovations* (pp. 287-335). Cambridge, MA: The MIT Press.

Rosenberg, N. (1963). Mandeville and Laissez-Faire. *Journal of the History of Ideas, 24* (2), 183-196.

Sachs, J. D. (2000). *Tropical underdevelopment***.** Center for International Development, Harvard University.

Sala-i-Martin, X. (2002). 15 years of new growth economics: what have we learnt? *Central Bank of Chile Working Paper 172*, 22p.

Schmoolkler, J. (1966). *Invention and economic growth*. Harvard University Press.

Schneider, P. H. (2005). International trade, economic growth and intellectual property rights: A panel data study of developed and developing countries. *Journal of Development Economics*, *78*(2), 529-547.

Sedgley, N. (2006). A time series test of innovation-driven endogenous growth. *Economic Inquiry*, *44*(2), 318-332.

Stern, S., Porter, M.E, & Furman, J. (2002). The determinants of national innovative capacity. *Research Policy*, *31*, 899–933.

Tebaldi, E. & Elmslie, B. (2008). Institutions, innovation and economic growth. *Journal of Economic Development*, *33*(2): 1-27.

UNESCO Statistical Yearbook (Several years). *UNESCO Institute for Statistics*. Quebec, Canada.

United States Patent and Trademark Office – USPTO. (Several years). *Patent counts by country/state and year all patents, all types*. Washington, DC.

TABLES

Table 1: Simple Correlation of Institutional Measures and Patent Count

Institutional Measure	*Patent Count 1970-2003 USPTO*	*Patent Count, 1995-2001 World Bank*
Regulatory Quality[a]	0.60	0.42
Rule of Law[a]	0.68	0.58
Control of Corruption[a]	0.69	0.56
Risk of Expropriation[b]	0.80	0.76
Average Institutional Index[c]	0.72[a]	0.55[b]

Source: Author's calculations; [a] number of countries=133; [b] number of countries =85, [c] the *average institutional index* was calculated as the average of Regulatory quality, Rule of Law, Control of corruption, Voice and Accountability, Political Stability, and Government Effectiveness .

Table 2: OLS Regressions of the Determinants of Institutions (Control of Corruption and Rule of Law)

Explanatory Variables	Control of Corruption				Rule of law			
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
ln Human Capital Density in the early 20th century	0.0440	0.0332	0.0900***	0.0424	0.0547*	0.0598	0.0920***	0.0503*
	[1.33]	[0.76]	[2.83]	[1.27]	[1.89]	[1.56]	[3.49]	[1.73]
Legal Origin – Socialist	-0.983***	-0.997***	-0.961***	-0.950***	-1.027***	-1.003***	-0.987***	-0.991***
	[-4.37]	[-4.15]	[-4.20]	[-4.21]	[-4.53]	[-4.15]	[-4.39]	[-4.35]
Legal Origin – French	-0.356**	-0.329**	-0.351**	-0.337**	-0.387***	-0.376**	-0.381**	-0.371**
	[-2.50]	[-2.24]	[-2.33]	[-2.36]	[-2.74]	[-2.57]	[-2.62]	[-2.61]
Legal Origin – German	0.0894	0.0563	0.328	0.0858	0.277	0.303	0.481	0.275
	[0.27]	[0.16]	[0.97]	[0.26]	[0.84]	[0.87]	[1.44]	[0.83]
Legal Origin – Scandinavian	0.221	0.202	0.678*	0.263	0.0311	0.0191	0.412	0.0689
	[0.57]	[0.51]	[1.72]	[0.67]	[0.080]	[0.048]	[1.07]	[0.18]
Prop. land within 100 km of the sea coast	0.620***	0.657***	0.318	0.660***	0.544**	0.600**	0.279	0.601***
	[2.73]	[2.75]	[1.43]	[2.90]	[2.46]	[2.57]	[1.34]	[2.73]
Absolute Latitude	1.010	1.012	2.770***		1.069*	1.069*	2.631***	
	[1.61]	[1.60]	[6.31]		[1.72]	[1.69]	[6.49]	
Mean Temperature	-0.0708***	-0.0731***		-0.0932***	-0.0612***	-0.0614***		-0.0855***
	[-3.82]	[-3.86]		[-7.54]	[-3.32]	[-3.26]		[-7.13]
ln Population Density in the early 20th century		0.0220				-0.0169		
		[0.31]				[-0.25]		
Constant	1.244**	1.349**	-0.507**	1.938***	1.112**	1.022*	-0.412**	1.849***
	[2.47]	[2.18]	[-2.38]	[7.37]	[2.22]	[1.70]	[-2.12]	[7.06]
Sample size	104	100	107	104	110	106	114	110
Adjusted R-squared	0.67	0.67	0.62	0.67	0.66	0.65	0.62	0.65

Notes: *** p<0.01, ** p<0.05, * p<0.1; t-ratios are reported between parentheses; all regressions were estimated with standard errors robust to arbitrary heteroskedasticity.

Table 3: OLS Regressions of the Determinants of Institutions (Regulatory Quality and Risk of Expropriation)

Explanatory Variables	Regulatory Quality				Risk of Expropriation				Average Institutional Quality			
	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14	Model 15	Model 16	Model 17	Model 18	Model 19	Model 20
ln Human Capital Density in the early 20th century	0.0783**	0.0801*	0.126***	0.0817**	0.164**	0.147	0.265***	0.165**	0.0821**	0.0887**	0.126***	0.0813**
	[2.30]	[1.73]	[4.05]	[2.41]	[2.30]	[1.60]	[3.85]	[2.31]	[2.56]	[2.10]	[4.16]	[2.54]
Legal Origin – Socialist	-0.853***	-0.841***	-0.756***	-0.881***	-1.064**	-1.109**	-0.786	-1.069**	-0.784***	-0.770***	-0.733***	-0.769***
	[-3.27]	[-2.99]	[-2.89]	[-3.39]	[-2.04]	[-2.00]	[-1.46]	[-2.06]	[-3.61]	[-3.31]	[-3.37]	[-3.57]
Legal Origin - French	-0.0972	-0.0843	-0.117	-0.110	-0.562*	-0.519*	-0.558*	-0.566*	-0.209	-0.198	-0.206	-0.201
	[-0.60]	[-0.50]	[-0.69]	[-0.68]	[-1.86]	[-1.68]	[-1.75]	[-1.89]	[-1.53]	[-1.39]	[-1.44]	[-1.47]
Legal Origin - German	-0.0910	-0.0814	0.142	-0.0900	0.262	0.181	0.708	0.264	0.106	0.116	0.302	0.104
	[-0.24]	[-0.20]	[0.37]	[-0.24]	[0.38]	[0.25]	[0.98]	[0.38]	[0.34]	[0.34]	[0.93]	[0.33]
Legal Origin - Scandinavian	-0.0173	-0.0250	0.416	-0.0470	-0.251	-0.361	0.646	-0.259	0.116	0.0893	0.508	0.135
	[-0.039]	[-0.054]	[0.93]	[-0.10]	[-0.30]	[-0.43]	[0.78]	[-0.32]	[0.31]	[0.23]	[1.35]	[0.36]
Prop. land within 100 km of the sea coast	0.521**	0.558**	0.224	0.475*	0.258	0.400	-0.376	0.251	0.475**	0.535**	0.196	0.493**
	[2.05]	[2.07]	[0.93]	[1.89]	[0.54]	[0.80]	[-0.80]	[0.53]	[2.17]	[2.31]	[0.92]	[2.27]
Absolute Latitude	-0.839	-0.838	1.261***		-0.200	-0.130	3.338***		0.461	0.488	2.032***	
	[-1.18]	[-1.15]	[2.65]		[-0.15]	[-0.098]	[3.58]		[0.76]	[0.79]	[4.87]	
Mean Temperature	-0.0781***	-0.0784***		-0.0591***	-0.142***	-0.148***		-0.138***	-0.0626***	-0.0624***		-0.0728***
	[-3.66]	[-3.56]		[-4.25]	[-3.60]	[-3.68]		[-5.26]	[-3.50]	[-3.40]		[-6.17]
ln Population Density in the early 20th century		-0.00627				0.0283				-0.0163		
		[-0.079]				[0.19]				[-0.24]		
Constant	1.928***	1.890***	-0.0572	1.350***	10.30***	10.44***	6.810***	10.16***	1.224**	1.130*	-0.319	1.541***
	[3.34]	[2.66]	[-0.25]	[4.47]	[9.61]	[7.96]	[15.1]	[18.2]	[2.52]	[1.88]	[-1.57]	[6.14]
Sample size	109	105	113	109	101	97	103	101	104	100	107	104
Adjusted R-squared	0.46	0.46	0.40	0.46	0.52	0.53	0.47	0.52	0.63	0.63	.058	0.63

Notes: *** p<0.01, ** p<0.05, * p<0.1; t-ratios are reported between parentheses; all regressions were estimated with standard errors robust to arbitrary heteroskedasticity.

Table 4: IV Regressions of accumulated Patent Counts on Control of Corruption

Explanatory Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Control of Corruption	1.527***	1.770***	2.069***	1.781***	1.722***	2.118***	1.764***
	[9.06]	[9.84]	[5.22]	[9.90]	[10.5]	[5.83]	[10.4]
ln Personnel engaged in R&D	0.487***	0.491***	0.509***	0.491***	0.472***	0.476***	0.463***
	[4.27]	[5.19]	[5.16]	[4.90]	[5.08]	[4.91]	[4.86]
ln Share of book production	0.553***	0.467***	0.452***	0.461***	0.494***	0.475***	0.499***
	[4.17]	[4.32]	[3.67]	[4.23]	[4.81]	[4.08]	[4.87]
Prop. land within 100 km of the sea coast			-0.226			-0.409	
			[-0.41]			[-0.82]	
Mean Temperature			0.0414			0.0471	
			[0.93]			[1.17]	
Population (aged 15-64) density				0.317			-0.0672
				[0.073]			[-0.019]
Population (aged 15-64) density Squared				-0.148			0.0717
				[-0.064]			[0.038]
Constant	3.286**	2.815**	1.844	2.771**	3.021***	2.085	3.118***
	[2.31]	[2.44]	[1.21]	[2.35]	[2.69]	[1.48]	[2.77]
Method	OLS	2SLS	2SLS	2SLS	GMM	GMM	GMM
Observations	76	76	76	76	76	76	76
R-squared	0.86	0.85	0.84	0.85	0.85	0.84	0.85
Hansen J statistic	-	6.557	5.669	7.258	6.557	5.669	7.258
p-value	-	0.256	0.129	0.202	0.256	0.129	0.202
Anderson-Canonical Correlation Test	-	77.58	24.90	87.66	77.58	24.90	87.66
p-value AR	-	0	0.0001	0	0	0.0001	0

Notes: The dependent variable is the natural logarithm of the USPTO patent count between 1970 and 2003. *** p<0.01, ** p<0.05, and * p<0.1. t-ratios are reported between parentheses; all regressions were ran with standard errors robust to arbitrary heteroskedasticity. All IV and GMM regressions are estimated using the following set of instruments for institutions: ln human capital density in the early 20th century, dummies for the origin of the legal system, mean temperature, and proportion of land within 100 km of the seacoast.

Table 5: IV Regressions of Accumulated Patent Counts on Rule of Law

Explanatory Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Rule of Law	1.644***	1.890***	2.098***	1.928***	1.889***	2.323***	1.947***
	[9.39]	[10.6]	[5.26]	[10.7]	[11.2]	[6.33]	[11.4]
ln Personnel engaged in R&D	0.412***	0.405***	0.408***	0.414***	0.420***	0.419***	0.413***
	[3.69]	[4.46]	[4.43]	[4.26]	[4.75]	[4.57]	[4.50]
ln Share of book production	0.601***	0.527***	0.526***	0.522***	0.510***	0.494***	0.520***
	[4.70]	[4.62]	[4.12]	[4.55]	[4.76]	[4.04]	[4.83]
Prop. land within 100 km of the sea coast			-0.143			-0.474	
			[-0.27]			[-0.94]	
Mean Temperature			0.0281			0.0505	
			[0.63]			[1.23]	
Population (aged 15-64) density				-1.736			-1.749
				[-0.44]			[-0.50]
Population (aged 15-64) density Squared				0.997			1.013
				[0.48]			[0.54]
Constant	4.019***	3.689***	3.135**	3.655***	3.441***	2.516*	3.576***
	[2.92]	[3.22]	[2.18]	[3.10]	[3.13]	[1.88]	[3.23]
Method	OLS	2SLS	2SLS	2SLS	GMM	GMM	GMM
Observations	76	76	76	76	76	76	76
R-squared	0.86	0.86	0.85	0.86	0.86	0.84	0.86
Hansen J statistic	-	5.328	4.361	5.699	5.328	4.361	5.699
p-value	-	0.377	0.225	0.337	0.377	0.225	0.337
Anderson-Canonical Correlation Test	-	77.62	27.59	83.02	77.62	27.59	83.02
p-value AR	-	0	0	0	0	0	0

Notes: The dependent variable is the natural logarithm of the USPTO patent count between 1970 and 2003. *** p<0.01, ** p<0.05, and * p<0.1. t-ratios are reported between parentheses; all regressions were ran with standard errors robust to arbitrary heteroskedasticity. All IV and GMM regressions are estimated using the following set of instruments for institutions: ln human capital density in the early 20th century, dummies for the origin of the legal system, mean temperature, and proportion of land within 100 km of the seacoast.

Table 6: IV Regressions of Accumulated Patent Counts on Regulatory Quality

Explanatory Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Regulatory Quality	1.614***	2.051***	2.053***	2.018***	2.128***	2.311***	2.222***
	[8.37]	[9.33]	[3.42]	[8.76]	[10.2]	[3.98]	[10.4]
ln Personnel engaged in R&D	0.420***	0.410***	0.402***	0.423***	0.432***	0.419***	0.453***
	[3.54]	[3.62]	[3.58]	[3.66]	[4.04]	[3.99]	[4.16]
ln Share of book production	0.791***	0.710***	0.678***	0.717***	0.710***	0.633***	0.690***
	[6.15]	[6.06]	[5.60]	[6.16]	[6.48]	[5.38]	[6.36]
Prop. land within 100 km of the sea coast			-0.252			-0.769	
			[-0.34]			[-1.08]	
Mean Temperature			-0.0138			0.000216	
			[-0.27]			[0.0045]	
Population (aged 15-64) density				-1.193			-4.867
				[-0.29]			[-1.45]
Population (aged 15-64) density Squared				0.792			2.659
				[0.37]			[1.50]
Constant	4.720***	4.311***	4.606***	4.290***	4.067***	4.124***	3.951***
	[3.25]	[3.04]	[2.81]	[3.00]	[3.04]	[2.78]	[2.98]
Method	OLS	2SLS	2SLS	2SLS	GMM	GMM	GMM
Observations	76	76	76	76	76	76	76
R-squared	0.84	0.83	0.83	0.83	0.83	0.82	0.82
Hansen J statistic	-	8.934	6.033	7.193	8.934	6.033	7.193
p-value	-	0.112	0.110	0.207	0.112	0.110	0.207
Anderson-Canonical Correlation Test	-	56.71	17.28	64.75	56.71	17.28	64.75
p-value AR	-	0	0.002	0	0	0.002	0

Notes: The dependent variable is the natural logarithm of the USPTO patent count between 1970 and 2003. *** p<0.01, ** p<0.05, and * p<0.1. t-ratios are reported between parentheses; all regressions were ran with standard errors robust to arbitrary heteroskedasticity. All IV and GMM regressions are estimated using the following set of instruments for institutions: ln human capital density in the early 20th century, dummies for the origin of the legal system, mean temperature, and proportion of land within 100 km of the seacoast.

Table 7: IV Regressions of Accumulated Patent Counts on Expropriation Risk

Explanatory Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Expropriation Risk	0.965***	1.421***	1.316***	1.433***	1.411***	1.226***	1.459***
	[7.68]	[7.41]	[3.48]	[7.61]	[7.63]	[3.30]	[8.10]
ln Personnel engaged in R&D	0.313**	0.241**	0.258**	0.236*	0.292***	0.305***	0.272**
	[2.49]	[2.08]	[2.52]	[1.90]	[2.66]	[3.06]	[2.31]
ln Share of book production	0.605***	0.372**	0.418**	0.358**	0.368**	0.425**	0.336**
	[4.20]	[2.30]	[2.40]	[2.21]	[2.34]	[2.47]	[2.12]
Prop. land within 100 km of the sea coast			0.529			0.592	
			[0.87]			[1.00]	
Mean Temperature			-0.00423			-0.0188	
			[-0.092]			[-0.42]	
Population (aged 15-64) density				1.105			0.868
				[0.26]			[0.21]
Population (aged 15-64) density Squared				-0.518			-0.378
				[-0.22]			[-0.17]
Constant	-1.715	-5.434**	-4.771	-5.592**	-5.848***	-4.243	-6.199***
	[-0.94]	[-2.36]	[-1.18]	[-2.48]	[-2.66]	[-1.07]	[-2.88]
Method							
Observations	75	75	75	75	75	75	75
R-squared	0.83	0.80	0.81	0.80	0.80	0.82	0.79
Hansen J statistic	-	10.13	9.677	10.67	10.13	9.677	10.67
p-value	-	0.0716	0.0215	0.0583	0.0716	0.0215	0.0583
Anderson-Canonical Correlation Test	-	38.78	11.13	40.72	38.78	11.13	40.72
p-value AR	-	0	0.025	0	0	0.0252	0

Notes: The dependent variable is the natural logarithm of the USPTO patent count between 1970 and 2003. *** p<0.01, ** p<0.05, and * p<0.1. t-ratios are reported between parentheses; all regressions were ran with standard errors robust to arbitrary heteroskedasticity. All IV and GMM regressions are estimated using the following set of instruments for institutions: ln human capital density in the early 20th century, dummies for the origin of the legal system, mean temperature, and proportion of land within 100 km of the seacoast.

Table 8: IV Regressions of Accumulated Patent Counts on Average Institutional Quality

Explanatory Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Average Institutional Quality	1.780***	2.022***	2.695***	2.053***	1.989***	2.962***	2.059***
	[9.55]	[11.1]	[4.76]	[11.1]	[11.7]	[6.06]	[11.9]
ln Personnel engaged in R&D	0.430***	0.426***	0.430***	0.443***	0.436***	0.493***	0.477***
	[3.88]	[4.31]	[4.13]	[4.26]	[4.61]	[4.74]	[4.78]
ln Share of book production	0.622***	0.558***	0.520***	0.555***	0.567***	0.449***	0.550***
	[4.96]	[5.16]	[3.91]	[5.21]	[5.47]	[3.55]	[5.38]
Prop. land within 100 km of the sea coast			-0.779			-0.974	
			[-1.11]			[-1.64]	
Mean Temperature			0.0652			0.085*	
			[1.15]			[1.76]	
Population (aged 15-64) density				-2.076			-4.403
				[-0.50]			[-1.23]
Population (aged 15-64) density Squared				1.306			2.555
				[0.60]			[1.35]
Constant	4.020***	3.720***	2.473	3.635***	3.704***	1.205	3.404***
	[2.95]	[3.12]	[1.43]	[2.99]	[3.26]	[0.77]	[2.93]
Method							
Observations	76	76	76	76	76	76	76
R-squared	0.86	0.86	0.84	0.86	0.86	0.81	0.86
Hansen J statistic	-	7.328	4.009	6.243	7.328	2.563	6.243
p-value	-	0.197	0.261	0.283	0.197	0.464	0.283
Anderson-Canonical Correlation Test	-	74.16	18.44	80.02	74.16	22.18	80.02
p-value AR	-	0	0.001	0	0	0.0002	0

Notes: The dependent variable is the natural logarithm of the USPTO patent count between 1970 and 2003. *** p<0.01, ** p<0.05, and * p<0.1. t-ratios are reported between parentheses; all regressions were ran with standard errors robust to arbitrary heteroskedasticity. All IV and GMM regressions are estimated using the following set of instruments for institutions: ln human capital density in the early 20th century, dummies for the origin of the legal system, mean temperature, and proportion of land within 100 km of the seacoast.

Table 9: Sensitivity Analysis

Explanatory Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Institution (Average Index)	1.869***	2.580***	1.841***	1.703***	1.686***	2.275***
	[7.70]	[4.21]	[4.09]	[3.52]	[3.15]	[10.4]
Institution Squared	0.226					
	[0.92]					
ln Personnel engaged in R&D	0.445***	1.173***	0.415***	0.412***	0.743***	0.381***
	[4.28]	[4.92]	[4.19]	[3.61]	[7.53]	[2.86]
ln Share of book production	0.527***	0.0869	0.606***	0.564***		0.530***
	[4.50]	[0.30]	[5.48]	[4.66]		[3.88]
OECD * Institution			-1.714			
			[-0.77]			
OECD			1.258			
			[0.86]			
ln gdp per capita				0.405	0.854*	
				[0.84]	[1.70]	
Constant	3.239**	-5.738*	4.036***	0.590	-8.267**	2.585*
	[2.31]	[-1.94]	[3.30]	[0.14]	[-1.98]	[1.66]
Method	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Sample Size	76	22	76	67	72	69
Centered R-Squared	0.86	0.87	0.86	0.87	0.84	0.81

Notes: The dependent variable in models 1 thru 4 is the natural logarithm of the USPTO patent count between 1970 and 2003 and the dependent variable in model 5 is the natural logarithm of the USPTO patent count between 1995 and 2001. *** p<0.01, ** p<0.05, and * p<0.1. t-ratios are reported between parentheses; all regressions were ran with standard errors robust to arbitrary heteroskedasticity. All 2SLS regressions are estimated using the following set of instruments for institutions: ln human capital density in the early 20th century, dummies for the origin of the legal system, mean temperature, and proportion of land within 100 km of the seacoast. For model 1, the squared measure of institutions are instrumented by using the set of instruments listed above plus the squared values of the following variables: ln human capital density in the early 20th century, mean temperature, and proportion of land within 100 km of the seacoast.

FIGURES

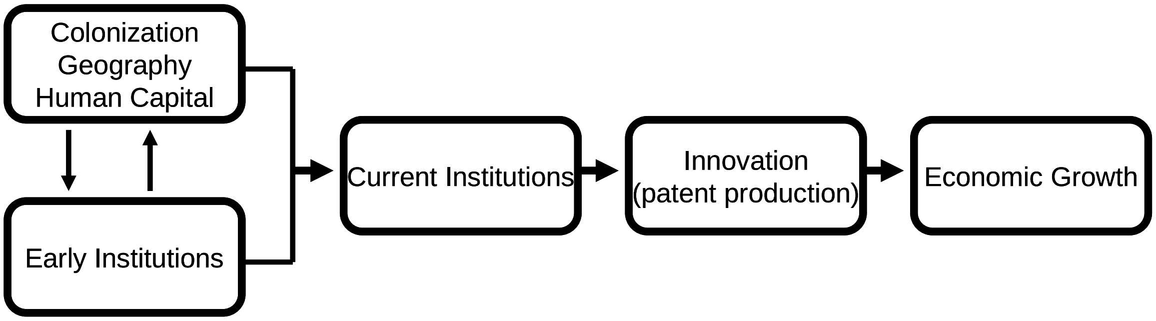

Figure 1: Institutions and Economic Growth

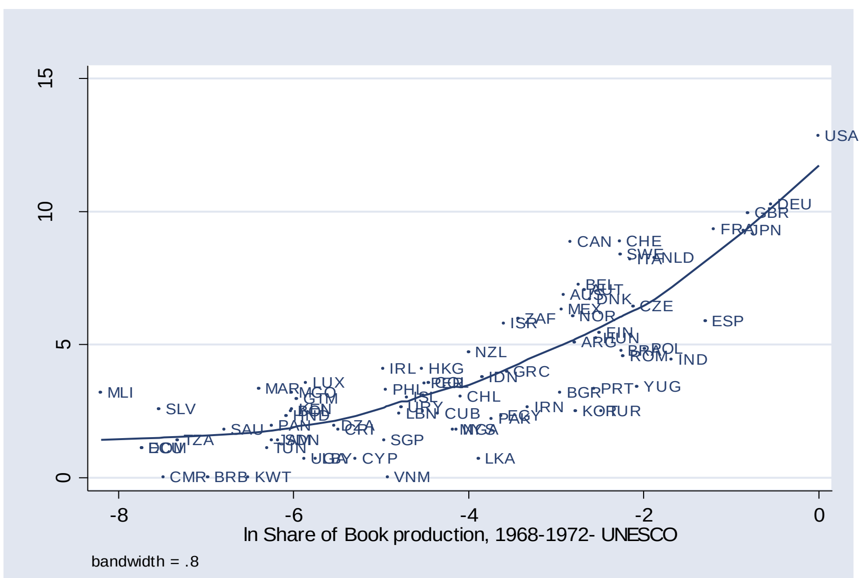

Figure 2: Book Production and Initial Stock of Patents