Labor Scarcity, Land Tenure, and Historical Legacy: Evidence from Mexico

Emily A. Sellars* Jennifer Alix-Garcia†

Abstract

How does labor scarcity impact institutional development? We examine the long-term institutional impacts of Mexico's 16th-century demographic collapse, which reduced the indigenous population by between 70 and 90 percent. We show that the collapse facilitated land concentration by the elite and the rise of a landowner class that dominated Mexican political economy for centuries. Though land institutions were transformed after the Revolution of 1910, the 20th-century agrarian reform was concentrated in areas experiencing a steeper population decline between 1570 and 1650. Disease-impacted areas now have more land in common-property institutions (*ejidos*) with limited property rights. The collapse thus had a persistent impact on the political economy of the country, but the nature of this impact has changed radically over time. Our identification strategy exploits climatic shocks associated with an epidemic in the 1570s that roughly halved the indigenous population. We discuss the implications of our findings for the study of the institutional impacts of population scarcity and historical persistence and change.

Acknowledgements: We are grateful to Barbara Forrest for pointing out the existence of *cocolitzli*. For detailed comments on this manuscript, we thank Francisco Garfias, Scott Gehlbach, Marcus Kurtz, Horacio Larreguy, Laura Schechter, and Felicity Vabulas. We also thank seminar participants at the American Political Science Association Annual Meeting, the Bureau for Research and Economic Analysis of Development Conference, the Midwest International Economic Development Conference, the Triangle Research and Environmental Economics (TREE) seminar, Oregon State University, the Harris School Political Economy Workshop, and the University of Wisconsin Development Workshop and Comparative Politics Colloquium.

*The Harris School of Public Policy, University of Chicago, 1155 E. 60th St., Chicago, IL, USA 60637, email: esellars@uchicago.edu, phone: 315-382-3756

†Department of Agricultural and Applied Economics, University of Wisconsin, 429 Taylor Hall, Madison, WI, USA 53706, email: alixgarcia@wisc.edu, phone: 608-262-4499

1 Introduction

Historical events and institutions can have long-term political consequences, and these consequences evolve over time. In one of the most famous arguments on path dependence and institutional development, Acemoglu, Johnson and Robinson (2002) connect population density at the start of colonial rule to contemporary patterns of development. They argue that colonial powers established perverse institutions aimed at resource extraction in population-dense areas, while sparsely settled areas received virtuous institutions aimed at attracting settlement (p. 1265–6). The quality of colonial institutions, as demonstrated by this work and others, had long-reaching political and economic effects (e.g., Engerman and Sokoloff 1997; La Porta et al. 1997; Nunn 2009).

In this paper, we trace the impact of Mexico's 16th-century demographic collapse on land institutions through the present day. Mexico has been cited as a classic example of pre-colonial population density leading to dysfunctional colonial institutions and long-term underdevelopment (e.g., Engerman and Sokoloff 1997; Acemoglu, Johnson and Robinson 2002; Acemoglu and Robinson 2012). However, this generalization masks profound variation in population dynamics, institutions, and development. While central Mexico was densely populated at the time of the Conquest, by most accounts the population fell by between 70 and 90 percent in the first century of colonial rule due to disease and famine (Cook and Borah 1971; Hassig 1985; Cook 1998).[1] As a comparison, the Black Plague—central to many explanations of European political development—is thought to have reduced Europe's population by between 30 and 60 percent (Levine 2006; Voigtländer and Voth 2013*b*). The cataclysmic demographic collapse dramatically altered Mexico's economic and political trajectory. The *encomienda*, the extractive institution highlighted in many accounts of Mexican underdevelopment (e.g., Acemoglu and Robinson 2012), had essentially disappeared in central Mexico less than 50 years after the Conquest (Gibson 1964; Hassig 1985; Knight 2002). In its place, the Spanish adopted a series of institutions aimed at controlling land and labor in a context of severe population scarcity. According to numerous historians of Mexico, the labor scarcity caused by the demographic collapse—not merely labor abundance at the time of Conquest—led to the development of perverse political and economic institutions (e.g., Borah 1951; Gibson 1964; Knight

[1] For a summary of colonial population estimates, see Appendix Table A1.

2002).

We argue that the 16th-century demographic collapse had a significant and persistent impact on Mexican land tenure and political economy. Population scarcity caused by the collapse lowered the costs and increased the benefits of acquiring land from indigenous holders. This led to the development of large estates (*haciendas*), which controlled much of the countryside until the Revolution of 1910. After the Revolution, Mexican land tenure and rural political economy was transformed by the far-reaching agrarian reform. Though the hacienda system no longer exists, this reform was concentrated in areas where large estates had been most dominant prior to the Revolution and where populations had fallen most dramatically in the early colonial period. As a result, disease-impacted areas now have more land in *ejidos*, the common-property institution created through the reform that has been associated with limited property rights, political corruption, and economic underdevelopment. The legacy of the population collapse thus continues to impact development, though the nature of this impact has changed dramatically since 1910.

Using subnational data on the number of people paying tribute to the Spanish Crown in the 16th and 17th centuries, we show that more of the rural population lived on haciendas prior to the Revolution, more land redistribution took place over the 20th century, and more land area is in ejidos today in places where the population declined more precipitously between 1570 and 1650. Identifying the causal impact of the demographic collapse on land institutions is difficult, as factors like soil quality, the disease environment, and topography could influence population dynamics, agricultural productivity. To address this concern, we use an instrumental-variables empirical strategy based on a peculiar feature of a massive epidemic in the 1570s. This epidemic, which roughly halved the indigenous population, is believed to have been caused by a rodent-transmitted pathogen that emerged after several years of drought were followed by a period of above-average rainfall (Acuña Soto et al. 2002). We construct proxies for these climate conditions using precipitation data dating back to the early 1500s extracted from tree-ring chronologies (Cook and Krusic 2004). Using these proxies, we predict the fall in population between the first complete tribute count in 1570 and the nadir of the indigenous population in 1650. The climate data capture deviations from normal conditions for each point in space and are arguably orthogonal to geographic, political, and economic confounds.

Our findings illustrate that the legacy of the population collapse continues to influence the political economy of Mexico. Landholding patterns structure political relationships as well as economic ones (e.g., Herbst 2000; Baland and Robinson 2008, 2012; Boone 2014). The specific land institutions that we examine—the ejido and the hacienda—have independently been linked with poor political outcomes, such as clientelism, economic corruption, and electoral manipulation (Engerman and Sokoloff 1997; Sokoloff and Engerman 2000; Dell 2012; Larreguy 2013; Albertus et al. 2015). However, the mechanisms through which each institution is thought to impede development are different. Whereas the hacienda system is blamed for fostering the rise of a dominant landowner class that perpetuated severe economic and political inequality for centuries, the ejido, introduced in part to address this inequality, is thought to impede development through the incomplete property rights granted to the smallholder *ejidatarios* that now control much of the nation's land. By connecting both of these institutions with the much earlier population collapse, we demonstrate both that the roots of observed underdevelopment may go back farther than is sometimes argued and that it is difficult to assess the long-term impact of these institutions in isolation given their interdependence.

More broadly, our work contributes to the literatures on the political economy of population and institutional change and evolution. The institutional impacts of population scarcity have been used to explain the end of serfdom and state development in Europe, long-term underdevelopment in Africa, the emergence and decline of slavery, and the evolution of property-rights protection (e.g., Domar 1970; Epstein 2000; Herbst 2000; Boone 2003; Levine 2006). We call attention to a different mechanism through which labor scarcity can have long-term institutional impacts. Building on work by Conning (2007) and others, we argue that labor scarcity can both enable and encourage land concentration by reducing the cost of acquiring land and increasing the benefits of controlling local labor markets. Whereas in Europe depopulation due to the Plague increased the bargaining power of labor and set the continent on a trajectory of prosperity (e.g., Levine 2006; Voigtländer and Voth 2013*b*), landowners in Mexico used their political and economic power to force workers onto employment in large estates. Through tying the collapse with the establishment of the hacienda, our paper provides empirical support for theoretical models examining the interplay between labor scarcity, outside labor options, and coercive institutions (e.g., Domar 1970; Acemoglu and Wolitzky

2011). Furthermore, our findings demonstrate that even short-term labor scarcity may produce long-term institutional consequences.

An additional contribution of our work, however, is that we continue to trace the impact of the demographic collapse after the long-running political control of landowners had been subverted. It is perhaps unsurprising that a massive demographic shock during a formative period of institutional development would have long-reaching impacts when even small shocks at critical moments can have persistent consequences (Acharya and Lee 2015). What is perhaps more interesting is that, while the collapse continues to shape development, the character of its impact has evolved. While the Mexican Revolution arguably represents an example of rapid and "discontinuous" change in political institutions (Roland 2004), the land reform implemented during and after the Revolution targeted long-running agrarian grievances. Because of this, the reform was tied to the legacy of the hacienda. Land tenure in disease-impacted areas has therefore changed dramatically since 1910, but landholding patterns remain distinct from areas that suffered comparatively little during the collapse. Our work therefore supports Douglass North's contention that even an abrupt, revolutionary change like Mexico's agrarian reform is "seldom as discontinuous as it appears on the surface" (North 1990, p. 90).

Moreover, while it is often argued that the impact of historical events deteriorates over time (e.g., Banerjee and Iyer 2005; Nunn 2008), this study illustrates that this is not the only possible trajectory. As we discuss, early patterns of land concentration during the 16th-century collapse were amplified in during the 19th-century economic boom before being radically transformed by the agrarian reform. Many papers on historical legacy associate a long-ago historical event or institution with a contemporary outcome, but this article highlights why one should be cautious in assuming the continuity of political and institutional processes between two snapshots of history. This has important implications for studying colonial institutions and development (e.g., La Porta et al. 1997; Acemoglu, Johnson and Robinson 2001; Acemoglu and Robinson 2008; Mahoney 2010) and institutional persistence and change more generally (e.g., North 1990; Pierson 2000; Roland 2004; Thelen 2004; Mahoney and Thelen 2010).

Our paper proceeds as follows. In Section 2, we outline our argument linking population scarcity to the rise of large estates and describe our empirical setting. Section 3 presents and summarizes the main data sources on colonial population and pre-Revolution land tenure. We describe our identification strategy in Section 4. Section 5 presents our empirical results on the population collapse and the hacienda. In Section 6, we discuss the relationship between the collapse, the hacienda, and Mexico's 20th-century land reform and present results associating the collapse with 20th-century land redistribution and land tenure. We conclude in Section 7.

2 Population Scarcity and Institutional Development

When the Spanish arrived in Mexico in 1519, they encountered a large and advanced society of around 20 million people (Gibson 1964; Cook and Borah 1971; Hassig 1985). After the fall of the Triple Alliance (Aztec Empire), the Spanish usurped the Alliance's tribute system, enabling them to obtain food and other resources without acquiring land or taking part in agricultural production (Hassig 1985; Knight 2002). The institution of the *encomienda*, which awarded colonizers and indigenous nobles the right to tribute and labor from specific villages, has often been cited as an important contributor to long-term underdevelopment in Latin America (Engerman and Sokoloff 1997; Acemoglu and Robinson 2012).

By the middle of the sixteenth century, a combination of disease, drought, and famine had decimated Mexico's population by up to 90 percent. The size and significance of Mexico's demographic collapse has few parallels in history. Arguably the closest comparison is the depopulation of Europe during the Black Plague. The Plague led to massive institutional change in Europe. The drop in Europe's population is argued to have caused or contributed to the decline in feudalism, the rise of urbanization, and the creation of modern states (e.g., Epstein 2000; Levine 2006; Voigtländer and Voth 2013*a*,*b*). Existing empirical evidence suggests that Europe's demographic losses raised per capita incomes, labor-market bargaining power, and consumption among those who survived. In fact, Voigtländer and Voth (2013*a*,*b*) argue the cycle of warfare and depopulation started by the Plague set Europe on a track of long-term growth and explains the sizable gap in European per capita incomes relative to the rest of the world by 1500. In their model, the depopulation of Europe set it on a course of long-run growth and relative prosperity not merely through the short-term demographic shock, but also through the cycle of warfare and state building brought on by the

Plague.

If the institutional impacts of the Plague in Europe were beneficial for economic growth and the bargaining power of labor, the same could not be said of 16th-century Mexico. Unlike Europe, Mexico's depopulation took place in a context of colonization and conquest. The encomienda was designed to extract as much tribute and labor from the population as possible. The collapse placed heavy burdens both on the colonizers, who could no longer live off of the surplus production of indigenous villages, and indigenous workers, who were increasingly unable to economically support both themselves and the Spanish elite (Hassig 1985, p. 185). From their position of power, the colonizers were forced to transition to other economic strategies. As acquiring wealth through tribute extraction became untenable, colonists transitioned to harnessing land and labor for agricultural production and commercialization (Borah 1951; Gibson 1964; Cook and Borah 1971; Knight 2002). The Mexican hacienda—the large agricultural estate that came to dominate local land and labor markets—grew out of this transition. The hacienda has been called "the chief single determinant of [colonial Mexico's] pattern of development" (Knight 2002, p. 74).

The demographic collapse shaped the rise of the hacienda in several ways. Most directly, the depopulation of the countryside facilitated land acquisition by colonial elites. The colonial government offered incentives for land acquisition in population-scarce areas, establishing a system through which landowners could request grants of nearby "unoccupied" land, much of which had been depopulated during major epidemics (Gibson 1964; Knight 2002). The program of forced indigenous resettlement in disease-impacted areas freed up further land for Spanish acquisition (Gibson 1964; Hassig 1985). By contrast, in areas that retained their population, indigenous holders were able to legally contest colonists' land claims, leading to lengthy disputes that could be very costly for the elite (Gibson 1964; Knight 2002, p. 120–1). Beginning in 1570, a colonial land-titling program provided legal protections for those who had acquired land, making it difficult for indigenous holders to contest usurpations when populations rebounded in the 17th century (McBride 1923; Gibson 1964; Knight 2002).

The labor-market impacts of the collapse also contributed to the rise of the hacienda. As the population plummeted, institutions introduced to ration labor between government and private use, such as the labor draft or *repartimiento*, were gradually abandoned in central Mexico. Landowners sought to bypass these institutions and hire labor directly, though labor scarcity made this difficult (Gibson 1964; Hassig 1985; Knight 2002). Debt peonage, through which peasant families took up residence in large estates and became tied to them through debt contracts, became synonymous with the hacienda (e.g., Knight 1986*a*; Florescano 1984; Knight 1986*b*, p. 85–9). The manipulation of local land markets offered a mechanism for landowners to reduce the cost of labor recruitment and retention. By encroaching on nearby indigenous landholdings, landowners were able to drive down the outside economic options of rural laborers, forcing them into employment on large estates (Florescano 1984, 1987; Conning 2007). In much of the country, the annexation of indigenous lands constituted "one of the *hacendados'* principal strategies for acquiring workers" (Florescano 1987, p.267). This amplified the relationship between population scarcity and land concentration by incentivizing land encroachment in areas around large estates. Through these interdependent forces, a small number of landowners increasingly dominated local land and labor markets. The political and economic dominance of a small number of landowners had important long-term impacts on the political economy of the countryside (Chevalier 1963; Conning 2007; Engerman and Sokoloff 1997; Knight 2002).

2.1 Persistence and change in the hacienda

The hacienda remained the central institution of the countryside after Mexico gained independence from Spain in 1821. Ongoing instability prevented the new national government from implementing major reforms until the second half of the 19th century. Under the rule of Porfirio Díaz (1884–1911), a series of policy changes were adopted that exacerbated landholding inequality and amplified the political and economic power of landowners. The development of railroads and economic policies favoring export agriculture raised land values and the profitability of agricultural production (McBride 1923; Coatsworth 1974). Limits placed on communal landholding by the Lerdo Law of 1856 provided a legal basis for the seizures of millions of hectares of indigenous and public lands by large estates to take advantage of these new economic opportunities (McBride 1923; González Navarro 1957; Meyer 1973; Coatsworth 1974; Knight 1986*b*). The result was a marked increase in land alienation and concentration.

By the eve of the Revolution in 1910, up to four-fifths of rural communities and half of the rural population lived on haciendas and were "subject to the political and social control of the estate" (Knight 1986*b*, p. 96). Agrarian concerns were a central motivation for the Mexican Revolution (1910–21).[2] During this time, the government began a land reform program that would last until 1991 and transform landholding patterns across the nation (Knight 1986*b*; Sanderson 1984; Markiewicz 1993). The common-property institutions created through this reform (*ejidos*) now hold more than half of the country's land area and have themselves been associated with adverse economic and political outcomes (see Section 6).

As we discuss below, while the hacienda no longer dominates rural life as it did prior to the Revolution, it shaped the pattern of land redistribution over the twentieth century and thus landholding arrangements today. Because more land was redistributed in places where the hacienda had been strongest prior to the Revolution, the areas that experienced a steeper decline in population between 1570 and 1650 have more land in ejidos today. The 16th-century demographic collapse thus continues to influence the political economy of Mexico, though the character of this impact has changed. In the remainder of the paper, we systematically examine the relationship between the collapse, landholding patterns prior to the Revolution, and land tenure today.

3 Data

This section describes the sources and characteristics of our data on colonial population levels and changes, the economic and social importance of haciendas, and the covariates that may be correlated with population dynamics, agricultural productivity, and institutions.

[2]Both the size of the hacienda sector in the early twentieth century and the role of land conflict in the Revolution remain the subject of considerable debate. Scholars in the second half of the 20th century argued that work byMcBride (1923); Tannenbaum (1929), and others overstated the role of the hacienda in the economy (e.g., Meyer 1973, 1986, 1991). A more recent literature has offered a modest corrective to the arguments of the revisionists (Knight 1986*b*, 1991, 2006). Even if census figures overstate the size of the hacienda sector as some argue, revisionist Jean Meyer agreed that "the hacienda...directly or indirectly leaves its mark on all of rural life" (1986, p. 483).

3.1 Colonial Population Density

Our measures of colonial population density were digitized from Gerhard's exhaustive *A Guide to the Historical Geography of New Spain* and *The North Frontier of New Spain* (Gerhard 1993*a*,*b*). The first book contains data on the colonial governorship of New Spain, which contained most of central Mexico, and the second includes data on the governorships of Nueva Galicia, Nueva Vizcaya, and Sinaloa and Sonora in the north. The remainder of modern Mexico was either conquered by the Spanish later than the time period that we examine (the northeast and Baja California) or does not have the necessary tree-ring data for us to implement our empirical strategy (the southeast). The study area is shown in Figure 1.

Figure 1: New Spain and its North Frontier, 1786

Population statistics in Gerhard are reported at the level of 1786 political divisions, outlined with the thicker gray lines in Figure 1. We examine a snapshot of population in each of these districts in two periods: around 1570, the first year in which Gerhard's data are relatively complete, and around 1650, the sample year closest to the nadir of Mexico's indigenous population.[3] Of the 191 districts in the Gerhard books, data are available in both sample periods for 177 units. Gerhard constructs the population data using a large number of historical sources, relying primarily on the *relaciones geográficas*, a set of questionnaires distributed to local officials by King Phillip II beginning in the late 1560s. The questionnaires were completed by Spanish administrators in consultation with local indigenous elders and nobles (Cook and Borah 1971). In addition to detailed information about local geography, history, and customs, these documents record the number of individuals paying tribute to the Spanish Crown in different years. These tribute counts are the primary measures of population available for this time period.

In general, each married male Indian and his immediate family constituted one tributary, and unmarried men and women over 15 and widows were considered "half-tributaries" (Gerhard 1993*a*). Because indigenous nobles and clergy were exempt from paying tribute, tribute counts can not be translated directly into indigenous population. Cook and Borah (1960*a*, 1971) estimate that the ratio of tributaries to total indigenous Mexicans was about 2.8 in the late 16th century, but this ratio is thought to have increased over time. Following Cook and Borah (1960*a*), we convert data recorded as "individuals" or "families" in Gerhard to tribute counts using an average of 5 members per family and 2.8 individuals per tributary.[4] After converting all units to be comparable, we calculate the municipal-level tribute population density in 1570 and 1650 by overlaying Gerhard's 1786 divisions with 20th-century municipal boundaries[5] and calculating the space-weighted average tributary density. The principal outcomes used in this paper are recorded at the level of the municipality.

There is some debate regarding the reliability of Spanish tribute data (Borah 1976). Though

[3]Hassig (1985) and Knight (2002) both estimate that the low point in population was during the 1630s.

[4]In the data appendix (Appendix A.2), we discuss these conversions and present estimates excluding the regions where the data had to be converted.

[5]Given that municipal boundaries change over time, we use the smallest unit at which all of our data are available. This is generally the municipality as of 1900 except when the 1900 municipality was later absorbed into a larger unit. For the state of Oaxaca, we aggregate the data to the level of the district in 1900.

there is some evidence that colonists had incentives to systematically over- or undercount population at different times, Cook and Borah (1971) conclude that the Crown was able to limit fraud during the period under study. After 1530, counts were conducted by a single Spanish official for each town and then subjected to a review through which interested parties, colonists, and indigenous nobles could challenge the results. The system of tributary counts and collection had been harmonized by the study period, and exemptions from tribute are not thought to vary systematically across administrative units (Cook and Borah 1960*a*). The late 16th-century numbers are thought to be particularly reliable, as they marked the beginning of Phillip II's efforts to extract resources from his dominion in order to pull Spain out of bankruptcy.[6] Cook and Borah (1960*b*) call the population evaluations of the late 16th century "remarkable for their abundance and reliability" (p. 1). The size of the collapse as measured our data is similar in magnitude to the estimates of Cook and Borah (1979) based on a comparison of village-level data for a subset of towns.

Nevertheless, measurement error remains a concern. Under classic errors-in-variables assumptions, econometric results using these data as an explanatory variable would be attenuated towards zero. However, it is possible that error varied systematically across space. For example, given that the nobility and clergy were exempt from tribute demands, it is plausible that our data would underestimate population in high-density areas where such groups were concentrated. Given that we are hypothesizing an inverse relationship between population density and hacienda formation, which took place mainly in rural areas, this potential source of bias would lead to an underestimate of the causal impact. Our best defense against the threat to identification posed by measurement error, however, is our use of an instrumental-variables estimation strategy to address these and other endogeneity concerns. We discuss this strategy in Section 4.

3.2 The Mexican Hacienda

The primary source of data on the hacienda is the Mexican Historical Archive of Localities (AHL), maintained by Mexico's National Institute of Statistics and Geography (INEGI). The AHL records the population of every locality in Mexico in each census or national count beginning in 1900 and continuing at approximately 10-year intervals. The AHL also records the political category

[6] Knight (2002) quipped that Phillip II "liked paperwork almost as much as he disliked rebellion" (p. 57).

of the locality in every census year as determined by the government. Using this information, we extract the number of localities that are described as "haciendas" (haciendas), "ranchos" (ranches), or "fincas" (farms/estates) and the total population living in these localities in 1900 and 1910 for each municipality.[7]

We adopt a relatively expansive measure of the "hacienda" for several reasons. Mexican scholars have highlighted regional and political differences in whether large estates were classified as "haciendas" or "ranchos" and have noted the lack of distinction between the various estate classifications (e.g., González Navarro 1969, p. 120; Gómez-Serrano 2000, p. 464). In our data, a significant number of localities are classified as haciendas in one census year (either 1900 or 1910) and ranchos or fincas in the other. To investigate the accuracy of our measure, we consulted Southworth (1910), which provides a list of large haciendas immediately prior to the Revolution. It is not possible to comprehensively match these data with that of the AHL as Southworth does not always include the geographic location of estates.[8] In general, the AHL listings appear to be more comprehensive than those of Southworth, though this relationship varies in space.[9] For the state of Aguascalientes in central Mexico, we were able to match a relatively large number of AHL villages to directory listings by name. Among properties listed in both sources, those classified by the AHL as haciendas had an average size of 15,555 hectares, those classified as ranchos 13,000 hectares, those that changed between rancho and hacienda in the two censuses 12,400 hectares, and those in other categories 8,667 hectares. These are all much larger than the typical size of smallholdings in Mexico.[10] Our estimates of the proportion of rural population living on estates are smaller than the state-level estimates in McBride (1923) and Tannenbaum (1929), though they are highly correlated

[7]Municipalities in Mexico are akin to counties in the United States and contain many villages and towns.

[8]For some states, the directory lists the address of the owner, often in Mexico City or in regional capitals, rather than the location of the estate.

[9]Some of the more problematic areas are outside of our study area in Chiapas and the Yucatán peninsula.

[10]According to the government's Ejidal Census of 1991, the average size of holdings was about 2.1 hectares/farmer in the ejidal sector.

with both of these.[11]

3.3 Control variables

We include a series of geographic control variables to account for factors that might explain both the population decline and the expansion of large estates. The key controls are average elevation, the area of the municipality, distance to Mexico City, distance to railroad line in the late 19th century, an indicator of whether a municipality is in the malarial zone (has a minimum elevation of less than 1000 m), and a measure of soil quality. Average elevation and slope are calculated at the administrative level using a digital-elevation model (DEM) with a resolution of 30 meters. Nineteenth-century railroad access is approximated by drawing straight lines between the major cities linked by these railroads (see Appendix Appendix A.5). Distances are extracted using GIS software. Our measure of soil quality is the space-weighted potential low-input maize productivity from the Food and Agriculture Organization's Global Agro-Ecological Zones (IIASA/FAO 2012). The FAO potential productivity measure is calculated using information about local climate, soil types, slope, and rainfall, and comes in a layer with 5-arcminute resolution (about 10 km cells). Maize is the main staple crop in Mexico, and potential maize productivity is highly correlated with that of many other grain and vegetable crops in the GAEZ data.

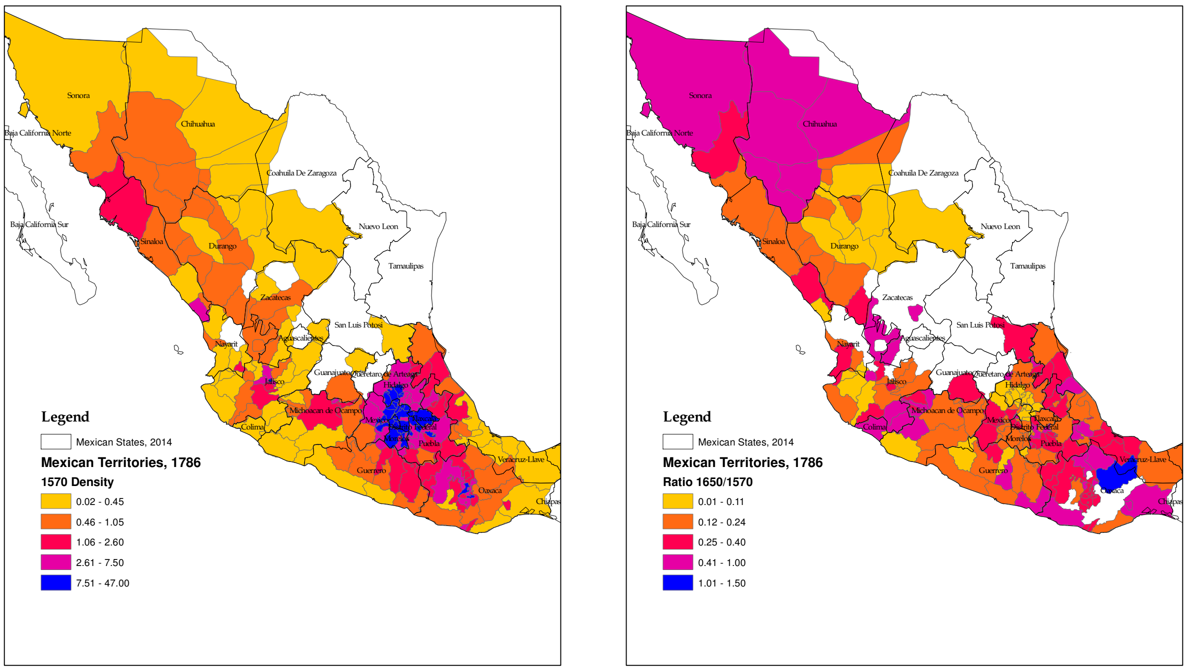
3.4 Summary statistics

The means and standard deviations of all included variables are available in the Appendix (Table A1). Figure 2 shows the density of the tributary population in 1570 and the change between 1570 and 1650 as expressed by the ratio of the latter to the former. Unsurprisingly, the area with the highest population density is the region around Mexico City, the heart of the Triple Alliance. Populations are also relatively dense in the Mixtecan region of Oaxaca in the south and in the Tarascan region of Michoacán in the center-west, both of which were major population centers prior to the Conquest (Cook and Borah 1971). The right side of Figure 2 illustrates the precipitous

[11] The correlation coefficients between our measures of the number of large estates and the proportion of the population living on them and those of Tannenbaum (1929) are .958 and .959 respectively. The correlation between our estimate of the number of hacienda villages and that of McBride (1923) is .613.

population decline between 1570 and 1650. Only two districts, both in Oaxaca, grew between these years. Tributary densities declined by over 70% on average among the 177 districts for which we have data.

Figure 2: Tributary Density, 1570 and 1650



(a) 1570

(b) Density 1650/Density 1570

In Figure 3 we show a simple log-linear fit of the percentage of each municipality's rural population living on haciendas in 1900 against the ratio of 1650 to 1570 tributary population. This model includes fixed effects for the four governorships in our sample (Nueva España, Nueva Galicia, Nueva Vizcaya, and Sinaloa and Sonora) to account for potential regional differences in population dynamics and the accuracy of how tribute data were collected. The correlation is negative, which is consistent with our argument that hacienda expansion was greater in areas where the population decline was more severe.[12]

[12]This correlation persists if we first aggregate the district-level colonial population data to the level of the state and match this information with McBride's (1923) measure of the percentage of the rural population living in hacienda communities by state in 1910. McBride's measure of hacienda importance has been cited

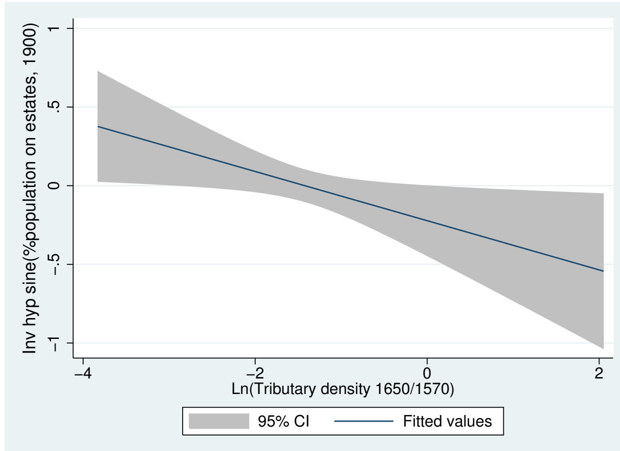


Figure 3: % 1900 population in haciendas against population collapse

Table 1 shows summary statistics for municipalities that experienced higher and lower than median population loss between 1570 and 1650 as measured by the ratio of 1650 to 1570 tributary density, as well as the normalized differences of the means of these variables. The size of the population collapse is considerable in both groups. The first column shows statistics for municipalities that experienced a more pronounced population decline (i.e., the ratio of 1650 to 1570 density was smaller than the median). Municipalities experiencing a larger decline in population had significantly higher population density in 1570, whereas the density in both groups was approximately equal in 1650. Although there is no apparent difference in the percentage of the rural population in estate villages in 1900 across the columns, the areas with greater collapse have a higher number estates per area on average. These municipalities also tend to be at slightly higher elevation, are less like to be in "malarial" areas, exhibit higher potential maize productivity, and are closer to 19th-century railroad lines relative to the less-affected municipalities. All of these geographic variables could potentially influence incentives for land acquisition and concentration. The majority of the municipalities in our sample were located in the governorship of Nueva España in central Mexico, where contemporary municipalities are small and numerous. The collapse appears to have been more severe on average in the center of the country (Nueva España and Nueva Galicia) relative to the north (Nueva Vizcaya and Sinaloa y Sonora).

in several influential studies of the Mexican Revolution (e.g., Tannenbaum 1929; Knight 1986*b*), though it has been criticized by later scholars as overstating the size of the hacienda sector (e.g., Meyer 1991).

Table 1: Summary statistics by tributary ratio, 1650/1570

	(1) High collapse	(2) Low collapse	(3) Norm diff
Municipal tributary density, 1570	3.917	1.747	-0.379
Municipal tributary density, 1650	0.621	0.623	0.001
1650/1570 municipal tributary density	0.158	0.480	0.653
% population in estates, 1900	25.543	25.603	0.002
> median estates/km^2	0.561	0.482	-0.113
Mean municipal elevation	1721.742	1457.056	-0.250
Minimum elevation < 1000 m	0.398	0.518	0.172
Mean municipal maize productivity	0.846	0.705	-0.154
Ln(Municipal area, 1900)	5.892	6.130	0.127
Ln(Population density, 1900)	2.762	2.412	-0.183
% population in communities > 5000	0.055	0.064	0.034
Ln(Km to Mexico City)	5.149	5.660	0.366
Nueva España	0.795	0.720	-0.125
Nueva Galicia	0.113	0.101	-0.029
Nueva Vizcaya	0.072	0.085	0.034
Sinaloa y Sonora	0.019	0.095	0.232
Observations	513	517	1030

An additional empirical challenge is that 1570 and 1650 colonial population density and growth may have been influenced by pre-existing institutions or by different patterns of Spanish rule during the first 50 years of colonization. For example, if areas that received more coercive labor institutions in the early colonial period experienced greater population decline prior to 1570, this might induce a negative correlation between early population decline and later strength of the hacienda. We propose an empirical strategy to address these concerns in the next section.

4 Identification strategy

We adopt an instrumental-variables estimating strategy to identify the causal impact of the collapse on land tenure. Specifically, we construct an instrument for the decline in tributary population between 1570 and 1650 using information on rain and drought during the 1570s. As discussed in Section 2, drought and disease were major causes of the indigenous population collapse in central Mexico. Three severe epidemics in particular—those of 1519–21, 1545–8, and 1576–80—had catastrophic effects on the native population and on colonial institutions (Gibson 1964; Hassig 1985;

Knight 1986*b*). While there is wide agreement that the first of these epidemics was caused by smallpox, the pathogen responsible for the second two remained unknown until recently. Recent scholarship has uncovered evidence that these later epidemics, recorded by their Nahuatl name "*cocoliztli*," were caused by an indigenous hemorrhagic fever that was transmitted by rodents (Acuña Soto, Calderon Romero and Maguire 2000; Marr and Kiracofe 2000; Acuña Soto et al. 2002). Using drought reconstructions from northern Mexico, Acuña Soto et al. (2002) show that the 1545 and 1576 cocoliztli epidemics emerged in years of above average rainfall following prolonged periods of drought, much like other rodent-transmitted diseases such as hantavirus. Following these authors we construct climate measures reflecting drought, rainfall abundance, and the dependence between the two to predict the 1570–1650 population decline as described below.

We extract subnational data on rainfall and drought using information from the *North American Drought Atlas* (Cook and Krusic 2004). The Atlas reconstructs estimates of historical drought conditions using evidence recovered from a network of tree-ring chronologies covering much of North America. Using the information from the chronologies, Cook and Krusic (2004) estimate the annual summer Palmer Drought Severity Index (PDSI) in each 2.5 degree by 2.5 degree grid point using a principal-components framework. The PDSI is a normalized measure of soil moisture that captures deviations from typical conditions at a given location. Abnormally dry conditions are recorded using negative numbers, and positive numbers represent abnormally wet conditions. We create a grid (2.5 km cells) of distance-weighted PDSI estimates between points in the Atlas. We overlay the 1900 municipal boundaries on this grid and extract average PDSI for each district and municipality in each year of the 1570s. While the Atlas covers most of Mexico, it does not have information on the far southeast and Yucatán peninsula due to the absence of usable tree rings in that area. To check whether our interpolated PDSI measures capture meaningful deviations in rainfall, we compared contemporary data from the Atlas with rainfall measures from CONAGUA, Mexico's National Water Commission. CONAGUA calculates a rainfall index similar to PDSI for hundreds of stations within the country, 287 of which overlap with our study area. We find that the correlation between our measure and that of CONAGUA is 0.36 between 1979 and 1990, and the correlation coefficient between the first differences of the measures is 0.48. This suggests that these data do pick up variation in weather, and are especially responsive to temporal changes.

The excluded instruments used in our analysis are the sum of the two lowest consecutive PDSI values between 1570 and 1575 (more negative numbers indicate more severe, prolonged drought), the maximum PDSI between 1576 and 1580 (which captures excess rainfall), and the ratio of the former to the latter. This parameterization is based upon the observation, discussed above, that cocolitzli was more likely to emerge when drought was followed by heavy rainfall. Using this approach, we estimate the impact of the population decline between 1570 and 1650 on measures of the importance of the hacienda sector in 1900. Specifically, we estimate the following system of equations:

$$H_{ma} = \alpha_0 + \alpha_1 T_{ma} + \alpha_{\mathbf{2}} \cdot \mathbf{G_m} + \epsilon_{ma} \tag{1}$$

$$T_{ma} = \beta_0 + \beta_{\mathbf{1}} \cdot \mathbf{Z_{ma}} + \beta_{\mathbf{2}} \cdot \mathbf{G_m} + \mu_{ma} \tag{2}$$

The dependent variable (H_{ma}) is measured in two ways: the inverse hyperbolic sine of the percent rural population living in hacienda communities in 1900[13] and an indicator variable for whether the municipality had an above-median number of estates per square kilometer that year. The key explanatory variable, the population decline in municipality m within 1786 administrative region a (T_{ma}), is measured using the ratio of 1650 to 1570 density. $\mathbf{Z_{ma}}$ represents the vector of instruments. We also control for a series of geographic features related to population and agriculture ($\mathbf{G_m}$). These include mean elevation, an indicator variable for whether any part of the municipality has an elevation lower than 1000 meters (i.e., is in the malarial zone), mean maize productivity, natural log of rural population in 1900, natural log of the municipal area, percentage of the population living in large cities in 1900, distance to Mexico City, and the distance to the railroad in 1900. We also include indicator variables for three of the four governorships included in our sample. We conduct estimations both with and without controlling for 1570 population density. To control for the fact that consistently high variation in rainfall may independently affect both tributary density and agricultural productivity, we include the standard deviation of PDSI over all the years of the

[13]The inverse hyperbolic sine transformation can be interpreted similarly to a log transformation. It is preferable to a log transformation given that it can handle values of zero and has a number of other useful properties (e.g., Burbage, Magee and Robb 1988). Results are robust to using the alternative transformation $ln(x + 1)$ (upon request).

data, though excluding this variable does not change any of the results. All standard errors are clustered at the level of the 1786 region.

5 Empirical results on the hacienda

Table 2 presents the first-stage and reduced-form relationships between the climate variables and the decline in population between 1570 and 1650 or measures of hacienda presence in 1900. Columns (1) and (2) show the effects of the climate instruments on the population collapse. Decreases in the dependent variable indicate a greater decline. As expected based on Acuña Soto et al.'s (2002) analysis, the size of the collapse increases in the severity of drought in the early 1570s as well as in the measure of excess rainfall in the late 1570s. The estimation reported in column (2) includes the full set of covariates. Columns (3) and (4) show qualitatively similar results including a control for the level of tribute density in 1570 with and without the additional controls. Columns (5) and (6) present reduced-form estimations of the instruments on the measures of hacienda importance. The signs of all of the coefficients are the opposite of what they are in columns (1–4), which is consistent with our argument that landed estates became more important in areas with epidemic-induced population loss.

Table 3 reports the instrumental-variables estimates of the impact of the 1570–1650 population collapse on hacienda importance in 1900. These estimations use only the area of the sample for which we have both colonial population density data and reconstructed drought data.[14] The first four columns report OLS and instrumental variables estimates of the impact of the collapse on the inverse-hyperbolic-sine-transformed rural population living on estates using the ratio of 1650 to 1570 population with and without baseline 1570 tributary density as a control. In columns (2) and (4), we instrument for the decline in population using the climate variables described in the previous section. The second four columns repeat the analysis using an indicator variable for whether the municipality had greater than median number of estate villages per square kilometer in 1900 as the dependent variable, with IV estimates in columns (6) and (8). We use the same

[14] We lose 43 municipalities from the original tribute density data as there is no climate data for the far southeast of the study area.

Table 2: First-stage regressions: Climate and the 1570-1650 Population Collapse

	Dependent variable 1650/1570 density				% pop estates	> median estates/km^2
	(1)	(2)	(3)	(4)	(5)	(6)
Drought 1570	0.413***	0.408***	0.285***	0.412***	-0.288***	-0.026
	(0.082)	(0.081)	(0.064)	(0.083)	(0.088)	(0.029)
Maximum PDSI, 1576-1580	-0.591***	-0.583***	-0.201*	-0.570***	0.459**	0.281***
	(0.149)	(0.150)	(0.111)	(0.144)	(0.198)	(0.065)
Maximum rainfall/Drought	-1.019**	-1.050**	-0.276	-1.054**	2.342***	1.310***
	(0.454)	(0.448)	(0.429)	(0.431)	(0.616)	(0.199)
Ln(Tributary density, 1570)			-0.162***	-0.224***	-0.013	-0.045*
			(0.046)	(0.051)	(0.110)	(0.025)
Other covariates	no	yes	no	yes	yes	yes
Observations	1030	1030	1030	1030	1030	1030
Adjusted R^2	0.234	0.287	0.257	0.336	0.398	0.309

Unit of observation is the municipality. Standard errors are in parentheses and are clustered at the 1786 administrative level. * p< 0.10, ** p<0.05, *** p < 0.01. These are partial results. Other covariates include mean elevation, elevation < 1000 meters, mean GAEZ maize productivity, 1/km to nearest railroad line, ln(municipal area), standard deviation of PDSI, % population in large cities, ln(km to Mexico City), ln(rural population in 1900), and governorship dummies.

geographic control variables as above.

As would be predicted by our argument, the point estimates of the coefficient on the explanatory variable of interest are negative and significant. Their magnitudes are large; a one standard deviation decrease in the 1650 to 1570 population ratio causes a .30 standard deviation increase in the percentage of the population living on estates in 1900. This is equivalent to about a seven percentage point increase in the outcome variable. The same decrease in the tributary ratio increases the likelihood of having a higher density of estates per area in 1900 by 14 percentage points. In appendix Appendix A.4, we complete a counterfactual exercise to further explore the substantive impact implied by our point estimates. In addition, we provide evidence that the empirical link between the population collapse and the hacienda is most pronounced in areas adjacent to rail lines in 1905, as would be predicted based on historical arguments linking rail development and land seizures by large estates in the Porfirian era (Section 2.1). Interestingly, the 1570 population density variable in the IV regressions of number of hacienda villages is negative and significant. One explanation for this is that land concentration was more extensive where population density was lower prior to 1570 due to the high cost of acquiring land in high-density areas.

Turning attention to the reliability of the IV estimation strategy, the F-statistics of our excluded instruments are 11.3, with a partial R squared in the first stage of .156. We conclude that the instruments are relevant. However, the IV estimates are considerably larger in magnitude than the OLS coefficients. This is potentially concerning given the possibility of weak instruments. We argue that the differences in magnitude can be plausibly explained by a combination of measurement error and omitted-variables bias in the OLS estimates. For these results to be explained by omitted-variables bias, important omitted variables in the OLS estimation would have to influence population density and the establishment of the hacienda in the same direction (i.e., either have a positive effect on both population and landed estates or a negative one). An example of such an omitted variable might be agricultural productivity, which might be positively correlated with pre-colonial population density and subsequent incentives for land concentration. One potential correlate of productivity has a positive and significant coefficient in both first-stage and reduced-form regressions: the standard deviation of PDSI across sample years. The point estimate of the coefficient on this variable in the first stage is .45 and in the second stage is approximately 1. Ma-

Table 3: Population Decline and Hacienda Importance

	Dependent variable							
	% pop on estates				> median estates/km^2			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1650 density/1570 density	-0.207**	-0.747**	-0.236***	-0.833***	-0.049	-0.193**	-0.068**	-0.217**
	(0.085)	(0.296)	(0.085)	(0.284)	(0.030)	(0.086)	(0.032)	(0.090)
Ln(1570 tributary density)			-0.113	-0.238*			-0.074***	-0.105***
			(0.129)	(0.143)			(0.026)	(0.033)
Other covariates	yes	yes	yes	yes	yes	yes	yes	yes
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Observations	1029	1029	1029	1029	1029	1029	1029	1029
Adjusted R^2	0.337	0.297	0.338	0.292	0.209	0.174	0.218	0.183

Unit of observation is the municipality. Standard errors are in parentheses and are clustered at the 1786 administrative level. * p< 0.10, ** p<0.05, *** p < 0.01. These are partial results. Other covariates include elevation < 1000 meters, mean GAEZ maize productivity, 1/km to nearest railroad line, mean elevation, ln(municipal area), standard deviation of PDSI, % population in large cities, ln(km to Mexico City), ln(rural population in 1900), and governorship dummies.

nipulating the standard equation describing omitted variables bias, the product of these two gives us an estimate of the bias that might occur were this variable excluded from the regression. The difference between the OLS and IV estimates in columns (1) and (2) is .54, which is very close to .45.

A second potential explanation for the difference in magnitudes is attenuation bias resulting from measurement error. The OLS coefficient is about one third of the size of the IV estimate. By way of comparison, Card (2001) finds that attenuation bias due to measurement error in the literature on the returns to education may exceed 30%.[15] In a simple univariate regression model, where a variable x^* is measured with error v: $x = x^* + v$, it can be shown that the noise to signal ratio required to generate a difference of this magnitude would be: $s = \frac{\sigma_v^2}{\sigma_{x^*}^2} = (1/.7 - 1)^{-1} = 2.4$. In other words, the variance of the population decline measurements would need to be 2.4 times greater than the true variance of the population decline.

In the online appendix (Appendix A.3), we discuss a variety of robustness checks using outcomes from different censuses, alternative measures of hacienda significance, different aggregations of data, and different subsets of the data intended to shed light on potential data measurement challenges in the tributary data. We also discuss a falsification test and alternative estimations of the standard errors. We find our results robust to all of these tests.

6 Haciendas, Land Reform, and the Ejido

The previous section established a relationship between Mexico's 16th-century population collapse and the strength of the hacienda sector prior to the Revolution of 1910. This finding is important given well known arguments linking persistent underdevelopment in Latin America with colonial landholding inequality and the institutional impacts of pre-colonial population density. Beyond these effects, however, we argue that one less recognized long-ranging impact of the hacienda (and thus the population collapse) was in how it structured the 20th-century agrarian reform in Mexico and the contemporary land tenure and rural political organization of the country.

The fight over control of the land was one of the principal catalysts of the Revolution (e.g.,

[15] For a more extensive discussion, see Hausman (2001).

Knight 1986*a*). Partially in response to ongoing agrarian agitation, the federal government began a massive land reform program 1915, which lasted until 1991 (Sanderson 1984; Markiewicz 1993). This program radically altered the political and economic landscape of the nation (Knight 1991). Through the land reform program, peasant villages could jointly petition the government to expropriate and redistribute land from nearby large estates. By 1940, millions of hectares of land had been transferred through the program. The hacienda, which had dominated Mexican rural political economy for centuries, was virtually "killed off" by the middle of the twentieth century (Sanderson 1984; Knight 1991, p. 96). Today, over 50% of the country's surface area is held by *ejidos*, the land institution created through the land reform. Around half of Mexico's rural population lives in an ejidal community, and ejidos act as important conduits between the Mexican government and rural citizens (de Janvry, Gordillo and Sadoulet 1997; Fox 2007; Larreguy 2013).

The structure of the land reform program and the ejido have been heavily criticized. Redistribution was carried out collectively at the level of the community, and beneficiaries (*ejidatarios*) were given limited property rights over their individual plots. Until the end of the 20th century, land usage rights could not be transferred or sold, and they could be lost if the land went unused. A significant portion of ejidal land today is managed as common property. Many studies have described how the peculiar structure of the reform harmed development by inhibiting access to credit and depressing agricultural investment (e.g., Yates 1981; de Janvry, Gordillo and Sadoulet 1997; Dell 2012). The limited property rights granted by the reform are also thought to have enabled pernicious political outcomes by keeping smallholders dependent on government resources and protections. For example, there is a good deal of empirical evidence linking the ejido with clientelism, voter suppression, cronyism, and vote buying (e.g., Cornelius 1990; Larreguy 2013; Albertus et al. 2015). Reforms implemented post-1991 have arguably mitigated these effects by providing greater property-rights protections (e.g., de Janvry, Gonzalez-Navarro and Saudolet 2014; Castañeda Dower and Pfutze 2015). However, there has been little full privatization of ejidal land, and the ejido remains a crucial political and economic institution in the countryside (Gordillo 2014).

The link between haciendas and the land reform was direct. Only land in large[16] properties

[16]During the early years of the reform, the cutoff was around 500 hectares, with some adjustments for soil

was eligible for expropriation and redistribution, and this concentrated reform in the areas surrounding large plantations (Sanderson 1984; Markiewicz 1993). In addition, communities with proof of alienation of lands by large estates were able to apply to have these lands restored directly (Tannenbaum 1929; Sanderson 1984). Finally, though resident workers on haciendas were ineligible to receive lands during the early years of the program, these workers became an important constituency demanding reform beginning in the mid-1930s when eligibility for ejidal grants was expanded under pressure from the citizenry (Sanderson 1984, p. 45–6). Therefore, even though the land institutions that grew directly out of the 16th-century population collapse faded from importance, disease-impacted areas continued on a different development trajectory through the present day as a result of these forces.

To investigate the evidence for this argument, we use the same measures of colonial population and hacienda strength as described in Section 3. Our measures of agrarian reform come from several sources. The first of these is Sanderson's (1984) Agrarian Reform Database. This dataset contains information on every definitive presidential resolution on land reform from 1916 to 1976.[17] We examine the four major categories of resolutions on land redistribution: outright grants of land to communities, restitutions of lost village lands, expansions of ejidal grants, and grants distributing land for new agricultural villages. We aggregate these resolution-level data to the municipal level, recording the number of positive land resolutions, the number of reform beneficiaries, and the amount of land redistributed through the program in each municipality over this time period.[18]

In addition, we examine two measures of contemporary landholdings in Mexico. The first of these is the total contemporary land area held in social property[19] in a municipality, as measured using GIS calculations from a layer measured after 1992. The social property sector in Mexico is made up of two different institutions: *ejidos*, the property institutions created through the land

and crop types. This shifted slightly over time. See Sanderson (1984) for a discussion.

[17]The presidential resolution was the final stage of the reform process and finalized the land grants.

[18]While the reform continued until 1991, the vast majority of land redistribution took place prior to 1976. The time period of most intense redistribution was the 1930s.

[19]Mexican law recognizes three types of property: public, private, and social.

reform, and *comunidades*, communal lands that were recognized by the government but were not established by government action. While there are similarities, ejidos and comunidades are managed somewhat differently, and their historical antecedents are very different.[20] Whereas ejidos were created through government expropriation and redistribution of land, comunidades typically entered the ejidal sector through "sector additions," which recognized pre-existing communal landholding. To construct our second measure of contemporary land tenure, we subtract from the first Sanderson's measure of the land area that entered the social-property sector through sector additions, as would be true of comunidades. Though redistributive land transfers account for most social property in Mexico (80% of the total land in the GIS layer), sector additions account for a significant portion of land in some areas, particularly in the southern states of Oaxaca, Guerrero, and Chiapas.

Several confounding factors could contaminate our estimation of the relationship between the hacienda system prior to the Revolution and the subsequent land reform, such as underlying differences in agricultural productivity, climate, or even political culture. To address this concern, we adopt a similar instrumental-variables estimation strategy as in Section 4. Building on the empirical findings described earlier, we instrument for the strength of the hacienda system in 1900 using our sixteenth-century climate shock instruments (the sum of the two lowest consecutive PDSI values between 1570 and 1575, the maximum PDSI between 1576 and 1580, and the ratio of the former to the latter). The first stage F-statistic for excluded instruments is 19 in the regressions using the percentage population measure of hacienda importance and 32 for regressions using an indicator of whether a municipality had above-median estates per area in 1900. These estimations also include all of the geographic correlates used in our earlier analysis. Standard errors are again clustered at the 1786-district level. The results of these estimations are reported in Table 4.[21]

The correlations between the Sanderson land reform measures and those of hacienda importance are all positive and statistically distinguishable from zero. The point estimates suggest a close

[20]For a summary of the legal differences between ejidos and comunidades, see Ruiz Massieu (1990).

[21]Results using OLS estimations are similar with somewhat smaller point estimates and standard errors. Alternative IV specifications instrumenting for the decline in colonial population directly show a negative relationship, as would be expected based on Section 5 (available upon request).

Table 4: Agrarian Reform and Haciendas – instrumental variable regressions

	Dependent variable									
	Resolutions		Beneficiaries		Land distributed					
					Sanderson		GIS area		GIS area - title	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
> median estates/km^2		1.560***		2.033***		1.987**		-0.535		0.500
		(0.291)		(0.599)		(0.805)		(0.386)		(0.363)
% population estates	0.562***		0.724***		0.734***		-0.063		0.232**	
	(0.100)		(0.197)		(0.240)		(0.144)		(0.108)	
Other covariates	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	1029	1029	1029	1029	1029	1029	1029	1029	1029	1029
Adjusted R^2	0.361	0.349	0.278	0.230	0.378	0.340	0.495	0.492	0.514	0.490

Unit of observation is the municipality. All dependent variables are inverse hyperbolic sine transformed. Standard errors are in parentheses and are clustered at the 1786 administrative level. * p< 0.10, ** p<0.05, *** p < 0.01. These are partial results. Other covariates include elevation < 1000 meters, mean GAEZ maize productivity, 1/km to nearest railroad line, mean elevation, ln(municipal area), standard deviation of PDSI, % population in large cities, ln(km to Mexico City), ln(population in 1900), & governorship dummy variables. Excluded instruments are as above.

relationship between the agrarian redistribution and earlier hacienda concentration. The beta coefficient on the percentage of the rural population living on large estates in 1900 (column (3)) is .53, indicating that an increase of one standard deviation in the percentage of the population living on estates in 1900 results in an increase of about half of a standard deviation in the number of beneficiaries of redistributed land. The beta coefficients on this variable in columns (1) and (5) are .75 and .45, respectively. In the regressions using an indicator for whether a municipality had above-median large estates per square kilometer in 1900, the beta coefficients are .59, .43, and .35.

The last four columns show results using the different measures of the current distribution of social property in Mexico. The first of these is the land area per municipality in the social property sector. Here the correlations with measures of hacienda strength are negative; areas with more haciendas in 1900 had less total land in the social-property sector in the 1990s. However, once we subtract off land that entered the social-property sector through sector additions (resolutions that recognized de facto communal holdings and did not involve government expropriation and redistribution), the sign of the correlations reverses. The relationship is statistically significant in the estimations measuring hacienda strength as the percentage of the rural population on large estates.

The different relationships implied by the two measures of contemporary landholdings make sense given differences in the origin and management of lands in the social property sector. Sector additions recognized pre-existing communal landholding arrangements that often pre-dated the reform. Given that villages would have been less likely to maintain control of communal lands in areas of high hacienda presence, we should expect to see fewer sector additions and more redistributive grants in these areas. Consistent with this argument and our earlier findings, the number of sector additions is positively correlated with the ratio of 1650 to 1570 population density (i.e., more sector additions were granted in areas where the colonial population collapse was less severe). The relationship is reversed for redistributive grants.

These results demonstrate that areas that experienced a more precipitous population decline continue to follow a different development trajectory in Mexico. Though the hacienda no longer exists, there was more land redistributed through the land reform program in areas where the hacienda was strongest prior to the Revolution, and more land in those areas is held by ejidos today.

This finding calls attention to the complex relationship between the hacienda and the ejido. Though both institutions are thought to have harmed political and economic development, the mechanisms through which they are thought to have done so are distinct. The land reform radically changed the rural political landscape, and the ejido largely replaced the hacienda as the prototypical institution of the Mexican countryside. While Mexico's land reform undermined the political and economic dominance of the landowner class, scholars blame the incomplete property rights granted through this reform for continued underdevelopment. Our work suggests the possibility that longer-term effects are at work when we observe a correlation between the agrarian reform and later economic outcomes. Given that haciendas were the major source of land for ejidal grants, it is possible that informal institutions associated with the hacienda have persisted through the ejido (e.g., Knight 1991). Moreover, the dependence between the ejido and the hacienda highlights the difficulty of constructing the right counterfactual to assess the impact of land reform. Mexico's social-property institutions came into being in very different ways, and assessing whether they have helped or hindered development requires consideration of these historical antecedents. In observing a negative relationship between ejidal land tenure and political outcomes, it is difficult to discern whether this relationship can be explained by the institutional features of the ejido, the contentious process of land reform, the historical legacy of the hacienda, or some combination of these.

The results linking the ejido to the population collapse also have implications for understanding historical legacy more generally. Much of the empirical work on historical dependence examines one or two snapshots in history, often associating a long-ago shock or event with a contemporary outcome. If one had assessed the long-run impact of the collapse on land tenure in 1900, the conclusions would be very different. While disease-impacted areas were characterized by the political and economic dominance of large estates in 1900, today these areas are notable for the number of smallholders tied to the common-property ejido. Though both institutions have been argued to impede long-term development, they are thought to do so through different mechanisms and with different consequences. Historical legacies evolve, and even well established historical patterns started by significant events—like the centuries-long dominance of the hacienda following the population collapse—may be responsive to major change. Observing a correlation between a contemporary outcome and a long-ago event thus does not imply that this positive or negative empirical relationship has always existed, nor does it imply that it will persist into the future.

7 Conclusion

The evidence presented in this paper suggests that the population collapse of the early colonial period has shaped Mexican land institutions through the present day. A greater proportion of the rural population lived on large estates prior to the Revolution in areas that experienced a steeper population decline between 1570 and 1650. Though land tenure arrangements were overhauled after the Revolution, we show that the collapse continues to shape political economy through its impact on the 20th-century land reform, which determined Mexico's contemporary land institutions and distribution. Our identification strategy uses information on climate conditions that were related to a significant epidemic in the late sixteenth century that decimated the indigenous population of the country.

This paper makes several important contributions to the study of institutions, historical persistence, and the political economy of population. To the best of our knowledge, this represents the first attempt to systematically test the hypothesis of Borah (1951) and other scholars of Mexican history that the sixteenth-century population collapse shaped the emergence of the hacienda. We also build on the significant body of work the political and economic consequences of the hacienda by highlighting its relationship with the ejido, arguably the institution that most directly impacts rural political economy today. A great deal of work on Mexico has sought to assess the political and economic effects of this unique common-property institution. By showing that the historical antecedents of the modern ejido date back to the 16th century, we highlight the challenge of assessing which "perverse" institution—the hacienda or the ejido—should be blamed for long-term rural underdevelopment.

More generally, our work contributes to understanding the political impacts of population scarcity. While the relative density of Mexico at the time of the Conquest shaped early institutional development, colonial elites were forced to adapt to the challenges of profound labor scarcity soon after. The catastrophic population decline fueled land concentration and the development of inefficient institutions, which favored small segments of society and misallocated resources for generations. While in some contexts, such as in Europe following the Black Plague, sharp reductions in population led to an increase in the bargaining power of labor, the concentration of land and power by elites following the collapse in Mexico set the country on a trajectory of long-term inequality. While demographic shocks on the level of *cocolitzli* are rare, these findings can help us to understand the institutional aftermath of epidemics and civil conflict, which may also produce a concentrated and sharp decline in local populations.

Our findings also have important implications for understanding historical persistence and institutional change. Path-dependent arguments about institutions and development are sometimes criticized as being overly deterministic. As Banerjee and Duflo (2013) argue, the recognition that historical events can have long-run consequences does not imply that current conditions are unchangeable. Areas affected by disease in the early colonial period continue to follow a different development trajectory in Mexico, but the way in which these areas differ changed dramatically after the Revolution. Demography was not destiny in Mexico. While the modern ejido arguably owed its existence to the legacy of the hacienda, the land reform represented a complete economic and political reorganization of the countryside. This illustrates that even institutions with long historical antecedents may be responsive to policy change. Historical dependence does not imply stasis.

References

Acemoglu, Daron and Alexander Wolitzky. 2011. "The Economics of Labor Coercion." *Econometrica* 79(2):555–600.

Acemoglu, Daron and James A. Robinson. 2008. "Persistence of Power, Elites, and Institutions." *American Economic Review* 98(1):267–293.

Acemoglu, Daron and James A. Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty.* New York, NY: Crown Business.

Acemoglu, Daron, Simon Johnson and James A Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91(5):1369–1401.

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

Acharya, Avidit and Alexander Lee. 2015. "Path Dependence in European Development: Medieval Politics, Conflict, and State-Building." *Working Paper* .

Acuña Soto, Rodolfo, David W. Stahle, Malcolm K. Cleaveland and Matthew D. Therrell. 2002. "Megadrought and Megadeath in 16th Century Mexico." *Emerging Infectious Diseases* 8(4):360–362.

Acuña Soto, Rodolfo, Leticia Calderon Romero and James H. Maguire. 2000. "Large epidemics of hemorrhagic fevers in Mexico, 1545-1815." *American Journal of Tropical Medicine and Hygiene* 62:733–739.

Albertus, Michael, Alberto Diaz-Cayeros, Beatriz Magaloni and Barry Weingast. 2015. "Authoritarian Survival and Poverty Traps: Land reform in Mexico." *World Development* . Forthcoming.

Baland, Jean-Marie and James A. Robinson. 2008. "Land and Power: Theory and Evidence from Chile." *American Economic Review* 98(5):1737–1765.

Baland, Jean-Marie and James A. Robinson. 2012. "The Political Value of Land: Political Reform and Land Prices in Chile." *American Journal of Political Science* 56(3):601–619.

Banerjee, A. and E. Duflo. 2013. "Under the Thumb of History? Political Institutions and the Scope for Action." MIT Working Paper 14-02.

Banerjee, Abhijit and Lakshmi Iyer. 2005. "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India." *American Economic Review* 95(4):1190–1213.

Boone, Catherine. 2003. *Political Topographies of the African State*. Cambrudge, UK: Cambridge University Press.

Boone, Catherine. 2014. *Property and Political Order in Africa: Land Rights and the Structure of Politics*. Cambridge, UK: Cambridge University Press.

Borah, Woodrow. 1976. The historical demography of aboriginal and colonial America: an attempt at perspective. In *The Native Population of the Americas in 1492*, ed. William Denevan. University of Wisconsin Press.

Borah, Woodrow Wilson. 1951. *New Spain's Century of Depression*. Berkeley, CA: University of California Press.

Burbage, John B., Lonnie Magee and A. Leslie Robb. 1988. "Alternative Transformations to Handle Extreme Values of the Dependent Variable." *Journal of the American Statistical Association* 83(401):123–7.

Card, David. 2001. "Estimating the Returns to Schooling: Progress on Some Persistent Econometric Problems." *Econometrica* 69(5):1127–1160.

Castañeda Dower, Paul and Tobias Pfutze. 2015. "Vote Suppression and Insecure Property Rights." *Journal of Development Economics* 114:1–19.

Chevalier, François. 1963. *Land and Society in Colonial Mexico: The Great Hacienda*. Berkeley, CA: University of California Press.

Coatsworth, John. 1974. "Railroads, Landholding, and Agrarian Protest in the Early Porfiriato." *Hispanic American Historical Review* 54(1):48–71.

Conning, Jonathan. 2007. "On 'The Causes of Slavery or Serfdom' and the Roads to Agrarian Capitalism: Domar's Hypothesis Revisited.". Working Paper, Hunter College, Downloaded 7/2013.

Cook, E. R. and P. J. Krusic. 2004. *The North American Drought Atlas.* New York, NY: Lamont-Doherty Earth Observatory and National Science Foundation.

Cook, Noble David. 1998. *Born to Die: Disease and New World Conquest, 1492-1650.* New York, NY: Cambridge University Press.

Cook, Sherburne Friend and Woodrow Wilson Borah. 1960*a*. *The Indian population of central Mexico, 1531-1610.* Vol. 44 University of California Press.

Cook, Sherburne Friend and Woodrow Wilson Borah. 1960*b*. *The population of central Mexico in 1548: an analysis of the Suma de vistas de pueblos.* University of California Press, Berkeley.

Cook, Sherburne Friend and Woodrow Wilson Borah. 1971. *Essays in population history: Mexico and the Caribbean.* Vol. 1 Univ of California Press.

Cook, Sherburne Friend and Woodrow Wilson Borah. 1979. *Essays in population history: Mexico and California.* Vol. 3 Univ of California Press.

Cornelius, Wayne A. 1990. "Mexico and Salinas: PRI at the Crossroads." *Journal of Democracy* 1(3):61–70.

de Janvry, A, Gustavo Gordillo and Elisabeth Sadoulet. 1997. *Mexico's Second Agrarian Reform: household and community responses, 1990-1994.* Center for US-Mexican Studies.

de Janvry, Alain, Marco Gonzalez-Navarro and Elisabeth Saudolet. 2014. "Are land reforms granting complete property rights politically risky? Electoral outcomes of Mexico's certification program." *Journal of Development Economics* 110:216–225.

Dell, Melissa. 2012. "Path Dependence in Development: Evidence from the Mexican Revolution.". Working paper, Harvard University Department of Economics.

Domar, Harold. 1970. "The Causes of Slavery and Serfdom: A Hypothesis." *Journal of Economic History* 30(1):18–32.

Engerman, Stanley L and Kenneth L Sokoloff. 1997. Factor endowments: institutions, and differential paths of growth among new world economies: a view from economic historians of the United States. In *How Latin America Fell Behind: Essays from the Economic Histories of Brazil and Mexico, 1800-1914*, ed. Stephen Haber. Stanford, CA: Stanford University Press.

Epstein, S. R. 2000. *Freedom and Growth: The Rise of States and Markets in Europe, 1300–1750.* London, UK: LSE/Routledge.

Florescano, Enrique. 1984. The formation and economic structure of the hacienda in New Spain. In *Cambridge History of Latin America, Vol. II*, ed. Leslie Bethell. Cambridge University Press pp. 153–188.

Florescano, Enrique. 1987. The hacienda in New Spain. In *Colonial Spanish America*, ed. Leslie Bethell. Cambridge University Press pp. 250–285.

Fox, Jonathan. 2007. "Rural democratization and decentralization at the state/society interface: What counts as 'local' government in the Mexican countryside?" *Journal of Peasant Studies* 34(3-4):527–559.

Gerhard, Peter. 1993*a*. *A guide to the Historical Geography of New Spain: Revised Edition.* University of Oklahoma Press.

Gerhard, Peter. 1993*b*. *The North Frontier of New Spain.* Norman, OK: University of Oklahoma Press.

Gibson, Charles. 1964. *The Aztecs Under Spanish Rule: A History of the Indians of the Valley of Mexico, 1519-1810.* Palo Alto, CA: Stanford University Press.

Gómez-Serrano, Jesús. 2000. *Haciendas y ranchos de Aguascalientes: Estudio regional sobre la tenencia de la tierra y el desarrollo agrícola en el siglo XIX.* Aguascalientes, Aguascalientes: Universidad Autónoma de Aguascalientes.

González Navarro, Moisés. 1957. El Porfiriato. La Vida Social. In *Historia Moderna de México*, ed. Daniel Cosío Villegas. Vol. IV of *Historia Moderna de México* México, DF: Editorial Hermes.

González Navarro, Moisés. 1969. "La tenencia de la tierra en México." *Caravelle* 12:115–134.

Gordillo, Gustavo. 2014. "El Campo: Contra-Hechos." *La Jornada* August 16.

Hassig, Ross. 1985. *Trade, tribute, and transportation: The sixteenth-century political economy of the Valley of Mexico.* Norman, OK: University of Oklahoma Press.

Hausman, Jerry. 2001. "Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left." *Journal of Economic Perspectives* 15(4):57–67.

Herbst, Jeffrey. 2000. *States and Power in Africa: Comparative Lessons in Authority and Control.* Princeton, NJ: Princeton University Press.

IIASA/FAO. 2012. "Global Agro-ecological Zones (GAEZ v3.0).". IIASA, Laxenburg, Austria and FAO, Rome, Italy.

Knight, Alan. 1986*a*. "Mexican Peonage: What Was It and Why Was It?" *Journal of Latin American Studies* 18(1):41–74.

Knight, Alan. 1986*b*. *The Mexican Revolution.* Lincoln, NE: University of Nebraska Press.

Knight, Alan. 1991. "Land and Society in Revolutionary Mexico: The Destruction of the Great Haciendas." *Mexican Studies/Estudios Mexicanos* 7(1):73–104.

Knight, Alan. 2002. *Mexico: The Colonial Era.* New York, NY: Cambridge University Press.

Knight, Alan. 2006. "Patterns and Prescriptions in Mexican Historiography." *Bulletin of Latin American Research* 25(3):340–366.

La Porta, Rafael, Florencio Lopez-De-Silanes, Andrei Shleifer and Robert W. Vishny. 1997. "Legal Determinants of External Finance." *The Journal of Finance* 52(3):1131–1150.

Larreguy, Horacio. 2013. "Monitoring Political Brokers: Evidence from Clientelistic Networks in Mexico.". Manuscript, Harvard Kennedy School of Government.

Levine, David. 2006. Why and How Population Matters. In *The Oxford Handbook of Contextual Political Analysis*. New York, NY: Oxford University Press.

Mahoney, James. 2010. *Colonialism and Postcolonial Development: Spanish America in Comparative Perspective*. New York: Cambridge University Press.

Mahoney, James and Kathleen Thelen. 2010. A Theory of Gradual Institutional Change. In *Explaining Institutional Change: Ambiguity, Agency, and Power*. Cambridge University Press.

Markiewicz, Dana. 1993. *The Mexican Revolution and the Limits of Agrarian Reform, 1915-1946*. Boulder, CO: Lynne Rienner Publishers.

Marr, John S. and James B. Kiracofe. 2000. "Was the Huey Cocoliztli a Haemorrhagic Fever?" *Medical History* 44:341–362.

McBride, George McCutchen. 1923. *The Land Systems of Mexico*. New York: American Geographical Society.

Mexican Central Railway. 1900. *Facts and Figures about Mexico and Her Great Railroad, The Mexican Central*. Mexico, DF: The Mexican Central Railway Company, Ltd.

Meyer, Jean. 1973. *Problemas campesinos y revueltas agrarias (1821-1910)*. México, DF: Secretaría de Educación Pública.

Meyer, Jean. 1986. "Haciendas y ranchos, peones y campesinos en el porfiriato. Algunas falacias estadísticas." *Historia mexicana* 35(3):477–509.

Meyer, Jean. 1991. *La Revolución Mexicana, 1910-1940*. México, DF: Editorial Jus.

North, Douglass C. 1990. *Institutions, Institutional Change, and Economic Performance*. Cambridge University Press.

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

Nunn, Nathan. 2009. "The Importance of History for Economic Development." *Annual Review of Economics* 1:65–92.

Pierson, Paul. 2000. "Increasing Returns, Path Dependence, and the Study of Politics." *American Political Science Review* 94(2):251–267.

Roland, Gerard. 2004. "Understanding institutional change: fast-moving and slow-moving institutions." *Studies in Comparative International Development* 38(4):109–131.

Ruiz Massieu, Mario. 1990. *Derecho Agrario*. México, DF: Universidad Autónoma de México.

Sanderson, Susan R Walsh. 1984. *Land reform in Mexico, 1910-1980*. New York, NY: Academic Press.

Sokoloff, Kenneth L. and Stanley L. Engerman. 2000. "History Lessons: Institutions, Factor Endowments, and Paths of Development in the New World." *Journal of Economic Perspectives* 14(3):217–232.

Southworth, John R. 1910. *El Directorio Oficial de las Minas y Haciendas de México/The Official Directory of the Mines and Estates of Mexico*. México, DF: Southworth, FRGS.

Tannenbaum, Frank. 1929. *The Mexican Agrarian Revolution*. Washington, DC: The Brookings Institution.

Thelen, Kathleen. 2004. *How Institutions Evolve: The Political Economy of Skills in Britain, the United States, and Japan*. New York, NY: Cambridge University Press.

Voigtländer, Nico and Hans-Joachim Voth. 2013*a*. "Gifts of Mars: Warfare and Europe's Early Rise to Riches." *Journal of Economic Perspectives* 27(4):165–186.

Voigtländer, Nico and Hans-Joachim Voth. 2013*b*. "The Three Horsemen of Riches: Plague, War, and Urbanization in Early Modern Europe." *Review of Economic Studies* 80:774–811.

Yates, P. Lamartine. 1981. *Mexico's Agricultural Dilemma*. Tucson, AZ: University of Arizona Press.

Appendix A Additional tables and data appendix (for online publication only)

Appendix A.1 Additional tables and figures

Table A1: Various estimates of Mexico's population collapse

Author	Population (millions) 1519	1595	Percent decline (%)
"Mexico"			
Rosenblat (1935)	4.5	3.5	22
Aguirre-Beltran (1946)	4.5	2.0	56
Zambardino (1980)	5-10	1.1-1.7	64-89
Mendizabal (1939)	8.2	2.4	71
Cook and Simpson (1948)	10.5	2.1-3.0	71-80
Cook and Borah (1960)	18-30	1.4	78-95
Central Mexican Region			
Sanders (1976)	2.6-3.1	0.4	85-87
Valley of Mexico			
Whitmore (1992)	1.3-2.7	0.1-0.4	69-96
Gibson (1964)	1.5	0.2	87
128 towns			
Kubler (1942)	0.2	0.1	55

source: Robert McCaa
http://www.hist.umn.edu/ rmccaa/noncuant/democatt.htm

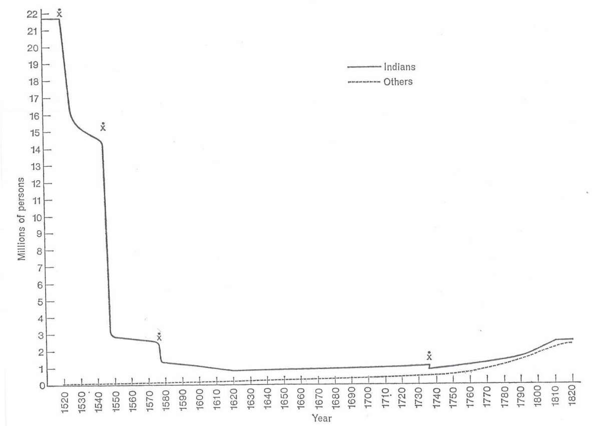


Figure A1: Decline of native population of Mexico

Table A2: Summary Statistics

	mean	sd	max	min
Municipal tributary density, 1570	2.828	4.191	46.155	0.037
Municipal tributary density, 1650	0.622	1.374	35.386	0.005
1650/1570 municipal tributary density	0.320	0.384	7.735	0.022
% population in estates, 1900	25.573	25.855	100.000	0.000
> median estates/km^2	0.521	0.500	1.000	0.000
Mean municipal elevation	1588.885	760.093	3225.858	12.107
Minimum elevation < 1000 m	0.458	0.498	1.000	0.000
Mean municipal maize productivity	0.775	0.649	3.257	0.000
Minimum km to railroad line	74938.280	97441.348	538305.625	0.000
Ln(Municipal area, 1900)	6.011	1.332	10.297	2.617
Ln(Population density, 1900)	2.586	1.359	6.515	-3.960
% population in communities > 5000	0.060	0.182	1.000	0.000
Ln(Km to Mexico City)	5.405	1.020	7.479	1.233
Observations	1030			

Appendix A.2 Data appendix

In order to generate as complete a sample of tributary data from Gerhard (1993*a*,*b*) as possible, we needed to convert raw population data in several forms into comparable units. In cases where Gerhard reports population in terms of "individuals" rather than tributaries, we converted the figure to tribute counts by dividing the number of individuals by 2.8. This figure is based on the discussion in Cook and Borah (1960*a*, 1971) relating tribute counts to population data. Also following these sources, where the population was reported as "families" or "vecinos" (neighbors), we use these figures as a proxy for the tribute count directly. This second conversion was only necessary in eight cases. We examine whether our main results change if we use only the divisions for which population data were recorded as tribute counts in both 1570 and 1650. Table A3 reports the results from the baseline instrumental-variables regressions using this subset of data. We lose a significant number of observations but the results remain qualitatively similar and strongly significant.

Table A3: IV-Estimation using only data initially reported as tributaries

	Dependent variable			
	% pop on estates		> median estates/km^2	
	(1)	(2)	(3)	(4)
1650 density/1570 density	-2.236*	-2.656**	-1.124**	-1.173***
	(1.171)	(1.055)	(0.476)	(0.423)
Ln(1570 tributary density)		-0.687**		-0.346***
		(0.327)		(0.111)
Other covariates	yes	yes	yes	yes
	IV	IV	IV	IV
Observations	886	886	886	886
Adjusted R^2	-0.090	-0.212	-1.202	-1.103

Unit of observation is the municipality. Standard errors are in parentheses and are clustered at the 1786 administrative level. * p< 0.10, ** p<0.05, *** p < 0.01. These are partial results. Other covariates include elevation < 1000 meters, mean GAEZ maize productivity, 1/km to nearest railroad line, mean elevation, ln(municipal area), standard deviation of PDSI, % population in large cities, ln(km to Mexico City), governorship dummies, ln(rural population in 1900), and governorship dummies.

We also needed to group together tribute data that were recorded in different years. While Gerhard's data are recorded in several waves, the specific year at which data are available for different colonial divisions varies by political division. The most problematic issues would arise if data were reported during the epidemic (1576-1580) as that could lead us to over- or understate the population collapse in some areas. None of our data were recorded in this window. Some data do come from before the period before the first cocoliztli epidemic in 1545. In these cases, we may be overestimating the population collapse by including the fall in population due to the first epidemic. Table A4 shows results from estimations that are limited to municipalities where at least part of the 1570 population data comes from data recorded before 1545. Our results are unaffected by this modification.

Table A4: IV-Estimation using only data with tributary observations after 1545

	Dependent variable			
	% pop on estates		> median estates/km^2	
	(1)	(2)	(3)	(4)
1650 density/1570 density	-0.782***	-0.870***	-0.205**	-0.229**
	(0.302)	(0.290)	(0.089)	(0.093)
Ln(1570 tributary density)		-0.248*		-0.108***
		(0.144)		(0.034)
Other covariates	yes	yes	yes	yes
	IV	IV	IV	IV
Observations	1023	1023	1023	1023
Adjusted R^2	0.289	0.284	0.172	0.181

Unit of observation is the municipality. Standard errors are in parentheses and are clustered at the 1786 administrative level. * p< 0.10, ** p<0.05, *** p < 0.01. These are partial results. Other covariates include elevation < 1000 meters, mean GAEZ maize productivity, 1/km to nearest railroad line, mean elevation, ln(municipal area), standard deviation of PDSI, % population in large cities, ln(km to Mexico City), governorship dummies, ln(rural population in 1900), and governorship dummies.

One might also think that merging data from time periods that were significantly removed from the dates of interest could drastically affect our outcomes. To test whether this is the case, we limit the data window to ten years around each of 1570 and 1650 (i.e., restrict the data to counts in 1560-1580 and 1640-1660). Table A5 shows that the results are unaffected, though the sample size is reduced considerably.

Table A5: IV-Estimation using only data within 10 year window of 1570 and 1650

	Dependent variable			
	% pop on estates		> median estates/km^2	
	(1)	(2)	(3)	(4)
1650 density/1570 density	-0.838***	-0.929***	-0.161**	-0.174**
	(0.234)	(0.225)	(0.075)	(0.079)
Ln(1570 tributary density)		-0.381***		-0.109***
		(0.129)		(0.035)
Other covariates	yes	yes	yes	yes
	IV	IV	IV	IV
Observations	671	671	671	671
Adjusted R^2	0.232	0.240	0.281	0.300

Unit of observation is the municipality. Standard errors are in parentheses and are clustered at the 1786 administrative level. * p< 0.10, ** p<0.05, *** p < 0.01. These are partial results. Other covariates include elevation < 1000 meters, mean GAEZ maize productivity, 1/km to nearest railroad line, mean elevation, ln(municipal area), standard deviation of PDSI, % population in large cities, ln(km to Mexico City), governorship dummies, ln(rural population in 1900), and governorship dummies.

Appendix A.3 Robustness checks

This appendix discusses the robustness of these findings and evaluate the assumptions underlying our empirical strategy. Our reported results are robust to using outcomes measured from the 1910 census in place of 1900, and to using alternative measures of hacienda significance, including the number of hacienda villages in a given municipality. Our findings are also unchanged when using the log of 1650 tributary density as the main explanatory variable and controlling for baseline 1570 density. As an additional check, we also estimate models excluding observations from the Federal District and Mexico State, which is the area of highest population density in our sample and potentially not representative of the country as a whole. We also aggregate the hacienda outcomes (measured at the level of the municipality) up to the 1786 division level and re-estimate our system of equations at that larger unit of analysis. The results are similar in magnitude and direction, though the instrument loses some power as a result of the large decrease in the number of observations. As a final check, we estimate a series of spatial lag models to account for potential spatial autocorrelation between units. The direction and significance of the coefficients remain unchanged, but the magnitude of the effects decreases somewhat in these specifications.[22] In the subsection below (Appendix A.2), we replicate the analysis using different subsets of our data to address potential challenges with data measurement and analysis.

One addition concern might be that weather shocks from the 1570s could have been correlated with important omitted geographic or economic characteristics (for example, if climate conditions are highly persistent across time and space). Given that PDSI measures are normalized to account for climatically appropriate conditions in each grid point, the potential for geographic confounds should be minimized. In Table A6, we repeat the analysis shown in Table 2 using drought and rain

[22] All these results are available upon request.

variables constructed from data in a later period (1670–80 rather than 1570–80).[23] The coefficients on the climate variables in the placebo regressions take the opposite sign as that of the results in Table 2. The coefficient on the drought variable is statistically significant, but the coefficients on the other two covariates are not, with the exception of rainfall in column (6).

We conclude that there is possibly concerning spatial/temporal correlation between the total drought and rain intensity variables, but not the ratio term. As a robustness check, we therefore repeat the estimation shown in Table 3 using only the ratio as an excluded instrument and including the separate 1570 weather terms as regressors. These results are shown in Table A7. The coefficients of interest are larger in magnitude than in the original estimates. The first stage F-statistic is just over 5, and the partial R-squared of the excluded instrument has decreased to .0229. The single instrument is clearly weak, although the Anderson-Rubin test of joint significance of endogenous regressor in the main equation clearly rejects the null of being equal to zero, with a p-value of 0.0001, and in all cases the Anderson-Rubin confidence intervals are complete and do not contain zero.

[23]We lose one observation due to the drought variable being equal to zero for one of the municipalities, thus making the ratio of rain to drought undefined.

Table A6: Falsification tests

	Dependent variable 1650/1570 density				% pop estates	> median estates/km^2
	(1)	(2)	(3)	(4)	(5)	(6)
Drought 1670	-0.209***	-0.220***	-0.140**	-0.178**	0.150	0.039
	(0.079)	(0.080)	(0.062)	(0.073)	(0.120)	(0.034)
Maximum PDSI, 1676-1680	-0.081	0.019	-0.087	-0.031	0.703***	0.127***
	(0.072)	(0.068)	(0.066)	(0.062)	(0.088)	(0.037)
Maximum rainfall/Drought	0.003	0.003	0.002	0.002	0.012	-0.001
	(0.004)	(0.003)	(0.003)	(0.003)	(0.007)	(0.002)
Other covariates	no	yes	no	yes	yes	yes
Observations	1029	1029	1029	1029	1029	1029
Adjusted R^2	0.121	0.171	0.202	0.244	0.385	0.229

Unit of observation is the municipality. Standard errors are in parentheses and are clustered at the 1786 administrative level. * $p< 0.10$, ** $p<0.05$, *** $p < 0.01$. These are partial results. Other covariates include elevation < 1000 meters, mean GAEZ maize productivity, 1/km to nearest railroad line, mean elevation, ln(municipal area), standard deviation of PDSI, % population in large cities, ln(km to Mexico City), ln(rural population in 1900), and governorship dummies.

Table A7: IV-Estimation using only drought-rainfall interaction as excluded instrument

	Dependent variable			
	% pop on estates		> median estates/km^2	
	(1)	(2)	(3)	(4)
1650 density/1570 density	-2.231**	-2.144**	-1.249**	-1.194**
	(0.997)	(0.896)	(0.557)	(0.505)
	[-6.14, -1.01]	[-5.66, -1.04]	[-3.43, -.65]	[-3.18, -.61]
Ln(1570 tributary density)		-0.802*		-0.497***
		(0.446)		(0.190)
Other covariates	yes	yes	yes	yes
	IV	IV	IV	IV
Observations	1030	1030	1030	1030

Unit of observation is the municipality. Standard errors are in parentheses and are clustered at the 1786 administrative level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. These are partial results. Other covariates include elevation < 1000 meters, mean GAEZ maize productivity, 1/km to nearest railroad line, mean elevation, ln(municipal area), standard deviation of PDSI, % population in large cities, ln(km to Mexico City), governorship dummies, and ln(rural population in 1900). Anderson-Rubin confidence intervals in [].

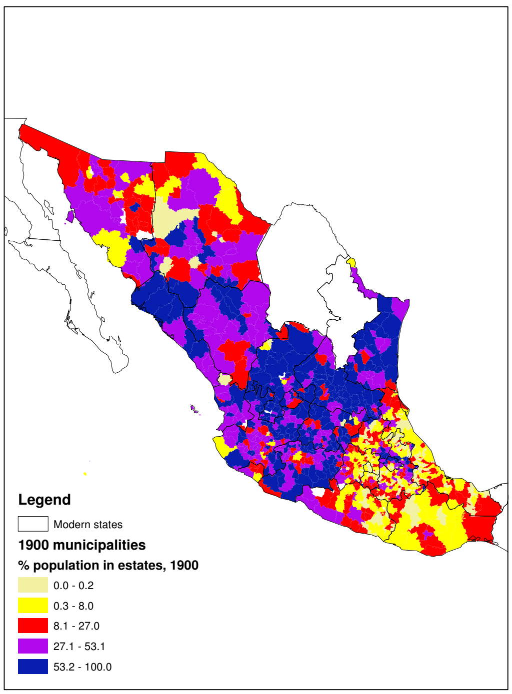
Appendix A.4 Counterfactual analysis of the causal impact

What would the distribution of hacienda population have looked like in the absence of the demographic collapse? This appendix proposes one potential response. Exploiting the instrumental-variables estimates, we explore what the variation in hacienda distribution would have looked like had there been no collapse of the population. To do this, we subtract off the predicted marginal effect of the population change in each municipality from the actual 1900 outcome using the estimates in column (2) of Table 3[24]. We then transform these estimates back into levels. The distribution of hacienda population changes substantially; the median actual percent of 1900 population in haciendas in the municipalities used in our estimation 16.7. When we remove the impact of the population collapse, this decreases to 5.6 percent. The change is even larger at the 75th percentile, which in 1900 levels is 44 percent, while our counterfactual estimate is 16.3 percent.

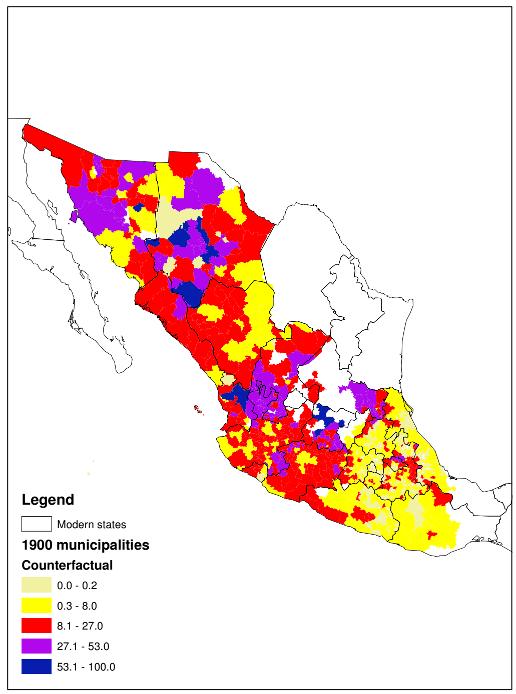
The maps in figure A2 shows a map of the actual distribution of hacienda population in percent of total population per municipality in the left panel. On the right is the hacienda population adjusted to take away the impact of the population changes in the data. The colors are coded according to the actual 1900 hacienda population quintiles. The differences are particularly striking across the northwest and just northeast of Mexico City. There is much less variance in the counterfactual estimates of the 1900 hacienda population; the reduction in this variance is about 47 percent. This suggestions that the demographic collapse alone had a significant effect in generating the observed distribution of individuals living in hacienda villages in 1900.

[24]This exercise follows the logic of the counterfactual experiment in Bohn and Pugatch 2013. We bounded our estimates between 0 and 100.

Figure A2: Change in hacienda population distribution from demographic collapse



(a) Actual



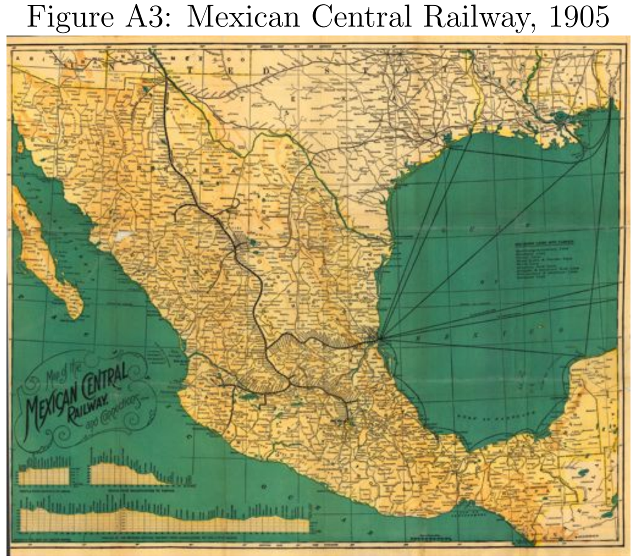
(b) Counterfactual

Appendix A.5 Heterogeneity in impact of the demographic collapse

Many observers have shown that a significant acceleration in land concentration in Mexico took place in the late 19th century following the development of transportation networks and the adoption of export-oriented policies. Coatsworth (1974) in particular has documented a significant period of peasant land alienation and hacienda expansion following the construction of railroads between 1876 and 1910. Though peasant land alienation was enabled by earlier legislation, Coatsworth shows that massive land transfers did not take place until the development of railroads, which played a key role in altering the returns to land and in increasing access to previously remote areas. The significance of railroad development to late 19th-century hacienda expansion suggests that the impact of the population collapse on hacienda importance measured in 1900 might be greater in municipalities along the railway.

Given the confound between transport infrastructure and productivity (see Donaldson 2010 and Duflo 2012, among others), estimating these these impacts is not straightforward. We adopt the approach of Duflo 2012 in proxying for rail links using straight lines between major cities along the railway. Figure A3 shows a map of the Mexican Central Railway at the turn of the twentieth century (Mexican Central Railway 1900). We approximate the railroad network using straight lines between 12 major Mexican cities (from North to South): Ciudad Juarez, Acuña, Monterrey, Torreón, Aguascalientes, Guadalajara, Manzanillo, Mexico City, Veracruz, Coatzalcoalcos, and Oaxaca. We extract the closest straight-line distance from any part of the 1900 municipality to the simulated railroad and estimate whether or not there is heterogeneity in the impact of distance to railroad by interacting this variable with the ratio of tributary densities.

Figure A3: Mexican Central Railway, 1905



The OLS and instrumental variables estimations are shown in Table A8. The coefficients on the key variables of interest are not statistically different from zero in the OLS regressions, but the IV estimates of the effect of the collapse are negative as before. In the IV regressions, the coefficient on the interaction term between the distance to railroad and severity of the collapse is positive, which suggests that the land tenure effect of the population collapse was intensified in areas that were closer to the railroad lines. In both cases, marginal effects calculated at 25th and 75th percentiles of railroad distance are positive and significant, where the effects at the 75th percentile are about half of those estimated for the 25th. Figure A4 shows the marginal effect of the tributary drop across the range of the distances from the railroad using the estimates from column (2). These results are robust to dropping municipalities that include the connecting cities. This evidence is consistent with the accounts of Coatsworth (1974) among others that stress the role of declining transport costs in fueling the 19th-century hacienda expansion, and supports the discussion of Section 2.3.

Figure A4: Marginal effect of population collapse across distance to railroad straight line

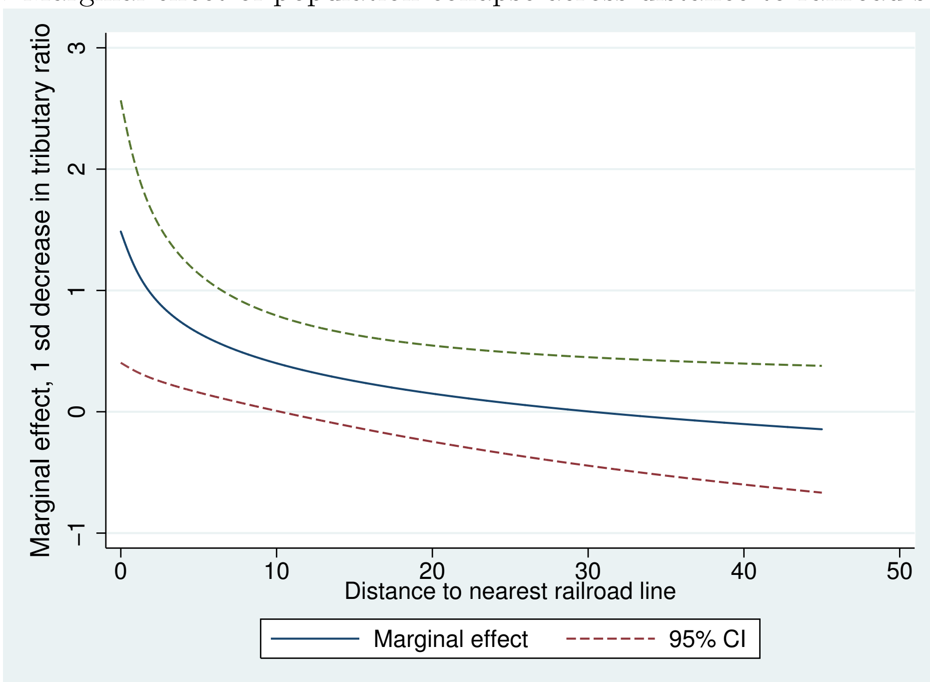


Table A8: Heterogeneity in impact of population decline by distance to railroad line

	Dependent variable					
	% pop on estates			> median estates/km^2		
	(1)	(2)	(3)	(4)	(5)	(6)
1650 density/1570 density	-0.194	-2.064**	-2.535**	-0.008	-0.453*	-0.728*
	(0.137)	(0.913)	(1.149)	(0.056)	(0.266)	(0.388)
Inv hyp sine(100s km to railroad line)	-0.105	0.607*	0.625	-0.074**	0.104	0.143
	(0.085)	(0.363)	(0.417)	(0.029)	(0.112)	(0.158)
Ln(Tributaries 1650/1570) x km to railroad line	-0.005	0.503*	0.583*	-0.019	0.103	0.149
	(0.046)	(0.257)	(0.314)	(0.019)	(0.077)	(0.114)
Other covariates	yes	yes	yes	yes	yes	yes
	OLS	IV	IV	OLS	IV	IV
Observations	1029	1029	920	1029	1029	920
Adjusted R^2	0.332	0.174	0.135	0.196	0.104	-0.039

Unit of observation is the municipality. Standard errors are in parentheses and are clustered at the 1786 administrative level. * p< 0.10, ** p<0.05, *** p < 0.01. These are partial results. Other covariates include elevation < 1000 meters, mean GAEZ maize productivity, 1/km to nearest railroad line, mean elevation, ln(municipal area), standard deviation of PDSI, % population in large cities, ln(km to Mexico City), ln(rural population in 1900), and governorship dummies. Columns (3) and (6) drop municipalities that contain railroad connection towns.