Essays in Economic Development

by

Alex Pierre Rivadeneira Acosta

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2019 by the
Graduate Supervisory Committee:

Gustavo Ventura, Chair
Todd Schoellman,
Galina Vereshchagina

ARIZONA STATE UNIVERSITY

May 2019

ABSTRACT

This dissertation consists in two chapters. In the first chapter I collected and digitized historical tax records from the Spanish colonial regime in Ecuador to estimate the long-run effects of a forced labor institution called *concertaje* on today's economic performance. This institution allowed landlords to retain indigenous workers due to unpaid debts, and forced them to work as peasants in rural estates known as haciendas. In order to identify the causal effects of *concertaje*, I exploit variation in its intensity caused by differences in labor requirements from the crops a region could grow. I first report that an increase in 10 percentage points in *concertaje* rates is associated with a 6 percentage points increase in contemporary poverty. I then explore several channels of persistence. Districts with higher *concertaje* rates have been historically associated with higher illiteracy rates, lower school enrollment, and populations with fewer years of education. I also report that *concertaje* is associated with a higher fraction of people working nowadays in the agricultural sector.

In the second chapter I use administrative data on the ownership, management, and taxes for the universe of all firms in Ecuador to study the role of family management in firm dynamics and its implications for aggregate productivity. A novel finding I document is that family-managed firms grow half as quickly as externally-managed firms. This growth differential implies that family-managed firms account for half of employment, despite comprising 80% of firms. I construct a general equilibrium model of firm dynamics that is consistent with these facts. Entrepreneurs choose whether to utilize family members as managers or hire external managers. External managers allow firms to scale up production, but their efficiency is a affected due to contractual frictions. Changes in the contractual environment that lead to a drop in the presence of family-managed firms by half could increase output on the order of 6%, as firms that abandon family management enjoy rapid growth.

DEDICATION

To my parents, who made all of this possible. To my sister, for always believing in me. And to Andrea, my great support over all these years.

ACKNOWLEDGMENTS

I am deeply grateful to my advisors Todd Schoellman, Gustavo Ventura and Galina Vereshchagina for extremely valuable comments and extraordinary support through the course of this project. This work also benefited from helpful suggestions and discussions with participants at seminars at Arizona State University, in particular Francesco Agostinelli, Alex Bick, Domenico Ferraro, Berthold Herrendorf and Bart Hobijn. I also thank the Servicio de Rentas Internas (SRI) and the Superintendencia de Companias Valores y Seguros (SCVS) for kindly helping me by providing access to their data. All errors are mine.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

Chapter 1

ATTACHED ONCE, ATTACHED FOREVER: THE PERSISTENT EFFECTS OF CONCERTAJE IN ECUADOR

1.1 Introduction

A major, yet open question in economics is why some regions are poorer than others. Although no consensus has emerged, a recent literature has focused on economic institutions as a key element in explaining contemporary development (Acemoglu et al. (2005)). Notably, this research agenda has placed special attention to the role of historical institutions in shaping modern ones and consequently, economic development. Explaining development through the lens of this theory demands to answer why some regions developed different institutions than others, but more important, to unfold the mechanisms by which they shaped long-run economic growth.

However, considering that the set of institutions that promote growth is potentially ample, it is difficult to point out directly what are the most immediate mechanisms that explain its long-run and persistent effects. Moreover, the identification of those mechanisms is challenging in a cross-country setting due to presence of several confounding factors. Nevertheless, the variation of historical institutions within a country could potentially address this difficulty by creating a more comparable setup to identify both its effects on contemporaneous development as well as its persistence channels.

Along these lines, this paper focuses on a specific forced labor institution created by the Spanish in the Ecuadorian highlands named *concertaje*. This institution forced indigenous workers to provide labor to landlords to pay off past debts. Because

indebted workers were usually unable to pay their obligations, this situation implied a lifetime of service to the rural estates (*haciendas*) where they worked. Additionally, this condition was transmitted over generations via inherited debt. These mechanisms allow the institution to persist for a long period of time. *Concertaje* was legally abolished in 1918, but survived through other forms of forced-labor relations until 1964.

Although I focus on the Ecuadorian case, *concertaje* is not very different from other types of forced labor relations in Latin American countries[1]. Furthermore, forced labor has been one of the most common labor relations through history (Acemoglu and Wolitzky (2011)). However, empirical evidence about the effects of forced labor systems is scarce due to data limitations. I circumvent this challenge by collecting and digitizing more than 2000 pages of historical tax records from the Spanish colonial regime around 1800 to construct a unique dataset of free and *hacienda* workers, covering nearly all the colonial districts in the Ecuadorian highlands where the institution existed.

In order to assess the effects of *concertaje* on contemporary development, I exploit the fact that its intensity, measured by the rate of indigenous labor permanently attached to *haciendas*, varied across colonial districts. I identify a causal relation guided by the Ecuadorian historiography and the well known hypothesis that natural endowments influenced the formation of forced labor institutions (Engerman and Sokoloff (1997)). Due to the peculiar characteristics of the Ecuadorian highlands, there are non-negligible variations both in crop suitability and productivity even

[1]The *conciertos* of Ecuador are comparable to *terrazgueros* in Colombia, *inquilinos* in Chile, *yanacunas* in Peru or *colonos* in Bolivia (Oberem (1978)). The system of getting indebt workers during agricultural activities was also common in Mexico and known as "tienda de raya" (Villegas (2008)).

within short distances. These variations in turn produced different incentives in landlords to coerce labor.

Data available from the GAEZ-FAO project allow me to compute for each district the relative productivity of crops with different labor requirements and use it as an exogenous variation of *concertaje*. Crucially, crop productivity is computed from estimates of potential yields based on agro-climatic, instead of agro-ecological conditions. Those agro-climatic conditions are orthogonal to human intervention. Specifically, I use the relative productivity of maize, a labor-intensive crop characteristic of the Ecuadorean highlands, to potatoes, as an instrument of the intensity of *concertaje*. Both are among the most common crops of the region. I document that on average, regions where potatoes were relatively more productive to grow than maize coerced less indigenous labor. Further details of the instrument are discussed below[2].

Using contemporary census data, under the instrumental variable framework I find that an increase in one standard deviation of *concertaje* around 1800 increased 1990's extreme poverty rate of a locality in 15%. These results are conditional on controlling for geographical variations, and robust to the inclusion of colonial province fixed effects, distance to capital, or the exclusion of certain regions and colonial province capitals. Moreover, I also find a negative effect of the institution on satellite night light intensity, a proxy for economic development (Henderson et al. (2012)). However, I also report that the effects of *concertaje* have reduced over time if we use extreme poverty rates as a measure of economic performance, but remain almost the same if we consider its impact on average night light intensity. Overall these results are

[2]Easterly (2007) also employs data from the GAEZ-FAO project to test the causality between inequality and economic development by using the log of suitability of land for wheat versus sugarcane as an instrument of inequality, without studying its direct effect on the formation of forced-labor institutions. Lagerlöf (2005) instead tests the effect of geographic characteristics as a determinant of slave use in the United States.

consistent with the work of Acemoglu et al. (2012) and Dell (2010) who document the negative long-run effects of other forced labor institutions in Latin America such as slavery in Colombia and Peru's mining *mita*, respectively, and with the work of Banerjee and Iyer (2005), who report a negative association between landlord system and development in India.

I then disentangle the mechanisms behind the detrimental effects of *concertaje* by first understanding why the institution itself persisted for so long, even after colonization ended. This exercise sheds light on how the intensity of the institution shaped economic incentives across districts. I document that around 1800 *concertaje* was strongly associated with average size of *haciendas* (measured by employment). That is, places where *concertaje* was higher also exhibited, in average, larger *haciendas*, implying agriculture was a profitable activity. Furthermore, a hundred of years later, the total value of land in places where *concertaje* prevailed was also higher, reflecting how the value of land was shaped by the number of attached workers. These two facts suggest that landlords retained workers as a business strategy to both profit and to maintain the value of their land. However, later in 1960 *concertaje* started to be associated with higher illiteracy rates and longer distance from farms to main roads. Finally, districts with higher rates of *concertaje* also report nowadays a large fraction of their workers in the agricultural sector and lower density of road networks.

This sort of reversal of fortune suggests that during the transition from an agrarian economy, *concertaje* became a major obstacle to development because it created significant barriers to human capital accumulation and labor mobility. First, the intergenerational transmission of debts gave landlords a tool to retain the offsprings of *conciertos* in order to keep a stable labor force. This forced children of *hacienda* workers to substitute schooling for early work, which is reflected in the high rates of illiteracy that predominate in districts where *hacienda* had a major presence.

Lastly, due to landlords' effort to retain workers, *concertaje* become also a barrier to labor mobility, specially from agriculture to other economic activities, generating a possible misallocation of labor. Regions with higher rates of *concertaje* induced generations of workers to specialize in agriculture and probably slowed down the process of structural transformation. This situation could have generated very persistent distortions in the labor market that in turn could have harmed economic development. I show a strong relationship between the rates of *concertaje* around 1800 and the fraction of individuals working in agriculture nowadays. This suggests that one of the most important persistence mechanisms of *concertaje* was the creation of barriers in the allocation of labor.

Studies that have investigated the historical origins of underdevelopment in the Americas have stressed the negative association between labor coercion and economic development, although through different channels. Notoriously, Engerman and Sokoloff (1997) hypothesis sees land inequality as the first-order consequence of labor coercion and the crucial channel that determined long-run economic growth. Specially because forced labor relations were seemingly more common in large plantations, or *haciendas*. The authors suggest that higher rates of inequality induced by forced labor systems hampered the formation of wider, deep markets, so they failed to create the right incentives to invest in public goods. Instead, Dell (2010) evidences a negative long-run effect of coercion from the mining *mita*, but finds a positive association between *haciendas* and development for the case of Peru, arguing that regions which did not contribute labor to the mining *mita* allowed the expansion of large landowners, which in turn had secure property rights, and so were more likely to invest and lobby for public goods.

The negative effects of coercion from *haciendas* reported in this paper states doubts about a generalization of Dell's hypothesis related to the distinct effects of

coercion coming from *mita* and *haciendas*, and rather emphasizes the general negative effects of all forms of forced labor. Moreover, the empirical results of this paper provide support to the idea that historical labor coercion, instead of land inequality by itself, is a key element in explaining contemporary development. Principally because labor coercion does not have to be necessarily accompanied with land inequality. In lieu, labor coercion was typically accompanied with education restraints that lowered the levels of human capital, and subsequent restrictions of labor mobility, which altogether provide a better explanation of its detrimental effect in the long-run.

This paper is organized as follows. The next section offers a historical background of *haciendas* and *concertaje*. Section 1.3 describes the identification strategy and the data used. Section 1.4 presents the results for the long-run impact of *concertaje* on contemporary poverty. Section 1.5 studies the mechanisms of persistence. Finally, Section 1.6 concludes.

1.2 Historical Background

1.2.1 Concertaje

The distribution of indigenous labor in the Ecuadorian highlands during colonial times was mainly organized via a pre-hispanic system known as *mita*. The *mita* was a system of forced but paid labor which assigned one-fifth of the male population between the ages of 18 and 55 to work on a specific task for a certain period of time (Cushner, 1982, p. 119). Its enforcement was mainly in charge of the local chiefs, who could lose their power or be forced to pay with their own wealth in case the solicited workers were not supplied (Cushner, 1982, p. 121). Given the absence of gold and silver mines in the Ecuadorean highlands, the *mita* was mainly distributed to landlords for agricultural and textile activities[3].

Concerns about their ability to secure a stable labor force led landlords to develop mechanisms to retain workers (Cushner, 1982, p. 128). The most important one was to force indigenous workers to get indebted with the *hacienda* during their *mita* period. Once workers were employed by the *hacienda*, landlords deducted from their wages a combination of benefits that included a small plot of land to cultivate, the guarantee to pay their tribute, and advanced provisions of food and clothes. For instance, the average wage for a worker ranged between 18-25 pesos, but tribute payments averaged 5-6 pesos and the cost of food and clothes charged by the *hacienda* an additional 5-10 pesos (Andrien, 1995, p. 122). Under this cost structure, it was not very difficult to make a worker finish the *mita* period indebted. In fact, it was more an exception, rather than a rule, that *mita* workers did not end up indebted at the end of their

[3]Deposits of precious metals were rapidly depleted during the first colonial years (Andrien, 1995, p. 18). Since many landlords combined in their *haciendas* both activities (Oberem, 1981, p. 347), in practice its distinction is not clear.

turn (Oberem, 1978, p. 314).

Except for the tribute cost, landlords had the power to adjust the extent of the benefits of their workers in order to retain them. This could be done either by increasing the provision of food and clothes or by charging higher prices for them. For instance, the parcel of land provided was small enough so its benefits could not meet any subsistence requirements, and the prices of food and clothes charged to workers exceeded those of the market (Oberem, 1978, p. 314-315; Cushner, 1982, p. 134), while the loss of livestock during shepherding was added directly to their debt with the *hacienda* (Andrien, 1995, p. 123). All of the previous strategies point out *mita* workers could be easily converted into indebted workers if the landlord decided to. Indeed, under the typical wage, workers were usually forced to extend their original contracting period with the *hacienda* mainly because they could not afford to pay their debts. An example of this is provided by Cushner (1982, p. 123-128), who documented the story of an indigenous worker in a jesuit *hacienda* who started with a debt of 27 pesos in 1701, and continued to work there for 15 more years just to end up accumulating a debt of 48 pesos by 1716.

Precisely, the process of attaching a worker to an *hacienda* created the institution named *concertaje*, a system of debt peonage that forced indigenous workers to be employed for extended periods of time. These *hacienda* workers, known as *conciertos*, were not considered slaves because landlords did not formally own them. However, the existence of debtor's prison guaranteed landlords a legal mechanism to retain them in their *haciendas*. And because *conciertos*' debt was considered an asset of the property, it allowed landlords to buy and sell them with the *hacienda* (Oberem, 1978, p. 310; Baud, 2007, p. 76). Therefore, by holding workers in their properties landlords not only gained a stable labor force, but also increased the value of their land.

1.2.2 Persistence

One of the main features of *concertaje* was its persistence. Once a worker was attached to an *hacienda*, the possibilities to escape were low. Those workers who tried to abandon the property were chased, and only in a few cases they were not recaptured. Furthermore, the intent to escape implied not only a harsh punishment, but also an increase in the amount indebted, since all the costs of prosecution were charged to the worker (Oberem, 1978, p. 316). In addition, their descendants ended up sharing the same fate since the debts of the parents were inherited by their children. Accordingly, many *haciendas* reported two or three generations of one family living and working together on the same estate (Alchon, 1991, p. 85). The difficulty to escape and the intergenerational transmission of debts attached entire generations of families to *haciendas*[4].

Persistence of the institution during colonial times can be noticed in Table 1.1, which shows the differences between attached-to-*hacienda* and free indigenous workers for two regions, Cayambe and Otavalo, in a similar exercise as Valarezo (2002). It is clear, as Valarezo notices, that even two centuries later, the percentages of indigenous workers attached to *haciendas* had barely changed in each zone: while Otavalo only had one third of its workers as *conciertos*, Cayambe had almost 80 percent of them under that condition.

The same pattern is corroborated using the data I collected. For the colonial administrations of Ibarra and Otavalo, which have data for the longest span of time, Figure 1.1 shows the evolution of *concertaje* from 1785 to 1830. Except for a few

[4]In 1833 the intergenerational transmission of debt was prohibited, however, it did not eliminate the intergenerational transmission of the *concierto* status, as in many cases children were forced to obtain a debt when young in order to guarantee their labor supply by the time they became adults (Oberem, 1978, p. 315).

Table 1.1: Free and Attached Indigenous Workers for Two Regions

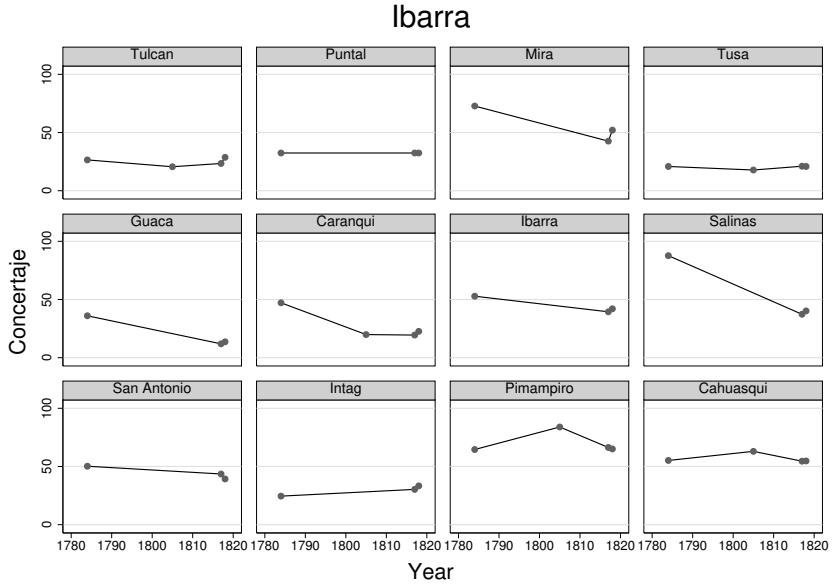
	Cayambe		Otavalo	
Year	Free	Attached to *hacienda*	Free	Attached to *haciendas*
1830	22.4%	77.6%	69.1%	30.9%
1820	25.2%	74.8%	66.5%	33.5%
1805	21.6%	78.4%	62.6%	37.4%
1785	NA	NA	56.1%	43.9%
1645/1685	20.7%	79.3%	66.3%	33.7%

Notes: Otavalo includes the colonial districts of San Pablo, Atuntaqui, Cotacachi and Otavalo, while Cayambe includes the colonial districts of Tabacundo, Tocachi, Cayambe and Cangagua. Attached to *haciendas* is the percentage of workers living in *haciendas*. Data for 1645/1685 is from Valarezo (2002). See Appendix A for a detailed explanation of the source of the rest of the data.

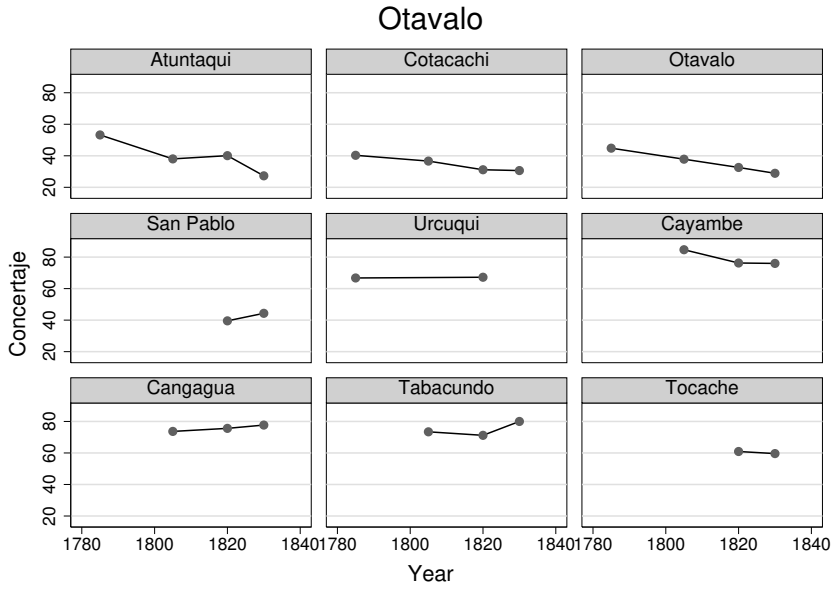
districts, the fraction of workers attached to *haciendas* is relatively stable[5]. Also, for those towns where we can observe its *concertaje* rates for at least two periods, Figure 1.2 shows the correlation between *concertaje* in period t with its immediate predecessor, in $t-1$. With the exception of a few outliers, the correlation is very strong, altogether suggesting that the institution was very persistent.

Data limitations do not allow us to observe the evolution of *concertaje* after Ecuador's independence, but the system persisted until the twentieth century despite its several reforms. For instance, in 1918 *concertaje* was legally abolished via the

[5]Mira, Pimampiro and Salinas reported to have a significant amount of slaves in the census of 1782, which might explain its substitution with *concertaje* (Salmoral (1994)).



(a)



(b)

Figure 1.1: *Concertaje* Rates for Different Years

Notes: See Appendix A for a detailed explanation of the source of the *concertaje* data. The town of Puntal after 1784 includes the information of the towns Puntal and El Angel.

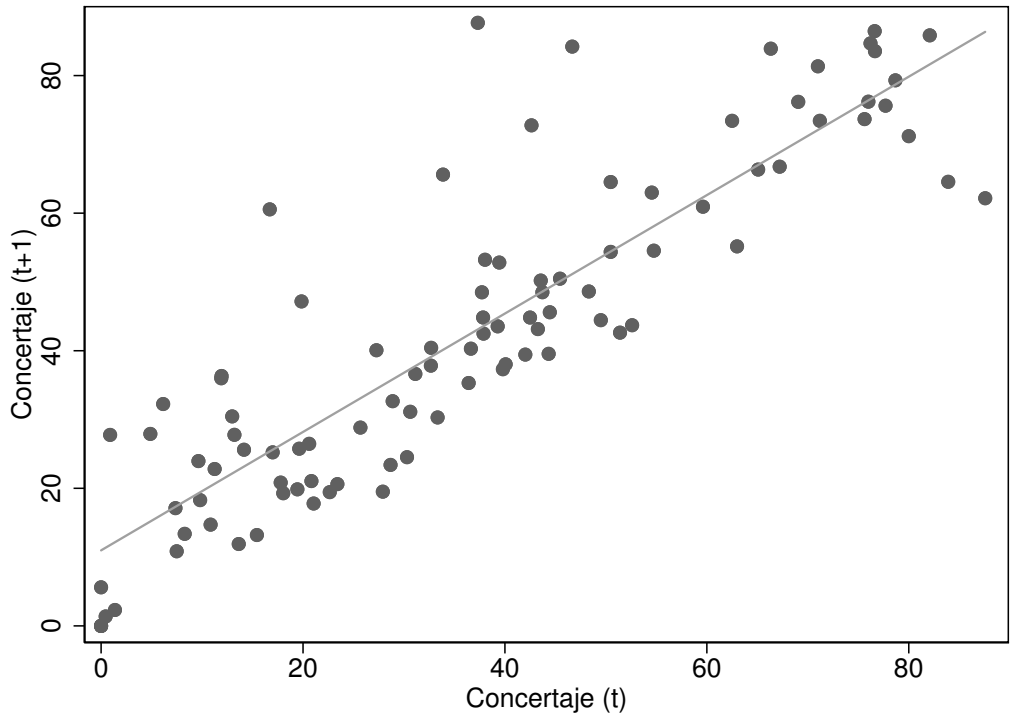


Figure 1.2: Correlation Between *Concertaje* in t and *Concertaje* in $t+1$

Notes: Sample includes only those towns where Is possible to observe *concertaje* rates for at least two periods. See Appendix A for a detailed explanation of the source of the *concertaje* data.

prohibition of prisons' debt and the impediment of inheritance of a parent's debts. However, once that *de jure* mechanism to retain indigenous workers disappeared, landlords continued to coerce them *de facto* by giving them the right to farm a small plot of land in an *hacienda* in return of labor obligations, a variation of *concertaje* known as *huasipungo* (Oberem, 1978, p. 321-322).

Precisely, one way to see the persistence of *concertaje* is by comparing its presence during colonial times with information about the *huasipungo* system 160 years later. Figure 1.3 shows the correlation between the distribution of *concierto* and *huasipungo* workers across several provinces. As can be seen, the provinces that had the greater concentration of *conciertos* circa 1805 continue to show the larger presence

of *huasipungo* workers in 1959. Ecuador's rural workers just witnessed the disappearance of any residuals of *concertaje* in 1964, when the agrarian reform abolished all forms of labor coercion. Saying that, it is possible to conclude that *concertaje* was one of the most persistent economic institutions in the Ecuadorian highlands.

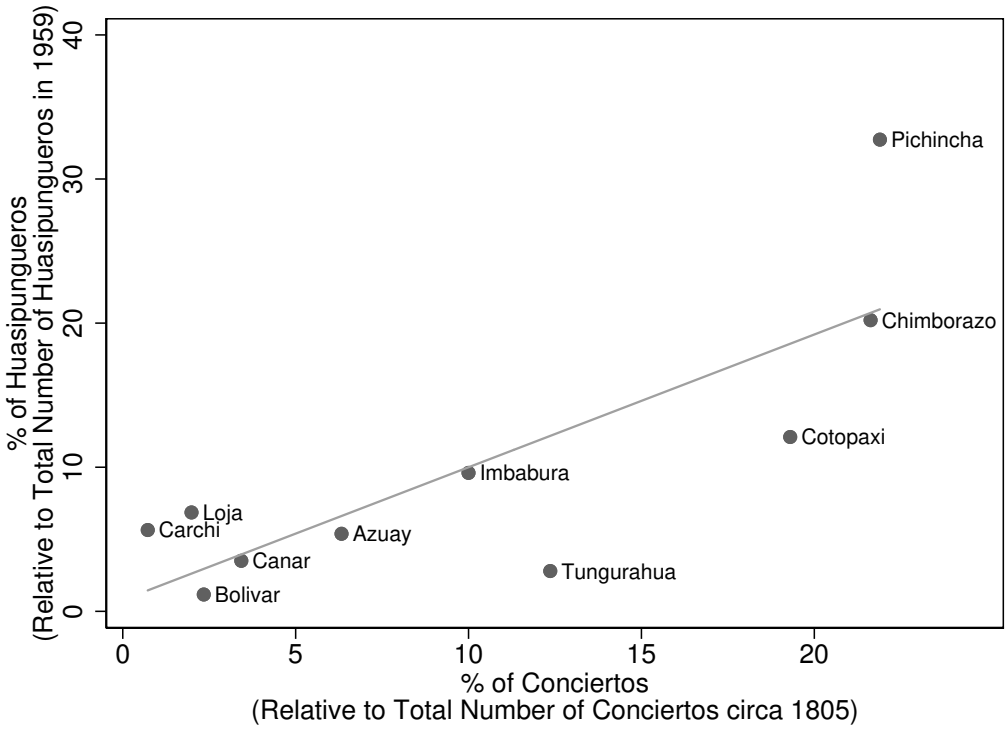


Figure 1.3: Relation Between *Concertaje* and *Huasipungo* System

Notes: Data about conciertos *circa* 1805 is from Oberem (1981), except for the provinces of Azuay and Canar, that comes from the historical tax records of the colonial province of Cuenca, described in Appendix A. Data about huasupingueros in 1959 comes from Barsky (1984) (p.73). The match between historical and modern provinces has been done following the criteria explained in Appendix A.

1.3 Labor Requirements and Coercion

In order to study the effect of *concertaje* on long-run economic development, I use as an instrument for *concertaje* the relative productivity of crops that differ on its

labor intensity. Historical analysis suggest that business considerations determined the decision of landlords to coerce labor, since debt peonage was also costly (Andrien, 1995, p. 124). However this hypothesis has not been formally tested. In the absence of profitable activities other than agriculture, the demand for labor was likely to be directed by the types of crops a region could grow. Anecdotical evidence suggests that was the case. For instance, the book records of a Jesuit *hacienda* located in the town of San Antonio report a high number of indebted workers, as well as a large production of maize, a labor intensive crop. Conversely, a Jesuit *hacienda* in the town of Santa Rosa failed to retain indebted workers, where instead no production of maize is mentioned (Andrien, 1995, p. 124). This is indicative of the economic motives behind labor coercion, as the same owners show a different behavior depending on the natural endowment of the region.

The previous historical observations are also consistent with Engerman and Sokoloff (1997) hypothesis that initial factor endowments determined the development path of a region, although their theory pays special attention to large scale plantations as sugarcane. However, the evidence discussed above suggests that not only did large scale plantations induced the formation of forced labor institutions, but in general, plantations with high labor requirements contributed to its existence, a point also made by Earle (1992).

Therefore, it is expected that *haciendas* located in regions with natural endowments for crops that require relatively fewer days of attention or are less labor-intense, such as potatoes or wheat, might not have attached many indigenous workers. On the other side, regions suitable for crops with high labor requirements such as sugarcane or maize might have experienced higher rates of *concertaje*[6]. In the next section I

[6]Sugarcane is widely known for being a labor-intensive crop, as it was typical of slave plantations, while maize requires extensive labor demand, specially during the three to four months of

discuss how I test this hypothesis by estimating the relationship between the relative potential yield of crops with different labor requirements and *concertaje*.

1.3.1 Identification Strategy

As suggested before, local endowments seems to have influenced the decision of landlords to attach labor to their properties by forcing indigenous workers to go into debt. I test this relationship by running the following regression,

$$concertaje_d = \gamma_0 + \gamma_1 ratio + \theta \mathbf{X}_d + \nu_d \tag{1.1}$$

where $concertaje_d$ is defined as the fraction of workers in colonial town d living in *haciendas* relative to the total working population in the district *circa* 1800. The variable *ratio* is the ratio between potential yield of highland maize relative to potential yield of potatoes, which captures the relative productivity of one crop versus the other. Then $\mathbf{X}$ includes several exogenous characteristics, in particular geographical ones like elevation, longitude and latitude.

Finally, I use *ratio* as an instrument for *concertaje* to quantify its effects on contemporaneous economic outcomes by running the following regression,

$$y_d = \alpha_0 + \alpha_1 concertaje_d + \beta \mathbf{X}_d + \epsilon_d \tag{1.2}$$

where, y_d is the outcome of interest in district d. All variables are discussed in detail below.

The instrumental variable framework requires that first, the ratio is correlated with *concertaje*, and second, that the ratio must be uncorrelated with other variables affecting our outcome of interest. While the first assumption is testable via the F

the plowing and tillage period (Andrien, 1995, p. 124), and had an extended crop cycle that lasted about ten months (Cushner, 1982, p. 73).

statistic in the first stage, the second is not. For instance, geographic factors that influence the suitability of certain crops, like elevation, might also have a direct effect on economic development. Therefore, the assumption for this case requires that, conditional on other exogenous factors, the suitability of two crops is uncorrelated with variables affecting the studied outcome. Consequently, I include as many geographical exogenous controls as possible to separate for any confounding effect of geography different from the ratio of potential yields. The section below discusses thoroughly the exogeneity of the proposed instrument.

1.3.2 Data

Crops

Due to the presence of the Andes Mountains and the location of the region near the Equator, the Ecuadorian highlands have non-negligible climate variations within short distances (Basile, 1974, p. 19). Those variations are translated into important differences of potential crops yields. I exploit those variations to predict the distinct agricultural labor requirements of a district by using spatial data reported by the GAEZ-FAO project. The GAEZ project provides estimates of potential crop yield (ton per hectare) at a grid resolution of $5' \times 5'$ (approximately 100 km^2). Yields of each crop could be computed for low, medium and high inputs. Moreover, one can select the source of water supply between rain-fed and irrigation. In the interest of replicating historical conditions, I selected a combination of low inputs and rain-fed water supply[7].

One important feature of the GAEZ project is that it reports potential (rather than actual) yields based on agro-climatic conditions, which, as discussed by Galor and Ozak (2016), are arguably exogenous to human intervention. None of the es-

[7]Results are robust to the selection of intermediate input and irrigation.

timates of potential yield reported by the GAEZ project use information on actual or historical production, but instead, they are based on very detailed models that incorporate agro-climatic conditions to predict the potential yield of a crop. Climate characteristics taken into account include, among others, temperature, radiation and moisture regimes. Further explanation of these data can be found in Nunn and Qian (2011) and Costinot and Donaldson (2016).

While using the GAEZ data to replicate historical conditions, an implicit assumption is that those conditions have not changed significantly in the last 300 years. The use of agro-climatic conditions instead of agro-ecological ones mitigates this concern as it is plausible to argue that the first have not changed drastically in the last centuries. Even if those conditions have somehow been modified, as long as they have affected more or less equally all crops, those concerns should be minor as I focus on the relative productivity of crops, instead of the absolute one. Therefore I assume that any effect of climate change has not been big enough to reverse the comparative advantage of one crop to another.

The crops I focus in this paper, and reported in the GAEZ project, are maize and potatoes[8], which represented some of the most important crops of the region (Cushner, 1982, p. 72). In order to compute the relative productivity of two crops in a district, I first compute the average potential yield of each crop using geospatial software. Formally, the relative productivity of two crops in a region d is,

$$z_d = \frac{\bar{y}_d^i}{\bar{y}_d^j}, i \neq j, i \in \{maize\}, j \in \{potato\} \tag{1.3}$$

where, $\bar{y}_d^i$ is the average potential yield of crop i in region d[9]. For the results pre-

[8]In particular, I use highland maize and the corresponding variation of maize that grows in the region.

[9]To avoid problems relative to places where a crop is not suitable at all, I follow a strategy similar to Easterly (2007) and add a constant (1) to the average potential yields and finally divide

sented below, I focus on the ratio of maize relative to potatoes, although results are similar if I use wheat instead of potatoes. Additional technical details are discussed in Appendix A.

Relative agro-climatic productivity is plausibly exogenous, however, one concern of using this instrument is that it might be capturing other effect instead of the institution itself. For instance, differences in crop suitability might reflect instead differences in caloric diet, which might have a direct effect on workers' productivity. While possible, this statement requires that one crop substantially outperforms the other in caloric terms. But maize and potato are both high-calorie food sources (US Department of Agriculture (2016)[10]). Moreover, the ubiquity of both crops in the Ecuadorian highlands allowed most regions to produce both crops, although in some places it was more suitable to produce one of them.

Differences in crop productivity may also have created differences in population, which could have lead to differences in economic organization. Figure 1.4 shows the relationship between tributary population around 1600 and the proposed instrument, namely, Maize to Potato productivity. As one can notice, there is no clear relation between the relative productivity of both crops and the size of tributary population at the beginning of colonial times. Although this result should be taken with caution, since early data for population is of low quality (Tabla Ducasse (1980)), it nonetheless is indicative of the initial conditions of the region.

A final concern is that differences in relative productivity between crops may be capturing differences in other geographic features that might have a direct impact on economic development. Hence, in order to distinguish for any other confounding fac-

one by the other to obtain a ratio.

[10]Using wheat instead potatoes, which has almost the same amount of calories than maize, produce very similar results.

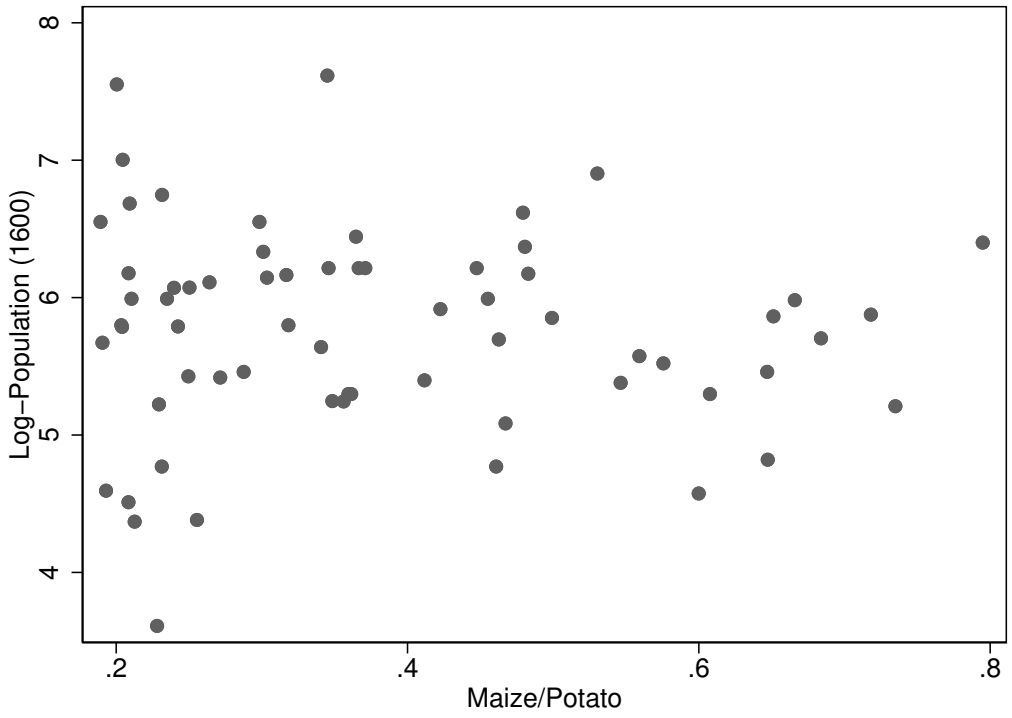


Figure 1.4: Population (1600) vs Maize to Potato Productivity.

Notes: See Appendix A for a detailed explanation of the source of the population data. Maize to Potato is computed as explained in Equation 1.3.

tor, I control for a comprehensive set of exogenous variables, which include elevation, latitude and longitude. All geographic variables are discussed below.

Concertaje

The data of *concertaje* come mainly from colonial tax records. Spanish authorities typically collected taxes by visiting each town twice a year and reported the individual taxes paid by every indigenous worker from 18 to 55 years old[11] (Villegas (2008)). Each of these records were filed in separate books organized by tax administrations or

[11]Taxes were paid in "San Juan" (June) and "Navidad" (December). Women, people with disabilities and the elderly were exempted of tribute.

colonial provinces. Some of those books are lost, but others are located in the National Archive of History in Ecuador. I collected at least one tax record from 9 of the 10 tax administrations that comprised the Ecuadorian Highlands. These records correspond to the years 1780-1830, the last period of the colonial administration and the first years after independence, when *concertaje* was already a consolidated system[12].

Every book clearly distinguishes between free workers and those living in *haciendas*, and they were usually organized as follows. First, authorities reported a list of the names of free workers and then, described the workforce of each *hacienda*. Hence, records show for each *hacienda* in the town the name of the owner, the names of the indigenous workers living there and the tax paid. The tax paid did not differ between free and *hacienda* workers. I digitize those records and compute at a town level how many free and attached workers lived at the time, as well as the number of *haciendas*. Figures 1.5, 1.6 and 1.7 are examples of how the tax books look like.

Historians as Tyrer (1988) have used some of these records to report aggregate rates of *concertaje*, but never reported a full list of the rates at the town level. A notable exception is Oberem (1981), who reported a partial list of the *concertaje* rates of 1805 at a town level, but only for those with the three highest and lowest rates at each tax administration, and using a different source that seems to be lost. I included that information in my dataset to complement it and have a comprehensive database for almost all the colonial districts of the time. Poloni (2006) also reported a list of the *concertaje* rates for the tax administration of Cuenca, although he used the same source as I do.

Tyrer (1988) mentions that the quality of the data is better for earlier years, since there were stronger incentives to know the real number of workers, as well as their

[12]The region gained its independence in 1822 and became part of The Gran Colombia until 1830 when it separated and named Ecuador.

exact location. The availability of books for some years allowed me to observe if there were dramatic differences in the number of workers and rates from year to year. Overall, as discussed in Section 1.2, the *concertaje* rates are similar over time, but seem to be higher for earlier years. Some of the differences in the *concertaje* rates seems to reflect the fact that officers collected taxes seasonally, and sometimes deferred its collection for the next visit[13]. In those cases where I observe the town for several years, I select the earliest source of data, following the observation of Tyrer (1988), but results are generally robust to this selection. This is also important because some of the books overlap with the independence wars, which might have affected the tax reports. Nevertheless, in some specifications I also include as a regressor the decade from which *concertaje* information was taken, to control for any time discrepancies in the quality of data or trends. Details about the source of every book and data selection can be found in Appendix A.

Since colonial data is organized at a town level and modern outcomes are aggregated at a district (parish) level, I matched *concertaje* data with contemporary outcomes mainly by comparing the names of modern districts with those reported in the tax books. Except for very few of them, almost all modern districts retain their colonial town name. This matching process leaves a sample of 127 modern districts. Further details about the matching are discussed in Appendix A. As mentioned in the introduction, almost all the districts studied in this paper belong to the Ecuadorian highlands, since the phenomena of *concertaje* was almost exclusive to this region[14]. Figure 1.8 shows a map of the studied region, the location of each tax administration, and displays the intensity of *concertaje* in each district.

[13]Tax records show plenty of cases where an indigenous worker pays the tribute for an old period.

[14]Data from population census reported by Salmoral (1994) at the end of the colony shows that population was heavily concentrated in the highlands.

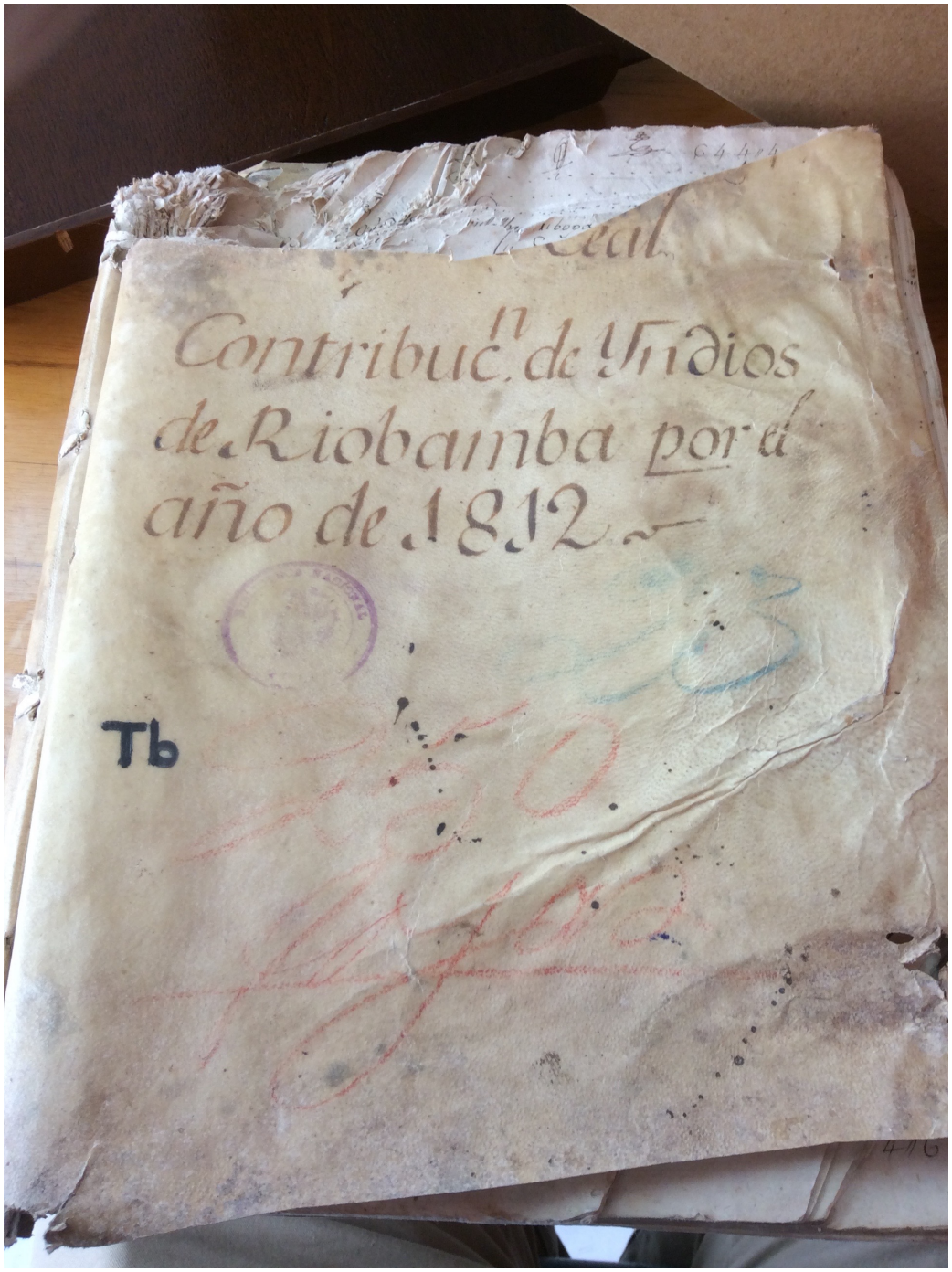


Figure 1.5: Example of the Cover of a Book of Tax Records

Notes: *Contribution of the indigenous workers from the province of Riobamba, year 1812.* See Appendix A for a detailed explanation of the source of this book.

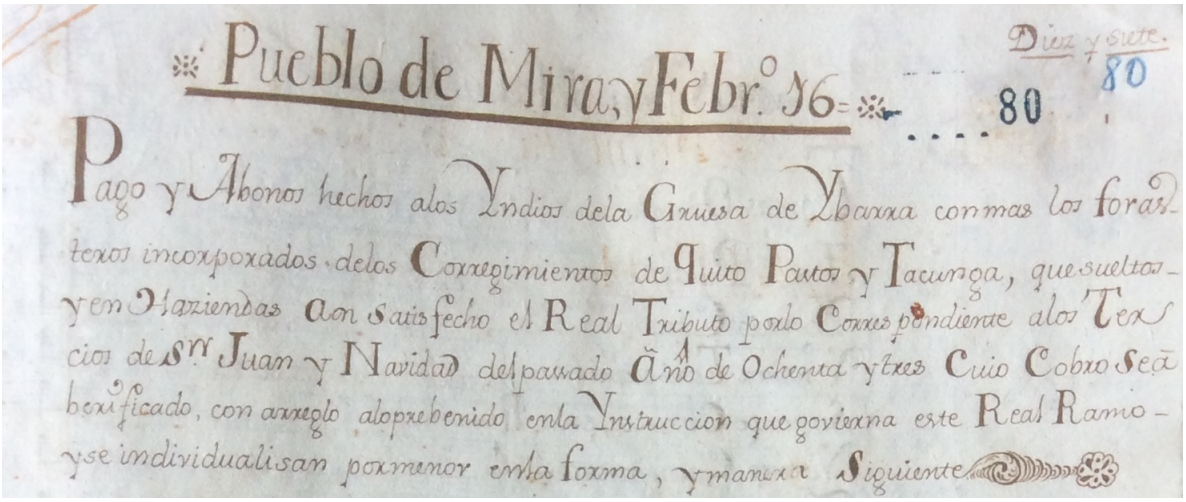


Figure 1.6: Example of the Description of the Tax Records

Notes: ... Town: *Payments by indigenous workers ...* ***free*** *and in* ***Haciendas***. From the book of Ibarra, year 1784.See Appendix A for a detailed explanation of the source of of this book.

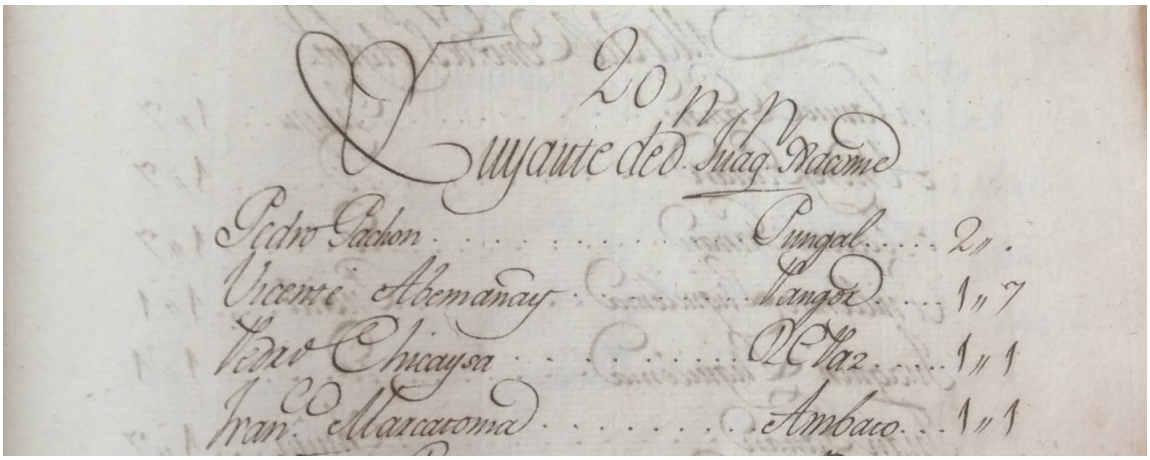


Figure 1.7: Example of the Tax Payments

Notes:*Hacienda 20 ...*: ***owned*** *by ...* with the following workers ... tax payment. From the book of Riobamba, year 1812. See Appendix A for a detailed explanation of the source of of this book.

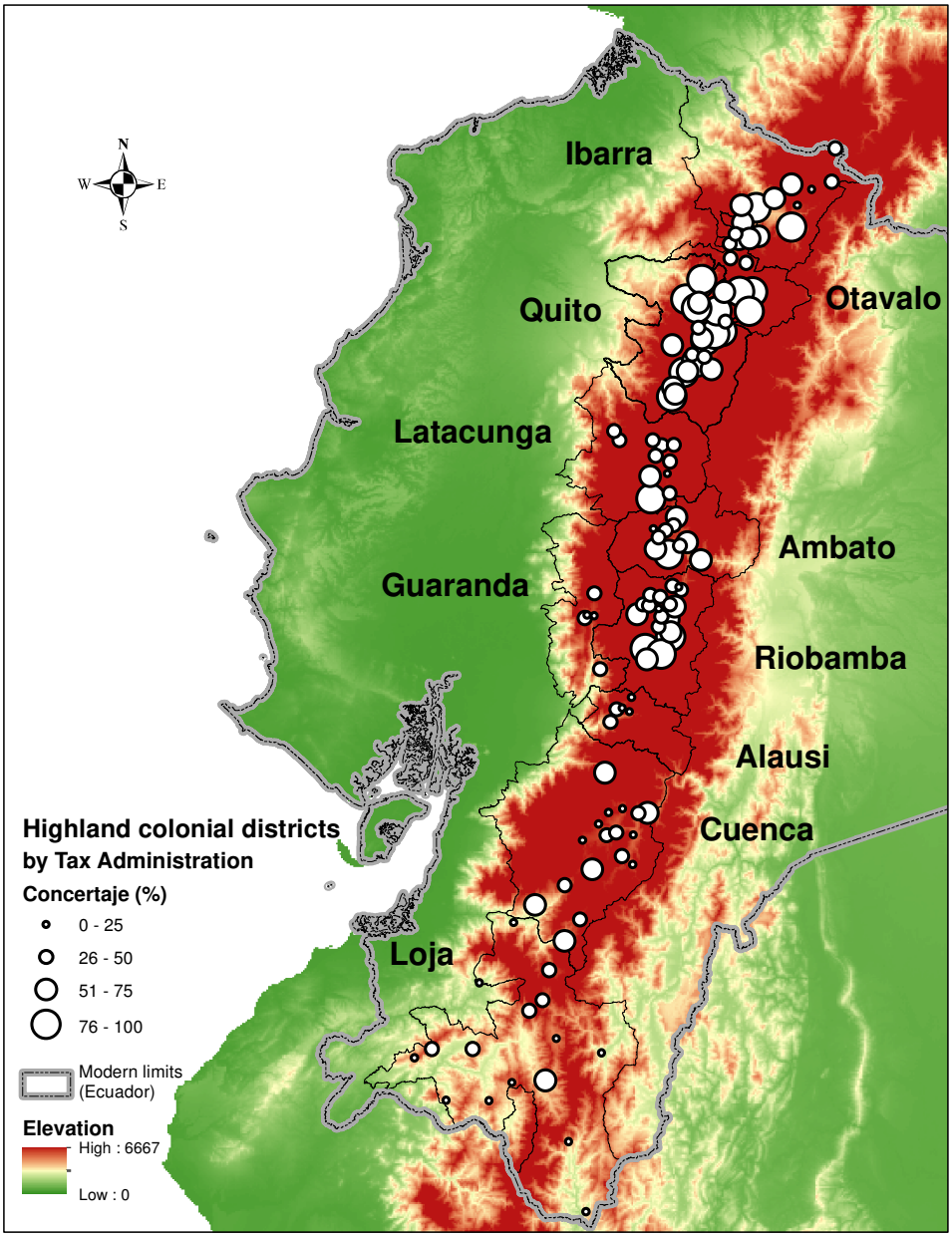


Figure 1.8: Map of the Studied Region with Its Colonial Tax Administrations

Geographic variables

Data of geographic controls are obtained by using geospatial software and GIS files reported by the Geographical Military Institute of Ecuador. For instance, the coordinates of towns (latitude and longitude) are recovered from a GIS file that contains the point-location of all modern districts in Ecuador. To compute elevation and slope, I overlay a map of Ecuador's parishes on 30 arc second (1 km) resolution elevation data produced by NASA's Shuttle Radar Topography Mission (National Aeronautics and Space Administration and the National Geospatial Intelligence Agency (2010)). Further details about geographical variables are provided in Appendix A.

Contemporary outcomes

Data for contemporaneous economic outcomes is obtained through the national Ecuadorian census, which is conducted every 10 years and includes information disaggregated at the district level. The principal variable I use to measure today's economic development at the district level is a poverty index called Unsatisfied Basic Needs (*Necesidades Basicas Insatisfechas* in Spanish) which is commonly reported in the census and used in several Latin American countries. In order to be considered as poor under that definition, a household should satisfy one of the following criteria.

1. Dirt floor or the house is made of precarious building materials
2. Lack of running water, sewers or toilets
3. Number of people per bedroom is greater than 3
4. Children between 7 and 11 years old do not attend school
5. Head of the household has less than fourth grade education and has more than three dependents

If a household satisfies two or more of the conditions above it is considered as extremely poor. The variable used is then the fraction of households in a district considered extremely poor. Other variables obtained from the census and used later to explore the channels of persistence include illiteracy, average years of school, secondary enrollment and sectoral composition. Summary statistics for all the variables discussed above are presented in Table 1.2.

Table 1.2: Summary Statistics

	Mean	S.D	p10	p90
	(1)	(2)	(3)	(4)
Concertaje	43.41	23.86	13.38	76.92
Maize/Potato	0.40	0.19	0.21	0.69
Altitude	25.73	4.67	19.06	30.84
Extreme Poverty (1990)	60.32	18.58	33.30	85.90
Night Light Intensity (1992)	4.91	7.36	0.00	14.02
Illiteracy (1960)	42.36	16.90	24.10	68.59
Functional Illiteracy (1990)	36.85	15.42	17.50	60.60
Years of School (1990)	5.48	1.95	3.00	8.30
Secondary School Enrollment (1990)	57.58	14.03	39.10	75.60

Notes: *Concertaje* is the percentage of workers living in *haciendas circa* 1800. Maize/Potato is the relative productivity of maize respect to potatoes as described in Equation (1.3). Altitude (hundreds) is the elevation in m^2 of the district's capital. Extreme Poverty (1990) is the percentage of population living in Extreme Poverty as defined in Section 1.3.2. Illiteracy (1990) is the percentage of population above 15 years who can't write or read. Functional Illiteracy (1990) is the percentage of population above 15 years with 3 or less years of education. Years of School (1990) is the average years of education of the population above 24 years. Secondary School Enrollment (1990) is the population between 12 and 17 years assisting to secondary school. Night Light Intensity (1992) is the average night light intensity of a district. Sources for all variables are described in the text.

1.4 Results

Table 1.3 shows both the OLS and IV estimates of equation (1.2) using as a dependent variable extreme poverty rates in 1990. First, one can notice that the F statistic reported in Panel B ranges between 11-20, depending on the specification. These results suggest a solid first stage, reducing concerns about a weak instrument[15].

Table 1.3: IV Results: Extreme Poverty (1990)

	Base Model	Time fixed effect	Without Southern Districts	Province fixed effect	Distance to Capital	No Capitals	Fraction Indigenous (1800)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Concertaje	.63***	.74***	.52***	.54**	.41**	.56***	.66***
	(.22)	(.25)	(.18)	(.28)	(.18)	(.21)	(.24)
Panel B: First Stage							
Maize/Potato	48.55***	47.07***	58.86***	36.11***	48.68***	47.57***	46.86***
	(11.74)	(12.16)	(12.96)	(10.8)	(12.59)	(12.05)	(12.29)
F	17.12	14.98	20.62	11.18	14.95	15.57	14.55
Panel C: OLS							
Concertaje	.09	.12	.1	.23***	.07	.05	.09
	(.08)	(.07)	(.08)	(.08)	(.07)	(.07)	(.09)
Geo. Controls	Yes	Yes	Yes	No	Yes	Yes	Yes
Time Controls	No	Yes	No	No	No	No	No
Province fixed effect	No	No	No	Yes	No	No	No
Observations	126	126	110	126	126	119	112
Mean Dep. Var.	60.32	60.32	59.7	60.32	60.32	62.05	58.92

Notes: Robust standard errors are in parentheses. Geographic controls include elevation, latitude and longitude.

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

[15]I report the F statistic from the Montiel Olea and Pflueger (2013) test, which is robust to the presence of heteroskedasticity (Pflueger and Wang (2014)).

The IV coefficient reported in the first column of Panel A suggest that the long-run effect of *concertaje* on extreme poverty is around 0.6. The result is statistically significant, while economically, it suggests that in average, an increase in 25 percentage points in the rate of *concertaje* circa 1800 (almost one standard deviation), increased extreme poverty in 1990 by 15 percentage points. To put it another way, this result implies that on average, a district that had all of its workers attached to *haciendas* around 1800 nowadays has an extreme poverty rate that is 60 percentage points higher than a district where *concertaje* was fully absent.

Column (2) includes as a control a time fixed effect for the decade data was taken (1780-1830) in order to control for discrepancies in the quality of the data due to potential changes in the collection of taxes at the end of the colonial period related with the independence struggle or simply trend effects. As can be seen, the magnitude of the effect of *concertaje* increases, although the F statistics is smaller in part because the time dummies reduce the variation of crops within provinces. Column (3) excludes the southernmost administration of Loja, which, because of its remoteness from Quito, was rarely inspected, and taxes were often ten to twenty years out of date (Newson, 1995, p. 243). Additionally, the region is the only one among the studied which reported important mining activities (Newson, 1995, p. 237). However, its exclusion does not change the results fundamentally, although improves the first stage.

Columns (4)-(7) of Panel A report the estimates when other controls are added. Overall, the point estimates for the IV coefficients remain close to the base model. Column (4) includes a colonial province fixed effect in order to capture other time-invariant unobservables such as historical regional differences, like being located in the same province as the nation's capital, or being in a bordering province, but excludes geographical controls to avoid overfitting the model. More important, since tax books

were organized by colonial province, adding a fixed effect will help to control for any systematic misreport in the level of *concertaje* at a province level. However, the inclusion of a fixed effect reduces the strength of the first stage relationship, as crop variation within some provinces is smaller.

Furthermore columns (5)-(6) report the results when I test whether being close to the capital of the province or excluding the colonial capital districts affect the results. Although the estimates are smaller, they are still statistically significant. This also suggests that the estimates are not driven by the inclusion of historically important administrative centers. Finally, column (7) controls for the fraction of indigenous population circa 1780 in order to observe if *concertaje* is actually capturing the effect of the institution instead of differences in the colonial distribution of indigenous population. The estimates reported are bigger from the previous ones, perhaps reflecting a direct effect of ethnic distribution, although note that this regression includes less observations because population data is missing for some districts.

Table 1.4 mimics the estimates presented in Table 1.3, but instead use average night light intensity of a district as a dependent variable, which has been shown as a good proxy for income. Hence, it is used as a robustness check for the long-run effects of *concertaje*. Night light intensity is obtained from US Air Force weather satellites, and is available in the National Centers for Environmental Information (National Geophysical Data Center (2010)). The data is reported as a six-bit digital number (DN) between 0 (no light) and 63, for every 30 arc-second output pixel. More details about this data and its relation with income is provided in Henderson et al. (2012). I use the satellite light data from 1992 (the earliest available), in order to make the results comparable to those presented in Table 1.3. The estimates suggest that an increase in 10 percentage points of *concertaje* in 1990 reduces average night light intensity in 2 digital numbers (DN). Interpretation of the magnitude of the effects

has to be taken with certain caution, due to the mapping between light intensity and income. However, all the IV estimates of Table 1.4 report the expected sign and are statistically significant.

Table 1.4: IV Results: Night Light Intensity (1992)

	Base Model	Time fixed effect	Without Southern Districts	Province fixed effect	Distance to Capital	No Capitals	Fraction Indigenous (1800)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Concertaje	-.2**	-.23**	-.17**	-.21**	-.13*	-.15*	-.25**
	(.09)	(.11)	(.09)	(.09)	(.08)	(.08)	(.11)
Panel B: First Stage							
Maize/Potato	48.55***	47.07***	58.86***	36.11***	48.68***	47.57***	46.86***
	(11.74)	(12.16)	(12.96)	(10.8)	(12.59)	(12.05)	(12.29)
F	17.12	14.98	20.62	11.18	14.95	15.57	14.55
Panel C: OLS							
Concertaje	-.02	-.02	-.02	-.1***	-.01	0	-.03
	(.03)	(.03)	(.03)	(.03)	(.03)	(.03)	(.03)
Geo. Controls	Yes	Yes	Yes	No	Yes	Yes	Yes
Time Controls	No	Yes	No	No	No	No	No
Province fixed effect	No	No	No	Yes	No	No	No
Observations	126	126	110	126	126	119	112
Mean Dep. Var.	4.91	4.91	5.46	4.91	4.91	4.43	5.34

Notes: Robust standard errors are in parentheses. Geographic controls include elevation, latitude and longitude.

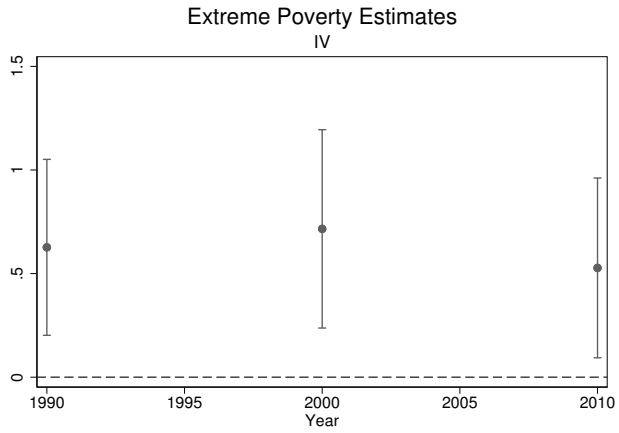
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

In all columns from Table 1.3 and 1.4, the IV estimates are larger than the OLS ones reported in Panel C. One reason why this can occur is measurement error at different levels. First, keep in mind that *concertaje* data come mainly from tax records organized at a province or tax administration level from more than 200 hundred years

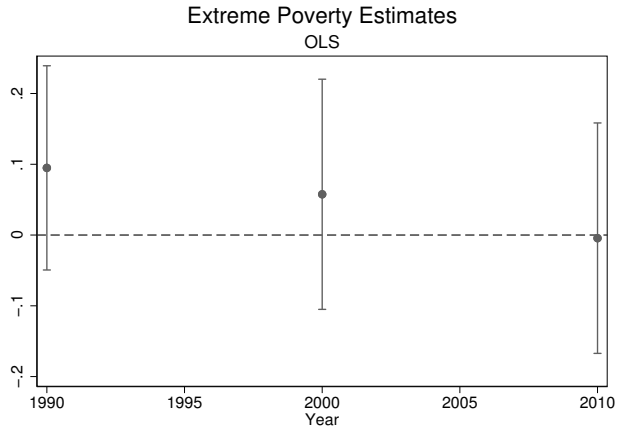
ago. Quality of the data might have differed by province, or worse, by town. Even though the Spanish had an organized system to collect taxes, nothing exempts the process from include accidental or deliberate misreports since each province had its own tax collector. One can suspect this is the case given that the OLS estimates that include province fixed effects are the largest (and statistically significant). More plausibly, by mixing data different sources, and different years, I could be introducing measurement error. It could also be the case that there are omitted variables that explain the process of *concertaje*. Although historical analysis suggests economic motives due to differences in types of crops, no ultimate explanation has been accepted.

As a robustness check, I also estimate the effects of the institution on extreme poverty and average night light intensity for different years in order to understand if the effect of *concertaje* has vanished over time. Figures 1.9 and 1.10 show the estimates of the base model for extreme poverty and average night light intensity respectively. While the effect of *concertaje* on extreme poverty appears to be reduced over time, the effect of *concertaje* on night light intensity remained almost the same. Considering the government's policies to reduce poverty in the last decades, one can expect that their efforts were more concentrated in those places with higher rates, some of them historically affected by *concertaje*. On the other side, the effect on night light intensity, as a proxy for income, shows a stable pattern, less affected by the efforts on poverty reduction.

In summary, the estimates reported in this section reveal a long-run and quantitatively important effect of *concertaje* on today's economic prosperity. These results are also consistent with the historical narrative of several scholars that have discussed the pervasive effects of *haciendas* in the rural workers of Ecuador. For instance, Oberem (1978) reports that the annual wage of a free agricultural worker, compared to that of *huasipunguero* (including the non-pecuniary benefits) was around 80% higher. To



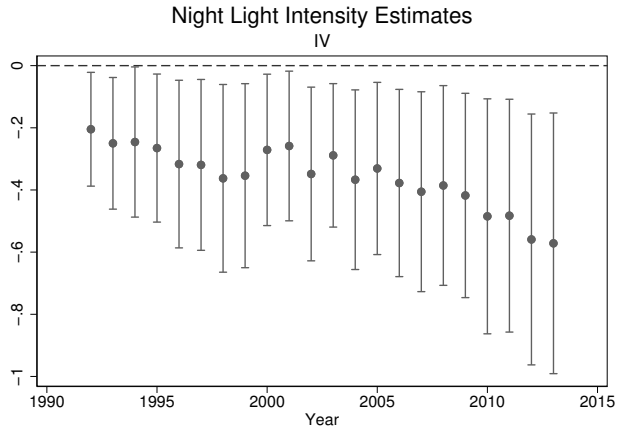
(a)



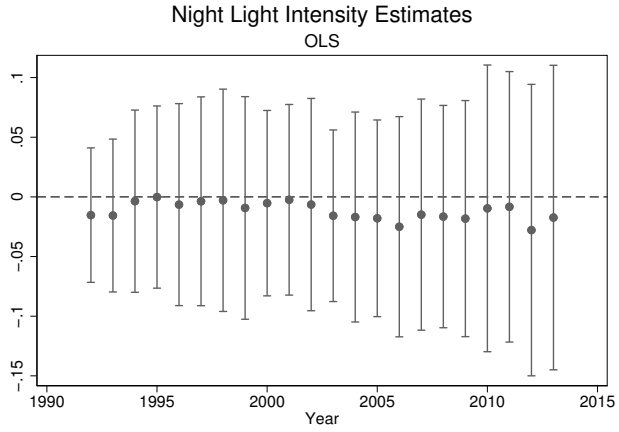
(b)

Figure 1.9: Extreme Poverty Estimates for Different Years

Notes: regressions as Base Model (1) in Table 1.3. Confidence interval: 95%. See Section 1.3.2 for definition of extreme poverty rates. Source: Author's calculations based on census data.



(a)



(b)

Figure 1.10: Night Light Intensity Estimates for Different Years

Notes: regressions as Base Model (1) in Table 1.4. Confidence interval: 95%. Source: Author's calculations based on satellite night lights data from National Geophysical Data Center (2010).

some extent, these results may not very surprising as coercing labor through forced debt or granting land usage was a common practice in the region until the agrarian reform. However, they confirm that the persistence of the institution generated effects that are reflected in the well being of today's population. In the next section I discuss the mechanisms by which *concertaje* generated its long-run effects.

1.5 Mechanisms

As discussed in Section 1.2.2, *concertaje*, and its subsequent variations, persisted for almost two hundred years greatly in part because it gave landlords the means to retain cheap labor, and to include workers as part of a valuable asset in their properties. However, given that the incentives to indebt workers were closely linked to the suitability of certain crops, not all regions developed under the shadows of this institution. In this section I study the mechanisms that explain how variations in the incidences of *concertaje* lead to a divergence in the development paths across regions.

The main argument is that the long-run economic effects of labor coercion are better understood if they are seen as a barrier to both accumulate human capital and reallocate labor from agriculture. In that sense, the negative effects of *concertaje* are explained because it was profitable for landlords to maintain the system while agriculture was the main economic activity, although by doing so they blocked and neglected the required investments (like the provision of public school) to develop other economic activities and limited workers' mobility. Precisely because the reallocation of workers would have supposed a reduction in their economic power.

For instance, Figure 1.11 documents that around 1800 *concertaje* was strongly associated with average size of *hacienda* (measured by employment). That is, places were *concertaje* was higher also exhibited, in average, larger *haciendas*, who usually are more productive (Adamopoulos and Restuccia (2014)). Furthermore, Figure 1.12

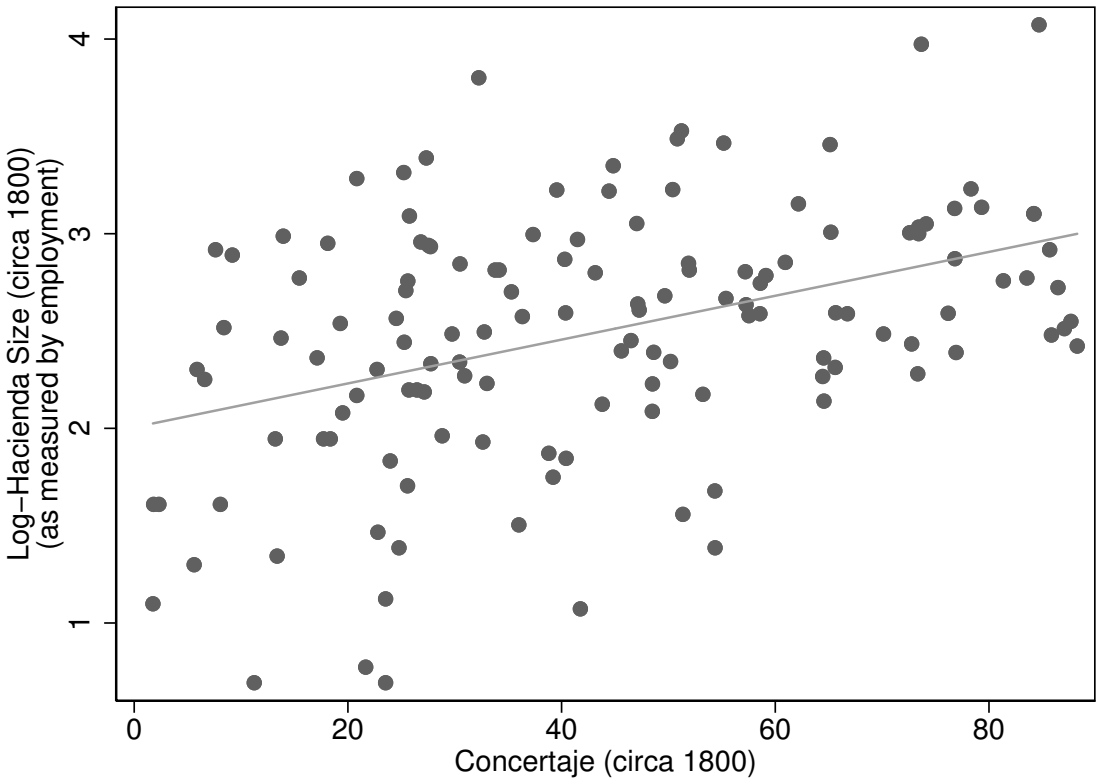


Figure 1.11: Log-hacienda Size (circa 1800) and *Concertaje*

Notes: See Appendix A for a detailed explanation of the source of the data.

shows that the total value of land in 1900 in those places with higher rates of *concertaje* was also higher. These two facts suggests not only that *concertaje* was a profitable activity for landlords, but also that regions with a high prevalence of it were likely to be more productive (as they were larger), and perhaps even richer in terms of land value. However, later in 1960 *concertaje* started to be associated with higher illiteracy rates and lower levels of educations that persist until today.

This sort of reversal of fortune suggests that during the transition from an agrarian economy, *concertaje* may have played a major distortion in the labor market, reducing the incentives to provide school and affecting labor reallocation. As landlords retained

workers in their *haciendas*, they created a barrier to accumulate human capital and reduced the incentives to provide public goods. Second, I also document that because of the barrier to mobility, most workers in regions with high prevalence of *concertaje* remain in the agricultural sector.

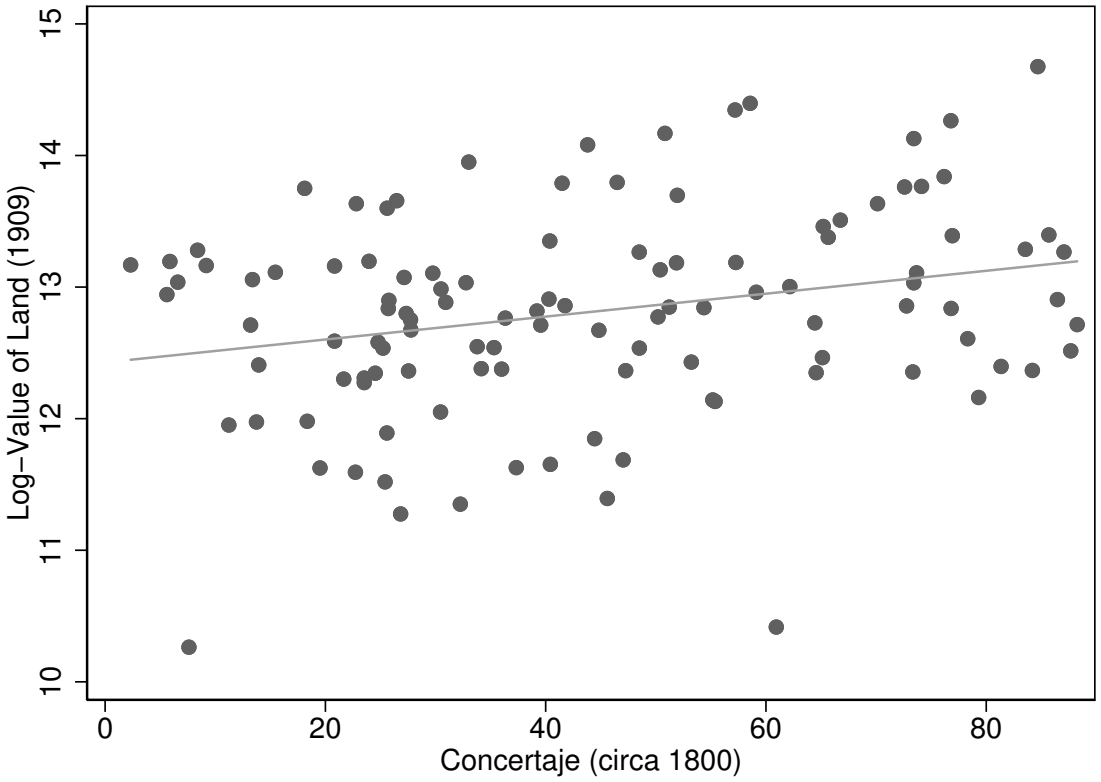


Figure 1.12: Log-value of Land (1909) and *Concertaje*

Notes: See Appendix A and A for a detailed explanation of the source of the data.

I also test other mechanisms discussed by the literature like inequality and public good provision, presenting mixed results. Although land inequality is strongly linked to *concertaje*, inequality by itself is not associated with lower levels of contemporary economic development. Hence, I interpret the results as a generalization of the overall negative effects of coercion, instead of its operation through inequality. Moreover, I

also show that public good provision, in terms of roads, seems to be negatively related to *concertaje*.

1.5.1 Human Capital Accumulation

In addition to the attached workers, their wives and sons were also part of the labor force of the *haciendas* (Oberem, 1981, p. 352). Women were usually employed as housekeepers for landlords, while sons were sent at a very early age to work. This early disposition of children to work limited their education, if any was available in the district. Accordingly, in 1934, 80% of indigenous workers attached to *haciendas* were illiterate, as opposed to the 40% of the rest of rural workers (Oberem, 1978, p. 323). In that spirit, I test as one of the potential channels of persistence of *concertaje* its effects on human capital by using the illiteracy rates, functional illiteracy rates and the average years of school of a district. Illiteracy rates come from the 1960 census, the first year for which this information available. The rest of the data come from the 1990 census. I omit the results of the first stage since they are similar or the same as those reported in Table 1.3.

The results for illiteracy are reported in Table 1.5. All of the estimates are statistically significant except when controlling for province fixed effects, and vary little for the different specifications. Economically, they suggest that in average, an increase of 10 percentage points in *concertaje* increased the illiteracy rate of a district by almost 4 percentage points in 1960. These results are considerable since they suggest that in average, a district with all of its workers attached to *haciendas* has illiteracy rates that are 40 percentage points higher than districts with no attached workers. I also estimate the effects of *concertaje* on a robust measure of illiteracy, namely, functional illiteracy, defined as the fraction of population 15 years and above with less than 3 years of primary education for the year 1990. Table 1.6 shows the results for

Table 1.5: IV Results: Illiteracy (1960)

	Base Model	Time fixed effect	Without Southern Districts	Province fixed effect	Distance to Capital	No Capitals	Fraction Indigenous (1800)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Concertaje	.39**	.48**	.44***	.26	.34*	.39**	.34*
	(.18)	(.19)	(.16)	(.18)	(.17)	(.18)	(.2)
Panel C: OLS							
Concertaje	.15**	.19**	.17**	.18**	.14*	.14*	.12
	(.07)	(.08)	(.08)	(.07)	(.07)	(.08)	(.08)
Geo. Controls	Yes	Yes	Yes	No	Yes	Yes	Yes
Time Controls	No	Yes	No	No	No	No	No
Province fixed effect	No	No	No	Yes	No	No	No
Observations	122	122	109	122	122	115	108
Mean Dep. Var.	42.36	42.36	43.77	42.36	42.36	43.11	41.57

Notes: Standard errors in parentheses. Geographic controls include elevation, latitude and longitude.

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

functional illiteracy, evidencing a similar effect than the one for illiteracy, confirming the conclusion about the negative relation of *concertaje* on education, except when dropping the capitals and accounting for being close to one.

Given that illiteracy means zero human capital accumulation via formal education, these results imply a strong relationship between institutions and human capital accumulation. The elevated rates of illiteracy observed in districts where *concertaje* had more presence, reflect both the poor incentives any individual attached to an *hacienda* had to go to school, as well as the impediment they faced due to work obligations. In that sense, the results presented previously give a quantitative relation between institutions and human capital. In particular they stress the negative effects

of labor coercion on education.

Table 1.6: IV Results: Functional Illiteracy (1990)

	Base Model	Time fixed effect	Without Southern Districts	Province fixed effect	Distance to Capital	No Capitals	Fraction Indigenous (1800)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Concertaje	.3*	.44**	.45***	.4**	.2	.26	.37*
	(.16)	(.18)	(.16)	(.19)	(.17)	(.17)	(.19)
Panel C: OLS							
Concertaje	.06	.09	.06	.19***	.05	.04	.08
	(.07)	(.07)	(.07)	(.06)	(.07)	(.07)	(.08)
Geo. Controls	Yes	Yes	Yes	No	Yes	Yes	Yes
Time Controls	No	Yes	No	No	No	No	No
Province fixed effect	No	No	No	Yes	No	No	No
Observations	126	126	110	126	126	119	112
Mean Dep. Var.	36.85	36.85	38.27	36.85	36.85	37.84	35.06

Notes: Standard errors in parentheses. Geographic controls include elevation, latitude and longitude.

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

I obtain the same negative relationship between *concertaje* and human capital if instead I use as a dependent variable years of school, like reported in Table 1.7. Overall, the results show a negative impact of the institution on the average years of school in a district. The point-estimate of *concertaje* is around -0.05 which means that an increase in 10 percentage points of *concertaje* around 1800 lead to a reduction of half a year of education in 1990. These effects are quantitatively important given that the mean of years of school in a district is around 5.5 years.

Additionally, Figures 1.13 replicates the exercise done in the previous section by estimating the effects of *concertaje* on illiteracy rates for different years, starting

Table 1.7: IV Results: Years of School (1990)

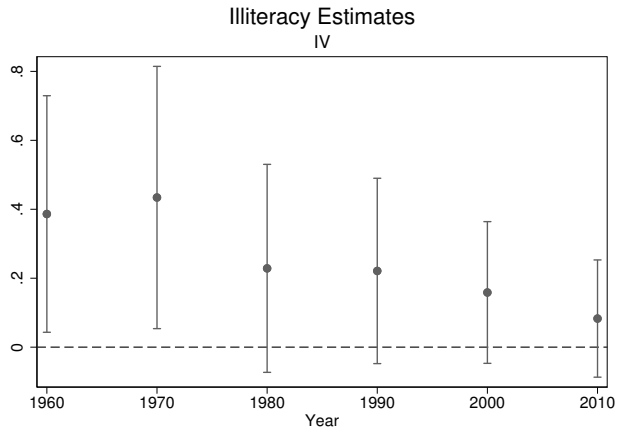
	Base Model	Time fixed effect	Without Southern Districts	Province fixed effect	Distance to Capital	No Capitals	Fraction Indigenous (1800)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Concertaje	-.04**	-.06***	-.06***	-.05**	-.03	-.04*	-.06**
	(.02)	(.02)	(.02)	(.02)	(.02)	(.02)	(.02)
Panel C: OLS							
Concertaje	-.01	-.01	-.01	-.02***	-.01	0	-.01
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
Geo. Controls	Yes	Yes	Yes	No	Yes	Yes	Yes
Time Controls	No	Yes	No	No	No	No	No
Province fixed effect	No	No	No	Yes	No	No	No
Observations	126	126	110	126	126	119	112
Mean Dep. Var.	5.48	5.48	5.3	5.48	5.48	5.3	5.66

Notes: Standard errors in parentheses. Geographic controls include elevation, latitude and longitude.

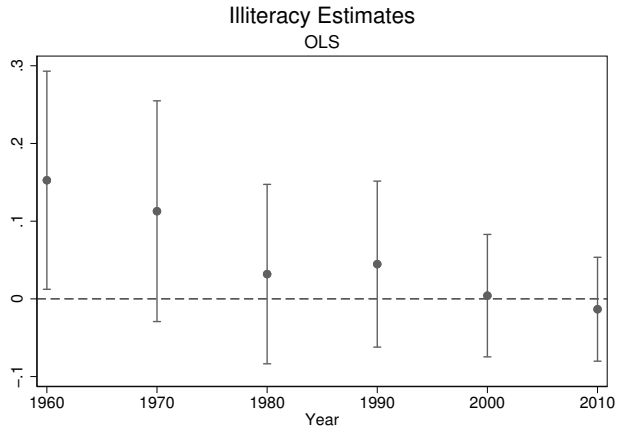
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

1960. Results have to be taken with caution, as the reference population for illiteracy differ depending on the census (see Appendix A). However, the exercise is useful to understand how the effect of *concertaje* on illiteracy has evolved over time. The results show a stronger effect of *concertaje* in the earliest decades, which decrease afterwards. For example, the estimates for 2010 are statistically indistinguishable from zero. Again, these decreasing effects could be explained by the government's efforts to abolish illiteracy. Moreover, they also suggest that the effects of *concertaje* may no longer be reflected in illiteracy rates, but in years of school, as reported in Figure 1.14.

The concerns about the low education of workers in *haciendas* is not new, as



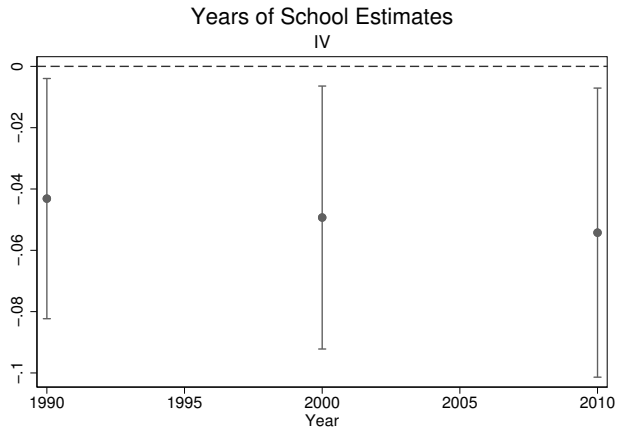
(a)



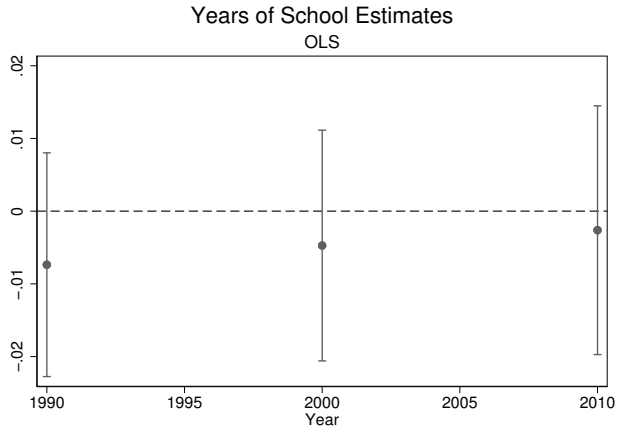
(b)

Figure 1.13: Illiteracy Estimates for Different Years

Notes: regressions as Base Model (1) in Table 1.5. Confidence interval: 95%. See Appendix A and A for definition of illiteracy rates. Source: Author's calculations based on census data.



(a)



(b)

Figure 1.14: Years of School Estimates for Different Years

Notes: regressions as Base Model (1) in Table 1.7. Confidence interval: 95%. See Appendix A for definition of years of school. Source: Author's calculations based on census data.

historically there have been several legal initiatives to improve this condition. For example, in 1899 the Congress passed a law that demanded that each *hacienda* had to send all children between 10 and 14 years old to school. Furthermore, it also required each *hacienda* with more than 20 *conciertos* to build a school (Oberem, 1978, p. 324). But in practice, the power landlords exerted over local authorities made it difficult to accomplish all these resolutions. Using census data from secondary enrollment in 1990, I also observe that regions with high rates of *concertaje* have a reduced school attendance, as reported in Table 1.8.

Table 1.8: IV Results: Secondary School Enrollment (1990)

	Base Model	Time fixed effect	Without Southern Districts	Province fixed effect	Distance to Capital	No Capitals	Fraction Indigenous (1800)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Concertaje	-.29*	-.48**	-.34**	-.4*	-.21	-.25	-.35**
	(.15)	(.19)	(.14)	(.21)	(.15)	(.16)	(.18)
Panel C: OLS							
Concertaje	-.02	-.04	-.01	-.07	-.01	.01	0
	(.06)	(.06)	(.06)	(.06)	(.06)	(.06)	(.06)
Geo. Controls	Yes	Yes	Yes	No	Yes	Yes	Yes
Time Controls	No	Yes	No	No	No	No	No
Province fixed effect	No	No	No	Yes	No	No	No
Observations	126	126	110	126	126	119	112
Mean Dep. Var.	57.58	57.58	57.64	57.58	57.58	56.42	58.51

Notes: Standard errors in parentheses. Geographic controls include elevation, latitude and longitude.

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

The landlord elite also had strong incentives to limit the education of indigenous workers as a strategy to preserve labor coercion. Only in 1978 were illiterate citizens

given the right to vote (Corkill (1985)). Hence, the provision of public goods, as roads and schools, may have been lower due to the lack of involvement of their citizens in public affairs. Moreover, illiterate workers could be easily deceived through the alteration of their debt accounts, which again gave an advantage to landlords to keep controlling *conciertos*. So on top of reducing *hacienda* workers' to access to education, the *concertaje* system also generated strong incentives to the landlord elite to block their formal education, reinforcing a circle of low education-low income.

Altogether, the results of this section make a case to explain the variations of human capital within a country, a mechanism also emphasized by Mariscal and Sokoloff (2000). It is possible to think that extractive institutions such as *concertaje* either create barriers to education or distort the incentives to invest in human capital, which in turn have negative effects on economic development. Additionally, *concertaje* also generated incentives for the landlord elite to block education of workers both as a method to perpetuate the institution, and perhaps also to maintain political control.

1.5.2 Sectoral Composition

Nowadays, the major disparities in productivity across countries are in the agricultural sector (Caselli (2005b), Restuccia et al. (2008)). At the same time, as documented by Restuccia et al. (2008), poor countries allocate more employment in agriculture than developed ones. One of the main justifications behind this empirical finding is the Schultz hypothesis (Schultz (1953)), which argues that poor countries face a "food problem", meaning that those countries allocate much of their resources to produce food for subsistence needs because of their low productivity. Although the food hypothesis seems plausible, another reason behind the important allocation of labor in agriculture in developing countries could come from historical labor coercive institutions. Given the close relation of *concertaje* with agriculture, a mechanism by

which the institution may have had persistent negative effects is by creating distortions precisely in that sector. Although those distortions might include a wide range of inefficiencies, a more direct effect of the institution could have been blocking the movement of workers to other economic activities.

Since *conciertos* were a valuable asset to the *hacienda*, landowners had strong incentives to retain both attached workers and their families for generations. Precisely, the mechanisms developed by the *haciendas* to control workers impede labor mobilization and might crowd out the development of other economic sectors. If in addition to this we include the lack of education of coerced workers evidenced in the previous section, it is not surprising this labor system generated an environment where workers remained attached to the agricultural sector, with little options to escape from it. As workers could not move to other occupations, this process might have negatively affected the economic prospects of a district.

The relationship between low income and share of labor allocated in agriculture observed in a cross-country setting also holds for the case of Ecuador. Figure 1.15 shows the positive relationship between the share of employment in agriculture and the poverty rates of the districts studied in this paper. Evidently, the immediate question is what are the reasons behind this relationship. Considering that *concertaje* might have provoked an effect in distorting labor markets, I hypothesize that there is a causal relationship between both of them, meaning that the institution could have had a direct effect in the labor composition of a district.

I test this hypothesis in the 2SLS framework by estimating the effects of *concertaje* in sectoral composition. As before, I estimate equation (1.2), but use A_d, the fraction of labor working in agricultural activities as a dependent variable. The results of these estimates are shown in Table 1.9. It is noticeable how the rate of *concertaje* is a strong predictor of the contemporary economic activity. Through all our different

specifications, the effect of *concertaje* is statistically significant, implying a very close relation between both: on average, an increase of 10 percentage points in the rate of *concertaje* have an effect of an increase between 6 to 10 percentage points in the share of workers allocated in agriculture.

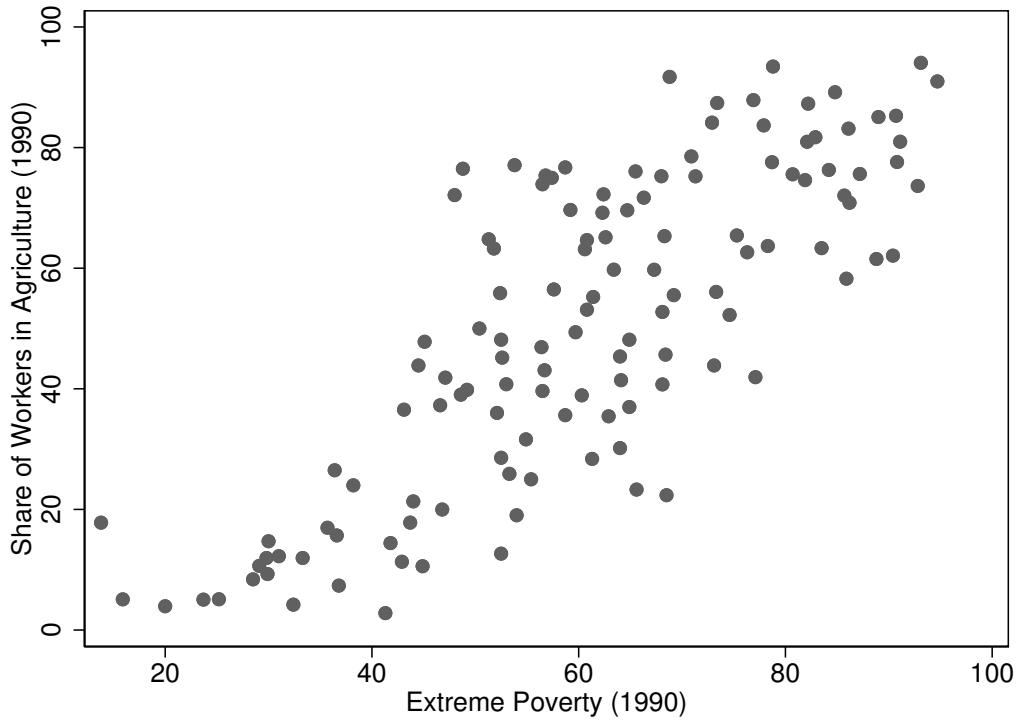


Figure 1.15: Share of Employment in Agriculture (1990) Vs Extreme Poverty (1990)

These results are also consistent with historical analysis showing a slow process of industrialization in the Ecuadorian highlands (Saint Geours (1994)). In particular, they reflect the persistent effects of *haciendas* and *concertaje* on agriculture dependence. Moreover, the results document how institutions may delay the process of structural transformation by creating barriers to allocate resources in the economy, in particular labor. If coercion dominates the labor market, not only the mobility of workers is limited, but the incentives to start new type of business could be reduced. A counterfactual then implies that the lack of labor coercion could facilitate

Table 1.9: IV Results: Share of Workers in Agriculture (1990)

	Base Model	Time fixed effect	Without Southern Districts	Province fixed effect	Distance to Capital	No Capitals	Fraction Indigenous (1800)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Concertaje	.83***	.95***	.8***	.88***	.6**	.73**	1.01***
	(.29)	(.32)	(.26)	(.32)	(.27)	(.29)	(.33)
Panel C: OLS							
Concertaje	.18	.23**	.16	.46***	.15	.12	.22*
	(.11)	(.11)	(.12)	(.1)	(.11)	(.11)	(.12)
Geo. Controls	Yes	Yes	Yes	No	Yes	Yes	Yes
Time Controls	No	Yes	No	No	No	No	No
Province fixed effect	No	No	No	Yes	No	No	No
Observations	126	126	110	126	126	119	112
Mean Dep. Var.	50.16	50.16	49.6	50.16	50.16	52.5	48.86

Notes: Standard errors in parentheses. Geographic controls include elevation, latitude and longitude.

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

the mobility of workers between sectors. A notable case is the town of Otavalo, where the minor presence of big *haciendas* and landlords contributed to the formation of trade specialists, who nowadays are renown because of their entrepreneurial talent (Valarezo (2002)). On the other side, the excessive burden landlords imposed over the indigenous workers in the neighbor town of Cayambe limited the formation of such traders. These differences in sectoral composition between both towns seemed to occur as early as the mid XIX century. For instance, Valarezo (2002) reports that census data of 1862 reveals important differences in the kind of occupations the population of each town had. While Otavalo reported to have 38% of its population working in the agricultural sector, Cayambe had 82%.

The results presented in this section also reflect the fact that historically, labor coercion has been intimately related with agriculture. Even nowadays, most forced labor relations occur in low-technology, labor-intensive activities as agriculture (Andrees and Belser (2009)). In the context of Latin America, for instance, labor coercion has been historically concentrated in agricultural and mining activities. While most mines controlled by the Spanish were depleted as early as possible, coercion in the agricultural sector remained for a longer period of time. Accordingly, the persistence of the institution seems to have created barriers in the allocation of labor that later on was reflected on important differences in economic performance.

1.5.3 Inequality

Engerman and Sokoloff (1997) suggested that the colonial strategy imposed in South America by the Spanish generated a very unequal society, a situation which later on affected the economic prospects of the region. Due to the distribution of land claims and privileges to a small group of settlers, a reduced economic and political elite emerged, who extracted rents from natives via forced labor or other forms of taxation. Inevitably, over time this strategy could have led to a high concentration of wealth. Ecuador seems not to be an exemption to this analysis, as the existence of *concertaje* is a reflection of the control of the indigenous population by a small, racially different elite. Consequently, it is expected to find a positive relation between *concertaje* and inequality.

Specifically, I focus on the distributional effects of *concertaje* on land inequality as the institution was heavily supported on the existence of big rural states. The data about land distribution in Ecuador at a parish level comes from the agricultural census of 1974. It is important to mention that this is the only source that reports land data in such disaggregated level. The census provides data about the number

of agricultural units and classifies them under 15 categories based on size. Using that information I build a Gini coefficient of land inequality similar to Nunn (2008). Additional details about the data and the construction of the Gini coefficients is provided in Appendix A.

I report the results between land inequality and *concertaje* in Table 1.10. Notably, from the OLS to the IV estimates, all specifications reflect a positive and statistically significant relationship between *concertaje* and land inequality. For instance, the IV estimates suggest that an increase in 10 percentage points in *concertaje* in 1800 increased Land Gini in 1979 in almost 5 percentage points. These estimates are consistent with the Engerman and Sokoloff (1997) hypothesis that forced labor institutions like slavery or *concertaje* induced high levels of inequality, and aligned with the empirical evidence presented by Nunn (2008) between the positive correlation between slavery and inequality in the United States. From a historical perspective, the results reported in Table 1.10 should not be unexpected. First, their correspond to the year of 1974, a period in which the agricultural reform was just starting to take effect, and so they reflect much of the colonial heritage conditions of land distribution until that moment.

Nevertheless, in order to consider inequality as a channel of persistence, it should be negatively related with economic prosperity. Although the evidence on inequality and development at country level is mixed[16], several papers have also used micro-data to explore this relationship. Acemoglu et al. (2008) used data from the region of Cundinamarca in Colombia, and reported a positive correlation between economic inequality and contemporaneous economic development. Instead Nunn (2008) reported

[16] Banerjee and Duflo (2003) report a U-shaped relationship between inequality and economic growth, arguing that changes in inequality in any direction are correlated with a decrease in growth in the future.

Table 1.10: IV Results: Land Gini (1974)

	Base Model	Time fixed effect	Without Southern Districts	Province fixed effect	Distance to Capital	No Capitals	Fraction Indigenous (1800)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Concertaje	.47**	.52**	.37*	.74***	.41**	.47**	.48**
	(.2)	(.21)	(.19)	(.26)	(.2)	(.21)	(.22)
Panel C: OLS							
Concertaje	.22***	.21***	.22***	.23**	.21***	.22***	.23***
	(.07)	(.08)	(.08)	(.09)	(.07)	(.08)	(.08)
Geo. Controls	Yes	Yes	Yes	No	Yes	Yes	Yes
Time Controls	No	Yes	No	No	No	No	No
Province fixed effect	No	No	No	Yes	No	No	No
Observations	125	125	109	125	125	118	112
Mean Dep. Var.	72.92	72.92	73.69	72.92	72.92	72.91	73.21

Notes: Standard errors in parentheses. Geographic controls include elevation, latitude and longitude.

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

a weak correlation between inequality and today's economic income.

The relation between land inequality and extreme poverty can be seen in Figure 1.16. There is no apparent correlation between both variables, suggesting that inequality by itself may not be a persistence mechanism, although the results should be taken with caution. For instance, inequality could not only appear in the land distribution across districts, but also at a country level, by creating a very hierarchical society where a small elite controls de political and economic power of a nation. For example, Acemoglu et al. (2008) emphasizes the role of political inequality instead of economic inequality in explaining differences in economic performance. This narrative fits the analysis of many historians who have studied the region, and have

identified landlords as a ruling elite that blocked the political participation of the rest of the population (Quintero López (1987)).

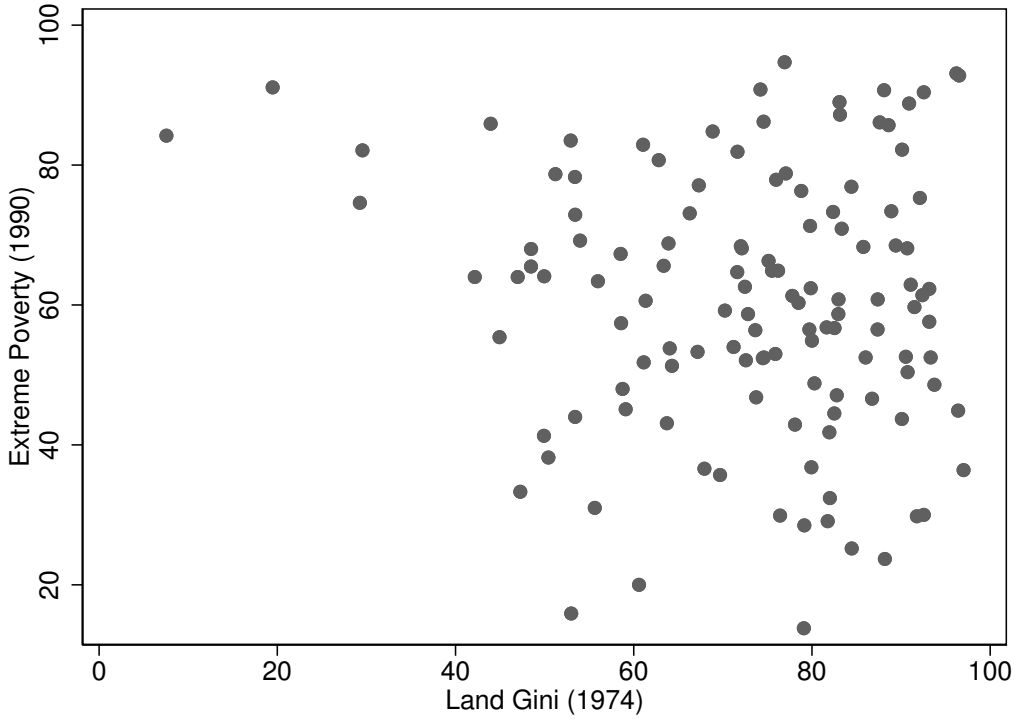


Figure 1.16: Land Inequality (1979) vs Extreme Poverty (1990)

These results also seem to suggest that although *haciendas* created a high concentration of land, its negative repercussions might come from the effects of coercion through human capital. Galor et al. (2009) argues that land inequality could have adverse effects in economic development by blocking human-capital promoting institutions, but the results presented in this section emphasize not only the role of land inequality by itself, but the negative effects of labor coercion, which might or might not be accompanied with land inequality. For example, the coast of Ecuador, although not predominantly influenced by mechanisms of labor coercion as *concertaje*, exhibited land concentrations very similar to those of the highlands[17]. Because

[17]Barsky (1984) reports a Gini coefficient for the highlands in 1979 of 0.64, while the one of the

of a small labor force and the risks of tropical diseases, the coast was not heavily populated, or controlled, by the Spanish. Although the ports were of economic relevance, agricultural activities there were not of special interest until the prices of cacao and the availability of some labor force generated enough conditions to attract new colonizers. In that case, land availability and a reduced population created a landlord class that claimed and appropriate the land because of its availability. For different historical reasons, the coastal region also showed a very unequal distribution of land, but not the negative effects of labor coercion. Probably the inconclusive empirical relationship between inequality and economic development is due to the lack of understanding of the mechanisms that lead to inequality. In some cases, inequality might be capturing the effect of coerce labor institutions, while in others, it might be the reflect of other historical processes.

1.5.4 Public Goods Provision

Instead of suggesting a negative relationship between *haciendas*, *concertaje*, and contemporary economic development, Dell (2010) states that under the presence of the mining *mita*, *haciendas* in Peru protected indigenous workers by offering an escape from the cruel conditions of working in the Potosi silver mines. Hence, she proposes an alternative hypothesis to Engerman and Sokoloff (1997) theory to explain the development path in the Americas, where she emphasizes the role landlords had on providing public goods, like roads, and secure property rights. Accordingly, I test if places with higher rates of *concertaje* also show higher presence of public goods, measured by the density of local roads.

In order to measure the density of local roads, I follow the same strategy as Dell (2010), and use a GIS road map of Ecuador. Roads are classified by two dimensions,

coast was of 0.62

either as local, secondary or primary, and according to the type of material they are made off: paved, not paved, temporal or other. Using that classification, I compute the total length of roads and divide them by the district surface area.

Table 1.11 shows the results between *concertaje* and local road density. Although the coefficients are not precisely estimated, all the specifications show a negative relation between the institution and the provision of roads. These results are contrary to those presented by Dell (2010), who hypothesizes that large landowners with secure property rights had higher incentives to invest on public goods. In contrast, I observe that public goods, reflected in road density, seem to be negatively related to the existence of forced labor in *haciendas*. If landlords in the highlands of Ecuador produced to sell to local markets, a question then is why they did not invest or lobby for better or more roads?

One possibility is that the data on roads, which belongs to 2010, might reflect quite different conditions from those in the past due to the recent important investments in infrastructure by the government. To see if that is the case, I use data from the 1974 agricultural census where there is information about the number of farms that are further from certain distance to transitable roads. The classification is a farm being less than 1km, between 1 and 5km, between 5 and 10 km and further than 10km. I compute for each district the fraction of farms that are further than 5km from a transitable road. The results can be observed in Table 1.12. Once again the estimates suggest that higher rates of *concertaje* are associated with farms being further from a transitable road.

Another explanation could be that the narrowness of the Andes in Ecuador, which is quite differently from the extensions found in countries like Peru or Bolivia (Basile (1974)), gives little variation in the road density. Then, the discrepancy in conclusions could simply reflect conditions inherently specific to each country. However, an

Table 1.11: IV Results: Local Road Density

	Base Model	Time fixed effect	Without Southern Districts	Province fixed effect	Distance to Capital	No Capitals	Fraction Indigenous (1800)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Concertaje	-24.63*	-31.07*	-18.24	-24.55*	-13.46	-15.61	-31.35*
	(14.64)	(16.84)	(13.64)	(14.79)	(12.83)	(12.63)	(16.68)
Panel C: OLS							
Concertaje	.74	.68	.12	-14.68**	2.18	2.08	-.8
	(5.29)	(5.52)	(5.65)	(6.44)	(5.02)	(4.82)	(6.19)
Geo. Controls	Yes	Yes	Yes	No	Yes	Yes	Yes
Time Controls	No	Yes	No	No	No	No	No
Province fixed effect	No	No	No	Yes	No	No	No
Observations	126	126	110	126	126	119	112
Mean Dep. Var.	1111.91	1111.91	1242.96	1111.91	1111.91	1045.82	1198.87

Notes: Standard errors in parentheses. Geographic controls include elevation, latitude and longitude.

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

economic explanation may lie in the fact that *conciertos* were an asset, and so, landlords minimized their chances to escape, limiting the availability of workers within a district. For instance, during the construction of a road system near the town of Otavalo, state officials complained about the *hacienda* system as the principal obstacle for labor recruitment (Williams, 2007, p. 51). Although it may have been in their best interest to improve the road system close to their properties, landlords also seemed to be careful about retaining the workers they had. If *hacienda* workers had to go and be part of the pool of road workers, the likelihood of escaping or that somebody else appropriated them might have been risks landlords did not wanted to take. Then, under the presence of such uncertainties, it is not clear that a major provision

of public goods could exist in districts with higher presence of *haciendas*. Also, the availability of plenty of labor allowed landlords to use its workers to transport objects, without the need to invest on transportation, as exemplified in 1925, when electric equipment was mobilized for 150 km. by 3000 indigenous workers (Albornoz, 1971, p. 80).

Table 1.12: IV Results: Percentage of Farms Further than 5 Km. (1974)

	Base Model	Time fixed effect	Without Southern Districts	Province fixed effect	Distance to Capital	No Capitals	Fraction Indigenous (1800)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Concertaje	.63**	.46	.52**	.65*	.31	.63**	.6**
	(.28)	(.29)	(.21)	(.38)	(.25)	(.3)	(.3)
Panel C: OLS							
Concertaje	-.04	-.08	.02	.08	-.07	-.06	-.06
	(.09)	(.09)	(.09)	(.1)	(.08)	(.1)	(.1)
Geo. Controls	Yes	Yes	Yes	No	Yes	Yes	Yes
Time Controls	No	Yes	No	No	No	No	No
Province fixed effect	No	No	No	Yes	No	No	No
Observations	122	122	107	122	122	116	108
Mean Dep. Var.	17.78	17.78	13.57	17.78	17.78	18.41	16.28

Notes: Standard errors in parentheses. Geographic controls include elevation, latitude and longitude.

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Perhaps also a larger presence of labor coercion affected the state capacity. Notably, Table 1.13 show that *concertaje* is associated with a lower presence of workers in the public sector. Hence, the provision of public goods may have been more difficult to accomplish in places where landlords have larger control. This could have happened if landlords wanted to exert more influence over local affairs by having a re-

duced number of government workers. At the same time they might have been afraid that a larger presence of the state may implied certain forms of regulation. Hence, the *concertaje* system could have not necessarily given landlords incentives to provide more public goods. Instead, it seems to be the case that the opposite happened.

Table 1.13: IV Results: Percent of Workers in Public Sector (1990)

	Base Model	Time fixed effect	Without Southern Districts	Province fixed effect	Distance to Capital	No Capitals	Fraction Indigenous (1800)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2SLS							
Concertaje	-.2**	-.28**	-.23***	-.24*	-.15*	-.19**	-.27**
	(.09)	(.11)	(.08)	(.12)	(.08)	(.1)	(.11)
Panel C: OLS							
Concertaje	-.07**	-.08**	-.05*	-.08**	-.07**	-.05	-.08**
	(.03)	(.03)	(.03)	(.04)	(.03)	(.03)	(.04)
Geo. Controls	Yes	Yes	Yes	No	Yes	Yes	Yes
Time Controls	No	Yes	No	No	No	No	No
Province fixed effect	No	No	No	Yes	No	No	No
Observations	126	126	110	126	126	119	112
Mean Dep. Var.	10.05	10.05	9.31	10.05	10.05	9.26	10.42

Notes: Standard errors in parentheses. Geographic controls include elevation, latitude and longitude.

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

1.6 Conclusion

This paper documents the long-run effects of *concertaje* in the Ecuadorian highlands by exploiting the variation in its intensity across districts. I first provide evidence of the persistence of the institution both during colonial times and afterwards to emphasize its importance in the economic development of the region. Then I use

as an instrument for *concertaje* the ratio of crops with different labor requirements. I state that this instrument could be considered as a plausibly exogenous variation of *concertaje* supported in the hypothesis that geographic conditions influenced the incentives to coerce, which in turn affected the type and severity of colonial institutions stablished in different regions. I estimate that in average, an increase of 10 percentage points in the rate of *concertaje* around 1800 increased the rate of extreme poverty in 1990 by 6 percentage points and reduced average night light intensity in 2 digital numbers (DN).

Due to the close relationship between *concertaje* and the labor market, I hypothesize that the presence of this labor coercive institution created barriers to the process of human capital accumulation and generated distortions that affected the sectoral composition of labor. I estimate that on average, an increase in 10 percentage points in the rate of *concertaje circa* 1800 increased the illiteracy rate of a district in 1960 by 3 percentage points and the share of workers in agriculture by 8 percentage points. These results plausibly document the role historical institutions as a cause of persistent market distortions.

Using data about land distribution around 1970 I also document a strong relationship between the institution and land inequality, favoring the Engerman and Sokoloff hypothesis about the relationship between endowments, institutions and inequality. However, I found no relation between inequality and economic development. Instead, I suggest that labor coercion by itself, which sometimes might be accompanied with land inequality, was a detrimental factor to economic development. These suggests the relevance of understanding the underlying reasons behind the historical inequality in order to asses the question between its relation with economic development.

Also, using data of contemporary roads, I report that there is a negative relation between *concertaje* and the provision of public goods. These results are opposite to

those of Dell (2010), who highlighted a positive role of *haciendas* in Peru in providing access to roads. Instead, the results presented in this paper do not favor Dell's hypothesis that landlords and *haciendas* shield individuals against extractive institutions. For the case of Ecuador, where the presence of gold and silver mines during colonial times was minimal, it is possible to have a more homogenous setting to study the effects of coercive labor institutions promoted by landlords. Within that setting, there seem to be more evidence favoring the historical analysis that *haciendas* constituted extractive institutions that affected the process of economic growth. Overall, I emphasize that is labor coercion by itself an important institution that generated divergences in the economic development across regions.

Under the lens of the evidence presented in this paper, the nature of coercion is what explains the negative effects of forced labor institutions. Coercion, when profitable, usually demands that workers are retained both in a certain activity (like agriculture) and in a certain place, becoming a barrier for labor reallocation across sectors and regions. To perpetuate this condition, coercers have to make sure that workers' outside option is low enough. This is achieved via legal punishment (debtor's prison), and by limiting and reducing the incentives the acquire other forms of human capital than those required by the coerced activity. This last point is achieved by either blocking politically attempts to invest in the provision of public goods like schooling, or by plainly denying the access to education. Furthermore, in order to keep a stable labor force over time, the next generation needs to be coerced from early ages, which induces a substitution of school for labor. As illiterate workers were relegated from political participation, like voting, this perpetuates the lower provision of public goods and reinforces the barriers to escaping from coercion.

A more robust conclusion of the results presented in this paper requires obtaining data of *concertaje* for the republican period, as well as the finding of intermediate

outcomes in the 19th century in order to better understand at which point institutions started to generate divergent development paths across regions. Future topics of research might consider the findings in this paper as a motivation to develop models where institutions explicitly play a role in the process of structural transformation by creating barriers or generate a misallocation of resources. Moreover, further contribution of empirical evidence about the relationship between institutions and human capital might help us to disentangle the puzzle about the fundamental causes of economic growth.

Chapter 2

ALL IN THE FAMILY: FIRM DYNAMICS AND FAMILY MANAGEMENT

2.1 Introduction

Firm dynamics in developing countries are quite different from those of developed ones. For instance, Hsieh and Klenow (2014) provide evidence that 40-year-old plants in India and Mexico are 40% and 100% respectively larger than new plants, whereas plants in the United States are 600% larger. These life-cycle differences in developing countries could reflect barriers that firms face to raise their productivity, which in turn may have important consequences for Total Factor Productivity (TFP). Therefore, understanding the forces behind these growth differentials appears to be relevant for explaining income variations across countries[1].

One channel that may explain these life-cycle differences is contract enforcement frictions that hinder the hiring of external managers in developing countries (Bloom et al. (2013)). Instead, firms' owners in developing countries are much more likely to rely on family members to manage their business (La Porta et al. (1999), Claessens et al. (2000)). This is relevant because managerial practices are associated with firm-level productivity and they tend to be particularly poor in family-managed firms (Bloom and Van Reenen (2007)). However, further exploration of this channel has been limited due to unavailability of firms' data where one can observe at the same time both family relations between owners and managers and their corresponding dynamics.

[1]See Caselli (2005a) and Jones (2016) for a review of the role of TFP in explaining cross-country income differences.

In this paper I overcome this challenge by using longitudinal data from administrative tax records for the universe of all firms in Ecuador, merged with a complete national registry of upper managers and shareholders. Thanks to a combination of information about true family linkages and inference through surnames, I can identify if a firm is managed by a family member. More concretely, I can tell if the upper management (e.g. CEO for a Corporation) is related to any of the shareholders of the firm. Consequently, I can classify each firm as Family-Managed (FM) or Externally-Managed (EM) and use their corresponding tax records to provide several stylized facts about each type of firm.

The main finding, depicted in Figure 2.1, is that the revenue life-cycle profile is twice as steep in Externally-Managed firms as in Family-Managed firms. That is, by the age of 20, EM firms are 200% larger if compared to firms of the same type at age 1, while FM firms are 100% larger than their corresponding younger counterparts. A consequence of these life-cycle differences is that despite representing 80% of total firms, Family-Managed firms only account for almost 40% of total revenue and 50% of total employment in the economy. Accordingly, in contrast to the overall distribution of family-managed firms, 80% of large firms (i.e. more than 500 workers) are externally-managed. I show that these results are robust to various classification criteria of family management and to different sample selections.

The magnitude of the life-cycle gap between FM and EM firms I present here is similar to the cross-country differences in the life cycle of plants documented in Hsieh and Klenow (2014). Since family management is thought to be a widespread phenomena in developing countries, this channel appears to be quantitatively relevant to explain cross-country differences in firm dynamics and its implications for economic development[2]. In that sense, although precise statistics of the overall presence of FM

[2]Economic growth could also be affected by both differences in firm dynamics (Akcigit (2017))

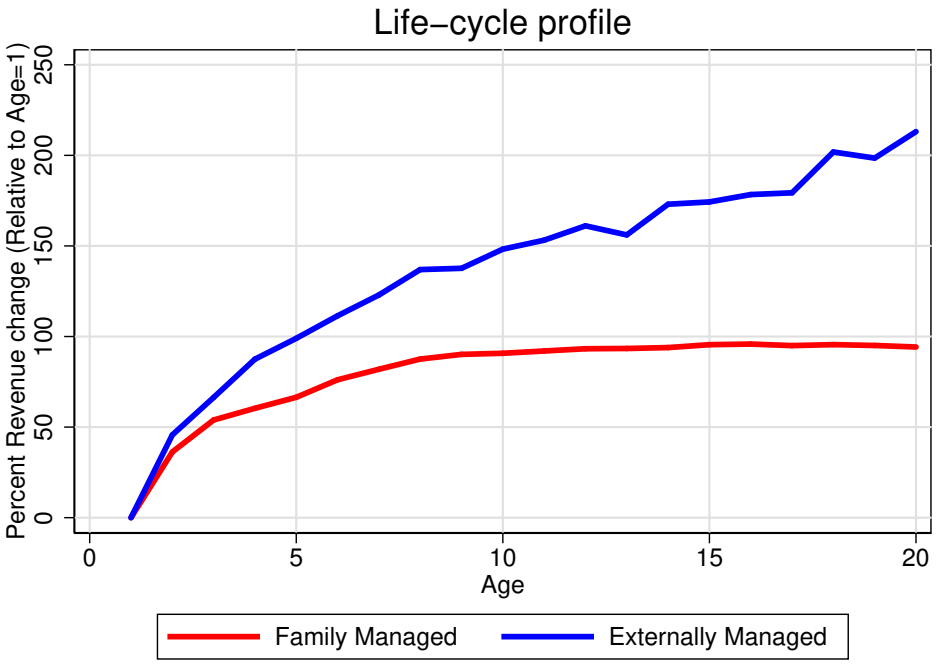


Figure 2.1: Life-cycle Profile of Family-Managed and Externally-Managed Firms

Notes: The sample includes the universe of all formal firms in Ecuador, except the financial sector, for the period 2009-2016. Age equal one is the base category. Revenue is deflated using the CPI. Data sources: Servicio de Rentas Internas (SRI) and Superintendencia de Companias, Valores y Seguros (SCVS) of Ecuador.

firms are rare for other countries, comparable samples to survey data from Bloom and Van Reenen (2007) suggest that the share of FM firms in Ecuador is around 2 times that of the United States and other developed countries.

I also report through a case-event study that there is an expansion in revenue and employment when a family-managed firm switches to external management. Although this transition is infrequent, when it happens, firms show in average gains in revenue and employment of 20% and 10% respectively two years after the switch, that persist over time. On the contrary, when firms change management but remain

and a large concentration of inherited firms (Morck et al. (2000)).

as a family-managed firm, they exhibit smaller gains that vanish faster years after the replacement.

Even though I cannot tell the identity of the successor within the family, altogether these findings are consistent with causal evidence that a firm's performance drops after a primogeniture succession (Bennedsen et al. (2007)), and that managerial practices are worse in family managed firms, especially when administered by the first-born (Bloom and Van Reenen (2007)). Therefore, one interpretation of the underperformance of family-managed firms over the life cycle is the selection of critical managerial positions from a small pool of talent (Perez-Gonzalez (2006))[3].

I also discuss two pieces of evidence that indicate the abundance of Family-Managed firms is likely to be the symptom of a deficient contractual environment (Laeven and Woodruff (2007), Bloom et al. (2009)). First, I show that the presence of family management is similar across industries, and varies little over age and time, revealing an overall barrier to the hiring of external managers. Moreover, contractual distortions could make difficult the diversification of ownership, making it more likely that the control of a firm falls in the hands of a family member (Burkart et al. (2003)). Consistent with that interpretation, I note that most firms in Ecuador have very few owners. For example, in 85% of the firms there are less than 5 owners.

Based on the stylized facts discussed above, I build a general equilibrium model of firm dynamics à la Hopenhayn (1992) and Hopenhayn and Rogerson (1993) where entrepreneurs choose to operate either with their endowed family managerial talent or by hiring external management services on the market. Management is modeled as a factor of production, and certain levels of productivity induce firms to operate only

[3]Empirical studies have also provided further understanding of mechanisms that explain differences in performance of family managed firms, like a reduced supply of time of family managers (Bandiera et al. (2017)) or how the family structure shape a business (Bertrand et al. (2008)).

through family management. Firms that choose to be family managed rely on the fixed pool of talent given by blood, which implies a decreasing returns to scale production function in the remaining factors. Firms that choose external management can hire more managers and hence have a constant returns to scale production function. However, contracting frictions make external managers less efficient (as they need to be incentivized and monitored).

I calibrate the model to the Ecuadorian economy and then use it as a framework to study counterfactuals and policy analyses. The model is able to replicate the new stylized facts presented here, including the life cycle differences between firms, the prevalence of FM firms and their relative share in aggregate employment. In the main counterfactual experiment, I find that changes in the contractual environment that lead to a drop in the presence of family-managed firms by half could increase output on the order of 6%. These gains operate through both extensive and intensive mechanisms. First, the share of Family Managed firms is reduced by almost half, so the larger presence of Externally Managed firms increases the average growth of firms by 50%. Second, already-EM firms become more efficient in the use of managerial services, so they demand more inputs and become larger.

Finally, I use the detailed administrative data at hand to provide some suggestive evidence on one managerial practice in which FM firms seem to differ: workforce selection. I join the tax records of every classified firm with those of their corresponding workers to create a full matched employer-employee dataset. I show first that workers in FM firms are on average paid 10% less after controlling for individual fixed effects. I then decompose worker's earnings using the AKM model (Abowd et al. (1999)) and document that worker quality and firm productivity (the respective fixed effects of the estimation) are more strongly correlated for externally-managed than family-managed firms. This seems to suggest that potentially productive firms, by

restricting their management to their family talent, fail to hire and retain otherwise good-match workers that could enhance firms' productivity.

This paper builds on recent models of management and firm dynamics that try to explain life cycle differences across countries. In particular, it closely follows on the work of Akcigit et al. (2016) in which firms can optimally decide when to hire external management in order to avoid the diminishing returns of size as in Lucas (1978). However, a key difference in this paper is that I explicitly categorize firms according to their family management status. Similarly, the work of Cole et al. (2016) focus on firms' life-cycle differences across countries through the eyes of financial frictions and its implications for technological adoption, but without emphasizing the role of management. Caselli and Gennaioli (2013) also stress the importance of financial frictions as an important mechanism to explain the abundance of family firms in developing countries, although abstracting from the firm dynamics[4]. Guner et al. (2018) instead study how the earnings life cycle of managers differs across countries, emphasizing the role of human capital accumulation, similar to Roys and Seshadri (2014) and Bhattacharya et al. (2013), while Alder (2016) discusses the aggregate implications of mismatch between managers and firms.

On the other side, understanding the sources of cross-sectional differences in the allocation of factors across countries has been the focus of recent papers (Guner et al. (2008); Midrigan and Xu (2014)) due to the aggregate consequences for Total Factor Productivity (Restuccia and Rogerson (2008), Hsieh and Klenow (2009), Bartelsman et al. (2013))[5]. With those works mainly centered on cross-sectional differences in size

[4]The role of financial frictions for economic development is also studied by Buera et al. (2011) and Moll (2014).

[5]See Restuccia et al. (2013), Restuccia and Rogerson (2013) and Hopenhayn (2014) for a review of the recent literature.

and less scope for the life-cycle dynamics of firms[6]. Yet, the cross-sectional implication of family management studied in this paper is that larger firms are typically those that have delegated outside management. Since family management is suspected to be abundant in developing countries, this is consistent with large cross-country differences in the size of establishments (Bento and Restuccia (2017)).

Finally, studies stressing the importance of management for firm's growth can be traced back to Penrose (1959), and more recently to Bertrand and Schoar (2003) and the research agenda of measuring managerial practices across countries (Bloom and Van Reenen (2010), Bloom et al. (2014)). There is also a large literature in corporate finances studying the particularities of family firms, with special focus on family successions (Villalonga and Amit (2006)), and performance (Morck et al. (1988); Yermack (1996))[7]. From a theoretical perspective, the effect of family management on firm's performance is ambiguous (Donnelley (1964)), as family managers could face higher incentives due to non-pecuniary benefits (Kandel and Lazear (1992)), while a negative aspect is the selection from a small pool of talent. In that sense, the empirical evidence presented in this paper puts more weight to the latter.

This paper is organized as follows. The next section describes the data and the methodology used to classify firms. Section 2.3 provides several stylized facts of Family-Managed and Externally-Managed firms. Section 2.4 describes the model, while Section 2.5 discusses the calibration strategy and Section 2.6 performs several counterfactual analysis of changes in the contractual environment. Section 2.7 returns to provide more empirical evidence about some managerial practices in which both types of firms differ. Section 2.8 concludes.

[6]See Bartelsman et al. (2009) for several facts about cross-country differences in firm dynamics.

[7]See Bertrand and Schoar (2006) for a review of the literature

2.2 Data

The main empirical exercise of the paper is to classify firms according to their family management status and then analyze if there are differences in performance associated with it. To do so, I rely on administrative records from two sources: (i) tax records and (ii) a national registry of shareholders and upper managers. The latter allows me to identify if a firm is managed by a family member, while the first one lets me observe their performance over the life-cycle and across several other dimensions.

Tax records, of restricted access, were provided by the Ecuadorian Internal Revenue Service (Servicio de Rentas Internas (SRI)), while a complete registry of shareholders and upper managers was provided by the Bureau of Companies (Superintendencia de Companias, Valores y Seguros (SCVS)). In both sources I have the unique firm's ID, which allows me to join datasets. Furthermore, both datasets combine the universe of nearly all formal firms (around 150,000 by year), and includes both Sole Proprietors and Limited Liability Companies, the two main forms to organize a business in the country. The only sector excluded from my analysis is the financial one, since it is regulated by a different entity and is not required to provide information about either managers or owners.

In addition I have both the unique citizen's ID of both shareholders and managers, so I can use their corresponding tax records in a similar fashion as firms. Finally, I can also link workers (beyond managers) to firms to create a complete matched employer-employee dataset because workers' tax records are mainly reported by their employers. This is particularly relevant to explore one channel in which FM firms differ: workforce selection, a point discussed in detail in Section 2.7.

It is noteworthy to emphasize that the registry of shareholders covers every formal firm regardless of size or condition, and it includes the records of every shareholder,

even for public-listed firms. What is more, I have the complete history of managers and shareholders from the moment a firm started operating, although I only have access to their corresponding tax records starting for 2009. Hence, I limit the analysis of the data for the period 2009-2016. Finally, all relevant variables were deflated using the CPI deflator provided by the National Institute of Statistics (INEC).

2.2.1 Tax Records

Firms

Tax records of firms include detailed financial information mainly composed by the firm's balance sheet and its income statement, accompanied by data of several firms' characteristics such as the year they where legally founded, 6-digit industry code and their geographic location. From the Balance Sheets I mainly use the firm's book value measure of capital stock, while from the Income Statement I obtain information about firm's revenues and costs.

Data about total employment do not appear directly in the tax records of firms, however, given that I also have access to the tax records of workers, which report the employer ID, I compute the total number of workers from that source and then merge it to the firms' database. Moreover, information about total employment for Limited Liability Companies was also provided by the Bureau of Companies, as firms were asked to self-report it while updating the information about ownership and management.

Individuals

Individuals report taxes in Ecuador mainly through their employers. Hence, it is possible to connect the IDs of both parts to create a matched employer-employee longitudinal dataset. For those who declare income from other sources different from

employed labor (rents, dividends, business income, self-employment income and capital income), a second report of taxes is used. The principal variable recovered from these records is labor earnings. Further details about the specificities of this information can be found in the Appendix B.

2.2.2 Ownership and Management

Data for ownership and management for Limited Liability Companies are available because by law, every Limited Liability Company constituted in Ecuador, regardless of its size, must report to the Bureau of Companies during the first four moths of the year a full list of its legal representatives, upper managers (discussed below), and shareholders. Any change during the year should be updated within three months after the event (Art. 20, Ley de Compañías (1999)). For any firm, the legal representative is the manager on whom falls the legal burden of the firm, so it is a category aside from any management position[8].

For Sole Proprietors, ownership and management are not reported to the Bureau of Companies. However, in this case ownership is trivial, as there is only one residual claimant. Since at the same time the law defines the residual claimant to be the legal representative of the firm, I consider that person to be the main manager of the firm. Instead, for Limited Liability Companies, the identity of managers and owners was divided in two parts. I describe below the treatment of each dataset.

Managers appointments

Starting from the moment of the firm's legal foundation, I was provided with a full history of upper-managerial appointments. This includes both the starting and ending date and the corresponding position. I use these spells to build a panel of these

[8]Foreign companies are only required to provide the identity of their legal representative.

managers for every firm, cut at the last day of the year. This was done to later be consistent with the timing of the information presented in the tax records. Management positions are self-reported to the Bureau of Companies.

Upper management in a Limited Liability Company overall is defined as the individuals in charge of leading it and responding to shareholders. Nevertheless, from all management positions, I mainly focus on the figure of Legal Representative (which in most of the cases is the CEO) since by law a Limited Liability Company requires a legal representative in order to operate (Art.6, Ley de Compañías (1999)). I also focus on that position because although firms report other managers (like CFO, President, members of the board or plant managers if applicable), this information has not been standardized, so the hierarchy of these positions is not completely clear. Although it is possible that the legal representation of the firm may be separate from the the actual managerial duties, anecdotal evidence suggest this is not the case. Nevertheless, in some cases I also consider the rest of upper managers, particularly the CEO, top manager and President. Details about the selection of upper manager's and its implications to the classification of firms are discussed in detail in Section 2.2.3.

Table 2.1 displays the frequency of different categories of managers in the dataset. As it can be observed, the main categories of upper management are CEO (Gerente General)/Manager and President/Executive President. Moreover, one can see that in 94% of the cases, the Legal Representative of the firm is either the CEO/Top Manager or the President of the firm, the most common case (69%) being the CEO/Top Manager. Therefore, as discussed above, focusing on the Legal Representative as the leading manager of the firm is a plausible assumption.

Table 2.1: Managerial Positions in Limited Liability Companies

	(1)	(2)
Position	Unconditional distribution	Conditioning on being the Legal Representative
Manager/CEO	54%	69%
President/Executive President	40%	25%
Other	6%	6%
N	665,395	516,230

Notes. Distribution of managerial positions for the years 2007-2016.

Data Source: Superintendencia de Companias, Valores y Seguros.

Ownership

Similar to the case of manager's appointments, I was provided with a full history of the stock's transactions of every firm from the moment of its foundation, accompanied with the identity of the person involved on it. This includes transactions like the initial amount of capital, purchases, sells, transfers, reevaluations, and new capital injections. Knowing all these details allows me to know the distribution of shareholders at any point of time. For example, when a transaction includes two parties, for instance, a buyer and seller, I have information about the identity of both. Hence, I can build at any period the stock of capital of every shareholder.

I compute at the end of each year the distribution of stocks of every shareholder and update it according to the flows provided by the stock's transactions. Additionally, since I have the complete list of owners of every firm, I can track the final shareholder for those cases when a firm is owned by another firm. This tracking can only be done for firms stablished in the country, as foreign firms are not required to reveal the full composition of their shareholders.

Figure 2.2 shows the distribution of the number of shareholders by firm for the year 2016. As it can be seen, ownership of most firms is highly concentrated. Around 85% of firms have fewer than 5 shareholders, while less than 1% of firms have more than 100 shareholders[9]. This lack of dispersion is consistent with the interpretation of weak rule of law and less effective legal protection of minority shareholders in developing countries (La Porta et al. (1997), La Porta et al. (1998), La Porta et al. (1999)).

In that sense, the concentration of ownership is likely to be related to the incidence

[9]The mode of the distribution is two since this is the minimum number of individuals required to start a limited liability company

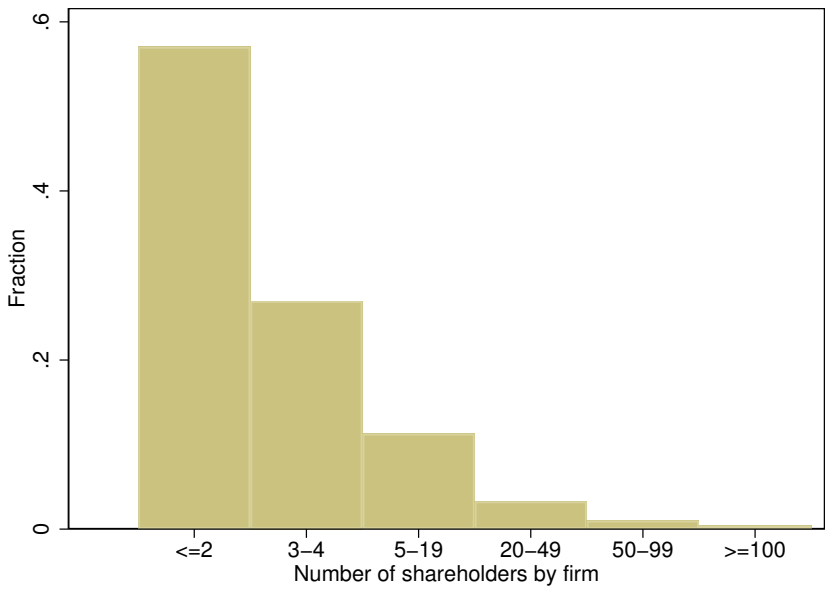


Figure 2.2: Frequency of Shareholders in Limited Liability Companies

Notes: The sample includes the universe of all non-sole proprietors firms in Ecuador, except the financial sector, for the period 2016. Data source: Superintendencia de Companias, Valores y Seguros (SCVS) of Ecuador.

of Family-Managed firms. If minority investor rights are not well protected, it may be challenging to diversify ownership, reducing the checks among shareholders and making it more likely that delegation falls on the hands of a family member. Instead, in environments with better protection rights, widely held corporations where professional managers exercise a vast control of the firm could arise. Hence, the decision to delegate is likely to be shaped by the legal environment, so family management could be seen as second best when there is difficulty to diversify ownership due to a weak protection of investors' right (Burkart et al. (2003)).

2.2.3 Classification of Firms

The ultimate goal of having the information of both managers and owners is to identify if a firm is managed by someone of the family or instead has an external manager. The identification of family linkages between the upper management and the shareholders within each firm was done by using a combination of two approaches: inference via surnames and observation of true linkages using administrative data. I explain both methodologies below.

Surnames

My benchmark approach to classify family-managed firms relies on identifying whether firm owners and firm managers are related through their surnames. Surnames offer two advantages for making such linkages. First, individuals sharing surnames are very likely to be related, particularly because the distribution of surnames in Western societies (including Ecuador) is skewed, meaning that although there are few individuals with really common surnames, most have infrequent ones, making it more likely that sharing a surname implies familiarity[10]. Moreover, as I am not inferring family linkages via surnames of two randomly drawn individuals, but instead focusing on shared surnames in a very specific place, a firm, it is more likely to accurately infer a family relation using this methodology.

Second, Ecuador is a Spanish-speaking country, so surnames encode information about both the mother's and father's families, making it possible to identify family relations from both sides until the fourth degree of consanguinity (cousins). One drawback, however, is that individuals in Spanish-speaking countries as Ecuador do

[10]For instance, the use of surnames has been used to provide estimates of intergenerational mobility in the absence of true family linkages (Güell et al. (2015)).

not change surnames when they get married, making it unfeasible to identify affinity relations. I partially overcome this problem by using administrative records of true linkages for a subset of individuals, as described below.

The surnames of managers and owners of Limited Liability Companies come from the lists provided by the Bureau of Companies. In order to recover current surname(s) of an individual, I split the name string of each manager and shareholder and separate the first name(s) from the surname(s). As mentioned above, Ecuadorians, as most Spanish-speaking people, usually have two surnames. The first one associated to the father, and the second one to the mother. Since it is also very common for people to have two first names, the typical string in the dataset is organized as follows,

[Surname 1][Surname 2][Name 1][Name 2]

and each piece of the name is separated by spaces. Of course, there are cases in which an individual only has one surname, or only one name. I take into consideration all the possible cases and develop a procedure to accurately separate each surname and name.

I first split each part of the string if separated by a space. The baseline case is two surnames and two first names, as I know then with great certainty that the first two words are surnames and the last two are first names. For the rest of the cases, the algorithm classifies the split parts into surnames or first names given its position in the string and how likely they are to be a surname or a first name, considering the names and surnames in the baseline case. I also consider Spanish conventions of surnames to take into account composed surnames that are not a unique word, like the surname "De la Cruz."

Once I have recovered the surnames, I first compare if the managers share surnames with the owners, and then quantify the magnitude of family involvement in

the management of the firm, measured by the fraction of ownership. The following example illustrates the methodology.

Suppose a firm has three individuals: two shareholders and a CEO, with the corresponding ownership an surnames as described above,

$$\overbrace{\text{Garcia}\quad\quad\text{Sanchez}}^{\text{owner 1: }\quad 40\%},\qquad \overbrace{\text{Perez}\quad\quad\text{Garcia}}^{\text{owner 2: }\quad 60\%}\qquad \overbrace{\text{Perez}\quad\quad\text{Acosta}}^{\text{manager}}$$

In this case, the CEO is related to one of the shareholders as both share the surname "Perez". Additionally, this surname is associated with 60% of the control of the firm, so potentially it could be considered as a Family-Managed firm, as the principal manager is related to the major shareholder of the company.

More generally, for every firm j, let S_o^j be the set of shareholders' surnames, S_m^j be the set of CEO's surnames and $\omega_{i,s}^j$ the ownership of individual i with surname s. To classify firms I use the following cut-off rule,

$$FM_j = \begin{cases} 1 & \text{if } \sum_{s\in\left\{S_o^j \bigcap S_m^j\right\}} \omega_{i,s}^j > \bar{c} \\ 0 & \text{otherwise} \end{cases} \tag{2.1}$$

where the baseline sets $\bar{c} = 50\%$. Under this criteria, the firm of the example will be effectively a Family-Managed firm. The decision to choose this cut-off was done for two reasons. First, it represents simply majority control of the firm, and second, it allows me to compare my estimates with similar findings in the literature (for instance, the classification of Family Ownership and Management in Bloom and Van Reenen (2007) is binary). Nevertheless, I provide several robustness checks of the empirical results linked to this definition by setting different thresholds.

Finally, I do not recover sole-proprietor's surnames, as this type of firms are considered FM since they satisfy the criteria defined in (2.1) given that the ownership of the firm is concentrated in one person, which at the same time is the Legal Representative of the firm (the manager legally accountable for the firm's actions.). Moreover,

this group of firms represent 45% of total firms, and so are an important group to be analyzed. Additionally, research by Bloom and Van Reenen (2007) have suggested that similar to Family-Managed firms, Founder-Managed firms (a category in which Sole Proprietors fall) have poor managerial practices. Nevertheless, I perform several robustness checks to see whether the inclusion of this group changes the results significantly.

True family linkages

To complement the identification of family linkages between managers and shareholders, I also use information of the true linkages using data provided by the Internal Revenue Service of Ecuador (SRI in Spanish). Due to control purposes, the SRI has developed a network of family connections using information provided by the Civil Registry. The core of the network is built on information about parents and couples, if applicable. In principle, once those links are known, the program can map every possible family relationship of any individual.

Unfortunately, this data is incomplete, as the network is on an early phase, so it only includes information of individuals who have renewed their citizens ID to one with a new format[11]. Due to this limitation I use this data more as a complement of the surnames methodology, in particular to identify linkages related to marriage and the corresponding family-in-law, for which there are no shared surnames but a family connection exists.

By using this network data, I complement the classification of FM firms applying the criteria defined in (2.1), and treating the existence of a relation as if

[11]In 2009 the Government changed the format of the national identification to increase its security. However, since it was not mandatory to renew it, many individuals do not appear in these dataset, and show with missing family linkages.

the parts shared a surname. Finally, after firms were classified as Family-Managed or Externally-Managed, I returned the list of firms to the tax office so they could anonymize their IDs in order to merge it with the corresponding tax records.

2.3 Empirical Analysis

In this section I use the methodology described above to classify each firm in the tax records database as family-managed or externally-managed. With that classification at hand, I document new stylized facts about the presence of family-managed firms, the life-cycle differences with respect to externally-managed firms, and the consequent implications for cross-sectional employment and size distribution.

2.3.1 Prevalence of FM firms

Family-Managed firms are the most common type of firms throughout the Ecuadorian economy. They represent around 80% of the total number of firms, with little variation across sectors and over time. Figure 2.3 shows how the distribution of FM firms is similar across the four biggest sectors (in terms of employment), while Figure 2.4 evidences the constant presence of FM firms over time.

Moreover, the prevalence of FM firms seem to be constant across age, although there is a slightly smaller fraction of FM firms in both the youngest and oldest cohorts, as displayed in Figure 2.5. This considerable number of family-managed firms evidences that their presence is a widespread phenomena and depict it as a structural feature of the economy.

To put these numbers in context, I compare the magnitude of the prevalence of FM firms with evidence provided from survey data in Bloom and Van Reenen (2007). Using survey data for medium-sized manufacture firms (firms between 50 and 10,000 workers) the authors present a summary of the prevalence of this type of firms for

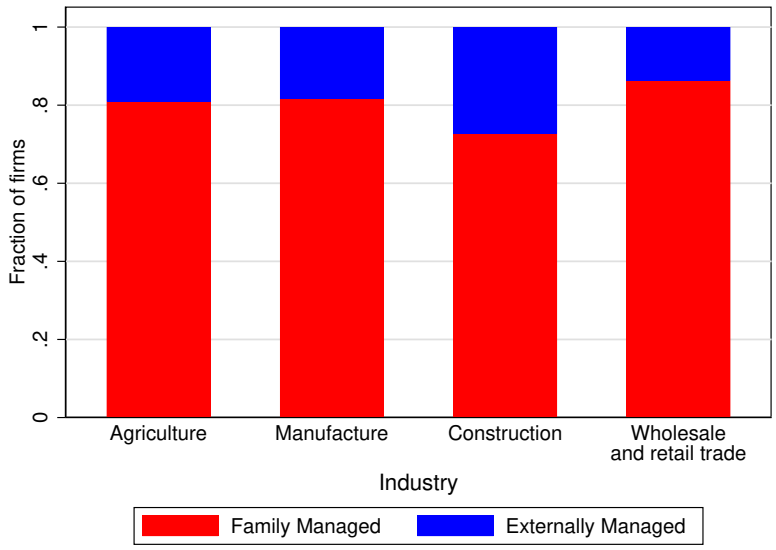


Figure 2.3: Frequency of Family-Managed and Externally-Managed Firms Across Industries

Notes: Firms were classified according to the cut-off rule defined in (2.1) with $\bar{c} = 50\%$. The sample includes the universe of all formal firms in Ecuador, both Limited Liability Companies and Sole Proprietors, except the financial sector, for the period 2009-2016.

a set of countries, including the United States. Although the definition of FM firm from that study is not exactly the same as the one used in this paper (see notes of Table 2.2), it is the most comparable sample.

Table 2.2 shows the fraction of FM firms in Ecuador using the sampling criteria described above accompanied with the cross-country data of Bloom and Van Reenen (2007). As one can see, the presence of this type of firms is higher in Ecuador as compared to all of the other countries. It is around 12 percentage points larger than the UK (the country of the sample with the largest share of FM firms) and more than 2 times larger than the U.S. This is consistent with the idea that the typical form of business organization in a developing country crucially relies on a family network to

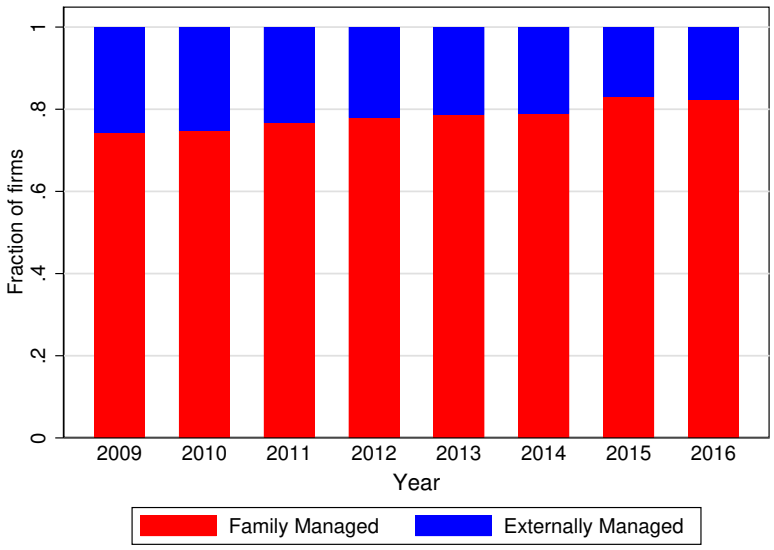


Figure 2.4: Frequency of Family-Managed and Externally-Managed Firms Across Years

Notes: Firms were classified according to the cut-off rule defined in (2.1) with $\bar{c} = 50\%$. The sample includes the universe of all formal firms in Ecuador, both Limited Liability Companies and Sole Proprietors, except the financial sector, for the period 2009-2016.

operate. (La Porta et al. (1999), Bloom et al. (2013))[12].

In order to provide a robustness check and bound the estimates of the participation of FM firms in the economy, Table 2.3 summarizes the prevalence of FM firms according to different criteria. When Sole Proprietors are considered, the classification of FM firms when using different cut-off rules do not change its frequency substantially. The difference between the less restrictive case ($\bar{c} = 25\%$) and the most restrictive one ($\bar{c} = 75\%$) is 13 percentage points. However, when only Limited Liability Companies are considered, the magnitude of this difference goes up to 26 percentage points. The reason, as mentioned above, is that Sole Proprietors account for 45% of the total firms. Due to these differences, I provide several robustness checks in Appendix B for

[12]Developed economies seem to have a larger fraction of family firms (Faccio and Lang (2002)) than traditionally assumed (Berle et al. (1932)), although to lesser extent than developing ones.

Table 2.2: Cross-country Differences in Family-Managed Firms (Manufacture Sector)

Country	Prevalence of Family-Managed Firms (%)
France	36
Germany	12
United Kingdom	35
United States	19
Ecuador	48

Notes. This Table compares the prevalence of Family-Managed firms for different countries. The data for countries different than Ecuador come from Bloom and Van Reenen (2007). For Germany, France, UK and the US I consider a Family-Managed firm to be that where one of the following criteria apply: a) a family is the largest shareholder and the CEO belongs to the family as well; b) a family is the largest shareholder and the CEO is chosen by primogeniture; and c) the CEO is the founder and largest shareholder. The sample includes medium-sized firms (between 50 and 10,000 workers) in the manufacture sector. The sample for Ecuador fits the same criteria for comparison purposes.

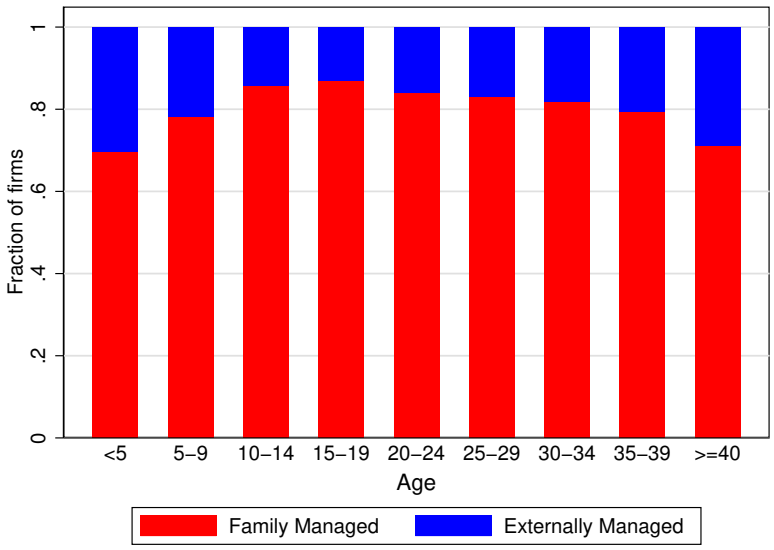


Figure 2.5: Frequency of Family-Managed and Externally-Managed Firms Across Age

Notes: Firms were classified according to the cut-off rule defined in (2.1) with $\bar{c} = 50\%$. The sample includes the universe of all formal firms in Ecuador, both Limited Liability Companies and Sole Proprietors, except the financial sector, for the period 2009-2016.

the results described below.

2.3.2 The Life-cycle of Family and Externally-Managed Firms

In this section I present one of the main empirical exercises of the paper: the life-cycle differences between Family-Managed and Externally-Managed firms. First, in order to compute the life-cycle profiles, I exploit the longitudinal characteristics of the data and estimate the following regression,

$$\log(y_{j,t}) = \alpha + \theta_j + \delta_t + \sum_{a \in A} \sum_{k \in FM_{j,t}} \lambda_{a,k} D_{j,t}^{a,k} + \text{error}_{j,t} \tag{2.2}$$

where $y_{j,t}$ is the outcome of interest (e.g. revenue) of firm j, at time t; θ_j represents a firm fixed effect; δ_t is a vector of time period dummies, while $D_{j,t}^{a}$ is a dummy

Table 2.3: Frequency of Family-Managed Firms

According to Different Criteria

	(1)	(2)	(3)
Cut-off Rule	$\bar{c} = 25\%$ (%)	$\bar{c} = 50\%$ (%)	$\bar{c} = 75\%$ (%)
Includes Sole Proprietors			
Yes (N=150,558)	87	79	74
No (N= 81,345)	77	60	51

Notes. This Table compares the prevalence of Family Managed firms using the methodology defined by (2.1) according to different thresholds. Column (1) uses a cut-off of $c = 25\%$, Column (2) report the baseline case, $c = 50\%$, and Column (3) uses a cut-off of $c = 75\%$.

variable that takes the value of one if the firm's age belongs to the category $a \in A = \{1, 2, \ldots, 20, 21+\}$, and,

$$FM_{j,t} = \begin{cases} 1 & \text{if firm } j \text{ at } t \text{ is FM} \\ 0 & \text{otherwise} \end{cases}$$

I omit the category of age 1, so the results should be interpreted relative to that age. Moreover, this age bin was normalized for each group separately in order to account for differences in the size of entry of both classes of firms. I also collapsed the age above 21 as a single bin because the number of firms in this category starts to be relatively small.

This specification captures age effects in a nonlinear way to account for any possible decrease in the growth rates of a firm over time (Haltiwanger et al. (2013)). It also captures any time-invariant unobserved characteristic of the firm, as it uses the within-variation in the dependent variable $y_{j,t}$[13]. Nevertheless, as it is well-known in the literature, even with the panel structure of the data, this specification cannot separately identify age, cohort and time effects[14]. However, it is parsimonious and has the strength to follow firms over time while absorbing any unobserved characteristic of the firm, including cohort effects. For presentation purposes, all results shown below were transformed from log-points to percentages to facilitate their interpretation.

The age coefficients λ_a, with $a \in \tilde{A} = \{1, \ldots, 20\}$ for each type of firm are plotted in Figure 2.6. The figure shows the life-cycle profiles of both types of firms using revenue as a dependent variable[15]. The results jump-out immediately: at age 20,

[13]Recently Kueng et al. (2014) uses a similar specification to document the forces that drive the life-cycle of firms using data from Canada.

[14]See Deaton (1997) and Schulhofer-Wohl (2018) for a recent discussion about the topic.

[15]I focus on revenue instead of employment due to concerns of the tax authority that their information of employment is misreported, particularly for the years before 2012, while this is not the case for revenue. In Appendix **??** I show the estimates of the life-cycle profile resultant

family-managed firms double in revenues, but externally-managed triple. The faster expansion of EM starts happening early in the life of the firm, as one can see before the age of 5 the patterns start to diverge. Moreover, after age 10, the life-cycle profile of family-managed firms remain flat and do not change for the next 10 years, which is not the case for the externally-managed ones. The latter keep expanding over time, so at the age of 20, in average they are 200% larger than their counterparts at age 1. Instead, by the age 20, family-managed firms look the same as at age 10, and are only 100% times larger than their younger counterparts at age 1.

The life-cycle profiles presented in Figure 2.6 include all the firms in the dataset, which combine entrants and quitters, thus are influenced by a selection effect. For example, FM firms may be more prone to stay in business for some non-pecuniary motive, or instead EM firms could identify easier if it is advantageous to stay in business, so they could shrink faster and exit. One way to address these concerns is to focus on the firm dynamics of firms that have stayed for all the periods data is available. Figure 2.7 shows the results when restricting the sample to fit that criteria. In this case the life-cycle profiles of both type for firms become steeper, however, the size of the gap between FM and EM firm at age 20 remains of the same magnitude as before: EM firms are 230% larger than their younger parts at age 1, while FM firms are 130% larger.

These results suggest that firm dynamics in developing countries are greatly influenced by the disproportionate presence of FM firms, which in average do not expand as they age. To further see this, Figure 2.8 shows the life-cycle profile of all firms in Ecuador, without separating them by the type of management. In this case the age profile closely follows that of FM firms, as at age 20 firms are in average 120% larger

of (2.2) when using employment as a dependent variable, and compare those results with the findings of this section and those reported by Hsieh and Klenow (2014).

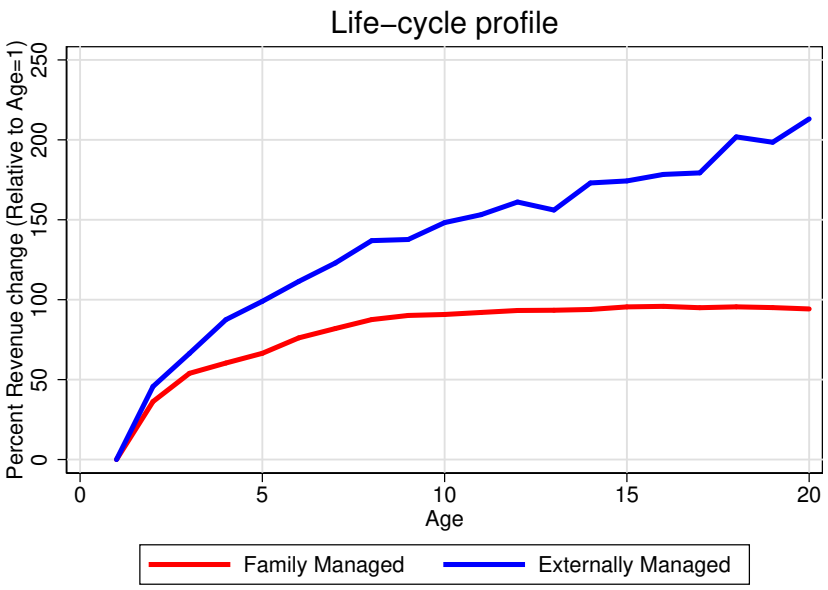


Figure 2.6: Life-cycle Profile of Family-Managed and Externally-Managed Firms

Notes: Estimated by (2.2). Firms were classified according to the cut-off rule defined in (2.1) with $\bar{c} = 50\%$. The sample includes the universe of all formal firms in Ecuador, both Limited Liability Companies and Sole Proprietors, except the financial sector, for the period 2007-2016. Age equal one is the base category. Revenue is Real Revenue, deflated using the CPI.

than younger firms.

Moreover, the gap in the life-cycle profiles between family-managed and externally-managed firms documented here mimics that reported by Hsieh and Klenow (2014) between Mexico and the United States, and hence provides a further explanation of why firm dynamic differ in developing countries: the abundance of family managed firms[16]. Indeed, there are firms in developing countries that expand as they age, however, those are typically the ones not managed by a family member.

[16] Although the mentioned paper uses cross-sectional data, in Appendix B I provide further comparison to their results using a similar methodology, arriving to the same conclusion.

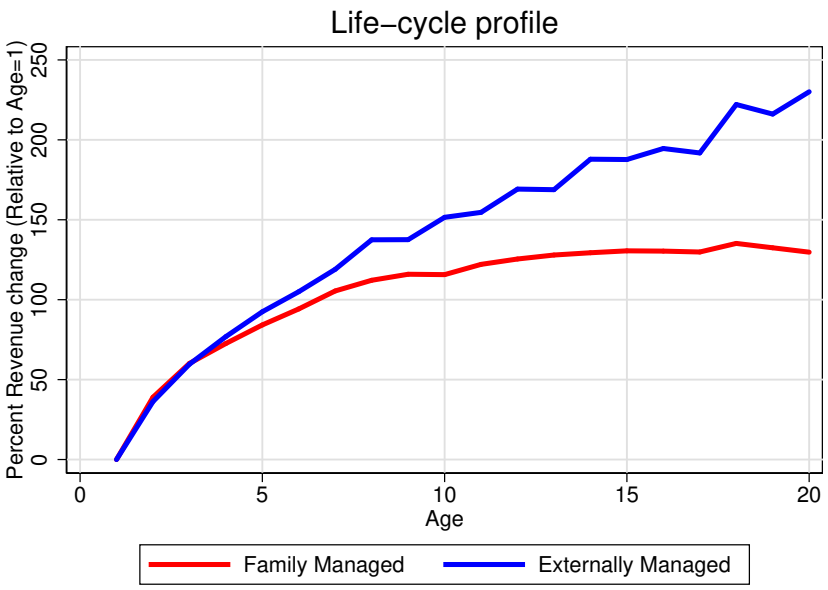


Figure 2.7: Life-cycle Profile of Family-Managed and Externally-Managed Firms (Stayers)

Notes: Estimated by (2.2). Firms were classified according to the cut-off rule defined in (2.1) with $\bar{c} = 50\%$. The sample includes the universe of all formal firms in Ecuador, both Limited Liability Companies and Sole Proprietors, except the financial sector, for the period 2007-2016. Age equal one is the base category. Revenue is Real Revenue, deflated using the CPI.

2.3.3 Cross-sectional Differences

The cross sectional implication of these life-cycle differences is that EM firms are typically larger. Figure 2.9 shows the distribution of firms according to their size. As one can see, larger firms are composed in a great part by externally-managed firms, while small firms are typically concentrated as family-managed firms. Although there are large firms that continue to be family-managed, that case seems to be more the exception than the rule. For instance, the largest categories of employment, which refers to firms with more than 500 workers, are composed in around 80% of EM firms.

Furthermore, these differences in expansion have also non-negligible aggregate

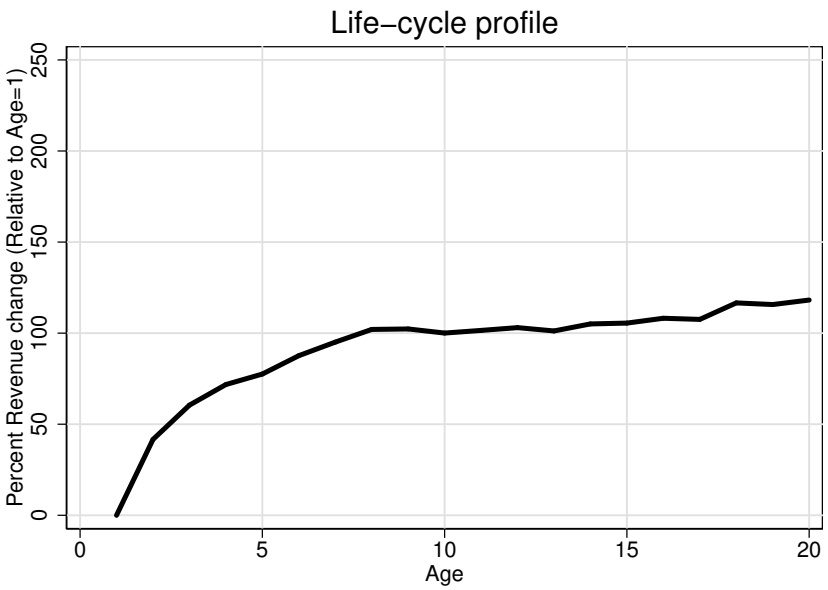


Figure 2.8: Life-cycle Profile of All Firms

Notes: Estimated by (B.1). The sample includes the universe of all formal firms in Ecuador, both Limited Liability Companies and Sole Proprietors, except the financial sector, for the period 2009-2016. Age equal one is the base category. Revenue is Real Revenue, deflated using the CPI.

implications due to the importance of each type of firms for the economy. In that sense, even though FM firms are disproportionately represented, they account for a smaller fraction of total revenue and employment. Table 2.4 shows that family-managed firms, despite being 80% of firms, only represent 49% of total employment and 41% of total revenue, consistent with the previous finding that EM firms are typically larger. Hence, forces that difficult the delegation of management to an outsider have critical aggregate implications, as they push towards a larger fraction of firms to keep the administration within the family, which in turn contribute to the existence of smaller firms that also do not expand as their age.

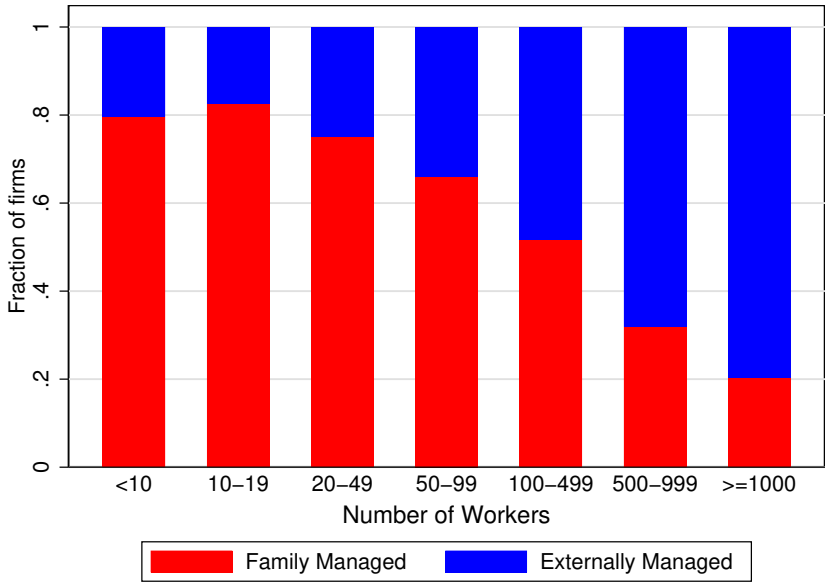


Figure 2.9: Composition of the Size Distribution of Firm of Family-Managed and Externally-Managed

Notes: Firms were classified according to the cut-off rule defined in (2.1) with $\bar{c} = 50\%$. The sample includes the universe of all formal firms in Ecuador, both Limited Liability Companies and Sole Proprietors, except the financial sector, for the period 2009-2016.

2.3.4 Switchers

Why do Family-Managed firms grow less? This is a point that will be continuously discussed throughout the paper, but several pieces suggest a mix of bad managerial practices (Bloom and Van Reenen (2007)) likely to be related to the supply of lower managerial talent from family members. However, it could simply reflect some further unobserved heterogeneity beyond management.

If instead firms can face rapid growth when delegating to outsiders, there should be gains in revenue or employment when firms switch. Precisely, I address this issue by studying what happens to firms when they switch and become EM. To do so I

Table 2.4: Total Participation in the Economy of Family and Non-Family Managed Firms

	(1) Employment (%)	(2) Revenue (%)
Family-Managed	49	41
Externally-Managed	51	59

Notes. This Table reports the share of Employment and Revenue of both Family-Managed and Externally-Managed firms for the overall economy.

restrict the sample of switchers to those firms that have switch to be an EM firm in such a way that they remain like that for the rest of the studied period. This sample restriction enables to clearly distinguish switchers, instead of simply reflecting spurious changes due to movements around the FM baseline threshold. This transition, although infrequent in the studied period, sheds light about the gains of expansion due to outside management.

The specification I estimate for this case-event study is the following,

$$\log(y_{j,t}) = \alpha + \theta_j + \delta_t + \sum_{\tau=1}^{\bar{\tau}} \beta_\tau D_j^\tau + \text{error}_{j,t} \tag{2.3}$$

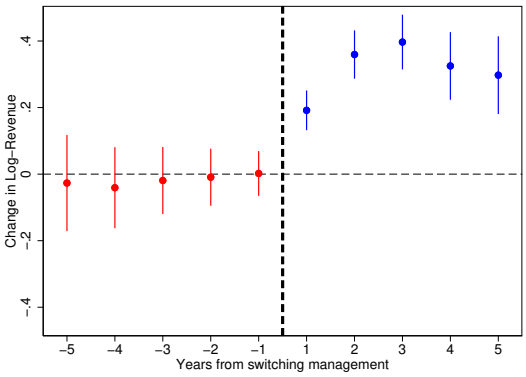
where $y_{j,t}$ is the outcome of firm j, at time t; θ_j represents its fixed effect; δ_t is a vector of time period dummies, while $D_{j,t}^\tau$ is a dummy variable that takes the value of one

if the firm's age belong to the category $a \in \bar{\tau} = \{-5, -4, \ldots, 0 \ldots 5\}$. The coefficient normalized is that of the last year a firm was FM, so the results should be interpreted relative to that period. The results of this exercise are reflected in Figure 2.10a and Figure 2.10c, for the outcomes of revenue and employment, respectively. The figures report the estimates of β_τ for the bins described above. In average, firms show an increase in 17% in revenue and 10% in employment one year after the switch, which continues to increase for the subsequent years, although decay a little 5 years afterwards for the case of revenue. Interestingly, there seem to be not much of a trend in a firm's revenue or employment before the switching decision.

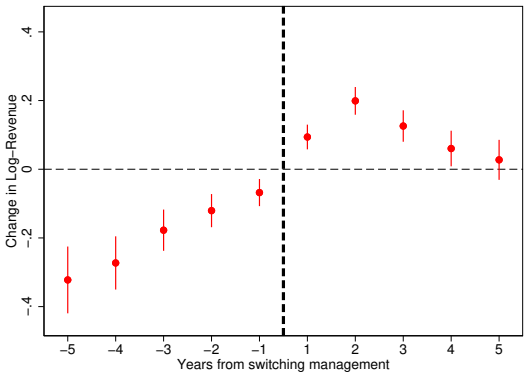
To see the opposite case, I consider firms that have changed management, but continued to operate as FM firm during the studied period. In order to see what occurs in the FM-FM transition, I estimate (2.3) restricting the sample to firms that have changed the top manager only once, but keep management within the family. The results are plotted in Figure 2.10b and Figure 2.10d for revenue and employment, respectively.

There are three observations to be noticed in this case. First, there are also positive gains after changing a manager, but the size of these gains are smaller than those reported when the transition is to an external manager. For instance, the gains in revenue are around 10% and in employment 5% relative to the last year a different family-manager operated. Second, those gains vanish faster compared to the FM-EM transition, as 5 years later the gains fade away for the case of revenue. Lastly, there seems to be a trend before switching, which may suggest a premeditated family transition. This could be the case if the CEO plans its retirement, so reorganizes the business in such a way to leave it "ready" for the successor.

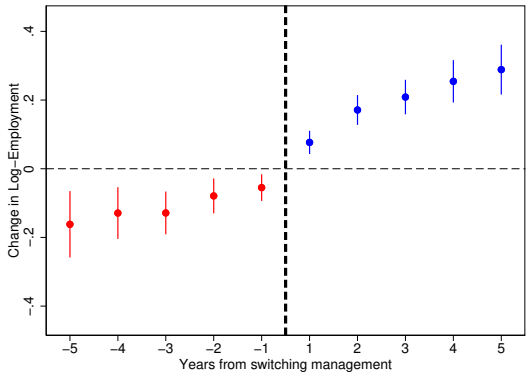
Although I cannot tell the motives to switch from EM to FM or to keep the business in the family, the results are consistent with Bennedsen et al. (2007), who



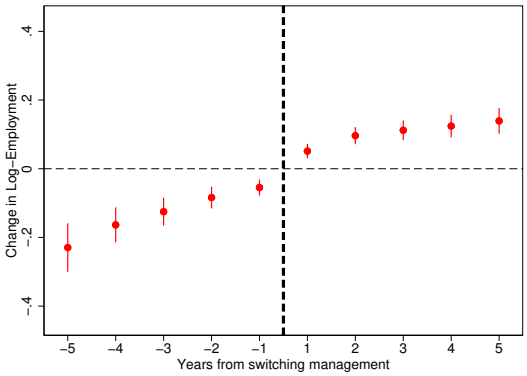
(a) FM-EM transition: revenue



(b) FM-FM transition: revenue



(c) FM-EM transition: employment



(d) FM-FM transition: employment

Figure 2.10: Case Event Study: Switchers

Notes: Figures 2.10a and 2.10c report the case-event estimates of (2.3) for firms that switch from FM to EM, while Figures 2.10b and 2.10d show the estimates for firms that changed management, but remain as a FM firm, using revenue and employment respectively as outcomes of interest. Firms were classified according to the cut-off rule defined in (2.1) with $\bar{c} = 50\%$. The FM-EM switchers are defined as those firms that (i), change from being FM to EM and (ii), stay as EM for the rest of the periods. Instead, the FM-FM are defined as those firms that (i), change management and (ii), continue as FM for the rest of the periods. The sample includes switchers in every sector, except the financial one, for the period 2007-2016. All results were normalized relative to the last year before a firm switch. Revenue is Real Revenue, deflated using the CPI and Employment is the number of workers formally registered computed from tax returns.

provided casual estimates that family successions have a negative effect on a firm's performance, and stress the value of professional outside managers to operate a family-owed firm. However, the positive improvement in terms of revenue and employment for firms that change management but continue to operate as a FM firm suggest that perhaps an indistinct rotation in management could lead to an increase a firm's performance, at least in the short-run.

Another interpretation is that not all family successions necessarily lead to a lower performance of the firm. For instance, the literature has stressed the negative effects of primogeniture succession (Bennedsen et al. (2007)) and management (Bloom and Van Reenen (2007)). However, the latter also provides evidence that companies that select CEO from family members other than the first-born child are no worse managed than other firms. Since I cannot differentiate the type of family bondage between top managers, I confound all type of family successions, so it leaves open the possibility that some transitions within the family may actually lead to positive effects for the firm. Moreover, the results presented here could also be thought as a lower bound of family-management transitions that could be larger for specific family successions, like the primogeniture one.

Overall, the evidence presented in this section supports previous findings that Family-Managed firms perform worse than their counterparts, but document a new fact: they also do not expand as their age. Given that these firms seem to represent the majority of firms in a typical developing country, understanding the forces that limit firms to keep management inside the family and the implications for economic development. The next section advances in that direction by studying and quantifying how distortions in the delegation environment could induces a larger mass of firms to keep operating only with their own managerial endowment.

2.4 Model

I build a general equilibrium model of firm dynamics à la Hopenhayn (1992) to analyze how different forces interact in the decision of a firm to hire external management, and in particular how distortions in the contractual environment affect that decision. The key modification of the model is to allow firms to optimally choose to be either managed by the family member or instead hire external management. The parameters of the model are identified using the data and several of the new stylized facts reported in the previous section. I then use the calibrated model to quantify the aggregate implications of changes in the contractual environment. The details of the model are described below.

2.4.1 Firms

There is a continuum of ex ante identical entrepreneurs who may start a firm and produce a homogenous final product. Production requires managerial input (h) and labor (n), and is subject to idiosyncratic productivity shocks (z) every period. The shocks are independent across firms but come from the same stochastic process. Every entrepreneur is endowed with $\bar{h}$ managerial efficiency units that cannot be traded in the market, which is interpreted as their family endowment of managerial talent. Managerial input, however, could also be hired in the market at a compensation rate w_h. If the entrepreneur decides to run a firm using only their managerial endowment, $\bar{h}$, the firm is labeled as a Family-Managed (FM) firm, in which case output is given by

$$y_{FM} = \left(z^{\gamma}\bar{h}^{1-\gamma}\right)^{1-\alpha} n^{\alpha}$$

Instead, if the entrepreneur rents managerial services in the market, that firm is considered as an Externally-Managed (EM) one. These external managerial services

are added to the initial endowment, and so expand the scope of production. In that case output is given by,

$$y_{EM} = \left(z^{\gamma}(\bar{h} + \phi h)^{1-\gamma}\right)^{1-\alpha} n^{\alpha}$$

where $0 < \alpha < 1$ represents the returns to labor, while $1 - \gamma$ $(0 < \gamma < 1)$, adjusted by $(1 - \alpha)$, governs the contribution of management to production. The parameter γ plays a critical role in the dynamics of the firm, since an expansion in productivity z will need to be paired eventually with an increase in managerial services. I discuss its implications for the life-cycle of the firm below. Note also that the difference between both types of firms is modeled as a *technological* one, in line with Bloom et al. (2016). Finally, the parameter ϕ represents the managerial efficiency of external management, which reflects in a reduced-form way the contractual environment, and is explained in detailed below.

Since firms face idiosyncratic shocks z every period t, there is a dynamic environment with the following timing. Every incumbent firm enters at period t with state z_{t-1} and first is hit by an exogenous shock with probability λ that forces it to exit, in which case it will receive zero profits in all future periods. If the firm stays in the market it receives a new shock z_t. Then the entrepreneur decides to run the firm as either a FM or an EM, and finally it hires labor and managerial services accordingly. Entrants on the other side, pay an entry cost, c_e, and draw a productivity shock from a distribution that is also *i.i.d*, and start operating immediately with the drawn z, so from that moment their timing is the same as incumbents.

The more natural way to think about this model is to start backwards, thinking first the problem at the last stage, which is the decision of choosing the optimal factor demands given the type of firm, to then discuss the optimal choice of either being a FM or EM firm, to finally study the exit/entry dynamics. The discussion below

follows that order.

Static problem

At the last stage, an entrepreneur that has chosen to run a FM firm, in the sense that it only uses its managerial endowment, $\bar{h}$, selects labor to maximize profits taking the price of labor w_n as given

$$\pi_{FM}(z; w_n) = \max_n \left\{ \left(z^\gamma \bar{h}^{1-\gamma} \right)^{1-\alpha} (n)^\alpha - w_n n \right\}$$

Note in this case the firm does not have to pay for the management it uses, since it operates with the given endowment. On the other side, an entrepreneur that has decided to delegate and run the firm as EM chooses to hire both labor and managerial services, h, compensated at rate w_h to maximize profits taking prices as given

$$\pi_{EM}(z; w_n w_h) = \max_{n,h>0} \left\{ \left(z^\gamma (\bar{h} + \phi h)^{1-\gamma} \right)^{1-\alpha} (n)^\alpha - w_n n - w_h h \right\}$$

However, due to contractual frictions, hiring management services implies to devote part of the managerial endowment to monitor them. These contractual frictions reduce the managerial efficiency and are captured by the parameter ϕ. This is a reduced-form way to account for a variety of distortions that mitigate external delegation in developing countries.

Akcigit et al. (2016) uses a similar specification interpreting it as a parameter that is country-specific, related to the contractual environment, the level of technology available or the level of financial development. It is also consistent with the evidence presented in Section 2.2.2 about the concentration of ownership and its interpretation as symptom of a weak rule of law, and with the fact discussed in Section 2.3 that the presence of FM firms is similar across sectors, suggesting an overall deficient contractual environment.

The corresponding demand of labor of a FM firm is given by

$$n = z^{\gamma}\bar{h}^{1-\gamma}\left(\frac{\alpha}{w_n}\right)^{\frac{1}{1-\alpha}} \tag{2.4}$$

While the demand for labor of an EM firm is

$$n = z\left(1-\alpha\right)^{\frac{1-\gamma}{\gamma}}\left(\frac{\phi(1-\gamma)}{w_h}\right)^{\frac{1-\gamma}{\gamma}}\left(\frac{\alpha}{w_n}\right)^{\frac{\gamma+\alpha(1-\gamma)}{\gamma(1-\alpha)}} \tag{2.5}$$

And the demand of managerial services of an EM is

$$h = \frac{z}{\phi}\left(1-\alpha\right)^{\frac{1}{\gamma}}\left(\frac{\phi(1-\gamma)}{w_h}\right)^{\frac{1}{\gamma}}\left(\frac{\alpha}{w_n}\right)^{\frac{\alpha}{\gamma(1-\alpha)}} - \frac{\bar{h}}{\phi} \tag{2.6}$$

Given the optimal demand for inputs, it can be shown that substituting (2.4) into the profit function of a FM firm leads to,

$$\pi^{*}_{FM}(z, w_n) = z^{\gamma}\left(1-\alpha\right)\left(\frac{\alpha}{w_n}\right)^{\frac{\alpha}{1-\alpha}}\bar{h}^{1-\gamma} \tag{2.7}$$

Equivalently substituting the optimal demand of labor (2.5) and managerial services (2.6) in the profit function of a EM firm gives the following expression,

$$\pi^{*}_{EM}(z, w_n, w_h) = z\gamma\left[\left(1-\alpha\right)\left(\frac{\alpha}{w_n}\right)^{\frac{\alpha}{(1-\alpha)}}\left(\frac{\phi(1-\gamma)}{w_h}\right)^{1-\gamma}\right]^{\frac{1}{\gamma}} + \frac{w_h\bar{h}}{\phi} \tag{2.8}$$

As can be observed in equations (2.7) and (2.8), $\pi_{FM}(z; w_n)$ is concave in z, while $\pi_{EM}(z; w_n, w_h)$ is linear for EM firms. This comes from the fact that outside delegation allow firms to scale up production, overcoming the diminishing returns of operating uniquely with their fixed endowment. This property generates an endogenous unique cutoff which defines the extensive margin of the firm's managerial decision. Below certain productivity level it is optimal for entrepreneurs to only uses their family managerial endowment to run the firm, while above this threshold it is optimal to switch and hire external management.

Formally, using (2.7) and (2.8) one can show that the productivity level at which it becomes optimal to hire external management is,

$$z^*(w_n, w_h) = \left[\left(\frac{1}{1-\alpha}\right)\left(\frac{w_n}{\alpha}\right)^{\frac{\alpha}{1-\alpha}}\left(\frac{w_h}{\phi(1-\gamma)}\right)\right]^{\frac{1}{\gamma}}\bar{h} \tag{2.9}$$

Note that the expression above is an equilibrium object which additionally to the parameters of the model, depends on both the equilibrium price of labor w_n and the equilibrium compensation rate of managerial services, w_h. A graphical representation of the cut-off rule is shown in Figure 2.11. Before certain threshold, firms will optimally demand no outside managerial services and operate just as a FM firm. Above the threshold, however, it is no longer optimal for them to continue to be FM, as higher levels of productivity require higher inputs of managerial services that are not fulfilled by the initial endowment, due to the complementarity between both.

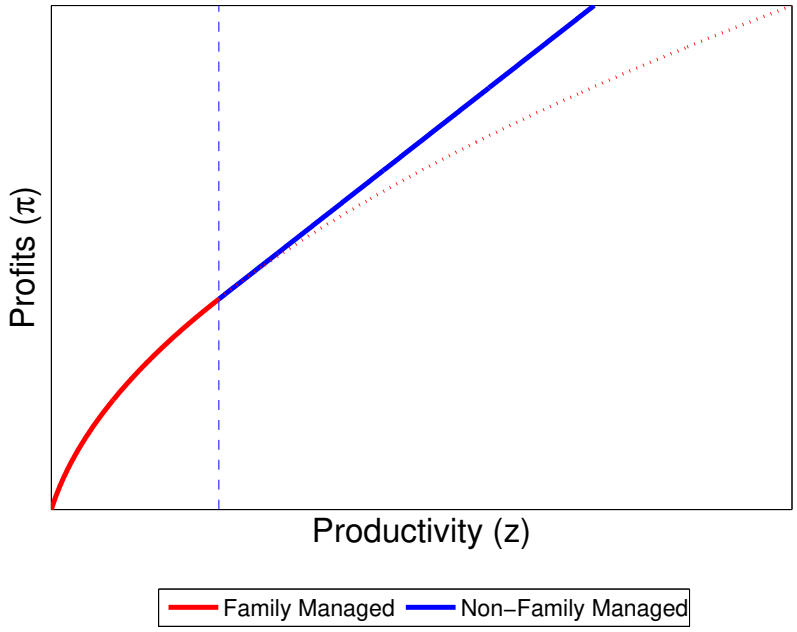


Figure 2.11: Profit's Function for Family-Managed and Externally-Managed Firms

Notes: Profit's function according to equations (2.7) and (2.8) respectively. The vertical dashed line represents the productivity level z^* that induces firms to switch from FM to EM.

Now, observe that this switching threshold critically depend also on the param-

eters of the model, including the magnitude of the managerial efficiency, ϕ and the endowment of managerial talent, $\bar{h}$. In particular, note that everything else fixed, it is decreasing in ϕ. That is, a reduction in the managerial efficiency, caused for example by a worsen in the contractual environment, increases the productivity level that induces a firm to switch, making the presence of FM firms more abundant. Only firms with really high productivity draws do the transition and hire external managerial services. Instead, when there are improvement in the managerial efficiency firms are more likely to delegate outside management as the benefits of doing so increase. Note also that the equilibrium cutoff is increasing in $\bar{h}$ and w_h, reflecting the intuition that both higher family managerial endowment and higher manager's compensation reduce the incentives of firms to hire external management.

Furthermore, notice from (2.9) that even when the managerial efficiency is very high, there will still exist a positive fraction of firms being FM. This speaks to the fact that even in countries with good contractual environment, like the United States, we observe a non-negligible fraction of firms being family-managed. It follows from the fact that $\bar{h}$ managerial efficiency units comes for free. As low-productive firms do not need to rely on outside management, it is optimal for them to simply use their family managerial talent to run the firm. Contractual distortions affecting the managerial efficiency more critically affects the marginal firm which is close to switching. Potentially highly productive firms, in an otherwise better contractual environment, can't expand and grow at higher rates since they have to keep relying on their family endowment to run the firm instead of hiring better suited managerial talent in the market.

Consequently, the parameter governing the managerial efficiency, ϕ, together with the endowment $\bar{h}$, will be critical in the calibration, as captures several forces that shape delegation, like the contractual environment, and is also policy relevant in the

sense that improvements in the rule of law or on the enforcement of contracts have direct implications for the expansion of firms via increasing the incentives to delegate and scale up production. In that sense, having information about the share of firms that delegate is critical.

Finally, given the endogenous cut-off, z^*, one can characterize the optimal demand of managerial services, and hence, the decision to be either a FM or EM firm, as

$$h(z; w_n, w_h) = \begin{cases} 0, & \text{if } z < z^* \\ \\ \frac{z}{\phi}(1-\alpha)^{\frac{1}{\gamma}} \left(\frac{\phi(1-\gamma)}{w_h}\right)^{\frac{1}{\gamma}} \left(\frac{\alpha}{w_n}\right)^{\frac{\alpha}{\gamma(1-\alpha)}} - \frac{\bar{h}}{\phi}, & \text{if } z \geq z^* \end{cases} \tag{2.10}$$

Analogously, the optimal labor demand of a firm will be given by,

$$n(z; w_n, w_h) = \begin{cases} z^{\gamma} \bar{h}^{1-\gamma} \left(\frac{\alpha}{w_n}\right)^{\frac{1}{1-\alpha}}, & \text{if } z < z^* \\ \\ z(1-\alpha)^{\frac{1-\gamma}{\gamma}} \left(\frac{\phi(1-\gamma)}{w_h}\right)^{\frac{1-\gamma}{\gamma}} \left(\frac{\alpha}{w_n}\right)^{\frac{\gamma+\alpha(1-\gamma)}{\gamma(1-\alpha)}}, & \text{if } z \geq z^* \end{cases} \tag{2.11}$$

Entry/Exit dynamics

The remaining stage to be considered is that which deals with the entry and exit dynamics. Recall firms face an idiosyncratic productivity shock every period, which I assume follows a first-order Markov process characterized by the conditional distribution $F(z_{t+1} \mid z_t)$, and is *i.i.d* across firms. Accordingly, given a shock z, the Bellman equation of an incumbent is,

$$\mathrm{V}(z, w_n, w_h) = \max \left\{ \max \left\{ \pi_f^*, \pi_n^* \right\} + \beta(1-\lambda) \int_{z'} \mathrm{V}(z', w_n, w_h) \mathrm{dF}(\mathrm{d}z' \mid z) \right\}$$

where λ denotes the probability of exit and both π_f^* and π_n^* are the optimal profits function of FM and EM firms defined by (2.7) and (2.8) respectively. This Bellman

equation captures today's benefit to operate, accompanied by the expected value of tomorrow's return given the state z, discounted by β, and the probability to survive to the next period, $1 - \lambda$.

Entrants, on the other side, pay an entry cost before knowing their initial productivity, which is drawn from the distribution $\nu(z)$ that is also $i.i.d$ across firms. Since conditional on paying the entry cost, c_e, the problem of an entrant looks the same as that of an incumbent, the present discounted value of a potential entrant is

$$V^e = \int V(z, w_n, w_h)\nu(z)dz - c_e \tag{2.12}$$

so in an equilibrium with firm entry, V^e will be zero.

Finally, the evolution of the distribution of firms is characterized by

$$\mu'(Z_0) = (1 - \lambda) \int_{z' \in Z_0} \mu(Z_0)dF(z' \mid z) + M\nu(dz)$$

where the distribution of firms tomorrow will be composed by the fraction of firms that survive, $1 - \lambda$, plus the distribution of entrants $\nu(z)$, accommodated by the corresponding mass, M.

2.4.2 Households

There is a unit measure of households that supply both labor and managerial services and collect profits, if any, from firms. Similar to Akcigit et al. (2016), I simplify the labor supply problem by assuming individuals are endowed with one efficiency unit of production labor, and h units of managerial talent which is drawn from a Pareto Distribution,

$$P(h > h_0) = \left(\frac{\theta - 1}{\theta}\mu_h\right)^{\theta} h_0^{-\theta} \tag{2.13}$$

where the parameter μ_h represents the average level of managerial talent relative to workers, and can be thought as the overall human capital of external managers. The parameter θ instead controls the dispersion in managerial talent. Accordingly, given labor rate w_n, managerial rate w_h and endowments, an individual chooses to be a manager by simply following a comparative advantage rule,

$$\underbrace{w_h h}_{\text{Manager's compensation}} \geq \underbrace{w_n}_{\text{Worker's compensation}}$$

Therefore, the fraction of managers in the economy is,

$$\begin{aligned} P(w_h h > w_n) &= P(h > \frac{w_n}{w_h}) \\ &= \left(\frac{\theta-1}{\theta}\mu_h\right)^{\theta} \left(\frac{w_h}{w_n}\right)^{\theta} \end{aligned} \tag{2.14}$$

And the total supply of managerial efficiency units is,

$$H^s = \left(\frac{\theta-1}{\theta}\mu_h\right)^{\theta} \left(\frac{w_h}{w_n}\right)^{\theta-1} \frac{\theta}{\theta-1} \tag{2.15}$$

Note the supply of managerial services is increasing in the relative compensation, with elasticity $\theta - 1$, which is positive as long as $\theta > 1$.

Finally, the supply of labor is given by,

$$N^s = 1 - \left(\frac{\theta-1}{\theta}\mu_h\right)^{\theta} \left(\frac{w_h}{w_n}\right)^{\theta} \tag{2.16}$$

2.4.3 Equilibrium

I focus on the stationary competitive equilibrium. A stationary competitive equilibrium in this economy is a list $\{\mu^*, M^*; w_n^*, w_h^*\}$ s.t,

- Labor market clears,

$$N^s(\mu^*, M^*; w_n^*, w_h^*) = \int n(z; w_n, w_h) d\mu(z)$$

- Manager's market clears,

$$H^s(\mu^*, M^*; w_n^*, w_h^*) = \int h(z; w_n, w_h) d\mu(z)$$

- There is an invariant distribution of firms,

$$\mu^* = T(\mu^*, M^*, w_n^*, w_h^*)$$

- The free entry condition holds,

$$c_e \geq \int V(z, w_n^*, w_h^*) \nu(z) dz$$

Further details of the equilibrium are described in Appendix B.

2.4.4 Life-cycle of Firms

The static decision of the firm to delegate has an implication to the the life-cycle of the firm. To see this, denote the subindex a as the age of the firm and suppose z evolves deterministically over time, and consequently, over age as well. Then, a FM firm with age a produces

$$y_{FM,a} = z_a^{\gamma} \left(\frac{\alpha}{w_n} \right)^{\frac{\alpha}{1-\alpha}} \left(\bar{h} \right)^{1-\gamma}$$

while output of an EM with age a is,

$$y_{EM,a} = z_a \left(\frac{\alpha}{w_n} \right)^{\frac{\alpha}{1-\alpha}} \left(\frac{\bar{h}}{z^*} \right)^{1-\gamma}$$

where the last expression uses the optimal cut-off (2.9). Let g_z be the growth rate of z conditional on survival. Therefore, the growth rate of an incumbent FM firm at age a is,

$$
\begin{aligned}
g_{FM,a} &\approx \log(y_a) - \log(y_{a-1}) \\
&= \gamma \left(\log(z_a) - log(z_{a-1})\right) \\
&\approx \gamma g_z
\end{aligned}
\tag{2.17}
$$

equivalently, the growth rate of a EM is,

$$
g_{EM,a} \approx g_z \tag{2.18}
$$

Hence, as discussed above, γ has an implication for the life cycle of firms. The growth rate of both FM and EM firms will depend on the growth rate of z. However, for any growth rate, g_z, the growth rate of a FM firm will be constrained by γ, while an EM firm will grow at the same rate of productivity. Therefore, distortions that reduce the efficiency of external management puts more weight on firms that cannot fully exploit the complementarities with productivity, and so grow at a slower rate.

2.4.5 Aggregate Productivity

The presence of Family-Managed firms in has an implication to the overall productivity of the economy. To see this, observe that one can write aggregate output as,

$$
Y = \left(\frac{\alpha}{w_n}\right)^{\frac{\alpha}{1-\alpha}} \bar{h}^{1-\gamma} \left[\int_{z_-}^{z^*} z^{\gamma} d\mu(z) + \int_{z^*}^{z_+} {z^*}^{\gamma-1} z d\mu(z)\right]
$$

where the expression in brackets reflects a weighted sum of firm's productivities, which can be thought as the aggregate productivity of the economy. Since externally-managed firms can scale up production, the sum of the productivities of this group enters in a linear way. This is not the case for family-managed firms, whose sum of productivities is affected by the parameter γ. Hence, the composition of firms in the

economy determines endogenously the aggregate productivity, as the cut-off level z^* changes the weights towards the left or right side of the distribution of productivities.

2.5 Calibration Strategy

To simplify the analysis and emphasis the role of family management on firm dynamics, I assume a firm's productivity, z, evolves deterministically at a growth rate g_z, and discretize accordingly the grid points. Note that this process could also be written as a Markov one with $\Pr(Z_{t+1} = z_{t+1} \mid Z_t = z_t) = 1$, so it reduces the uncertainty of tomorrow's outcome to only the probability of survive, $1 - \lambda$. Furthermore, this specification also induces to a stationary distribution given the assumption of exogenous exit.

I also assume that the distribution of initial draws of productivities, $\nu(z)$, is Pareto, with scale parameter z_m that is normalized to 1, shape parameter a. This parametric assumption is made since the entrants size in the data is well approximated by a Pareto distribution. With that in mind, the model described so far has 11 parameters:

$$\left\{\beta, \lambda, \alpha, g_z, \gamma, \bar{h}, \phi, a, \mu_h, \theta, c_e\right\}$$

Some of the parameters are standard, while others are to some extent new and related to management. In terms of identification, the first five parameters are calibrated separately, and the remaining six, $\bar{h}, \phi, a, \theta, \mu_h, c_e$, are jointly calibrated by targeting several moments from the data, including the new stylized facts documented in this paper. I consider a period to be one year to use the same frequency as the data.

Regarding the calibration of the standard parameters, I first calibrate the discount rate β to match a real interest rate of 5% typically used in the literature. The probability of exit, λ, is chosen to match the mean firm exit rate for the years 2009-

2016. The parameter α is the labor's share of income, and is assumed to be 2/3, following standard estimates.

As discussed in Section 2.4.4, the parameters g_z and γ govern the life-cycle of the firm, and so are chosen to match the growth rate of EM and FM firms. I chose g_z to match the relative size of an EM firm at age 20 respect to an entrant of the same type, while γ is chosen to match the gap between FM and EM firms at the same age, as documented in Section 2.3.2.

Finally, each of the last parameters are jointly calibrated by targeting some specific moments under the procedure I describe below. First, recall that there is a unit measure of households, so the average size of firms is given by

$$\text{Average size} = \frac{(1 - F_H)}{M/\lambda} \tag{2.19}$$

where F_H is the fraction of *external* managers in the economy. So using data about the fraction of external managers, average size and exit rate λ, I pick the mass of entrants that is consistent with the expression above. Then, for a given entrants distribution, $\nu(z)$ and mass of entrants, M, I compute the stationary distribution of firms, $\mu(z)^*$, and pick the productivity threshold, z^*, that gives the share of Family-Managed firms that is consistent with that observed in the data.

I then use the equilibrium condition in the labor market to compute $\bar{h}$. To see this, note that one can rewrite the equilibrium condition in the labor market as follows

$$1 - F_H = \left(\frac{\alpha}{w_n}\right)^{\frac{1}{1-\alpha}} \bar{h}^{1-\gamma} \left[\int_{z_-}^{z^*} z^{\gamma} d\mu(z) + \int_{z^*}^{z_+} {z^*}^{\gamma-1} z d\mu(z)\right] \tag{2.20}$$

so I normalize $w_n = 1$ and obtain the value of $\bar{h}$. I then use the equilibrium condition in the managerial market to get the parameter that governs the efficiency of external managers, ϕ. Since the equilibrium condition in the managerial market can be written as

$$(F_H)^{\theta-1}\mu_h = \frac{\bar{h}}{\phi}\int_{z^*}^{z_+}\left(\frac{1}{z^*}-\frac{1}{z}\right)z d\mu(z) \tag{2.21}$$

one can get ϕ using for a given μ_h, so I normalize this parameter[17]. This normalization does not affect the counterfactuals in the following section, since for different values of μ_h, ϕ simply escalates. I then recover the equilibrium managerial rate, w_h, that is consistent with (2.9) and adjust a to ensure market clears. I also adjust θ in such a way that targets the share of output of family-managed firms in the economy. The cost of entry, c_e is finally obtained from the free entry condition (2.12).

2.5.1 Results

The results of the calibration exercise are presented in Table 2.5. Column (3) describe the interpretation of each parameter, as discussed above, while column (2) show their values. Although some of the parameters were jointly calibrated, column (5) shows the key moments targeted by each of them as explained in the previous section. I discuss briefly the magnitude and implications of some of these parameters.

First, the growth rate, g_z, is of 0.06, as it was chosen to match the life-cycle of EM firms, so in average, a firm of this type that has survived 20 periods will be close to 3.2 times (1.06^{20}) larger than an EM entrant. The value of γ is 0.50, as it targets the fact documented in this paper that family-managed firms grow as half as quickly as those managed by a non-family member. So for the growth rate of z of 0.06, a FM that have survived 20 periods as such will be in average close to 1.8 times (1.03^{20}) as large than the typical FM entrant. Interestingly, note that $(1-\gamma)(1-\alpha)$ also represents the importance of managerial input in the production function, which

[17]This parameter is also calibrated for India by Akcigit et al. (2016), but comparing the differences in human capital between immigrants before and after coming to the United States using a similar methodology as Hendricks and Schoellman (2018).

given the parameters above, is 0.17. This means that 17% of the returns to scale comes from management. This is a similar value as that reported Bloom et al. (2016), which comes from their experimental work on the importance of management (Bloom et al. (2013)).

On the other side, columns (5) and (6) in Table 2.5 display the moments generated by the model and compare them to their empirical counterparts in the data. As one can observe, the model match relatively well the targeted moments. Moreover, Figure 2.12a show the life-cycle of both type of firms from the simulated model. Although the shape of the life-cycle is not captured, the relative gap between both types is, showing at age 20 that EM are in average 220% larger than their counterparts, while FM are in average 180% larger than the typical FM entrant. Figure 2.12b instead display the simulated life-cycle of all the firms. So similar to what we observe on the data and is reported in Section 2.3.2, the model shows that the abundance of FM firms implies that when seen altogether, the overall life-cycle profile of firms appears to be flat.

2.6 Changes in the Managerial Efficiency

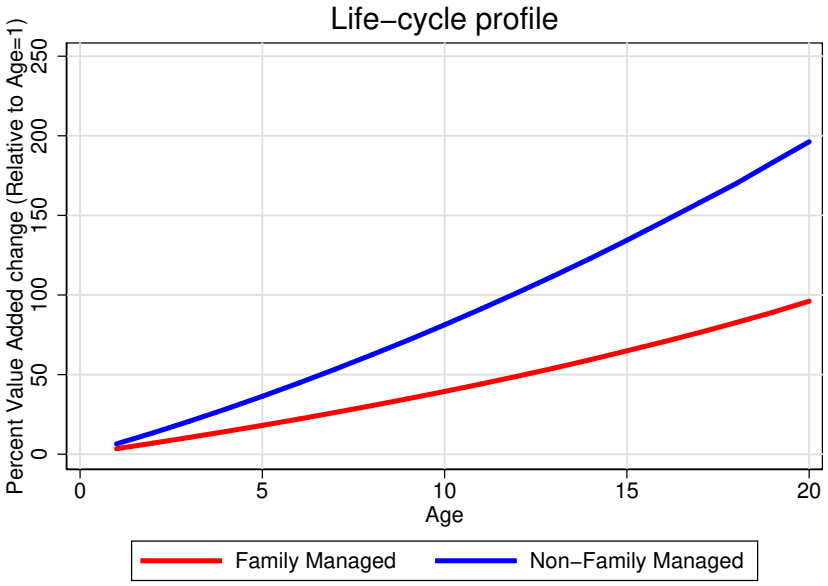
Using the calibrated version of the model, I quantify the aggregate implications of changing the size of the managerial efficiency, ϕ, and its consequences for the overall life-cycle of firms. As discussed above, from a policy perspective this could be thought as an improvement in the contractual environment or reforms looking to strength the rule of law. The first set of experiments is to increase ϕ in such a way that the fraction of FM firms reduce to 0.75, 0.60 and 0.40 respect to the baseline estimate. The results are shown in Table 2.6.

The first column of the Table reports the baseline aggregates given the parametrization described above. Output in this case was normalized to 100 to facilitate the

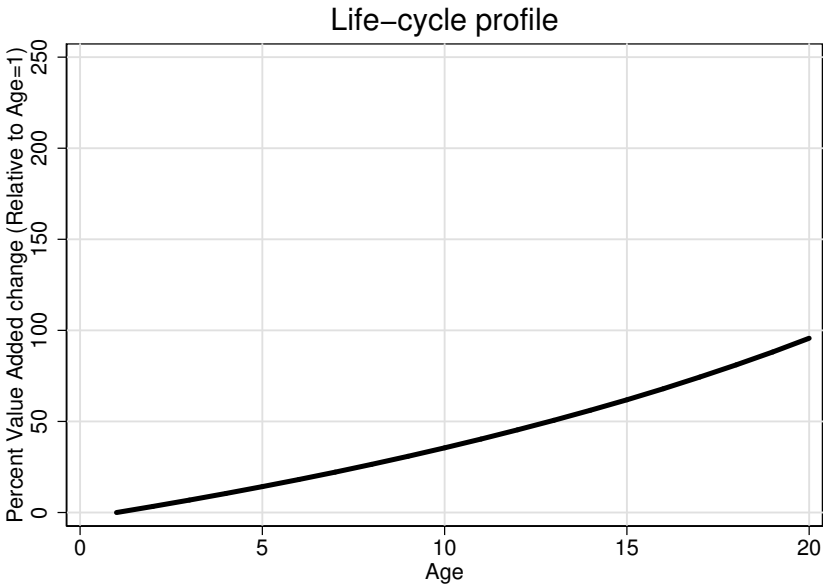
Table 2.5: Parameter Values

Parameter	Value	Interpretation	Target	Data	Model
(1)	(2)	(3)	(4)	(5)	(6)
g_z	0.06	Growth rate of productivity	Life-cycle EM (Age 1 vs 20)	2.13	2.20
γ	0.50	Complementarity btw management and productivity	Life-cycle differences (Age 20)	2.26	1.77
$\bar{h}$	72.73	Managerial endowment	Average size (workers)	13	13
ϕ	48.55	Managerial efficiency	% of FM	80%	81%
μ_h	0.2	Average managerial skill	Normalization		
a	1.48	Pareto shape (entrants)	% of External Managers	4.3%	4.3%
θ	1.61	Pareto shape (managerial talent)	Share Total Revenue FM	40%	41%
λ	0.10	Exit rate	Exit rate	10%	10%
α	0.65	Labor share	Labor compensation	—	—
c_e	28.21	Entry cost	—–		
β	0.96	Discount factor	Real interest rate	—	—

Notes. This Table report the estimates of the parameters used to calibrate the model. See Section 2.5 for details about the calibration procedure.



(a)



(b)

Figure 2.12: (a) Life-cycle profile of Family-Managed and Externally-Managed firms, according to the parameters reported in Table 2.5. Age equal one is the base category. The number of periods simulated is 50. (b) As for (a) but for the overall economy.

Table 2.6: Counterfactual Analysis: Increase in Managerial Efficiency

	Baseline Economy			
	(1)	(2)	(3)	(4)
Level of Managerial Efficiency (ϕ)	(ϕ_1)	(ϕ_2)	(ϕ_3)	(ϕ_4)
Output	100	101	103	106
Fraction of FM firms	0.81	0.75	0.60	0.40
Share of output (FM)	0.40	0.32	0.21	0.11

Notes. This Table reports the effects of improving the managerial efficiency from the baseline estimate. Column (1) depicts the baseline economy, where output has been normalized to 100 for comparison purposes.

interpretation of the experiments. Column (2) report the results when ϕ improves the managerial environment in such a way that the fraction of FM firms drop from 0.80 to 0.75. In this case output increases 1%, and the participation of FM firms in total output decreases in 6 percentage points, as EM firms become more efficient and capture a larger share of the market.

Column (3) and (4) in Table 2.6 show the results when the managerial efficiency improves to reduce the fraction of FM firm to 0.60 and 0.40. In these cases the gains on output correspond to an increase of 3% and 6% when compared to the benchmark economy in column (1). At the same time, the participation of FM firms in the economy drops in 19 and 29 percentage points, respectively. These exercises suggest that as the managerial efficiency increases, for instance, due to an improvement in the contractual environment, output expands as a result of both a greater selection of firms that move away from family management, and an overall increase in the efficiency of those that were already externally managed, which reduces the share of the participation in output of FM firms.

More generally, Figure 2.13 show how improvements in the managerial efficiency generate output gains (left axis), and how at the same time these changes reduce the fraction of FM firms in the economy (right axis). As it can be seen, increasing the managerial efficiency monotonically augments output, although the relation is not linear. For instance, when managerial efficiency leads to a drop of family-managed firms to 10%, output increases 12% relative to the benchmark economy. Now recall from Table 2.2 that the fraction of FM firms in developed countries is lower than in Ecuador, and in average the share of FM firms from that group is twice as large. So using them as a benchmark, it suggests that changes in the contractual environment that lead to a drop in the presence of family-managed firms by half could increase output on the order of 6%.

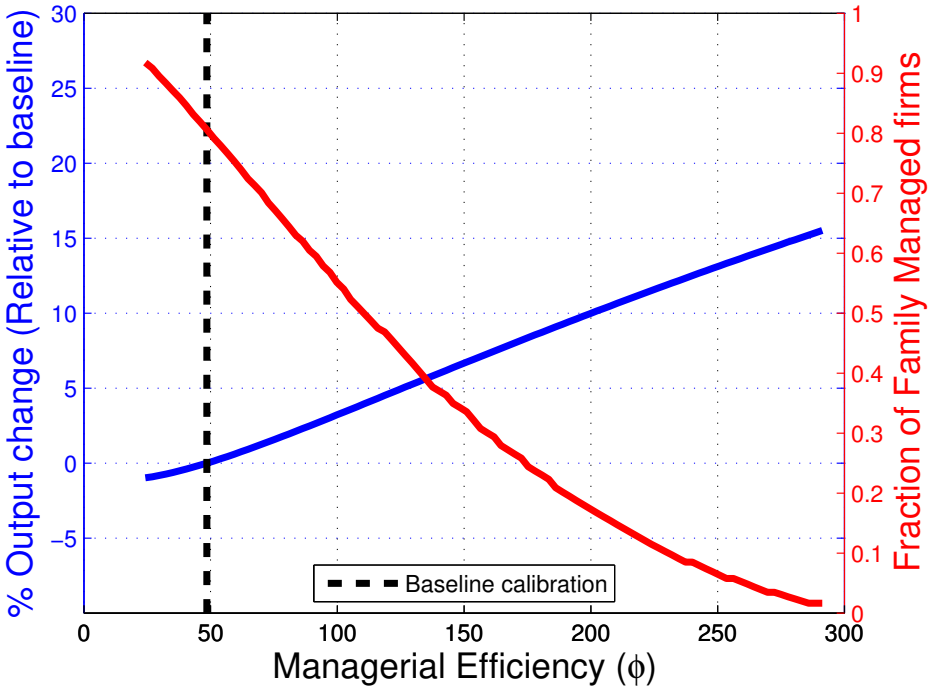


Figure 2.13: Changes in Output and Composition of Firms for Different Levels of ϕ

Notes: Changes in output (left) and composition of firms (right) for different levels of ϕ. Output was normalized to the benchmark economy. The dashed line represent the baseline estimate of ϕ. The rest of the parameters are the same as reported in Table 2.5 for each counterfactual.

Note that given the parameters g_z and γ, the differences in life-cycle between both FM and EM will not be affected. However, due to changes in composition, the *overall* firm dynamics will be altered. To illustrate this point, Figure 2.14 show the life-cycle profiles for the baseline economy accompanied with the corresponding ones generated for three different economies with different managerial efficiency levels. As the value of ϕ increases, when seen altogether, the life cycle of firms improves due to a composition effect. For instance, the baseline case closely follows that of FM firms displayed in Figure 2.12b, similar to the empirical evidence shown in Figure 2.8. Instead, when ϕ increases in such a way that the fraction of FM firms drop by half,

the life-cycle of firms improve in such a way that at age 20, firms are in average 150% larger that the typical entrant at age 1.

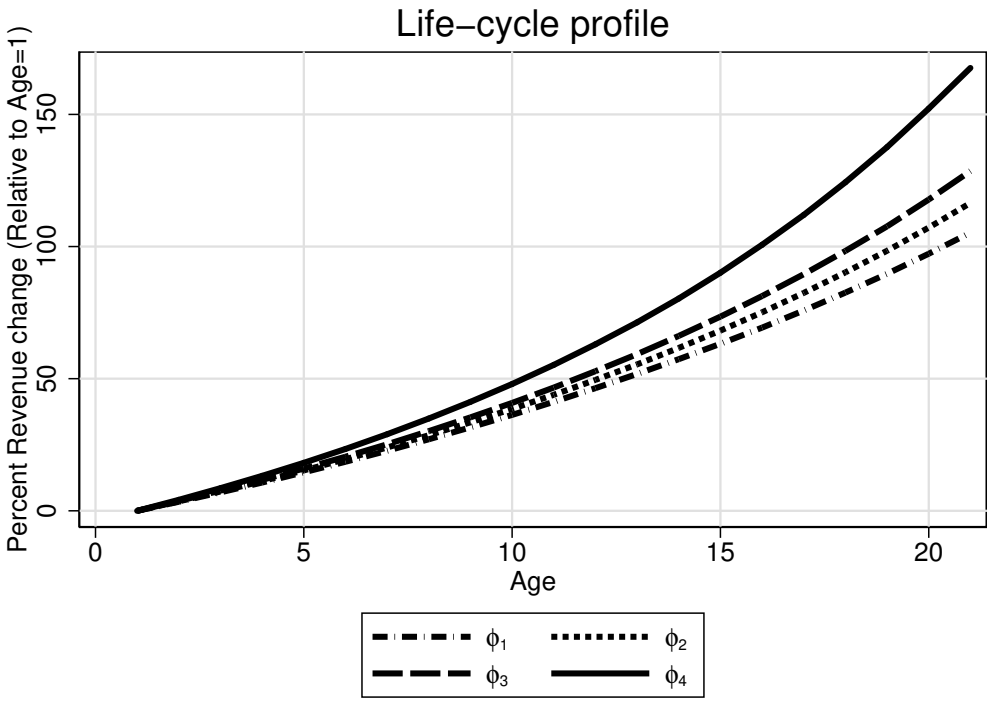


Figure 2.14: Simulated Life-cycle Profile of Family Managed and Non-Family Managed Firms

Notes: Estimated by (2.2), according to the parameters reported in Table 2.5, except for ϕ, which is altered to the values portrayed in Table 2.6. Age equal one is the base category. The number of periods simulated is 50.

In summary, the experiments performed in this section suggest that improvements in the managerial efficiency, for instance, linked by a reduction of distortions in the contractual environment, create incentives to firms to more frequently hire external management. These improvements in turn leads to important gains in output. These gains operate in two ways. First, induces a larger fraction of medium-sized firms, which an otherwise contractual environment will stay as FM firms, to switch and

start expanding at earlier ages. Consequently, in the cross section, there is not only a larger group of EM firms, but also firms that were already EM become bigger in terms of output and employment, as they scale up production thanks to a greater ease to delegate. Finally, as EM firms become more efficient, they also capture a larger share of the market.

In terms of the life-cycle, the relative gap between both type of firms is not altered, however, when seen altogether, firms seem to expand in average at higher rates. For instance, when the level of ϕ is such that the managerial efficiency drops the fraction of FM firms by half, in average firms are 150% larger than the typical entrant. This model, in consequence, stresses the importance of the composition of firms in terms of the type of management to account for differences in the life-cycle across countries. To the extent that developing countries seem to have a larger fraction of family-managed firms, as documented in this paper, it suggests that their firm dynamics are greatly influenced by the overrepresentation of FM firms. Through the lens of the model, the relatively flat life-cycle profile of FM firms occur because firms can't scale up production by hiring the required management to expand, critically influenced by the low levels of managerial efficiency, so they end up operating only with their fixed endowment, facing eventually a span of control problem.

2.7 Managerial Differences

The previous section explores the potential aggregate gains of increasing the managerial efficiency. Through the lens of the model, this improvement leads to gains in output as they decrease the participation of FM firms in the economy by giving more incentives to delegate and scale up production. At the same time the model gives a quantitative assessment of how the composition of firms is reflected in the overall life-cycle of firms. However, the model abstract from particularities in the set

of managerial skills that may account for the differences in performance between both type of firms documented in this paper. In this section I provide further empirical evidence along those lines of future extensions of the model that may amplify the effects of improving the delegation environment.

One of the four categories of managerial practices measured by Bloom and Van Reenen (2007) is labeled as *incentives*, and refers to a firm's ability to attract and retain human capital, as well as fixing and rewarding employees according to their performance. Given I also have access to employees records, I explore management differences between FM an EM through this channel, by studying differences in compensation and workforce selection. This exercise is similar to Bender et al. (2018) who link survey data of managerial practices with administrative data of Germany, although I have the advantage to explore a larger dataset of firms and focus specifically on the role of family management.

To perform this analysis I match workers to their correspondent employers at an annually basis using the reported employer ID in worker's tax records (Form 107). I restrict the sample to be for full time workers (earn more than the minimum salary) and exclude from these set of workers individuals reported as managers in the administrative registries.

I start documenting that the earning's distribution of workers in Family-Managed firms is skewed to the left, while the opposite is true for the corresponding distribution in Externally-Managed firms. The results are displayed in Figure 2.15. The Figure simply shows the raw distribution of earnings. At first glance, it suggests that workers in EM firms tend to have higher earnings, although the distribution of workers' earnings in EM firms do not fully dominate that of FM firms. Perhaps more interesting, top earners mainly work in firms with external managers.

To further quantify if there is an earning's premium in Externally-Managed firms,

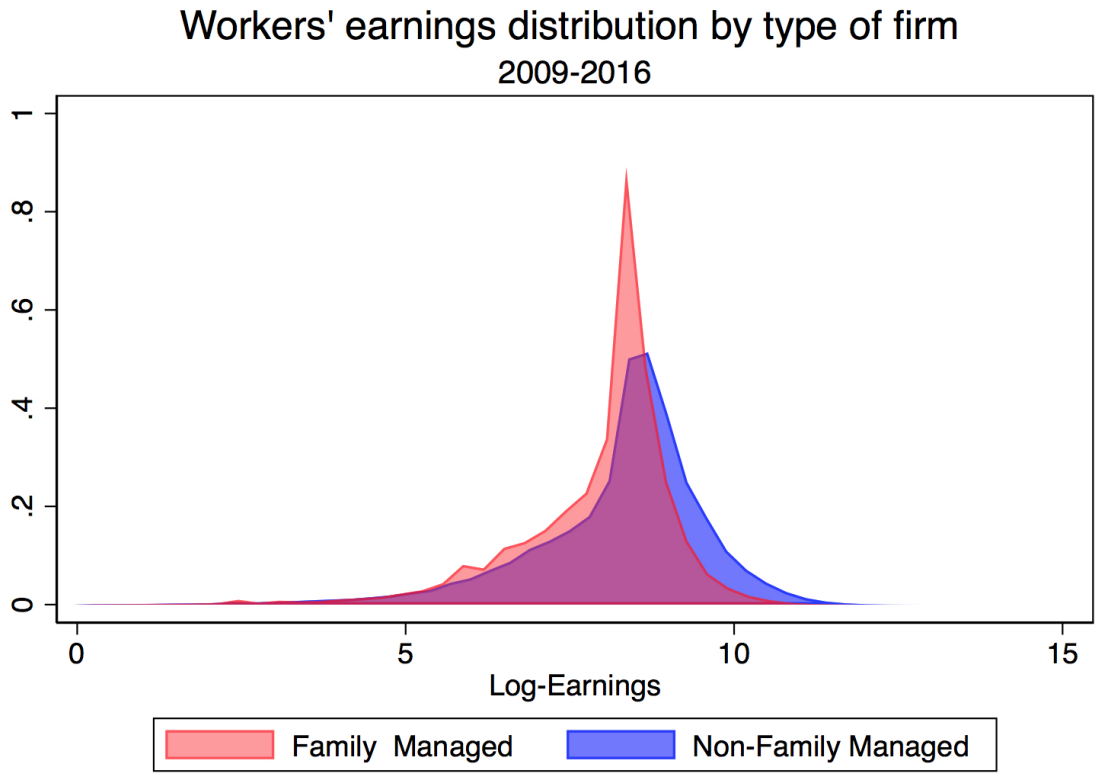


Figure 2.15: Annual Earnings Distribution for Workers Employed in FM and EM Firms

Notes: Firms were classified according to the cut-off rule defined in (2.1) with $\bar{c} = 50\%$. The sample includes the universe of matched workers on formal firms in Ecuador, both Limited Liability Companies and Sole Proprietors, except the financial sector, for the period 2009-2016. Earnings is Real Earnings, deflated using the CPI.

I estimate the following regression,

$$\log(y_{i,t}) = \mu + \alpha_i + \delta_t + \mathbf{FM_{i,t}} + \text{error}_{i,t} \tag{2.22}$$

where $y_{i,t}$ is the annual earning of a worker i, at year t; the coefficients α_i correspond to a time invariant fixed effects, δ_t denotes a year fixed effect and $\mathbf{FM_{i,t}}$ is a variable that correspond to 1 if the individual was working for a FM firm in period t, and 0 if that was not the case; while $\text{error}_{i,t}$ is the residual. Unfortunately, it is not possible to

Table 2.7: Earning's Differences Between Workers in FM and EM Firms

	(1)	(2)
FM		
(Employed in a Family-Managed firm)	-0.10	-0.097
	(.001)	(.001)
Controls		
Individual fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Include Sole Proprietors	Yes	No
Number of Observations	11,783,008	10,030,409

Notes. This Table report the estimates of (2.22). FM, is a categorical variable that takes the value of 1 if a worker is employed in a Family-Managed firm, as classified by (2.1) with $\bar{c} = 50\%$. Column (1) includes workers from all firms. Column (2) only uses workers in Limited Liability Companies. Standard errors in parenthesis.

control for demographic characteristics of workers, as this information is not available in the tax records. Hence, these estimates abstract for important time-varying forces that could explain differences in compensation, as experience, although other time-invariant characteristics, like education, may be in part captured in the fixed effects. The results are displayed in Table 2.7.

Column (1) show that in average, FM firms pay 10% less than EM ones. These results are statistically significant and suggest a positive premium for working in a

EM firm. Also the magnitude is economically relevant, as it suggest a non-negligible difference in earnings. Now, recall that in the group of FM firms are also included Sole Proprietors. Therefore, in order to see if the inclusion of these firms is driving the results, Column (2) redo the exercise, but exclude employees that work for Sole Proprietors. As one can observe, the results change little, since most matched-workers appear to be employed in Limited Liability Companies. This finding is consistent with the fact that EM firms are generally larger, so in part may reflect the premium for working for a large business (e.g. Troske (1999)).

One further explanation of this difference in the distribution of worker's earnings could also be attributed to the selection of more skilled or talented workers in EM firms. For instance, Sraer and Thesmar (2007) report that public-listed firms in France controlled by a family pay in average lower wages, in part because they employ less-skilled employees. To see if that is the case, I decompose worker's earnings in both firm's and worker's components via the AKM model (Abowd et al. (1999)), recover the workers effect and use it as a proxy measure of the quality of the workforce of each type of firm. More specifically, I follow the literature and assume that log-annual earnings ($y_{i,t}$) can be written in the following way,

$$\log(y_{i,t}) = \alpha_i + \psi_{j(i,t)} + \delta_t + \mathbf{X_{i,t}} + \text{error}_{i,t} \tag{2.23}$$

where $i = 1 \dots N$ denotes workers and $j = 1 \dots J$ firms. The coefficients α_i and $\psi_{j(i,t)}$ correspond to time invariant unobserved heterogeneity at the worker and firm level (fixed effects), and typically interpreted by the literature as worker's "ability" and firm's "productivity". Finally, δ_t denotes year fixed effects and $\mathbf{X_{i,t}}$ are time-varying observable on workers; $\text{error}_{i,t}$ is the residual.

However, as mentioned above, unfortunately tax records do not contain demographic information about workers, like age or gender, making this piece of informa-

tion missing in the estimation. Hence, I regress a modified version of (2.23), where worker's characteristics, $\mathbf{X_{i,t}}$, are omitted. I also do not include the classification of worker's firm (FM vs EM) as a control variable. Instead, after I recover the fixed effects estimates, I associate them with their correspondent firm status in that year,

$$fm_{J(i,t)} = \begin{cases} 1 & \text{if individual } i \text{ at } t \text{ was working in a FM-firm} \\ 0 & \text{if individual } i \text{ at } t \text{ was working in a EM-firm} \end{cases}$$

Finally, I estimate (2.23) using the methodology developed by Correia (2016)[18].

As way to measure differences in workforce selection, I simply compute the correlation between workers and fixed effects for both type of firms. The results of this exercise are depicted in Figure 2.16, where I collapsed the information into 20 equal-sized bins to facilitate the visualization of the graph[19]. As one can observe, there is a flat correlation between worker's and firm's fixed effects for FM firms. This is not the case for EM firms. For the latter, more productive firms tend to hire more productive workers, as proxied by their individual fixed effects.

Although these results do not necessarily imply differences in assortative matching (Eeckhout and Kircher (2011)), they are suggestive about the hiring patterns of both type of firms. If workers in FM firms expect that upper positions will be held by family members of the owners, instead of being rewarded in a more meritocratic way, it is likely that many of them will abandon the firm in look for other opportunities. This could particularly be more hurtful for the firm if the workers that leave are those more talented.

Another interpretation of this finding is that managers in FM firm may fail to reward and properly recognize outside talent. Since one of the tasks of managers is

[18]This estimator has been implemented in STATA in the "reghdfe" package.

[19]To estimate the fixed effects a normalization relative to one was, so the scale in both axis has been normalized relative to some arbitrary level.

to recruit the adequate workforce for the firm, lacking the ability to do so could have important implications for the productivity of the firm. In fact, as mentioned above, properly giving incentives to workers is one of the managerial practices that Bloom and Van Reenen (2007) report are associated with better firm's performance. In that sense, allowing the model to capture this additional channel will likely to amplify the effects of changes in the managerial efficiency.

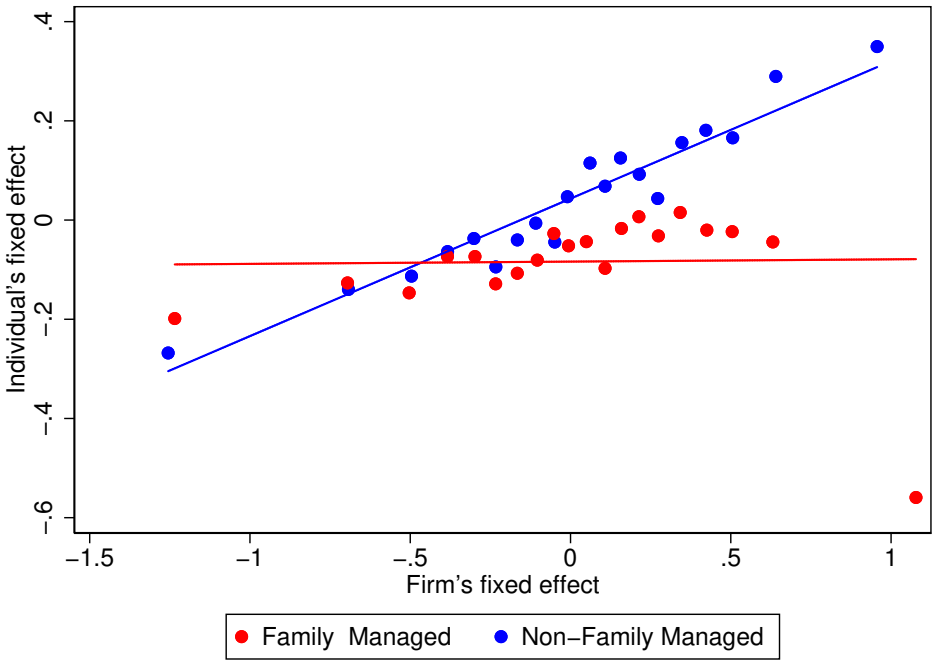


Figure 2.16: Correlation Between Workers and Individuals Fixed Effects for Family and Externally Managed Firms

Notes: Estimated by (2.23). Firms were classified according to the cut-off rule defined in (2.1) with $\bar{c} = 50\%$. The sample includes the universe of matched workers on formal firms in Ecuador, both Limited Liability Companies and Sole Proprietors, except the financial sector, for the period 2009-2016.

2.8 Conclusion

This paper uses novel administrative data of family-relations between owners and managers within a firm and their tax records to provide several stylized facts about differences in performance between Family-Managed and Externally-Managed firms. One of the main new findings is that firms managed by a family member grow as half as those that do not. Given the large number of firms that is suspected to operate under this condition in developing countries, these channel appear to be a quantitatively relevant to explain cross-country differences in the life-cycle of firms, like those documented by Hsieh and Klenow (2014). For instance, I report that these family-managed firms represent around 80% of the total number of firms in Ecuador and comparable samples suggest that this share is around 2 times that of a developed economy.

A consequence of this life-cycle differences is that despite being 80% of firms, these family-managed firms only account for almost 40% of total revenue and 50% of total employment in the economy. Moreover, the vast majority of these firms are associated with lower levels of employment. I also provide some evidence via a case-event study that there are gains of 20% in revenue and 10% in employment when firms switch from being a FM firm to a EM one. However, this transition is really infrequent. The apparent benefits of operating with external management leads to the question of why there is there is a large concentration of Family-Managed firms? In that sense, I provide some pieces of evidence that suggest that a deficient contractual environment may be behind the low levels of outside delegation. I show for example that ownership is quite concentrated, reflecting the difficulty to cope with outsiders. I also report that family-management is a widespread phenomena across sectors, suggesting it is a structural feature of the economy.

I then write a model of firm dynamics to study the stylized facts documented in this paper. In the model entrepreneurs endogenously decide if they operate as a family-managed firm, or delegate and hire external managerial services. A family-managed firm in the model is characterized as one that only uses the endowment of managerial talent of the entrepreneur, while a externally-managed firm is the one that hires external management in the market. The contractual environment affect the managerial efficiency and shapes the decision of a firm to remain as family-managed. I calibrate the model to the Ecuadorian economy by using several of the new empirical findings documented here. The model is able to replicate most of the new stylized documented in this paper. I then perform several counterfactuals associated with an increase in the managerial efficiency. Through the lens of the model changes in the contractual environment that lead to a drop in the presence of family-managed firms by half could increase output on the order of 6%. These gains operate through two channels. First, they induce more firms to hire external management and scale up production, and second, they improve the productivity of firms that were already externally-managed, leading them to capture a larger share of the market.

I also provide empirical evidence related to workforce selection and compensation between FM and EM firms that can be associated with differences in managerial practices. I report that EM in average pay more, consistent with the fact that they are typically larger. I then decompose workers' earnings via the AKM model (Abowd et al. (1999)) and show that the correlation between firm's fixed effects and worker's fixed effects is more pronounced for externally-managed firms, and close to flat for family-managed firms. This suggests first that the average quality of workers is higher in EM firms, but also that relatively higher productive family-managed firms tend to have lower quality workers. Two potential explanations to this finding is the self-selection of workers to firms where they can professionally grow and achieve upper

positions that are not necessarily reserved to family members. Another is that family-managers display poor managerial practices like failing to recognize and incentivize outside talent.

In that sense, an important future extension of the model is to include worker heterogeneity, as it could amplify the effects of distortions in the contractual environment. Another relevant extension will be to add heterogeneity to the talent of managers in FM firms. This will account for the fact that there are successful family-managed firms. But more important, this extension will add a further scope of policies to be studied. As improvements in the delegation environment may take longer time to execute, another relevant policies to analyze is what type of incentives or training could be given to family-managed firms in order to increase their productivity.

It will also be interesting to understand further hybrid mechanism of control inside a firm. In this paper I focus specifically on the top management positions, however, firms operate with different layers of management (Garicano and Rossi-Hansberg (2015)). Understanding if family members or outside workers are part of these layers will help to disentangle the sources of managerial productivity and model richer micro-foundations of delegation. In that sense, learning about the mechanism of control for outside managers in these different layers on environments with contractual frictions will expand our knowledge of the forces behind the selection of family managers. To do so, it will be required to collect further information about managerial practices of firms and join them with their administrative records.

More generally, future studies of management in both developed and developing countries should focus on the role of these family-managed firms for firm dynamics due to their suspected large presence. That will allow us confirm and contrast the findings in this paper and to generate more comparable data across countries. For example, one important fact to be confirmed in coming studies is if the life-cycle gap

between FM and EM documented in this paper is *smaller* or *larger* in environments with better contractual enforcement. Yet is to be verified in the future once data of the same nature used in this paper is available for other countries.

REFERENCES

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, 67(2):251–333.

Acemoglu, D., Bautista, M. A., Querubin, P., and Robinson, J. A. (2008). Economic and political inequality in development: the case of cundinamarca, colombia. *Institutions and economic performance*, pages 181–246.

Acemoglu, D., García-Jimeno, C., and Robinson, J. A. (2012). Finding eldorado: Slavery and long-run development in colombia. *Journal of Comparative Economics*, 40(4):534–564.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2005). Institutions as a fundamental cause of long-run growth. *Handbook of economic growth*, 1:385–472.

Acemoglu, D. and Wolitzky, A. (2011). The economics of labor coercion. *Econometrica*, 79(2):555–600.

Adamopoulos, T. and Restuccia, D. (2014). The size distribution of farms and international productivity differences. *American Economic Review*, 104(6):1667–97.

Akcigit, U. (2017). Economic growth: The past, the present, and the future. *Journal of Political Economy*, 125(6):1736–1747.

Akcigit, U., Alp, H., and Peters, M. (2016). Lack of selection and limits to delegation: firm dynamics in developing countries. Technical report, National Bureau of Economic Research.

Albornoz, O. (1971). *Las luchas indígenas en el Ecuador*. Editorial Claridad.

Albuja Mateus, A. E. (1998). *Doctrinas y parroquias del obispado de Quito en la segunda mitad del Siglo XVI*. Abya-Yala, Quito, Ecuador, 1. ed.. edition.

Alchon, S. A. (1991). *Native society and disease in colonial Ecuador*, volume 71. Cambridge University Press, New York;Cambridge [England];.

Alder, S. D. (2016). In the wrong hands: Complementarities, resource allocation, and tfp. *American Economic Journal: Macroeconomics*, 8(1):199–241.

Andrees, B. and Belser, P. (2009). *Forced labor: coercion and exploitation in the private economy*. Lynne Rienner Publishers.

Andrien, K. J. (1995). *The kingdom of Quito, 1690-1830: the state and regional development*, volume 80. Cambridge University Press, Cambridge;New York;.

Bandiera, O., Lemos, R., Prat, A., and Sadun, R. (2017). Managing the family firm: evidence from ceos at work. *The Review of Financial Studies*, 31(5):1605–1653.

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

Banerjee, A. V. and Duflo, E. (2003). Inequality and growth: What can the data say? *Journal of economic growth*, 8(3):267–299.

Barsky, O. (1984). *La reforma agraria ecuatoriana*, volume 3. Corporacion Editora Nacional, Quito.

Bartelsman, E., Haltiwanger, J., and Scarpetta, S. (2009). Measuring and analyzing cross-country differences in firm dynamics. In *Producer dynamics: New evidence from micro data*, pages 15–76. University of Chicago Press.

Bartelsman, E., Haltiwanger, J., and Scarpetta, S. (2013). Cross-country differences in productivity: The role of allocation and selection. *American Economic Review*, 103(1):305–34.

Basile, D. G. (1974). *Tillers of the Andes: farmers and farming in the Quito Basin*, volume no. 8. University of North Carolina at Chapel Hill, Dept. of Geography, Chapel Hill.

Baud, M. (2007). Liberalism, indigenismo, and social mobilization in late nineteenth-century ecuador. In *Highland Indians and the state in modern Ecuador*.

Bender, S., Bloom, N., Card, D., Van Reenen, J., and Wolter, S. (2018). Management practices, workforce selection, and productivity. *Journal of Labor Economics*, 36(S1):S371–S409.

Bennedsen, M., Nielsen, K. M., Perez-Gonzalez, F., and Wolfenzon, D. (2007). Inside the family firm: The role of families in succession decisions and performance*. *The Quarterly Journal of Economics*, 122(2):647–691.

Bento, P. and Restuccia, D. (2017). Misallocation, establishment size, and productivity. *American Economic Journal: Macroeconomics*, 9(3):267–303.

Berle, A. A., Means, G. C., et al. (1932). The modern corporation and private property.

Bertrand, M., Johnson, S., Samphantharak, K., and Schoar, A. (2008). Mixing family with business: A study of thai business groups and the families behind them. *Journal of financial Economics*, 88(3):466–498.

Bertrand, M. and Schoar, A. (2003). Managing with style: The effect of managers on firm policies*. *The Quarterly Journal of Economics*, 118(4):1169–1208.

Bertrand, M. and Schoar, A. (2006). The role of family in family firms. *Journal of economic perspectives*, 20(2):73–96.

Bhattacharya, D., Guner, N., and Ventura, G. (2013). Distortions, endogenous managerial skills and productivity differences. *Review of Economic Dynamics*, 16(1):11–25.

Bloom, N., Eifert, B., Heller, T. C., Jensen, E., and Mahajan, A. (2009). Contract enforcement and firm organization: Evidence from the indian textile industry. Technical report, Center on Democracy, Development, and The Rule of Law.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., and Roberts, J. (2013). Does management matter? evidence from india *. *The Quarterly Journal of Economics*, 128(1):1–51.

Bloom, N., Lemos, R., Sadun, R., Scur, D., and Reenen, J. (2014). The new empirical economics of management. *Journal of the European Economic Association*, 12(4):835–876.

Bloom, N., Sadun, R., and Van Reenen, J. (2016). Management as a technology? Technical report, National Bureau of Economic Research.

Bloom, N. and Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries*. *The Quarterly Journal of Economics*, 122(4):1351–1408.

Bloom, N. and Van Reenen, J. (2010). Why do management practices differ across firms and countries? *Journal of Economic Perspectives*, 24(1):203–24.

Buera, F. J., Kaboski, J. P., and Shin, Y. (2011). Finance and development: A tale of two sectors. *American Economic Review*, 101(5):1964–2002.

Burkart, M., Panunzi, F., and Shleifer, A. (2003). Family firms. *Journal of Finance*, 58(5):2167–2201.

Caselli, F. (2005a). Accounting for cross-country income differences. *Handbook of economic growth*, 1:679–741.

Caselli, F. (2005b). Chapter 9 accounting for cross-country income differences. volume 1, Part A of *Handbook of Economic Growth*, pages 679 – 741. Elsevier.

Caselli, F. and Gennaioli, N. (2013). Dynastic management. *Economic Inquiry*, 51(1):971–996.

Claessens, S., Djankov, S., and Lang, L. H. (2000). The separation of ownership and control in east asian corporations. *Journal of Financial Economics*, 58(1):81–112.

Cole, H. L., Greenwood, J., and Sanchez, J. M. (2016). Why doesn't technology flow from rich to poor countries? *Econometrica*, 84(4):1477–1521.

Compañía Guía del Ecuador (1909). *El Ecuador: guia comercial, agricola, e industrial de la Republica*. E. Rodenas.

Corkill, D. (1985). Democratic politics in ecuador, 1979-1984. *Bulletin of Latin American Research*, 4(2):63–74.

Correia, S. (2016). Reghdfe: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects.

Costinot, A. and Donaldson, D. (2016). How large are the gains from economic integration? theory and evidence from us agriculture, 1880-1997.

Cushner, N. P. (1982). *Farm and factory: the Jesuits and the development of agrarian capitalism in colonial Quito, 1600-1767*. State University of New York Press, Albany.

Deaton, A. (1997). *The analysis of household surveys: a microeconometric approach to development policy*. The World Bank.

Dell, M. (2010). The persistent effects of peru's mining mita. *Econometrica*, 78(6):1863–1903.

DIVA-GIS (2016). Diva-gis. `http://www.diva-gis.org/Data`. Accessed: 2016-11-06.

Donnelley, R. G. (1964). Family business. *Harvard Business Review*, 42:93–106.

Earle, C. (1992). To enslave or not to enslave: Crop seasonality, labor choice, and the urgency of the civil war.

Easterly, W. (2007). Inequality does cause underdevelopment: Insights from a new instrument. *Journal of Development Economics*, 84(2):755 – 776.

Eeckhout, J. and Kircher, P. (2011). Identifying sorting in theory. *The Review of Economic Studies*, 78(3):872–906.

Engerman, S. L. and Sokoloff, K. L. (1997). Factor endowments, institutions, and differential paths of growth among new world economies. *How Latin America Fell Behind*, pages 260–304.

Faccio, M. and Lang, L. H. (2002). The ultimate ownership of western european corporations. *Journal of Financial Economics*, 65(3):365–395.

Galor, O., Moav, O., and Vollrath, D. (2009). Inequality in landownership, the emergence of human-capital promoting institutions, and the great divergence. *The Review of Economic Studies*, 76(1):143–179.

Galor, O. and Ozak, O. (2016). The agricultural origins of time preference. *American Economic Review*, 106(10):3064–3103.

Garicano, L. and Rossi-Hansberg, E. (2015). Knowledge-based hierarchies: Using organizations to understand the economy. *Annual Review of Economics*, 7:1–30.

Güell, M., Rodrguez Mora, J. V., and Telmer, C. I. (2015). The informational content of surnames, the evolution of intergenerational mobility, and assortative mating. *The Review of Economic Studies*, 82(2):693–735.

Guner, N., Parkhomenko, A., and Ventura, G. (2018). Managers and productivity differences. *Review of Economic Dynamics*, 29:256–282.

Guner, N., Ventura, G., and Xu, Y. (2008). Macroeconomic implications of size-dependent policies. *Review of Economic Dynamics*, 11(4):721 – 744.

Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2013). Who creates jobs? small versus large versus young. *Review of Economics and Statistics*, 95(2):347–361.

Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, 102(2):994–1028.

Hendricks, L. and Schoellman, T. (2018). Human capital and development accounting: New evidence from wage gains at migration. *The Quarterly Journal of Economics*, 133(2):665–700.

Hopenhayn, H. and Rogerson, R. (1993). Job turnover and policy evaluation: A general equilibrium analysis. *Journal of Political Economy*, 101(5):915–938.

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica: Journal of the Econometric Society*, pages 1127–1150.

Hopenhayn, H. A. (2014). Firms, misallocation, and aggregate productivity: A review. *Annual Review of Economics*, 6(1):735–770.

Hsieh, C.-T. and Klenow, P. (2014). The life cycle of plants in india and mexico. *The Quarterly Journal of Economics*, 129(3).

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly journal of economics*, 124(4):1403–1448.

Jenkins, S. P. et al. (2015). Ineqdec0: Stata module to calculate inequality indices with decomposition by subgroup. *Statistical Software Components*.

Jones, C. I. (2016). The facts of economic growth. In *Handbook of macroeconomics*, volume 2, pages 3–69. Elsevier.

Kandel, E. and Lazear, E. (1992). Peer pressure and partnerships. *Journal of political economy*, 100(4):801–817.

Kueng, L., Yang, M.-J., and Hong, B. (2014). Sources of firm life-cycle dynamics: Differentiating size vs. age effects. Working Paper 20621, National Bureau of Economic Research.

La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (1999). Corporate ownership around the world. *The journal of finance*, 54(2):471–517.

La Porta, R., Lopez-de Silanes, F., Shleifer, A., and Vishny, R. W. (1997). Legal determinants of external finance. *The journal of finance*, 52(3):1131–1150.

La Porta, R., Lopez-de Silanes, F., Shleifer, A., and Vishny, R. W. (1998). Law and finance. *Journal of political economy*, 106(6):1113–1155.

Laeven, L. and Woodruff, C. (2007). The quality of the legal system, firm ownership, and firm size. *The Review of Economics and Statistics*, 89(4):601–614.

Lagerlöf, N.-P. (2005). Geography, institutions and growth: the united states as a microcosm.

Ley de Compañías, E. (1999). Quito. *Registro Oficial*, 312.

Lucas, R. E. (1978). On the size distribution of business firms. *The Bell Journal of Economics*, pages 508–523.

Mariscal, E. and Sokoloff, K. L. (2000). Schooling, suffrage, and the persistence of inequality in the americas. *Political Institutions and Economic Growth in Latin America: Essays in Policy, History, and Political Economy*, page 159.

Midrigan, V. and Xu, D. Y. (2014). Finance and misallocation: Evidence from plant-level data. *American Economic Review*, 104(2):422–58.

Moll, B. (2014). Productivity losses from financial frictions: Can self-financing undo capital misallocation? *American Economic Review*, 104(10):3186–3221.

Montiel Olea, J. L. and Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, 31(3):358–369.

Morck, R., Schleifer, A., and Vishny, R. (1988). Management ownership and market valuation - an empirical analysis. *Journal of financial economics*, 20(1-2):293–316.

Morck, R., Stangeland, D. A., and Yeung, B. (2000). Inherited wealth, croporate control, and economic growth : the canadian disease? In *Concentrated corporate ownership*, pages 319–369.

National Aeronautics and Space Administration and the National Geospatial Intelligence Agency (2010). Shuttle radar topography mission 30 arc second finished data. `https://webmap.ornl.gov/wcsdown/wcsdown.jsp?dg_id=10008_1`. Accessed: 2015-12-10.

National Geophysical Data Center (2010). Version 4 dmsp-ols nighttime lights time series. `http://www.ngdc.noaa.gov/dmsp/downloadV-4composites.html`. Accessed: 2016-11-06.

Newson, L. A. (1995). *Life and death in early colonial Ecuador*, volume 214. University of Oklahoma Press, Norman.

Nunn, N. (2008). Slavery, inequality, and economic development in the americas. *Institutions and economic performance*, pages 148–80.

Nunn, N. and Qian, N. (2011). The potato's contribution to population and urbanization: Evidence from a historical experiment*. *The Quarterly Journal of Economics*, 126(2):593.

Oberem, U. (1978). Contribución a la historia del trabajador rural de américa latina: "conciertos" y "huasipungueros" en ecuador.

Oberem, U. (1981). "indios libres" e "indios sujetos a haciendas" en la sierra ecuatoriana a fines de la colonia. *Revista Pendoneros*, (20).

Penrose, E. T. (1959). The theory of the growth ofthe firm. *New York: Sharpe*.

Perez-Gonzalez, F. (2006). Inherited control and firm performance. *American Economic Review*, 96(5):1559–1588.

Pflueger, C. E. and Wang, S. (2014). A robust test for weak instruments in stata.

Poloni, J. (2006). *El mosaico indígena: movilidad, estratificación social y mestizaje en el corregimiento de Cuenca (Ecuador) del siglo XVI al XVIII*, volume 184. Editorial Abya Yala.

Quintero López, R. (1987). El estado terrateniente de ecuador (1809-1895). *JP Deler e Y. Saint-Geours (comp.), Estados y naciones en los Andes*, 2:397–417.

Restuccia, D. et al. (2013). Factor misallocation and development. *The New Palgrave Dictionary of Economics, Online Edition, Eds. Steven N. Durlauf and Lawrence E. Blume, Palgrave Macmillan*.

Restuccia, D. and Rogerson, R. (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic dynamics*, 11(4):707–720.

Restuccia, D. and Rogerson, R. (2013). Misallocation and productivity. *Review of Economic Dynamics*, 1(16):1–10.

Restuccia, D., Yang, D. T., and Zhu, X. (2008). Agriculture and aggregate productivity: A quantitative cross-country analysis. *Journal of Monetary Economics*, 55(2):234–250.

Roys, N. and Seshadri, A. (2014). Economic development and the organization of production.

Saint Geours, Y. (1994). La sierra centro norte (1830-1925).? in historia y región en el ecuador.(1830-1930), edited by juan maiguashca, 143-188. *Quito: Corporación Editora Nacional, FLACSO and CERLAC*.

Salmoral, M. L. (1994). La población del reino de quito en la época de reformismo borbónico: circa 1784. *Revista de Indias*, 54(200):33.

Schulhofer-Wohl, S. (2018). The age-time-cohort problem and the identification of structural parameters in life-cycle models. *Quantitative Economics*, 9(2):643–658.

Schultz, T. W. T. W. (1953). *The economic organization of agriculture.* McGraw-Hill, New York.

SIISE (2011). Sistema integrado de indicadores sociales del ecuador. `http://www.siise.gob.ec/siiseweb/`.

SNI (2011). Sistema nacional de informacion. `http://sni.gob.ec/inicio`. Accessed: 2016-11-12.

Sraer, D. and Thesmar, D. (2007). Performance and behavior of family firms: Evidence from the french stock market. *Journal of the European Economic Association*, 5(4):709–751.

Tabla Ducasse, J. O. d. L. (1980). La poblacion ecuatoriana en la epoca colonial: Cuestiones y calculos. *Anuario de Estudios Americanos*, 37:235–277.

Troske, K. R. (1999). Evidence on the employer size-wage premium from worker-establishment matched data. *The Review of Economics and Statistics*, 81(1):15–26.

Tyrer, R. B. (1988). *Historia demografica y economica de la Audiencia de Quito: poblacion indigena e industria textil, 1600-1800*, volume 1. Banco Central del Ecuador, Quito.

US Department of Agriculture, Agricultural Research Service, N. D. L. (2016). Usda national nutrient database for standard reference. `http://www.ars.usda.gov/ba/bhnrc/ndl`. Accessed: 2016-11-06.

Valarezo, G. R. (2002). Relaciones interetnicas en la sierra norte: la diversa evolucion de cayambe y otavalo en el transito a la republica. *Procesos*, (19):23.

Villalonga, B. and Amit, R. (2006). How do family ownership, control and management affect firm value? *Journal of Financial Economics*, 80(2):385–417.

Villegas, L. F. B. (2008). Espacio, cuestión agraria y diferenciación cultural en chimborazo, ecuador. una aproximación histórica. *Gazeta de Antropología*, 24(1).

Williams, D. (2007). Administering the otavalan indian and centralizing governance in ecuador, 1851-1875. *Highland Indians and the State in Modern Ecuador*.

Yermack, D. (1996). Higher market valuation of companies with a small board of directors. *Journal of Financial Economics*, 40(2):185–211.

APPENDIX A

SUPPLEMENTARY MATERIAL FOR CHAPTER 1

Concertaje

The data of *concertaje* and *haciendas* come from two sources. The main one is the colonial tax records located in the National Archive of History of Ecuador. The second one is from Oberem (1981). I describe the treatment of the data below.

National Archive of History

As mentioned in the paper, the information of *concertaje* was obtained via colonial tax records. Those records are at the level of colonial districts (*parishes*) and kept in books organized by colonial tax administration. I collected as many books as possible, trying to obtain at least one record by administration. The only administration for which I did not find any record was that of Guaranda. I took photographies of those records and digitize them. The exact location of each record in the Archive is reported below.

Since the collection of taxes was in charge to the authority of each colonial tax administration, and later kept separately, the books differ in their quality, completeness, and raise questions about their comparability. These are the three main concerns regarding these books that I try to minimize by employing the following criteria. First, although most of the time the quality of the books is good, in the sense that it is relatively easy to read them, in some cases the ink has vanished and is hard to recover the information written there. If other sources for the data are available, I disregard the data coming from books difficult to read in an effort to minimize transcription errors.

Second, even when in most of the cases the books include complete information for all districts, some of them do not have information of certain districts due to partial destruction or time deterioration. For those reasons, I use as a reference the population census *circa* 1784 reported by Salmoral (1994) in order to identify potential missing districts. Consistently, each table below mentions if the *concertaje* information of that district is complete, and reports if the district appears in the population census.

Finally, most of the books, in particular the earliest ones, seems to be more comparable between them, as they explicitly mention the payments of indigenous workers, free and living in *haciendas* for the periods of *San Juan* and *Navidad*. The books after 1822, when the region gained its independence, tend to omit the tax collection period or reflect partial payments for a different period. Moreover, I am particularly cautious with the information after 1822 since the independence wars could have temporarily affected the distribution of workers and the new political control of landlords could have allowed them to skip the taxation of their workers. Hence, to validate the truthfulness of some observations after 1822, especially when no other source is available, I refer to the information of Oberem (1981). I discuss in detail all of these caveats for each book below.

Alausi

Book 1 (1826)

Location:

Archivo Nacional de Historia. Seccion: Empadronamientos. Caja: 12. Documento: 11.

	Colonial District	Description	Census
1	Tixan	Complete	No
2	Alausi	Complete	Yes
3	Guasuntos	Complete	No
4	Chunchi	Complete	No
5	Sibambe	Complete	No

Missing districts from census: None.
Notes:

1. This book does not specify if the payments belong to the periods of *San Juan* and *Navidad*. This book is also difficult to read and does not specify the location of any of the *haciendas* mentioned. Hence, I used other sources to identify their location. This procedure recover rates of *concertaje* similar to those presented by Oberem (1981). However, due to concerns related to its quality, I disregard the data from sample. Instead I use the data presented in Oberem (1981).

2. The census data provides aggregate information for Alausi, so none of the other 4 districts appear on it.

Ambato

Book 1 (1823)

Location:

Archivo Nacional de Historia. Seccion: Tributarios. Caja: 35.

	Colonial District	Description	Census
1	Pillaro	Complete	Yes
2	Patate	Complete	Yes
3	Pelileo	Complete	Yes
4	Quero	Complete	Yes
5	Tisaleo	Complete	Yes
6	Santa Rosa	Complete	Yes
7	Quisapincha	Complete	Yes
8	Izamba	Complete	Yes
9	Ambato	Complete	Yes

Missing districts from census: (1) Banos.
Notes:

1. This book does not specify if the payments belong to the periods of *San Juan* and *Navidad*. The number of indigenous workers in this book is inferior to those reported by Oberem (1981), but the *concertaje* rates are relatively similar (see-Figure A.1). Hence, I use all the information of the book.

Book 2 (1825)

Location:

Archivo Nacional de Historia. Seccion: Empadronamientos. Caja: 29. Documento: 21.

	Colonial District	Description	Census
1	Banos	Complete	Yes

Missing districts from census: Read notes below.
Notes:

1. This book does not specify if the payments belong to the periods of *San Juan* and *Navidad*. This book also contains data about other districts of the region, but it is very difficult to read since the ink has vanished. Consequently, I did not recover information from it. I only extracted the information of Banos in order to have all the districts mentioned in the population census of 1780, although I omit this observation due to concerns about the quality of data.

Cuenca

Book 1 (1792)

Location:

Archivo Nacional de Historia. Seccion: Presidencia de Quito. Caja: 132. Libro: 305. Documento: 8612.

	Colonial District	Description	Census
1	Banos	Complete	Yes
2	Cumbe	Complete	No
3	Giron	Complete	Yes
4	Canaribamba	Complete	Yes
5	Ona	Complete	Yes
6	Nabon	Complete	No
7	San Bartolome	Complete	Yes
8	Paccha	Complete	Yes
9	Jadan	Complete	No
10	Sigsig	Complete	No
11	Gualaceo	Complete	Yes
12	Guachapala	Complete	No
13	Paute	Complete	Yes
14	Azogues	Complete	Yes
15	Canar	Complete	Yes
16	Deleg	Complete	Yes
17	Sidcay	Complete	No

Missing districts from census: (1) Cuenca, (2) Sayausi.
Notes: This is the only source of *concertaje* for the colonial province Cuenca.

Guaranda

Notes:

1. No tax records have been found for this province. Hence, the sample for this province come from Oberem (1981). The town of Simiatug is not included because of missing information (see below).

Ibarra

Book 1 (1784)

Location:

Archivo Nacional de Historia. Seccion: Presidencia de Quito. Caja: 83. Libro: 207. Documento: 8607.

	Colonial District	Description	Census
1	Tulcan	Complete	Yes
2	Guaca	Complete	Yes
3	Tusa	Complete	Yes
4	Puntal	Complete	Yes
5	Mira	Complete	Yes
6	Salinas	Complete	Yes
7	Cahuasqui	Complete	Yes
8	Intag	Complete	Yes
9	San Antonio de Caranqui	Complete	Yes
10	Asuncion de Caranqui	Complete	Yes
11	Pimampiro	Complete	Yes
12	Ibarra	Complete	Yes
13	Otavalo	Complete	Yes

Missing districts from census: None.
Notes:

1. The district of Otavalo is included in the records, although it didn't belong to the colonial administration of Ibarra. Seems that the records of Otavalo reflect unpaid taxes for several districts of its jurisdiction. I did not use those records, as they seem to aggregate information from various places.

Book 2 (1817)

Location:

Archivo Nacional de Historia. Seccion: Tributarios. Caja: 29. Documento: 27.

	Colonial District	Description	Census
1	Tulcan	Complete	Yes
2	Guaca	Complete	Yes
3	Tusa	Complete	Yes
4	Puntal	Complete	Yes
5	El Angel	Complete	No
6	Mira	Complete	Yes
7	Salinas	Complete	Yes
8	Cahuasqui	Complete	Yes
9	Intag	Complete	Yes
10	San Antonio de Caranqui	Complete	Yes
11	Asuncion de Caranqui	Complete	Yes
12	Pimampiro	Complete	Yes
13	Ibarra	Complete	Yes

Missing districts from census: None.
Notes:

1. Apparently Puntal split into Puntal and El Angel after 1784 since some of the names of the *haciendas* reported in 1784 appear in the location of El Angel in 1817.

Book 3 (1818)

Location:

Archivo Nacional de Historia. Seccion: Tributarios. Caja: 35. Documento: 12.

	Colonial District	Description	Census
1	Tulcan	Complete	Yes
2	Guaca	Complete	Yes
3	Tusa	Complete	Yes
4	Puntal	Complete	Yes
5	El Angel	Complete	No
6	Mira	Complete	Yes
7	Salinas	Complete	Yes
8	Cahuasqui	Complete	Yes
9	Intag	Complete	Yes
10	San Antonio de Caranqui	Complete	Yes
11	Asuncion de Caranqui	Complete	Yes
12	Pimampiro	Complete	Yes
13	Ibarra	Complete	Yes

Missing districts from census: None.
Notes:

1. Apparently Puntal split into Puntal and El Angel after 1784 since some of the

names of the *haciendas* reported in 1784 appear in the location of El Angel in 1817.

Latacunga

Book 1 (1828)

Location:

Archivo Nacional de Historia. Seccion: Presidencia de Quito. Caja: 259. Libro: 645. Documento: 13228.

	Colonial District	Description	Census
1	Alaquez	Complete	Yes
2	Mulalo	Complete	Yes
3	Tanicuchi	Complete	Yes
4	Toacazo	Complete	Yes
5	Saquisili	Complete	Yes
6	Poalo	Complete	No
7	Pujili	Complete	Yes
8	Cusubamba	Complete	Yes
9	San Miguel	Complete	Yes
10	San Felipe	Complete	Yes
11	Sigchos	Complete	Yes
12	Chugchilan	Complete	No
13	Isinlivi	Complete	Yes
14	Angamarca	Complete	Yes
15	San Sebastian*	Incomplete	Yes
16	Latacunga*	Incomplete	Yes

Missing districts from census: None.
Notes:

1. This book seems to be a partial list of the tax collected for the second semester of the year. Opposite to the headlines of most of the other books, this one refers to the "contribucion personal" (personal contribution), instead of the collection of the taxes of San Juan and Navidad. One can also note that this is an incomplete and partial list since the *concertaje* rates are systematically lower than those reported by Oberem (1981) (see Figure A.4).

2. San Sebastian, although incomplete, is preserved. The only information missing for that district are two pages from the *hacienda* section. It is assumed that each page reported 20 workers, which is the average per page.

3. The data of Latacunga is missing. However, at the beginning of the book the index indicates that the data of that district only comprises two pages. Its absence does not appear to be a major loss.

4. Apparently Sigchos split into Sigchos and Chugchilan after 1805, while Saquisili split into Saquisili and Poalo. Due to the closeness of those district I merged them back to both compare with the information from Oberem (1981) and have

a more accurate report of the *concertaje* rates given the partial report of the book. Similarly, the information of Angamarca is omitted since the *concertaje* rates are way below what is reported by Oberem (1981).

Loja

Book 1 (1786)

Location:

Archivo Nacional de Historia. Seccion: Presidencia de Quito. Caja: 96. Libro: 233. Documento: 8610.

	Colonial District	Description	Census
1	San Sebastian de Loja	Complete	Yes
2	San Juan del Valle	Complete	Yes
3	Saraguro	Complete	Yes
4	Catacocha	Complete	Yes
5	Santiago	Complete	Yes
6	Malacatos	Complete	Yes
7	Chuquiribamba	Complete	Yes
8	Gonzanama	Complete	Yes
9	Cariamanga	Complete	Yes
10	Sozoranga	Complete	Yes
11	Guachanama	Complete	Yes
12	Celica	Complete	Yes
13	Zamora	Complete	Yes
14	Valladolid	Complete	Yes
15	Chito	Complete	Yes
16	Zaruma	Complete	Yes
17	Yulug	Complete	Yes

Missing districts from census: (1) Loja. Read below.
Notes:

1. Loja, the capital, does not appear in the tax records by itself.

Book 2 (1792)

Location:

Archivo Nacional de Historia. Seccion: Presidencia de Quito. Caja: 132. Libro: 305. Documento: 8611.

	Colonial District	Description	Census
1	San Sebastian de Loja	Complete	Yes
2	San Juan del Valle	Complete	Yes
3	Saraguro	Complete	Yes
4	Catacocha	Complete	Yes
5	Santiago	Complete	Yes
6	Malacatos	Complete	Yes
7	Chuquiribamba	Complete	Yes
8	Gonzanama	Complete	Yes
9	Cariamanga	Complete	Yes
10	Sozoranga	Complete	Yes
11	Guachanama	Complete	Yes
12	Celica	Complete	Yes
13	Zamora	Complete	Yes
14	Valladolid	Complete	Yes
15	Chito	Complete	Yes
16	Zaruma	Complete	Yes
17	Yulug	Complete	Yes

Missing districts from census: (1) Loja. Read below.
Notes:

1. Loja, the capital, does not appear in the tax records by itself.

Otavalo

Book 1 (1785)

Location:

Archivo Nacional de Historia. Seccion: Presidencia de Quito. Caja: 83. Libro: 207. Documento: 8608.

	Colonial District	Description	Census
1	Tocachi	Incomplete	Yes
2	Tabacundo	Incomplete	Yes
3	Cayambe	Incomplete	Yes
4	San Pablo	Incomplete	Yes
5	Cotacachi*	Incomplete	Yes
6	Urcuqui	Complete	Yes
7	Atuntaqui	Complete	Yes
8	Ibarra	Complete	Yes
9	Otavalo	Complete	Yes

Missing districts from census: (1) Tumbabiro.
Notes:

1. As with the case of Otavalo in Ibarra, the district of Ibarra is included in the records, although it didn't belong to the tax administration of Otavalo. For the same reasons mentioned before, I don't use the data of this district.

2. Cotacachi is incomplete since the book starts with the *hacienda* section. However I use it, since one can recover the number of free workers given that at the end the book it's reported the total number of workers.

Book 2 (1820)

Location:

Archivo Nacional de Historia. Seccion: Tributarios. Caja: 35.

	Colonial District	Description	Census
1	Tocachi	Complete	Yes
2	Tabacundo	Complete	Yes
3	Cangagua	Complete	No
4	Cayambe	Complete	Yes
5	San Pablo	Complete	Yes
6	Atuntaqui	Complete	Yes
7	Urcuqui	Complete	Yes
8	Cotacachi	Complete	Yes
9	Otavalo	Complete	Yes

Missing districts from census: (1) Tumbabiro.
Notes:

1. Apparently Cayambe split into Cayambe and Cangagua after 1784.

Book 3 (1830)

Location:

Archivo Nacional de Historia. Seccion: Indigenas. Caja: 174. Documento: 7.

	Colonial District	Description	Census
1	Tocachi	Complete	Yes
2	Tabacundo	Complete	Yes
3	Cangagua	Complete	No
4	Cayambe	Complete	Yes
5	San Pablo	Complete	Yes
6	Atuntaqui	Complete	Yes
7	Cotacachi	Complete	Yes
8	Otavalo	Complete	Yes

Missing districts from census: (1) Tumbabiro, (2) Urcuqui.
Notes:

1. Apparently Cayambe split into Cayambe and Cangagua after 1784.

2. This source is just a summary of *haciendas*, *conciertos* attached to them, and free workers.

Quito

Book 1 (1785)

Location:

Archivo Nacional de Historia. Seccion: Presidencia de Quito. Caja: 96. Libro: 235. Documento: 8610.

	Colonial District	Description	Census
1	Chimbacalle	Complete	Yes
2	Maria Magdalena	Complete	Yes
3	Chillogallo	Complete	Yes
4	Aloag	Complete	Yes
5	Aloasi	Complete	Yes
6	Machachi	Complete	Yes
7	Uyumbicho	Complete	Yes
8	Amaguana	Complete	Yes
9	Sangolqui	Complete	Yes
10	Pintag	Complete	Yes
11	Alangasi	Complete	Yes
12	Conocoto	Complete	Yes
13	Guapulo	Complete	Yes
14	Cumbaya	Complete	Yes
15	Tumbaco	Complete	Yes
16	Puembo	Complete	Yes
17	Yaruqui	Complete	Yes
18	Quinche	Complete	Yes
19	Guayllabamba	Complete	Yes
20	Perucho	Complete	Yes
21	San Antonio	Complete	Yes
22	Calacali	Complete	Yes
23	Pomasqui	Complete	Yes
24	Zambiza	Complete	Yes
25	Cotocollao	Complete	Yes
26	Santa Clara	Complete	Yes
27	San Blas*	Complete	No
28	Santa Barbara*	Complete	No
29	San Roque*	Complete	No
30	San Sebastian*	Complete	No
31	San Marcos*	Complete	No
32	Catedral*	Incomplete	No

Missing districts from census: (1) Quito, (2) Guali, (3) Canzacoto, (4) Nanigal, (5) Mindo. Read below.

Notes:

1. Quito, the capital, does not appear in the tax records by itself. It is assumed that San Blas, Santa Barbara, San Roque, San Sebastian and San Marcos were the parishes that comprised Quito. Catedral is also assumed to be part of Quito, although its records are incomplete.

2. Guali, Canzacoto, Nanigal and Mindo belong to the subtropical region of Quito.

Riobamba

Book 1 (1778)

Location:

Archivo Nacional de Historia. Seccion: Presidencia de Quito. Caja: 83. Libro: 207. Documento: 8606.

	Colonial District	Description	Census
1	Riobamba	Incomplete	Yes
2	Cajabamba	Complete	Yes
3	Licto	Complete	Yes
4	Guamote	Complete	Yes
5	Punin	Complete	Yes
6	Chambo	Complete	Yes
7	Quimiag	Complete	Yes
8	Yaruquies	Complete	Yes
9	San Luis	Complete	Yes
10	Penipe	Complete	Yes
11	Guanando	Complete	Yes
12	Ilapo	Complete	Yes
13	Cubijies	Complete	Yes
14	Guano	Complete	Yes
15	San Andres	Complete	Yes
16	Lican	Complete	Yes
17	Calpi	Complete	Yes
18	Pungala	Complete	Yes

Notes:

1. The data of several districts seem incomplete. It refers to the collection of taxes of indigenous workers belonging to certain encomiendas, instead of mention the payments of the taxes of San Juan and Navidad. Hence, I dismiss this information.

Book 2 (1812)

Location:

Archivo Nacional de Historia. Seccion: Tributarios. Caja: 34.

	Colonial District	Description	Census
1	Riobamba	Complete	Yes
2	Lican	Complete	Yes
3	Yaruquies	Complete	Yes
4	San Luis	Complete	Yes
5	Licto	Complete	Yes
6	Pungala	Complete	Yes
7	Punin	Complete	Yes
8	Guamote	Complete	Yes
9	Cebadas	Complete	Yes
10	Columbe	Complete	Yes
11	Cicalpa	Complete	Yes
12	Cajabamba	Complete	Yes
13	Calpi	Complete	Yes
14	San Andres	Complete	Yes
15	Guano	Complete	Yes
16	Ilapo	Complete	Yes
17	Cubijies	Complete	Yes
18	Penipe	Complete	Yes
19	Guanando	Complete	Yes
20	Quimiag	Complete	Yes
21	Chambo	Complete	Yes

Notes:

1. The district of Guamote does not appear in the population census of 1780.

Oberem (1981)

In his essay, he published three partial lists relative to *concertaje* and *haciendas* circa 1805. The first one includes, for every tax administration, districts with their respective *concertaje* rates. However, the rates of *concertaje* are only reported for districts with the three highest and three lowest rates, relative to their respective administration. The second one reports a partial list of districts with a different measure of the incidence of the institution, namely *conciertos* per *hacienda*, accompanied by the number of *haciendas*. As with the first one, this measure is only reported for a small group of districts, those with the three highest and three lowest rates of *conciertos* per *hacienda* relative to the administration they belonged. The third one is a general list for the whole region where he only reported the districts that had rates of *concertaje* above 75% or below 25%, but did not report any specific number. The three lists do not fully overlap. Additionally, he reported the aggregate number of *conciertos*, *haciendas* and tributaries at a colonial tax administration level. Oberem mentions that this information was obtained from a private archive which, despite my best efforts, has not been possible to locate. Nevertheless, when I compare the data

with that from the National Archive of History, the numbers are very close, which gives me confidence about its truthfulness.

For both measures of *concertaje*, i.e, either *conciertos/tributaries* or *conciertos/haciendas*, I have districts where I observe i) both the numerator and denominator, ii) only the numerator or iii) none. Given these limitations, I address the problem of incomplete information by doing the following procedure. First I classify districts with missing observations under three categories: a) those where I can infer the number of *conciertos*, but not the total number of tributaries, b) those that I cannot infer either the number of *conciertos* or the number of tributaries but are mentioned in some list and c) those that I cannot infer either the number of *conciertos* or the number of tributaries and are not mentioned in any part of the essay, but are assumed to be in the original source based on the list of districts reported in the population census of 1784. For all cases, what I can impute is the rate of *concertaje* under certain boundaries, which allows me to have an educated guess of the number of *conciertos* and tributaries, given all the extra information provided. This procedure is only used as a robustness check. Now, I describe it with an example.

Consider Table 2 below. It shows the districts of the tax administration of Guaranda. The text mentions that this administration had 9 districts. From those 9, we know the *concertaje* rates of 6 of them (3 lowest, 3 highest). The lowest of the highest *concertaje* rates was 0.39 and the highest of the lowest was 0.22. Then, we have missing information for the following districts: Chapacoto, Asancoto and Simiatug, that belong to type a), b) and c), respectively. First, it is mentioned that Chapacoto had 2 *conciertos* per *hacienda* and 6 *haciendas*, that is, the district had 12 *conciertos*, as reported in brackets. Given those 12 *conciertos*, the rate of *concertaje* for the district was in the interval [0.22, 0.39]. Furthermore, it is mentioned that Chapacoto had a *concertaje* rate below 0.25. Hence, we could narrow the interval to [0.22, 0.25]. Given those 12 *conciertos*, it should have been the case that the district had between 48 and 55 tributaries. Now we only have missing data for the two other districts. Let's move to Asancoto, a district for which we only know that had a *concertaje* rate below 0.25. Again, the rate of *concertaje* for the district was in the interval [0.22, 0.25]. In that case we can't actually tell the number of *conciertos* and tributaries. Nevertheless, given that we know the total number of workers in the administration, we can tell the number of *conciertos* to be distributed among the two districts with missing information. In this case, it was reported to be 586 *conciertos* in the whole administration, and so far the other districts sum a total of 506 workers. Hence, both districts should add up 80 *conciertos*. I split those 80 evenly between both, and then recover the respective interval of tributaries. For the case of Simiatug, that is not reported in any list, I assume it to be one of the districts that belonged to the administration because it appears in the population census of 1784. In that case, the only thing we know is that its *concertaje* rate, if reported, must have been between 0.25 and 0.39. Since I already assigned 40 *conciertos* to it, I can compute the interval of tributaries ([102, 160]). Off course, the number of *conciertos* for Asancoto and Simiatug could be very different, but their *rates* are consistent with all the information available.

Lastly, I follow the exact procedure to impute the number of *haciendas* when they are not reported. That is, given that for every colonial tax administration I know the districts with the three highest and three lowest *conciertos* per *hacienda*, I can

deduce the number of *haciendas* a district in that administration had. Tables A.1-A.8 reports the data of this exercise for each tax administration. The information imputed is in brackets. Additionally, Figures A.1 and A.4 summarizes the data by tax administration and compare the *concertaje* rates of 1805 with the information obtained from the National Archive of History.

Table A.1: *Concertaje* Rates 1805: Ambato

Colonial District	Modern District	Modern Province	Tax Administration	Conciertos	Tributarios	Concertaje	Haciendas	Conciertos by Hacienda
Quero	Quero	Tungurahua	Ambato	400	475	0.84	18	22
Pillaro	Pillaro	Tungurahua	Ambato	475	724	0.66	47	10
Patate	Patate	Tungurahua	Ambato	191	296	0.65	[16, 19]	[10, 12]
Santa Rosa[a]	Santa Rosa	Tungurahua	Ambato	[420]	[649, 857]	[0.49, 0.65]	35	12
Izamba[a]	Izamba	Tungurahua	Ambato	[140]	[215, 286]	[0.49, 0.65]	20	7
Tisaleo[a]	Tisaleo	Tungurahua	Ambato	[215]	[331, 439]	[0.49, 0.65]	43	5
Ambato	Ambato	Tungurahua	Ambato	546	1123	0.49	[46, 55]	[10, 12]
Pelileo	Pelileo	Tungurahua	Ambato	378	876	0.43	23	16
Quisapincha	Quisapincha	Tungurahua	Ambato	138	806	0.17	[12, 14]	[10, 12]
Imputed				2903	[5495, 5882]	[0.49, 0.53]	[260, 274]	[11, 11]
Reported				2892	5627	0.51	270	11

Notes The number of districts in the tax administration is not mentioned. The district of Banos, although mentioned in the population census, is not included because 1) is not mentioned in the text and 2) the total sum of *conciertos* already exceed the total reported, so it's assumed that the district was not included in the original list.

Table A.2: *Concertaje* Rates 1805: Guaranda

Colonial District	Modern District	Modern Province	Tax Administration	Conciertos	Tributarios	Rate	Haciendas	Conciertos by Hacienda
Pallatanga	Pallatanga	Chimborazo	Guaranda	38	91	0.42	[6, 19]	[2, 6]
San Miguel	San Miguel	Bolivar	Guaranda	69	176	0.39	12	6
Guaranda	Guaranda	Bolivar	Guaranda	242	620	0.39	25	10
Chapacoto[a]	Magdalena	Bolivar	Guaranda	[12]	[48, 55]	[0.22, 0.25]	6	2
Asancoto[b]	Asuncion	Bolivar	Guaranda	[40*]	[160, 182]	[0.22, 0.25]	[7, 20]	[2, 6]
Simiatug[c]	Simiatug	Bolivar	Guaranda	[40*]	[102, 160]	[0.25, 0.39]	[7, 20]	[2, 6]
Guanujo	Guanujo	Bolivar	Guaranda	122	557	0.22	22	6
Chimbo	Chimbo	Bolivar	Guaranda	13	60	0.22	6	2
Santiago	Santiago	Bolivar	Guaranda	10	89	0.11	5	2
Imputed				586	[1903, 1990]	[0.29, 0.31]	[96, 135]	[4, 6]
Reported				586	1951	0.3	108	5

Notes The number of districts reported in the tax administration is 9. I conserve the data from the imputation of Asancoto and Chapacoto because the *concertaje* rates are in a very close interval. That is not the case for the town of Simiatug.

Table A.3: *Concertaje* Rates 1805: Ibarra

Colonial District	Modern District	Modern Province	Tax Administration	Conciertos	Tributarios	Rate	Haciendas	Conciertos by Hacienda
Pimampiro	Pimampiro	Imbabura	Ibarra	73	87	0.84	5	15
Cahuasqui	Cahuasqui	Imbabura	Ibarra	80	127	0.63	[6, 10]	[8, 13]
El Angel	El Angel	Carchi	Ibarra	65	134	0.48	[5, 8]	[8, 13]
Guaca[a]	Guaca	Carchi	Ibarra	[8]	[32, 40]	[0.20, 0.25]	1	8
Puntal[a]	Puntal	Carchi	Ibarra	[40]	[83, 160]	[0.25, 0.48]	5	8
Ibarra[a]	Ibarra	Imbabura	Ibarra	[294]	[613, 1176]	[0.25, 0.48]	21	14
Mira[c]	Mira	Imbabura	Ibarra	[70*]	[146, 280]	[0.25, 0.48]	[5, 9]	[8, 13]
Salinas[c]	Salinas	Imbabura	Ibarra	[70*]	[146, 280]	[0.25, 0.48]	[5, 9]	[8, 13]
Intag[c]	Intag	Imbabura	Ibarra	[70*]	[146, 280]	[0.25, 0.48]	[5, 9]	[8, 13]
San Antonio[c]	San Antonio	Imbabura	Ibarra	[70*]	[146, 280]	[0.25, 0.48]	[5, 9]	[8, 13]
Tulcan	Tulcan	Carchi	Ibarra	27	131	0.20	4	7
Caranqui	Caranqui	Imbabura	Ibarra	27	136	0.20	2	13
Tusa	Tusa	Carchi	Ibarra	29	163	0.18	[2, 4]	[8,13]
Imputed				923	[2090, 3274]	[0.28,0.44]	[71, 96]	[10, 13]
Reported				925	2175	0.43	76	12

Notes The number of districts reported in the tax administration is 13.

Table A.4: *Concertaje* Rates 1805: Latacunga

Colonial District	Modern District	Modern Province	Tax Administration	Conciertos	Tributarios	Rate	Haciendas	Conciertos by Hacienda
Cusubamba	Cusubamba	Cotopaxi	Latacunga	387	508	0.76	[18,39]	[10, 21]
Pujili	Cahuasqui	Cotopaxi	Latacunga	956	1302	0.73	46	21
San Sebastian	San Sebastian	Cotopaxi	Latacunga	746	1232	0.60	[36, 75]	[10, 21]
Saquisili[a]	Saquisili	Cotopaxi	Latacunga	[810]	[1350, 2250]	[0.36, 0.60]	15	54
Tanicuchi[a]	Tanicuchi	Cotopaxi	Latacunga	[250]	[417, 694]	[0.36, 0.60]	25	10
San Felipe[a]	San Felipe	Cotopaxi	Latacunga	[160]	[267, 444]	[0.36, 0.60]	16	10
Alaquez[b]	Alaquez	Cotopaxi	Latacunga	[205*]	[320, 533]	[0.36, 0.60]	[9, 21]	[10, 21]
San Miguel[b]	San Miguel	Cotopaxi	Latacunga	[205*]	[320, 533]	[0.36, 0.60]	[9, 21]	[10, 21]
Angamarca[c]	Angamarca	Cotopaxi	Latacunga	[205*]	[320, 533]	[0.36, 0.60]	[9, 21]	[10, 21]
Mulalo[c]	Mulalo	Cotopaxi	Latacunga	[205*]	[320, 533]	[0.36, 0.60]	[9, 21]	[10, 21]
Toacazo	Toacazo	Cotopaxi	Latacunga	105	289	0.36	[5, 11]	[10, 21]
Isinlivi	Isinlivi	Cotopaxi	Latacunga	179	555	0.33	4	45
Sigchos	Sigchos	Cotopaxi	Latacunga	103	371	0.28	10	10
Imputed				4516	[7571, 9777]	[0.46, 0.60]	[196, 323]	[14, 23]
Reported				4515	8282	0.55	242	19

Notes The number of districts in the tax administration is not mentioned.

Table A.5: *Concertaje* Rates 1805: Loja

Colonial District	Modern District	Modern Province	Tax Administration	Conciertos	Tributarios	Rate	Haciendas	Conciertos by Hacienda
Chuquiribamba	Chuquiribamba	Loja	Loja	99	200	0.49	5	20
Malacatos	Malacatos	Loja	Loja	40	88	0.45	[5, 10]	[4, 8]
Loja	Loja	Loja	Loja	50	132	0.38	13	4
Saraguro[a]	Saraguro	Loja	Loja	[84]	[336, 1200]	[0.07, 0.25]	6	14
Guachanama[a]	Guachanama	Loja	Loja	[8]	[32, 114]	[0.07, 0.25]	1	8
Gonzanama[b]	Gonzanama	Loja	Loja	[18*]	[72, 257]	[0.07, 0.25]	[2, 5]	[4, 8]
Yulug[b]	Yulug	Loja	Loja	[18*]	[72, 257]	[0.07, 0.25]	[2, 5]	[4, 8]
Cariamanga[b]	Cariamanga	Loja	Loja	[18*]	[72, 257]	[0.07, 0.25]	[2, 5]	[4, 8]
Catacocha[c]	Catacocha	Loja	Loja	[18*]	[47, 72]	[0.25, 0.38]	[2, 5]	[4, 8]
Celica[c]	Celica	Loja	Loja	[18*]	[47, 72]	[0.25, 0.38]	[2, 5]	[4, 8]
Zaruma[c]	Zaruma	Loja	El Oro	[18*]	[47, 72]	[0.25, 0.38]	[2, 5]	[4, 8]
Valladolid[c]	Valladolid	Loja	Zamora	[18*]	[47, 72]	[0.25, 0.38]	[2, 5]	[4, 8]
Chito[c]	Chito	Loja	Zamora	[18*]	[47, 72]	[0.25, 0.38]	[2, 5]	[4, 8]
Zamora[c]	Zamora	Loja	Zamora	[18*]	[47, 72]	[0.25, 0.38]	[2, 5]	[4, 8]
San Juan del Valle	Toacazo	Loja	Loja	20	267	0.07	[3,5]	[4, 8]
Santiago	Santiago	Loja	Loja	2	41	0.05	1	2
Sozoranga	Sozoranga	Loja	Loja	1	220	0.005	1	1
Imputed				466	[1814, 3465]	[0.13, 0.27]	[53, 87]	[5, 9]
Reported				465	2832	0.16	64	7

Notes The number of districts reported in the tax administration is 16.

Table A.6: *Concertaje* Rates 1805: Otavalo

Colonial District	Modern District	Modern Province	Tax Administration	Conciertos	Tributarios	Rate	Haciendas	Conciertos by Hacienda
Cayambe	Cayambe	Pichincha	Otavalo	470	555	0.85	8	59
Cangagua	Cangagua	Pichincha	Otavalo	319	433	0.74	6	53
Tabacundo	Tabacundo	Pichincha	Otavalo	221	301	0.73	11	20
San Pablo[a]	San Pablo	Imbabura	Otavalo	[222]	[304, 584]	[0.38, 0.73]	6	37
Tocachi[a]	Tocachi	Pichincha	Otavalo	[144]	[197, 379]	[0.38, 0.73]	9	16
Urcuqui[c]	Urcuqui	Imbabura	Otavalo	[112*]	[153, 295]	[0.38, 0.73]	[3, 6]	[20, 37]
Atuntaqui	Atuntaqui	Imbabura	Otavalo	81	213	0.38	7	12
Otavalo	Otavalo	Imbabura	Otavalo	720	1903	0.38	[19, 36]	[20, 37]
Cotacachi	Cotacachi	Imbabura	Otavalo	448	1223	0.37	[12, 22]	[20, 37]
Imputed				2737	[5282, 5886]	[0.46, 0.52]	[81, 111]	[25, 34]
Reported				2737	5705	0.48	97	28

Notes The number of districts reported in the tax administration is 9.

Table A.7: *Concertaje* Rates 1805: Quito

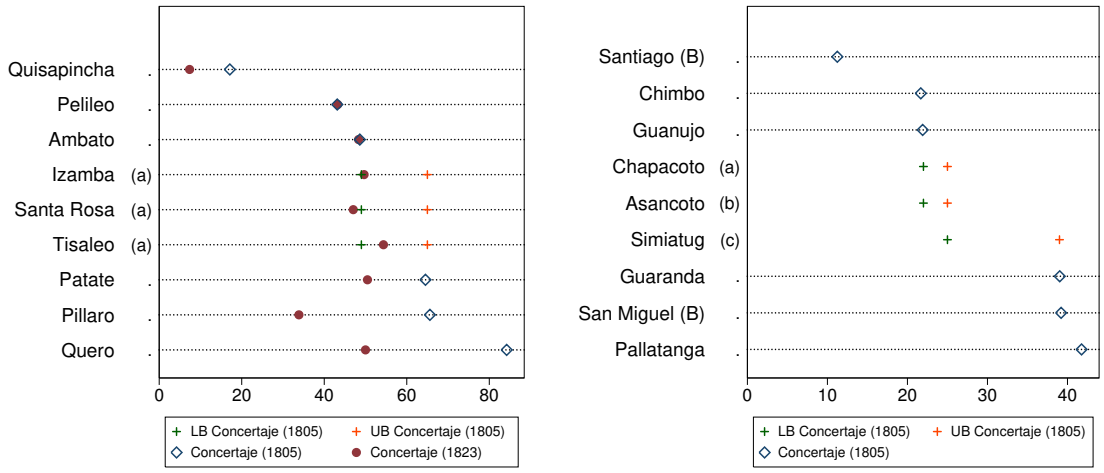
Colonial District	Modern District	Modern Province	Tax Administration	Conciertos	Tributarios	Rate	Haciendas	Conciertos by Hacienda
Cotocollao	Cotocollao	Pichincha	Quito	284	346	0.82	[12, 41]	[7, 23]
Tumbaco	Tumbaco	Pichincha	Quito	303	346	0.81	13	23
Perucho	Perucho	Pichincha	Quito	107	136	0.79	[5, 15]	[7, 23]
Yaruqui[b]	Yaruqui	Pichincha	Quito	[400*]	[506, 533]	[0.75, 0.79]	[17, 57]	[7, 23]
Amaguana[a]	Amaguana	Pichincha	Quito	[338]	[463, 1300]	[0.26, 0.75]	13	26
Pintag[a]	Pintag	Pichincha	Quito	[276]	[378, 1062]	[0.26, 0.75]	12	23
Chillogallo[a]	Chillogallo	Pichincha	Quito	[224]	[309, 862]	[0.26, 0.75]	32	7
Guapulo[a]	Guapulo	Pichincha	Quito	[28]	[38, 108]	[0.26, 0.75]	4	7
Magdalena	Magdalena	Pichincha	Quito	56	218	0.26	8	7
Alangasi	Alangasi	Pichincha	Quito	43	253	0.17	[2, 7]	[7, 23]
Zambiza	Zambiza	Pichincha	Quito	87	670	0.13	[4, 12]	[7, 23]
Imputed				2146	[,]	[,]	[,]	[,]
Reported				3968	7422	0.53	323	12

Notes The number of districts in the administration is not mentioned.

Table A.8: *Concertaje* Rates 1805: Riobamba

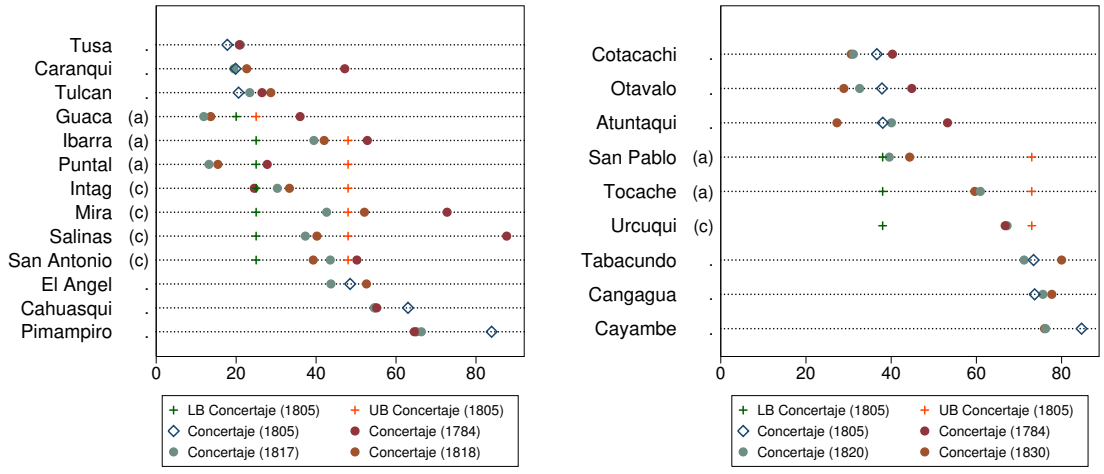
Colonial District	Modern District	Modern Province	Tax Administration	Conciertos	Tributarios	Rate	Haciendas	Conciertos by Hacienda
Pungala	Pungala	Chimborazo	Riobamba	198	229	0.86	[7, 18]	[11, 27]
Columbe	Columbe	Chimborazo	Riobamba	320	383	0.83	[12, 29]	[11, 27]
Cebadas	Cebadas	Chimborazo	Riobamba	205	252	0.81	[8, 19]	[11, 27]
Licto[a]	Licto	Chimborazo	Riobamba	[629]	[839, 1797]	[0.35, 0.75]	17	37
Calpi[a]	Calpi	Chimborazo	Riobamba	[476]	[635, 1360]	[0.35, 0.75]	17	28
Cubijies[a]	Cubijies	Chimborazo	Riobamba	[54]	[72, 154]	[0.35, 0.75]	2	27
Penipe[a]	Penipe	Chimborazo	Riobamba	[44]	[59, 126]	[0.35, 0.75]	4	11
San Luis[a]	San Luis	Chimborazo	Riobamba	[33]	[44, 94]	[0.35, 0.75]	3	11
Guanando[a]	Guanando	Chimborazo	Riobamba	[27]	[36, 77]	[0.35, 0.75]	3	9
Punin[b]	Punin	Chimborazo	Riobamba	[255*]	[340, 729]	[0.35, 0.75]	[9, 23]	[11, 27]
Lican[b]	Lican	Chimborazo	Riobamba	[255*]	[340, 729]	[0.35, 0.75]	[9, 23]	[11, 27]
Riobamba[b]	Riobamba	Chimborazo	Riobamba	[255*]	[340, 729]	[0.35, 0.75]	[9, 23]	[11, 27]
Cicalpa[c]	Cicalpa	Chimborazo	Riobamba	[255*]	[340, 729]	[0.35, 0.75]	[9, 23]	[11, 27]
Quimiag[c]	Quimiag	Chimborazo	Riobamba	[255*]	[340, 729]	[0.35, 0.75]	[9, 23]	[11, 27]
Cajabamba[c]	Cajabamba	Chimborazo	Riobamba	[255*]	[340, 729]	[0.35, 0.75]	[9, 23]	[11, 27]
Guamote[c]	Guamote	Chimborazo	Riobamba	[255*]	[340, 729]	[0.35, 0.75]	[9, 23]	[11, 27]
Chambo[c]	Chambo	Chimborazo	Riobamba	[255*]	[340, 729]	[0.35, 0.75]	[9, 23]	[11, 27]
Cicalpa[c]	Cicalpa	Chimborazo	Riobamba	[255*]	[340, 729]	[0.35, 0.75]	[9, 23]	[11, 27]
San Andres	San Andres	Chimborazo	Riobamba	149	422	0.35	[6, 14]	[11, 27]
Guano	Guano	Chimborazo	Riobamba	126	492	0.26	[5, 11]	[11, 27]
Yaruquies	Yaruquies	Chimborazo	Riobamba	115	629	0.18	[4, 10]	[11, 27]
Imputed				4671	[7112, 12576]	[0.37, 0.66]	[169, 354]	[13, 27]
Reported				4673	8816	0.53	229	20

Notes The number of districts in the tax administration is 21.



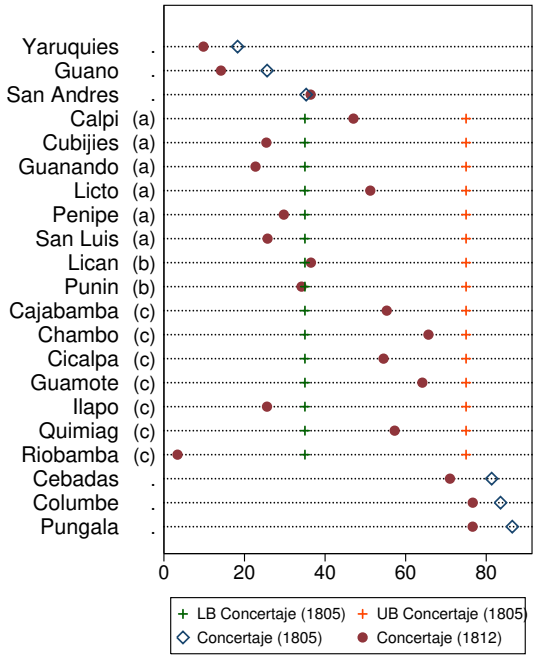
(a) Ambato (b) Guaranda

Figure A.1: Ambato and Guaranda: Comparison of *Concertaje* Rates for Different Years

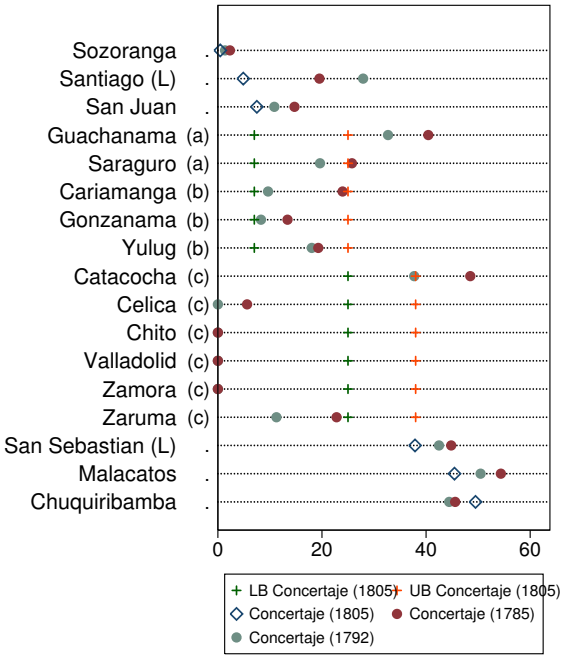


(a) Ibarra (b) Otavalo

Figure A.2: Ibarra and Otavalo: Comparison of *Concertaje* Rates for Different Years

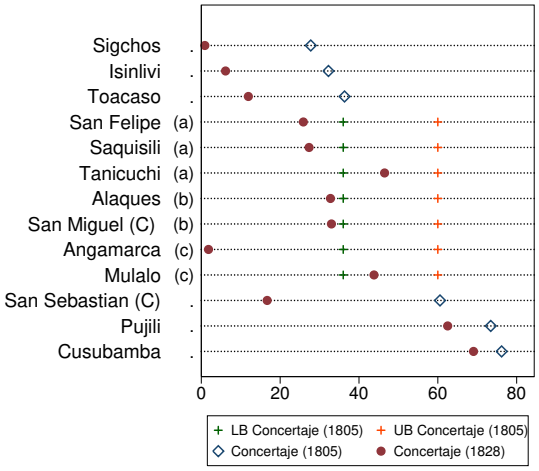


(a) Riobamba

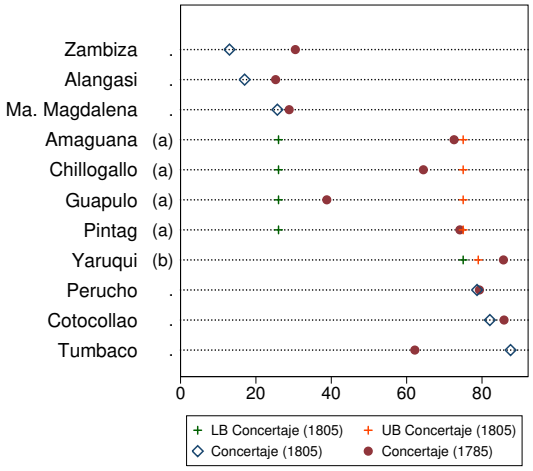


(b) Loja

Figure A.3: Riobamba and Loja: Comparison of *Concertaje* Rates for Different Years



(a) Latacunga



(b) Quito

Figure A.4: Latacunga and Quito: Comparison of *Concertaje* Rates for Different Years

Measurement Error

Since the data come from several sources, I tried to minimize the measurement error by employing the following procedure. When two or more sources exist for one province, I kept those with the highest quality, defined by its completeness and readability. So I dismiss the books of Alausi (1826), Ambato (1825), and Riobamba (1778). On top of that, I selected only books before 1830, that is, before Ecuador's independence. This is because during the first years of the republican period the nation had a reduced administrative capacity (Saint Geours (1994)), so its records are less trustworthy.

Summary

To sum up, I collected *concertaje* data for all of the 10 colonial tax administrations of the Ecuadorian highlands. Excluding the central tax districts, from the 127 districts reported in the population census of 1784, I have data of *concertaje* for 121 of them. I don't have information of Mindo, Guali, Nanigal and Cansacoto for the province of Quito; Sayausi for the province of Cuenca; and Tumbabiro for the province of Otavalo. Additionally, I have information for 6 out of 10 central tax districts. I don't have information of Latacunga, Loja, Cuenca and Quito, although I have information of 5 parishes of Quito.

Additionally, I have information of *concertaje* for the 13 following districts that do not appear in the census: El Angel, in the administration of Ibarra; Cangagua in the administration of Otavalo; Guamote, in the administration of Riobamba; Cumbe, Nabon, Jadan, Sigsing, Guachapala and Sidcay in the administration of Cuenca; and Tixan, Guasuntos, Chunchi and Sibambe in the administration of Alausi. That gives a total of 145 districts. In the next section I describe the matching of colonial districts with contemporary ones.

Matching

From the 145 districts for which I have *concertaje* data, I disregard the information of Chimbacalle, Maria Magdalena, Guapulo, Cotocollao, Santa Clara, San Blas, Santa Barbara, San Roque, San Sebastian de Quito and San Marcos that would have matched with the contemporary parish of Quito for the following reasons. First, its tax records are incomplete both in the records of 1794 and in the list of Oberem. More important, Oberem (1981) mentions that the numbers reported for the district of Quito are of low quality, due that many indigenous workers camouflaged as *mestizos* (mixed between Spanish and Indigenous workers). Finally, given that Quito has been historically the main administrative center of the region, *concertaje* was minimal, so including it in the sample could overestimate the effects of the institution. That leave us to match 135 districts.

I am able to match those 135 districts with 129 contemporary parishes by proceeding as follows. First, I used the official political division of Ecuador to merge districts that have joined over time to form a modern parish. That is the case of 4 places: Guanujo and Guaranda, which belong to the modern parish of Guaranda; Caranqui and Ibarra, which formed the parish of Ibarra; and Sicalpa and Cajabamba, which are now the parish of Villa del la Union, and Riobamba and Yaruquies, which are now the parish of Riobamba. Then, even when I do not have information of *concertaje* for the districts of Latacunga and Loja, I assume that San Sebastian de Latacunga and San Felipe, and San Sebastian de Loja and San Juan del Valle, represent their respective *concertaje* data since they are parishes of those places. Finally, 120 districts (93% of the sample) are matched simply by name. The remaining 5 districts are matched via other sources. From those 129 districts, three of them are dropped: Banos (T), Angamarca and Simiatug for the reasons explained above. I then give a code to each colonial district using today's administrative codes. The table below summarizes the matching process.

	Colonial District		Modern District	Tax Admin.	Modern Province	Matching Criteria
1	Alausi	1	Alausi	Alausi	Chimborazo	Name
2	Guasuntos	2	Guasuntos	Alausi	Chimborazo	Name
3	Sibambe	3	Sibambe	Alausi	Chimborazo	Name
4	Tixan	4	Tixan	Alausi	Chimborazo	Name
5	Chunchi	5	Chunchi	Alausi	Chimborazo	Name
6	Ambato	6	Ambato	Ambato	Tungurahua	Name
7	Banos	7	Banos	Ambato	Tungurahua	Name
8	Izamba	8	Izamba	Ambato	Tungurahua	Name
9	Patate	9	Patate	Ambato	Tungurahua	Name
10	Pelileo	10	Pelileo	Ambato	Tungurahua	Name
11	Pillaro	11	Pillaro	Ambato	Tungurahua	Name
12	Quero	12	Quero	Ambato	Tungurahua	Name
13	Quisapincha	13	Quisapincha	Ambato	Tungurahua	Name
14	Santa Rosa	14	Santa Rosa	Ambato	Tungurahua	Name
15	Tisaleo	15	Tisaleo	Ambato	Tungurahua	Name
16	Azogues	16	Azogues	Cuenca	Canar	Name
17	Banos	17	Banos	Cuenca	Azuay	Name
18	Canar	18	Canar	Cuenca	Canar	Name
19	Canaribamba	19	Santa Isabel	Cuenca	Azuay	Other
20	Cumbe	20	Cumbe	Cuenca	Azuay	Name
21	Deleg	21	Deleg	Cuenca	Canar	Name
22	Giron	22	Giron	Cuenca	Azuay	Name
23	Guachapala	23	Guachapala	Cuenca	Azuay	Name
24	Gualaceo	24	Gualaceo	Cuenca	Azuay	Name
25	Jadan	25	Jadan	Cuenca	Azuay	Name
26	Nabon	26	Nabon	Cuenca	Azuay	Name
27	Ona	27	Ona	Cuenca	Azuay	Name
28	Paccha	28	Paccha	Cuenca	Azuay	Name
29	Paute	29	Paute	Cuenca	Azuay	Name
30	San Bartolome	30	San Bartolome	Cuenca	Azuay	Name
31	Sidcay	31	Sidcay	Cuenca	Azuay	Name
32	Sigsig	32	Sigsig	Cuenca	Azuay	Name
33	Asancoto	33	Asuncion	Guaranda	Bolivar	Name
34	Chapacoto	34	Magdalena	Guaranda	Bolivar	Name
35	Chimbo	35	Chimbo	Guaranda	Bolivar	Name
36	Guanujo			Guaranda		
		36	Guaranda		Bolivar	Merged
37	Guaranda			Guaranda		
38	Pallatanga	37	Pallatanga	Guaranda	Bolivar	Name
39	San Miguel	38	San Miguel	Guaranda	Bolivar	Name
40	Santiago	39	Santiago	Guaranda	Bolivar	Name
41	Simiatug	40	Simiatug	Guaranda	Bolivar	Name
42	Cahuasqui	41	Cahuasqui	Ibarra	Imbabura	Name
43	El Angel	42	El Angel	Ibarra	Carchi	Name

Continued...

	Colonial District		Modern District	Tax Admin.	Modern Province	Matching Criteria
44	Huaca	43	Huaca	Ibarra	Carchi	Name
45	Caranqui			Ibarra		
		44	Ibarra		Imbabura	Merged
46	Ibarra			Ibarra		
47	Intag	45	Imantag	Ibarra	Imbabura	Other
48	Mira	46	Mira	Ibarra	Carchi	Name
49	Pimampiro	47	Pimampiro	Ibarra	Imbabura	Name
50	Puntal	48	Bolivar	Ibarra	Carchi	Other
51	Salinas	49	Salinas	Ibarra	Imbabura	Name
52	San Antonio	50	San Antonio	Ibarra	Imbabura	Name
53	Tulcan	51	Tulcan	Ibarra	Carchi	Name
54	Tusa	52	San Gabriel	Ibarra	Carchi	Other
55	Alaques	53	Alaques	Latacunga	Cotopaxi	Name
56	Angamarca	54	Angamarca	Latacunga	Cotopaxi	Name
57	Cusubamba	55	Cusubamba	Latacunga	Cotopaxi	Name
58	Isinlivi	56	Isinlivi	Latacunga	Cotopaxi	Name
59	San Felipe			Latacunga		
		57	Latacunga		Cotopaxi	Merged
60	San Sebastian			Latacunga		
61	Mulalo	58	Mulalo	Latacunga	Cotopaxi	Name
62	Pujili	59	Pujili	Latacunga	Cotopaxi	Name
63	San Miguel	60	San Miguel	Latacunga	Cotopaxi	Name
64	Saquisili	61	Saquisili	Latacunga	Cotopaxi	Name
65	Sigchos	62	Sigchos	Latacunga	Cotopaxi	Name
66	Tanicuchi	63	Tanicuchi	Latacunga	Cotopaxi	Name
67	Toacaso	64	Toacaso	Latacunga	Cotopaxi	Name
68	Cariamanga	65	Cariamanga	Loja	Loja	Name
69	Catacocha	66	Catacocha	Loja	Loja	Name
70	Celica	67	Celica	Loja	Loja	Name
71	Chito	68	Chito	Loja	Zamora Ch.	Name
72	Chuquiribamba	69	Chuquiribamba	Loja	Loja	Name
73	Gonzanama	70	Gonzanama	Loja	Loja	Name
74	Guachanama	71	Guachanama	Loja	Loja	Name
75	San Sebastian			Loja		
		72	Loja		Loja	Merged
76	San Juan			Loja		
77	Malacatos	73	Malacatos	Loja	Loja	Name
78	Santiago	74	Santiago	Loja	Loja	Name
79	Saraguro	75	Saraguro	Loja	Loja	Name
80	Sozoranga	76	Sozoranga	Loja	Loja	Name
81	Valladolid	77	Valladolid	Loja	Zamora Ch.	Name
82	Yulug	78	Yulug	Loja	Loja	Name
83	Zamora	79	Zamora	Loja	Zamora Ch.	Name
84	Zaruma	80	Zaruma	Loja	El Oro	Name
85	Atuntaqui	81	Atuntaqui	Otavalo	Imbabura	Name

Continued...

	Colonial District		Modern District	Tax Admin.	Modern Province	Matching Criteria
86	Cangagua	82	Cangagua	Otavalo	Pichincha	Name
87	Cayambe	83	Cayambe	Otavalo	Pichincha	Name
88	Cotacachi	84	Cotacachi	Otavalo	Imbabura	Name
89	Otavalo	85	Otavalo	Otavalo	Imbabura	Name
90	San Pablo	86	San Pablo	Otavalo	Imbabura	Name
91	Tabacundo	87	Tabacundo	Otavalo	Pichincha	Name
92	Tocache	88	Tocache	Otavalo	Pichincha	Name
93	Urcuqui	89	Urcuqui	Otavalo	Imbabura	Name
94	Alangasi	90	Alangasi	Quito	Pichincha	Name
95	Aloag	91	Aloag	Quito	Pichincha	Name
96	Aloasi	92	Aloasi	Quito	Pichincha	Name
97	Amaguana	93	Amaguana	Quito	Pichincha	Name
98	Calacali	94	Calacali	Quito	Pichincha	Name
99	Chillogallo	95	Lloa	Quito	Pichincha	Other
100	Conocoto	96	Conocoto	Quito	Pichincha	Name
101	Cumbaya	97	Cumbaya	Quito	Pichincha	Name
102	Guayllabamba	98	Guayllabamba	Quito	Pichincha	Name
103	Machachi	99	Machachi	Quito	Pichincha	Name
104	Perucho	100	Perucho	Quito	Pichincha	Name
105	Pintag	101	Pintag	Quito	Pichincha	Name
106	Pomasqui	102	Pomasqui	Quito	Pichincha	Name
107	Puembo	103	Puembo	Quito	Pichincha	Name
108	Quinche	104	Quinche	Quito	Pichincha	Name
109	Sangolqui	105	Sangolqui			
110	San Antonio	106	San Antonio	Quito	Pichincha	Name
111	Tumbaco	107	Tumbaco	Quito	Pichincha	Name
112	Uyumbicho	108	Uyumbicho	Quito	Pichincha	Name
113	Yaruqui	109	Yaruqui	Quito	Pichincha	Name
114	Zambiza	110	Zambiza	Quito	Pichincha	Name
115	Cajabamba			Riobamba		
		111	Villa de la Union		Chimborazo	Merged
116	Sicalpa			Riobamba		
117	Calpi	112	Calpi	Riobamba	Chimborazo	Name
118	Cebadas	113	Cebadas	Riobamba	Chimborazo	Name
119	Chambo	114	Chambo	Riobamba	Chimborazo	Name
120	Columbe	115	Columbe	Riobamba	Chimborazo	Name
121	Cubijies	116	Cubijies	Riobamba	Chimborazo	Name
122	Guamote	117	Guamote	Riobamba	Chimborazo	Name
123	Guanando	118	Guanando	Riobamba	Chimborazo	Name
124	Guano	119	Guano	Riobamba	Chimborazo	Name
125	Ilapo	120	Ilapo	Riobamba	Chimborazo	Name
126	Lican	121	Lican	Riobamba	Chimborazo	Name
127	Licto	122	Licto	Riobamba	Chimborazo	Name
128	Penipe	123	Penipe	Riobamba	Chimborazo	Name
129	Pungala	124	Pungala	Riobamba	Chimborazo	Name

Continued...

	Colonial District		Modern District	Tax Admin.	Modern Province	Matching Criteria
130	Punin	125	Punin	Riobamba	Chimborazo	Name
131	Quimiag	126	Quimiag	Riobamba	Chimborazo	Name
132	Riobamba			Riobamba		
		127	Riobamba		Chimborazo	Name
133	Yaruquies			Riobamba		
134	San Andres	128	San Andres	Riobamba	Chimborazo	Name
135	San Luis	129	San Luis	Riobamba	Chimborazo	Name

The 2010 census reports 452 parishes in the Ecuadorian Highlands, while I've been able to match 125 of them (the parishes of Zamora, Chito and Valladolid belong to the Amazon province of Zamora Chinchipe, while Zaruma belongs to the coastal province of El Oro). That represents 27% of the total number of highland parishes. Figures A.5 to A.11 provides the location of the towns matched and the tax route during colonial times according to the order they appear in the tax records.

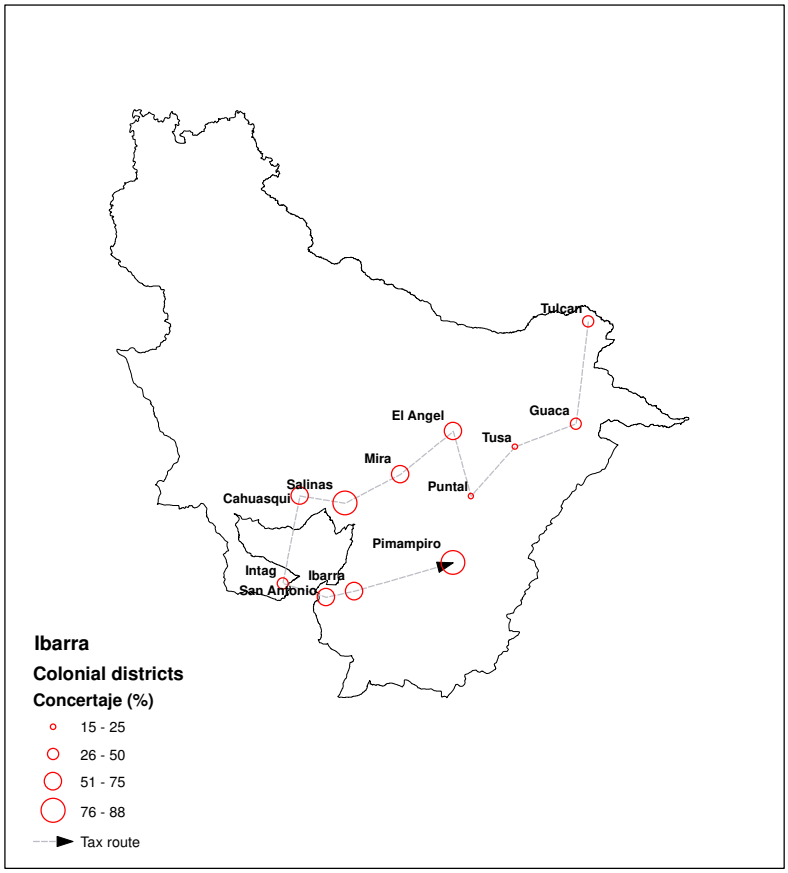


Figure A.5: Ibarra: Towns and Tax Route

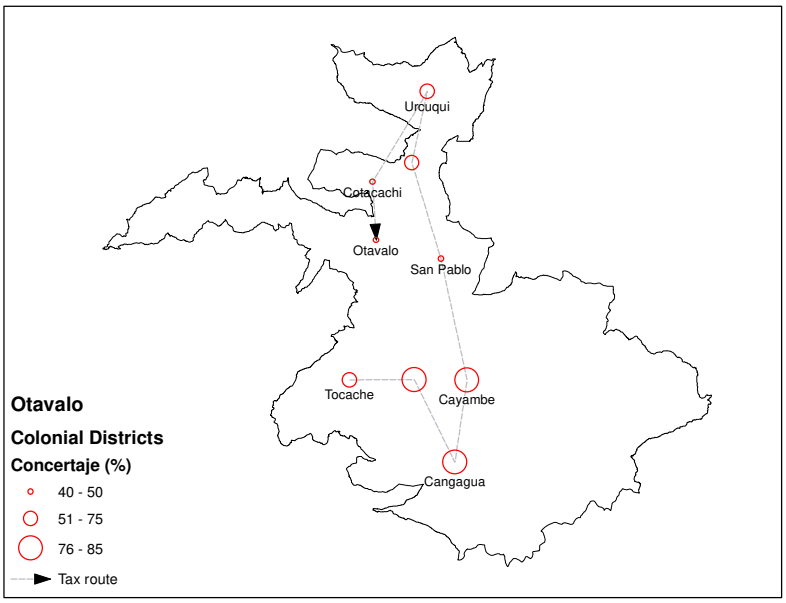


Figure A.6: Otavalo: Towns and Tax Route

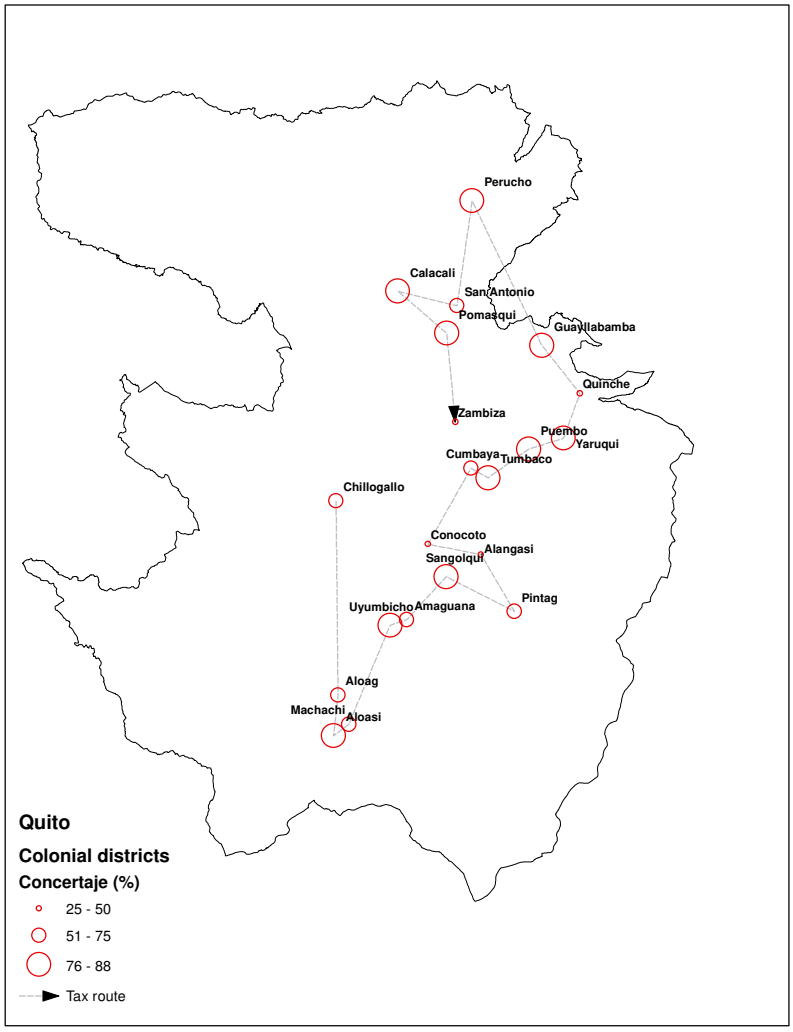


Figure A.7: Quito: Towns and Tax Route

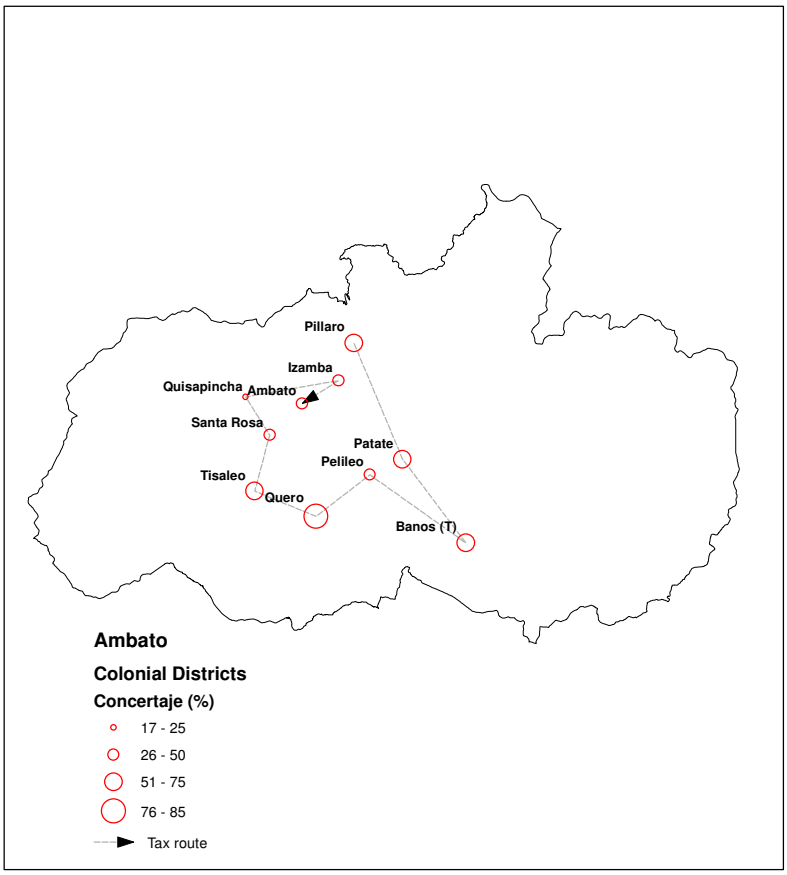


Figure A.8: Ambato: Towns and Tax Route

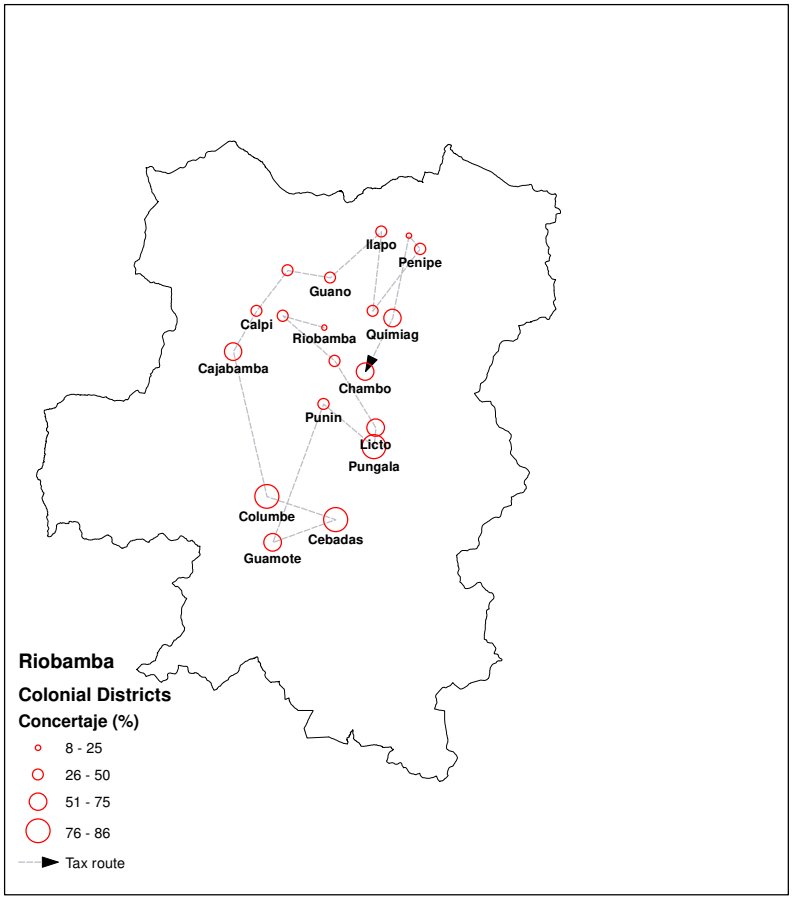


Figure A.9: Riobamba: Towns and Tax Route

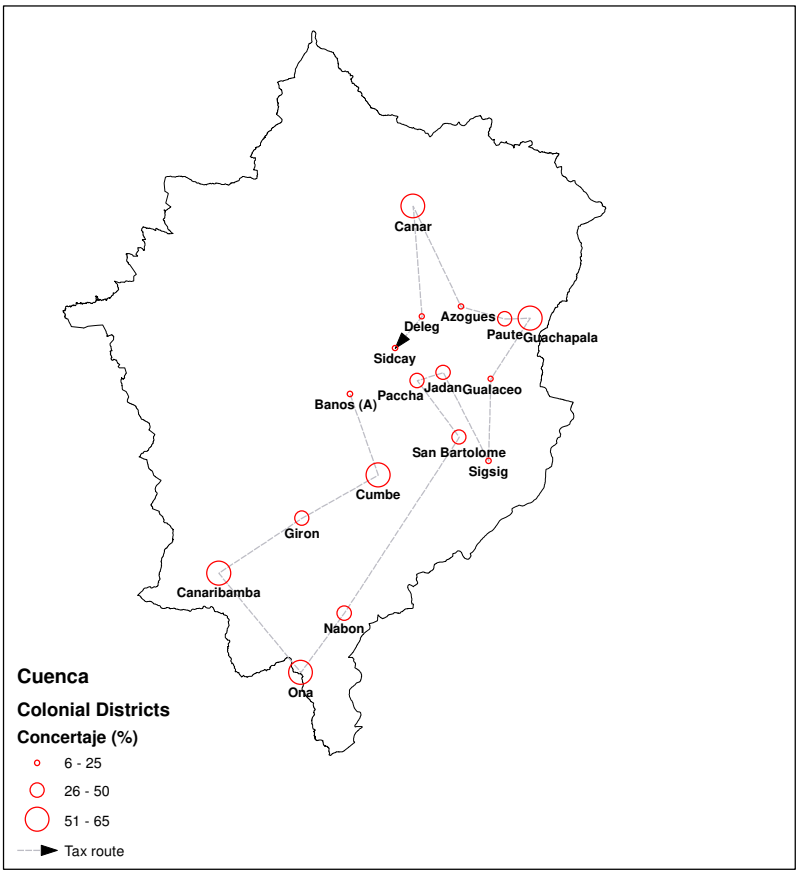


Figure A.10: Cuenca: Towns and Tax Route

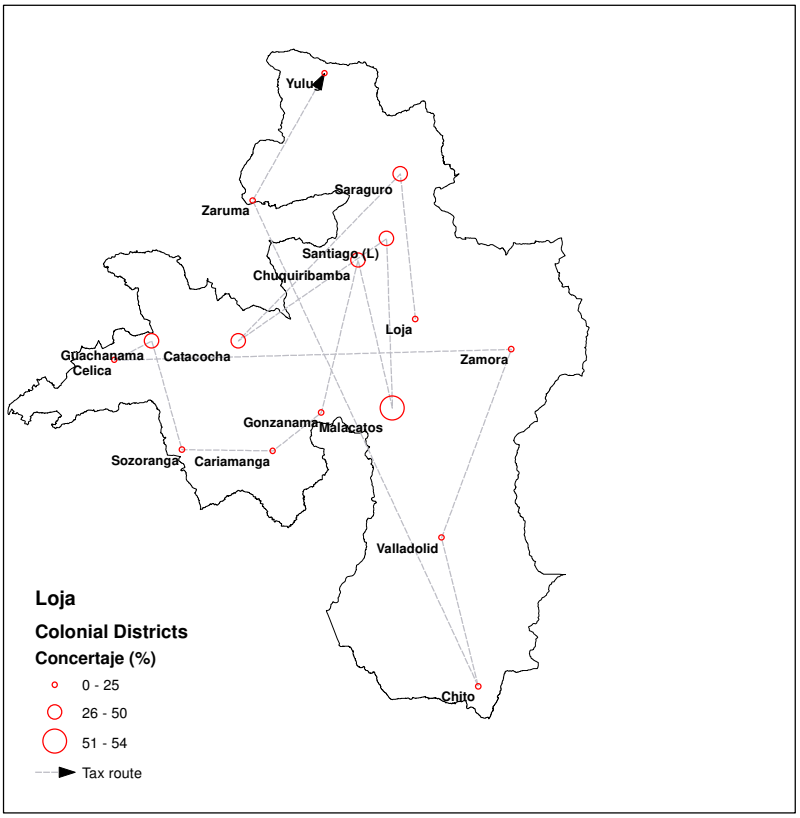


Figure A.11: Loja: Towns and Tax Route

Colonial Data

Population (1590-1600)

Population data for the first years of the colonial period comes from several sources. Tyrer (1988) report the data from 1590 and 1591 using as sources "La relacion de Zaruma" and "El censo de Morales Figueroa", so I use the numbers from there. Data for 1598 comes from Albuja Mateus (1998). I combine all those sources and compute the average population *circa* 1600. The matching with other records was done by comparing names of towns, similar to the procedure explained above, although keep in mind that some of those towns may incorporate the information of several others.

Population (1784)

Data for colonial population around the time of the *concertaje* data comes from Salmoral (1994). In her article she collected population data around 1784, summarize it, and reported the most trustable records. I use precisely the latter to obtain the population of each district.

Geographic Controls

I obtain a map of the point location of the district's center from the Military Institute of Ecuador. From there, I used geospatial software to recover their coordinates. A GIS map with district administrative boundaries (2012) was obtained from the INEC. To compute distance to the capital, I calculate the Euclidean distance of each district's center to the modern province capital they belong.

Elevation data are from the Shuttle Radar Topography Mission (SRTM). Details about it can be read in Dell (2010). I use the SRTM data to obtain the point-elevation of the the district's center. I also used the data to compute the average elevation and slope within each district using today's administrative boundaries. In order to accurately compute elevation and slope, water bodies were removed. A GIS file of the waterbodies of Ecuador was obtained from DIVA-GIS (2016).

Crops

Data for potential yield was obtained from the GAEZ-FAO project. A very detailed explanation of how to obtain the raw files could be found in the *readme* document attached to the crop's data. In short, in the GAEZ application I selected, for the crops highland maize, potatoes and wheat, their respective Agro-climatically attainable yield under rain-fed water supply and low input, and download several raster files. Then, using geospatial software I project those rasters to the system of coordinates UTM 17S. While projecting, I resample the rasters in order to make them suitable to the size of districts[1]. Finally, using today's boundaries, I compute the average potential yield in each district. Figure A.12 shows the relative productivity differences in maizes versus potatoes for the Ecuadorian highlands.

[1]In ArcGIS, I employed the option *"BILINEAR"* in the Project Raster function given that the variable is continuous.

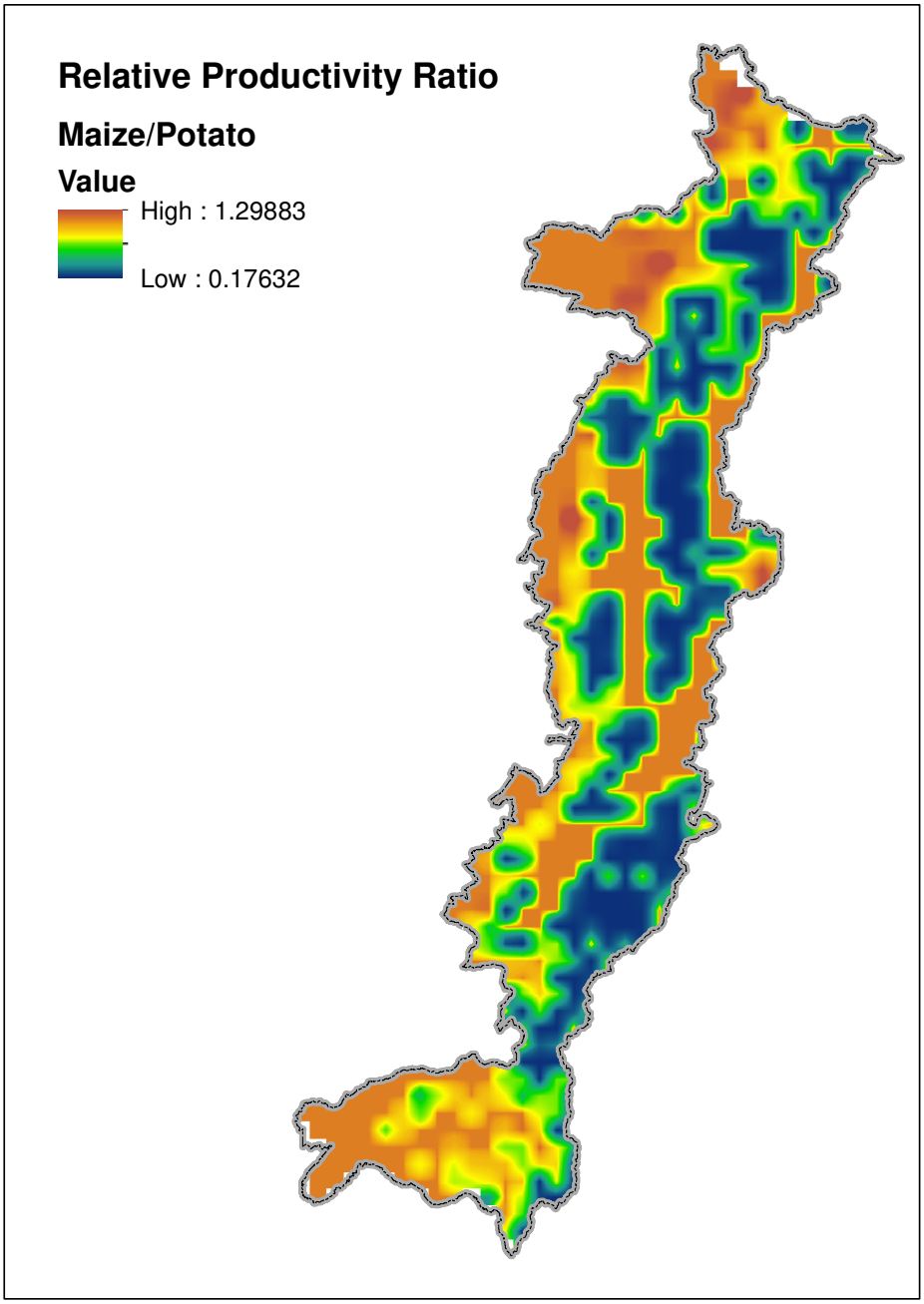


Figure A.12: Relative Productivity Differences in Maize and Potatoes

Notes: Ratio of productivity in highland maize (in tonnes/ha) relative to productivity in potatoes (in tonnes/ha). Source: Author's calculations based on data from the FAO GAEZ project.

Intermediate Outcomes

Land Value

Land value for 1909 is reported in the *Guia Comercial Agricola e Industrial de la Republica del Ecuador (1909)* Compañía Guía del Ecuador (1909). The book contains

detailed information about the country at the beginning of the century. It includes information, by district, of the total value of land as reported in the official cadastral records. This data has been scanned by the jesuit library *Aurelio Espinosa Polit*, and distributed in CD's. I transcribed the data into an spreadsheet.

Land Distribution

The information relative to land comes from the 1974 Agricultural Census. This data is only available in books, so I digitized them. The unit of observation used by the census is the Agricultural Unit of Production (Unidad de Produccion Agropecuaria (UPA) in Spanish), which is defined as any unit of land devoted partially or totally to agricultural production, regardless of size, location, type of ownership or legal status. This is the only Agricultural Census that reports information at a parish level.

Land Gini.

The Gini coefficient of land inequality is computed using data about the number and size of holdings. The size of each holding is classified as follows: (1) less than 0 hectares, (2) 0.01 to 0.5 hectares, (3) 0.5-1 hectares, (4) 1-2 hectares, (5) 2-3 hectares, (6) 3-4 hectares, (7) 4-5 hectares, (8) 5-10 hectares, (9) 10-20 hectares, (10) 10-20 hectares, (11) 20-50 hectares, (12) 50-100 hectares, (13) 100-500 hectares, (14) 500-1000 hectares, (15) 1000-2500 hectares, (16) more than 2500 hectares. I follow the same procedure as Nunn (2008), and calculate the Gini coefficient using the program *ineqdec0* (Jenkins et al. (2015)).

Distance to farm.

In the 1974 Agricultural Census there is information if the production unit (UPA) is close to a certain distance to a road that is accessible during the whole year ("Distancia a la via transitable todo el ao"). The categories are less than 1 km, between 1 and 5 km, between 5 and 10 km and above 10 km. I compute the fraction of production units on these different categories.

Illiteracy

Illiteracy data before 1990 comes from the population census of several years. Once again, I digitized the books since they are only available physically. I discuss in details of each census below.

Illiteracy (1962)

Illiteracy is reported relative to the population of 6 years and above that can't read or write (Table 2). Data for Quisapincha and the Oriental districts of Zamora, Chito, Valladolid is not reported.

Illiteracy (1974)

Illiteracy is reported relative to the population of 10 years and above that can't read or write (Table 21).

Illiteracy (1982)

Illiteracy is reported relative to the population of 10 years and above that can't read or write (Table 27). Data for Alausi and Sibambe is not reported.

Contemporary Outcomes

Data for contemporary outcomes come from the population census of 1990, 2001 and 2010, obtained from the *Sistema Integrado de Indicadores Sociales del Ecuador* (SIISE (2011)) and *Sistema Nacional de Informacion* (SNI (2011)).

Extreme Poverty

The definition of Extreme Poverty is described in the text (SIISE; Section: *Desigualdad y Pobreza*; Subsection: *Pobreza*; Variable: Extrema pobreza por necesidades basicas insatisfechas (NBI); Source: Censo de Poblacion y Vivienda; Years: 1990, 2001, 2010; Disaggregation: Parroquia.). Data for Lican in 2001 is not reported.

Illiteracy

Illiteracy is reported relative to the population of 15 years and above that can't read or write (SIISE; Section: *Educacion*; Subsection: *Educacion de la poblacion*; Variable: Analfabetismo; Source: Censo de Poblacion y Vivienda; Years: 1990, 2001, 2010; Disaggregation: Parroquia.). Data for Lican in 2001 is not reported.

Functional Illiteracy

Functional illiteracy is defined as the fraction of population 15 years and above with less that 3 years of primary education (SIISE; Section: *Educacion*; Subsection: *Educacion de la poblacion*; Variable: Analfabetismo funcional; Source: Censo de Poblacion y Vivienda; Year: 1990; Disaggregation: Parroquia.). Data for Lican in 2001 is not reported.

Years of School

Is the average years of school (approved) for the population above 24 years (SIISE; Section: *Educacion*; Subsection: *Cobertura y acceso a la Educacion*; Variable: Escolaridad; Source: Censo de Poblacion y Vivienda; Years: 1990, 2001, 2010; Disaggregation: Parroquia.). Data for Lican in 2001 is not reported.

Secondary Enrollment.

Secondary enrollment is defined as the fraction of population between 12 and 17 years that assist to school (SIISE; Section: *Educacion*; Subsection: *Cobertura y acceso a la Educacion*; Variable: Tasa bruta de asistencia; Source: Censo de Poblacion y Vivienda; Reference Population: Tasa de asistencia secundaria; Years: 1990, 2001, 2010; Disaggregation: Parroquia.). Data for Lican in 2001 is not reported.

Share of Workers in Different Occupations

Population in the Agricultural Sector

The working population in the agricultural sector is defined as the number of individuals that declared to be working in that sector. (SNI; Section: *Economia*; Subsection: *Ver Economia*; Variable: Poblacion ocupada en agricultura, silvicultura, caza y pesca; Source: Censo de Poblacion y Vivienda; Year: 1990; Disaggregation: Parroquia.)

Population in the Public Sector

The working population in the public sector is defined as the number of individuals that declared to be working in that sector. (SNI; Section: *Economia*; Subsection: *Ver Economia*; Variable: Poblacion ocupada en Administracion publica y defensa; Source: Censo de Poblacion y Vivienda; Year: 1990; Disaggregation: Parroquia.)

Working Population

The working population in is defined as the number of individuals above 10 years that during the census week declared a) worked at least one hour or b) had a job but didn't work (SNI; Section: *Economia*; Subsection: *Ver Economia*; Variable: Poblacion ocupada; Source: Censo de Poblacion y Vivienda; Year: 1990; Disaggregation: Parroquia.).

Satellite Lights

Night Light Intensity, as described in the text, is obtained from the National Centers for Environmental Information. I use geospatial software to compute the average night light intensity in a district by proceeding as follows. I first project the raster data to the coordinate system UTM 17S and then overlay a political map of Ecuador to calculate the average radiation in each district.

Road Density

Road density is defined as the total length of roads (considering changes in elevation) divided by the surface area of a district. Total length is computed using a GIS road network map of Ecuador available from the Military Geographic Institute of Ecuador. Roads are classified by type as National (3), Local (4), Primary (14), or Secondary (15), and by material: paved (1), not paved (2), temporary (3) or other (4). Data of surface come from the administrative boundaries provided by INEC and described in the geographic controls section.

APPENDIX B

SUPPLEMENTARY MATERIAL FOR CHAPTER 2

Data

Tax Forms

In this section I describe some details about the tax forms in Ecuador.

Firms

The two ways to operate a business in Ecuador is by either creating a Limited Liability Company or by registering as as Sole Proprietor. Limited Liability Companies have to fill the Form 101, while Sole Proprietors the Form 102. More properly, the 102 form is filled by individuals who report gross incomes above USD $300,000 other than labor earnings, which could include that coming from business activities. Hence, I make sure to only include in my sample individuals who report business income, as they are considered a firm for tax purposes. More specifically, I consider as Sole Proprietors the individuals who have filled the Form 102, and report non-missing business revenue in the field labeled as "Total Ingresos", which correspond to the lines 699 for the years 2009-2013 and 6999 to the years 2014-2016.

There is also a sub category within Sole Proprietors that is not consider in the analysis, referred as Regimen Impositivo Simplificado, which has been designed to incentivize informal firms to pay taxes. They are not considered in this study as tax records for this group has very limited information, given that the payments are a lump-sum.

Individuals

Individual income taxes in Ecuador are reported through two different forms. The first one, called Form 107, corresponds to people who only have incomes from dependent employment, and so their taxes are deducted directly from their monthly salaries. The Form 102 instead is filled individually by every person who has income from other sources different from employed labor (aside from business income, it includes self-employment, rents, dividends and capital income).

The Form 102 has a shorter version, namely the Form 102A, for individuals who report gross income less than USD $300,000. An individual that has both income from dependent employment and other sources must consolidate all the information in the 102 form.

Matching of Records

Regarding to the matching between the information of owners and managers and their corresponding tax records, Firms' ID in the tax records are anonymized due to privacy concerns. Hence, I first do the classification of firms as FM and EM using the records of the Bureau of Companies and then give the tax authority the classified list, so they can anonymize the firm's ID to properly join with their tax records.

Classification of Firms

Surnames

A significant part of the identification of family management inside a firm relies on the comparison of surnames between upper-managers and owners. Therefore, it is important to assure that this inference can be done with minimal error. The assumption used in this paper is that managers and owners that share a surname are family-related. This will be the case if surnames unequivocally reflect a family relation. Although one share surnames with a family member, surnames are not unique to each family. However, as long as they are distributed in such a way that there is a sufficient number of rare surnames (to avoid mislabeling a family relation), the identification through this method is a good approximation.

Consequently, this inference will be affected by the shape of the distribution of surnames in the society. Figure B.1 provide evidence that the distribution of surnames in Ecuador (for all individuals) is effectively skewed in the sense that although there are a few individuals with really common surnames, most have unusual ones. Recently Güell et al. (2015) have stressed how this property of the distribution of surnames in Western societies facilitate the identification of family connections, when true family linkages are unknown, and use it to recover estimates of intergenerational mobility from cross-sectional data.

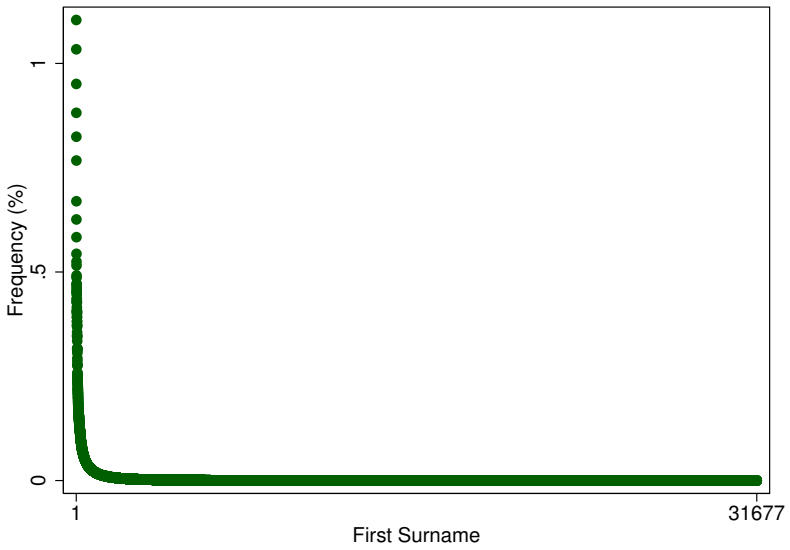


Figure B.1: Distribution of First Surnames

Notes: Surnames were obtained from individual tax records publicly reported by the tax office in their webpage for the period 2007-2016. Each number in the horizontal axis represent a surname, so in total there are displayed 31677 different surnames.

Since the focus of this paper is to identify relations between managers and owners within a very specific place, a firm, assuming the existence of a family linkage when a surname is shared is more plausible in this context, than say, making the same

inference for two randomly selected individuals. In addition, due to the access of administrative records of family linkages, one can effectively verify to which extent the surname methodology can identify family relations. The next section discuss the results of this exercise.

True Family Linkages

Because of privacy concerns, I do not have direct access to the network data. Instead, the SRI processed the information for me and return a list of family linkages between manager's and owners using the records of the Bureau of Companies. The information was anonymized accordingly, and in addition was codified in such a way that I cannot tell explicitly the relationship between two individuals. Instead, the relationships where labeled into three categories. The first one refers to immediate family, which includes parents, offsprings and spouses. The second refers to relations built from the first category, which includes siblings, grandparents and mothers and fathers in law. The third one includes cousins and uncles and more distant relatives as great-grand parents.

In the text I comment that the network data has missing linkages due that it has only included the individuals that have renewed their ID. However, another limitation of using this network for the future is that even if today everybody will renew their ID, individuals who passed away while having the previous ID will never show up in the network built by the SRI. For example, if the parents of an individual never renewed their ID and already passed away, that individual will appear in the network as an orphan, making it impossible to connect with her siblings or further relatives.

Prevalence of Family-Managed Firms

Life-cycle Differences

Raw Data

In order to provide a more transparent depiction of the data, I first show averages of firms' size across ages for Family and Non-Family managed firms, in terms of revenue and employment, by simply pooling firms for all the years data is available. Figure B.2 show the average size of both class of firms as measured by employment. As one can see, bin by bin, EM firms are not only in average larger, but their expansion across bins is also of a higher magnitude when compared to FM firms. At age 20, the gap between both type of firms is of an order of 5. Figure B.3 redo the exercise but considering revenue as a measure of size. In this case the results reflect the same pattern of employment, although the size of the age-differences is larger.

Further, Figure B.4 show the average size of Family and Non-Family managed firms but without considering Sole-Proprietors. This exercise shows that Sole-Proprietors push the average size of Family-Managed firms across age downwards, although the gap with Non-Family Managed firms still remain, which is now in the order of 2.5 at age 20. This occur despite I only consider relatively large Sole Proprietors (Form 102), which according to the tax authority are those that have reported annual revenue or costs higher than 300,000 USD, or have declared capital higher than 200,000

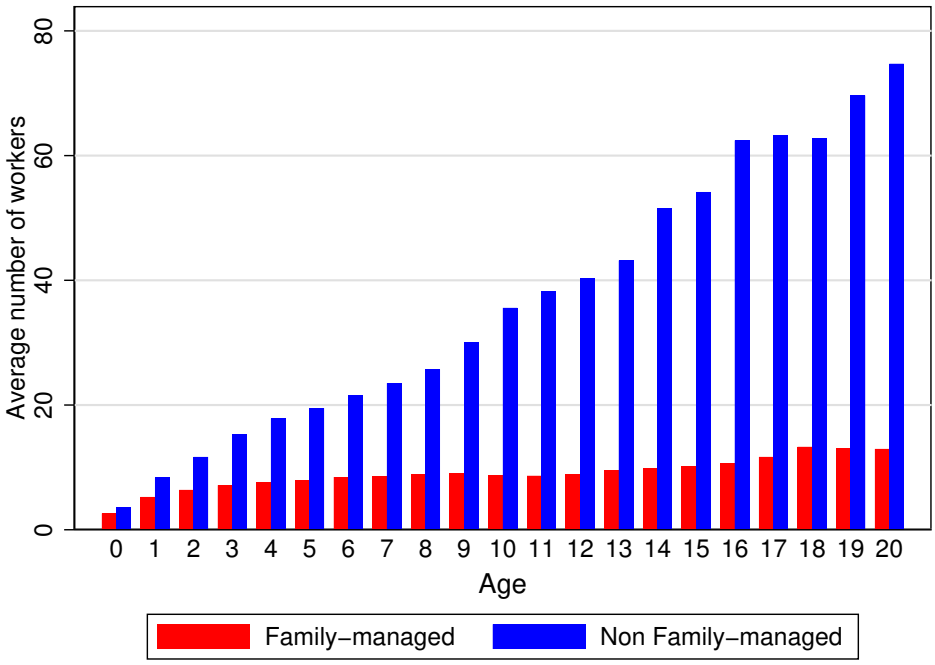


Figure B.2: Average Size of Firms Across Age in Terms of Employment by Type of Management

Notes: Firms were classified as FM and EM according to the cut-off rule defined in (2.1) with $\bar{c} = 50\%$. The sample includes Limited Liability Companies and Sole Proprietors in every sector, except the financial one, for the period 2009-2016. The averages were computed by simply pooling data in each age category. Employment data was computed by adding up the total number of workers from individual tax records (Form 107) by employer id.

USD[1].

As mentioned in the text, there are discrepancies in the information of employment when computed from the tax records of workers and that provided by the Bureau of Companies. These discrepancies can be noted by comparing Figure B.4, in which employment data is taken from tax records, with Figure B.4, which uses self-reported data from the Bureau of Companies. Note that this comparison is done only for Limited Liability Companies, as the Bureau does not have information of Sole Proprietors. Although there are still difference in size across ages, EM firms appear to be somehow smaller when using employment data from the Bureau of Companies.

In summary, simple exploration of the data through computation of averages by age in the cross-section already suggests that EM firms are larger, and face a rapid expansion as they age. The purpose of a more detailed estimation of the life-cycle patterns of both type of firms, as done in Section 2.3, is to control for several other forces that could be driving these gaps.

[1]A Sole Proprietor that exports, despite its size, is also considered in this group.

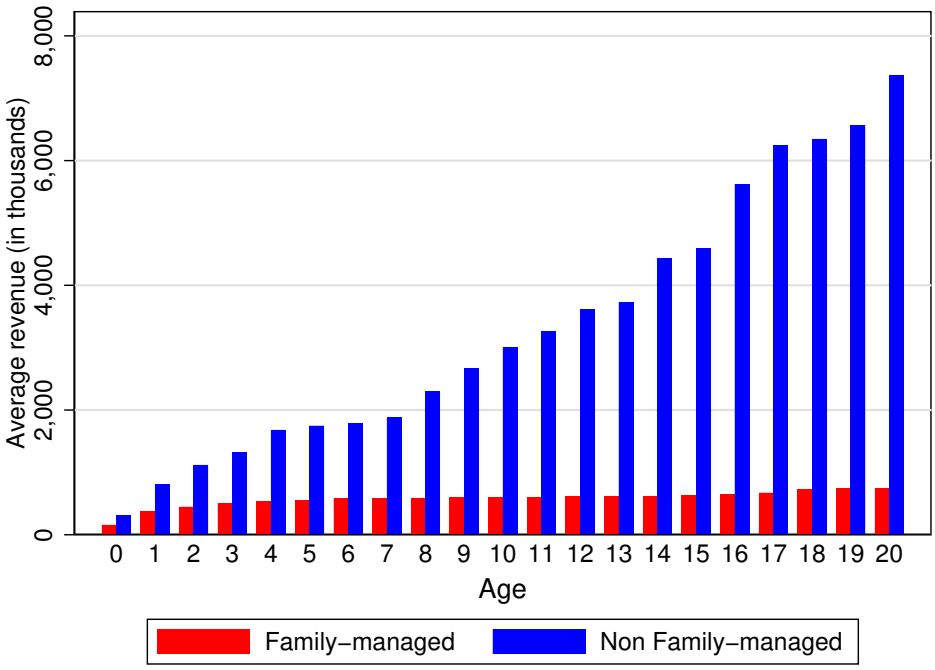


Figure B.3: Average Size of Firms Across Age in Terms of Revenue by Type of Management

Notes: Firms were classified as FM and EM according to the cut-off rule defined in (2.1) with $\bar{c} = 50\%$. The sample includes Limited Liability Companies and Sole Proprietors in every sector, except the financial one, for the period 2009-2016. The averages were computed by simply pooling data in each age category. Revenue is Real Revenue, deflated using the CPI.

Comparison with Hsieh and Klenow (2014)

For comparison purposes, I follow a similar methodology as Hsieh and Klenow (2014) and compute the life-cycle profile of firms treating the data as repeated cross sections, but controlling for 6-industry codes, and so exploiting the variation within industry. To do so, I estimate the following specification for each type of firm,

$$\log(e_{j,t}) = \alpha + \theta_c + \delta_t + \sum_{b \in B} \lambda_b D^b_{j,t} + \text{error}_{j,t} \tag{B.1}$$

where $e_{j,t}$ is the number of worker of firm j, at time t; θ_c represent a vector of 6-digit industry fixed effects; δ_t is a vector of time period dummies, while $B^a_{j,t}$ is a dummy variable that takes the value of one if the firm's age belong to the category $b \in B = \{< 5, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40+\}$, which are the same categories age categories used by Hsieh and Klenow (2014). The results of this exercise are plotted in Figure B.6.

Interestingly, this exercise produce results very similar to those reported in Hsieh and Klenow (2014). That is, the gap in life cycle gap profiles between EM and FM firms are on the same order of magnitude as the gap between USA and Mexico both

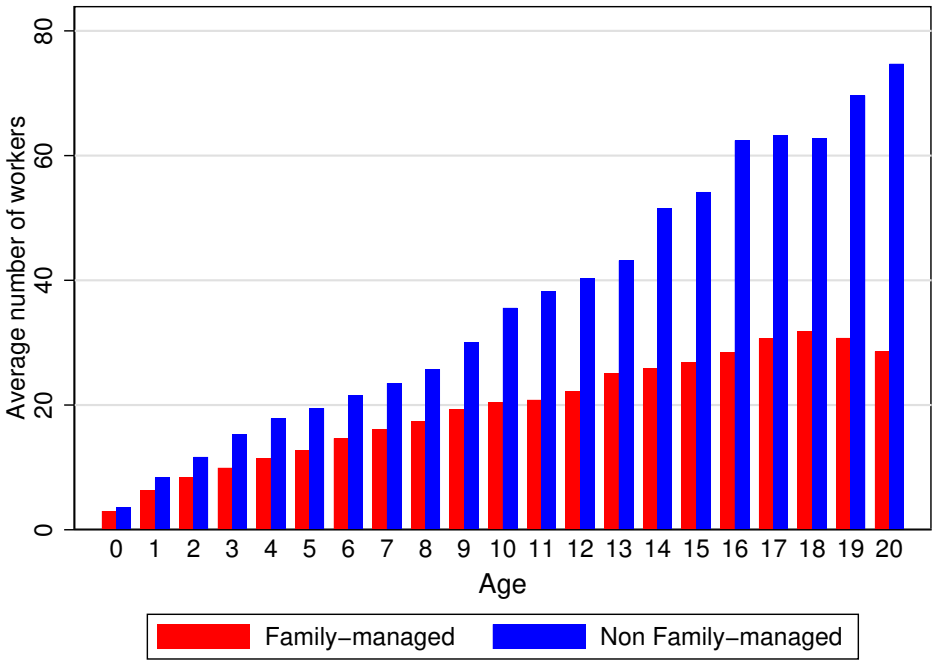


Figure B.4: Average Size of Firms Across Age in Terms of Employment by Type of Management (Only LLC)

Notes: Firms were classified as FM and EM according to the cut-off rule defined in (2.1) with $\bar{c} = 50\%$. The sample includes only Limited Liability Companies in every sector, except the financial one, for the period 2009-2016. The averages were computed by simply pooling data in each age category. Employment data was computed by adding up the total number of workers from individual tax records (Form 107) by employer id.

at age 20-24 and above 40. Note also that in this case the shape of the age-differences display a convex form, instead of the concave one that appear when exploiting the panel structure of the data.

Robustness of the Results to Sample Selection

As discussed above, the universe of firms considered in this study includes Sole Proprietors, which by definition are classified as Family-Managed firms, since ownership falls in one individual which at the same time is its Legal Representative, or the principal manager of the firm. Given that these firms represent 46% of the total number of observations, they may disproportionately affect the estimation of the life-cycle profile of FM firms. In order to see if that is the case, I estimate (2.2) but without considering this group. Figure B.7 show the results. Although the gap between both class of firms is still present, now its size reduces to 1.5 at age 20. Hence, even when Sole Proprietors push the gap upward (as anticipated by Figure B.4), the life-cycle differences continue to show.

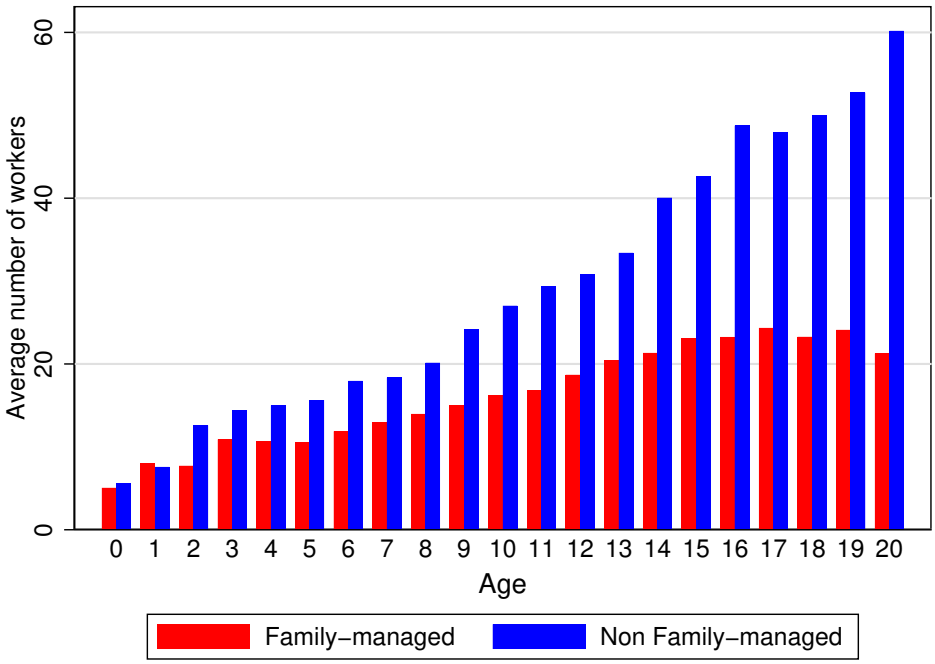


Figure B.5: Average Size of Firms Across Age in Terms of Employment by Type of Management (Only LLC)

Notes: Firms were classified as FM and EM according to the cut-off rule defined in (2.1) with $\bar{c} = 50\%$. The sample includes only Limited Liability Companies in every sector, except the financial one, for the period 2009-2016. The averages were computed by simply pooling data in each age category. Employment data come from self-reported information provided by the Bureau of Companies.

Robustness of the Results to Firm's Classification

In this section I reestimate the life-cycle profiles employing the specification (2.2), but changing the criteria used to classify FM firms as defined in (2.1) by employing a different cut-off. Figure B.8 show the results when the cut-off is changed to $\bar{c} = 25\%$. That is, now a firm is labeled as Family Managed if the principal manager is related with a family that controls more than 25% of the ownership of the firm. This is a more relaxed threshold and so will now include a larger group of FM firms, as shown in Table 2.3. When using that classification criteria, the gap at age 20 between both type of firms shrinks relative to the baseline estimates, and is of the order of 1.5. This is likely to reflect the inclusion of firms where ownership is not fully diversified, and so a manager still has some non-mayority control of the firm. However, the change is not substantial due to the overall little dispersion of ownership seen in the data.

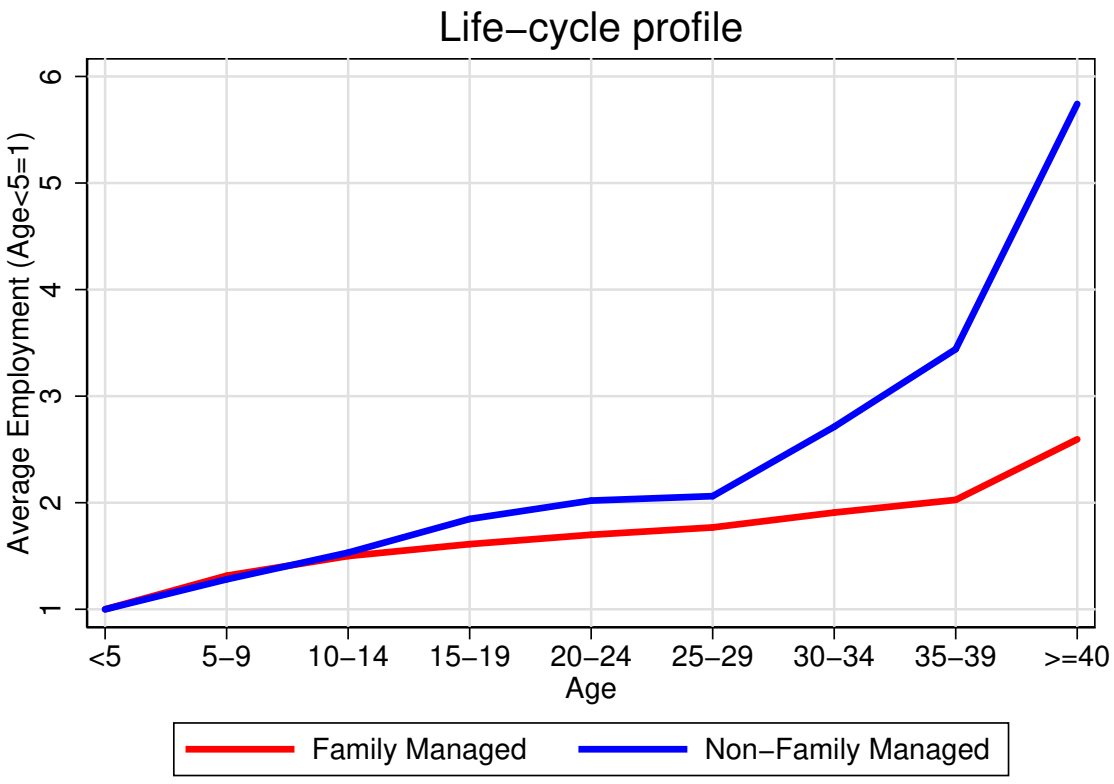


Figure B.6: Life-cycle Profile of Family Managed and Non-Family Managed Firms (Manufacture)

Notes: Estimated by (B.1). Firms were classified according to the cut-off rule defined in (2.1) with $\bar{c} = 50\%$. The sample includes the universe of all formal firms in Ecuador, both Limited Liability Companies and Sole Proprietors, except the financial sector, for the period 2007-2016. Age equal one is the base category. Employment is the number of workers formally registered with the firm, that is, those that reported being employed by the firm in the Form 107.

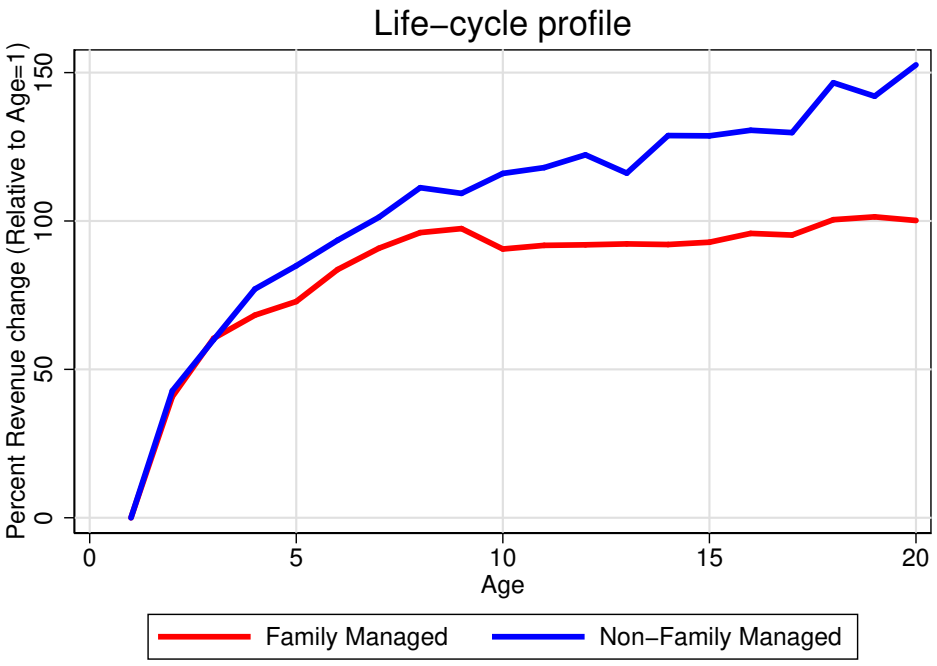


Figure B.7: Life-cycle Profile of Family Managed and Non-Family Managed Firms (Only LLC)

Notes: Estimated by (2.2). Firms were classified according to the cut-off rule defined in (2.1) with $\bar{c} = 50\%$. The sample includes the universe of all Limited Liability Companies, except the financial sector, for the period 2007-2016. Age equal one is the base category. Revenue is Real Revenue, deflated using the CPI.

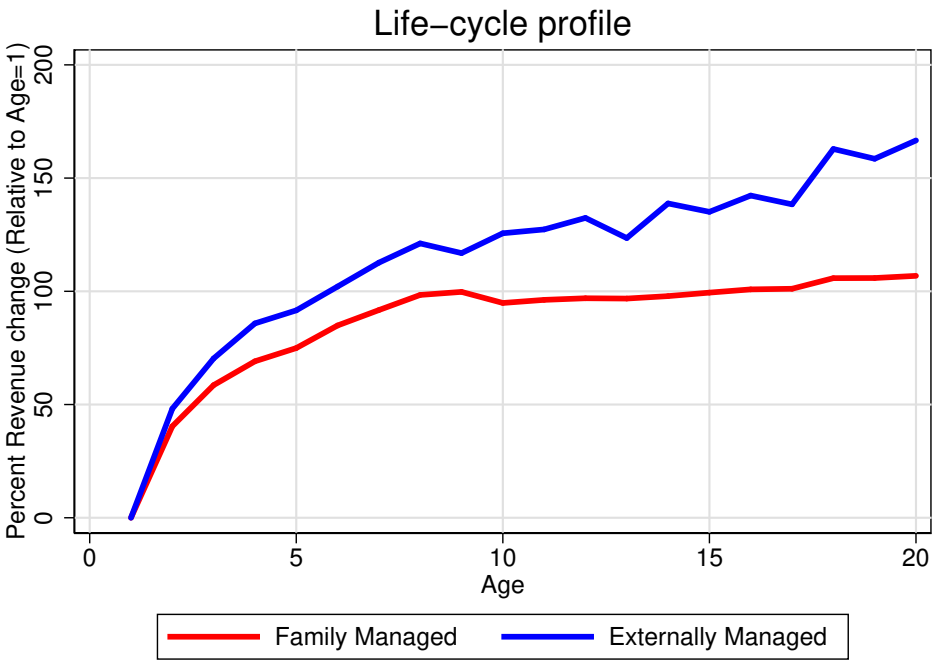


Figure B.8: Life-cycle Profile of Family Managed and Non-Family Managed Firms (25% Cut-off)

Notes: Estimated by (2.2). Firms were classified according to the cut-off rule defined in (2.1) with $\bar{c} = 25\%$. The sample includes the universe of all formal firms in Ecuador, both Limited Liability Companies and Sole Proprietors, except the financial sector, for the period 2009-2016. Age equal one is the base category. Revenue is Real Revenue, deflated using the CPI.

Equilibrium

This section describes in detail the equilibrium of the model. First, introduce the notation of z^+ and z^- as the upper and lower levels of the productivity distribution. Now, using equations (2.10) and (2.15), the equilibrium in the managerial market is defined by,

$$\begin{aligned}\left(\frac{\theta-1}{\theta}\mu_h\right)^{\theta}\left(\frac{w_h}{w_n}\right)^{\theta-1}\frac{\theta}{\theta-1} &= \int_{z^*}^{z_+}\left(\frac{z}{\phi}(1-\alpha)^{\frac{1}{\gamma}}\left(\frac{\phi(1-\gamma)}{w_h}\right)^{\frac{1}{\gamma}}\left(\frac{\alpha}{w_n}\right)^{\frac{\alpha}{\gamma(1-\alpha)}}-\frac{\bar{h}}{\phi}\right)d\mu(z)\\ &= \frac{\bar{h}}{\phi}\int_{z^*}^{z_+}\left(\frac{1}{z^*}-\frac{1}{z}\right)zd\mu(z)\end{aligned} \tag{B.2}$$

where the last line uses the optimal cut-off (2.9). Equivalently, by using the equilibrium cut-off (2.9), from equations (2.11) and (2.16), the equilibrium in the labor market is given by

$$1-\left(\frac{\theta-1}{\theta}\mu_h\right)^{\theta}\left(\frac{w_h}{w_n}\right)^{\theta} = \left(\frac{\alpha}{w_n}\right)^{\frac{1}{1-\alpha}}\bar{h}^{1-\gamma}\left[\int_{z_-}^{z^*}z^{\gamma}d\mu(z)+\int_{z^*}^{z_+}{z^*}^{\gamma-1}zd\mu(z)\right] \tag{B.3}$$

Observe that in an equilibrium with entrants, the free entry condition is,

$$c_e = \int V(z,w_n,w_h)\nu(z)dz \tag{B.4}$$

Now, let Γ be the transition matrix for the productivity z, of size $n_z \times n_z$, with n_z being the number of productivity levels. The elements of this matrix are: $P_{i,j} = \Pr(Z_{t+1} = z_j \mid Z_t = z_i)$. Since z grows deterministically, $P_{i,j=i+1} = 1$ for $j < n_z$ and $P_{n_z,n_z} = 1$. In the stationary equilibrium, the distribution of firms is given by,

$$\begin{aligned}\mu(z) &= (1-\lambda)\Gamma'\mu(z) + M\nu(z)\\ &= \mu(z)\tilde{\Gamma} + M\nu(z)\end{aligned} \tag{B.5}$$

where $\tilde{\Gamma} = (1-\lambda)\Gamma'$. Therefore, the stationary distribution of firms can be written as,

$$\mu(z) = M\left(I-\tilde{\Gamma}\right)^{-1}\nu(z) \tag{B.6}$$

Then, by using (B.6) as the distribution, equations (B.2), (B.3) and (B.4) are a system of equations that solve the equilibrium mass of entrants, M, the managerial compensation rate, w_h, and the wage rate, w_n.

Computation of Equilibrium

The algorithm used to compute the equilibrium is quite standard. I first start with a guess of the values of w_n, w_h and M, and then solve the static problems of both type of firms, the optimal decision of what type of management they chose, as well as the dynamic program by iterating in the value function. I then verify that both labor and managerial markets clear, and that the free entry condition holds, and continue updating prices until that is the case.