Angeles, L. and Elizalde, A. (2017) Pre-colonial institutions and socioeconomic development: The case of Latin America. *Journal of Development Economics*, 124, pp. 22-40. (doi:10.1016/j.jdeveco.2016.08.006)

This is the author's final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

http://eprints.gla.ac.uk/128869/

Deposited on: 22 September 2016

Pre-Colonial Institutions and Socioeconomic Development: The Case of Latin America.

Luis Angeles[*]and Aldo Elizalde[†]

August 23, 2016

Abstract

We study the effects of pre-colonial institutions on present-day socioeconomic outcomes for Latin America. Our thesis is that more advanced pre-colonial institutions relate to better socioeconomic outcomes today. We advance that pre-colonial institutions survived to our days thanks to the existence of largely self-governed Amerindian communities in rural Latin America. Amerindians groups with more advanced institutional capacity would have been able to organize and defend their interests in front of national governments; leading to better development outcomes for themselves and for the population at large. We test our thesis with a dataset of 324 sub-national administrative units covering all mainland Latin American countries. Our extensive range of controls covers factors such as climate, location, natural resources, colonial activities and pre-colonial characteristics - plus country fixed effects. Results strongly support our thesis.

Keywords: Latin America; Institutions; Pre-colonial factors; Amerindians; Long-run development.

JEL classification: O11, O43.

[*]Economics, Adam Smith Business School, University of Glasgow. Glasgow G12 8QQ, UK. Email: luis.angeles@glasgow.ac.uk.

[†]Economics, Adam Smith Business School, University of Glasgow. Glasgow G12 8QQ, UK. Email: a.elizalde-archundia.1@research.gla.ac.uk

1 Introduction

While much research effort has been directed over the last two decades to the role of institutions in long-run economic development, most of the literature has followed the lead of Acemoglu et al. (2001) in focusing on the institutional structures implanted throughout the world by European nations during the colonial period. While this approach has certainly advanced our understanding of the process of long-run development, recent research has also uncovered a substantial role for pre-colonial institutional factors - in particular for the African case (Gennaioli and Rainer 2007, Michalopoulos and Papaioannu 2013, 2014). This paper aims to take the literature forward by analyzing the role of pre-colonial institutions on present-day socioeconomic outcomes for Latin America. To the best of our knowledge, this is the first attempt of this kind for the Latin American case.

Our study employs as its unit of analysis the largest administrative divisions of each Latin American country below the national level, which we refer to as sub-national states.[1] We collect socioeconomic and geographic data for 324 sub-national states from all 17 countries in mainland Latin America - including data on the ethnic composition of the population within each state. This ethnic structure of the population is used to construct, for each state, an index of pre-colonial institutional advancement - by linking each ethnic group to its level of political complexity as reported in George Peter Murdock's *Ethnographic Atlas* (1967). The resulting measure of pre-colonial institutions is then used as a predictor of present-day measures of socioeconomic development such as education, health, or income per person. Through a large battery of robustness tests, we show that our prior of a positive relationship between pre-colonial institutions and present-day socioeconomic development holds while controlling for a number of alternative determinants of economic success.

Our paper relates to several strands of the literature on economic de-

---

[1] The actual name given to these administrative divisions changes from country to country: *provincias* in Argentina, *departamentos* in Bolivia, *regiones* in Chile, *estados* in Mexico, and so on.

velopment over the very long-run. A number of papers have uncovered how historical phenomena have persistent effects on economic development, in particular for Africa. Nunn (2008) and Nunn and Wantchekon (2011) relate the intensity of the African slave trade, most of which takes place over the pre-colonial period, to current levels of income per capita and interpersonal trust. Huillery (2011) argues that the attitudes of pre-colonial African states towards Europeans have an influence on current development outcomes as colonizers invested more in the areas where Africans were less hostile. Closest to our paper are the works of Gennaioli and Rainer (2007) and Michalopoulos and Papaioannu (2013, 2014), who also measure pre-colonial institutions using the degree of political complexity of different ethnic groups from Murdock (1967). We differ from these last three works by using sub-national states as the unit of analysis, an approach which has been increasingly employed in the wider literature on growth and development (Acemoglu and Dell 2010, Gennaioli et al. 2014).[2]

Our paper also relates to a number of recent works focusing on the determinants of long-run development in Latin America. Bruhn and Gallego (2012) study the effects of different colonial activities, such as mining or plantation agriculture, using sub-national states as the unit of analysis. This offers a relevant set of controls for our own study, as we seek to factor out colonial influences from pre-colonial ones. Controlling for colonial activities appears particularly relevant in the light of another recent study, by Arias and Girod (2014), presenting evidence that colonial policies were to a large extent determined by pre-colonial factors - in particular the existence or absence of forced labour. Also of relevance, Maloney and Valencia (forthcoming) study the persistence of economic success from pre-colonial times to the present across the Americas - a theme that was first investigated by Acemoglu et al. (2002) in their "reversal of fortune" thesis. We

---

[2] Gennaioli and Rainer (2007) use nations as their unit of analysis, which would result in too few observations in Latin America and does not allow for the inclusion of country fixed effects. Michalopoulos and Papaioannu (2013) use ethnic groups as their unit of analysis, which is feasible for Latin America but comes with its own special set of challenges (see Elizalde 2016 for an analysis using this alternative approach).

add evidence to this side of the literature by controlling for the initial level of development in some of our regressions. Finally, it is important to mention that a number of papers have studied the long-run effects of specific historical episodes in Latin America, such as the *mita* labour system of Peru and Bolivia (Dell 2010), the exploitation of sugarcane and gold in different areas of Brazil (Naritomi et al. 2012), or the early establishment of political inequality in 19th century Colombia (Acemoglu et al. 2008).[3]

The rest of the paper is organized as follows. Section 2 offers an historical overview, formulates our main hypothesis and presents historical evidence in its favour. Section 3 discusses the data and presents our empirical methodology. Sections 4 and 5 present our empirical analysis, extensions and robustness checks. Section 6, finally, offers some concluding remarks.

2   Historical overview and main hypothesis

The paucity of research on the role of pre-colonial institutions in socioeconomic development for the Latin American case may be explained by the overpowering importance of the colonial experience in this region. Indeed, colonialism was not only much longer lasting than in other regions of the world, about three centuries for most Latin American nations, but it was also accompanied by a massive transformation of the ethnic structure of the population. Suffering the consequences of a new disease environment, the aboriginal population of the Americas (henceforth *Amerindians*) was decimated over the hundred years following first contact with Europeans. In its place, a society of "whites" (European descendants), mestizos (people with mixed Amerindian and European ancestry), Amerindians and Africans, these last ones brought to the continent as slave labour, took over. Amerindians have always occupied the lower ranks of the new social hierarchy established with the colonial conquest, while European descendants, and later on mestizos, have dominated political and economic affairs ever since.

---

[3]For a more comprehensive review of the literature on how historical factors determine long-run development see Nunn (2014).

While correct, these central features of Latin American history should not lead us to conclude that pre-colonial culture was simply wiped out following the European arrival and had no chance of influencing the present. Western culture may well be the dominant element in present-day Latin America, as best exemplified by the almost universal adoption of the Christian religion and the languages from Portugal and Spain, but in numerous cases this dominance contains important elements of Amerindian culture. Pre-colonial rites, such as offerings to the Mother-Earth, are still common among large sections of the Latin American population - and do not stop those who perform them from attending the Sunday Mass. And pre-colonial languages such as *quechua* or *nahuatl* (the lingua francas of the former Inca and Aztec empires) can be easily heard in the food markets of cities and towns across the Andes and central Mexico. Elements of pre-colonial culture have survived five centuries of colonial and post-colonial regimes, and the present section discusses how this took place and why it could matter for the analysis of current socioeconomic outcomes.

While Amerindians remained at the fringes of the economic and political power ever since the European conquest, they played a crucial role as the main source of labour, and therefore principal factor of production, in two of the most important sectors of the colonial economy: mining and agricultural production for the local market.[4] Amerindian labour was the main source of wealth for Spanish settlers in the Americas, as best summarized by the aphorism *"Sin Indios no hay Indias"* ("Without Indians there is no Indies") - attributed to 16th century Spanish settlers when defending the granting of rights over Indian labour against accusations by the Crown of excessive exploitation.

The extraction of this Amerindian labour relied on the use of aboriginal structures of power and organization. While the growing class of mestizos lived in towns and cities and collaborated closely with the European elite,

---

[4] Agricultural production for the export market, focused on crops such as sugarcane, tobacco and cotton, employed African slave labour. See Angeles (2013) for an analysis of the factors determining the flow of slaves to the Americas.

Amerindians by and large retired to their rural communities where they lived a separate cultural life from the rest of society. Spanish governors referred to this network of Indian villages, where no Europeans lived permanently and where Amerindians were free to organize internal affairs according to their traditional law, as the "Republic of Indians". This "Republic" should not be understood as an independent political entity with well-defined frontiers and a recognized legal status within the colonial society. Instead, it was a name used to group all areas of Amerindian self-rule, and followed from the colonial policy of keeping Amerindians separated from the European elite yet accessible as a source of labour.

WIthin the Republic of Indians, villages were compelled to pay taxes and supply tribute in the form of labour, but colonial authorities were otherwise uninterested in dictating the day to day life of Amerindians. A number of schemes such as the *encomienda*, *repartimiento* or *mita* were put in place to recruit, transport and monitor labour tribute, but the administration of these systems typically relied on local leaders or headmen, who enjoyed privileges such as the private ownership of land and exception from taxation. In this way, as James Lang put it, "The Spanish enterprise in the New World rested on an indigenous social order" (Lang 1975, p. 7).

This separate existence of Amerindians and their freedom to organize their internal affairs was put in place to facilitate the colonizers' access to Amerindian labour, but had the unintended consequence of preserving many features of pre-colonial culture and pre-colonial institutions over the long run. The existence of a "Republic of Indians", inadvertently, kept alive the notion of large pre-colonial nations in the places where these existed. Amerindian groups which had once developed advanced political organizations were able to maintain some of that institutional capacity throughout the colonial period.

It is the central thesis of this paper that such institutional capacity was finally put to use with the advent of independence, and in particular during the 20th century, and resulted in a positive association between present-

day socioeconomic outcomes and pre-colonial institutional development. An important aspect of this thesis is that the institutional capacity of pre-colonial groups was far from uniform across the Americas. At the time of first contact with Europeans, Amerindian institutional complexity varied from the multi-layered bureaucracy administering the vast Inca Empire to the numerous small chiefdoms with no political organization beyond the village level. Figure 1 presents the situation by plotting our measure of pre-colonial institutions (to be described in detail in the next section) across all sub-national states in Latin America. The contours of the Inca and Aztec empire are in evidence, as are the relatively advanced areas of the Yucatan peninsula and central Colombia (where groups such as the Maya and the Muisca were organized at the level of confederacies of city-states). The variation observed in figure 1 implies that the effect of pre-colonial institutions on socioeconomic development would be sizeable in some regions and negligible in others.

[Figure 1]

Which, however, would be the specific mechanisms explaining a link from more advanced pre-colonial institutions, preserved in Latin America's rural communities, and socioeconomic development as measured today? While several mechanisms may be at play, we believe the historical evidence clearly supports the existence and importance of one in particular: institutionally-advanced Amerindian groups were able to organize themselves and defend their interests in front of colonial and post-colonial governments. This would have resulted in more favorable policies towards regions inhabited by Amerindians with advanced institutional capacities, in particular in the areas of education, health, and public infrastructure, and better access to natural resources, in particular agricultural land. Indeed, it is on this last point, the ownership and redistribution of agricultural land, where we believe the evidence for this general mechanism is most clearly in evidence - as we document in what follows.

The ownership of land has always been a hotly contested issue in Latin

7

America and a topic of utmost importance for Amerindians, for whom agriculture is typically the main economic activity. Already during colonial times it is possible to observe that members of the most institutionally-advanced pre-colonial groups were sometimes able to use the judicial system to challenge the dominant classes for the ownership of land. Of far more relevance, however, is the achievement of land reforms across many countries in Latin America throughout the 20th century. A land reform is one of the largest political and economic transformations that a government can try to put in place - its hallmark feature being the redistribution of land from large and rich landowners to people with little or no land. Land reforms do not take place without the concerted and sustained organization of the groups who stand to benefit from it, in this case Amerindians.

If the mechanism we put forward is correct, a look at figure 1 would be enough to predict the existence and scope of land reforms across Latin America. The most far reaching land reforms would be expected to take place in the countries occupying the core of the former Aztec and Inca empires - Mexico, Peru and Bolivia.[5] These may be followed by countries in the peripheries of such empires, such as Ecuador and Chile, and countries where other advanced pre-colonial groups existed, such as Guatemala and Colombia. Finally, countries like Argentina, Brazil, Uruguay and Paraguay, where pre-colonial groups were not organized beyond the village level, would have seen the least activity in terms of land reform. As it turns out, the pattern just hypothesized matches the historical record on land reforms almost perfectly.

The two most far-reaching land reforms of the continent took place in Mexico and Bolivia, in both cases following periods of armed conflict in which Amerindians played a leading role (the Mexican revolution of 1910 and the Bolivian revolution of 1952). About four-fifths of agricultural land was redistributed in Bolivia, and as much as half in the case of Mexico (Kay

---

[5]The capital of the Inca empire was Cuzco, located in the highlands of Peru, but as figure 1 shows a higher share of present-day Bolivia is inhabited by descendants from the Inca empire as compared to present-day Peru.

1998, p. 17). These two reforms, the first ones to take place in the continent, served as an inspiration for similar, but less successful ones, which spread across Latin America over the decades of the 1960s and 1970s.[6] For instance, while the proportion of peasants and rural workers benefiting from land redistribution was about three quarters in Bolivia and close to a half in Mexico, it was a third in Peru, one fifth in Chile, and around one tenth in Ecuador and Colombia (Kay 1998, p. 17). Other reforms taking place in Central America had varying levels of success but always below the Bolivian and Mexican cases. Finally, it is also revealing to consider the countries in which land reforms did not take place at all or at no significant degree. Argentina is the only case in Latin America where land reform and land redistributions have been completely absent, while Brazil, Uruguay and Paraguay have seen minor amounts of land redistribution but no attempts at comprehensive land reform as in the rest of the continent (Kay 1998).

It is important to note at this point that we are silent about the ultimate success of land reforms as a policy to induce development. Many land reforms in Latin America were subsequently diluted or reversed, and the formal empirical work that exists on the matter shows a negative association between land redistribution and present-day development outcomes.[7] What we do believe, however, is that the pattern and depth of land reforms across the continent provides clear evidence of the different capacities of Amerindian groups to make their interests heard; and that this capacity seems, at first sight, related to their pre-colonial levels of political organization. Thus, we regard the above as evidence of a broader pattern of success in political demands, which would result in beneficial policies towards Amerindians in areas such as health, education and poverty reduction.

Finally, while the above mechanism offers a clear linkage from precolonial institutions to socioeconomic success of Amerindian groups, we must emphasize that most of our empirical analysis will use as dependent vari-

---

[6] A third case of land reform which was also highly influential was that of Cuba (1959). The Cuban case is of course different, and Amerindians did not play any role in it.

[7] See, in particular, Dell (2012) for the case of Mexico.

ables measures of socioeconomic success for the whole population of each sub-national state, not for its Amerindian component. Of course, since Amerindians are counted within the total population their success will be reflected at the aggregate level; but there are reasons to believe that more than this simple composition effect is at play. Indeed, poverty and ignorance rarely benefit anybody - including those who are not poor. Policies to improve the human capital of Amerindians may benefit the mestizo and white population by supplying them with more able and better educated employees and business partners. The rationale that sustains the delivery of free and universal education and health services in most developed nations, namely that the externalities of these activities benefit the whole society, should also apply to the provision of education and health to minority groups such as Amerindians in Latin America.

The following sections describe how we set out to test the consequences of the above hypothesis.[8]

3 Data and methodology

As our unit of analysis is the sub-national state, the first step in our methodology involves the construction of a measure of pre-colonial institutional advancement at this level of disaggregation. We follow Gennaioli and Rainer (2007) and Michalopoulos and Papaioannu (2013) in using the level of political complexity of each ethnic group as reported in Murdock's *Ethnographic Atlas* as our measure of institutions. Before continuing, a few comments are in order regarding this data source and this particular variable.

---

[8]It is important to note that other mechanisms linking pre-colonial institutions and socioeconomic development may be at play. For example, ethnic groups with experience of large-scale political organization may be able to organize the delivery of public services at the local level, without the help of a national government, and may be better at integrating within a complex market economy. For the case of Africa, Gennaioli and Rainer (2007) emphasize that higher levels of political organization may increase the accountability of local chiefs - for instance by rendering them accountable to courts or higher-level authorities within the ethnic group. We chose not to emphasize these alternative mechanisms as evidence for their existence is lacking - unlike the case for the mechanism working via political influence at the national level.

The Ethnographic Atlas summarizes a vast anthropological literature into a single work by coding more than 60 variables for 1267 ethnic groups from around the world.[9] The variables cover a vast array of social and economic aspects such as the way in which families are organized, how political leaders are elected, and the importance of economic activities such as agriculture. While the Atlas has been an important source of information in studies of long-run development for some time now, it is important to note that most of its sources are anthropological works which involved direct observation of the ethnic groups in question between the late 19th century and the postwar period. Thus, while many African ethnic groups would have been observed during the early phases of colonialism in Africa, ethnic groups in the Americas were observed long after the establishment of European colonial empires in or around their territories. In other words, it is more difficult to claim that the information in the Atlas reflects the status of pre-colonial institutions for ethnic groups in the Americas than for ethnic groups in Africa.

While the above limitation of the data is certainly important, a number of additional considerations suggest that its use will nevertheless be informative. First, while most sources of the Atlas are anthropological works, Murdock also uses historical sources written at the time of first contact with Europeans whenever available. This is the case for the largest Amerindian groups, such as the Aztecs or the Incas, for which well-documented descriptions of their societies during the early 16th century exist. As an example, variables relating to the people of the former Inca empire use Pedro Cieza de Leon's Cronicas del Peru (1554) as a source, written by one of the earliest Spanish conquistadors in South America (Murdock 1981, p. 84). The use of historical sources guarantees that variables relating to political complexity adequately reflect pre-colonial institutions, at least for the major groups.[10]

---

[9] The Ethnographic Atlas has been expanded and updated several times since its publication. The current version, used in this paper, is due to Gray (1999) and can be accessed at http://intersci.ss.uci.edu/wiki/index.php/Ethnographic_Atlas

[10] The ethnic groups in Latin America for which 16th century sources are used in-

11

For smaller groups, we may worry that some of them lost political complexity between the colonial conquest and the time when anthropologists visited them. While this possibility cannot be discarded, it is unlikely to concern an important number of groups since the vast majority of them were organized in small chiefdoms, tribes or bands to begin with - forms of political organization which are at the bottom of the classification system used in the Atlas anyway. Furthermore, and as described in the previous section, colonial policy in Latin America typically provided large autonomy to ethnic groups in their internal affairs (the "Republic of Indians"), which would have preserved pre-colonial institutions as long as they posed no threat to colonial rulers.

In the Ethnographic Atlas, the degree of political complexity of each group is measured by the variable "Jurisdictional Hierarchy beyond the local community level". The variable takes discrete values between 0 and 4, where the value represents the number of levels of political organization above the local community. Murdock assigns a value of 0 to groups organized in bands or single-village tribes, a value of 1 to chiefdoms comprising a few villages or a single city-state, and a value of 2 for large chiefdoms with many cities or confederacies of city-states. Values 3 and 4 are reserved for states with several levels of intermediate bureaucracy between its ruler and the local community (provinces, municipalities and so on). These categories are somewhat related to the standard classification of political complexity in anthropological studies, as first formulated by Elman Service, which classifies societies into bands, tribes, chiefdoms and states (Service 1971). As discussed by Diamond (1997), the level of political complexity is closely related to technological advancement, which is needed in order to support an ever larger class of non-food producers.

For the Americas, the only pre-colonial group that achieves the maximum value of 4 in Murdock's classification is the Incas. Indeed, the Inca Empire is

---

clude: Chichimec, Tarascan, Aztec, Maya, Arawak, Carib, Chibcha (or Muisca), Inca, and Coastal Tupi (Murdock 1967).

well-recognized as the most sophisticated political and administrative structure developed in the Western Hemisphere before the European conquest (Burkholder and Johnson 1998, p. 19).[11] Perhaps surprisingly, the Aztec Empire of central Mexico, the only pre-colonial state comparable to the Inca Empire in terms of extension and population, is only assigned a value of 2. This, arguably, is due to the Aztec's particular political organization, which has been described as hegemonic or indirect. Kingdoms conquered by the Aztecs remained independent in all internal affairs, their rulers were typically not removed, and representatives of the Aztec Emperor, such as provincial governors, were largely absent. We follow Murdock's choice and use a value of 2 for the Aztec ethnic group through most of the paper; but we also subject our results to robustness checks where either the Aztec group is assigned the maximum value of 4 or the Inca group is assigned a lower value of 2. Most other Amerindian groups are assigned a value of 0 or 1 in Murdock's scale, with the exception of a few groups organized in confederacies of city-states such as the Muisca of central Colombia or the Zapotecs of southern Mexico.

We combine the above variable with data on the ethnic structure of the population for each sub-national state in Latin America to construct a population-weighted average of Murdock's Jurisdictional Hierarchy index for all states. The variable is constructed using only the population and institutional data for Amerindian ethnic groups, and as such reflects the average level of institutional complexity among the Amerindian population residing in each state. We refer to this variable as "pre-colonial institutions" in the remainder of the paper.

It is worth pointing out that the construction of this variable requires the matching of ethnic groups listed under two different datasets. Indeed, the data on population shares comes from national censuses, which we accessed individually through each nation's statistical agency, while the data

---

[11]We also assign the value of 4 to the *Aymaras*, a large Amerindian group which was part of the Inca empire and who were not assigned a value of Jurisdictional Hierarchy in Murdock (1967). Our results are not dependent on this choice.

on the characteristics of each ethnic group comes from the *Ethnographic Atlas*. While a majority of groups receive the same name in the national censuses and in the *Atlas*, there are a number of cases where the names assigned in these two sources differ. We used a diversity of additional material in order to make sure that as many ethnic groups as possible were matched - please refer to table A1 in the Appendix for details. By this procedure we were able to increase the number of Amerindian ethnic groups matched between these two sources to 102.[12]

These 102 matched Amerindian groups represent 71% of the total Amerindian population of Latin America - albeit this percentage varies significantly from country to country. The fact that almost 30% of the Amerindian population could not be matched is to be expected given that the *Ethnographic Atlas* does not offer an exhaustive list of all groups but rather a survey of the groups for which anthropological work is available. For the Amerindian groups that could not be matched, we assign the minimum value of Jurisdictional Hierarchy under the assumption that small and less organized groups were more likely to remain unresearched by anthropologists. The assumption is supported by the fact that all groups present in the Atlas with a value of Jurisdictional Hierarchy equal to 1 or higher were matched to our census data. As a robustness check, we also experiment assigning non-matched groups a value equal to the average of all matched groups within the same state. Figure 1, introduced in the previous section, offers a visual overview of our measure of pre-colonial institutions across all sub-national states in the continent.

With our measure of pre-colonial institutions at hand, we investigate its influence on present-day socioeconomic outcomes in Latin America using

---

[12]Our data uses the most recent census available for each country, as previous versions would have a less comprehensive coverage of the Amerindian population (for instance not recording the exact ethnic group). It is worth mentioning that the percentage of Amerindians in the total population may change significantly between two censuses, as census questions are modified and social attitudes towards Amerindians evolve. While we cannot say much about how our results would be affected if different census years were used, we did try using two different censuses for Bolivia (2001 and 2012) as the data was sufficiently detailed in these two cases. Our results were not affected in any material way.

the following econometric specification:

$$Y_{s,c} = \alpha_c + \beta PCI_{s,c} + \theta AmPop_{s,c} + \gamma X_{s,c} + \epsilon_{s,c} \qquad (1)$$

In equation (1), $Y_{s,c}$ is an outcome variable such as a measure of schooling, health or economic well-being. Subscript s denotes sub-national states, subscript c denotes countries, and α_c is a set of country-specific fixed effects. $PCI_{s,c}$ is the measure of pre-colonial institutions described above and $X_{s,c}$ is a set of variables controlling for state characteristics such as population density, geography, and a number of colonial and pre-colonial factors potentially affecting socioeconomic outcomes. Finally, $AmPop_{s,c}$ is the share of Amerindians in the total population of the state today, a control variable that we single out for its importance.

The share of Amerindians in the total population today is positively related to pre-colonial institutions, as the most advanced pre-colonial states were also the most densely populated (the correlation coefficient between $AmPop$ and PCI is 0.35). This leads us to control for the share of Amerindians today since this variable will have a separate effect on socioeconomic outcomes quite distinct from the one we are trying to estimate in this paper.

Indeed, Amerindians of all ethnic groups have traditionally suffered from discrimination within Latin American societies. They usually find it difficult to be accepted in non-manual employment and are given a low priority by national governments when it comes to investing in education, health or public infrastructure. Under such circumstances, it is not surprising that regions with a larger share of Amerindian population are usually characterized by lower levels of socioeconomic success.[13]

---

[13]See Psacharopoulos and Patrinos (1994) for a detailed analysis of this issue.

Our aim in this paper is to estimate the effect of pre-colonial institutions on socioeconomic development within the context of societies where the descendants of pre-colonial populations are subject to discrimination. It is therefore important to factor out the negative effect of a larger Amerindian population, which results from a larger share of people being discriminated against, from the positive effect of pre-colonial institutions that we advance. In a way, we hypothesize that while all Amerindians are subject to discrimination, those belonging to groups with more developed pre-colonial institutions would be able to overcome it by organizing and successfully defending their rights in front of the national government and society at large. It is this latter effect we try to isolate, as it is the only one directly related to the quality of pre-colonial institutions.

Most of our analysis will include the share of Amerindians in the population as a control variable but, for completeness, we also run our baseline regressions without it. The overall conclusions of our analysis do not change in this case, albeit the size and statistical significance of the coefficients of interest is uniformly weaker (see section 4.4).

All our regressions include country fixed effects as these control for a wealth of characteristics shared by all states within the same nation. In particular, a number of colonial and post-colonial institutional factors will be common to all sub-national units, such as the written laws and the constitution, the organization of public health and education, the balance of power between the different branches of the government, and so on. As we search to isolate the effect of pre-colonial institutions, accounting for as many colonial and post-colonial institutional factors as possible is important. Of course, it may still be the case that some of these factors display variability at the sub-national level: some aspects of the law, for instance, may be applied more stringently in the capital city as compared to far-off provinces. We cannot control directly for such effects, but note that if the

way in which national institutions such as legal codes apply locally corre-
lates with the local level of pre-colonial institutions, a potential explanation
may be that more advanced pre-colonial institutions facilitate the operation
of national law. In that case, our regressions would simply be capturing an
additional indirect effect of pre-colonial institutions, one that works via a
more effective functioning of the post-colonial state.

We will consider as dependent variables three indicators of education
(percent of the population who completed primary education, percent who
completed secondary education, average years of schooling), one indicator of
public health (infant mortality rate), two indicators of economic well-being
(percent of the population with access to drinking water, percent with access
to electricity), and two indicators of overall economic development (GDP
per capita and poverty rates). The battery of control variables at the state
level will be discussed in the following section, as they are progressively
introduced. The sources and precise definitions of all variables used in the
paper can be found in table A2 in the Appendix.

Finally, it is worth pointing out that, given the nature of Murdock's
dataset, the variable $PCI_{s,c}$ may suffer from measurement error. This, how-
ever, would only result in a bias towards zero in our estimates of coefficient
β. As most of our results rely on β being different from zero, we may say
that conclusions would be stronger if this variable could be measured with
more precision.

Our empirical analysis proceeds in two phases. First, section 4 analyses
the relationship between our measure of pre-colonial institutions and so-
cioeconomic success for the whole population, while section 5 will consider
finer partitions of the population within each sub-national state. The reason
for this procedure is that the number of dependent variables and the set of
controls available is much richer for the specification using the whole popu-
lation of each state. We aim to provide a more complete picture by focusing
initially on the whole population, and turning to a finer level of analysis at
the cost of less available data later on.

Before turning to the empirical analysis, tables 1 and 2 offer an overview of our data. Table 1 lists all countries in our dataset together with their total population, the percentage of Amerindians in their population, and the percentage of their Amerindian population being matched to the Ethnographic Atlas. Our data covers all 17 countries in continental Latin American for a total of 324 sub-national states.[14] Table 2 presents descriptive statistics for all variables used in our analysis.[15]

[Table 1]

[Table 2]

4 Empirical analysis: whole population

4.1 Baseline results

Throughout this section our dependent variable measures socioeconomic success for the whole population of each sub-national state. We begin our analysis with a set of regressions that illustrate how our coefficient of interest changes as different sets of control variables are added. All regressions include country fixed effects, and results are reported in table 3. The dependent variable for this initial analysis is the percentage of the population who finished secondary schooling, in logarithmic form, but similar results

---

[14] For Brazil the population shares of different Amerindian groups is only available at the level of regions (groups of 3 to 9 states). We assign to each Brazilian state the population shares of the region it belongs to. For Argentina the data is available at the state level but gives only a partial breakdown, with the population of only the main Amerindian groups of each state being available. We complete the missing data for Argentina using national totals for each group and assumptions about the distribution of each group outside the states where they are most numerous. For Uruguay we do not have data on different Amerindian groups, only the population share of all Amerindians in each state. This, however, is not a problem for the construction of our measure of Pre-Colonial Institutions for Uruguay as we know that all Amerindian groups in Uruguay have a value of zero for Jurisdictional Hierarchy. For all other 14 countries we have a complete dataset giving population shares for all Amerindian groups in every state.

[15] A matrix of bivariate correlations can be found as table A3 in the appendix. Correlations confirm that the share of Amerindians in the population is negatively related to all measures of socioeconomic success.

are obtained using our other measures of well-being. Standard errors are clustered at the country level throughout the paper.

[Table 3]

The first column of table 3 reports our most simple regression where only the measure of pre-colonial institutions is included alongside fixed effects. The coefficient on pre-colonial institutions is already statistically significant at the 5% level and takes a value of 0.0376. The coefficient doubles in magnitude to 0.0893 and its statistical significance increases to 1% in the second column, where the share of Amerindians in the total population of each state is added. As expected, this last variable has a strong negative influence on our outcome measure and its absence was responsible for a negative bias in the effect of pre-colonial institutions. We are thus confirming previous research as to the overall poorer socioeconomic outcomes of the Amerindian population, but adding a new result whereby areas where Amerindians groups were characterized by more advanced pre-colonial institutions have better outcomes.

The next four columns of table 3 add a large number of state-specific characteristics which may have an effect on socioeconomic outcomes and whose absence could create an omitted variable bias. In column 3 we control for the present-day population density of each state - as the provision of education (and public services in general) may be more costly in less densely settled territories. As expected, areas of higher population density tend to have better outcomes, but the effect of pre-colonial institutions continues to be large and statistically significant.

Columns 4 to 6 deal with the important issue of geography, including aspects such as climate, location and natural resources. Following Diamond (1997), advanced pre-colonial institutions may be expected to arise in regions with favorable geographic characteristic such as an abundance of agricultural land, proximity to the coast, or a mild climate. If that was the case, higher socioeconomic outcomes today may be due to the permanence

of such geographic characteristics, and not to the superior institutions that the groups living in these areas developed.

With this in mind, column 4 augments our baseline regression with three indicators of climate (latitude, altitude, and temperature), an index of terrain ruggedness (taking higher values when land elevation is irregular), and an index of Malaria prevalence. Column 5 adds to this list the area of the state in question plus three indicators of its locational advantage: distance to the capital, distance to the sea, and a dummy for landlocked states. Column 6, finally, directly measures the most important forms of natural resource wealth by adding an index of land suitability for agriculture and dummy variables indicating the presence of oil or gas fields, gold or silver mines, and any other mines.

The main result of these three columns is that the coefficient on pre-colonial institutions remains statistically significant at the 1% level in all cases and its magnitude is not much affected. In column 6, when all controls are included, the coefficient takes a value of 0.0775 - a similar magnitude as in column 2. Since our dependent variable is measured in logarithmic form, this coefficient indicates that an increase in the average level of pre-colonial institutions by 1 unit is associated with an increase in secondary school achievement of around 8%. This is a large effect when we consider an average value of secondary school achievement across all states of 41% with a standard deviation of 15% - passing from a pre-colonial population of tribesmen to one of multi-city chiefdoms (increase of 2 units) would lead to a one standard deviation change in secondary schooling.

Turning to our state-specific indicators of geographic advantage, latitude, terrain ruggedness, distance to the capital and distance to the sea all appear to have a statistically significant relationship with our outcome variable. Latitude and ruggedness seem to pick up most of the effect from climatic factors, as neither malaria prevalence, temperature or altitude are statistically significant in their presence. Somewhat surprisingly, distance to the capital is positively related to education while being landlocked appears

20

to carry no additional penalty once distance to the coast has been accounted for. Finally, none of our four indicators of natural resource wealth exerts a statistically significant effect on secondary school achievement. Arguably, the gains from having these resources at hand are counteracted by opposite effects much discussed in the 'natural resource curse' literature.

The results of table 3 may be reproduced over our full array of socioeconomic indicators, as shown in table 4. This table takes as its baseline the regression reported in the last column of table 3, with all state characteristics considered so far, and considers as dependent variable each of the eight outcome measures at our disposal. Remarkably, our index of pre-colonial institutions is consistently related with better outcomes for all of them: it is positively related with measures of education, drinking water, electricity and GDP per capita, and negatively related with infant mortality and poverty rates. In all cases the relationship is statistically significant at the 5% level or better.

[Table 4]

The effect of pre-colonial institutions is not only statistically significant, the magnitude of the effect is also large. Since all dependent variables are used in logarithmic form, coefficients may be interpreted directly as semi-elasticities. Interestingly, the largest effects are observed for our measures of overall economic development. A 1-unit increase in the index of pre-colonial institutions is associated with a 19% increase in GDP per capita and a 12% decrease in the poverty rate. The effects for all other dependent variables are in the 3 to 8% range for a 1-unit increase, in all cases a sizeable change.

4.2   Controlling for colonial economic activity

While our results so far have identified a positive relationship between pre-colonial institutions and current socioeconomic development, the effect may not be a direct one, as hypothesized so far, but an indirect one working via the colonial process. In accordance

with the work of Arias and Girod (2014), pre-colonial institutions may be a major explanatory factor behind the choice of economic activity put in place during the colony. By controlling for the different types of economic activity during colonial times we test for the existence of this indirect channel, challenging the notion that pre-colonial institutions are important in their own right with the notion that they matter mainly because of their influence in the subsequent colonial period. In other words, this section offers an exploration into one potential channel linking pre-colonial institutions and current development which is different from the one emphasized until now.

To test for this alternative explanation we take advantage of the work of Bruhn and Gallego (2012), who investigate the role of different economic activities put in place during colonial times on economic development in Latin America. Bruhn and Gallego (2012) use states as the unit of analysis, which renders their dataset compatible with ours. They classify states into four mutually exclusive groups according to the main economic activity taking place in their territory during the colony. These four groups are:

a) Mining. In particular the gold mines of Brazil, the silver mines of Mexico, Peru and Bolivia, and the associated mines producing mercury for the process of silver extraction through amalgamation.

b) Plantations. Places dedicated to the cultivation of high-value cash crops for the export market, in particular sugarcane, tobacco and cotton. Plantations relied essentially on slave labour.

c) Other colonial activities. Places where the dominant economic activity was agricultural production for the local market (from Amerindian lands or from latifundia) and industry.

d) No colonial activities. Places where the colonial state had marginal or no influence, like remote parts of the Amazonian rainforest and the

22

extreme south of Argentina and Chile.[16]

We incorporate in our regressions dummy variables for the first three types of economic activities, leaving the case of no colonial activities as our excluded category. If the effects of pre-colonial institutions only works via colonial activities, the inclusion of these controls would render our coefficient of interest small and not significant. Results are reported in table 5, where all regressions include country fixed effects and our full range of baseline controls.

[Table 5]

Table 5 is strongly supportive of our thesis. Indeed, our measure of pre-colonial institutions continues to have a positive and statistically significant effect on the eight dependent variables we consider. The size of the coefficients is not much affected with respect to table 4, only the effects on educational achievement and electricity provision are reduced to some extent. This indicates that the relationship between pre-colonial institutional development and present-day outcomes is largely *not* mediated by the type of economic activity put in place during the colony - in accordance with our thesis.

Turning to the effects of colonial activities on present-day outcomes, table 5 gives us a mixed picture. The effect seems clearest on overall measures of economic development, as states associated with mining and plantation agriculture have lower levels of GDP per capita than states left untouched by the colonial economy. This is in line with Bruhn and Gallego (2012), who base most of their analysis on the effects on GDP per capita. For other measures of socioeconomic development, however, the evidence is less conclusive. Areas where slave-based plantations were located are indeed characterized

---

[16]Bruhn and Gallego (2012) combine the information on the type of economic activity in each state with data on pre-colonial population density to produce a classification into three types of colonial activities which they refer to as "bad", "good" and "ugly". We don't follow their approach as it incorporates value judgements as to what is believed to be "good" or "bad" (let alone "ugly"). The classification of colonial activities into mining, plantations, and others is much less likely to be affected by our own beliefs.

by lower secondary education and higher infant mortality (statistically significant at the 10% level), while areas where mining activities took place have higher poverty rates (statistically significant at the 5% level). On the other hand, no further statistically significant effects are estimated for any of the other socioeconomic outcomes we consider - albeit estimated coefficients have usually the expected sign. Overall, while colonial activities may well play a role in determining current development outcomes, our results show that their consideration does not diminish the importance of pre-colonial institutions.

4.3   Controlling for other pre-colonial characteristics

If the results so far clearly point towards a persistent role of pre-colonial institutions on current socioeconomic development, one may still argue that pre-colonial features other than institutional complexity may explain our findings. As we mentioned briefly, institutional complexity usually correlates with economic development, and it is possible that richer pre-colonial societies were able to adapt better and take advantage of the new colonial environment simply because of their wealth. Furthermore, the *Ethnographic Atlas* provides a large array of cultural and economic practices of the societies it surveys. We are therefore in a position to control for a number of pre-colonial characteristics other than the complexity of their political structure - and we do so in what follows.

   We start with overall economic development in pre-colonial times. Clearly, measures of income per head are not available for this time period in the Americas, but we may follow much of the relevant literature and rely on estimates of population density as a proxy for overall economic development (see, for instance, Acemoglu et al. 2002). The data on pre-colonial population density at the state level comes from Bruhn and Gallego (2012), and table 6 adds this variable as an additional control to the regressions reported in table 5.[17] Furthermore, the variable on pre-colonial population density is

---

[17]We note that table 6 is thus controlling for both pre-colonial and present-day population density. The correlation between these two variables is positive but not too high

interesting in its own right, as it is a means to test the "reversal of fortune" hypothesis, whereby areas which were richer than average prior to the arrival of Europeans would have a tendency to be poorer than average today (Acemoglu et al. 2002).

[Table 6]

Once again, results are fully consistent with the thesis of this paper. The coefficient on pre-colonial institutions is hardly affected by the inclusion of this variable and remains statistically significant for all the dependent variables we consider. It is not the case, then, that areas of high institutional development are better off because of an initial advantage in terms of wealth.

Turning to the "reversal of fortune" hypothesis, this is only partially supported by our results - the coefficient on pre-colonial population density having the expected negative effect on socioeconomic development only for our measures of education, GDP per capita and poverty rates. Statistical significance is reached only for GDP per capita, in accordance with previous findings in the literature.[18]

In table 7 we take an additional step and control for nine social and economic characteristics of pre-colonial societies other than their institutional complexity. These characteristics are the fraction of the population dedicated to gathering, hunting, fishing and agriculture; their typical pattern of settlement (from fully nomadic to compact and permanent settlements); their degree of class stratification; a dummy for the existence of slavery; a dummy for the existence of elections in determining leader succession and, finally, a dummy for the existence of inheritance rules for property (see table A2 in the Appendix for detailed definitions). To construct each of them,

---

(0.47). While Europeans did settle in larger numbers near the areas of high Amerindian population during colonial times, the trend reversed strongly following independence. Some of the most densely populated regions in present-day Latin America, such as the areas around the cities of Sao Paulo, Rio de Janeiro, Buenos Aires and Santiago, were very sparsely populated in the year 1500.

[18]For alternative views on the reversal of fortune hypothesis see Chanda et al. (2014) and Maloney and Valencia (forthcoming). We note that, contrary to the rest of the literature, we are testing the reversal of fortune only among Latin American countries.

we proceed as for our measure of pre-colonial institutions: we calculate the population-weighted average among all Amerindian groups present in the state.[19]

It is interesting to note that variables such as the existence of elections (a measure of proto-democracy) and the existence of inheritance rules (a measure of proto-property rights) are essentially uncorrelated with our measure of pre-colonial institutions (bivariate correlation of -0.04 with the first one and 0.10 with the second one). Thus, political complexity is truly a separate dimension along which societies may be described, and its effects may be quite distinct from those of democracy or property rights.

Each of these additional pre-colonial characteristics is considered separately in the columns of table 7, with the exception of the first four measures, all relating to the economic activity of the population, which are included simultaneously in column 2.[20] The regressions also control for the different colonial activities as in table 5 and for pre-colonial population density as in table 6, besides all the state-specific characteristics that have been included all along. The dependent variable is the percentage of the population with secondary education.

[Table 7]

As table 7 makes clear, the inclusion of these additional pre-colonial characteristics does not challenge the importance of pre-colonial institutions. In all regressions the coefficient on our measure of pre-colonial institutions remains positive and statistically significant at the 1% level. The magnitude of the coefficient is remarkably consistent, fluctuating closely around the value of 0.0700 in all but one case (column 2, where the coefficient equals 0.0918). Thus, the coefficient is usually very similar to what is obtained before any of

---

[19]For the Amerindian groups that could not be matched to the Atlas we assign a value equal to the average value of all other groups within the state. Uruguay is excluded from table 8 as we don't have enough data to calculate these additional variables for it.

[20]These four variables do not sum up to 1, as a fraction of the population may be counted in more than one of them, and sometimes in none of them.

these additional pre-colonial characteristics is controlled for (first column of table 8). In most cases, the additional pre-colonial characteristic considered turns out to have no statistically significant effect on development. In particular, nascent forms of democracy (column 6) and property rights (column 7) are not associated with better outcomes today. This reinforces our thesis, which regards institutional complexity as the crucial aspect of pre-colonial societies affecting socioeconomic development.

The exercise of table 7 may be reproduced using the other seven dependent variables considered previously. While we do not report these results for conciseness, we have carried them out and the importance of pre-colonial institutions is never challenged. The sign and statistical significance of pre-colonial institutions carries through for all seven alternative outcome variables and in essentially all specifications considered in table 7 (results are available upon request).

4.4   Robustness checks

We have carried out a number of robustness checks on the above results, some of which we have already referred to. First, we assign to all non-matched Amerindian groups from our census data a value of the index of Jurisdictional Hierarchy equal to the average value of all matched groups within the state (instead of a value of 0). Second, we have tried assigning different values of Jurisdictional Hierarchy to the Aztec and Inca groups, the two largest in Latin America. One variation assigns the maximum value of 4 to the Aztecs, bringing them in line with the Incas. Another variation brings the Inca value down to 2, in line with the Aztecs and other ethnic groups. This last change de facto reduces the range of our measure of institutional development from 0-4 to 0-2. Finally, we have tried excluding Brazil and Argentina from our regressions, as these two countries do not have complete data on the ethnic composition of their population at the state level. In all cases, our results carry through all these checks unchallenged with only minor quantitative changes in the estimated coefficients.

Following our previous discussion about the role of the share of Amerindians in the total population as a control variable, we have also estimated all baseline regressions without this control. The results, reported in table A4 in the appendix, are still supportive of our thesis as the coefficients on pre-colonial institutions continue to be associated with higher levels of socioeconomic development across the board. On the other hand, and given that the negative effect of discrimination is not factored out, the coefficients are smaller in magnitude and reach statistical significance in only two cases (secondary education and poverty rates).

There is also a point in carrying out our regressions using a more parsimonious approach to the set of controls included. In the preceding sections, we have always expanded our control set when additional variables are considered - a correct approach if all the controls are exogenous but problematic if some of them happen to be endogenous. An alternatively approach would see us adding each additional control variable separately to our baseline regression, and we have also proceeded this way. Results are reported in table A5 in the appendix, and demonstrate that the conclusions of our analysis are unchanged by this procedure.

5 Empirical results: extensions

Our results so far provide solid evidence for a positive effect of pre-colonial institutions on the socioeconomic development of the whole population at sub-national level. While these results are in accordance with our thesis, they would also confirm alternative stories linking the pre-colonial past with the present - for instance, a different mechanism intermediating between pre-colonial institutions and outcomes, or a different explanation for the persistence of these institutions. The objective of the present section is to offer additional evidence which narrows down the set of acceptable explanations for the results presented so far, and further confirm the plausibility of the thesis we advance. Within each sub-national state, we compare results for urban as opposed to rural areas and for Amerindians as opposed to

non-Amerindian people. While data at this level of aggregation is available only for a subset of the variables considered so far, the results are worth highlighting.

5.1   Comparing rural and urban regions

As we have argued in section 2, our thesis emphasizes the survival of pre-colonial institutional capacity thanks to the existence of a "Republic of Indians" throughout rural Latin America. If this hypothesis is correct, we should find that the positive relationship between pre-colonial institutions and socioeconomic development is stronger among rural areas: Amerindians living in rural areas would have a larger capacity to organize effectively and have their demands addressed.

To test this aspect of our thesis, we have gathered data allowing us to run separate regressions for the rural and urban regions of Latin America at the level of sub-national states. We were able to find separate values for the rural and urban regions of each state for four dependent variables (primary education, secondary education, access to drinking water, access to electricity) and for the ethnic composition of the population, which allows us to calculate the percentage of Amerindians in the total population and to construct distinct measures of pre-colonial institutions for the rural and urban regions of each state. We have also included our control set for geography, climate and natural resources, but the values of these controls do not change between the rural and urban area of any given state.[21]

As it turns out, the results of this exercise clearly support our thesis - as reported in table 8. We consider the four dependent variables mentioned above, and for each case run separate regressions using only rural areas and only urban areas. The effect of pre-colonial institutions is positive and statistically significant for the four cases covering rural areas, while only two of the four cases covering urban areas reach statistical significance. More

---

[21]Argentina is omitted from this exercise, as there is no information on the distribution of its Amerindian population between urban and rural areas.

important, the coefficient on pre-colonial institutions is always much larger for rural areas - between two and three times larger for primary education, secondary education and drinking water, and larger still for electricity. As an example, a one-unit increase in the measure of pre-colonial institutions would be associated with an increase of 8.6% in secondary education in rural areas against an increase of 3.2% in urban areas. We conclude that these results further solidify our thesis.

[Table 8]

5.2   Comparing Amerindians with non-Amerindians

As we discussed in section 2, the mechanism we put forward would explain why Amerindians benefit from more advanced pre-colonial institutions, and hypothesized that the externalities of the process of development mean that non-Amerindians would benefit as well. So far there has been no way of knowing whether the positive effect on development that we find takes place only among Amerindians, among both Amerindians and non-Amerindians, and how the magnitudes of these two channels compare.

To shed some light on this issue, we have managed to gather separate measures of primary and secondary education for the Amerindian and non-Amerindian population at the level of sub-national states for all Latin American countries except Brazil. We run separate regressions using only Amerindians and only non-Amerindians, and control for all state-specific aspects considered so far. Our results are reported in table 9 and reveal a number of points.

[Table 9]

First, table 9 brings support to the idea that the positive effect of pre-colonial institutions concerns both Amerindians and non-Amerindians: the coefficient of interest is positive in all cases and statistically significant in three of them. Thus, the notion that some form of positive externality favouring non-Amerindians is in place seems reinforced. Second, and

30

as expected, the estimated effect is larger for Amerindians than for non-Amerindians, although the difference is perhaps not as large as one may have anticipated (the coefficient for Amerindians is about 60% larger for primary education and about 20% larger for secondary education). Further investigation into the mechanisms behind these results is clearly an avenue for future research, but at this point the hypothesis that pre-colonial institutional development is beneficial for the population at large can be upheld.

Finally, as an additional test to our findings, we have looked for evidence of the positive effect of pre-colonial institutions on the total population at an earlier time period. The data requirements for this exercise cannot be met for Latin America as a whole, but an analysis using only Mexican states has been possible. We collected historical data on literacy rates and on the ethnic composition of the population using the Mexican censuses of 1921, 1930 and 1950. After constructing our measure of pre-colonial institutions using the ethnic composition of the population in 1930 and 1950 (the years for which it is available), we run a set of regressions similar to those performed so far in the paper using literacy rates as the dependent variable. The results, which can be found in table A6 in the appendix, show that a positive relationship is indeed in place - although the number of observations at our disposal is too limited for coefficients to reach standard levels of statistical significance.

Overall, the additional findings reported in this section substantially increase our confidence on the validity of the thesis put forward all along this paper.

6   Concluding remarks

If one thing has been learned from the last two decades of research on economic development over the very long run it is that the past cannot be easily cast aside. Every society builds on the successes and mistakes of its predecessors, and inherits a set of rules and institutions that are usually modified only gradually. While this seems obviously true for the "winners" of economic history, the European nations that colonized the world, it is

also the case for the "losers", those nations being colonized. What came out of the colonizing process throughout the world was not a mirror image of European society but a new reality where pre-colonial culture and institutions survived, often below a layer of official or dominant culture. These two layers interact and modify each other, and both of them ought to be considered in the study of today's developing countries.

This paper brings support to the above assertions, and adds to the substantive evidence already in place for the case of Africa. As our empirical results show, Latin American pre-colonial institutions - and more precisely the degree of political complexity - are powerful predictors of present-day measures of socioeconomic development. Several aspects render our evidence particularly convincing. First, our results are obtained controlling for country fixed effects, thus factoring out many institutional factors playing a role at the national level. Second, we introduce a large array of controls for geographic factors including climate, location and the presence of natural resources. Third, we consider additional historical forces such as the type of economic activity in place during the colony and the economic and social profile of pre-colonial societies (besides their institutional complexity). Finally, we show how the influence of pre-colonial institutions is far stronger in rural areas, which is in accordance with the historical account we give for the transmission of pre-colonial factors, and that both Amerindians and non-Amerindians appear to benefit.

The present paper, together with the literature it contributes to, enhances our understanding of how developing countries got to where they are now. Understanding this is important in its own right, and increases the chances of making the right decisions when considering where they head to in the future.

References

Acemoglu, D. and Dell, M., 2010, Productivity differences between and within countries, American Economic Journal: Macroeconomics 2, 169-188.

Acemoglu, D., Johnson, S. and Robinson, J. A, 2001, The colonial origins of comparative development: an empirical investigation, American Economic Review 91 (5), 1369-1401.

Acemoglu, D., Johnson, S. and Robinson, J. A, 2002, Reversals of Fortune: Geography and Institutions in the making of the modern world income distribution, Quarterly Journal of Economics.

Acemoglu, D., Bautista, M. A., Querubin, P. and Robinson, J. A., 2008, Economic and political inequality in development: the case of Cundinamarca, Colombia, in: Helpman, E. (ed.), Institutions and Economic Performance, Cambridge, MA: Harvard University Press.

Acemoglu D., Gallego F. A, and Robinson, J. A., 2014, Institutions, Human Capital and Development, Annual Review of Economics 6, 875–912.

Angeles, L., 2013, On the Causes of the African Slave Trade, Kyklos 66 (1), 1-26.

Bockstette, V., Chanda, A. and Putterman, L., 2002, States and markets: the advantage of an early start, Journal of Economic Growth 7, 347-369.

Bruhn, M. and Gallego, F. A., 2012, Good, bad, and ugly economic activities: do they matter for economic development?, Review of Economics and Statistics 94 (2), 433-461.

Burkholder, M. A. and Johnson, L. L., 1998, Colonial Latin America, Third Edition, Oxford: Oxford University Press.

Chanda, A., Cook, C. J. and Putterman, L., 2014, Persistence of fortune: Accounting for population movements, there was no post-columbian reversal, American Economic Journal: Macroeconomics 6, 1-28.

Comin, D., Easterly, W. and Gong, E., 2010, Was the wealth of nations determined in 1000 BC?, American Economic Journal: Macroeconomics 2, 65-97.

Dell, M., 2010, The persistent effects of Peru's mining mita, Econometrica 78 (6), 1863-1903.

Dell, M., 2012, Path dependence in development: Evidence from the Mexican Revolution, mimeo, Harvard University.

Diamond, J., 1997, Guns, germs, and steel, London: W. W. Norton & Company.

Elizalde, A., 2016, Pre-colonial institutions and long-run development in Latin America, PhD. dissertation, University of Glasgow.

Gennaioli, N and Rainer, I., 2007, The modern impact of precolonial centralisation in Africa, Journal of Economic Growth 12, 185-234.

Gennaioli, N., La Porta, R., Lopez de Silanes, F. and Shleifer, A., 2014, Growth in regions, Journal of Economic Growth 19, 259-309.

Gray, J. P., 1999, A corrected Ethnographic Atlas, World Cultures Journal 10 (1), 24-85.

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

Huillery, E., 2011, The impact of European settlement within French West Africa: Did pre-colonial prosperous areas fall behind?, Journal of African Economies 20 (2), 263-311.

Kay, C., 1998, Latin America's agrarian reform: lights and shadows, Land Reform, Land Settlement and Cooperatives 1998/2, 8-31.

Lang, J., 1975, Conquest and Commerce: Spain and England in the Americas. New York: Academic Press.

Maloney, W. F. and Valencia Caicedo, F., forthcoming, The persistence of (subnational) fortune, The Economic Journal, doi: 10.1111/ecoj.12276.

Michalopoulos, S. and Papaioannou, E., 2013, Pre-colonial ethnic institutions and contemporary African development, Econometrica 81 (1), 113-152.

Michalopoulos, S. and Papaioannou, E., 2014, National institutions and subnational development in Africa, Quarterly Journal of Economics 129, 151-213.

Murdock, G. P., 1967, Ethnographic Atlas: A Summary, Ethnology 6

(2), 109-236.

Murdock, G. P., 1981, Atlas of World Cultures, Pittsburgh, PA: University of Pittsburgh Press.

Naritomi, J., Soares, R. R., and Assuncao, J. J., 2012, Institutional development and colonial heritage within Brazil, Journal of Economic History 72 (2), 393-422.

Nunn, N., 2008, The long-term effects of Africa's slave trades, Quarterly Journal of Economics 123 (1), 139-176.

Nunn, N., 2014, Historical Development, in (eds.): Aghion, P. and Durlauf, S. N., Handbook of Economic Growth, volume 2A, 347-402, Elsevier.

Nunn, N. and Wantchekon, L., 2011, The slave trade and the origins of mistrust in Africa, American Economic Review 101 (7), 3221-3252.

Psacharopoulos, G. and Patrinos, H. A., 1994, Indigenous People and Poverty in Latin America, Washington: The World Bank.

Service, E. R., 1971, Primitive social organization: an evolutionary perspective, 2nd edition, New York: Random House.

Weidmann, N. B., Rod, J. K. and Cederman, L.-E., 2010, Representing ethnic groups in space: A new dataset, Journal of Peace Research 47 (4), 491-499.

Figure 1
Pre-colonial institutions in Latin America

Pre-colonial Institutions

- 0.000 - 0.300
- 0.301 – 1.000
- 1.001 – 3.000
- 3.001 – 4.000
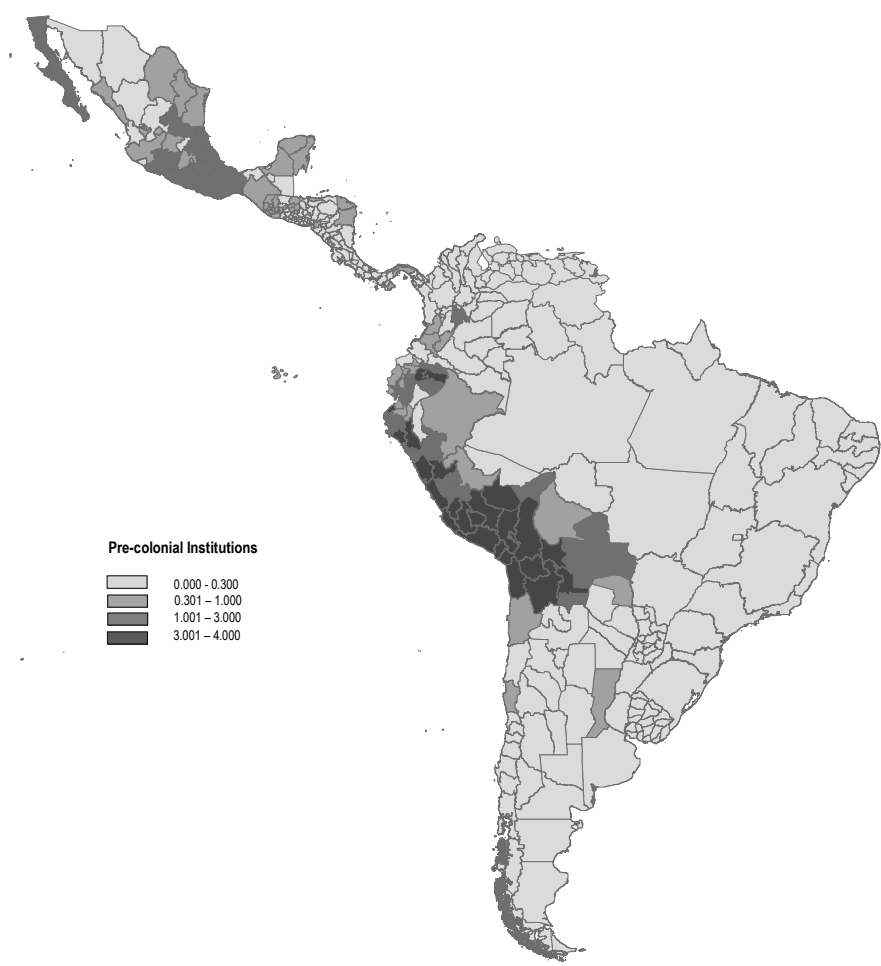
Table 1
Latin American countries and their Amerindian populations

Country	Total population	Amerindian population as % of total population	Amerindian population matched to Ethnographic Atlas as % of total Amerindian population	Number of states
Argentina	40,117,096	2.8%	25%	24
Bolivia	10,059,856	41%	84%	9
Brazil	190,755,799	0.4%	38%	27
Chile	15,116,385	4.6%	96%	13
Colombia	41,174,853	3.4%	47%	33
Costa Rica	4,301,712	1.6%	26%	7
Ecuador	14,451,115	7%	42%	24
El Salvador	57,44,113	0.2%	15%	14
Guatemala	11,237,196	39%	62%	8
Honduras	6,076,885	6.3%	96%	18
Mexico	103,263,388	5.7%	77%	32
Nicaragua	5,483,447	8%	27%	17
Panama	3,405,813	12%	29%	12
Paraguay	5,163,198	1.7%	55%	18
Peru	27,412,157	15%	96%	25
Uruguay	3,285,877	2%	100%	19
Venezuela	27,225,775	2.8%	79%	24
TOTAL	514,274,665	5%	71%	324

Table 2
Descriptive Statistics

	Observations	Mean	Std. Dev.	Min	Max
Dependent variables:					
Infant mortality rate	324	20.6	10.8	1.4	56.4
Years of schooling	319	5.98	2.13	1.03	11.45
Primary school achievement	324	0.81	0.16	0.29	0.96
Secondary school achievement	324	0.41	0.15	0.05	0.75
Drinking water	324	0.86	0.17	0.05	0.99
Electricity	324	0.84	0.19	0.03	0.99
PPP GDP per capita	300	5763.52	4722.30	1248.88	40449.09
Poverty rate	272	29.18	20.78	1.23	81.67
Main variable of interest:					
Measure of pre-colonial institutions	324	0.57	1.06	0	3.99
Baseline control variables:					
Share of Amerindians in total population	324	0.11	0.19	0.00001	0.96
Population density	324	394.9	3407.41	0.13	58706.88
Land Suitability Index	321	0.56	0.29	0.002	0.998
Malaria Stability Index	321	1.3	1.4	0	5
Latitude	324	16.02	10.73	0.015	54.33
Altitude (km.)	324	0.68	0.92	0	4.33
Temperature (Celsius)	319	20.72	5.28	4.7	27.77
Ruggedness Index	324	1.43	1.07	0	4.75
Land area (sq. km.)	324	63786.14	151196.8	44	1600000
Landlocked dummy	324	0.54	0.49	0	1
Distance to capital (km.)	324	464.08	477.69	0	2559.34
Inverse distance to coast	320	0.89	0.1	0.54	0.99
Oil & Gas dummy	324	0.16	0.36	0	1
Gold & Silver dummy	324	0.12	0.32	0	1
Other mines dummy	324	0.23	0.42	0	1
Additional control variables					
Colonial activities					
Mining	283	0.14841	0.356136	0	1
Plantations	283	0.081272	0.273737	0	1
Other colonial activities	283	0.650177	0.477759	0	1
No colonial activities	283	0.120141	0.325703	0	1
Pre-colonial characteristics					
Log of pre-colonial population density	301	0.51	2.42	-9.58	5.97
Gathering	324	0.09	0.07	0.02	0.43
Hunting	324	0.11	0.07	0.02	0.40
Fishing	324	0.10	0.06	0.02	0.31

Agriculture	324	0.47	0.21	0.02	0.93
Settlement Pattern	324	5.11	1.74	0.15	7.98
Class Stratification	324	2.03	1.19	0	4.99
Slavery	324	0.21	0.34	0	1
Elections	324	0.06	0.20	0	0.99
Inheritance Rules for Property	324	0.22	0.37	0	1

Table 3
Baseline results

Dependent variable: Percent of the population having completed Secondary education (in logs)

	(1)	(2)	(3)	(4)	(5)	(6)
Pre-Colonial Institutions	0.0376** [0.0148]	0.0893*** [0.0214]	0.0750*** [0.0237]	0.0780*** [0.0211]	0.0811*** [0.0199]	0.0775*** [0.0196]
Share of Amerindian population		-0.746*** [0.192]	-0.632** [0.222]	-0.586** [0.223]	-0.597** [0.223]	-0.582** [0.222]
Log population density			0.0279** [0.0105]	0.0366** [0.0137]	0.0533*** [0.0151]	0.0574*** [0.0123]
Latitude				0.00808** [0.00303]	0.00931*** [0.00281]	0.00916*** [0.00267]
Malaria Stability Index				-0.00733 [0.0139]	-0.0113 [0.0125]	-0.0102 [0.0122]
Temperature (Celsius)				-0.00456 [0.00436]	-0.00679 [0.00480]	-0.00557 [0.00463]
Altitude (km.)				-0.0241 [0.0159]	-0.0243 [0.0170]	-0.0277 [0.0178]
Ruggedness Index				-0.0610** [0.0227]	-0.0542** [0.0212]	-0.0534** [0.0217]
Land area (sq. km.)					3.56e-08 [8.39e-08]	1.73e-08 [8.10e-08]
Landlocked dummy					-0.0650 [0.0527]	-0.0572 [0.0505]
Distance to capital (km.)					6.80e-05*** [2.13e-05]	5.79e-05 [3.63e-05]
Inverse distance to coast					-0.856*** [0.274]	-0.822** [0.289]
Land Suitability Index						-0.0584 [0.125]
Oil & Gas dummy						0.00740 [0.0318]
Gold & Silver dummy						0.0457 [0.0425]
Other mines dummy						0.0159 [0.0334]
Country Fixed Effects	YES	YES	YES	YES	YES	YES
Observations	324	324	324	317	317	317
Adjusted R-squared	0.703	0.767	0.775	0.787	0.794	0.793

Notes: Cluster standard errors at country level are in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 4
Baseline results with 8 different dependent variables

Dependent variable	Infant Mortality (1)	Years of Schooling (2)	Primary education (3)	Secondary education (4)	Drinking water (5)	Electricity (6)	Log GDP per capita (7)	Poverty rates (8)
Pre-Colonial Institutions	-0.0631***	0.0510**	0.0248**	0.0775***	0.0562***	0.0827***	0.190**	-0.115***
	[0.0195]	[0.0176]	[0.0103]	[0.0196]	[0.0181]	[0.0265]	[0.0814]	[0.0379]
Controls included:								
Share of Amerindian pop.	YES	YES	YES	YES	YES	YES	YES	YES
Log population density	YES	YES	YES	YES	YES	YES	YES	YES
Geography, location and natural resources	YES	YES	YES	YES	YES	YES	YES	YES
Country fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	317	316	317	317	317	317	297	270
Adjusted R-squared	0.787	0.821	0.878	0.793	0.504	0.651	0.641	0.847

Notes: Cluster standard errors at country level are in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 5
Controlling for colonial activities

Dependent variable:	Infant Mortality (1)	Years of Schooling (2)	Primary education (3)	Secondary education (4)	Drinking water (5)	Electricity (6)	Log GDP per capita (7)	Poverty rates (8)
Pre-Colonial Institutions	-0.0643***	0.0394**	0.0162**	0.0642***	0.0489**	0.0563**	0.190**	-0.117**
	[0.0209]	[0.0140]	[0.00608]	[0.0170]	[0.0212]	[0.0228]	[0.0802]	[0.0436]
Other colonial activities	0.0454	-0.0188	-0.00608	-0.0321	-0.0519	-0.0249	-0.124	0.139
	[0.0501]	[0.0260]	[0.00792]	[0.0352]	[0.0353]	[0.0189]	[0.108]	[0.120]
Mining colonial activities	0.00516	0.00375	0.0160	0.0145	-0.0450	-0.00881	-0.302**	0.279**
	[0.0671]	[0.0510]	[0.0108]	[0.0611]	[0.0468]	[0.0216]	[0.125]	[0.116]
Plantation colonial activities	0.162*	-0.0796	-0.0139	-0.124*	-0.0504	-0.0631	-0.330*	0.335
	[0.0903]	[0.0597]	[0.0200]	[0.0643]	[0.0302]	[0.0449]	[0.159]	[0.224]
Controls included:								
Share of Amerindian pop.	YES	YES	YES	YES	YES	YES	YES	YES
Log population density	YES	YES	YES	YES	YES	YES	YES	YES
Geography, location and natural resources	YES	YES	YES	YES	YES	YES	YES	YES
Country fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	280	279	280	280	280	280	279	270
Adjusted R-squared	0.774	0.822	0.916	0.817	0.476	0.734	0.612	0.851

Notes: Cluster standard errors at country level are in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 6
Controlling for pre-colonial characteristics: population density

Dependent variable:	Infant Mortality (1)	Years of Schooling (2)	Primary education (3)	Secondary education (4)	Drinking water (5)	Electricity (6)	Log GDP per capita (7)	Poverty rates (8)
Pre-Colonial Institutions	-0.0634***	0.0402**	0.0167**	0.0664***	0.0473**	0.0558**	0.200**	-0.122**
	[0.0201]	[0.0140]	[0.00601]	[0.0175]	[0.0198]	[0.0228]	[0.0764]	[0.0464]
Pre-colonial population density	-0.00609	-0.00593	-0.00439	-0.0158	0.0114	0.00390	-0.0697*	0.0352
	[0.0223]	[0.00769]	[0.00432]	[0.0111]	[0.0102]	[0.00792]	[0.0351]	[0.0249]
Other colonial activities	0.0493	-0.0151	-0.00331	-0.0221	-0.0591	-0.0273	-0.0804	0.117
	[0.0572]	[0.0255]	[0.00832]	[0.0313]	[0.0410]	[0.0195]	[0.101]	[0.114]
Mining colonial activities	0.00783	0.00633	0.0180	0.0215	-0.0500	-0.0105	-0.272**	0.263**
	[0.0690]	[0.0504]	[0.0109]	[0.0578]	[0.0513]	[0.0212]	[0.120]	[0.113]
Plantation colonial activities	0.169	-0.0726	-0.00865	-0.105	-0.0640	-0.0678	-0.247	0.293
	[0.0982]	[0.0610]	[0.0209]	[0.0635]	[0.0386]	[0.0480]	[0.151]	[0.222]
Controls included:								
Share of ethnic population	YES	YES	YES	YES	YES	YES	YES	YES
Log population density	YES	YES	YES	YES	YES	YES	YES	YES
Geography, location and natural resources	YES	YES	YES	YES	YES	YES	YES	YES
Country fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	280	279	280	280	280	280	279	270
Adjusted R-squared	0.773	0.822	0.916	0.818	0.476	0.733	0.626	0.853

Notes: Cluster standard errors at country level are in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 7
Controlling for pre-colonial characteristics: socioeconomic factors

Dependent variable: Percent of the population having completed Secondary education (in logs)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Pre-Colonial Institutions	0.0673***	0.0918***	0.0723***	0.0657**	0.0674***	0.0685***	0.0682***
	[0.0177]	[0.0195]	[0.0187]	[0.0267]	[0.0177]	[0.0188]	[0.0197]
Pre-colonial population density	-0.0232	-0.0183	-0.0241	-0.0232	-0.0231	-0.0241	-0.0231
	[0.0138]	[0.0125]	[0.0140]	[0.0138]	[0.0137]	[0.0137]	[0.0137]
Other colonial activities	-0.0306	-0.0332	-0.0337	-0.0301	-0.0313	-0.0325	-0.0313
	[0.0354]	[0.0361]	[0.0361]	[0.0362]	[0.0387]	[0.0349]	[0.0364]
Mining colonial activities	0.0144	0.00658	0.00896	0.0155	0.0134	0.0125	0.0140
	[0.0575]	[0.0531]	[0.0559]	[0.0530]	[0.0644]	[0.0555]	[0.0583]
Plantation colonial activities	-0.107	-0.116*	-0.104	-0.106	-0.107	-0.110	-0.108
	[0.0702]	[0.0633]	[0.0749]	[0.0712]	[0.0723]	[0.0695]	[0.0717]
Population employed in:							
Gathering		0.783					
		[0.448]					
Hunting		-0.160					
		[0.365]					
Fishing		1.320**					
		[0.473]					
Agriculture		-0.208					
		[0.158]					
Settlement pattern			-0.0247				
			[0.0220]				
Class Stratification				0.00277			
				[0.0307]			
Slavery					-0.00543		
					[0.0563]		
Elections						-0.0818	
						[0.0908]	
Property rights							-0.0211
							[0.0600]
Controls included:							
Share of ethnic population	YES	YES	YES	YES	YES	YES	YES
Log population density	YES	YES	YES	YES	YES	YES	YES
Geography, location and natural resources	YES	YES	YES	YES	YES	YES	YES
Country fixed effects	YES	YES	YES	YES	YES	YES	YES
Observations	261	261	261	261	261	261	261
Adjusted R-squared	0.812	0.823	0.814	0.811	0.811	0.812	0.811

Notes: Cluster standard errors at country level are in brackets. *** p<0.01, ** p<0.05, * p<0.1
Uruguay has been removed from the sample for this table.

Table 8
Contrasting rural and urban areas

Dependent variable:	Primary education		Secondary education		Drinking water		Electricity	
	rural	urban	rural	urban	rural	urban	rural	urban
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pre-Colonial Institutions	0.0239**	0.00914**	0.0865***	0.0325**	0.0860**	0.0415	0.126***	0.00668
	[0.0100]	[0.00394]	[0.0226]	[0.0131]	[0.0379]	[0.0272]	[0.0382]	[0.00583]
Controls included:								
Share of Amerindian pop.	YES	YES	YES	YES	YES	YES	YES	YES
Log population density	YES	YES	YES	YES	YES	YES	YES	YES
Geography, location and natural resources	YES	YES	YES	YES	YES	YES	YES	YES
Country fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	291	290	291	290	291	290	291	290
Adjusted R-squared	0.886	0.934	0.828	0.886	0.615	0.229	0.694	0.618

Notes: Cluster standard errors at country level are in brackets. *** p<0.01, ** p<0.05, * p<0.1
Argentina has been removed from the sample for this table.

Table 9
Comparing Amerindians with non-Amerindians

Dependent variable:	Primary education		Secondary education	
	Amerindians	non-Amerindians	Amerindians	non-Amerindians
	(1)	(2)	(3)	(4)
Pre-Colonial Institutions	0.0233	0.0148*	0.0844**	0.0720***
	[0.0197]	[0.00794]	[0.0344]	[0.0222]
Controls included:				
Share of Amerindian pop.	YES	YES	YES	YES
Log population density	YES	YES	YES	YES
Geography, location and natural resources	YES	YES	YES	YES
Country fixed effects	YES	YES	YES	YES
Observations	252	253	252	253
Adjusted R-squared	0.814	0.896	0.697	0.853

Notes: Cluster standard errors at country level are in brackets. *** p<0.01, ** p<0.05, * p<0.1
Brazil is not in the sample for this table.

APPENDIX

Table A1: Matching of ethnic groups

Panel A: Ethnic groups with similar or same names in the Ethnographic Atlas and in the Census

Alacalufe, Apinayé, Aymara, Bacairi, Bororo, Botocudo, Bribri, Cakchique, Camayura, Caraja, Chacobo, Chamacoco, Chichimec, Chinantec, Chiriguan, Chorote, Chorti, Cocama, Cubeo, Curipaco, Guató, Huichol, Kuikuru, Lenca, Macusi, Mam, Mapuche, Mataco, Maya, Mazateco, Miskito, Mixe, Mixteco, Mundurucu, Nambicuar, Ona, Palikur, Papago, Paressi, Piaroa, Pima, Popoluca, Qhiche, Sanema, Shavante, Sherente, Seri, Siriono, Tarahumar Tapirapé, Tehuelche, Tenetehar, Toba, Totonac, Trumai, Tupinamba, Tzeltal, Umotina, Waiwai, Wuitoto, Yagua, Yanomamo, Yaqui, Zapotec, Zoque

Panel B: Ethnic groups with different names in the Ethnographic Atlas and in the Census

Name in Atlas	Name in Census	Source of matching
Aweikoma	Kaingang	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group mainly from Brazil. Amongst other ethnonyms Aweikoma is also known as Coroado, Cayapa, etc.
Aztec	Nahualt	Encyclopedia Britannica. Available at http://www.britannica.com/ Amerindian group from Central Mexico. Nahualt is the language spoken and disseminated by the Aztec Empire.
Caduveo	Kadiwéu, Guaikurú	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group from Paraguay and Brazil. Amongst other ethnonyms they are also known as Caduvi, Kaiwa, etc.
Campa	Ashaninka	Encyclopedia of World Cultures, Vol. VII –South America- and Ethnologue: *Language of the World* https://www.ethnologue.com/ Ashaninka belongs to one of the seven main groups of Amerindian group known as Campa. Ashaninkas tend to be found in remote areas in Peru and Brazil.
Cayapa	Chachi	Encyclopedia Britannica. Available at http://www.britannica.com/ Amerindian group from the west coasts of Ecuador that traditionally was known as Cayapas. Nowadays, they called themselves as Chachis.
Cayua	Guarani, Guarani Kaiowá, Guarani Mbya, Guarani Nhandeva, Ava-Guarani, Tupi-Guarani, Pai-Tavytera, Guarani Occident.	Métraux (1948). The Guaraní. In Steward, Julian H. (ed.), Handbook of South American Indians, Vol. 3; and Ethnologue: *Language of the World* https://www.ethnologue.com/ Guarani encompasses a wider range of alternative names. Guarani can be found across Argentina, Bolivia, Paraguay and Brazil; and in each country the name of this group is used in different ways. For example, Bolivia classifies all type of Guarani-related groups under a single name known simply as Guarani – at least as it is reported in the recent census of 2012-. In Brazil, instead, it can be found three types of Guarani groups – Kaiowá, Mbya and Nhandeva-. However, all these names came to be labelled in the 18th century. Even the very single name of Guarani was introduced by the colonisers to differentiate those Amerindians that assimilated quickly the colonial power from those who rejected it, the latters in which case were identified by their pre-colonial name known as Cayua. A problem arises when in modern times those Amerindians who live in rural areas, and therefore have had hardly any external influence in their local matters, like to be called also as Guaranis. To overcome such a confusing association it seems that within the anthropological literature Guaranis are identified as Cayuas. And it is precisely this way that Murdock (1967) labelled them. We then proceed to aggregate all type of Guarani-related groups into a single classification: Cayuas.

CONTINUED…

Panel B: Ethnic groups with different names in the Ethnographic Atlas and in the Census

Name in Atlas	Name in Census	Source of matching
Chibcha	Muisca	Encyclopedia Britannica. Available at http://www.britannica.com/ Amerindian group mainly from Colombia. It is regarded as the second most influential group outside the Inca Empire in South America.
Choco	Embera	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group that lives in Colombia, Ecuador and Panama. Amongst other ethnonyms they are also known as Cholo, Meme, Catio, etc.
Cocopa	Cucapá	Ethnologue: *Language of the World* https://www.ethnologue.com/ Amerindian group from the northern Mexico who is also known as Kikima, Cocopah, etc.
Cuna	Tule	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group that predominately lives in Panama but a few of them can also be found in Colombia. The name Cuna is more related to their cultural origins than the actual name they like to be used -Tule.
Goajiro	Wayuu	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group from Colombia and Venezuela. The name Goajiro is meant to have been introduced by Spanish colonisers.
Guahibo	Sikuani	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group from Colombia and Venezuela. The name Guahibo is from pre-colonial origins whereas Sikuani is a term labelled by them in recent times to encompass the whole groups under Guahibo heritage.
Inca	Quechua	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group that lives in Argentina, Chile, Bolivia, Ecuador and Peru. As its northern counterpart, Quechuas is the language spoken and disseminated by the Inca Empire. All these countries with Inca heritage identify this group as Quechuas.
Jivaro	Achuar, Shuar, Shiwiar	Encyclopedia of World Cultures, Vol. VII –South America; and Ethnologue: *Language of the World* https://www.ethnologue.com/ Amerindian group mainly from Ecuador. They are also known as Givari, Zibaro, etc.
Lengua	Enlhet, Enxet	Encyclopedia of World Cultures, Vol. VII –South America; and Ethnologue: *Language of the World* https://www.ethnologue.com/ One of the dozens hunter-gatherer bands of the Gran Chaco in Paraguay. The other ethnonyms used are Lengua-Sur and Lengua-Maskoi.
Paez	Nasa	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group that lives in Colombia. Paez is the language spoken by this group and instead they called themselves as Nasa.
Paraujano	Añu	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group from Venezuela. Nowadays they called themselves as Añu.
Piapoco	Tzase	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group from Colombia and Venezuela. Amongst other ethnonyms they are also known as Yapaco, Cuipoco, Deja, etc.

CONTINUED…

Panel B: Ethnic groups with different names between Atlas and Census

Name in Atlas	Name in Census	Source of matching
Tarasco	Purépecha	Encyclopedia of World Cultures, Vol. VIII –Middle America and the Caribbean. Amerindian group from western Central Mexico. The term Purépecha comes from pre-colonial times whereas Tarasco was introduced during contact with Spanish colonisers.
Taulipang	Pemon	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group from Venezuela and Brazil. Amongst other ethnonyms they are also known as Arekuna or Taurepan.
Tucuna	Tikuna,	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group that live nearby the Amazon in Brazil, Colombia and Peru. Tucuna is a term with foreign origin. There are other ethnonyms to identify this group such as Jaunas, Tocunas, etc.
Tunebo	U`wa	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group mainly from Colombia. Amongst other ethnonyms they are also known as Covari, Luna, Tame, etc.
Warrua	Warao	Ethnologue: *Language of the World* https://www.ethnologue.com/; and Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group mainly from Venezuela. Warao is a self-name meaning "lowland people". Amongst other ethnonyms they are also known as Guarauno or Tiuitiuia.
Yahgan	Yámana	Ethnologue: *Language of the World* https://www.ethnologue.com/ Amerindian group that lives in Chile and Argentina. This group is also identified as Tequenica.
Yaruro	Pumé	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group from Venezuela. Pumé is a term commonly used amongst people within this group. On the other hand, Yaruro is meant to be a term mostly used by non-Amerindians in Venezuela. They are also known as Capuruchano and Saruro.
Yupa	Yukpa	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group from Venezuela. The difference between Yupa and Yukpa is only based on dialectic. This group is also identified as Yuko.

Table A2: Definitions and sources of variables (baseline analysis)

Variable	Description
Dependent variables:	
Infant Mortality Rates	The number of deaths of children under 1 year old in a given year per 1000 live births in each state. *Source: Country's national statistics office*
Drinking Water	Proportion of households in each state that have access to drinking water. *Source: Redatam- Commission for Latin America and the Caribbean (ECLAC) and country's national statistics office*
Electricity	Proportion of households in each state that have electricity. *Source: Redatam- Commission for Latin America and the Caribbean (ECLAC) and country's national statistics office*
Average Years of Schooling	Average of years of schooling from primary level onwards of population aged 15 and above in each state. Most recent value available in period between 1990 and 2006. *Source: Gennaioli et al. (2013)*
Primary Education	Proportion of total population in each state that completed primary education. *Source: Redatam- Commission for Latin America and the Caribbean (ECLAC) and country's national statistics office*
Secondary Education	Proportion of total population in each state that completed secondary education. *Source: Redatam- Commission for Latin America and the Caribbean (ECLAC) and country's national statistics office*
Log GDP per capita	Annual Log Gross Domestic Product per capita in each state as calculated by *Bruhn and Gallego (2012)* & ***Acemoglu, D.; Gallego, F. Robinson, J. A (2014)***
Poverty rates	Annual Log Poverty rates as calculated by *Bruhn and Gallegos (2012).*
Main regressor of interest :	
Index of Pre-colonial Institutions	As described in text. *Source:* Gray (1999) , A Corrected *Ethnographic Atlas*; Redatam- Commission for Latin America and the Caribbean (ECLAC) and country's national statistics offices
Other independent variables:	
Share of Amerindian Population	Proportion of Amerindian groups in total population in each state. Source: Redatam- Economic Commission for Latin America and the Caribbean (ECLAC) and country's national statistics offices
Population Density	Total population of each state divided by its total state's surface area (sq. km). Source: Redatam- Commission for Latin America and the Caribbean (ECLAC) and country's national statistics offices
Latitude	Absolute latitude of the centroid of each state. *Source: Own calculation using Geographical Information System (GIS)*
Altitude	Average altitude of each state (km). Source: *Bruhn and Gallegos (2012) and Global Gazetteer Version 2.1*
Temperature	Average temperature in degree Celsius in each state during period 1950-2000. *Source: Gennaioli et al. (2013)*
Land Suitability Index	Average land quality for agriculture in each state. Variable takes values between 0 and 1, with higher values denoting more fertile land. Data drawn from Ramankutty et al. (2002), who developed a geospatial index of land suitability for agriculture by examining three major components: croplands, climate conditions and soil characteristics. The index represents the probability that a particular grid cell will be cultivated. Ramankutty et al. (2002) report such index at a 0.5 degree resolution. We compute this index by spatially averaging all the grid cells that fall within each state. Data computed with ArcGIS 10.1.
Malaria Stability Index	Average values of malaria index. Variable takes values from 0 to 5, with higher values denoting more prevalence of the various types of malaria. Data drawn from Kiszewski et al (2004). To calculate this index we averaged all the grid cells that fall within each state. Data computed with ArcGIS 10.1.

CONTINUED…

Variable	Description
Malaria Stability Index	Average values of malaria index. Variable takes values from 0 to 5, with higher values denoting more prevalence of the various types of malaria. Data drawn from Kiszewski et al (2004). To calculate this index we averaged all the grid cells that fall within each state. Data computed with ArcGIS 10.1.
Ruggedness Index	This index measures small-scale terrain irregularities by taking the differences of elevation of a grid cell with respect to the grid cells that surrounds it. The index is computed by averaging all the grid cells that fall within each state. Lower values denote areas at nearly level terrain whereas higher values represent highly rugged areas. Data computed with ArcGIS 10.1. Source: Nunn and Puga (2010) available at http://diegopuga.org/data/rugged/
Land Area	Total surface area of each state in sq. km. *Source: Country's national statistics offices*
Land Locked	A dummy variable that indicates whether states have access to the sea. *Source: Own calculation*
Distance to Capital	Distance between the centroid of each state and the capital city of the country. *Source: Own calculation*
Inverse Distance to Coast	Distance between the centroid of each state and the nearest coastline in thousands of kilometres. *Source: Gennaioli et al. (2013)*
Oil & Gas	Dummy variable that indicates the existence of oil or gas fields in each state. *Source: U.S. Geological Survey*
Gold & Silver	Dummy variable that indicates the existence of gold or silver mines in each state. *Source: U.S. Geological Survey*
Other Mines	Dummy variable that indicates the existence of any other mines (cooper, zinc, coal, etc.) in each state. *Source: U.S. Geological Survey*
Gathering	Percentage of the population dedicated to the collection of wild plants and small land fauna. *Source: Murdock (1967), Ethnographic Atlas, v1.*
Hunting	Percentage of population dedicated to hunting. *Source: Murdock (1967), Ethnographic Atlas, v2.*
Fishing	Percentage of population dedicated to fishing. *Source: Murdock (1967), Ethnographic Atlas, v3.*
Agriculture	Percentage of population dedicated to agriculture. *Source: Murdock (1967), Ethnographic Atlas, v5.*
Settlement Pattern	An ordered variable that takes a value between 1-8, where 1 indicates fully nomadic and migratory societies; 2 is assigned to societies with semi-nomadic characteristics; 3 identifies societies with semi-sedentary characteristics; 4 indicates societies that lived in compact an impermanent settlements; 5 is given to societies those in neighbourhoods of dispersed family homes; 6 corresponds to groups in separated hamlets forming a single community; 7 is given to societies living in compact and relatively permanent settlements; and 8 corresponds to groups residing in complex settlements. *Source: Murdock (1967), Ethnographic Atlas, v30.*
Class Stratification	An ordered variable that takes a value between 1-5, where 1 is given to "*absence of significant class distinction among freemen, ignoring variations in individual repute achieved through skill, valor, piety, or wisdom*"; 2 corresponds to "*wealth distinctions, based on the possession or distribution of property, present and socially important but not crystallized into distinct and hereditary social classes*"; 3 is assigned to "*elite stratification, in which an elite class derives its superior status from, and perpetuates it through, control over scarce resources, particularly land, and is thereby differentiated from a propertyless proletariat or serf class*"; 4 indicates "*dual stratification into a hereditary aristocracy and a lower class of ordinary commoners of freemen, where traditionally ascribed noble status is at least as decisive as control over scarce resources*"; and 5 is given to "*complex stratification into social classes correlated in large measure with extensive differentiation of occupational statuses*". *Source: Murdock (1967), Ethnographic Atlas, v67.*
Slavery	A binary index that takes a value of 1 for societies characterized by any type of slavery and zero otherwise. *Source: Murdock (1967), Ethnographic Atlas*, v70.
Elections	A binary index that takes a value of 1 for societies where succession was conducted through "*election or other formal consensus*", and zero otherwise. *Source: Murdock (1967), Ethnographic Atlas, v72.*
Inheritance Rules for Property	A binary index that takes a value of 1 for societies reporting any type of inheritance rule of real property (land), and zero otherwise. *Source: Murdock (1967), Ethnographic Atlas, v74.*

Table A3

Matrix of correlations

	Pre-Colonial Institutions	Share of Amerindian pop.	Infant Mortality Rates	Years of Education	Primary Education	Secondary Education	Drinking Water	Electricity	GDP per capita	Poverty Rates
Pre-Colonial Institutions	1									
Share of Amerindian pop.	.35	1								
Infant Mortality Rates	.02	.36	1							
Years of Education	.05	-.34	-0.75	1						
Primary Education	.15	-.33	-0.47	0.66	1					
Secondary Education	.27	-.17	-0.57	0.65	0.64	1				
Drinking Water	-.19	-.53	-0.41	0.41	0.33	0.24	1			
Electricity	-.16	-.55	-0.59	0.60	0.59	0.59	0.64	1		
GDP per capita	-.10	-.33	-0.54	0.51	0.44	0.56	0.40	0.59	1	
Poverty Rates	.19	.39	0.21	-0.05	-0.23	-0.12	-0.44	-0.39	-0.52	1

Table A4

Baseline regressions without controlling for the share of Amerindians in the population

Dependent variable	Infant Mortality (1)	Years of Schooling (2)	Primary education (3)	Secondary education (4)	Drinking water (5)	Electricity (6)	Log GDP per capita (7)	Poverty rates (8)
Pre-Colonial Institutions	-0.0282 [0.0215]	0.0188 [0.0170]	0.00490 [0.00506]	0.0427*** [0.0125]	0.0177 [0.0150]	0.0144 [0.0105]	0.143 [0.0829]	-0.0620* [0.0302]
Controls included:								
Share of Amerindian pop.	NO	NO	NO	NO	NO	NO	NO	NO
Log population density	YES	YES	YES	YES	YES	YES	YES	YES
Controls for geography, location and natural resources	YES	YES	YES	YES	YES	YES	YES	YES
Country fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	317	316	317	317	317	317	297	270
Adjusted R-squared	0.766	0.786	0.841	0.761	0.390	0.467	0.617	0.834

Notes: Cluster standard errors at country level are in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table A5

Regressions where controls are added separately to the baseline regression

Dependent variable	Infant Mortality (1)	Years of Schooling (2)	Primary education (3)	Secondary education (4)	Drinking water (5)	Electricity (6)	Log GDP per capita (7)	Poverty rates (8)
Panel A: Baseline + Colonial activities								
Pre-Colonial Institutions	-0.0643***	0.0394**	0.0162**	0.0642***	0.0489**	0.0563**	0.190**	-0.117**
	[0.0209]	[0.0140]	[0.00608]	[0.0170]	[0.0212]	[0.0228]	[0.0802]	[0.0436]
Panel B: Baseline + pre-colonial population density								
Pre-Colonial Institutions	-0.0624**	0.0383**	0.0156**	0.0646***	0.0469**	0.0536**	0.199**	-0.122**
	[0.0215]	[0.0150]	[0.00609]	[0.0181]	[0.0201]	[0.0218]	[0.0792]	[0.0415]
Panel C: Baseline + pre-colonial economic activities								
Pre-Colonial Institutions	-0.0710***	0.0757***	0.0308**	0.100***	0.0551**	0.0847***	0.214**	-0.147***
	[0.0212]	[0.0215]	[0.0113]	[0.0222]	[0.0256]	[0.0277]	[0.0856]	[0.0475]
Panel D: Baseline + pre-colonial Settlement pattern								
Pre-Colonial Institutions	-0.0677***	0.0556**	0.0256**	0.0831***	0.0448**	0.0847**	0.195**	-0.128***
	[0.0220]	[0.0196]	[0.0103]	[0.0204]	[0.0172]	[0.0307]	[0.0839]	[0.0358]
Panel E: Baseline + pre-colonial Class Stratification								
Pre-Colonial Institutions	-0.0974**	0.0782**	0.0354**	0.0808**	0.0847***	0.0944**	0.178*	-0.119**
	[0.0397]	[0.0324]	[0.0135]	[0.0305]	[0.0266]	[0.0331]	[0.0849]	[0.0508]
Panel F: Baseline + pre-colonial slavery								
Pre-Colonial Institutions	-0.0630***	0.0505**	0.0243**	0.0779***	0.0558**	0.0798***	0.188**	-0.115***
	[0.0210]	[0.0174]	[0.00961]	[0.0200]	[0.0198]	[0.0248]	[0.0840]	[0.0379]
Panel G: Baseline + pre-colonial elections								
Pre-Colonial Institutions	-0.0682***	0.0521**	0.0247**	0.0780***	0.0548***	0.0850**	0.191**	-0.114***
	[0.0208]	[0.0188]	[0.0102]	[0.0202]	[0.0176]	[0.0307]	[0.0815]	[0.0371]
Panel H: Baseline + pre-colonial property rights								
Pre-Colonial Institutions	-0.0620***	0.0514**	0.0259**	0.0787***	0.0581**	0.0784***	0.193**	-0.110**
	[0.0203]	[0.0198]	[0.0109]	[0.0220]	[0.0201]	[0.0249]	[0.0837]	[0.0371]

Note: All regressions include baseline controls (share of Amerindians, current population density, geography, location and natural resources) plus country fixed effects.

Table A6

Regressions for Mexican states, early and mid-20th century

Dependent variable:	Literacy rates in 1921		Literacy rates in 1930		Literacy rates in 1950	
	(1)	(2)	(3)	(4)	(5)	(6)
Pre-colonial institutions	0.161	0.155	0.0515	-0.0300	0.0456	0.0648
	[0.0975]	[0.111]	[0.0345]	[0.0373]	[0.0644]	[0.0913]
Share of Amerindians	-1.184	-2.021*	-0.0358	-0.488	-1.237*	-1.494
	[0.984]	[1.042]	[0.348]	[0.350]	[0.673]	[0.993]
Population density	-0.0185	-0.0175	0.0217	0.0306	0.0382	0.0256
	[0.0757]	[0.0619]	[0.0268]	[0.0208]	[0.0319]	[0.0414]
Controls for geography	yes	yes	yes	yes	yes	yes
Controls for location and natural resources	no	yes	no	yes	no	yes
Observations	28	28	28	28	32	32
R^2	0.409	0.669	0.118	0.554	0.621	0.501

Notes: pre-colonial institutions, share of Amerindians and population density are calculated using data from 1930 for the regressions explaining literacy rates in 1921 and 1930, and using data from 1950 for the regressions explaining literacy rates in 1950. Controls for geography are latitude, altitude, temperature, ruggedness, and malaria index. Controls for location and natural resources are land area, landlocked dummy, distance to capital, distance to coast, land suitability index, oil & gas dummy, gold & silver dummy, and other mines dummy.
Standard errors in brackets.