

Goel, Rajeev K.; Saunoris, James W.

**Article — Accepted Manuscript (Postprint)**

## The role of socio-historic factors and income inequality in global innovation

Economics of Innovation and New Technology

**Provided in Cooperation with:**

Kiel Institute for the World Economy – Leibniz Center for Research on Global Economic Challenges

*Suggested Citation:* Goel, Rajeev K.; Saunoris, James W. (2023) : The role of socio-historic factors and income inequality in global innovation, Economics of Innovation and New Technology, ISSN 1476-8364, Taylor & Francis, London, Vol. 32, Iss. 4, pp. 498-518, <https://doi.org/10.1080/10438599.2021.1959326>

This Version is available at:

<https://hdl.handle.net/10419/307992>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by-nc/4.0/>

## **The role of socio-historic factors and income inequality in global innovation**

**Rajeev K. Goel \***

Illinois State University and Kiel Institute for the World Economy

(E-mail: [rkgoel@ilstu.edu](mailto:rkgoel@ilstu.edu))

**James W. Saunoris**

Eastern Michigan University

(E-mail: [jsaunori@emich.edu](mailto:jsaunori@emich.edu))

---

\* Correspondence: Rajeev K. Goel, Department of Economics, Illinois State University, Normal, IL 61790-4200, USA. Comments of the Editor and referees are appreciated.

No potential competing interest was reported by the authors.

### **Abstract**

In order to better understand why some nations are more innovative than others, this paper considers the influence of socio-historic dimensions of colonialism and income inequality. Using data from 72 nations over the years 1997-2018, we find that ethnic income inequality (along with income inequality) and ethnolinguistic fractionalization of the population reduced the pace of innovation. These findings underscore the role of income distribution and other social factors in facilitating innovation. Furthermore, nations with a colonial past were more innovative, *ceteris paribus*, whereas nations with long histories (longevity) were no different from others. Finally, we found some support for presidential democracies fostering more innovation.

**Keywords:** innovation; patents; R&D; colonialism; technological change; genetic distance; fractionalization; income inequality; history; path dependence; presidential democracies

**JEL classification:** O33; O34; O38

## 1. Introduction

The pursuit of innovation has incentivized individuals and businesses for centuries as they tried to overcome personal challenges, improve standards of living, and outsmart competitors. Over time, governments have recognized the importance of innovation for self-sufficiency, global competitiveness, and growth. Yet, despite the global recognition of their importance, our understanding of the determinants of the variation across countries in terms of innovativeness is incomplete. This is evident anecdotally given the wide variations in the innovation (patenting) rates across nations. A part of this variation is due to different resource endowments across nations, institutions, and the inherent uncertainty with the pursuit of inventions. These factors, however, do not explain all of the story. It is possible that historical precedents (e.g., history of colonization) alter institutions (David (1994), Nelson (1993), North (1990), Whitley (2002)), resources and attitudes that set certain nations on (inert) paths that make them technologically dependent and relatively less innovative (David (1985, 2001), Mokyr (1992, 2010)).

This paper studies the drivers of innovation across nations, paying special attention to the historical factors. An understanding of the significant drivers of innovation would be important to understand why some nations are more innovative than others. Whereas recognition of historical precedents driving innovation is not new (Borcan et al. (2018), David (1985, 2001), Goel and Saunoris (2016), Mokyr (1992, 2010)), we include and compare the influences of some new measures (e.g., the genetic distance of a nation's elite, state history) that capture historical precedents and the social composition of nations in terms of their impact on innovativeness. Certain historical developments can put some nations on a course to greater innovativeness over time, while other nations might be held back due to their histories. Which historical precedents significantly impact innovation productivity across nations? Are nations with a colonial past or nations with long histories more innovative? Another contribution is to see the effects of the historical factors on nations with differing innovation rates. Do the factors that affect innovation similarly impact the most innovative nations and the least innovative nations?

The impacts of historical events on technological progress can conceivably be positive or negative. On the one hand, nations with extractive resources were historically attractive to colonizers and in such cases, the resources were exported resulting in relatively little domestic industrial development (see Acemoglu et al. (2001)). In fact, colonization likely killed some domestic industries (e.g., handloom manufacturing in India). Without significant domestic industrialization, colonizers had fewer incentives to develop/streamline institutions (e.g., patent laws, property protection laws in general). All this would inhibit innovation in colonies over time, even when they eventually became independent they would remain behind the technological frontier, relying on high-tech imports and unable to generally get a foothold into competitive international markets.

On the other hand, in some cases, colonizers themselves made colonies their home, sharing similar language and culture over time. This facilitated knowledge transfer (for example, through the formation of research joint ventures and associated knowledge spillovers). On the other hand, the discrimination against native languages and traditions by colonial rulers would likely be associated with a loss of useful knowledge, some of which would have led to useful innovations.

Furthermore, the non-extractive nature of resources in some colonies, e.g., weather, hydroelectric potential, etc., would again have prompted colonizers to transfer know-how and institutions to colonies (also see Acemoglu et al. (2001)). This would facilitate technical change.

Along another dimension, institutions in old nations (i.e., nations that have been in existence for a long time (as opposed to the newly formed transition nations, for example), whether colonized or otherwise, might evolve on their own, impacting the pace of innovation and technological change. Such nations would have long-standing institutions. Homegrown/evolved institutions (e.g., norms about property rights protection, etc.) can affect the incentives for innovation as well as technology transfer across nations.

Beyond the technical aspects (the complexity and transferability of technologies and know-how) and institutional aspects (David (1994)), the social aspects (e.g., ethnic and language composition of populations) are likely important in dictating how different nations are empowered over time to pursue or generate new technologies. Some of the measures of historical precedents that we use enable us to uniquely account for the socio-economic differences and examine their influence on innovation over time.

The distribution of income and innovation may be correlated across nations (Aghion et al. (2019), Antonelli and Gehringer (2017), Tselios (2011), Zweimüller (2000)). Income inequality can impact both the demand and supply of innovations. On the demand side, the rich and poor demand different classes of goods, with the poor not expressing demand for luxury products. This alters signals for demand-pull innovations. On the supply side, the presence of income inequalities likely affect the incentives for participation (and intensity of participation) in innovative activities. Furthermore, the income distribution in a nation might be rooted in its ethnic composition, with certain ethnicities having greater incomes (by being closer to the elite, for example) than others. The present research will uniquely examine the impact of ethnic inequality in terms of its impact on innovation.

Using the total patent grants per 100,000 population as the outcome variable and employing data from 72 nations over the years 1997-2018, we find that ethnic income inequality (along with income inequality), ethnolinguistic fractionalization of the population, and genetic distance reduced the pace of technological change. Quantitatively, a one percent increase in ethnic income inequality decreases the patents per capita by 1.2 percent. Furthermore, nations with a colonial past were more innovative, *ceteris paribus*, whereas nations with long histories (longevity) were no different from others. Finally, we found some support for presidential democracies fostering more innovation.

Besides adding to the literature, the findings are potentially useful for the design of effective technology policies. More broadly, the findings have implications for economic development and productivity (Acemoglu and Zilibotti (2001)) and the diffusion of knowledge (Antonelli and Link (2015)). If certain historical precedents turn out to significantly inhibit innovation in some nations, then policies could be altered to counter the latent effects to better foster innovation. Conversely, the inertia of certain historical aspects might limit the ability of policymakers to increase the pace of innovation, at least in the short and intermediate terms.

The organization of the rest of the paper includes a discussion of the literature and the model in the next section, followed by data and estimation, results, and conclusions.

## 2. Literature and Model

### 2.1 Literature

Nations' innovation systems might evolve due to historical precedents (Bockette et al. (2002), Borcan et al. (2018), Chanda and Putterman (2007), Freeman (1995), Nelson (1993)). History can shape institutions, including the development of the infrastructure supporting innovations, leading to a national innovation system (Freeman (1995), Nelson (1993)). An often-cited example is the QWERTY format of the keyboards, which have outlived their practicality in the electronic age but continue to be widely used (David (1985)). More broadly, Johnson (2014) cites six innovations (glass; refrigeration; recorded sound; water purification; clocks and artificial light) that led to the development of other useful technologies over time. This is especially pertinent given that technological change evolves over time (Antonelli (2009), Dosi and Nelson (2010), Nelson and Winter (1982), Ruttan (1997)). Technological change is often incremental, with new technologies seldom in a user-friendly form in initial innovation. As technologies evolve over time, they either build off earlier technologies or are intricately intertwined such that licensing of previous patents is required for further innovation and patents (Scotchmer (2004)). Besides the inertia, the underlying technological complexity might dictate the ability to stitch technologies in the future (Antonelli (2009)). The inert externalities from prior technology development and adoption can impact industry standardization over time, increasing switching costs to (better) future technologies (David (1985)). De Liso and Filatrella (2011) note that a number of factors, such as technology users' reluctance to switch to the new one, the degree of diffusion of the old technology, etc. contribute to slowing down the process of substitution/adoption of new for old technologies.

Within this landscape, historical events such as colonialization could crucially impact technological progress by either preventing or enabling access to complementary technologies. Alternatively stated, the prevalence and pace of knowledge spillovers can in part be determined by historical events (see Bockstette et al. (2002), Chanda and Putterman (2007)). Given the inertia or latency with certain technologies, such impacts can have lasting effects in setting nations on different trajectories of technological change. The literature on path dependence in technology is formidable – see Arthur (1994), Chandler (1992), David (1985, 1994, 2001), De Liso and Filatrella (2011), Fagerberg et al. (2008), Simme et al. (2008), Vergne and Durand (2010)). Studying the impact of colonialism on R&D spending in a different sample of nations over 1996-2012, Goel and Saunoris (2016) found the impact to be negative, with somewhat mixed statistical significance.

Beyond direct impacts on technological change and knowledge spillovers, historical precedents can shape institutions, which in turn affect technological change and innovation. Colonies inherit legal systems from colonizers, for example, which affect norms regarding property rights protection, among other things. In a well-cited work, Acemoglu et al. (2001), examined the

influence of colonial origins on the relative economic development of nations. The present research can be viewed in that context, with the focus here being on innovation productivity.

Institutions are accounted for by the degree of patent protection and via the form of democracy. In the political-economy context, this research ties to the role of the form of government by considering whether presidential democracies significantly influence the pace of technical change. It has been argued that presidential systems are politically more stable (by not having mid-term elections like parliamentary systems) and more efficient (by not encouraging the formation of political coalitions that slow down the legislative process in parliamentary democracies), (see Gerring et al. (2009), Kaminsky (1997)). Broadly speaking, the inclusion of presidential forms of government ties to the literature on the quality of government (La Porta et al. (1999)). Goel and Nelson (2020) considered different forms of government and their impacts on female entrepreneurs (also see Bjørnskov and Rode (2020)). They found that female owners of businesses prospered under presidential democracies.

Along another dimension, although not novel, this research ties to the literature on the link between R&D and patents, or the input and output of the research process.<sup>1</sup> Given the gestation period, both technological and institutional (David (1994)), between R&D and patenting, the impact of research on patents is felt over time (Hall et al. (1986)).<sup>2</sup> We take account of this aspect in the model we estimate.

The role of the historic and social environment in technical development has been noted by Mokyr (1992); also see Mokyr (2010)). Particularly, Mokyr (1992) examines the role of resistance in the French and British economies during technological progress related to the Industrial Revolution. Certain past events might influence/alter the pace and course of innovation. For example, as noted by Taalbi (2017), the energy and economic crises of the 1970s induced problem-driven innovation, whereas new legislations and policy interventions contributed to environmental innovation. More generally, building on the seminal contributions of Schumpeter, Antonelli (2015) notes that innovation is a creative response to changing contexts. In this context, the current internet and driver-less car revolutions are generating many innovations, and more related aspects will unfold in the coming years.

Income inequality or the distribution of income can impact innovation, and, in turn, innovation might also alter income inequality (Antonelli and Gehringer (2017)). The underlying idea is the poor would mainly demand basic goods, while the rich also demand luxury goods. The demands for these goods, contingent on the underlying distribution of income would affect relative demand-pull innovation (see, for example, Zweimüller (2000)). The present study, by considering the impact of role of the ethnic composition of a nation in its income distribution, further allows for the possibilities how demands for different goods might be different across ethnicities with different income levels. The literature linking innovation and income inequality has considered the direction of causality, both from income inequality to innovation (Tselios

---

<sup>1</sup> Patents are a useful and readily available indicator of innovation output. However, they are not a perfect indicator, suffering from the drawbacks of not being able to capture unpatented and unpatentable innovations.

<sup>2</sup> Hall et al. (1986) note that the composition of R&D (e.g., between basic and applied research) can vary between large and small firms and this might impact R&D's effect on patents.



(2011) and the reverse (Aghion et al. (2019)). In general, the correlation between income inequality and innovation is found to be positive, although the relation is sensitive to the measure/scope of income inequality (Aghion et al. (2019)). Our research tests the income inequality - innovation nexus by adding the social or ethnic dimension.

A discussion of the empirical model follows.

## 2.2 Model

We pose a simple linear model of innovation that draws on the extant literature and then focuses on the relative aspects of different historical dimensions. As discussed above, the impacts of various historical aspects on innovation are not clear a priori and we expect that the formal empirical analysis will better inform us. With the data at the country and year level, the general form of the estimated equation is:

$$\text{Innovation (PATENTS)} = f(\text{R\&D}_{-1}, \text{History/Social}_M, \text{Institutions}_X, Z_Y) \quad \dots(1)$$

$M = \text{EthnicINQ, IncomeINQ, EthnicFRAC, EthLingFRAC, COLONY, StateHIST, GeneticDIST}$

$X = \text{PatentRights, Presidential}$

$Y = \text{GDP}_{-1}, \text{EDUC, PopGrowth, LATITUDE, ISLAND}$

The dependent variable is the number of patents granted (weighted by population) in a country. While patents are an imperfect measure of innovation output (i.e., they do not capture unpatented and unpatentable innovations, for example), they remain arguably the most readily and widely available measure of innovation. In our sample of 72 nations over the period 1997-2018, there were on average about 15 patents granted per 100,000 population.

The main input in innovation is research spending and we include R&D spending (as a percentage of GDP) with a one-year lag (see Hall et al. (1986)).<sup>3</sup> This measure of research spending includes both research personnel and related equipment.

The key focus and the main contribution of this work, however, entails the consideration and comparison of the different dimensions of socio-historic-economic aspects in terms of their impacts on innovation. An understanding of these influences will inform policymakers about their abilities and limitations to facilitate technical change in the short term. Income inequality can impact innovation - nations with more unequal incomes have a larger economic divide between the haves and have-nots, which would likely inhibit innovation as a large section of the population that lacks resources or means to pursue uncertain innovation. Furthermore, income inequality could lead to political instability that would also dampen technological progress (via low expected rewards to innovation effort). We include two different measures of income inequality - EthnicINQ is a relatively new measure (from Alesina et al. (2016)) that tempers income inequality differences across ethnic groups. This is compared to IncomeINQ, a more

---

<sup>3</sup> In initial models (not reported), we also included the second and third lags of R&D but the resulting coefficients were statistically insignificant.



traditional measure of income inequality (based on disposable income, see Solt (2020)).<sup>4</sup> As noted above, a few studies have looked at the innovation-income inequality relation (Aghion et al. (2019), Tselios (2011)), but no study has considered the income inequality based on ethnic divisions in this context.

The social composition of populations is related to knowledge spillovers and collaborations or networking which can significantly impact innovation and the pace of technological change. For instance, the ruling party may use the state to enrich its ethnic group at the expense of public goods and services that may inhibit technological progress (see, e.g., La Porta (1999)). The importance of social factors is accounted for by including measures of both ethnic fractionalization (EthnicFRAC) and ethnolinguistic fractionalization (EthLingFRAC), (see Desmet et al. (2012)).

The variable COLONY captures nations with a colonial past. Colonized nations were subjected to exploitation (inhibiting innovation), while they also benefitted from a smoother adoption of a legal system (tied to property rights protection) and easier formation of research joint ventures and associated spillovers. Former colonies, especially those that inherited the English language legacy from colonizers (e.g., India, Pakistan, Nigeria, etc.) are likely benefitting more than some other former colonies, especially given that in recent years English has turned out to be the dominant language of the internet. This would reduce the transaction costs of information acquisition (and dissemination) for such nations.

Finally, the longevity of nations may be important in terms of gestation and evolution of institutions and the state history index, StateHIST, captures this aspect (see Borcan et al. (2018)). At a broader level, this research can be seen as examining the effects of history on the development of institutions. For instance, colonial rulers influenced educational attainment in their colonies, which ultimately would impact innovation (see Bennett et al. (2017)).<sup>5</sup>

Two dimensions of institutions are considered: PatentRights and Presidential. Patent rights involve the ability of innovators to appropriate the gains from innovation (see Park (2008)) - we would expect nations with stronger patent rights to have greater innovation (patenting), *ceteris paribus*. The form of government, captured via the presidential form of democracies, can impact innovation behavior. Presidential democracies, compared to parliamentary democracies, have been argued to be more stable and legislatively efficient (Goel and Nelson (2020)).<sup>6, 7</sup>

---

<sup>4</sup> Along a related dimension, Goel and Saunoris (2020) recently showed greater income inequality to facilitate entrepreneurship (not necessarily entrepreneurship that is driven by innovation) across nations.

<sup>5</sup> In some cases, historical precedents can shape democratic institutions over time. Furthermore, the colonial history of the colonies and colonizers might vary across nations. In many cases, the colonizers were attracted by resource endowments. An examination of the influence of different types of colonizers is beyond the scope of current work.

<sup>6</sup> As an alternative related angle, we also considered an index of state fragility, capturing states prone to coups and frequent government changes. The resulting coefficient was statistically insignificant.

<sup>7</sup> Admittedly, our use of presidential democracies enables distinction from parliamentary democracies, but also bunches other non-democratic regimes in the default. A key example of this would be China, which is now a major world inventor, despite being a non-democratic regime. Future studies could focus on political systems and their impacts on innovation.

The vector  $Z$  includes control variables that are used in every model we estimate. These controls help in identifying the influence of the focus variables on innovation. Lagged GDP accounts for economic prosperity, EDUC captures educational attainment, population growth is associated with market potential (expected payoff from innovation versus the status quo), while LATITUDE and ISLAND capture the fixed locational or geographic aspects of nations.<sup>8</sup>

Economic prosperity is associated with better institutions at a macro-level and with positive market potential or payoffs from innovation at a micro-level. Educational attainment can be viewed as a complementary input to research personnel employed to produce innovations. The geographic factors capture resource endowments, openness to trade (island nations somewhat challenged in this respect), and productivity. All these influence the production of innovations as well as other goods and services.

### 3. Data and Estimation

#### 3.1 Data

The data set used in the following analysis is an unbalanced panel of 72 countries from 1997 to 2018—see Table 1 for variable definitions and sources, Table 1B for summary statistics, and Table 1A in the Appendix for the countries used in the analysis. The majority of the sample countries (approximately 34%) come from the Europe and Central Asia regions. The number of countries included in the analysis is constrained by the availability of the property rights index, measured as the average from 1960-1990.

The main variable of interest used to proxy for innovation is the number of patents granted (total count by filing office), including direct and PCT national phase entries, scaled per 100,000 population (PATENTS). Although imperfect, PATENTS provide a readily available and consistent measure for innovation across countries and time. The country with the most innovation is S. Korea (255 patent granted per 100,000 population), while Uganda has the least innovation (0.003 patents granted per 100,000 population).

The main independent variables are used to capture different dimensions of socio-historic-economic impacts on innovation. We use two measures for ethnic diversity including ethnolinguistic diversity (EthLingFRAC) and ethnic income inequality (EthnicINQ). Ethnolinguistic fractionalization is measured as the probability that two randomly selected individuals belong to different ethnolinguistic (or religious) groups and reflects the ethnolinguistic heterogeneity in a particular country (Desmet et al. (2012)). A more recent measure of ethnic diversity (EthnicINQ) from Alesina et al. (2016) is designed to capture

---

<sup>8</sup> Many scholars studying the influence of institutions on economic growth etc. consider the influence of latitude of nations (see, for example, Acemoglu et al. (2001); and Tselios (2011) notes the importance of geography in the analysis of income inequality on innovation).

economic differences across ethnic groups. EthnicINQ is a Gini index measuring the degree of income inequality across different ethnic groups.<sup>9</sup>

EthnicINQ and EthLingFRAC are weakly positively correlated with a correlation coefficient of 0.25. The country that is the most diverse according to EthnicINQ (EthLingFRAC) is Sudan (India); and the least diverse countries include Iceland, S. Korea, Madagascar, Malta, Mauritius, and Singapore (S. Korea).

In addition to measures of social composition, we also account for the longevity, history, and genetic distance of the state. The state history index (StateHIST) is based on three aspects of state presence which capture the duration of the state's existence from its emergence to 2000 CE. A state's presence is determined across three dimensions: (1) government above the tribal level; (2) local- or foreign-based government; and (3) amount of territory ruled by this government. According to Borcan et al. (2018: p.1) "The presence of a state is one of the most reliable historical predictors of social and economic development." The second measure of state history that we employ in the following analysis is an indicator variable for countries that are former British, French, Spanish or Portuguese, or other colonies (COLONY).<sup>10</sup> We also employ a unique measure for the genetic distance (GeneticDIST) between the ruling-elite population and the ethnic majority in the year 1900 (Dharmapala (2021)).<sup>11</sup> Genetic distance would capture tendencies towards networking and abilities to influence policy (innovation policy or otherwise).<sup>12</sup> These three measures capture somewhat different dimensions of state presence.

The correlation between StateHIST and COLONY is -0.24, while GeneticDIST is positively correlated with COLONY (0.36) and uncorrelated with StateHIST (-0.005). Interestingly, EthnicINQ is negatively (-0.07) correlated with StateHIST and positively (0.16) correlated with COLONY, while EthLingFRAC is positively correlated with StateHIST (0.38) and negatively (-0.04) correlated with COLONY. Further, GeneticDIST is positively correlated with EthnicINQ (0.54) and EthLingFRAC (0.14).

Patent law and the structure of the government are two types of institutions that are directly relevant to technological development. To control for the strength of patent protection we include as a control variable Park's (2008) Patent Rights Index (PatentRights), using the average from 1960-1990 (see Table 1). Finally, an indicator variable is used to denote a presidential, rather than parliamentary, democracy (Presidential).<sup>13</sup>

---

<sup>9</sup> In general, the ethnic composition of nations varies very gradually over time. Thus, having these variables for a single year seems reasonable – see Table 1.

<sup>10</sup> Another related angle might be to consider nations that were involved in colonial wars. Such experiences could impact research investments, especially through heightened military spending.

<sup>11</sup> Given the wide range in the genetic distance measure, we use its logged value (Tables 1 and 1B).

<sup>12</sup> Some of these historical variables are more readily used/relevant in economic growth/development research (see, for examples, Acemoglu et al. (2001), Borcan et al., (2018), Chanda and Putterman (2007)) and their application to innovativeness in this paper seems novel.

<sup>13</sup> Future research could also consider other regime dimensions such as authoritarian regimes, etc.

In sum, the data sources used are all from reputed sources that are routinely used in the literature (except, some of those like GeneticDIST and EthnicINQ are relatively new). A discussion of the estimation strategy follows.

### 3.2 Estimation

To estimate the model given by equation (1) we first linearize the model and then estimate it using OLS. We account for regional-specific factors<sup>14</sup> and time-specific events by including regional and time dummy variables in each model estimated.<sup>15</sup> Standard errors are clustered at the country level. Furthermore, we employ quantile regression in Table 4 to determine whether the impacts of socio-historic factors on innovation change when the prevalence of innovation changes across nations. In contrast to OLS, the quantile regression allows one to estimate at any point of the conditional distribution and is also less sensitive to the influence of outliers. The  $R^2$ s (and pseudo  $R^2$ s) in Tables 2-4 are quite decent and instill confidence in the fit of the various models estimated.

The results are discussed in the next section.

## 4. Results

### 4.1 Baseline models

The baseline models in Table 2 report different variations of equation (1) above. The main goal is to determine the relative influence of different socio-historic dimensions of colonialism and income inequality on innovation across nations.

Turning first to the main variables of interest, the socio-historic and income inequality dimensions, we find that different dimensions of social influences, including EthnicINQ, EthLingFRAC, and GeneticDIST inhibit innovation. In terms of elasticity, a one percent increase in EthnicINQ (EthLingFRAC) decreases PATENTS by 1.2 percent (0.71 percent). Ethnic differences pertain to inabilities to form effective contracts and have less trust in general, and in the context of innovation in particular, they might inhibit the flow of knowledge. Such knowledge flows are especially relevant given the evolutionary and interdependent nature of different technologies that are often rooted in inertia.

Nations with colonial pasts seem to be benefitting from institutional and technological spillovers that foster innovation over time, such that nations with a colonial past are associated with on average 13 (the mean of the dependent variable is 15) more patent grants per 100,000 population

---

<sup>14</sup> The regions are: East Asia and Pacific; Europe and Central Asia; Latin America and Caribbean; the Middle East and North Africa; North America; South Asia; and Sub-Saharan Africa.

<sup>15</sup> Note that R&D and GDP are in lagged values, which mitigates concerns about reverse causality. Further, these issues would be absent in the case of historical variables.

relative to non-colonial nations. Conversely, a nation's longevity or StateHIST failed to have a significant effect.

Consistent with intuition, the effect of lagged R&D on patenting is positive and significant in all cases. The positive effect of economic prosperity is also consistent with intuition, but the statistical support is more modest in this case. Finally, the coefficient on education is positive and statistically significant in three of the five models, indicating that a more educated population contributes to more innovation, on average; however, the lack of statistical significance in some models may be due to the lack of data on the quality of education.<sup>16</sup>

Of the two institutional dimensions, unlike PatentRights, Presidential democracies mostly were correlated with higher innovation rates<sup>17</sup>. On the other hand, higher population growth nations had lower innovation rates - the resulting coefficient was negative and statistically significant in a majority of the cases. It could be the case that, in nations with rapid population growth, the market for existing products is expanding rapidly, which makes the payoffs from innovation less attractive.

Finally, the two geographic measures, LATITUDE and ISLAND, fail to find statical support - nations differing on these attributes were no different from others in terms of their impacts on technological progress.

#### 4.2 Additional considerations - considering EthnicFRAC and IncomeINQ

In Table 3, we include two additional variables, EthnicFRAC and IncomeINQ, as alternative regressors.<sup>18</sup> The idea is to provide a comparison of the effect of ethnic fractionalization with ethnolinguistic fractionalization (Table 2); and income inequality with ethnic inequality (Table 2). Further, COLONY is added to Models 3.3 and 3.4 to control for colonization for both these cases.

Regarding ethnic fractionalization, the impact of innovation is negative and similar to that of EthLingFRAC in Table 2; however, statistical support for the effect of EthnicFRAC is weaker.<sup>19</sup> Thus, it is the broader social dimension, including both ethnic and linguistic dimensions that significantly inhibit innovation. Accounting for language differences has implications for knowledge dissemination.

On the other hand, the negative effect of income inequality is in line with the negative effect of EthnicINQ found in Table 2. In terms of relative magnitudes, the elasticity of PATENTS with respect to IncomeINQ (= -2.2; Model 3.2) is almost double that with respect to EthnicINQ (= -

<sup>16</sup> However, some aspects of the education quality are likely captured via the R&D input.

<sup>17</sup> A part of the reason for the insignificance of the coefficient on PatentRights might be that this composite index is unable to qualitatively distinguish among the institutional norms (transaction costs) of patent protection across nations. Additionally, the available index has observations for different years for individual nations that we average over time (see Table 1).

<sup>18</sup> Additionally, as another dimension of income inequality, we considered the percentage of total income held by the top 5% of the population, and while the coefficient was negative, it was statistically insignificant. These results are available upon request from the authors.

<sup>19</sup> In terms of relative elasticities, the elasticity of PATENTS with respect to EthLingFRAC (Model 2.2) equals -0.71, and that with respect to EthnicFRAC (Model 3.1) equals -0.27.

1.15; Model 2.1), implying that a given percent change in income inequality has a more modest impact on innovation when the ethnic dimension is considered. However, IncomeINQ loses statistical significance when COLONY is included as a regressor in Model 3.4.

The results for the other variables are in general agreement with what was reported in Table 2. R&D and GDP positively impact innovation, while population growth negatively impacts innovation. The effects of education and patent rights are mostly statistically insignificant, and the coefficient on Presidential is statistically significant in two of the four models. Next, we consider another empirical dimension with an alternative estimation technique.

#### 4.3 Quantile regression: Impact of socio-historic variables across nations with differing innovation rates

Table 4 presents results from the quantile regression (see Koenker (2005) for related background). The quantile regression enables a determination of the influence of the explanatory variables across the distribution of the dependent variable. Further, quantile regression avoids any parametric distribution assumptions on the error. Do historical factors affect more innovative nations differently from less innovative nations? Three sets of quantiles are reported, q25, q50, and q75, with q50 being the median regression. The use of the quantile regression also is a way to account for the influence of outliers.

With regard to the two historical variables from Table 2, EthnicINQ and COLONY, the signs and statistical significance hold across the board in Table 4. The (positive) sign of the coefficient on COLONY is especially large at q75 – i.e., colonial past especially benefits some of the most innovative nations such as Canada, Denmark, Finland, and the USA.

Most of the results from the OLS regressions in Tables 2 and 3 are supported, with some variations across quantiles in some cases. The positive link from R&D to patenting is evident across the board, as one would expect. Further, consistent with intuition, the payoffs from R&D are the greatest in the most innovative nations.

Interesting results emerge with respect to the impact of patent rights - stronger patent rights fail to have a significant impact in the most innovative nations (at q75 in 4a.3 and 4b.3). It could be the case that patented technologies in more innovative nations are more complex and thus have an in-built safeguard against infringement. On the other hand, higher population growth is statistically insignificant in the least innovative nations (Models 4a.1 and 4b.1), whereas higher educational attainment especially benefits this group.

Finally, unlike Tables 2 and 3, island nations had lower innovation (Models 4a.1-4a.3), whereas nations more distant from the equator were more innovative (median regression in Model 4a.2).

#### 4.4 Robustness checks

##### 4.4a Using negative binomial estimation

Because the dependent variable PATENTS is a count variable based on the number of patents granted (per 100,000 population), a useful robustness check is to estimate the baseline models using either a Poisson or negative binomial regression that better accounts for count-based



dependent variables. To distinguish between the negative binomial and the Poisson regression, we conduct the likelihood ratio test, under the null that the overdispersion parameter is  $\alpha$  is equal to zero, thus favoring Poisson estimation. The test results, reported at the bottom of Table 5, suggests that the overdispersion parameter  $\alpha$  is different from zero, therefore negative binomial regression is preferred over Poisson estimation.

To this end, we estimate the baseline models from Table 2 using a negative binomial regression and report the results in Table 5—note that the reported coefficients are incidence rate ratios. Consistent with the baseline findings, EthnicINQ and COLONY remain statistically significant, while the coefficients on EthLingFRAC and GeneticDIST are negative but lack statistical significance. The control variables also show some interesting differences. For instance, R&D, GDP, and LATITUDE are positive and statistically significant across all models (except that LATITUDE is insignificant in Model 5.2). The coefficient on Presidential is positive, albeit statistically significant throughout.

#### 4.4b Dropping ISLAND and LATITUDE as regressors

Because both ISLAND and LATITUDE lack statistical significance in explaining innovation, we dropped these variables and re-estimated the models reported in Table 2 and report the results in Table 6. This provides a robustness check of our main results.

Overall, the results are mostly consistent with the baseline findings. Specifically, the coefficients on EthnicINQ and EthLingFRAC are negative and statistically significant and the coefficient on COLONY is positive and statistically significant. The coefficient on GeneticDIST is also negative but marginally statistically insignificant (p-value=0.106). More significantly, the impact of lagged R&D remains positive and statistically significant across all the models estimated.

#### 4.4c Considering both population and population growth

Not only the growth in the population matters for innovation, but the level of the population might also contribute to innovation. To account for the influence of large population sizes (or country size) on innovation, Table 7 reports the baseline model results after including the log of the population as an additional regressor. Although the coefficient on Population is statistically insignificant, except in Model 7.5, the remaining results support the baseline findings. Relatively speaking, there was greater (negative) statistical support for the effect of population growth than for population.<sup>20</sup> A plausible explanation is that in nations with growing populations, the market continues to expand, reducing the need for attracting customers via new products.

#### 4.4d Focusing on the subset of former colonies

Nations with a colonial past might be qualitatively different from other nations that some quantitative measures might not be able to distinguish. As an additional robustness check, we restrict the sample to only those countries that were former colonies and re-estimate the baseline

---

<sup>20</sup> We also tried re-estimating the models in Table 7 by keeping population and dropping population growth. The results were qualitatively very similar to what is reported in Table 7. These are not reported but are available upon request.



models (Table 1A in the Appendix identifies nations with a colonial past). The results are reported in Table 8.<sup>21</sup>

Interestingly, while the results are mostly the same as the baseline models, the one exception is that state history is now positive and statistically significant. Therefore, among colonized countries, those with a long history of state presence experience higher rates of innovation, on average. This is perhaps not surprising given the history of colonized countries. Additionally, we find that the level of development (GDP) and education are no longer statistically significant (except in Model 8.4). These results support the baseline findings while also revealing additional insights into the correlates of innovation for former colonies.

#### 4.4e Examining the effects of structural breaks

The severity of the recent global recession (2007 to 2009) may have adversely impacted the institutions that govern innovation. For example, the financial crisis may have led financial firms to be more stringent in loaning capital. Furthermore, since innovation is forward-looking, economic shocks can lead to greater uncertainty which would reduce R&D and innovation (Goel and Ram (2001)).

To account for this unique aspect, we introduced a dummy variable equal to one for the global recession (2007-2009) and zero otherwise (GlobalRec). This variable is included in the baseline models and the results are reported in Table 9. The coefficient on GlobalRec is negative and statistically significant across all models, indicating a decline in the average level of innovation during the global recession. Nonetheless, we continue to find that EthnicINQ, EthLingFRAC and GeneticDIST negatively impact innovation, while former colonies positively impact innovation. Finally, even accounting for the global recession, the positive contribution of R&D to innovation remains strong.

The concluding section follows.

## 5. Concluding remarks

In order to better understand why some nations are more innovative than others, this paper considers the influence of socio-historic dimensions of colonialism and ethnic/income inequality. Determining effective drivers of innovation and understanding why some nations are more innovative than others remains a challenge.

Using data from 72 nations over the years 1997-2018, we find that ethnic income inequality (along with income inequality) and ethnolinguistic fractionalization of the population reduced the pace of innovation. There are significant differences in the relative magnitudes of the impacts

---

<sup>21</sup> To examine the heterogenous impact of colonialism on innovation, we disaggregated COLONY into five dummy variables including British colony, French colony, Spanish and Portuguese colony, other colony, and no colony. We then re-estimated the baseline model including these colony dummy variables (no colony was the base case) and report the results in Table 2A in the Appendix. Interestingly, former British and French colonies have (on average) lower rates of innovation, while French and other former colonies show greater rates of innovation—note, however, that the coefficient on other colonies is the only statistically significant coefficient.

as well. In terms of relative magnitudes, a one percent increase in the traditional income inequality (IncomeINQ from Model 3.2) would be associated with a more pronounced (double) negative impact on innovation, compared to a similar increase in ethnic income inequality (EthnicINQ from Model 2.1). These differences underscore the importance of social effects and reveal that changes in income inequality have a more modest impact on innovation when the ethnic dimension is considered. These findings underscore the role of income distribution and other social factors in influencing technical change. While the influence of the social environment on technological progress has been noted by other scholars (Mokyr (1992)), our findings provide a somewhat different dimension via the role of the demographic distribution of the population (captured by the variables *EthnicFRAC* and *EthLingFRAC*). As expected, nations with greater research spending were more innovative and greater educational attainment generally fostered innovation.

Furthermore, nations with a colonial past were more innovative, *ceteris paribus*, whereas nations with long histories (longevity) were no different from others. Finally, we found some support for presidential democracies fostering more innovation, while a nation's geographic location and island nation status did not have an appreciable impact.

The main results withstood a battery of robustness checking, including alternative estimation, changing the composition of the control variables, and accounting for the influence of the global recession (Tables 5-9).

From a policy perspective with regard to influencing innovation and the pace of technological change, it is not the history or longevity of nations *per se* that matters, rather it is the experience with colonization and the social and income distribution. This insight seems new to the literature on the drivers of technical change. While altering some historical precedents is beyond the realm of current policies, many nations have income redistribution programs in place to address income inequities. Also, although very costly and rare, switching the form of government (e.g., from a parliamentary system to a presidential one or vice versa) might help counter some of the historical inertia.<sup>22</sup> When such initiatives recognize the spillovers on technological progress it is less clear.

On the other hand, innovation spillovers from public support of education and research are widely recognized by policymakers and our findings also bear this out. Future research could extend this line of inquiry by examining the difference in the types of innovation outputs (design versus utility patents, for instance).

Overall, our findings reveal that when it comes to innovation, both income distribution and ethnic distributions of populations matter, and both of these might have some historical inertia. While changing both is politically challenging in the short term, income changes are relatively more politically expedient than demographic changes. These results provide an alternative explanation for why some nations might be laggards in technical development - it may be some

---

<sup>22</sup> Many British colonies in Africa, independent in the 1950s and 1960s, began as parliamentary systems and have since switched over to presidential systems of government (Robinson and Torvik (2016)).

overlooked factors like historic demographic compositions of their populations that might have put them on a different trajectory of technical development. The key insight from this research for a scholar of innovation and technological change is that the higher innovativeness of some nations might not be due to concerted policy actions, but due to some favorable historical events.

## References

- Acemoglu, D., Johnson, S., Robinson, J.A., 2001. The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5), 1369–1401.
- Acemoglu, D., Zilibotti, F., 2001. Productivity differences. *Quarterly Journal of Economics*, 116(2), 563–606.
- Aghion, P., Akcigit, U., Bergeaud, A., Blundell, R., Hemous, D., 2019. Innovation and top income inequality. *Review of Economic Studies*, 86(1), 563–606.
- Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., Wacziarg, R., 2003. Fractionalization. *Journal of Economic Growth*, 8(2), 155–194.
- Alesina, A., Michalopoulos, S., Papaioannou, E., 2016. Ethnic inequality. *Journal of Political Economy*, 124(2), 428–488.
- Antonelli, C., 2009. The economics of innovation: From the classical legacies to the economics of complexity. *Economics of Innovation and New Technology*, 18(7), 611–646.
- Antonelli, C., 2015. Innovation as a creative Response. A reappraisal of the Schumpeterian legacy. *History of Economic Ideas*, 23(2), 99–118.
- Antonelli, C., Gehringer, A., 2017. Technological change, rent and income inequalities: A Schumpeterian approach. *Technological Forecasting and Social Change*, 115, 85–98.
- Antonelli, C., Link, A.N., 2015. *Routledge Handbook of the Economics of Knowledge*. Routledge, UK.
- Arthur, W.B., 1994. *Increasing Returns and Path Dependence in the Economy*. The University of Michigan Press, Ann Arbor.
- Bennett, D.L., Faria, H.J., Gwartney, J.D., Morales, D.R., 2017. Economic institutions and comparative economic development: A post-colonial perspective. *World Development*, 96, 503–519.
- Bjørnskov, C., Rode, M., 2020. Regime types and regime change: A new dataset on democracy, coups, and political institutions. *The Review of International Organizations*, 15, 531–551.
- Borcan, O., Olsson, O., Putterman, L., 2018. State history and economic development: Evidence from six millennia. *Journal of Economic Growth*, 23(1), 1–40.
- Bockstette, V., Chanda, A., Putterman, L., 2002. States and markets: The advantage of an early start. *Journal of Economic Growth*, 7, 347–369.
- Chanda, A., Putterman, L., 2007. Early starts, reversals and catch-up in the process of economic development. *Scandinavian Journal of Economics*, 109, 387–413.
- Chandler, A.D., 1992. Organizational capabilities and the economic history of the industrial enterprise. *Journal of Economic Perspectives*, 6, 79–100.
- David, P.A., 1985. Clio and the economics of QWERTY. *American Economic Review*, 75, 332–337.

- David, P.A., 1994. Why are institutions the ‘carriers of history’?: Path dependence and the evolution of conventions, organizations and institutions. *Structural Change and Economic Dynamics*, 5(2), 205–220.
- David, P.A., 2001. Path dependence, its critics and the quest for ‘Historical Economics’. In: Garrouste, P., Ioannides, S. (Eds.), *Evolution and Path Dependence in Economic Ideas: Past and Present*. Edward Elgar, Cheltenham, UK, pp. 15–40.
- De Liso, N., Filatrella, G., 2011. On delayed technological shifts. *Economics of Innovation and New Technology*, 20(6), 563–580.
- Desmet, K., Ortuño-Ortín, I., Wacziarg, R., 2012. The political economy of ethnolinguistic cleavages. *Journal of Development Economics*, 97, 322–338.
- Dharmapala, D., 2021. A new measure of foreign rule based on genetic distance. *Economic Inquiry*, 59, 622–647.
- Dharmapala, D., Hines Jr, J.R., 2009. Which countries become tax havens?. *Journal of Public Economics*, 93(9–10), 1058–1068.
- Dosi, G., Nelson, R.R., 2010. Technical change and industrial dynamics as evolutionary processes. In: Hall, B.H., Rosenberg, N. (Eds.), *Handbook of the Economics of Innovation* vol. 1. North-Holland, The Netherlands, pp. 51–127.
- Fagerberg, J., Mowery, D.C., Verspagen, B., 2008. Innovation-Systems, Path-Dependency and Policy: The Co-Evolution of Science, Technology and Innovation Policy in a Small, Resource-Based Economy. DIME working paper number 2008.1 (<http://www.dimeeu.org/files/active/0/FagMowVer.pdf>).
- Freeman, C., 1995. The 'National System of Innovation' in historical perspective. *Cambridge Journal of Economics*, 19(1), 5–24.
- Gerring, J., Thacker, S.C., Moreno, C., 2009. Are parliamentary systems better? *Comparative Political Studies*, 42(3), 327–359.
- Goel, R.K., 1999. *Economic Models of Technological Change*. Quorum Books, Westport, CT.
- Goel, R.K., Nelson, M.A., 2020. Presidential versus parliamentary systems: Where do female entrepreneurs thrive? *Social Science Quarterly*, 101(5), 1773–1788.
- Goel, R.K., Ram, R., 2001. Irreversibility of R&D investment and the adverse effect of uncertainty: Evidence from the OECD countries. *Economics Letters*, 71(2), 287–291.
- Goel, R.K., Saunoris, J.W., 2016. Institutional path dependence and international research intensity. *Economic Modelling*, 52, (PB), 851–858.
- Goel, R.K., Saunoris, J.W., 2020. Does income inequality sand or grease the wheels of entrepreneurial activity? International evidence. *Australian Economic Papers*, 59(2), 138–160.
- Hall, B.H., Griliches, Z., Hausman, J.A., 1986. Patents and R and D: Is there a lag? *International Economic Review*, 27(2), 265–283.

- Johnson, S., 2014. *How we Got to Now: Six Innovations that Made the Modern World*. Riverhead Books.
- Kaminsky, E. B. - Z., 1997. On the comparison of presidential and parliamentary governments. *Presidential Studies Quarterly*, 27(2), 221–228.
- Koenker, R., 2005. *Quantile Regression*. Cambridge University Press, New York.
- La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R., 1999. The quality of government. *Journal of Law, Economics, and Organization*, 15(1), 222–279.
- Marshall, M.G., Gurr, T.R., Jagers, K., 2014. Polity IV Project: Political Regime Characteristics and Transitions, 1800–2013. <http://www.systemicpeace.org/polity/polity4.htm>.
- Mokyr, J., 1992. Technological inertia in economic history. *Journal of Economic History*, 52, 325–338.
- Mokyr, J., 2010. The contribution of economic history to the study of innovation and technical change: 1750–1914. In: Hall, B.H., Rosenberg, N. (Eds.), *Handbook of the Economics of Innovation* vol. 1. North-Holland, The Netherlands, pp. 11–50.
- Nelson, R.R. (Ed.), 1993. *National Innovation Systems*. Oxford University Press, New York.
- Nelson, R.R., Winter, S.G., 1982. *An Evolutionary Theory of Economic Change*. Belknap Press, Cambridge.
- North, D.C., 1990. *Institutions, Institutional Change, and Economic Performance*. Cambridge University Press, Cambridge.
- Park, W.G., 2008. International patent protection: 1960–2005. *Research Policy*, 37, 761–766.
- Putterman, L., Weil, D.N., 2010. Post-1500 population flows and the long-run determinants of economic growth and inequality. *Quarterly Journal of Economics*, 125, 1627–1682.
- Robinson, J.A., Torvik, R., 2016. Endogenous presidentialism. *Journal of the European Economic Association*, 14(4), 907–942.
- Rosenberg, N., 1976. *Perspectives on Technology*. Cambridge University Press, Cambridge.
- Ruttan, V.W., 1997. Induced innovation, evolutionary theory and path dependence: Sources of technical change. *Economic Journal*, 107, 1520–1529.
- Scotchmer, S., 2004. *Innovation and Incentives*. MIT Press, Cambridge.
- Simmie, J., Carpenter, J., Chadwick, A., Martin, R., 2008. *History matters: Path dependence and innovation in British City–Regions*. Research Report. NESTA, London.
- Solt, F., 2020. Measuring income inequality across countries and over time: The standardized world income inequality database. *Social Science Quarterly*, 101(3), 1183–1199.
- Taalbi, J., 2017. What drives innovation? Evidence from economic history. *Research Policy*, 46(8), 1437–1453.

Tselios, V., 2011. Is inequality good for innovation? *International Regional Science Review*, 34(1), 75-101.

Treisman, D., 2000. The causes of corruption: A cross-national study. *Journal of Public Economics*, 76, 399–457.

Vergne, J.-P., Durand, R., 2010. The missing link between the theory and empirics of path dependence: Conceptual clarification, testability issue, and methodological implications. *Journal of Management Studies*, 47, 736–759.

Whitley, R., 2002. Developing innovative competences: The role of institutional frameworks. *Industrial and Corporate Change*, 11, 497–528.

Zweimüller, J., 2000. Schumpeterian entrepreneurs meet Engel's Law: The impact of inequality on innovation-driven growth. *Journal of Economic Growth*, 5, 185–206.



**Table 1: Variable definitions and data sources**

Variable	Description	Source
<i>PATENTS</i>	Total patent grants per 100,000 population, measured by total count by filing office.	WIPO
<i>R&amp;D</i>	Research and development expenditures as a percent of GDP.	The World Bank (2018)
<i>PatentRights</i>	Index of Patent Rights (average from 1960-1990), based on coverage, membership in international treaties, duration of protection, enforcement mechanisms, and restrictions. Higher numbers denote stronger patent protections.	Park (2008)
<i>GDP</i>	Log of per capita real GDP.	The World Bank (2018)
<i>EDUC</i>	Educational attainment. Tertiary school enrollment measured as a percent of gross enrollment.	The World Bank (2018)
<i>PopGrowth</i>	Population growth rate.	The World Bank (2018)
<i>Presidential</i>	Presidential democracies. Indicator variable equal to one if the form of democracy is presidential and zero otherwise.	Bjørnskov and Rode (2020)
<i>LATITUDE</i>	The (absolute) value of the latitude of a nation from the equator.	Alesina et al. (2016)
<i>ISLAND</i>	Island nation. Indicator variable equal to one for island nations and zero otherwise.	Dharmapala and Hines (2009)
<i>EthnicINQ</i>	Ethnic income inequality Gini index, which measures the difference in mean income across ethnic groups (for year 2000). The index is measured on a scale of 0 (equal) to 1 (unequal).	Alesina et al. (2016)
<i>IncomeINQ</i>	Gini index of income inequality based on household disposable (post-tax and post-transfer) income. The index is measured on a scale of 0 (equal) to 100 (unequal).	Solt (2020)
<i>EthnicFRAC</i>	Ethnic fractionalization, calculated as $EthnicFRAC_j = 1 - \sum_{i=1}^N s_{ij}^2$ where $s_{ij}$ is the share of ethnic group in country j. Higher values reflect more ethnic diversity [Range 0-1].	Alesina et al. (2003)
<i>EthLingFRAC</i>	Ethnolinguistic fractionalization, which reflects ethnic-linguistic heterogeneity and is measured as the probability that two randomly selected individuals belong to two different ethnolinguistic groups. Higher values reflect more ethnolinguistic diversity [Range 0-1].	Alesina et al. (2016). (Original source is Desmet et al. (2012))
<i>COLONY</i>	Colonial past. Dummy variable identifying former colonies and zero otherwise.	Treisman (2000)
<i>StateHIST</i>	State history index. Normalized aggregate state history index for the period 3500 BCE to 2000 CE (using a 0% discount rate), where state presence is based on whether there was a government above the tribal level, if the government if foreign or locally based, and how much of the territory of the modern country was ruled by this government. Higher values reflect greater state presence.	Borcan et al. (2018)
<i>GeneticDIST</i>	Genetic distance. Log of elite-population genetic distance, which measures the genetic distance of a country's ruling elite from its ethnic majority in year 1900.	Dharmapala (2021)

Note: unbalanced panel data for 72 countries from 1997 to 2018, except the following variables only vary by country: *PatentRights*, *LATITUDE*, *ISLAND*, *EthnicFRAC*, *EthLingFRAC*, *EthnicINQ*, *COLONY*, *StateHIST*, and *GeneticDIST*.

**Table 1B: Summary statistics**

	Observations	Mean	Standard Deviation	Minimum	Maximum
<i>PATENTS</i>	820	15.090	31.400	0.003	255.800
<i>R&amp;D</i>	726	1.302	1.053	0.015	4.553
<i>PatentRights</i>	820	2.068	0.832	0.000	4.140
<i>GDP</i>	820	9.477	1.295	5.850	11.610
<i>EDUC</i>	820	50.840	24.340	0.970	136.600
<i>PopGrowth</i>	820	0.935	0.886	-2.171	5.009
<i>Presidential</i>	820	0.501	0.500	0.000	1.000
<i>LATITUDE</i>	820	36.960	17.230	1.282	64.990
<i>ISLAND</i>	820	0.146	0.354	0.000	1.000
<i>EthnicINQ</i>	820	0.375	0.234	0.000	0.966
<i>IncomeINQ</i>	796	35.340	7.796	22.200	59.700
<i>EthnicFRAC</i>	820	0.305	0.238	0.002	0.930
<i>EthLingFRAC</i>	802	0.333	0.273	0.003	0.930
<i>COLONY</i>	774	0.659	0.474	0.000	1.000
<i>StateHIST</i>	799	0.241	0.154	0.021	0.748
<i>GeneticDIST</i>	820	2.974	3.025	0.000	7.738

**Table 2**  
**Socio-historic dimensions of colonialism and innovation:**  
**Baseline models**  
**Dependent variable: *PATENTS***

	(2.1)	(2.2)	(2.3)	(2.4)	(2.5)
<i>R&amp;D<sub>t-1</sub></i>	8.826** (4.238)	9.030** (3.610)	7.861* (4.189)	10.056** (4.393)	12.525** (5.332)
<i>PatentRights</i>	2.036 (3.306)	3.196 (3.528)	0.306 (3.226)	2.314 (3.169)	-0.990 (3.315)
<i>GDP<sub>t-1</sub></i>	4.009 (4.485)	7.151 (4.298)	10.353** (4.242)	3.158 (3.944)	3.254 (5.302)
<i>EDUC</i>	0.321* (0.163)	0.266 (0.184)	0.317* (0.188)	0.503** (0.232)	0.271 (0.170)
<i>PopGrowth</i>	-4.985 (3.257)	-4.152 (3.181)	-6.682* (3.653)	-5.758* (3.363)	-6.292* (3.527)
<i>Presidential</i>	11.513** (5.694)	8.307 (5.568)	11.348 (6.811)	12.695* (6.883)	16.147* (8.100)
<i>LATITUDE</i>	0.171 (0.234)	-0.209 (0.246)	0.027 (0.237)	-0.001 (0.261)	0.120 (0.252)
<i>ISLAND</i>	-9.393 (7.002)	-0.528 (5.878)	-1.724 (6.660)	4.451 (4.621)	6.111 (5.914)
<i>EthnicINQ</i>	-46.930*** (16.638)				
<i>EthLingFRAC</i>		-31.377** (15.474)			
<i>COLONY</i>			12.591** (6.249)		
<i>StateHIST</i>				29.948 (26.144)	
<i>GeneticDIST</i>					-4.313* (2.471)
Observations	820	802	774	799	820
R-squared	0.625	0.586	0.577	0.609	0.594
Year dummies	Y	Y	Y	Y	Y
Region dummies	Y	Y	Y	Y	Y

*Notes:* See Table 1 for variable details. Constants are included in each model, but not reported. Each model is estimated using OLS. Cluster-robust standard errors are in parentheses.

Asterisks denote the following significance levels: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , and \*  $p < 0.1$ .

**Table 3**  
**Socio-historic dimensions of colonialism and innovation:**  
**Additional considerations (*EthnicFRAC*; *IncomeINQ*)**  
**Dependent variable: *PATENTS***

	(3.1)	(3.2)	(3.3)	(3.4)
<i>R&amp;D<sub>t-1</sub></i>	8.929** (4.065)	8.071** (3.999)	7.773* (4.090)	7.502* (4.230)
<i>PatentRights</i>	0.024 (3.075)	0.197 (3.327)	-0.414 (3.017)	0.305 (3.373)
<i>GDP<sub>t-1</sub></i>	8.384* (4.493)	8.487* (4.289)	10.691** (4.294)	10.489** (4.212)
<i>EDUC</i>	0.295 (0.208)	0.289 (0.186)	0.287 (0.191)	0.298* (0.178)
<i>PopGrowth</i>	-6.513* (3.704)	-7.112* (3.773)	-5.973* (3.409)	-7.067* (3.694)
<i>Presidential</i>	10.707 (7.000)	13.035* (7.438)	9.467 (6.255)	12.191* (6.891)
<i>LATITUDE</i>	-0.069 (0.240)	-0.159 (0.247)	-0.129 (0.252)	-0.084 (0.246)
<i>ISLAND</i>	1.358 (5.299)	1.522 (5.271)	-4.082 (6.956)	-1.840 (6.851)
<i>EthnicFRAC</i>	-13.678 (12.296)		-22.824* (12.726)	
<i>IncomeINQ</i>		-0.976** (0.436)		-0.561 (0.379)
<i>COLONY</i>			13.654** (6.266)	11.522* (6.336)
Observations	820	796	774	755
R-squared	0.554	0.565	0.587	0.585
Year dummies	Y	Y	Y	Y
Region dummies	Y	Y	Y	Y

Notes: See Table 2.

**Table 4**  
**Determinants of innovation across the prevalence of innovation:**  
**Quantile regression**

	<b>Dependent variable: <i>PATENTS</i></b>					
	(4a.1) q25	(4a.2) q50	(4a.3) q75	(4b.1) q25	(4b.2) q50	(4b.3) q75
<i>R&amp;D<sub>t-1</sub></i>	4.452*** (0.423)	5.665*** (0.567)	8.082*** (1.401)	3.472*** (0.489)	5.913*** (0.651)	6.924*** (1.081)
<i>PatentRights</i>	2.214*** (0.304)	1.941*** (0.605)	0.413 (0.886)	1.404*** (0.514)	1.686*** (0.503)	1.076 (1.076)
<i>GDP<sub>t-1</sub></i>	0.420 (0.296)	0.883 (0.676)	1.161 (1.164)	1.423** (0.665)	1.734*** (0.519)	4.577*** (1.482)
<i>EDUC</i>	0.036*** (0.014)	0.013 (0.022)	0.081 (0.053)	0.039* (0.021)	0.004 (0.033)	0.072 (0.050)
<i>PopGrowth</i>	-0.213 (0.321)	-1.379*** (0.485)	-1.544* (0.817)	-0.541 (0.376)	-1.453*** (0.433)	-1.284 (1.054)
<i>Presidential</i>	1.770** (0.744)	2.459*** (0.576)	1.948 (1.216)	0.272 (0.645)	1.582* (0.808)	2.542** (1.122)
<i>LATITUDE</i>	0.009 (0.021)	0.068** (0.030)	0.095 (0.071)	-0.008 (0.028)	0.039 (0.032)	0.108 (0.102)
<i>ISLAND</i>	-3.104*** (0.764)	-3.794*** (0.915)	-3.270** (1.376)	-1.770* (1.032)	-1.315 (0.806)	-1.292 (1.589)
<i>EthnicINQ</i>	-8.475*** (1.458)	-10.955*** (1.853)	-10.342*** (2.968)			
<i>COLONY</i>				2.312*** (0.746)	2.430*** (0.664)	7.281*** (1.847)
Observations	820	820	820	774	774	774
Pseudo R-squared	0.245	0.330	0.393	0.235	0.314	0.390
Year dummies	Y	Y	Y	Y	Y	Y
Region dummies	Y	Y	Y	Y	Y	Y

*Notes:* See Table 1 for variable details. Constants are included in each model, but not reported. Each model is estimated using quantile regression where q25 is the 25<sup>th</sup> percentile, q50 is the 50<sup>th</sup> percentile (median), and q75 is the 75<sup>th</sup> percentile. Bootstrapped standard errors are in parentheses based on 50 replications.

Asterisks denote the following significance levels: \*\*\* p < 0.01, \*\* p < 0.05, and \* p < 0.1.

**Table 5**  
**Socio-historic dimensions of colonialism and innovation:**  
**Using a negative binomial regression**  
**Dependent variable: *PATENTS***

	(5.1)	(5.2)	(5.3)	(5.4)	(5.5)
<i>R&amp;D<sub>t-1</sub></i>	0.252*** (0.073)	0.275*** (0.093)	0.242*** (0.087)	0.300*** (0.090)	0.264*** (0.092)
<i>PatentRights</i>	0.059 (0.130)	0.085 (0.143)	0.061 (0.156)	0.131 (0.154)	0.028 (0.148)
<i>GDP<sub>t-1</sub></i>	0.853*** (0.139)	0.958*** (0.150)	0.990*** (0.131)	0.738*** (0.176)	0.975*** (0.145)
<i>EDUC</i>	-0.005 (0.003)	-0.006 (0.004)	-0.004 (0.004)	0.001 (0.005)	-0.005 (0.004)
<i>PopGrowth</i>	-0.395*** (0.122)	-0.415*** (0.132)	-0.425*** (0.129)	-0.349*** (0.116)	-0.461*** (0.128)
<i>Presidential</i>	0.177 (0.142)	0.161 (0.179)	0.170 (0.162)	0.237 (0.150)	0.238 (0.175)
<i>LATITUDE</i>	0.024*** (0.008)	0.012 (0.009)	0.019** (0.007)	0.017** (0.008)	0.019** (0.008)
<i>ISLAND</i>	-0.245 (0.163)	-0.089 (0.213)	-0.032 (0.186)	-0.025 (0.186)	-0.029 (0.173)
<i>EthnicINQ</i>	-1.089*** (0.294)				
<i>EthLingFRAC</i>		-0.570 (0.454)			
<i>COLONY</i>			0.392** (0.198)		
<i>StateHIST</i>				0.122 (0.745)	
<i>GeneticDIST</i>					-0.007 (0.041)
LR test ( $\alpha = 0$ )	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	820	802	774	799	820
Pseudo R-squared	0.2133	0.2080	0.2004	0.2198	0.2034
Year dummies	Y	Y	Y	Y	Y
Region dummies	Y	Y	Y	Y	Y

*Notes:* See Table 1 for variable details. Constants are included in each model, but not reported. Each model is estimated using negative binomial regression. The reported coefficients are incidence-rate ratios. Cluster-robust standard errors are in parentheses and probability values are in brackets. Asterisks denote the following significance levels: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , and \*  $p < 0.1$ .



**Table 6**  
**Socio-historic dimensions of colonialism and innovation:**  
**Excluding ISLAND and LATITUDE**  
**Dependent variable: *PATENTS***

	(6.1)	(6.2)	(6.3)	(6.4)	(6.5)
<i>R&amp;D<sub>t-1</sub></i>	10.216** (4.437)	8.304** (3.823)	8.152* (4.376)	9.702** (4.366)	12.043** (5.346)
<i>PatentRights</i>	2.110 (3.317)	3.351 (3.501)	0.343 (3.139)	2.320 (3.176)	-1.536 (3.370)
<i>GDP<sub>t-1</sub></i>	3.511 (4.635)	6.796* (3.951)	10.043*** (3.750)	3.579 (3.781)	4.993 (4.997)
<i>EDUC</i>	0.342* (0.173)	0.261 (0.181)	0.324* (0.191)	0.491** (0.225)	0.271 (0.164)
<i>PopGrowth</i>	-6.162* (3.163)	-4.013 (2.948)	-6.822* (3.647)	-5.165 (3.171)	-6.115* (3.523)
<i>Presidential</i>	10.535* (5.522)	8.307 (5.515)	11.081* (6.573)	13.498* (6.924)	16.821** (8.379)
<i>LATITUDE</i>	-----	-----	-----	-----	-----
<i>ISLAND</i>	-----	-----	-----	-----	-----
<i>EthnicINQ</i>	-40.976*** (15.388)				
<i>EthLingFRAC</i>		-28.833** (13.443)			
<i>COLONY</i>			12.418** (6.056)		
<i>StateHIST</i>				26.934 (24.915)	
<i>GeneticDIST</i>					-3.989 (2.438)
Observations	820	802	774	799	820
R-squared	0.617	0.583	0.577	0.607	0.589
Year dummies	Y	Y	Y	Y	Y
Region dummies	Y	Y	Y	Y	Y

*Notes:* See Table 1 for variable details. Constants are included in each model, but not reported. Each model is estimated using OLS. Cluster-robust standard errors are in parentheses.

Asterisks denote the following significance levels: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , and \*  $p < 0.1$ .

**Table 7**  
**Socio-historic dimensions of colonialism and innovation:**  
**Considering the impact of population size**  
**Dependent variable: *PATENTS***

	(7.1)	(7.2)	(7.3)	(7.4)	(7.5)
<i>R&amp;D<sub>t-1</sub></i>	8.731** (4.246)	8.922** (3.622)	7.769* (4.161)	9.795** (4.332)	12.381** (5.255)
<i>PatentRights</i>	3.448 (3.246)	4.558 (3.629)	1.543 (3.223)	3.410 (3.329)	1.124 (3.235)
<i>GDP<sub>t-1</sub></i>	2.768 (4.195)	5.804 (3.833)	9.116** (3.883)	2.521 (3.888)	1.290 (4.956)
<i>EDUC</i>	0.310** (0.152)	0.257 (0.172)	0.310* (0.182)	0.466** (0.212)	0.255 (0.156)
<i>Population</i>	-1.763 (1.417)	-1.843 (1.478)	-1.438 (1.438)	-2.466 (1.546)	-2.560* (1.461)
<i>PopGrowth</i>	-5.039 (3.153)	-4.324 (3.085)	-6.847* (3.595)	-5.994* (3.358)	-6.293* (3.367)
<i>Presidential</i>	11.872** (5.580)	8.879 (5.653)	11.884* (6.933)	12.714* (6.729)	16.641** (7.875)
<i>LATITUDE</i>	0.176 (0.230)	-0.187 (0.230)	0.031 (0.233)	0.026 (0.242)	0.133 (0.244)
<i>ISLAND</i>	-11.350 (7.890)	-2.734 (6.524)	-3.508 (7.886)	2.135 (5.295)	2.814 (6.044)
<i>EthnicINQ</i>	-45.679*** (16.246)				
<i>EthLingFRAC</i>		-29.593* (15.125)			
<i>COLONY</i>			11.076* (6.021)		
<i>StateHIST</i>				39.203 (27.594)	
<i>GeneticDIST</i>					-4.305* (2.388)
Observations	820	802	774	799	820
R-squared	0.629	0.590	0.579	0.616	0.603
Year dummies	Y	Y	Y	Y	Y
Region dummies	Y	Y	Y	Y	Y

*Notes:* See Table 1 for variable details. Constants are included in each model, but not reported. Each model is estimated using OLS. Cluster-robust standard errors are in parentheses.

Asterisks denote the following significance levels: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , and \*  $p < 0.1$ .

**Table 8**  
**Socio-historic dimensions of colonialism and innovation:**  
**Focusing on the subset of former colonies**  
**Dependent variable: *PATENTS***

	(8.1)	(8.2)	(8.3)	(8.4)	(8.5)
<i>R&amp;D<sub>t-1</sub></i>	14.767* (7.677)	14.992* (8.025)	13.900* (8.280)	16.647** (7.598)	20.180** (9.178)
<i>PatentRights</i>	-4.072 (4.109)	-1.867 (3.385)	-5.291 (4.260)	-2.490 (4.123)	-6.812 (4.283)
<i>GDP<sub>t-1</sub></i>	2.512 (7.051)	6.256 (6.105)	8.429 (7.016)	0.599 (6.463)	2.509 (7.464)
<i>EDUC</i>	0.160 (0.178)	-0.013 (0.162)	0.020 (0.166)	0.321** (0.146)	-0.026 (0.172)
<i>PopGrowth</i>	-8.126** (3.937)	-4.646 (3.654)	-11.616** (4.671)	-11.288*** (4.054)	-9.531** (4.240)
<i>Presidential</i>	14.152** (6.588)	3.820 (5.340)	11.868* (6.908)	10.048* (5.913)	18.495** (8.006)
<i>LATITUDE</i>	0.225 (0.286)	0.163 (0.315)	0.346 (0.371)	0.185 (0.280)	0.337 (0.362)
<i>ISLAND</i>	-9.025 (10.676)	-4.824 (9.081)	-2.689 (11.649)	5.596 (10.421)	5.075 (14.101)
<i>EthnicINQ</i>	-47.694*** (15.030)				
<i>EthLingFRAC</i>		-45.497*** (15.519)			
<i>COLONY</i>			-----		
<i>StateHIST</i>				72.243** (31.060)	
<i>GeneticDIST</i>					-4.556* (2.542)
Observations	510	510	510	489	510
R-squared	0.708	0.695	0.658	0.739	0.690
Year dummies	Y	Y	Y	Y	Y
Region dummies	Y	Y	Y	Y	Y

*Notes:* See Table 1 for variable details. Sample is restricted to former colonies (i.e. COLONY=1). Constants are included in each model, but not reported. Each model is estimated using OLS. Cluster-robust standard errors are in parentheses. Asterisks denote the following significance levels: \*\*\* p < 0.01, \*\* p < 0.05, and \* p < 0.1.

**Table 9**  
**Socio-historic dimensions of colonialism and innovation:**  
**Accounting for the global recession (2007-2009)**  
**Dependent variable: *PATENTS***

	(9.1)	(9.2)	(9.3)	(9.4)	(9.5)
<i>R&amp;D<sub>t-1</sub></i>	8.826** (4.238)	9.030** (3.610)	7.861* (4.189)	10.056** (4.393)	12.525** (5.332)
<i>PatentRights</i>	2.036 (3.306)	3.196 (3.528)	0.306 (3.226)	2.314 (3.169)	-0.990 (3.315)
<i>GDP<sub>t-1</sub></i>	4.009 (4.485)	7.151 (4.298)	10.353** (4.242)	3.158 (3.944)	3.254 (5.302)
<i>EDUC</i>	0.321* (0.163)	0.266 (0.184)	0.317* (0.188)	0.503** (0.232)	0.271 (0.170)
<i>PopGrowth</i>	-4.985 (3.257)	-4.152 (3.181)	-6.682* (3.653)	-5.758* (3.363)	-6.292* (3.527)
<i>Presidential</i>	11.513** (5.694)	8.307 (5.568)	11.348 (6.811)	12.695* (6.883)	16.147* (8.100)
<i>LATITUDE</i>	0.171 (0.234)	-0.209 (0.246)	0.027 (0.237)	-0.001 (0.261)	0.120 (0.252)
<i>ISLAND</i>	-9.393 (7.002)	-0.528 (5.878)	-1.724 (6.660)	4.451 (4.621)	6.111 (5.914)
<i>GlobalRec</i>	-11.005*** (3.341)	-11.365*** (3.828)	-13.734*** (3.565)	-13.913*** (3.266)	-10.951** (4.186)
<i>EthnicINQ</i>	-46.930*** (16.638)				
<i>EthLingFRAC</i>		-31.377** (15.474)			
<i>COLONY</i>			12.591** (6.249)		
<i>StateHIST</i>				29.948 (26.144)	
<i>GeneticDIST</i>					-4.313* (2.471)
Observations	820	802	774	799	820
R-squared	0.625	0.586	0.577	0.609	0.594
Year dummies	Y	Y	Y	Y	Y
Region dummies	Y	Y	Y	Y	Y

*Notes:* See Table 1 for variable details. Constants are included in each model, but not reported. Each model is estimated using OLS. Cluster-robust standard errors are in parentheses.

Asterisks denote the following significance levels: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , and \*  $p < 0.1$ .

## Appendix

**Table 1A**  
**Countries included in the analysis**

Algeria*^	Hungary*	Panama*^
Argentina*^	Iceland^	Peru*^
Australia*	India*	Philippines*^
Austria	Indonesia*^	Poland*^
Belgium*	Iran, Islamic Rep.^	Portugal^
Brazil*^	Ireland*^	Romania*^
Bulgaria*	Israel*	Rwanda*
Canada*	Italy	Saudi Arabia
Chile*^	Jamaica*	Singapore*
China	Jordan*	South Africa*^
Colombia*^	Korea, Rep.*^	Spain
Costa Rica*^	Luxembourg*	Sri Lanka*^
Cyprus*^	Madagascar^	Sudan*^
Denmark	Malaysia*	Sweden
Ecuador*^	Malta*	Switzerland^
Egypt, Arab Rep.*^	Mauritius*	Thailand
El Salvador*^	Mexico*^	Trinidad and Tobago*
Ethiopia^	Morocco*	Tunisia^
Finland*^	Mozambique*^	Turkey^
France^	Nepal	Uganda*^
Germany	Netherlands*	United Kingdom
Greece	New Zealand*	United States*^
Guatemala*^	Norway	Uruguay*^
Honduras*^	Pakistan*^	Vietnam*

N = 72. \* denotes a former British, French, Spanish or Portuguese, or other colony.

^ denotes presidential democracies.

**Table 2A**  
**Socio-historic dimensions of colonialism and innovation:**  
**Examining the heterogeneity of colonialism**  
**Dependent variable: *PATENTS***

	(2A.1)
<i>R&amp;D<sub>t-1</sub></i>	7.938** (3.907)
<i>PatentRights</i>	2.411 (4.684)
<i>GDP<sub>t-1</sub></i>	11.720*** (3.805)
<i>EDUC</i>	0.256** (0.121)
<i>PopGrowth</i>	-3.594 (3.573)
<i>Presidential</i>	10.179* (6.095)
<i>LATITUDE</i>	-0.097 (0.227)
<i>ISLAND</i>	7.649 (7.008)
<i>British Colony</i>	-3.933 (6.673)
<i>French Colony</i>	1.475 (3.681)
<i>Spanish &amp; Portuguese Colony</i>	-0.569 (7.968)
<i>Other Colony</i>	26.444** (11.145)
Observations	758
R-squared	0.617
Year dummies	Y
Region dummies	Y

*Notes:* See Table 1 for variable details. Constants are included in each model, but not reported. Each model is estimated using OLS. Cluster-robust standard errors are in parentheses.

Asterisks denote the following significance levels: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , and \*  $p < 0.1$ .