EL COLEGIO DE MÉXICO
CENTRO DE ESTUDIOS ECONÓMICOS

MAESTRÍA EN ECONOMÍA

TRABAJO DE INVESTIGACIÓN PARA OBTENER EL GRADO DE MAESTRO EN ECONOMÍA

THE COLONIAL LEGACY OF EXTRACTIVE INSTITUTIONS: EVIDENCE FROM MEXICO AND THE XVIII CENTURY

VÍCTOR HUGO DELGADO BARRERA

PROMOCIÓN 2015-2017

ASESOR:

SINAIA URRUSTI-FRENK

JULIO 2017

In memory of my mother Martha Barrera Villaseñor

Acknowledgement

I would like to thank my Advisor Sinaia Urrusti Frenk for her advice, mentoring and excellence in class. All the errors are my own.

I also want to thank my professors, especially Julen Berasaluce and Gabriela Calderón. My colleagues, Elizabeth de la Paz, Diana Jordán and Miriam González for their comments, and support. Students of econometrics lab, in particular to Isaac Lara and Xanath Rivera. Finally, I want to thank my family.

The Colonial Legacy of Extractive Institutions: Evidence from Mexico and the XVIII Century

Abstract

In this study, I exploit variation in extraction experience from different regions in Mexico, registered by the *Real Hacienda*, and population records of the XVIII century, to identify whether current economic performance can be traced back to the Colonial period. Using a unique database and instrumental variables as identification strategy, I found two opposite effects; municipalities with a high presence of an indigenous population in the XVIII century are less developed today, while direct extraction or tax collection by *Real Hacienda* has a positive effect on current economic performance. The results are robust for different samples and several measures of extraction and economic performance. As potential mechanisms, I found that collection has an adverse effect through public goods provision, which varies across regions and municipalities. When examining literacy as a proxy for access to education, places with a higher historical indigenous population present lower access to education, and extraction depicts a positive effect in the long-term. I found evidence of a colonial legacy that persists through the rules of redistribution, which varies across regions and municipalities.

Contents

I Introduction

Several studies have found evidence on how historical events can have persistent outcomes (Nunn, 2008a), and how historical institutions can partially explain differences across countries and within a country (Acemoglu et al., 2001; Dell, 2010). In this work, I consider that if we see differences across regions within a country, with a shared history and the predominance of extractive colonial institutions, it could be the case that extraction was not homogenous, and the differences across regions are due to past extraction differences. Specifically, I test if current economic performance in Mexico can be traced back to the colonial period. For such a task, I consider the way society formed in Mexico (Robinson, 2013), a measure of direct extraction based on tax data from the XVIII century and their effects in the long-term.

The principal challenge of this study is finding a reliable source of variability that allows us to identify any evidence of colonial legacy. Throughout an analysis of the historical background of the colonial period, such variability could come from extraction or tax collection itself and the social interaction that emerged between the indigenous population and Europeans. Historians have recently documented the activity of the *Real Hacienda*, an administrative organism in charge of collecting taxes across territories. Historians claim that understanding tax data can help us comprehend the economic environment in that time (Marichal, 2003). If this is the case, I consider that this information can also help us to understand the incentives underlying the system that lead to a legacy that can persist today. Regarding the indigenous population, records from the mid and late XVIII century, which if imprecise, are a good proxy for the structure rather the number of habitants. Therefore, the research question that naturally arises is: Does evidence exist of a colonial legacy within Mexico derived from its fiscal experience and population structure?

For the analysis, I build a novel database that combines historical and contemporary data. I consider the information of 23 *cajas reales* or royal treasuries that collected local taxes. I pay attention to three main taxes which are tributes, *diezmos* and *alcabalas*, focusing on the period from 1786 until 1810. The tribute is a tax on the indigenous population in exchange for protection and the new Catholic faith, the *diezmo* on silver mining and the *alcabalas* is a tax on trade (TePaske and Klein, 1986). The reason for focusing on these three taxes relies on their importance not only per

their magnitude regarding share but also to the extent that these taxes can reflect economic reality in the colonial period. Furthermore, I choose the 1786-1810 period for three reasons: 1) almost all 23 royal treasuries were functioning by this time; 2) I can assign to almost all municipalities today one extraction experienced; and 3) the data is likely to be more accurate. The population records that I use are mainly from the mid XVIII century, since I dispose both indigenous and total population at the jurisdiction level.

The first approach is estimating a direct relationship between economic performance, tributes, and native population with a linear regression model, I also consider additional covariates whose omission might bias the estimates such as geography, location, and size of the municipality today. One reason to start the analysis with tributes is that the tributary system was an institution itself inherited from pre-Hispanic societies. Evidence from the Ordinary Leas Squares (OLS) estimates suggests that two opposite effects exist. There is an adverse effect on current economic performance; places with abundant historical indigenous populations show a low per capita gross domestic product (GDP) today. In contrast, extraction shows a positive correlation between tributes collection and current per capita GDP.

Evidence from OLS estimates suggests a strong association between the past and the present; however, we cannot conclude causality. OLS estimates are not consistent due to the presence of endogeneity, which emerges due to two issues. The first comes from an alternative explanation that regions where the most economically underdeveloped municipalities are selected into a stronger extraction experience than others biasing our results toward zero. Analogously, indigenous populations could have been in initially less developed locations. The second source of endogeneity is measurement error since the number of native populations is likely to be underreported. Such a measurement error would bias our estimates toward zero. Therefore, the OLS estimates underestimate the effects.

I follow an instrumental variable strategy (IV) to identify a causal effect. I instrument tributes and the indigenous population normalized by the total population with distances. As instrumental variables for the tax collection, I propose the distance from Mexico to the royal treasuries plus the distance to municipalities' centroids (*SD*) and the distance from local royal treasuries to mu-

nicipalities' centroids interacting with local geography (WD). Concerning population, I use the ratio of these distances (RD) and a variable that proxies the distance to the Port of Veracruz. The idea behind these instruments is that the extraction activity depends on the distances from Mexico, royal treasuries' ubication and local geography. The historical evidence suggests that tax supply locations affected tax demand and not the other way. That is, if the *Real Hacienda* identified an important source of wealth in a particular place, then the *Real Hacienda* would build a royal treasury to improve extraction. In sum, it is this locating phenomenon that allows me to identify a causal effect of extractive colonial institutions on current economic performance.

The IV results are in line with OLS estimates, the more the indigenous people there were in a municipality, the less developed today; and the more the collection, the better the economic performance. One concern is that the instruments are weak for the population variable; therefore, limited information maximum likelihood estimators are reported, which are robust in the presence of weak instruments, and the effect prevails. Although the over-identification test fails to reject the null that the instruments are valid, strictly exogeneity might not be achieved due to a possible direct effect of the instruments; however, I show that if such bias exists, then the IV estimates would sub-estimate the effect. Furthermore, I performed robustness checks to determine the nature of the positive impact of extraction. Besides, the results are robust to different samples, current economic performance measures and tax collection. In sum, through the construction of an extraction index that considers *alcabalas* and tributes that highlight that collection matters, the positive effect is likely to come from some inequality that enhanced economic growth.

After analyzing the causal effect of the colonial period in Mexico, I use contemporary data and similar identification strategies to investigate two potential mechanisms of persistence that are relevant for the case of Latin American countries; these are paved roads as a proxy for public good provision and literacy as a proxy for access to education. When examining the distance to paved roads, I find that places with a higher collection have a lower public good provision and the indigenous population is not significant. The consistency of the OLS estimates allows me to observe different effects across municipalities. I observe that rural places in the south have lower access to paved roads. These disparities partially drive the economic difference that we see today.

About literacy as a proxy for access to education, municipalities with high levels of indigenous populations in the XVIII century have lower access to education today while collection has a positive effect in the long-term.

This work contributes in the literature in several ways. For the analysis, I use a unique database that compiles historical and contemporary data. My findings contribute to a substantial body of research that has studied the importance of history in economic development, through the persistence of extractive institutions. Besides, this is one of the first studies that attempts to trace back current differences across regions in Mexico to the colonial period, providing not only causal evidence but also potential channels through which the effect persisted. The results are in line with recent literature regarding Latin America, where we cannot reject the idea that initial inequality promoted growth. However, my results suggest that an adverse effect should not be discarded, since it seems to come from the institutions that shaped human interaction, not only in Mexico but also in Latin America.

The remainder of the work is organized as follows. Section II and III present a brief literature review and describe the historical context, respectively. Section IV discusses the database construction and descriptive statistics. Section V presents the baseline model and reports OLS estimates. Section VI explains the identification strategy and reports the causal results. Section VII provides robustness checks and addresses threats for the identification strategy. Section VIII discusses potential mechanisms, and Section IX concludes.

II Literature Review

In this section, I do a brief review of existing literature regarding the role of institutions as a mechanism through which history plays a relevant role in Mexico. The reason is that this study focuses on how extractive institutions established in the colonial period affect current performance within Mexico. Other mechanisms and contexts are in the literature showing that history is relevant to explain current development outcomes. Nunn (2009) and Banerjee and Duflo (2014) provide extensive surveys on different aspects, including geography (Nunn and Puga, 2012), social norms (Nunn and Wantchekon, 2011), small accidents (Dell, 2012), changes in formal rules (Beaman

et al., 2012), etc. However, the purpose of the literature review is contextualizing a problem that has recently arisen in the literature in the Mexican context.

In the recent years, extensive literature has emerged attempting to explain the roots of economic development and the mechanisms or channels through which they have a persistent impact today. History has played a role; the first studies examine the importance of colonial rule, an event that determined long-term economic development through the establishment of institutions that prevail today (Engerman and Sokoloff, 1997; Porta et al., 2008; Acemoglu et al., 2001). These studies rely on the importance of colonial institutions from different viewpoints. On the one hand, the identity of the colonizer matters, since it determined the legal systems of the time (Porta et al., 2008); and on the other hand, local characteristics such as endowments shaped former domestic institutions that ultimately affected development in the long-term (Engerman and Sokoloff, 1997; Acemoglu et al., 2001). For instance, Acemoglu et al. (2001) focused on the relationship between the disease environment and the first European settlements; this allows them to approximate settler mortality as a measure for former institutions because it incorporates the disease environment under which colonial institutions were born. Then, the authors follow an instrument variable strategy, using settler mortality as an instrument; finding a positive effect, the better the institutions, the higher the economic performance.

In the literature concerning the institutions as a mechanism through which history matters, we can identify two study types, depending on the level of observation for the study; these are macro (Acemoglu et al., 2001, 2002; Rodrick et al., 2004) or micro (Banerjee and Iyer, 2005; Dell, 2010) which complement one another since both have advantages and disadvantages. It is important to consider that the starting point is simple; some institutions are better in the long-term, and this applies to both countries and regions. For instance, in Acemoglu et al. (2001) seminal article, analyzing national data brings variability in other potential explanations that the authors discard as determinants of current performance, such as geography or culture, proving a good perspective of the general problem. However, measures on institutions vary between countries, and variables like culture might not capture a representative set of beliefs and values. The micro studies are more recent with a central objective to achieve internal validity by isolating one institution, giving

a better understanding of the mechanisms bearing the process through which extractive institutions persist (Banerjee and Iyer, 2005; Dell, 2010). About the Latin America case, Dell (2010) focuses on the long-term effect of *mita* in Peru. The *mita* was an institution that consisted of a forced labor system in the mining sector and was implemented in specific areas. The author, using regression discontinuity as identification strategy, finds that *mita* districts consume 32% less than *non-mita* districts today. Even though these studies share that an extractive colonial legacy leads to a worse economic outcome today, the channels of persistence are considerably different. Acemoglu et al. (2001) suggest that institutions persist and Dell (2010) finds that *mita* drove the allocation of public goods, land tenure and differences in human capital. The different channels do not contradict each other; however, when micro observations are considered, the channels of persistence can be clearer, which could help us extend the analysis to policy.

The case of Latin American countries is not far from Acemoglu et al. (2001) explanation. Previously, Engerman and Sokoloff (1997) hypothesized that one reason that could explain the underdevelopment in Latin America is that differences in initial endowments enhanced different institutions across these territories as well as unbalanced political power among individuals, resulting in a system that protects the privileges of the elite. There is opposing evidence regarding this hypothesis. For Latin American countries, the evidence proposes that inequality was beneficial for long-term economic development. Acemoglu et. al (2008) find that there is a positive relationship between land inequality and enrollment rates in Cundinamarca, Colombia. Moreover, Dell (2010) states that in *mita* districts in Peru, where people were induced into a forced labor system, a lower presence of large landowners in the form of haciendas, due to government restrictions, consequently resulted in a lower provision of public goods and low current performance. Additionally, Nunn (2008b) tested the hypothesis of Engerman and Sokoloff (1997) in the U.S.; the author finds a positive relationship between slavery and inequality, but the latter does not seem to have a robust correlation with current performance.

Explaining the differences in Latin America Acemoglu and Dell (2010) document that the determinants that account for the difference across countries could also explain the differences across municipalities. The authors find that disparities in human capital can explain variation in

economic outcomes; nevertheless, a component can be attributed to local institutions under the argument that *de jure* and *de facto* institutions vary within a country. Regarding the Mexican case Robinson (2013), in his conference to the Mexican Senate titled "Why regions fail?, exposes that differences observed at the national level, due to variation in institutions also apply to areas within a country. The author references the formation of inclusive and extractive institutions, where the inclusive creates a proper environment for creativity and entrepreneurship while the extractive does not. Then, Robinson (2013) provides an explanation for the difference between regions in Mexico as the South having more extractive economic and political institutions than the rest of the country, and these have interacted with national extractive institutions in unfortunate ways.

In sum, the literature presents a framework for studying the Mexican case, and it suggests considering the following facts. First, history plays a significant role in explaining differences through institutions that were extractive in Mexico. Second, the persistence correlates with mechanisms ultimately conveyed in economic outcomes; the literature proposes public goods. Third, if the factors that depict differences across countries apply to areas within a country, then we could explain differences within Mexico, due to the existence of more extractive institutions. Combining all these facts, I analyze that history matters across municipalities, through differences in extraction experienced during the colonial period, which endures under potential mechanisms, leading to a variety of economic outcomes today.

III Historical Background

The sudden change after the fall of the Aztecs was the complete elimination of the authorities of this empire. To keep some control over the new domains, Cortés redistributed the *señoríos*, indigenous states with a degree of autonomy, to his soldiers, through an institution called *encomiendas*, as a reward or "plunder" for the conquest (Miranda, 2005; García, 2008; Hauseberger, 2010). The *encomiendas* consisted on the right of the conquerors, namely the *encomenderos*, over the indigenous population in a particular *señorío*. The *encomenderos* had the right to charge a tribute in the form of goods and services, without the right over the land, and had the responsibility of keeping the military control over these territories (García, 2008). Besides this, the *encomienda* was

not hereditary and in exchange the *encomendero* offered protection and the Catholic faith (Yeager, 1995; Batchelder and Sanchez, 2013). This institution lent itself to abuses from the *encomenderos*; such issue was followed by attempts of the Crown to defend the native population (Gibson, 1967; Zavala, 1973). In this setting, the tribute would be the first antecedent of the fiscal system and the *encomienda* the first formal institution that dictated the relationship between the local population and the Spaniards; the institutions around extraction and the indigenous population would develop during three centuries (Marichal and Marino, 2001).

While the central government installed the Kingdom of New Spain, the population from Europe began relocating across the domains along with the conquerors; then the incoming population legitimized such settlements with the enacting of *cabildos* and trade networks (García, 2008). Nonetheless, among this new population came Catholic orders whose task was to spread the Catholic faith among the native population (García, 2008). Moreover, conflicts arose between the conquerors and the population because they looked for the best position in the government; however, the conquerors were displaced from critical positions and replaced by a civil government; then, the Crown achieved representation by the figure of Viceroy (García, 2008).

The integration of the indigenous population implied the accommodation of the *señoríos* which brought several changes like the formation of *cabildos* redefining the *señoríos* as *pueblos de indios* with a well-organized local government, assigning the title of *gobernador* to the *caciques* (García, 2008). In each *pueblo de indios*, there was an obligation to concentrate the population in urban centers characterized by a *central plaza*, church, and roads. Many localities were born with these characteristics, where the main locality was named *cabecera* (García, 2008).

As the Spanish Empire was expanding its territories, there was a need to ensure fiscal collection control over the new territories. The command came from a very well structured administration. One of the most significant elements was the *Real Hacienda* that functioned through a network of *cajas reales*, or royal treasuries, with the purpose of guaranteeing a fair amount was entitled to the Crown (TePaske and Klein, 1986; Jauregui, 1998). For each *caja*, there were royal officials in charge, standing out the accountant and the treasurer. The activities of the accountant and the treasurer were subject to norms and rules and had no distinction of power. The accountant was in

command of the administration while the treasurer commanded the operation. The Spanish Crown also established rules to avoid fraud from the royal administrators. Additionally, to promote an honest administration, Felipe III added a new organism, the *Tribunales de Cuentas*, where the *Contadores Mayores* watched the functioning of the royal treasure within the administration; this organism also analyzed the tax records before sending them to Spain (Jauregui, 1998; Rivarola, 2005).

In Mexico, the royal treasuries were born as the Spanish Empire expanded its territories; the foundation of a royal treasury depended on four criteria: 1) the presence of mines, 2) density of indigenous population, 3) ports and 4) military control. The first *caja* was created in Mexico (1521) right after the fall of *Tenochtitlan*, and the second one in Veracruz (1531) due to the importance of trading with the colonies and Spain. Other *cajas* such as *Zacatecas* and *Guanajuato* were established right after discovering valuable deposits of silver and gold. In the second half of the XVIII century followed by the Bourbon Reforms, *cajas* were located in strategic regions to ensure an efficient collection and to strengthen the royal authority (TePaske and Klein, 1986).

In the XVII century, New Spain lived a period of maturity reached until the second half of the XVII century. However, major changes occurred in XVIII century when the Crown passed to the Bourbons. After the second half of the XVIII century, the reforms canalized attention to the way of administrating the domains. These changes received the name of the Bourbon Reforms. In 1765, José Gálvez, under the command of Carlos III, visited all the tribunals and royal treasuries in New Spain, to manage and control the taxation of the *Real Hacienda*. Following Marino (2001), Gálvez paid attention to increase revenue through tributes; he advised promoting a higher control over the indigenous population and minor authorities, the homogenization of the tributes, and an increase of this to other *castas*. The administration triggered two royal decrees, one in 1786 and the other in 1794, which aimed to enhance the effectiveness of the collection system (Marino, 2001). By the end of the XVIII century, tribute payments were an obligation for the indigenous people. In theory, the heads of households, between 18 and 50 years old had to pay one peso and a half bush of corn (Marino, 2001; Marichal, 2003; Molina, 2004). There were also exceptions toward women, *caciques*, governors, and mayors (Marino, 2001), in the case of adverse environments, especially

epidemics. For an adequate collection, the administration set wages for the officials to avoid abuses and strict rules to develop an accurate census on the homemade population (Marino, 2001; Molina, 2004).

Another proposal from José Gálvez was the establishment of *Intendentes* whose job was to carry out a task of the Viceroy in a defined region or *Intendencia*. The *Ordenanza de Intendencias* was approved in 1786 and arrived in the Americas in 1787 (Commons, 1993). The duties were principally the collection of taxes, justice delivery, military control and the administration of the cities and localities within the corresponding jurisdiction (Commons, 1993; Jauregui, 1998). In sum, it is this attempt from the Crown to achieve a more efficient extraction through the *Intendencia* system that allows assigning a proxy for an extractive experience to each municipality today. The latter is because a principal royal treasury was assigned to one *Intendencia*, which had the responsibility of administrating a specific set of jurisdictions that can be matched to municipalities today.

Turning attention toward the understanding of the historical context of the tax data, historians have documented the functioning of the *Real Hacienda*, and how its study can help in understanding the economy in the colonial period (Klein, 1985). Then, if the Spanish Crown ruled through extractive institutions, the collection experience, conceived as a direct measure of extraction, could show us how extraction was under different contexts across regions in Mexico. Following Marichal (2003) and Marichal (2005), we can identify three central taxes: 1) tributes, 2) *alcabalas* (taxes on trade) and 3) *diezmos*. Given these taxes, I will pay attention to tributes and then I will briefly describe the rest of the taxes.

The Spanish adopted tributes from the pre-Hispanic society. In the pre-Hispanic system, there was already a clear structure of who received and paid tributes; this structure prevailed until the New Spanish government replaced it with its own system. Nevertheless, such a system was followed by abuse from the encomenderos and corregidores, responsible for collecting tributes (Miranda, 2005). For example, the *encomenderos* used to impose more tributes than authorized to the indigenous people, and they tended to occupy their lands (Marichal, 2003; Miranda, 2005). Therefore, the tribute is a tax to some extent because its foundations propitiated a behavior of pillage,

suggesting that this kind of behavior could prevail even today.

In the early days, the tributes took place in the *cabecera* that served as an administrative capital; this meant that most of the times, the natives needed to attend with their tributes. Besides, there was also a structure behind the process of *tasación*, a process in which the authority estimated, the “right" amount to charge. In the before-mentioned process, there was a visit, some tributaries were computed, and then the tributes were established according to how much the authorities could collect (Miranda, 2005). This tribute was received in money and goods and then offered in public auctions (Marino, 2001). In the beginning, the collection was a responsibility of the *corregidor*; then it was given to the Royal officials; from the total amount, they receive a proportion of salary. In towns where the tributes were leased, the contractor was responsible for the collecting. At the end of the XVI century, the Crown prohibited to *corregidores* to charge tributes, leaving this capacity to the royal officials or those with the king’s authorization; in towns under the *encomenderos*, there was a butler responsible for collecting such tributes (Miranda, 2005).

Regarding the tendency of tributes, historians have linked their dynamics with demographics (Marichal, 2003, 2005; Molina, 2004). For example, Marichal (2003) observes that the tributes decrease in the second half of the XVIII century, due to the Agrarian crisis of 1785, causing around 300,000 deaths in rural areas. Molina (2004) states that in the first half of the XVIII century before the Bourbon reforms, several crises and epidemics affected mostly the indigenous population, deteriorating their living standards by the end of the century. Despite this, collecting prevailed even in those times; tribute collectors continued collecting tributes, in extreme cases, by extortion (Molina, 2004). The evidence suggests that the increase in tributes might be not only because of the interest of the Crown’s to collect more but also demographic shocks on the indigenous population (Marichal, 2003, 2005; Molina, 2004).

Silver was a symbol of wealth, and it was extracted and processed through very complex techniques (Marichal, 2003). Bars of silver were translated to *Casas de Monedas* so that they could be transformed into coins known as *pesos reales* (Marichal, 2003). In the process, the *Real Hacienda* charged the *diezmo* a 10% tax (Marichal, 2003). The important aspect of silver production is its extraction, which employed native populations, affecting their living conditions. Finally, the

Crown also taxed extensive markets through the *alcabalas*, a fee charge to any object sold and paid by everyone. The importance of the *alcabala* relies on the issue that it can tell us about the general situation of the economy and the existence of important markets (Marichal, 2003).

I identify three features in the tax data that would be helpful in determining the existence of a legacy derived from the extraction experience. The first one is that despite the Crown's attempts to homogenize the tribute rates across all the regions, this was not entirely achieved, suggesting variability in collection across regions. Second, taxes are related to a specific population sector; the tributes affected the indigenous people while the *alcabalas* affected everyone. Thus, we can think that the experience also varied per societal structure. Third, taxes are also related to location and geographic conditions, since these are likely to determine to some extent the population settlements and environments for its growth.

IV Data

The empirical analysis and identification strategy relies on the construction and use of a unique database that combines historical records and contemporary data, at the jurisdiction of the XVIII level and municipality level today. The data includes historical variables on the native population, total population, the presence of the Aztec Empire, taxes and economic activities from the XVIII century.

I compile data on populations from different sources. The principal source for indigenous people and the total population in the mid XVIII century is available in Commons (2002); I complement this data with records from Gerhard (1986) for the territories in the New Spain, the north (Gerhard, 1982), and south of Mexico (Gerhard, 1979). The advantage of using the data compiled by Commons (2002) is that it computes the total number of total populations which allows normalizing the indigenous population. The resulting proportion would account for the social interaction between two groups; a second advantage is that during the first half of the XVIII century, the indigenous population faced several epidemics and agricultural crises. Consequently, the level of population would also depict disease environments. The data in Commons (2002) are obtained from the tributary lists and multiplied by a factor of 4.5; this is because the head of the household

oversaw paying the tribute and the average members ranged between four and five.

The data available for populations comes at the jurisdiction level of the XVIII century. Thus, to assign a level of historical population to each municipality today, I matched each municipality to a jurisdiction of the XVIII century. I performed such task using the historical and geographic information system of the Indias. A municipality is included in a jurisdiction when most of its surface belonged to a jurisdiction. If any place belonged almost equally to two different jurisdictions, I use the *cabecera municipal* today, to determine the jurisdiction. Once I matched every municipality to a jurisdiction, it was easy to determine the population level and the *Intendencia* to which it belonged. It is important to remember that each *Intendencia* had a central royal treasury, allowing assigning of royal treasury to each municipality.

The data for the north of Mexico is less accurate than data available for the South. In these regions, the population does not come at jurisdiction level, but at provincia level, adding up to more than one jurisdiction. For example, the *Provincia del Nuevo Reino de Leon*, today *Nuevo León*, is reported with $1,000$ indigenous people and $5,420$ non-indigenous people; the proportion of indigenous population is 0.18. In this case, such information is assigned to each of the municipality that belonged to the *provincia*. I produced the same produce for the state of *Chihuahua*. Commons (2002) does not offer information for the *Nuevo Santander*, today *Tamaulipas*. I used the data from Vazquez (2006), which lists the number of settlers between 1749 and 1750; then I approximated the indigenous population with data available for decades later (Gerhard, 1982). The data for the state of *Baja California* also considered California; I divided it by two and assigned it to the municipalities that conform this peninsula today. Lastly, I did not consider the territories that today are part of the United States.

The tax data are compiled by TePaske and Klein (1986); *El Colegio de México* has digitally published the data. The data contains information about the *cargo* (income) and *data* (expenditures) for every royal treasury at different points of time. I focus on the period of 1786-1810 because the number of observations is considerably higher, and local records are more likely to be registered. In total, there are 23 royal treasuries distributed across the country, and each one contemplates various sets of taxes. In these data, I pay attention to central treasuries and to three taxes

(tributes, *diezmos*, and *alcabalas*), which I discuss in the previous section. One needs to be careful when using these data, since many taxes were sent to Mexico, even time after the local *cajas* were founded; therefore, the performance of taxes does not appear in the local records for early periods (Klein, 1985). For example, the tributes collected in *Oaxaca* and *Puebla* were sent to Mexico before the creation of the local treasuries, and per the author this represented about the 90% of all tributes collected (Klein, 1985). It is worth mentioning that I do not intend to describe the economy from back then; however, but about how its structure or institutions are revealed through the fiscal system that could ultimately show an effect on current performance. Since these data consist of the amount collected in pesos of the XVIII century, I weight collection with population densities.

As a measure for extractive experience, I propose two extraction indexes. The first accounts for tributes collected within a royal treasury and the native population normalized by the total population in that jurisdiction. The variable *Tributes* is an average of tributes collected between 1786 and 1810.

$$EI_1 = Tributes_c \times \frac{Indigenous\ Population_j}{Total\ Population_j} \tag{1}$$

The second index accounts for the *alcabalas* and tributes, since both seem to be closely related. Nonetheless, the weight for *alcabalas* is the total population in the jurisdiction divided by the entire population within the royal treasury. The reason is that per historical evidence, the *alcabalas* were charged into every single product sold. The variable *Alcabalas* is an average of *alcabalas* collected between 1786 and 1810.

$$EI_2 = Alcabalas_c \times \frac{Total\ Population_j}{Total\ Population_c} + Tributes_c \times \frac{Indigenous\ Population_j}{Total\ Population_j} \tag{2}$$

Additionally, I consider data that captures local characteristics, to uncover the nature of the colonial legacy. Geographic controls are available from two sources, the National Institute of Statistics and Geography (INEGI) and the National Commission for the Knowledge and Use of Biodiversity (CONABIO). I considered longitude, altitude, average temperature, average rainfall, and elevation. From Acemoglu and Dell (2010) I obtained data on slopes for Mexican municipal-

ities. Besides geographic controls, I also incorporated location covariates and sizes fixed effect. For the location controls, I used a dummy variable that equals 1 when the municipality belongs to a state in the north, which considers the proximity to the U.S., and distances from municipalities' centroids to Mexico City. I incorporated two sizes fixed effect; the first equals 1 if a town in 2000 was found to be rural; that is, if it had less than 2,500 people, and the second one, if the city had a total population of more than 100,000. These data are available in the National Population Council (CONAPO) and CONABIO.

As historical controls, I add a dummy variable that equaled 1, if there was a presence of the Aztec Empire in the jurisdiction. Besides, I considered dummy variable for the presence of certain economic activities, in particularly agriculture, cattle, and mining at the jurisdiction level of the mid XVIII century. Data on the presence of the Aztecs is available in text format in (Gerhard, 1986), and the economic activities in the mid XVIII century are available from Commons (2002) and Gerhard (1979, 1982, 1986).

To measure current performance, I took the natural logarithm real per capita GDP at municipality level in 2000. These data was computed by the Mexican Office of Research in Human Development (OIDH). Finally, to test potential mechanisms I employ proximity to paved roads as a measure for the provision of public goods available in Acemoglu and Dell (2010), and literacy rates in 2000 as a proxy for access to education also available in OIDH.

Table 1 provides the descriptive statistics of the data used for the analysis of the colonial legacy in Mexico. The historical observations correspond to two periods. The population variable corresponds to the first half of the XVIII century and the tax data corresponds to the end of the same century. The descriptive statistics for the indigenous population highlight the points of places with high levels of indigenous populations; some places reached a maximum density of 1, meaning that everyone in those municipalities was indigenous. However, the variance of this variable is 0.25, indicating that the ratio was likely to be the same across jurisdictions. The descriptive statistics on tax data shows that there might be a meaningful difference across regions concerning extraction; some areas provide zero revenue to the government. What is surprising is that the two extraction indexes are quite similar, which suggests that both variables could be related to a common concept

Table 1: Descriptive statistics

	Mean	SD	Min	Max	N
Indigenous pop / Total Pop	0.758	0.251	0	1	2,388
Tributes	160,865	144,359	0	466,576	2,331
Alcabalas	262,143	365,369	0	1,148,000	2,331
Diezmos	25,983	46,968	.000	132,154	2,331
Extraction Index 1	140,272	135,872	0	465,704	2,264
Extraction Index 2	141,392	136,345	.0001	470,466	2,264
Per Capita GDP 2000	7,113	4,325	1,324	46,944	2,418
Ln Per Capita GDP 2000	8.701	0.589	7.188	10.76	2,418
Average Rainfall	1,120	643.2	102.5	4,000	2,443
Average Temperature	18.84	3.947	6.625	27	2,452
Latitude	198,312	33,379	144,046	323,828	2,456
Longitud	985,552	43,629	864,446	1,170,000	2,456
Elevation	1,272	823.3	-5	2,903	2,456
Slope	3.826	2.941	0.0235	17.22	2,423
Population w. Royal Treasury	6,476,000	4,798,000	0	13,800,000	2,478
Distance Veracruz	509.5	404.3	8.729	2,531	2,443
Distance Acapulco	576.5	393.1	22.12	2,422	2,443
Distance Mexico City	456.9	374.0	2.112	2,275	2,443
Agriculture Fixed Effect	0.910	0.287	0	1	2,100
Mining Fixed Effect	0.311	0.463	0	1	2,100
Cattle Fixed Effect	0.623	0.485	0	1	1,964
Aztecs	0.466	0.499	0	1	2,464
Rural	0.163	0.369	0	1	2,478
City	0.0960	0.295	0	1	2,478
North	0.113	0.317	0	1	2,478

Source: Own elaboration based on Gerhard (1979, 1982, 1986), Commons (2002), TePaske and Klein (1986), Acemoglu and Dell (2010), INEGI, CONABIO, CONAPO and OIDH.

Notes: Taxes are measured in pesos of the XVIII century. The per capita GDP are displayed in dollars (PPP) at constant prices of 2010. Average Rainfall in millimeters, Average temperature in ^{o}C and distances in Km, slope in degrees and elevation in meters.

of extraction or collection. The geographic controls seem to have small variations within Mexico. For instance, the statistics for average temperature presents a mean of nearly 19 degrees; however, the variance is relatively small, suggesting that average temperature is almost identical for some areas. The location controls that measure distances to ports are quite similar. The descriptive statistics of distance to Mexico City differ considerably to the distance to ports. The variables on historical economic activity fixed effects tell us that in almost all jurisdictions, people practiced agriculture as an economic activity; more than half had cattle but the presence of mining was more scarce, which is intuitive because it depends on the discovery of silver and gold deposits. Then, the fixed effect for the presence of the Aztec Empire suggests that almost half of the territories were in the domain of the Aztec Empire. Finally, in the sample, around 10% of the municipalities can be major cities, and around 16% were considered rural in 2000. The table also notes disparities in current economic performance; the mean for per capita GDP in 2000 was around $7,000$ dollars per year (PPP) with maximums up to $47,000$ dollars per year (PPP), both at constant prices of 2010.

V Basic Empirical Analysis: OLS Estimates

Here, I begin examining the relationship between indigenous population, the share of tributes, *diezmos* and a measure of current economic performance. Figure 1 shows a negative correlation between the natural log of indigenous populations normalized by the total population in the mid XVIII century, and the natural log real per capita GDP in 2000. Places, with a high proportion of indigenous people are associated with a lower real per capita GDP in 2000. Figure 2 presents the second key relationship between the share of tributes collected by the royal treasuries and the natural log real GDP per capita in 2000. The figure proposes that places that had a higher share of tributes had a lower per capita GDP in 2000. Nevertheless, many municipalities that lived a great tribute collection appears to have a high per capita GDP, which is also reasonable, since an alternative hypothesis is that the colonial period affected initial inequality that ultimately promoted growth.

Regarding the *diezmos*, a tax on silver production, Figure 3 shows the relationship between silver production in the colonial period and economic outcomes. Municipalities that belonged to

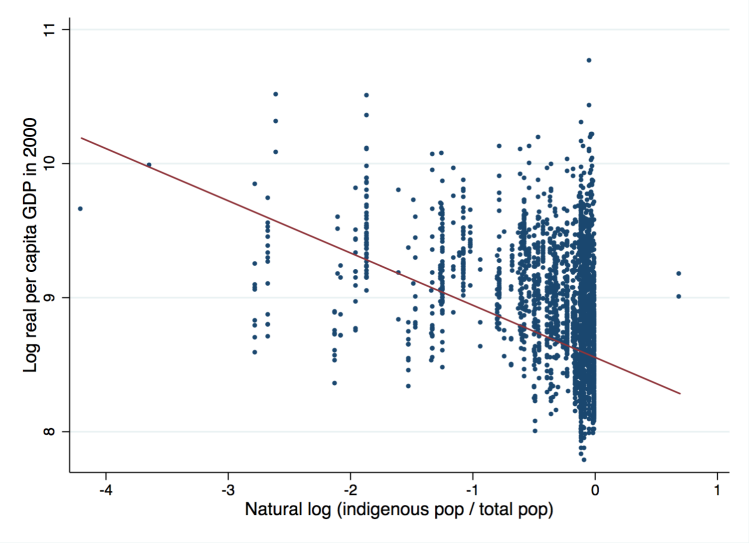

Figure 1: Historical indigenous population and current economic performance

a jurisdiction where mining was relatively significant seem to be more developed today. The last relationship corresponds to the argument of historians that silver brought technical knowledge, which helped the development of markets across territories, which are advantageous in the long-term. However, such argument is not conclusive because with more silver extraction would imply more explode population, which is not necessarily beneficial in the long-term.

One concern is that the shares do not reflect the actual relationship between per capita GDP and tax collection or extraction, because total collection may depend on other factors that are not observable. Consequently, it would be ideal to analyze how collection varied across domains in the XVIII century. Even though such data are not available, I suggest the creation of an extraction index, through the interaction between the amount of revenue collected in a given period and the population affected by the tax burden. Figure 4 displays the connection between the extraction index that accounts for tributes and per capita GDP in 2000. A first impression is that there is more variability in the index than in the share. The connection is negative, but other determinants could drive this correlation. Hence, to isolate the effect of the extraction index, I investigated the association using a linear regression model. I examined the relationship between performance and extraction by controlling for other characteristics that might be essential for the economic

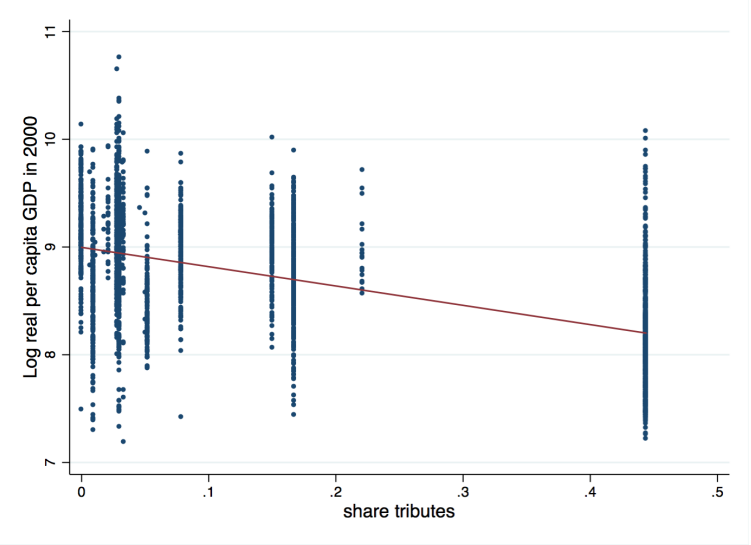

Figure 2: Tributes and current economic performance

development in the long-term. The baseline equation is:

$$ln\ y_{m,j,c} = \alpha + \beta\ ln(Ind_j) + \delta\ ln(Trib_c \times Ind_j) + \mathbf{X}'_{i,j,c}\Omega + \mathbf{X}'_j\Phi + \varepsilon_{m,j,c} \tag{3}$$

where m is an index for the municipality, j the jurisdiction of the XVIII century, and c the royal treasury. The dependent variable is the natural log of y_i, which is the real per capita GDP in municipality i in 2000. lnd_j is the total number of indigenous population normalized by the entire populations in the jurisdiction j in the mid XVIII century. The first extraction index is denoted by $ln(Trib_c \times Ind_j)$. $\mathbf{X}'_{i,j,c}$ is a vector of control variables that capture geographic and location difference across municipalities. $\mathbf{X}'_j$ is a vector of controls at jurisdiction level in the XVIII century. Considering that I am using aggregate variables, I allowed for within-cluster correlations by using robust standard errors clustered at jurisdiction level. The parameters of interest are β and δ which will help us to determine if history can explain current economic performance. In other words, the parameters will determine whether the extractive experience and the early interaction between the population can partially explain the economic difference today.

Table 2 provides the estimates for the baseline model in Column (1)-(4). In column (1) I include geographical controls that might be relevant for a municipality in the long-term. The geograph-

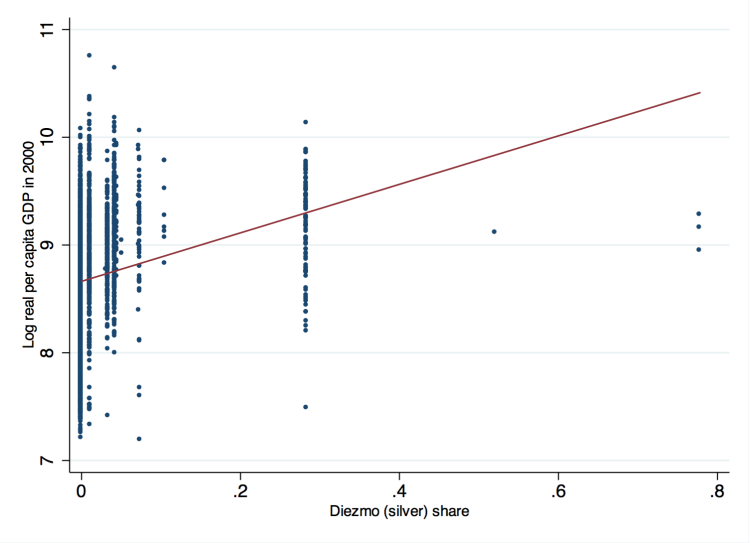

Figure 3: *Diezmo* and current economic performance

ical controls are average rainfall, average temperature, latitude, longitude, elevation, and slopes. Although there may not be enough variation in local geography, these factors are relevant for the presence of certain economic activities, for example, places with high levels of rainfall might be more suitable for agriculture. The model considers the inclusion of latitude and temperature to capture differences between the north and south. I control for elevation and slope, since living conditions might vary depending on the presence of mountain ranges. The importance of controlling for geography relies on the fact that population locates where conditions are favorable, which can bias the results. Once the geographical controls are added, the coefficients remain significant, proposing a negative relationship between the indigenous population and per capita GDP today, while the effect of the extraction index seems to be positive and statistically significant at the 1% level.

Column (2) reports the estimates with distance controls that might be critical for development in the long-term and even for current growth. I include the distance from each municipality centroid to Mexico City and to the Port of Acapulco. The coefficients remain significant. When fixed effect for a rural municipality in 2000 and fixed effects for cities with more than 100,000 habitats in 2000 are incorporated, the coefficient for the extraction and indigenous population remains

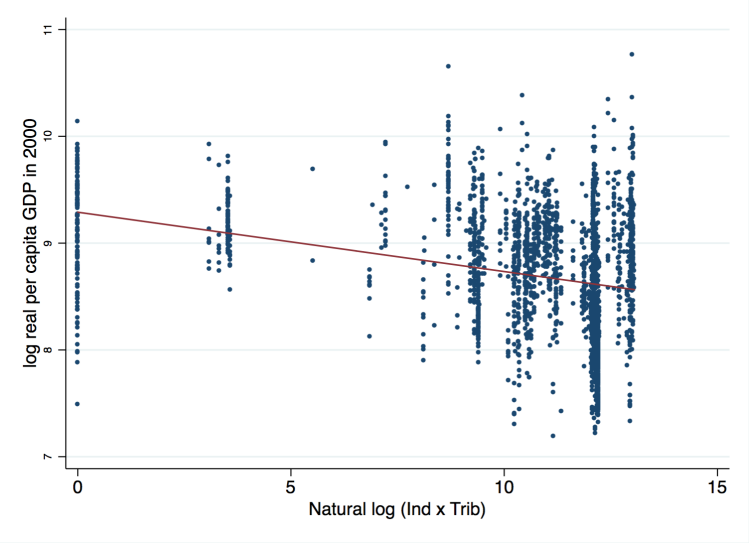

Figure 4: Extraction index and current economic performance

significant. These fixed effects attempt to capture characteristics within localities that can drive current economic performance. Besides, in column (4), I control for municipalities that belong to states in the north, since current economic performance can be due to the proximity of the municipalities to the U.S. The coefficient for the extraction index and indigenous population are significant at the 10 % and 5% level respectively.

In column (5), adding mining production, *diezmos* and historical economic activities, such as agriculture and cattle, does not alter the significance of the estimates. These covariates are added because we can think that population and extraction are likely to be related to the presence of some relevant economic activities. The coefficient becomes more negative for the indigenous population and more positive for the extraction index, which means that the OLS estimates in column (4) were sub-estimating the effect. The measure that accounts for silver production is the natural log of diezmos normalized by the total population in a jurisdiction. I also included the interaction between the mining fixed effect and the silver productions considering that extracting was different than processing. Besides, it turns that the coefficient for diezmos is negative, but it is not statistically significant (not reported).

The last variable introduced references pre-existing traits of societies that possibly induced

Table 2: OLS results

Dependent variable: Natural log real per capita GDP in 2000						
	(1)	(2)	(3)	(4)	(5)	(6)
$ln(Ind)$	-0.19***	-0.16***	-0.11**	-0.11**	-0.16***	-0.15***
	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)
$ln(Ind \times Trib)$	0.05***	0.02*	0.02**	0.02*	0.03*	0.03
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Observations	2,199	2,199	2,199	2,199	1,886	1,886
R-squared	0.48	0.52	0.56	0.57	0.58	0.58
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Distance Controls	No	Yes	Yes	Yes	Yes	Yes
City Fixed Effect	No	No	Yes	Yes	Yes	Yes
Rural Fixed Effect	No	No	Yes	Yes	Yes	Yes
North Fixed Effect	No	No	No	Yes	Yes	Yes
Agriculture Fixed Effect	No	No	No	No	Yes	Yes
Cattle Fixed Effect	No	No	No	No	Yes	Yes
Diezmo Control	No	No	No	No	Yes	Yes
Mining Fixed Effect**Diezmos*	No	No	No	No	Yes	Yes
Aztec Fixed Effect	No	No	No	No	No	Yes

Notes: The Table reports the OLS estimates. The unit of observation is municipality. Standard errors adjusted for clustering within jurisdiction of the XVIII century in parentheses. The geographic controls are average temperature, average rainfall, latitude, longitude, elevation and slope. The distance controls include distance from each municipality to Mexico City and to Acapulco. City is a fixed effect that equals one if the population in 2000 exceeded 100,000. Rural fixed effect account for municipalities with less that 2500 people in 2000. The variable *Ind* stands for indigenous population normalized by total population. The interaction is the extraction index $Ind \times Trib$. Coefficients that are significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.

populations into a more extractive system or initially under-developed areas. It is important to remember that the tribute system was inherited from the Aztecs; therefore, an alternative explanation is that societies under Aztec rule were selected into a more intensive extractive experience or less developed places. Thus, the effect that we observed can be due to the pre-Hispanic institutions. In column (6), I controlled for the presence of the Aztec Empire before the conquest, the coefficient for the extraction index becomes insignificant.

VI Identification Strategy

An important issue is endogeneity, since it is unclear that the Colonial rule in Mexico had a direct effect on current economic performance. I consider that there are mainly two endogenous variables; these are the indigenous population and the tributes. For the native population, an alternative story is that these were in initially poor places. Furthermore, if tributes were related to population, places with a higher rate of payment could show living conditions that were beneficial in the long run. The second source of endogeneity is measurement error, since it might be the case that local authorities underreported population and therefore tributaries. When doing so, officials could collect tributes for themselves. Regarding the *diezmos*, I will assume that these are to some extent exogenous, the reason is that they show geographical conditions.

The strategy that I pursue is to use instrumental variables that are correlated with the population distribution and the extraction index that considers tribute collection. The first advantage of using instrumental variables is that they will allow exogenous variation, which will determine the precise causal effect. The second benefit is that if the instruments are not correlated with the measurement error in the number of indigenous people or the tributes collected, then IV will yield consistent estimates. As instruments, I used the variation on distances from Mexico City to royal treasuries, and from local treasuries to municipality centroids and local geography. The distance instruments are valid under the assumption that the supply location affected the demand location and not the other way around [1]. On the one hand, if royal treasuries may have located in places with a high supply of tributes then the instruments are valid. On the contrary, if the location of royal treasuries also influenced the supply of tributes, in such manner that municipalities decided to locate, then the instruments are not valid. For the variable regarding the density of indigenous population, the intuition of the main instrument is straightforward. It assumes that non-indigenous populations located mainly in Mexico and the main cities. The further the city, fewer the number of non-indigenous people. Besides, it also considers that indigenous population concentrated in places around the main cities.

The instruments are based on historical evidence. Selection was important to determine what

[1] I consider that the *Real Hacienda* demanded taxes and communities are suppliers. See Nunn (2008a)

taxes to collect in each of the regions where a treasury was funded. Besides, in the second half of the XVIII century, new treasuries were born to increase the efficiency of collection. For this reason, I instrument collection with distances. I consider the distance from Mexico to each royal treasury, which reflects transportation costs, the longer the distance, the harder to collect. Furthermore, the instrument also accounts for the collection to vary within a treasury's jurisdiction, according to the local distances and local geography. The second endogenous variable is the natural log of indigenous population normalized by the total population. Such variable is instrumented by a ratio of distances and a variable that proxies the distance to the Port of Veracruz. This ratio considers that the whole population structure in the XVIII century varied according to how far the settlements were from Mexico City, and the main cites. Since the north territories were just discovered, naturally there were fewer populations than in the center or south of Mexico. The last instrument is the distance from one of the largest ports back of the time, *Veracruz*, which connected New Spain with the Atlantic Ocean. Besides, Veracruz is one of the first villages founded in New Spain; then we can expect that the structure of the population varied from that village.

The first stage regressions of the 2SLS procedure are:

$$\begin{aligned} ln(Ind_j) = \phi_0 + \phi_1 SD_{m,j,c} + \phi_2 WD_{m,j,c} + \phi_3 RD_{m,j,c} + \phi_4 VP_{m,j,c} \\ + \mathbf{X}'_{m,j,c}\Omega + \mathbf{X}'_j\Phi + u_{m,j,c} \end{aligned} \tag{4}$$

$$\begin{aligned} ln(Trib_c \times Ind_j) = \pi_0 + \pi_1 SD_{m,j,c} + \pi_2 WD_{m,j,c} + \pi_3 RD_{m,j,c} + \pi_4 VP_{m,j,c} \\ + \mathbf{X}'_{m,j,c}\Omega + \mathbf{X}'_j\Phi + v_{m,j,c} \end{aligned} \tag{5}$$

In this first stage, the *SD* instrument captures the total distance Mexico-royal treasury-municipality, WD the distances between royal treasuries and municipalities centroids, *RD* the ratio of distances and VP a proxy that measures the distance to the Port of Veracruz. The instruments interact with local geography for which I consider slope and elevation. The idea behind the interaction with local geography is that it can be a primary source of exogeneity and intuition. The population might have located in places that were more favorable for living, besides collection becomes harder when local geography imposes a challenge.

Table 3: IV results

	Dependent Variable: ln per capita GDP in 2000				
	(1)	(2)	(3)	(4)	(5)
$ln(Ind)$	-0.5345**	-0.4977*	-0.3269*	-0.3417**	-0.3420**
	(0.2680)	(0.2728)	(0.1684)	(0.1442)	(0.1383)
$ln(Ind \times Trib)$	0.1043**	0.0765*	0.0689**	0.0856***	0.0888**
	(0.0456)	(0.0436)	(0.0291)	(0.0331)	(0.0360)
$Over-identification\ Test$	.14	.12	.13	.21	.11
$Endogeneity\ Test$	.002	.002	.011	.003	.005
	First stage: Dependent variable: $ln(Trib \times Ind)$				
Sum Distances	-0.0038***	-0.0040***	-0.0054***	-0.0051***	-0.0048***
	(0.0004)	(0.0005)	(0.0010)	(0.0009)	(0.0009)
Ratio Distances	0.0005	0.0004	0.0004	0.0002	0.0003
	(0.0005)	(0.0004)	(0.0004)	(0.0005)	(0.0005)
Within Distance	0.0038***	0.0040***	0.0055***	0.0052***	0.0048***
	(0.0005)	(0.0005)	(0.0010)	(0.0010)	(0.0010)
Veracruz Port	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$F-stat$	43.51	41.93	13.32	10.13	12.9
	First stage: Dependent variable: $ln(Ind)$				
Sum Distances	0.0006***	0.0006***	0.0004*	0.0003	0.0003
	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0003)
Ratio Distances	0.0008***	0.0008***	0.0009***	0.0011***	0.0011***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Within Distance	0.0006***	0.0006***	0.0005**	-0.0003	-0.0004
	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0003)
Veracruz Port	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$F-stat$	4.2	3.9	6.3	6.14	8.48
Observations	1,893	1,893	1,893	1,886	1,886
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Aztecs Fixed Effects	Yes	Yes	Yes	Yes	Yes
Historical Economic Activity	Yes	Yes	Yes	Yes	Yes
Locations Fixed Effects	No	Yes	Yes	Yes	Yes
Distance Controls	No	No	Yes	Yes	Yes
Diezmo Control	No	No	No	Yes	Yes
Mining fixed effect $\times$ *Diezmos*	No	No	No	Yes	Yes

Notes: The Table reports the IV estimates. The unit of observation is municipality. Standard errors adjusted for clustering within jurisdiction of the XVIII century in parentheses. The geographic controls, the distance controls are specified as in the baseline model. The location fixed affects include city fixed effects, rural fixed effect and north fixed effect. The variable *Ind* stands for indigenous population normalized by total population and $Ind \times Trib$ the extraction index. In (1)-(4) the distance from royal treasury to municipalities' centroid interact with the slope as a measure for local geography. In Column (5), the instrument that account for the distance to the Port of Veracruz consider local elevation. Coefficients that are significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.

Table 3 reports the main findings regarding causality. The results are in line with the OLS estimates, the coefficient for indigenous population is negative, and for the tribute collection is positive. Both coefficients are statistically significant at different levels when adding several controls. The results suggest that there is the colonial legacy that can be divided into two opposite effects. Regarding the magnitude, the coefficient for indigenous population is more negative, indicating that OLS were sub-estimating the adverse effect. Similarly, the coefficients for tributes collection are larger than the OLS estimates. In column (1), I include geographical controls, Aztecs fixed effect and historical economic activity fixed effects. Columns (2)-(3) add location and distance controls respectively. Finally, in column (4) I incorporate the silver production, measured as an interaction between the *diezmos* normalized by the total population with the mining fixed effect; such coefficient is negative and significant at 10% (not reported), suggesting a negative effect derived form silver extraction. The last column (5) consider an alternative instrument to capture the proximity to the port of Veracruz, in this instrument the local distance interacts with the slope and elevation, the coefficients are similar in magnitude.

In the first stage, tax collection depends negatively on the distance from Mexico City to the royal treasury plus the distance from the royal treasury to each municipality centroid interacted by local geography factors, such as slope and height. What seems surprising is that the longer the distance and less favorable the local geography from royal treasuries to municipalities, the more tributes are collected. However, if tributes are more likely to be related to population and growth conditions, it might be the case that there were better conditions for the indigenous people far from the main cities. Regarding the indigenous population, the first stage tells us that it is positively associated with the instrument of relative distance. Nevertheless, for this variable, the distance to the Port of Veracruz does not seem relevant. Although the population structure is not meant to be instrumented with distances, the first stage suggests a positive relationship between the density of indigenous people and the distance to the local royal treasury. The results confirm that royal treasuries were located in places surrounded by communities with a high density of indigenous populations, since the further the municipalities from the royal treasury, the higher the indigenous people.

Table 4: LIML results

Dependent Variable: ln per capita GDP in 2000					
	(1)	(2)	(3)	(4)	(5)
$ln(Ind)$	-0.624	-0.615	-0.346*	-0.353**	-0.347**
	(0.402)	(0.412)	(0.194)	(0.156)	(0.144)
$ln(Trib \times Ind)$	0.094	0.063	0.069**	0.087**	0.089**
	(0.062)	(0.060)	(0.030)	(0.034)	(0.037)
Observations	1,893	1,893	1,893	1,886	1,886
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Aztecs Fixed Effects	Yes	Yes	Yes	Yes	Yes
Historical Economic Activity	Yes	Yes	Yes	Yes	Yes
Locations Fixed Effects	No	Yes	Yes	Yes	Yes
Distance Controls	No	No	Yes	Yes	Yes
Diezmo Control	No	No	No	Yes	Yes
Mining fixed effect $\times$ *Diezmos*	No	No	No	Yes	Yes

Notes: The Table reports the LIML estimates. The unit of observation is municipality. Standard errors adjusted for clustering within jurisdiction of the XVIII century in parentheses. The geographic controls, the distance controls, city fixed effect, rural fixed effect are as specified in the baseline model. The variable *Ind* stands for indigenous population normalized by total population and $Ind \times Trib$ the extraction index. Coefficients that are significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.

In Table 3, I also include the p-value for the over-identification test and endogeneity test of the regressors. Under the null hypothesis of the over-identification, the instruments are valid. The instruments are valid; the null is not rejected at 10% in all the specifications. Under the null of the endogeneity test, the regressors are considered exogenous, and therefore OLS estimates are consistent. Column (5) reports the results for an alternative instrument regarding the distance to Veracruz. In all columns, given the p-value of the endogeneity test, the null hypothesis is rejected, confirming that the use of IV is preferable. An important issue is that the over-identification test might be just in the margin. I combined and tested several sets of similar instruments; however, the ones that I report in the table are more likely to underestimate the effect in the case that they violate the exclusion restriction.

In the first stage, the instruments for the indigenous population are weak; the $F - statistic$ is at most 8.48. Some alternative estimators are robust in the presence of a weak instrument. Therefore,

Table 4 reports the same five specifications using a limited information maximum likelihood estimator. The primary results hold a negative effect that comes from the population structure and a positive from collection. The IV results suggest that there is something in the structure of the population that has been persistent in Mexican society rather to think that the presence of indigenous populations is something that is not beneficial. On the other hand, the effect due to the presence of the *Real Hacienda* and tribute collection can come from two different ways that do not contradict each other. If tributes are related to population, then collection itself would reflect a set of favorable conditions that allowed the payment of such taxes. The second one is that places with tribute gatherings experienced a higher level of inequality, which ultimately brought economic development in the long run. Since geographical controls are added in the model, then the positive effect is more likely to come from inequality. The latter argument correlates with recent literature on the consequences of the Colonial rule in Latin America. Nevertheless, evidence of this cannot discard the adverse effect that is likely to come from how people interact with one another.

VII Robustness Check

As mentioned above, one concern is that population data are not very accurate for some regions in the north of Mexico. As a robustness check, it is important to know if the results are biased due to these observations. Table 5 reports the IV estimates using the restricted sample, the principal results hold. There are two opposite effects: a negative effect derived from the population structure and a positive effect from the tribute collection. The results are statistically significant when adding controls and several fixed effects.

A second issue is the positive impact of the tribute collection. If it this references inequality besides favorable conditions for population growth, then it should be the case that we see a similar effect when using an alternative tax, such as the *alcabalas*. In Table 6, I show the results with the same identification strategy using *alcabalas* as explanatory variables. It is important to mention that in this case, I am not evaluating the effect of the alcabalas, since a proper identification of this tax would require interaction with the total population. In this case, I am using it as an alternative proxy for tributes. In column (1) and (3), I report the IV estimates for the whole and the restricted

Table 5: Robustness check-Result for restricted sample

	Dependent Variable: ln per capita GDP in 2000				
	(1)	(2)	(3)	(4)	(5)
$ln(Ind)$	-0.5370**	-0.5002*	-0.3244*	-0.3838***	-0.3723***
	(0.2693)	(0.2737)	(0.1700)	(0.1460)	(0.1427)
$ln(Trib \times Ind)$	0.1044**	0.0770*	0.0688**	0.1204***	0.1189***
	(0.0456)	(0.0434)	(0.0288)	(0.0385)	(0.0386)
Observations	1,890	1,890	1,890	1,886	1,883
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Aztecs Fixed Effects	Yes	Yes	Yes	Yes	Yes
Historical Economic Activity	Yes	Yes	Yes	Yes	Yes
Locations Fixed Effects	No	Yes	Yes	Yes	Yes
Distance Controls	No	No	Yes	Yes	Yes
Diezmo Control	No	No	No	Yes	Yes
Mining fixed effect $\times$ *Diezmos*	No	No	No	Yes	Yes

Notes: The Table reports the IV estimates for the restricted sample. The unit of observation is municipality. Standard errors adjusted for clustering within jurisdiction of the XVIII century in parentheses. The geographic controls, the distance controls, city fixed effect, rural fixed effect are as specified in the baseline model. The variable *Ind* stands for indigenous population normalized by total population and $Ind \times Trib$ the extraction index. Coefficients that are significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.

sample, respectively. Column (2) and (4) consider the presence of weak instruments and report the LIML estimates for the whole and limited sample. Column (5), uses the alternative instrument for the distance to the Port of Veracruz. The specifications use robust standard errors clustered at the jurisdiction level. The over-identification test suggests that the instruments are valid and the endogeneity test highlights the importance of the IV procedure. The results are very similar to those in Table 3. There are two opposite effects: a negative from the population and a positive from tax collection.

The fact that *alcabalas* are significant may suggest that tax collection itself is what drives the positive effect. Therefore, I proceed to the construction of an extraction index that incorporates the information from *alcabalas* and tributes. In this index, I add both taxes; however, the weights are different, since each tax aimed to different people. The *alcabalas* interact with the total population in a jurisdiction, normalized by the population within the royal treasury. The tributes interact with

Table 6: Robustness check- *Alcabalas* as proxy for tributes

	Dependent Variable: ln per capita GDP in 2000				
	(1)	(2)	(3)	(4)	(5)
$ln(Ind)$	-0.8620***	-0.8910***	-0.8718***	-0.9028***	-0.8199***
	(0.1966)	(0.2076)	(0.2005)	(0.2125)	(0.1950)
$ln(Alcab \times Ind)$	0.1932***	0.1970***	0.1949***	0.1989***	0.1858***
	(0.0651)	(0.0671)	(0.0657)	(0.0678)	(0.0643)
Observations	1,886	1,886	1,883	1,883	1,886
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Aztecs Fixed Effects	Yes	Yes	Yes	Yes	Yes
Historical Economic Activity	Yes	Yes	Yes	Yes	Yes
Locations Fixed Effects	Yes	Yes	Yes	Yes	Yes
Distance Controls	Yes	Yes	Yes	Yes	Yes
Diezmo Control	Yes	Yes	Yes	Yes	Yes
Mining fixed effect $\times$ *Diezmos*	No	No	No	Yes	Yes

Notes: The Table reports the IV-LIML estimates, using *alcabalas* an alternative measure for extraction. The unit of observation is municipality. Standard errors adjusted for clustering within jurisdiction of the XVIII century in parentheses. The geographic controls, the distance controls, city fixed effect, rural fixed effect are as specified in the baseline model. The variable *Ind* stands for indigenous population normalized by total population and $Ind \times Alcab$ is a proxy for the extraction index. Coefficients that are significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.

the indigenous people normalized by the entire population in that jurisdiction. One reason is that *alcabalas* had a more important presence in all the *Intendencia*, and since this tax reflects the importance of markets, it is desirable to know their local importance with respect to all markets in the administration. Furthermore, I also weighted the alcabalas only with the total population in the jurisdiction, and the results do not change, but the significance is slightly different.

In Table 7, the IV and LIML results for the extraction index are reported. The key results hold that there is a positive effect from extraction and a negative effect from the population structure. In column (1)-(3), the IV estimates for the whole and limited sample, respectively, are reported while in column (2)-(4), I consider the presence of weak instruments and report the LIML estimates. In Column (5), I use the alternative instrument for the distance to the Port of Veracruz. In all specifications, the coefficients are significant, the over-identification test suggests that instruments are valid and the endogeneity tests that the IV estimates are preferable to the OLS. When comparing

the estimates with those in the tributes in Table 3, we can observe that such coefficient is lower in the case of the indigenous population and higher for extraction, which is a very natural result, contemplating that one tax will underestimate the effect of the extraction.

To discard that alternative factors in 2000 are driving the results, I also follow the identification strategy using the natural log per capita GDP 2010 as the dependent variable at the municipality level. Although some geographical data are until 2000, the identification strategy is still reasonable, if we accept that on average geography does not change. The results do not vary when accounting for a different year of per capita GDP. The over-identification test fails to reject the null hypothesis that the instruments are valid. The endogeneity test rejects the null hypothesis that the OLS are consistent; therefore, IV is preferable. Additionally, I also consider alternative periods for measuring historical tax collection, and the effects persevere; nevertheless, for the variables that examine tax collection after 1790, it turns that the coefficient for silver production is not significant. [2].

A potential threat for the identification is that the instruments do not fulfill the exclusion restriction because they may have a direct effect on per capita GDP, independently of the local collection and population. In other words, distances to royal treasuries, Mexico City and Veracruz Port might be associated with distances to relevant locations that promoted economic growth. One way to test this issue is to estimate the reduced form between GDP and the instruments. Such estimation is shown in Table 8. In column (1), I use the whole sample and in column (2) the restricted sample. In curly brackets, I report the robust standard error, in square brackets standard errors clustering within jurisdiction are reported, and in parenthesis robust standard errors clustering within royal treasury are reported. In this table, two instruments seem to be significant, the distance to the Port of Veracruz and total distance $Mexico\ City - Royal\ treasury - municipality$. Such coefficient is only significant in the presence of robust standard errors, and they are very close to zero.

One possible explanation for the potential violation of the exclusion restriction is that being far from Mexico City and Veracruz Port could be harmful in the long-term. This location effect is natural, since high economic activity characterized these places. In the identification strategy, I control for distance to Mexico City, although the coefficient is negative and is not statistically different from zero. However, it turns that the direct relationship between per capita GDP and the

[2]These results are available upon request

Table 7: Extraction Index and current economic performance

	Dependent Variable: ln per capita GDP in 2000				
	(1)	(2)	(3)	(4)	(5)
$ln(Ind)$	-0.4293***	-0.4390***	-0.4310***	-0.4410***	-0.4418***
	(0.1470)	(0.1539)	(0.1501)	(0.1573)	(0.1471)
$ln(Extraction\ Index)$	0.1382***	0.1395***	0.1383***	0.1397***	0.1262**
	(0.0491)	(0.0496)	(0.0492)	(0.0497)	(0.0499)
Observations	1,886	1,886	1,883	1,883	1,886
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Aztecs Fixed Effects	Yes	Yes	Yes	Yes	Yes
Historical Economic Activity	Yes	Yes	Yes	Yes	Yes
Locations Fixed Effects	Yes	Yes	Yes	Yes	Yes
Distance Controls	Yes	Yes	Yes	Yes	Yes
Diezmo Control	Yes	Yes	Yes	Yes	Yes
Mining fixed effect $\times$ *Diezmos*	No	No	No	Yes	Yes

Notes: The Table reports the IV-LIML estimates, using an extraction index that account for tribute and *alcabalas*. The unit of observation is municipality. Standard errors adjusted for clustering within jurisdiction of the XVIII century in parentheses. The geographic controls, the distance controls, city fixed effect, rural fixed effect are as specified in the baseline model. The variable *Ind* stands for indigenous population normalized by total population. Coefficients that are significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.

instrument that considers the distance to Mexico City is positive. When the identification includes the distance to Veracruz, it is not significant but positive (not reported); in the reduced form the instrument that uses the distance to Veracruz also presents a positive correlation with per capita GDP in 2000. Therefore, if such bias exists, the IV estimates are lower bound.[3] Other threats consist of the associations of the instrument with the explanatory variables; this is not the case, since the relationships are not significant. Finally, the fact that the instruments are somehow weak suggests that further research should include a better identification of the population structure with better instruments or alternative methods that ensure internal validity.

[3]Miguel et al. (2004) follow a similar argument to justify that in case of a direct relationship between the instrument and the dependent variables, the IV estimates can be lower bound.

Table 8: Reduced form- current economic performance and excluded instruments

	(1)	(2)
Ratio Distances	-0.0001	-0.0001
	[.00009]	[.00009]
	{.0003}	{ .0003}
	(0.0005)	(0.0005)
Sum Distances	0.0001	0.0001
	[.00003]	[.00003]
	{.0001}	{.0001 }
	(0.0002)	(0.0002)
Within Distances	0.0000	0.0000
	[.00006]	[.00006]
	{.0002}	{ .0002}
	(0.0003)	(0.0003)
Veracruz Port Distance	0.0000	0.0000
	[.00006]	[.00000]
	{.00000}	{.0000 }
	(0.0000)	(0.0000)
Observations	2,282	2,239

Notes: The Table reports the direct correlation between the excluded instruments and per capita GDP in 2000 at municipality level. Below each coefficient three standard errors are reported. The first, reported in brackets, is a robust standard errors. The second, reported in curly brackets, is standard error adjusted for clustering within jurisdiction of the XVIII century. The third, reported in curly parenthesis, is standard error adjusted for clustering within royal treasury.

VIII Potential Mechanisms

Although there may be many potential mechanisms through which the Colonial period could have a persistent effect on current economic performance in Mexico, I analyzed the public goods provision, specifically, proximity to paved roads and literacy rates. In recent literature, the evidence suggests that disparities in institutions might be linked to different public goods provision in Latin America. Acemoglu and Dell (2010) describe such a pattern, arguing that differences across countries and within a country could be due to differences in institutions, which are also related to the gap in access to paved roads. Dell (2010) also finds evidence for the persistence of mita in Peru, through road networks around the mita boundary, arguing that the presence of large landowners provided a system that encouraged public goods provision.

The pattern tested is that if access to paved roads as a proxy for public goods provision is closely related to current economic performance in Mexico, then It should be the case that public goods provision is also correlated to a historical legacy. For this purpose, Table 9 examines the relationship between the natural log of distance to paved roads, population structure in the XVIII century and the extraction index that consider tributes and *alcabalas* collection.

The identification strategy is the same used for the causal identification of the historical legacy. It uses a set of four instruments referring to distances to Mexico City, royal treasuries, Veracruz and local geography. As a dependent variable, I considered the proximity to paved roads as a proxy for public goods provision. The population structure is negative but not statistically significant. In column (1)-(2), I report the OLS estimates, although column (1) does not add silver production as a control; in both columns, the extraction is positive and significant at the 5% level. Since the variable of interest could be endogenous, column (3) and (4) reports the IV estimates; the coefficient of the extraction index remains significant at the 10% level. The over-identification test suggests that the instruments are valid and the endogeneity test fails to reject that OLS are consistent. Finally, the last column presents the LIML estimates, which are more robust in the presence of weak instruments.

In sum, I find evidence that public goods provision could be a potential mechanism through which the Colonial legacy persists. Places, where collection or extraction was considerably higher,

have a higher distance to paved roads today; in other words, fewer public goods provision, which promoted economic development to some extent.

Additionally, given the results in Table 9, the OLS estimators are likely to be consistent. I performed a heterogeneous analysis to determine what is driving the results. Intuitively, we can think that the mechanism varies under different contexts, we can expect different impacts in rural municipalities, big cities, or in the south of Mexico. Table 10 examines the relationship between distance to paved roads, population structure, and extraction. Column (1) shows the baseline model for the mechanism. In column (2), we can observe that municipalities with less than 100,000 habitants, extraction has a greater effect, such municipalities have a longer distance to paved roads today. The population structure is not statistically significant (not reported). In column (3), I considered that the effect can be different in municipalities that in 2000 were considered rural. It turns that the population structure has a substantial impact on public goods provision; a rural municipality that had a high presence of indigenous population enjoys a lower level of public goods provision. Regarding extraction, rural municipalities have a larger coefficient than non-rural, meaning lower provision in these rural communities. In column (4), the estimates suggest that provision is lower in the south. When a municipality is considered rural and located in the south, public goods provision is lower than rural areas in the north. In summary, extraction hurts public goods provision; if non-rural places experience greater provision, then non-rural areas also experience more growth than rural areas, explaining to some extent the differences we observe today within Mexico.

The second mechanism relies on the fact that extraction and population structure lowered access to education, which is also related to public goods provision. Table 11 examines this potential mechanism. The independent variable is the literacy rate at municipality level in 2000. The only difference in the identification strategy is that I omit the instrument that references the distance to the Port of Veracruz. In column (1), the OLS estimates are reported, which are not significant. Column (2)-(3), examines the causal relationship using the whole and the restricted sample, respectively. The effects are very similar when per capita GDP was analyzed. Places with a high level of indigenous populations in the XVIII century have lower access to education today. Re-

Table 9: Distance to paved roads as potential mechanism

	Dependent Variable: ln distance to paved roads				
	(1)	(2)	(3)	(4)	(5)
$ln(Ind)$	0.0016	-0.0188	-0.0958	0.0908	0.0958
	(0.0789)	(0.0960)	(0.3207)	(0.2909)	(0.3008)
$ln(ExtractionIndex)$	0.1035**	0.1150**	0.1250**	0.1469*	0.1473*
	(0.0465)	(0.0534)	(0.0620)	(0.0860)	(0.0866)
$Overidentificationtest$	-	-	.56	.35	.35
$Endogeneitytest$	-	-	.66	.62	.62
	First stage: Dependent variable: $ln(Extraction\ Index)$				
Sum Distances	-	-	-0.0032***	-0.0031***	-.0028***
	-	-	(0.0006)	(0.0006)	(0.0005)
Ratio Distances	-	-	0.0011***	0.0010**	0.0001***
	-	-	(0.0003)	(0.0004)	(.0004)
Within Distance	-	-	0.0065***	0.0056***	0.0052***
	-	-	(0.0013)	(0.0015)	(.0014)
Veracruz Port	-	-	0.0033**	0.0026*	0.0024*
	-	-	(0.0013)	(0.0015)	(0.0013)
$F-stat$			26.07	12.55	12.55
	First stage: Dependent variable: $ln(Ind)$				
Sum Distances	-	-	0.0004*	0.0003	.0003
	-	-	(0.0003)	(0.0003)	(0.0003)
Ratio Distances	-	-	0.0009***	0.0012****	0.0012***
	-	-	(0.0002)	(0.0002)	(0.0002)
Within Distance	-	-	-0.0005	0.0006	0.0006
	-	-	(0.0007)	(0.0008)	(0.0008)
Veracruz Port	-	-	0.0000	0.0009	0.0009
	-	-	(0.0007)	(0.0007)	(.0007)
$F-stat$			6.36	6.55	6.55
Observations	1,893	1,893	1,893	1,886	1,886
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Aztecs Fixed Effects	Yes	Yes	Yes	Yes	Yes
Historical Economic Activity	Yes	Yes	Yes	Yes	Yes
Locations Fixed Effects	No	Yes	Yes	Yes	Yes
Distance Controls	No	No	Yes	Yes	Yes
Diezmo Control	No	Yes	No	Yes	Yes
Mining fixed effect $\times$ *Diezmos*	No	Yes	No	Yes	Yes

Notes: The Table reports the OLS-IV estimates for proximity to paved road as potential mechanism The unit of observation is municipality. Standard errors adjusted for clustering within jurisdiction of the XVIII century in parentheses. The geographic controls, the distance controls, city fixed effect, rural fixed effect are as specified in the baseline model. The variable *Ind* stands for indigenous population normalized by total population. In (3)-(4) the distance from royal treasury to municipalities' centroid interact with the slope as a measure for local geography. In Column (5), the instrument that account for the distance to the Port of Veracruz consider local elevation. Coefficients that are significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.

Table 10: Proximity to Paved Road-Heterogeneous effects

	Dependent Variable: ln distance to paved roads				
	(1)	(2)	(3)	(4)	(5)
$ln(EI)$	0.1150**				
	(0.0534)				
$0.city \times ln(EI)$		0.1165**			
		(0.0535)			
$1.city \times ln(EI)$		0.1107**			
		(0.0552)			
$0.rural \times ln(Ind)$			-0.0278		
			(0.0953)		
$1.rural \times ln(Ind)$			0.3033**		
			(0.1453)		
$0b.rural \times ln(EI)$			0.1150**		
			(0.0531)		
$1.rural \times ln(EI)$			0.1162**		
			(0.0528)		
$0.north \times ln(Ind)$				-0.0794	
				(0.1098)	
$1.north \times ln(Ind)$				0.0401	
				(0.1850)	
$0.north \times ln(EI)$				0.1276**	
				(0.0579)	
$1.north \times ln(EI)$				0.0378	
				(0.0611)	
$0.north \times 0.rural \times ln(EI)$					0.1232**
					(0.0573)
$0.north \times 1.rural \times ln(EI)$					0.1282**
					(0.0564)
Observations	1,904	1,904	1,904	1,904	1,904
R-squared	0.3905	0.3907	0.3920	0.3864	0.3882

Notes: The Table reports the OLS-heterogeneous effects for proximity to paved road as potential mechanism. The unit of observation is municipality. Standard errors adjusted for clustering within jurisdiction of the XVIII century in parentheses. All specifications include geographic controls, distance controls, city fixed effect, rural fixed effect, north fixed effect, historical economic activity fixed effect and Aztecs fixed effects. The variable *Ind* stands for indigenous population normalized by total population. *EI* stands for extraction index that account for tribute and *alcabalas*. Coefficients that are significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.

Table 11: Literacy as potential mechanism

Dependent Variable: Literacy rate in 2000 at municipality level				
	(1)	(2)	(3)	(4)
$ln(Ind)$	-0.0190	-0.3023***	-0.3164***	-0.4031**
	(0.0120)	(0.0998)	(0.1072)	(0.1764)
$ln(Extraction\ Index)$	0.0011	0.1386**	0.1472**	0.2031*
	(0.0068)	(0.0683)	(0.0725)	(0.1214)
$Over-identificationtest$	-	.17	.19	.32
$Endogeneitytest$	-	.038	.036	.038
Observations	1,904	1,904	1,901	1,904
Geographic Controls	Yes	Yes	Yes	Yes
Aztecs Controls	Yes	Yes	Yes	Yes
Historical Economic Activity Fixed Effect	Yes	Yes	Yes	Yes
Distance Controls	Yes	Yes	Yes	Yes
North Fixed Effect	Yes	Yes	Yes	Yes
City Fixed Effect	Yes	Yes	Yes	Yes
Rural Fixed Effect	Yes	Yes	Yes	Yes

Notes: The Table reports the OLS-IV estimates for literacy as potential mechanism. The unit of observation is municipality. Standard errors adjusted for clustering within jurisdiction of the XVIII century in parentheses. All specifications include geographic controls, distance controls, city fixed effect, rural fixed effect, north fixed effect, historical economic activity fixed effect and Aztecs fixed effects. The variable *Ind* stands for indigenous population normalized by total population. *EI* stands for extraction index that account for tribute and *alcabalas*. Coefficients that are significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.

garding extraction, we can observe a positive effect. Column (4) reports LIML estimates, both coefficients remain significant. The over-identification test fails to reject that the instruments are valid and the endogeneity test suggests that IV estimates are preferable. The results suggest that the mechanism of public goods is reasonable, the difference we see in Mexico can be explained by the disparities in public goods provision. Then, one possible explanation is that the Colonial legacy persist through the rules of redistribution, and that these have not been historically favorable for certain population and regions.

IX Concluding Remarks

This work has studied the Colonial legacy of extractive institutions, as a partial explanation of current differences in economic performance across regions within Mexico. In doing so, it is one of the first studies that consider that history matters through the institutions built in Mexico during the Colonial period. The study also contributes to the debate on how extractive institutions work and how they persist in the context of the Mexican case.

The evidence relies on a unique database that compiled historical and contemporary data. I used two variables that attempted to capture the extractive nature of the former institutions; these are the indigenous population and tax collection as a direct measure of extraction. Following an identification strategy based on instrument variables, I found two opposite effects. The evidence suggests an adverse effect in the long-term that comes from the institutions that shaped social interaction; if we accept that extractive institutions emerged where the indigenous population was abundant, then municipalities that presented a high density of native people are less developed today. The collection or extraction seems to have a positive effect, and although this can be attributed to favorable geographical conditions, it does not discard that tax collection brought inequality that ultimately promoted economic development in the long-term. The key assumption in the identification strategy is that the location of tax supply affected location demand and not the other way around; besides in case of any bias, the IV estimates would sub-estimate the effects. Therefore, this evidence brings to the table the positive effect of the Colonial legacy in Latin America; however, it also depicts that a negative effect should not be excluded.

I tested two potential mechanisms, through which the colonial institutions have a persistent effect. When examining the proximity to paved roads as a proxy for public goods provision, extraction has an adverse effect, that varies across regions and municipalities, which partially explains the differences today. The second mechanism tests the relation of access to education and history, which is closely related to public goods. As a proxy for access to education, I employed literacy rates. I found that a high density of indigenous populations in the XVIII century lowers access to education today; on the contrary, high extraction promotes high levels of access to education today. Therefore, the evidence suggest that the legacy persist through the rules of redistribution,

which have not been favorable in some regions within Mexico.

Further research should focus on both examining more precise data, other potential mechanisms such as violence, and follow them through different periods of time; this direction would close the gap between a seemingly deterministic past and actions for changing the present situation in Mexico.

References

Acemoglu, D. and Dell, M. (2010). Productivity differences between and within countries. *American Economic Journal: Macroeconomics*, 2(1):169–188.

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

Acemoglu, D., Johnson, S., and Robinson, J. A. (2002). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *Quarterly Journal of Economics*, 117(4):1231–1294.

Banerjee, A. and Duflo, E. (2014). Under the thumb of history? political institutions and the scope for action. *Annual Review of Economics*, 6:951–971.

Banerjee, A. and Iyer, L. (2005). History, institutions and economic performance: the legacy of colonial land in india. *American Economic Review*, 95(4):1190–1213.

Batchelder, R. W. and Sanchez, N. (2013). The encomienda and the optimizing imperialist: and interpretation of spanish imperialism in the americas. *Public Choice*, 156:45–60.

Beaman, L., Duflo, E., Pande, R., and Topalova, P. (2012). Female leadership raises aspiration and educational ataintment for girls: A policy experiment in india. *Science*, 335:581–86.

Commons, A. (2002). *Geografía histórica de México en el siglo XVIII : análisis del Theatro americano*. Universidad Nacional Autónoma de México, Instituto de Geografía, México.

Commons, A. C. (1993). *Las Intendencias de la Nueva España*. Universidad Nacional Autónoma de Méxic, Mexico City.

CONABIO (2017). Portal de geoinformación: Sistema nacional sobre biodiversidad.

CONAPO (2017). Consejo nacional de población: México en cifras.

Dell, M. (2010). The persistent effect of peru's mining mita. *Econometrica*, 78(2010):1863–1903.

Dell, M. (2012). Path dependece in development: Evidence from the mexican revolution. *Unpublished manuscript. Harvard University., Cambridge, MA*.

Engerman, S. and Sokoloff, K. (1997). Factor endowmnets, institutions and differential paths of growth among new wordl economies. In Haber, S., editor, *How Latin American Fell Behind, ed. by S. Haber*, pages 260–304. Stanford: Stanford University Press.

García, B. (2008). *Nueva Historia Mínima de México Ilustrada*, chapter La Época Colonial hasta 1750. El Colegio de México, México.

Gerhard, P. (1979). *The southeast frontier of New Spain*. University Press, Princeton.

Gerhard, P. (1982). *The North Frontier of New Spain*. Princeton University, Princeton.

Gerhard, P. (1986). *Geografía histórica de la Nueva España 1519-1821*. Instituto de Investigaciones Históricas, Universidad Nacional Autónoma de México, México.

Gibson, C. (1967). *Los aztecas bajo el dominio español (1519-1810)*. México: Siglo Veintiuno, Mexico.

Hauseberger (2010). *Historia Económica de México*, chapter The colony. El Colegio de México, México.

INEGI (2017). Instituto nacional de estadística y geografría.

Jauregui, L. (1998). *La real hacienda de Nueva España: su administración en la época de los Intendentes, 1786-1821*. México: Universidad Nacional Autónoma de México, Facultad de Economía.

Klein, H. S. (1985). La economía de la nueva españa, 1680-1809: un análisis a partir de las cajas reales. *Historia Mexicana*, 34(4):561–609.

Marichal, C. (2003). Los orígenes del sitema fiscal en méxico: del imperio azteca al imperio español, siglos xvi-xviii. In *Crónica Gráfica de los impuestos en México, siglos XVI-XX, Sistema de Administración Tributaria*, pages 17–50. Sistema de Administración Tributaria, México.

Marichal, C. (2005). El sistema fiscal del méxico colonial. In y Luis Jáuregui (coords), A., editor, *Penuria sin fin: historia de los impuestos en México siglos XVIII-XX*. Instituto Mora, México, 37-78 edition.

Marichal, C. and Marino, D. (2001). *De colonia a nación: impuestos y política en México, 1750-1860*. El Colegio de México, México.

Marino, D. (2001). *De colonia a nación: impuestos y política en México, 1750-1860*, chapter El afán de recaudar y la dificultad en reformar. El tríbuto indígena en la Nueva España tardocolonial. El Colegio de México, México.

Miguel, E., Satyanath, S., and Sergenti, E. (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of Political Economy*, 112(4):723–753.

Miranda, J. (2005). *El tributo índigena en la Nueva España durante el siglo XVI*. El Colegio de México, México, 2nd ed. edition.

Molina, A. (2004). Tributos y calamidades en el centro de la nueva españa, 1727-1762. los límites del impuesto justo. *Historia Mexicana*, 54(1):15–57.

Nunn, N. (2008a). The long-term effects of africa's slave trade. *Quaterly Journal of Economics*, 123(1):139–176.

Nunn, N. (2008b). Slavery, inequality, and economic development in the americas: an examination of slavery, inequality, and economic development in the americas: An examination of the engerman-sokoloff hypothesis. In *Institutions and Economic Performance*, pages 148–180. Harvard University Press.

Nunn, N. (2009). The importance of history for economic development. *The Annual Reviewof Economics*, 1:65–92.

Nunn, N. and Puga, D. (2012). Ruggedness: The blessing of bad geography in africa. *Review of Economics and Statistics.*, 94(1):20–36.

Nunn, N. and Wantchekon, L. (2011). The slave trade and the origins of mistrust in africa. *American Economic Review*, 101(7).

OIDH (2017). Oficina de investigación en desarrollo humano.

Porta, R. L., de Salinas, F. L., and Shleifer, A. (2008). The economic consequences of legal origins. *Journal of Economic Literature*, 46.

Rivarola, J. B. (2005). *La Real Hacienda*. Ediciones y Arte, Paraguay.

Robinson, J. A. (2013). Why regions fail: The mexican case. In *Harvard Presentations*.

Rodrick, D., Subramanian, A., and Trebbi, F. (2004). Institutions rule: The primacy of institutions over geography and integration in economic development,. *Journal of Economic Growth*, 9(2).

TePaske, J. J. and Klein, H. S. (1986). *Ingresos y egresos de la Real Haciend de Nueva españa*. Instituto Nacional de Antropología e Historia, México.

Vazquez, G. (2006). José de escandón y las nuevas poblaciones del nuevo santander. *Script Nova*, 10(218).

Yeager, T. J. (1995). Encomienda or slavery? the spanish crown's choice of labor organization in sixtenth-century spanish america. *The Journal of Economic History*, 55(4):842–859.

Zavala, S. A. (1973). *La encomienda Indiana*. Porrúa, México.