Locational fundamentals, trade, and the changing urban landscape of Mexico

Jennifer Alix-Garcia[*] Emily A. Sellars[†]

February 16, 2018

Abstract

Where do cities emerge and evolve? We examine persistence and change in the distribution of Mexico's urban population from the colonial era to the present, with emphasis on the country's 20th-century transformation. We demonstrate that while early trade patterns and historical persistence were instrumental in sowing the seeds of Mexico's contemporary city system, both technological change and policy significantly altered the trajectory of urbanization. The importance of locational fundamentals decreases over time, while the influence of trade-related geography increases, highlighting that political and economic decision-making shape the importance of geography for development. The findings suggest that although geographic advantage plays an important role in initial urbanization, geography is not destiny in urban development.

JEL codes:
Keywords: urbanization, geography, trade, historical persistence

We are grateful for comments from participants in workshops at the University of Wisconsin, Texas A& M, Oregon State University, and at the conference on Long Run Development in Latin America.

---

[*] Department of Applied Economics, Oregon State University, email: jennifer.alix-garcia@oregonstate.edu
[†] Bush School of Government & Public Service, Texas A&M University, email: sellars@tamu.edu

1 Introduction

Today, as in much of history, power and wealth are concentrated in urban areas. Urbanization entails the reorganization of land use, labor relations, economic behavior, and political organization, and because of this, the formation of cities has often been seen as inherent to economic development itself (Harris & Todaro 1970, Davis & Weinstein 2002, Glaeser 2014). A large literature has sought to explain how, where, and why cities emerge and grow. History and geography loom large over much of this work. Geographic factors—the richness of soils, access to navigable rivers, the ruggedness of terrain, and other persistent features of the landscape—are thought to determine a location's attractiveness for initial urban and economic development. Once established, areas of early population and wealth concentration may build on themselves over time through the benefits of local increasing returns, expectations of future settlement and investment, the positive relationship between economic development and well-functioning institutions, and other factors (Krugman 1991, Davis & Weinstein 2002, Bleakley & Lin 2012, Henderson et al. 2016). Considerable debate remains, however, over the relative importance of geography and history and the limitations of both of these theories. Are certain areas more inherently geographically advantaged than others or is there room for policy in shaping urban development? Moreover, how important was early population concentration in determining today's urban landscape?

This paper examines 400 years of subnational data from Mexico to assess the evolution of how time-invariant geographic features and historical precedent have influenced urbanization. Mexico's urban transformation constitutes an ideal setting to examine sources of persistence and change in urban development for several reasons. Mexico is a large and diverse country in terms of geography, trade access, history, and contemporary population. Though parts of central Mexico have been urbanized for over 500 years, the country remained predominantly rural until the 20th century. In 1900, around 12% of the population lived in cities exceeding 15,000 inhabitants. Today, Mexico's urbanization rate – 80% – matches that of the United States.[1] Some of this shift can be traced to Mexico's massive population growth during the 20th century. However, the country's urban transformation was far from uniform. Some areas that were relatively densely settled in 1900, such as much of the central altiplano, remained rural or even declined in population by 2010. By contrast, large cities like Tijuana and Cancún arose in sparsely settled areas far from historical population centers. Moreover, a disproportionate amount of Mexican urbanization was concentrated around large metropolitan areas as opposed to regional population centers. As a result, there are fewer small cities today than would have been expected based on overall population growth over the century.[2] By contrast, the largest cities, and most notably Mexico City, grew considerably faster than the population as a whole.

Mexico's urban transformation occurred during a century of tremendous economic and political changes, including the Revolution and civil war, a transformational agrarian reform, the rise and fall of one-party rule, and major shifts in trade, industrial, and demographic policies. We investigate whether and how these political and economic factors altered the effect of both geographic fundamentals and historical persistence on city emergence and growth.

Our analysis proceeds in several steps. We first examine the correlates of urbanization and population concentration at several intervals from the sixteenth century to the present. We do this by dividing the territory of Mexico into over 9,000 15-kilometer by 15-kilometer grid cells and extracting geographic covariates and population measures. The grid approach allows us to examine the emergence of urban areas over a long period of time without relying on political divisions, such as locality definitions or municipal boundaries, whose number and size has changed endogenously with population concentration in Mexico (González Navarro 1974, Unikel 1976). This is similar to the approach used in Bosker & Buringh (2015) in the European context.

Using the grid cells, we examine the geographic features—agricultural productivity, ele-

---

[1] Unless otherwise noted, cited statistics represent the authors' own calculations using census data.

[2] See Section 4.

vation, ruggedness, and proximity to waterways, the northern border, and Mexico City—that have been correlated with urbanization in different eras. We also examine the relationship between present and past population concentration in different time periods. We transform the grid cells into a panel to examine continuities and changes in urban development over the 20th century. This approach allows us to observe when and where new urban areas in Mexico emerged and how existing urban populations grew over time.

In line with prior work, we find that geographic fundamentals and path dependence shaped historical and contemporary urbanization. Today, as in the early colonial period, areas that have greater productive capacity and are nearer to Mexico City are more likely to be urbanized than areas in less attractive geographic settings. We also find evidence illustrating the importance of path dependence to urban development. The overall regional distribution of population in Mexico has remained largely persistent from the 16th century to the present in the face of major political and demographic shocks, such as the massive population collapse after the arrival of the Spanish, decades of civil war following independence from Spain, the Mexican Revolution, and the creation of the modern Mexican state.

However, we also demonstrate that the effects of both geographic advantage and historical persistence changed considerably over the last century due to government policy. As the century progressed, economic policies encouraging tourism and free trade shifted the urban population of Mexico northward and towards the coasts. Areas nearer to the northern border with the United States and to natural anchorage areas on the coast became more likely to urbanize and grow as the century progressed and as access to international trade networks increased in importance. By contrast, the relative importance of agricultural productivity and the disease environment declined, resulting in fewer cities in the fertile central areas of the country outside of malarial zones and closer to Mexico City. This calls attention to ways in which political choices intensify or diminish the importance of certain geographic endowments in urbanization either directly (as with policies regarding colonization or urbanization itself) or indirectly (as with policies regulating trade or transport). The causes of urban development are therefore as closely tied with political economy concerns, as are its consequences (e.g., Bates 1981, Ades & Glaeser 1995, Krugman & Elizondo 1996).

This paper contributes to a large literature on urban economics and economic geography, and especially to work on urbanization in developing countries (e.g., Overman & Venables 2005, Glaeser 2014). Geography has long occupied a central place in the study of urbanization, as it has in the literature on economic development more generally (e.g., Pirenne 1925, Weber et al. 1958, Bairoch 1988, De Vries 1984, Mellinger et al. 2000, Nunn & Puga 2012). Many geographic qualities of a location have been used to explain the emergence of cities, including resource abundance, agricultural productivity, waterway access, the presence of natural defenses, and the ruggedness of terrain (e.g., Pirenne 1925, Weber et al. 1958, Bairoch 1988, De Vries 1984). We build on this foundation by providing evidence on how the effects of these natural geographic advantages evolve over time and how this evolution depends on political choices made over trade and industrial policy.

This work also contributes to the literature on path dependence in economic development, both in general and specific to urbanization. Historical dependence has featured in the literature on urban development in two major ways. First, many scholars have noted that the location of existing cities may shape the formation of new ones, both positively (for example, by opening up possibilities for local trade) and negatively (by crowding out the potential for nearby urban growth) (Fujita & Mori 1996, 1997, Behrens 2007, Bosker & Buringh 2015). Second, historical dependence may shape the growth and longevity of existing cities as areas of early population concentration increase in size due to the benefits of increasing returns and agglomeration economies (Krugman 1993, Bleakley & Lin 2012, Henderson et al. 2016, e.g.,) or because of natural demographic growth from existing urban populations (Jedwab et al. 2014). We find evidence of both types of historical persistence in Mexican urbanization; areas with more dense colonial populations are more likely to be urban today, and the largest cities in the country received disproportionate growth during the 20th century through agglomeration.

However, the Mexican case also illustrates the limitations of historical persistence in explaining the contemporary city system. As trade and industrial policy drove settlement and investment northward, remote places like Tijuana (a ranch of less than 300 people in 1900) and Ciudad Juárez (a modest railroad town of around 8,000) grew to become some of the largest cities in the country, while important colonial cities and mining centers nearer to Mexico City, such as Guanajuato and Zacatecas, faded from importance. The role that increasing economic integration with the United States played in remaking Mexico's city system is similar to what has been found in work linking trade patterns and urbanization in Europe (e.g., Redding & Sturm 2008, Redding et al. 2011, Bosker et al. 2013). Geographic advantage is policy- and context-dependent, and policy changes can upend even long-running persistence in investment and urbanization.

A final contribution of this work is to the literature on urbanization in developing and especially middle-income economies. Urbanization tends to increase as countries develop, and more urban population raises the prospects for further economic growth (e.g., Overman & Venables 2005, Jedwab et al. 2014). Several scholars have argued that patterns of urbanization in countries that urbanized early were distinctive from those being observed in "late developers" in Africa, Latin America, or Asia today (Overman & Venables 2005, Jedwab et al. 2014, Henderson et al. 2016). The extent to which lessons from urbanization in medieval Europe or the Middle East can be applied to understanding urbanization in contemporary developing countries is thus unclear. We examine the urbanization trajectory of a country as it grew from low to middle income, a little-studied path that remains open to the majority of the world, but is not yet well understood. Our findings provide evidence on the role that policy can play in structuring urbanization in this setting, highlighting the potential for political actors to reshape regional development patterns.

Our paper proceeds as follows. We begin by reviewing existing theories on geography and path dependence in urbanization, focusing on sources of persistence and change in urban development. We then discuss the sources and construction of the datasets used in the paper. We then present evidence on the correlates of urbanization across time, comparing trajectories during the pre-industrial period (1570–1900) and the period of Mexico's urban transformation (1900–2010). We then use a panel strategy to examine changes in how geography and history shaped urbanization during this second period. We conclude by discussing the implications of the findings for the role of geography and history in development.

2 Geography and path dependence in urbanization

What constitutes an advantageous location for urban development and how has this changed over time? Geography features in the literature on urban economics in two main ways. First, a location's inherent geographic characteristics determine the attractiveness of initial settlement. These "first-nature" characteristics include agricultural productivity, access to mineral or other resources, access to rivers or transport networks, and natural defenses (e.g., Weber et al. 1958, Bairoch 1988, Bosker & Buringh 2015). The attractiveness of settlement also depends on a location's geographic position relative to other centers of population ("second-nature" geography). Locations that are too far from other cities may be too remote from trade networks to develop, while those that are too close may be crowded out by the urbanization of neighboring areas (e.g., Pirenne 1925, Bosker & Buringh 2015).

In both cases, trade costs play a central role in structuring a location's geographic advantage (e.g., Anas et al. 1998, Fujita & Mori 1996, Krugman 1993, Duranton 1999, Konishi 2000, Behrens 2007). They determine how "close" a location is to another one (e.g., Behrens 2007), a second-nature feature. However, trade costs also play an important role in first-nature geography. Urban areas house a higher population density than can be supported on the land they occupy and must import food to survive. When trade costs are high, cities are limited in size by the presence of food production within a small radius of the urban core. As technology advances and as barriers to trade fall, food can be brought in from farther away to support the urban population. Nearby land productivity, water resources, and the other natural features become less important to the placement of cities. This trajectory is eloquently analyzed in Bosker & Buringh (2015), who examine the role of geography in the establishment of the modern city system of Europe. A similar argument is made by Henderson et al. (2016), who show that agricultural productivity played a less important role in determining the location of cities in countries that developed later than those that developed earlier because of improvements in transport technology over time.

The importance of trade in urban development suggests a mechanism through which policymakers could—either purposefully or inadvertently—shape where cities end up and how much they grow. Trade and transit costs depend on available technology, but they also depend on policy. The construction of transit options, such as railroads or maritime ports, can reshape the geographic advantageousness of different locations (Coatsworth 1981, Fujita & Mori 1996, Jedwab et al. 2015). Trade policy can also determine second-nature geography, how advantageous a location is relative to nearby areas. Krugman & Elizondo (1996), for example, present a model to illustrate how trade protectionism and import-substitution incentivize urban primacy by directing industrial production to a concentrated domestic market. Political factors themselves may play a role in determining geographic advantage for urban development. Political instability and uncertainty can drive over-concentration capital cities (Ades & Glaeser 1995), and a classic literature has linked warfare with increased urbanization as people take shelter in densely settled areas (Tilly 1989, Warman 2001, Dincecco & Onorato 2016).

If the foundation and growth of cities are related to policy, this raises the question of how much agency political actors have to alter the path of urbanization. Most research on urbanization highlights that population concentration remains persistent over time. There are several reasons why this is could be the case. One is that some areas may be inherently geographically advantaged over others for supporting a dense population, and these remain attractive areas for investment and settlement regardless of technology or policy (e.g., Davis & Weinstein 2002). Another explanation is that spatial economies of scale and increasing-returns technology lead to economic and demographic concentration in places that urbanize early, even if these places are not or are no longer geographically advantaged. These spatial economies of scale could be due to pre-existing infrastructure or other sunk costs in urban areas, such as schools or housing. Another possibility, suggested in a classic paper by Krugman (1991) among others, is that these historical antecedents help resolve ambiguity about where investment and settlement will locate in the future. Because increasing returns models are characterized by multiple equilibria, the earlier presence of a city can signal to actors where settlement and investment are likely to be located in the future, mitigating "spatial coordination failures" wherein actors fail to co-locate or co-invest with others (Krugman 1991, Behrens 2007, Bleakley & Lin 2012, Jedwab et al. 2015).

Though all of these mechanisms would predict persistence in urbanization patterns, they make different predictions about the role that policy can play in altering the urban landscape. If some areas are inherently geographically advantaged over others, as is argued by Davis & Weinstein (2002), there is little that humans can do to alter spatial inequality. If, however, historical persistence is driven predominantly by spatial coordination and expectations, as both Bleakley & Lin (2012) and Jedwab et al. (2015) find in recent work, policy and technology can play a role in setting a new spatial equilibrium if the signal is strong enough to change beliefs about where settlement and investment will occur in the future. Moreover, if policy and technology can change the attractiveness of a location for initial settlement, choices about these things may have important and persistent consequences for regional development.

In the remainder of this paper, we examine the relative explanatory power of geography, history, and policy on city emergence and growth in Mexico from the 16th century to the present. We focus on continuity and change in urban development and the ways in which policy choices altered the nature of geographic advantage and the pull of historical precedent.

3 Data and Setting

We combine subnational data on population, urbanization, transport, geography, and policy from 1570 to 2010. This section describes the sources of our data and the methodology used to construct our key variables.

3.1 Population and "urbanization"

We use several datasets to trace population and urbanization over time. We rely most heavily on the Historical Archive of Localities (AHL), which is maintained by Mexico's National Institute of Statistics and Geography (INEGI). This dataset contains the population and political classification of the vast majority of localities in Mexico in each census or national count from 1900 until 2010. We use text analysis to link localities over time to their present geographic location. We also cross-reference this data with the hard copy of the locality-level census to ensure that we capture all urban localities in each census year. The data appendix provides further detail on adjustments made to this dataset to make it usable for this project.

Once we link localities to a geographic location, we aggregate information on total population and urban settlements to the level of 225 km^2 (15-by-15 km) grid cells. This is the average size of Mexican municipalities today. The grid cells are our primary unit of analysis. The grid approach addresses the problem that both contemporary and past administrative boundaries and metropolitan area distinctions were determined by urbanization patterns and government policy (González Navarro 1974, Unikel 1976). The grid approach follows the work of Bosker & Buringh (2015) on the emergence of European cities.

We code a grid cell as 1 if it contains one or more urban localities and 0 otherwise. There is no consensus in the literature regarding the definition of a city (Henderson 2005, Glaeser 2007). We therefore use three population thresholds to determine whether a given locality within a grid cell is "urban": a low threshold of 5,000, and alternative threshold of 15,000, and a restrictive threshold of 50,000.[3] We also examine grid-cell population density as an

---

[3] While the first threshold might seem small, INEGI's official cutoff for urban status is a population of only

alternative measure of population and economic concentration, following Davis & Weinstein (2002) and others. To control for potential congestion or agglomeration effects that might confound our analysis, we also calculate the presence of cities within 25- and 50-km radii of each grid cell and condition on these measures in some specifications. Though these spatial spillovers are not a primary focus of our analysis, it is important to include them as our measures of policy, geography, and history are also spatially correlated.

Figure 1 shows the distribution over time of the binary "urban" indicator variables. The number of cities of all sizes increases over the 20th century with the steepest increases occurring between 1940 and 1970. This corresponds with Mexico's demographic boom and the period of rapid industrialization under import-substitution.

3.2 Colonial population

Though comprehensive, locality-level data in Mexico are unavailable prior to the 20th century, we examine district-level data on population concentration beginning in 1570. We examine these data for two reasons. First, we expect that the determinants of population concentration in the colonial era may have been different given differences in technology, economic structure, and politics. Second, given the extensive literature on path dependence and historical persistence in urban economics, we expect that population density during this "baseline" time period may continue to shape urban development today.

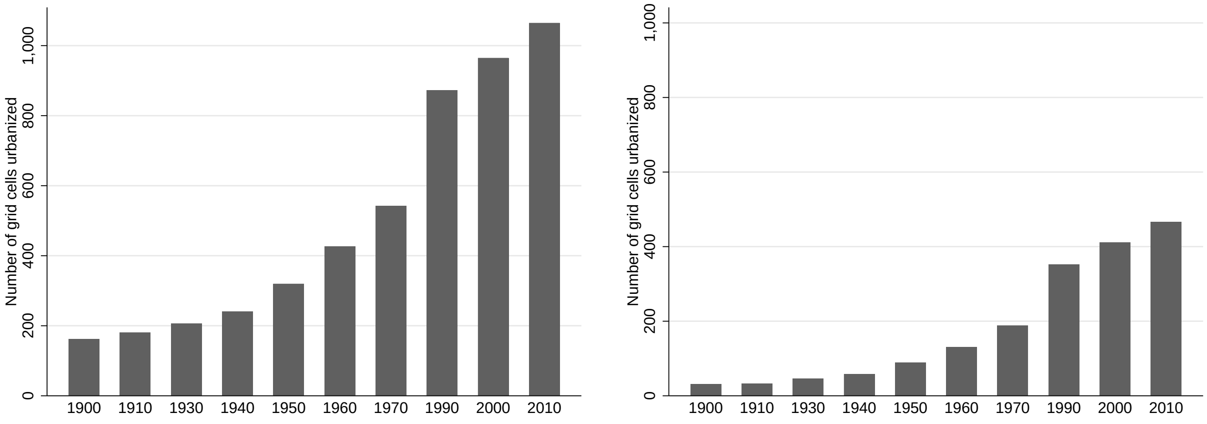
We digitize estimates of colonial population from Gerhard (1993*a*, 1993*b*, 1993*c*), who uses the *relaciones geográficas* as his primary source of information. We examine three snapshots of colonial population: in 1570, the first year at which we have data for much of the colony; in 1650, which is close to the nadir of indigenous population following Mexico's demographic collapse;[4] and in 1800, just before the War of Independence. Gerhard's data are reported at the level of 1786 political divisions and are typically recorded in terms of tributaries, or

---

2,500 people, and Bosker & Buringh (2015) use a 5,000-person cutoff in their analysis of urban development in Europe.

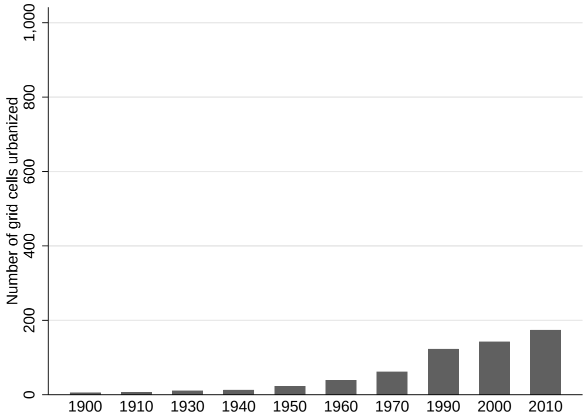
[4] Hassig (1985) and Knight (2002) both estimate that the low point in population was during the 1630s.

Figure 1: Three different ways of identifying “cities”



(a) > 5,000

(b) > 15,000



(c) > 50,000

Figures show the number of grid cells urbanized by decade according to the AHL data, using population cutoffs of 5,000 (a), 15,000 (b), and 50,000 (c). Y-axes are all on the same scale.

the number of individuals paying tribute to the Spanish Crown. Each married male counted as one tributary, and unmarried adults and widows counted as half of a tributary. The ratio of tributaries to total indigenous population was approximately 2.8 in the late 16th century (Gerhard 1993*a*, Cook & Borah 1960, 1971). In 1800, densities are given in total population (rather than tribute units). These data and their reliability are described in extensive detail in Sellars & Alix-Garcia (2017). Figure 2 in Section 4.1 presents the spatial extent of Gerhard's data for 1570 and 1650. Detailed population data are unavailable for much of northeastern Mexico in the early colonial period as the region had not yet come under solid Spanish control.

3.3 Location fundamentals and trade costs

We digitize several measures of locational attractiveness for urban development and trade access using geographic data. These measures include the standard deviation of elevation, to account for the ruggedness of the terrain, and an indicator representing whether the average elevation of a grid cell is lower than 1500 meters. This second measure is meant to proxy for the disease environment as malaria, yellow fever, and other diseases endemic inhibited colonization of low-lying areas (e.g., Hassig 1985). Both measures were extracted using a 30-meter digital elevation model maintained by INEGI. We also include a measure of potential agricultural productivity for low-input, rain-fed maize, which has been the main staple crop in Mexico for centuries. This measure was extracted from the Food and Agriculture Organization's Global Agro-Ecological Zones (GAEZ) dataset (IIASA/FAO 2012). The GAEZ potential productivity measure is calculated using exogenous information on climate, soil type, slope, and average rainfall calculated at a 5-arcminute resolution (about 10-km cells in Mexico). The potential productivity of maize is highly correlated with that of many other grain and vegetable crops. Because Mexico has no major navigable rivers, we include freshwater access among our proxies for geographic advantage (as opposed to trade advantage), calculated as distance to the nearest river or lake using contemporary spatial data from

INEGI.[5]

Because our study period spans a great number of years, direct proxies for transportation costs and internal trade are limited. We use qualitative data to examine changes in these factors in different time periods, as we discuss below. We use two measures of access to external trade: the Euclidean distance from each grid cell to the contemporary U.S. border and to the nearest "anchorage area" (NOAA 2014). Anchorage areas include both ports and natural harbors. Ports are places where infrastructure allowing for trade through shipping already exists, and natural harbors are areas with sufficient depth and protection to allow ships to drop anchor. These points are presented in Figure 2. We include distance to coast as a control variable to help isolate the effect of ports on city development. We also calculate the Euclidean distance to Mexico City. Given the long-run concentration of population in this area, this variable captures access to internal trade and political centrality. It also is a proxy for historical persistence, as we discuss below.

4 Persistence and change in population concentration

In the subsections that follow, we estimate the correlates of urban concentration over time using a series of cross-sectional regressions:

$$S_{us} = F(\alpha + \mathbf{X_u}\beta + \epsilon_{us}) \tag{1}$$

Our dependent variable is either a binary indicator (S_{us}) capturing whether a grid cell includes a locality of over 5,000, 15,000, or 50,000 population in a given year or a measure of population density. We use a probit model for the binary dependent variable models[6] ($F(\cdot) = \Phi(\cdot)$) and ordinary least squares in the second. The vector $\mathbf{X_u}$ includes the geographic

---

[5] There have been some shifts in the placement of rivers, streams, and lakes over time because of human intervention and other factors. One prominent example is the massive drainage project around Mexico City in 17th century.

[6] We use a probit model to enable us to calculate well defined predicted probabilities for urban development. The results are robust to using OLS for these specifications as well.

covariates discussed above. We include a spatially and temporally lagged dependent variable in some specifications to account for agglomeration and congestion in city development. We deliberately do not include state fixed effects in these specifications as we are interested in measures of geographic access and locational quality overall and do not wish to identify these effects off of within-state variation. Unless otherwise noted, standard errors are clustered at the level of the modern state.

4.1 Urban development in Mexico, 1519–1900

We begin by examining the correlates of population concentration in the colonial period. At the time of the Spanish Conquest, the Triple Alliance (Aztec Empire) was a large and urbanized society with as many as 20 million residents (Gibson 1964, Cook & Borah 1971, Hassig 1985). The center of population and power was the altiplano in the center of the country, which had several geographic advantages that encouraged population concentration. The area around Tenochtitlan (now Mexico City) is located outside of the tropical zone in an area where maize productivity is relatively high. Though this area is mountainous, Tenochtitlan was located in the middle of a large lake system. This was used for trade with hinterland areas to feed city residents and accumulate wealth (Hassig 1985, Ch. 2). The lake system was especially important given the high cost of overland transit. Prior to the arrival of the Spanish, there were no pack animals and the wheel was not used for transport. Overland trade took place almost entirely on foot, and the high cost of transport placed natural limits on the size of localities and the structure of the urban environment (Hassig 1985).

Existing demographic and trade patterns were upended following the Conquest in 1519–21. Colonial rule ushered in a new political and economic structure, though the Spanish maintained Mexico/Tenochtitlan as the political capital. The Conquest was also accompanied by one of the most dramatic demographic collapses in history. By 1650, a combination of disease, drought, and famine reduced the population by over 90% by some estimates (Cook

& Borah 1971, Hassig 1985). The magnitude of the collapse has few parallels in history, the closest historical comparison arguably being the depopulation of Europe during the Black Plague.

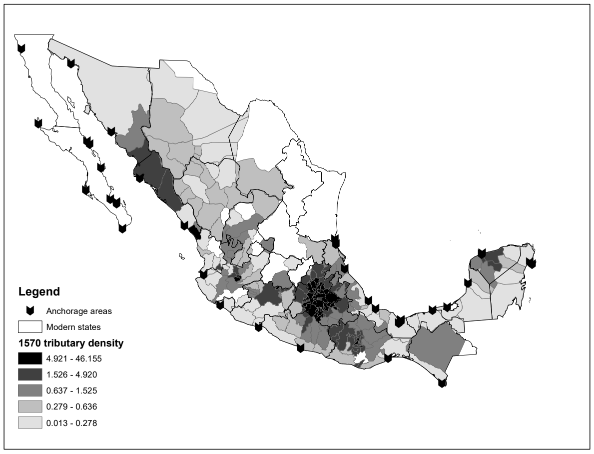
Figure 2 presents the geographic distribution of population concentration at three points during the colonial period: around 1570 (the first year in which our data are relatively complete), around 1650 (close to the nadir of Mexico's population), and 1800 (at the end of the colonial era). Darker colors indicate areas of greater population/tributary density. In 1570, the areas of greatest population concentration were predominantly around Mexico City. Some areas of contemporary Oaxaca, Yucatán, Michoacán, and Sinaloa also had relatively high population densities due to pre-Columbian settlement (Gerhard 1993*a*). While comprehensive population data from the time of the Conquest do not exist, the distribution in the 1570 map conforms with qualitative descriptions of the earlier time period in archival documents and other sources (e.g., Cook & Borah 1971, Hassig 1985). Though much of Mexico was severely affected by the population collapse of the 16th century, the area around Mexico City retained its relative dominance in 1650.[7]

As the colonial period progressed, population shifted towards mining and agricultural production in the center north and west that had not previously supported dense populations (Unikel 1976, Knight 2002). Between 1700 and 1800, regional centers like Guadalajara grew into cities, driven by new immigration from Spain and the emergence of a "merchant elite" benefiting from rising inter-colonial trade (Unikel 1976, Knight 2002). By royal policy, virtually all external trade was routed through the port cities of Veracruz and Acapulco, which led to a reorganization of transport networks to facilitate movement of goods and people between these cities and central Mexico (Unikel 1976, Hassig 1985, p. 169–171). Settlements along the roads between Mexico City and the ports thus saw considerable growth, while pre-Columbian commercial centers bypassed by the road, such as Texcoco, declined in population and importance (Hassig 1985, p. 174–5).

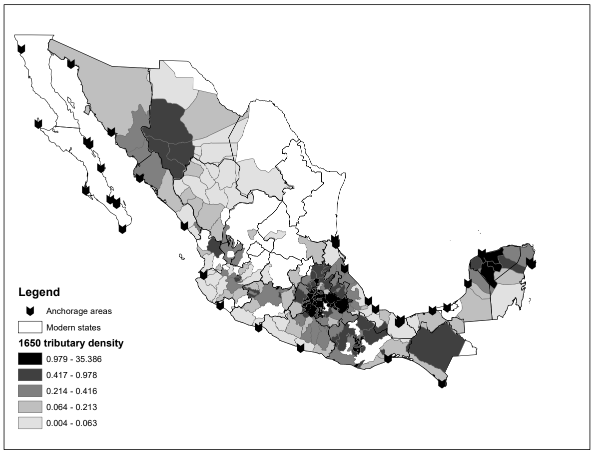
---

[7] Note that the scale of the maps is recorded in quantiles; the 1650 density numbers are of a much smaller magnitude than those of 1570 due to the collapse.

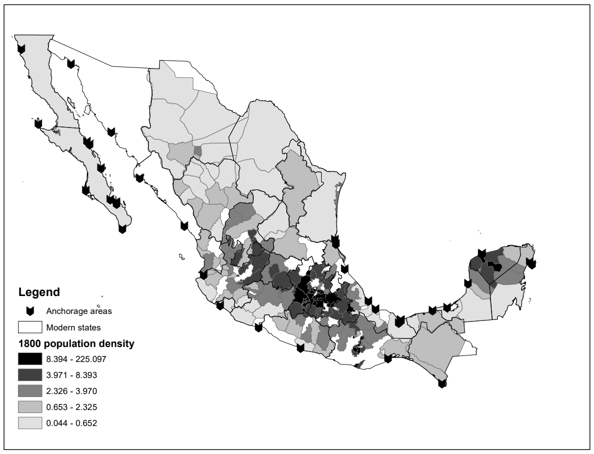
Figure 2: Colonial era population densities



(a) 1570



(b) 1650



(c) 1800

White spaces are missing data. 1570 and 1650 are measures of tributaries to the crown, while the 1800 density measures total population per square kilometer. Legend shows higher density quintiles in darker colors. Quintiles are within year.

However, despite the major economic, political, and demographic upheaval of the colonial era, what is most striking is the persistence in the distribution of population. Population density at the end of the colonial period in 1800 remained concentrated around Mexico City and highly correlated with population density in 1650 ($\rho = 0.70$) and 1570 ($\rho = 0.83$). In Table 1, we more systematically examine the relationship between population density at these three intervals and the geographic covariates described above. Reported are OLS estimates using the log of population density in 1570 (Columns (1-2)), 1650 (Columns (3-6)), and 1800 (Columns (7-11)) as the dependent variables.

The strongest determinant of population density in 1570 is distance to Mexico/Tenochtitlan, the economic and political center of the empire. This is also true of density in 1800, though not in 1650. This could be because the demographic collapse from disease, which reached its worst point around 1650, was more severe near Mexico City. More population density in 1570 is concentrated in higher elevations outside of malarial/disease zones, though this correlation diminishes in later time periods. It is difficult to identify the effect of maize productivity and it is highly correlated with distance to Mexico City. If we remove distance to Mexico City from our estimation, maize productivity has a large and positive effect on population density in 1570, but it is negatively correlated with density in 1800 once conditioning on earlier population density. More rugged terrain is also positively related to population concentration in 1570, consistent with the hypothesis that early settlements thrived in areas that were easier to defend. Terrain ruggedness is negatively related to late colonial density, however. This would be consistent with a decline in the importance of defensive location. Notably, distance to ports and anchorage areas had no discernible effect on population density in the early colonial period, but it is negatively related to density in 1800 in line with the increasing importance of maritime trade in late colonial era.

The most striking finding is the marked persistence in population density over time. After nearly 300 years of Spanish rule, population density in 1800 remained higher in areas of dense pre-colonial settlement. The coefficient estimates indicate strong historical persis-

Table 1: Correlations of colonial density, grid cells

	Ln(density, 1570)		Ln(density, 1650)				Ln(density 1800)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Elevation < 1500 m	-0.449**	-0.562***	-0.297	0.002	-0.335	0.023	-0.112	-0.348**	-0.009	-0.405	0.045
	(0.165)	(0.197)	(0.311)	(0.298)	(0.294)	(0.303)	(0.174)	(0.143)	(0.128)	(0.239)	(0.165)
Std grid cell elevation	0.001*	0.002**	0.001	-0.000	0.001	-0.000	-0.001**	-0.001**	-0.001***	-0.000	-0.001***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
Ln(maize productivity)	0.396	0.603**	0.405	0.071	0.479	0.012	-0.230	-0.516***	-0.401***	-0.200	-0.288**
	(0.263)	(0.264)	(0.301)	(0.327)	(0.314)	(0.318)	(0.164)	(0.163)	(0.115)	(0.188)	(0.110)
Ln(Km to border)	0.036	0.223	0.094	0.047	0.155	-0.013	0.377***	0.378**	0.280***	0.581**	0.315***
	(0.201)	(0.202)	(0.208)	(0.146)	(0.188)	(0.148)	(0.106)	(0.150)	(0.074)	(0.208)	(0.077)
Ln(Km to DF)	-0.492***		-0.168	0.189			-0.225	-0.686***	-0.270***		
	(0.173)		(0.208)	(0.221)			(0.139)	(0.120)	(0.090)		
Ln(km anchorage)	-0.137	-0.173	-0.245	-0.120	-0.259	-0.112	-0.178*	-0.249*	-0.219**	-0.301***	-0.227**
	(0.175)	(0.180)	(0.194)	(0.116)	(0.189)	(0.127)	(0.087)	(0.123)	(0.092)	(0.105)	(0.088)
Ln(km coast)	-0.012	0.023	0.035	0.033	0.046	0.022	0.035	0.113*	0.037	0.156*	0.033
	(0.092)	(0.099)	(0.084)	(0.065)	(0.083)	(0.067)	(0.053)	(0.064)	(0.052)	(0.082)	(0.056)
Ln(km water)	-0.032	-0.039	-0.033	0.005	-0.035	0.006	0.032	0.027	0.049*	0.024	0.052**
	(0.038)	(0.037)	(0.073)	(0.051)	(0.070)	(0.055)	(0.032)	(0.027)	(0.026)	(0.025)	(0.024)
Ln(Tributary density, 1570)				0.760***		0.725***	0.750***		0.657***		0.793***
				(0.160)		(0.126)	(0.114)		(0.075)		(0.087)
Ln(Tributary density, 1650)								0.496***	0.181**	0.528***	0.123
								(0.110)	(0.081)	(0.119)	(0.100)
Observations	6491	6491	5842	5736	5842	5736	4495	4017	3951	4017	3951
Adjusted R^2	0.309	0.242	0.139	0.498	0.132	0.489	0.724	0.729	0.817	0.618	0.805

Unit of observation is the 15 x 15 km grid cell. Clustered standard errors are in parentheses. * $p < 0.10$, ** $p<0.05$, *** $p < 0.01$.

tence in the interim as well; the elasticity of both 1650 and 1800 density to that of 1570 is around 0.75. This relationship remains even after conditioning on distance to Mexico City. Though consistent with much of the work on historical persistence in urbanization, this persistence is somewhat surprising given the context. Colonial rule forced a major reorganization of religious, social, and political institutions (e.g., Hassig 1985, Gerhard 1993*a*). This reorganization occurred alongside the dramatic demographic collapse discussed above, which reduced the indigenous population by over 90% by some estimates. Even in the face of this unprecedented population shock, however, populations rebounded in many of the same locations as they had occupied prior to the collapse. This suggests that the shift in societal organization induced by the Conquest and the collapse was not enough to alter the long-run spatial equilibrium of settlement, investment, and political power.

Several reasons may account for this persistence. First, colonial institutions were initially designed to extract wealth quickly by usurping the existing Aztec tribute system. This incentivized colonial investment and settlement in areas of dense pre-colonial population (Gibson 1964, Hassig 1985, Gerhard 1993*a*). Though strategies of wealth extraction shifted in the wake of the demographic collapse (e.g., Sellars & Alix-Garcia 2017), settlement patterns persisted, and most immigration from Spain was driven to pre-existing urban areas (Unikel 1976, Knight 2002, p. 61, p. 177-8). Second, despite the political and economic reorganization of the colony, taxation policy kept the population tied to specific villages, inhibiting internal migration (Knight 2002). Finally, because the cost of overland transit remained high, interregional trade was limited, and settlement patterns remained driven by nearby grain production in the central altiplano.

Settlement patterns changed little following the War for Independence (1810–21). Mexico experienced stagnant demographic and economic growth for most of the 19th century due to ongoing warfare, economic uncertainty, and political fragmentation.[8] The road system remained limited due to a lack of investment in infrastructure, inhibiting trade and move-

---

[8] These factors are described in depth in Coatsworth (1978) and Haber (1992), among others.

ment. As late as 1873, the country had less than 5-km of high-quality road (i.e., passable by four-wheeled carts) per 10,000 inhabitants, less than one-tenth of the U.S. figure, and the constant risk of banditry added to transportation costs (Haber 1992, Beatty 2001). Internal trade remained concentrated around Mexico City and in a series of disconnected regional markets (Unikel 1976).

There were few significant changes in infrastructure until the Porfirian era (1876–1910). During this time, railroad development increased dramatically, and freight costs dropped to around one-third of their pre-railroad levels (Coatsworth 1978). This period also coincided with an expansion of political support for export-oriented economic activity through a combination of tax relief and tariff relaxation. This led to a period of economic growth and land concentration, though a limited increase in international trade based on available statistic (Table B1). At the turn of the 20th century, Mexico remained predominantly rural. Around 65% of the working age population was employed in agriculture (Reynolds 1970), and nearly 90% of the population lived in localities of less than 15,000 people.

In Table 2, we examine correlations between city location in 1900 and our main geographic and historical covariates. Reported are the regressions using the 15,000-person cutoff for city location for consistency with later results. Results using the cutoff of 5,000—arguably more relevant for early urbanization (Bosker & Buringh 2015)—are presented in Appendix Table B3. In 1900—as in the late colonial period—grid cell elevation, distance to anchorage, and (in some specifications) agricultural productivity are positively related to population concentration. Distance to Mexico City is negatively related, but is only marginally significant when included with historical density. Historical population density is not itself significantly related with urbanization in these specifications, which is potentially surprising. However, we believe that this is a function of the restrictive definition of urbanization given the time period. Only 33 localities in Mexico had reached 15,000 inhabitants as of 1900, and the country remained predominantly rural. Results using a more expansive 5,000-person cutoff for urbanization (Table B3) indicate strong and positive correlations with 1570 density as

Table 2: Correlates of urban grid cells in 1900

	Dep. var: Has locality of > 15,000 population = 1								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Elevation < 1500 m	-0.006***	-0.006***	-0.006***	-0.005**	-0.006***	-0.007***	-0.005***	-0.007***	-0.006***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Std grid cell elevation	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(km coast)	0.000	0.000	-0.000	0.000	0.000	0.000	0.000	0.001	0.001
	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Ln(km water)	0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	0.000	-0.000
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
Ln(maize productivity)	0.003**	0.001	0.001	0.001	0.002	0.002	0.001	0.004***	0.002
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
Ln(Km to border)	-0.000	0.000	-0.000	-0.000	0.001	0.000	0.000	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(Km to DF)	-0.002***	-0.001*	-0.001*	-0.001*					
	(0.001)	(0.001)	(0.001)	(0.001)					
Ln(km anchorage)	-0.003***	-0.002***	-0.003***	-0.002***	-0.003***	-0.003***	-0.002***	-0.003***	-0.003***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(Tributary density, 1570)		0.001		-0.000	0.001**		0.001		0.001**
		(0.001)		(0.001)	(0.001)		(0.001)		(0.001)
Ln(Tributary density, 1650)			0.001	0.001		0.001*	0.001		
			(0.001)	(0.001)		(0.001)	(0.001)		
City of 15000, t-1, 50 km								0.000	-0.004**
								(0.002)	(0.002)
Observations	9342	6491	5842	5736	6491	5842	5736	9342	6491
Adjusted R^2									

Unit of observation is the 15 x 15 km grid cell. Clustered standard errors are in parentheses. * $p< 0.10$, ** $p<0.05$, *** $p < 0.01$. Coefficients robust to excluding Mexico City area. Reported coefficients are marginal effects from a probit.

well as a stronger relationship with maize productivity.

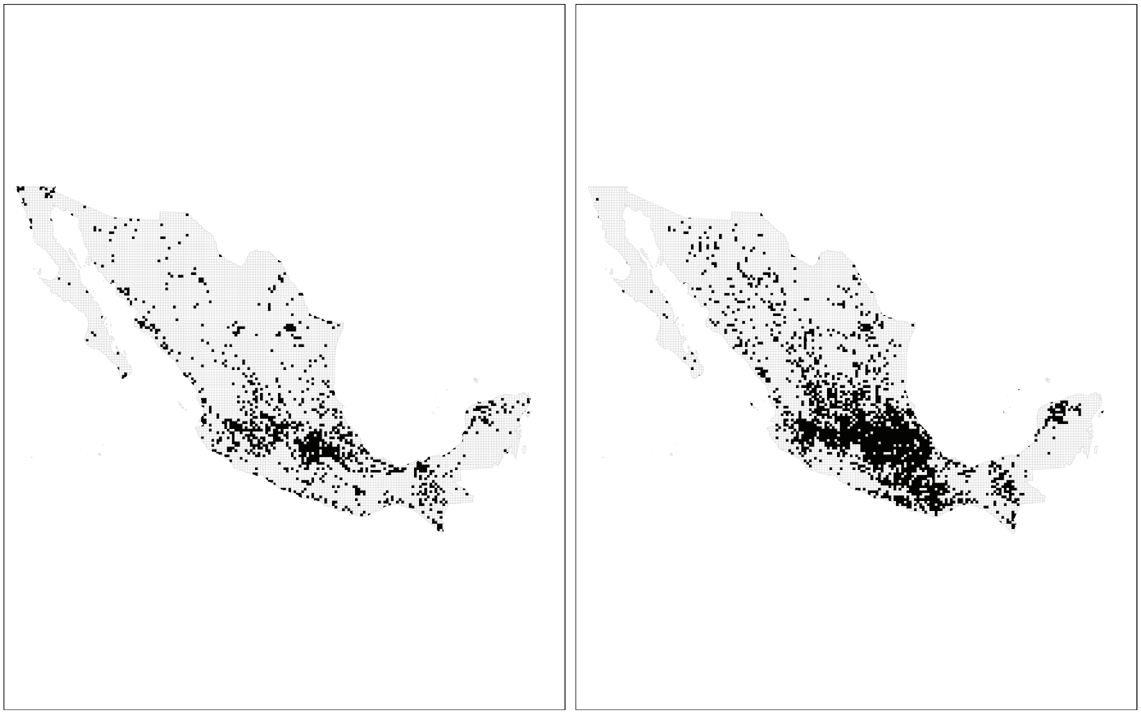
At the time of the Revolution in 1910, the distribution of Mexico's population strongly resembled that of the early colonial period. Most of the country remained rural. Though trade and economic restructuring had shifted populations somewhat toward the coasts and mining areas, settlement and investment remained concentrated on traditional population centers. There were only six cities of above 50,000 population in 1900—Guanajuato, Monterrey, Puebla, San Luís Potosí, Guadalajara, and Mexico City—and all had been important population centers during the colonial period. Over the next century, however, Mexico's demographic structure and city system changed dramatically. By 2000, the vast majority of Mexicans lived in cities, and some of the country's largest metropolitan areas, such as Tijuana or Ciudad Juárez, had grown up in sparsely settled areas far from the dense altiplano of the center. In the remainder of the paper, we examine sources of persistence and change in Mexico's urban transformation

4.2 Mexico's Urban Transformation

There are significantly more Mexican cities today than there were in 1900. Figure 3 presents the location of urban centers in contemporary Mexico. In the left column, dark grid cells indicate the cell contains at least one "urban" locality as of 2010 using the 5,000 (top) and 50,000 (bottom) population thresholds described above. Just over 10% of the grid cells contain a locality with a population of 5,000 or more. Urban areas under the 5,000-person definition cluster in the center of the country, along the coasts, and along the highways leading to border cities. There are over 200 localities of 50,000 people or more in Mexico, but they are clustered in fewer grid cells. While spread throughout the country, there are several clusters of large cities around the sizable metropolitan areas of Mexico City, Guadalajara, and Monterrey.

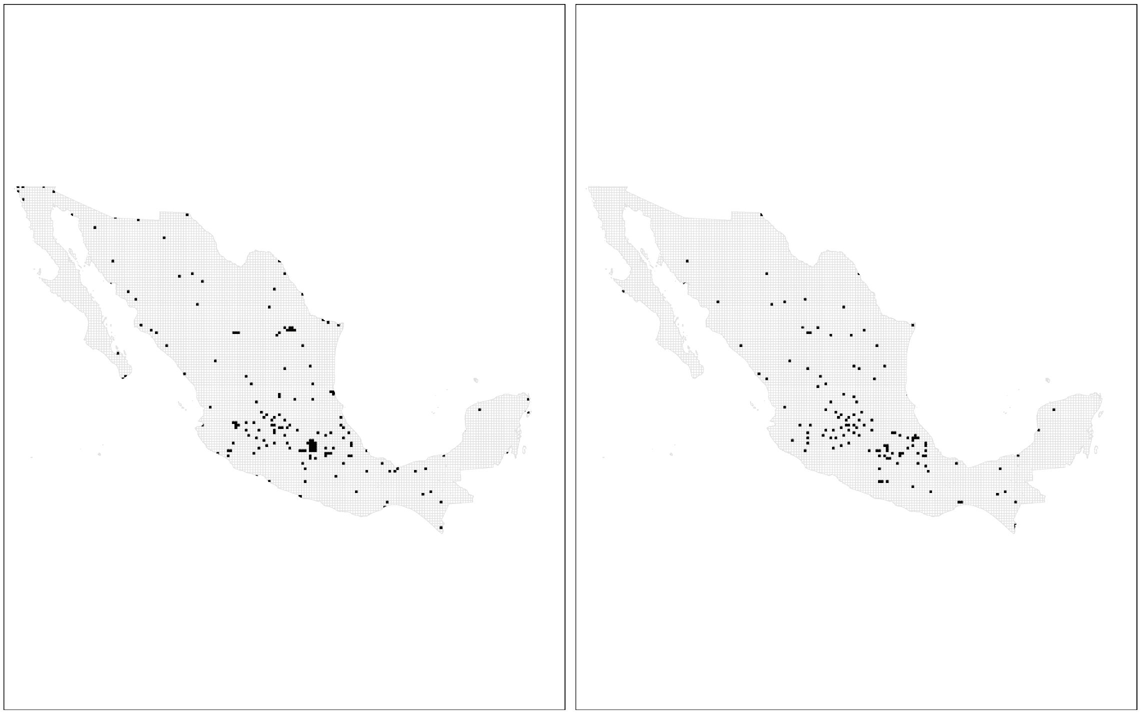
Demographic change can explain some of the country's rapid urbanization. The country grew by a factor of eight between 1900 and 2010. However, population growth alone cannot

Figure 3: Mexico's urban distribution: actual and counterfactual



(a) Actual: > 5,000

(b) Counterfactual: > 5,000



(c) Actual: > 50,000

(d) Counterfactual: > 50,000

Units on map are 15 x 15 km grid cells. Cells are black when population in a grid cell locality exceeds the relevant threshold. Left side panels are the actual city distribution in 2010. Right side panels present counterfactuals under the assumption that 1900 localities all grew at the same rate.

explain the geographic pattern of city growth. In the right column of Figure 3, we examine what Mexico's urban system would look like today had all localities in the country had grown at the same rate between 1900 and 2010. To create this counterfactual population, we use each locality's 1900 population as the baseline and multiply this by Mexico's 2010 population (112,336,538) over its 1900 population (13,607,259). Under the counterfactual, there would be many more—nearly twice as many—grid cells containing localities of 5,000-people or more as actually exist in Mexico. The distinction is especially evident in the center of the country, where many small towns failed to grow proportional to population. By contrast, there would be somewhat fewer grid cells today containing large, 50,000-person cities. The spatial distribution of large cities also differs from the counterfactual. More large cities are clustered around central metropolitan areas and more large cities are located along the U.S. border than would otherwise be expected.

Why did some areas of Mexico urbanize rapidly while others failed to grow? In this section, we discuss sources of persistence and change in Mexico's urban development during the 20th century. We divide the century into three eras: revolution and restructuring (1900–40), industrialization and import-substitution (1940–70), and the period of political and economic liberalization (1970–2010). We summarize important features of each era in Table B1 in the appendix. We build on this analysis in Section 5 by estimating a series of flexible panel regressions to examine continuities and changes in the effects of location fundamentals and historical persistence in urban development over time.

4.2.1 Revolution and restructuring: 1910–1940

The Mexican Revolution of 1910 and subsequent civil war severely disrupted urbanization patterns. Between 1910 and 1921, the population of Mexico declined by about 5% due to violence, disease, and emigration to the United States. Major transportation networks—particularly railroads—were destroyed by revolutionaries hoping to weaken the regime (Coatsworth 1978). Economic growth and trade faltered. The accumulated value of exports fell by 22 percent in 1921 relative to 1912 levels (Ficker 2004). Though much of the fighting had ended by the 1920s, political control of the country remained tenuous. Between 1911 and 1914, there were five presidents, and between 1914 and 1934, seven. Violence and unrest continued in much of the countryside, temporarily driving people into regional population centers and out of villages (González Navarro 1974, Warman 2001).

As the government reconsolidated power, trade and export growth began to recover and eventually exceeded 1910 levels. However, the economic crisis of 1929 provided another major shock to the Mexican economy and society by halting export growth and precipitating the return of hundreds of thousands of migrants from the United States (González Navarro 1974, Knight 1991, Haber et al. 2003, Ficker 2004). In the wake of the crisis, there was a leftward turn in politics, culminating in the presidential election of Lázaro Cárdenas in 1934. Under Cárdenas, the government expanded agrarian reform, centralized control of labor organizations, and nationalized the majority of forest, mineral, and oil resources. Political control was reestablished through the consolidation of one-party rule (Haber et al. 2008). This laid the foundation for the period of rapid growth and industrialization that followed.

4.2.2 Industrialization and import-substitution: 1940–1970

In 1940, the population remained rural, with 20 percent living in cities (Sobrino 2010) and 50 percent living in communities of less than 1,000 people (Unikel 1976). Sixty percent of population growth during the 1920s and 1930s had been concentrated in communities of less than 2,500 people (Unikel 1976, p. 27), and employment in agriculture was at its peak level of 70% of the working population (Reynolds 1970). These patterns changed dramatically over the next 30 years with the industrial boom during and after World War II.

During the war, worldwide demand for industrial products drove up wages and decreased urban unemployment in Mexico, stimulating rural-urban migration. Modest-sized border towns, such as Tijuana, Ciudad Juárez, Reynosa, and Matamoros, saw rapid growth as a result of both increasing U.S. demand for Mexican industrial products and the relocation of population towards military bases in the south of the country (Unikel 1976, p. 39). This early urbanization was fostered by the policies of President Ávila Camacho (1941–46), who encouraged industrial investment and rural-urban migration to support manufacturing. Mexico's domestic market for industrial products grew further with the introduction of President Alemán's program of import-substituting industrialization (ISI) in the late 1940s, which increased demand for Mexican manufacturing. To support this industrial growth, road construction and paving increased rapidly beginning in 1950. The share of manufacturing in GDP increased from 15 to 25 percent from 1940 to 1970 as a result of these and other factors.

This was also a time of demographic increase in Mexico, and most of the new growth was concentrated in cities. The urban population of Mexico increased by a factor of five between 1940 and 1970. Expressed in terms of total population growth, the growth in urban population accounted for 53–65 out of every 100 new Mexican citizens between 1940 and 1970 (Unikel 1976, p. 25). By 1970, 45% of the population lived in localities of greater than 15,000 people, and another 8% lived in moderate towns of between 5,000 and 8,000 people. For the first time in modern history, the majority of Mexicans lived in large towns and cities (Unikel 1976, p. 31). The transition from agricultural to non-agricultural employment mirrored this shift: from its peak of 70% in 1930, the percentage of employment in agriculture dropped to 65 in 1940, 58 in 1950, and 54 in 1960 (Reynolds 1970).

This period also saw important shifts in the regional distribution of Mexico's urban population. It was during this period that the north began to grow quickly. A linchpin of the ISI system had been the *maquiladora* program, which began in 1966 and encouraged the opening of labor-intensive export-processing plants along the northern border (Moreno-Brid et al. 2005). By 1970, northern cities like Tijuana and Juárez had grown to become some of the largest in the country (Unikel 1976).

4.2.3 Political and economic liberalization: 1970–present

In the 1970s, the “Mexican Miracle” of economic growth began to falter. The government took advantage of high oil prices to finance growth through public expenditures and ensure political survival. Oil-dependence exposed the country to external shocks. When the oil market collapsed in 1981, the government was forced to declare a moratorium on external debt service payments, and the period of expansion was over (Moreno-Brid et al. 2005). This ushered in nearly two decades of economic crises and stagnating growth (Haber et al. 2008). Starting in 1982, Mexico became an experiment in trade liberalization, deregulation, and privatization. Previous subsidies and fiscal incentives to exporting industries were eliminated. In 1985 the government signed a bilateral trade agreement with the U.S. that served as a precursor to NAFTA. The following year, it signed onto the General Agreement on Tariffs and Trade. Commitment to economic liberalization culminated with the start of NAFTA in 1994.

The percentage of exports in GDP rose significantly during this time period, from around 8% in 1970 to almost 30% currently. The composition of exports also shifted. The share of exports in manufacturing rose from 20 to 80 percent between 1980 and 2004, while the percentage of oil exports fell from about 60 to near 10 percent. Agricultural exports fell to less than 5 percent by 2004 with declines in agricultural employment (Moreno-Bird et al. 2004). Urbanization continued to climb, but the rate of growth leveled off somewhat as demonstrated in Figure 1 above.

By 2010, 80% of the country lived in urban areas. However, the geographic position of these areas differs somewhat from earlier periods. In Table 3, we examine the geographic-, policy-, and history-related correlates of contemporary city locations. These reveal interesting continuities and changes in urban development over the 20th century. Disease environment no longer seems to predict city location. However, distance to coast, the border, and anchorage areas all have substantial effects. When we control for spatial structure by including spatially and temporally lag existence of cities, the magnitude and significance of the other coefficients in the regression remain largely the same. Cross-equation tests of the coefficients of this estimation with that of 1900 reject equivalence with a p-value of 0.004.

As in the regressions on colonial population, distance to Mexico City remains immensely important in explaining urban development in 2010. However, this effect is difficult to substantively interpret. Mexico City continues to represent the center of political power in the country; this is the factor that has been stressed most frequently in the political economy of urbanization literature with reference to this city in particular (e.g., Ades & Glaeser 1995, Krugman & Elizondo 1996). However, it is also home to a very large consumer market, which is also an explanatory factor for agglomeration economies. Potentially most important, however, is the fact that Mexico City has been the center of population density for over 500 years. Population density in 1570 has a positive and large effect on the probability of a grid cell being urban today, consistent with some historical persistence. In the next section, we further probe the differences between 1900 and 2010 cities and examine the timing of city "emergence" in grid cells with different characteristics.

5 Continuities and change in urban development

A major theme of early research in the economic geography of urbanization is the development of "primate cities" – cities which dominate a country or region both in terms of size and political influence. Seminal works in this area emphasize the key role of internally-focused trade and high transport costs in the development of these cities (Ades & Glaeser 1995, Krugman & Elizondo 1996, Glaeser 2014). One implication of this work is that these metropolises should shrink as a country liberalizes. Mexico City is often highlighted as a canonical example of a primate megacity that developed as a result of numerous political (e.g., the centralization of power and one-party rule) and economic (e.g., ISI policies) factors.

As a first step in examining changes in urban development over the 20th century, we examine the trajectory of urban primacy in Mexico. We present the Gini coefficients of

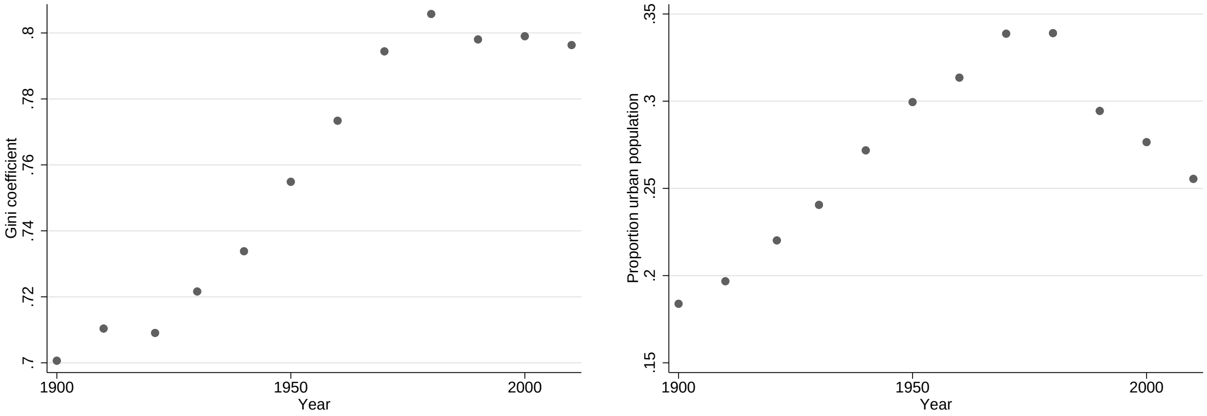
Table 3: Correlates of urban grid cells in 2010

	Dep. var: Has locality of > 15,000 population = 1								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Elevation < 1500 m	-0.021*	-0.020	-0.013	-0.012	-0.029*	-0.028	-0.022	-0.033**	-0.029*
	(0.012)	(0.013)	(0.012)	(0.012)	(0.017)	(0.019)	(0.018)	(0.015)	(0.015)
Std grid cell elevation	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(km coast)	0.007*	0.012**	0.011**	0.010**	0.016***	0.014***	0.015***	0.008**	0.014***
	(0.004)	(0.005)	(0.004)	(0.004)	(0.006)	(0.005)	(0.005)	(0.004)	(0.005)
Ln(km water)	-0.005**	-0.005*	-0.005	-0.004	-0.006*	-0.006*	-0.005	-0.005**	-0.005*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)
Ln(maize productivity)	0.033***	0.027***	0.025**	0.022**	0.040***	0.046***	0.034***	0.048***	0.037***
	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)	(0.011)	(0.011)	(0.009)	(0.010)
Ln(Km to border)	-0.010	0.006	0.004	0.005	0.024**	0.023**	0.021***	-0.001	0.016**
	(0.007)	(0.011)	(0.008)	(0.008)	(0.011)	(0.009)	(0.008)	(0.006)	(0.008)
Ln(Km to DF)	-0.038***	-0.034***	-0.036***	-0.033***					
	(0.006)	(0.006)	(0.006)	(0.006)					
Ln(km anchorage)	-0.025***	-0.029***	-0.029***	-0.028***	-0.033***	-0.033***	-0.033***	-0.022***	-0.029***
	(0.003)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.007)	(0.003)	(0.005)
Ln(Tributary density, 1570)		0.009***		0.007	0.019***		0.019**		0.017***
		(0.003)		(0.004)	(0.006)		(0.009)		(0.005)
Ln(Tributary density, 1650)			0.008**	0.004		0.015**	0.002		
			(0.004)	(0.005)		(0.007)	(0.007)		
City of 15000, t-1, 50 km								0.050***	0.037***
								(0.011)	(0.011)
Observations	9342	6491	5842	5736	6491	5842	5736	9342	6491
Adjusted R^2									

Unit of observation is the 15 x 15 km grid cell. Clustered standard errors are in parentheses. * p< 0.10, ** p<0.05, *** p < 0.01. Coefficients robust to excluding Mexico City area. Reported coefficients are marginal effects from a probit.

city size in Mexico by year, using the aggregated AHL data within areas defined as urban polygons by the Mexican government[9]. We also plot the percentage of urban population living in Mexico City for each year in our sample (Figure 4). The Gini coefficient is relatively flat until 1930. It increases steadily until 1980, when it levels off and even slightly decreases. The proportion of the population living in Mexico City rises increasingly steeply until 1970, and begins to drop significantly in 1990.

Figure 4: Gini coefficients of SUN cities and population in DF



(a) Gini

(b) Proportion urban population in Mexico City

The temporal pattern evident in the figure suggests a correlation between the relative closure of the economy and adoption of ISI policies (predominantly between 1940 and 1970) and the concentration of the population in and around Mexico City. Further examination of cross sectional correlations illustrates the evolving role of trade-related variables.

In figures 5 and 6, we graph out the partial effects of the locational fundamentals (elevation, maize productivity, distance to coast, and distance to fresh water) and trade-related variables (distance to border, anchorage area, and Mexico City) on the probability of observing cities across space in different eras.[10] These maps show the role of these variables

---

[9] This GIS layer, part of the SUN database, was created in 2012 to catalog urban development within Mexico.

[10] The partial effects come from cross-sectional probit regressions on all of our covariates, where we extract the probability of city presence determined by these key groups of variables, setting all other variables in the regression to their mean values. Map shading is for quintiles within each map.

in the 1900 predicted probability of finding a city over 15,000 in a grid cell, and the ratio of this probability in 2010 to that of 1900.

The first correlation to note is that fundamentals variables generate significant variation in 1900 (Figure 5 (a)), and that geography in 1900 favored the center of the country. The change in the prediction of areas favored for urban development by fundamentals is highest on the coasts and in the Yucatan Peninsula. The maps of trade variables demonstrate an even larger change in correlations – in 1900, trade-related variables predict cities in the center and south of the country, as well as around anchorage areas. The change in correlation for this variables in 2010 shows a shift towards the north east. The power of the variables in the estimations also varies over time – removing the geographic variables from the 1900 regression lowers the pseudo-likelihood by 7.5 percent (relative to the unrestricted model), while removing these same variables from the 2010 regressions lowers the pseudo-likelihood by 4.9 percent. For the trade variables, the effect is the opposite: the change in the pseudo-likelihood between the unrestricted regression and one that limits the estimation to non-trade variables lowers the pseudo-likelihood by 4.6 percent in 1900, and by 8.3 percent in 2010. It appears that trade variables have more predictive power in 2010 than in 1900, while the opposite is true of locational fundamentals. Below, we examine when these changes occurred using a series of flexible panel regressions.

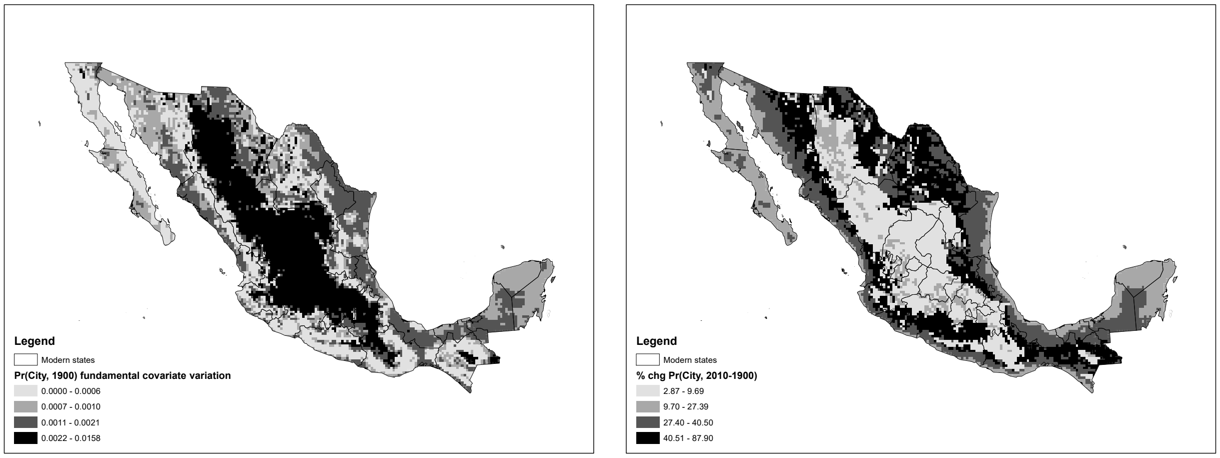
5.1 Estimation strategy

Our estimation intends to identify differential effects of key time-invariant characteristics. The main outcomes are indicators of achieving "city" status. We assign the value 1 to a cell when one or more localities in grid cell g in state s in decade t have passed one of the urbanization thresholds: a population greater than 5,000, 15,000, or 50,000. Through the census, we have data on urbanization every 10 years from 1900 to 2010.[11] In order to measure how covariate impact changes over the historical eras of interest, we interact these

---

[11] We report regressions including 1980 and 1921 even though there are well documented problems with census data in these years. Results are robust to excluding these data points (see Appendix tables).

Figure 5: Predicted cities in 1900 versus 2010 using fundamentals covariates

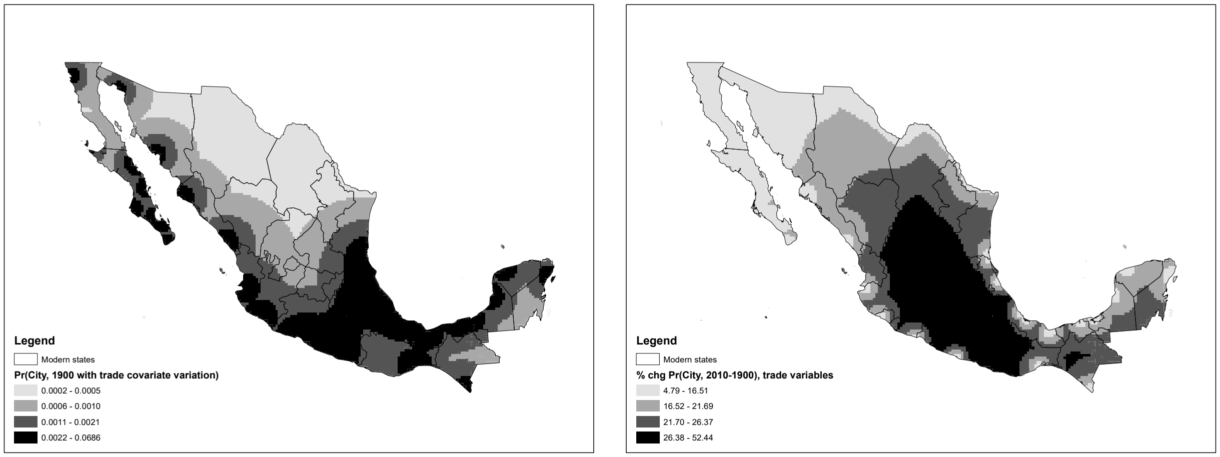


(a) 1900

(b) 1900 v 2010

Map (a) presents the predicted probability of a city in 1900, using only variation from elevation, maize productivity, distance to coast, and distance to fresh water. All other variables are set to their mean values. Map (b) presents the ratio of these same predictions in 2010 relative to 1900. Predictions come from the specifications in column (2) of Tables 2 and 3.

Figure 6: Predicted cities in 1900 versus 2010 using trade covariates



(a) 1900

(b) 1900 v 2010

Map (a) presents the predicted probability of a city in 1900, using only variation from distance to border, anchorage area and Mexico City. All other variables are set to their mean values. Map (b) presents the ratio of these same predictions in 2010 relative to 1900. Predictions come from the specifications in column (2) of Tables 2 and 3.

variables with dummies for the eras (γ_e): 1900, 1910-1940 (Revolution), 1950-1970 (ISI), and 1980-2010 (free trade). The coefficients estimated from these interactions indicate the change in the average effect for each of these periods relative to the omitted baseline of 1900.[12]. The $\mathbf{X}_{mg}$ matrix includes the natural log of distance to the US border, Mexico City, to the nearest water source, the coast, and anchorage point, tributary density in 1570, standard deviation of elevation, and maize productivity.

$$U_{gst} = \alpha_0 + \sum_m \sum_t \beta'_{mt} \mathbf{X}_{mg} \gamma_e + \gamma_g + \gamma_t + \epsilon_{gst} \qquad (2)$$

Fixed effects are included at the grid cell level g, and standard errors are clustered at the level of the state s. The reported results censor the dependent variable upon first urbanization. These regressions can therefore be thought of as explaining the likelihood of new city emergence for different urbanization thresholds. We also include, in some specifications, the lagged value of the presence of another city exceeding 15,000 inhabitants within a 50 km radius of the grid cell[13].

5.2 Results

The full set of regression results is available in the appendix. Figure 7 shows marginal effects of key covariates calculated for a one standard deviation change in the covariate. The impact of the standard deviation of elevation is statistically significant, but not shown due to the minuscule size of the effects. The coefficients should be interpreted relative to their impact on the probability of city location in 1900 – the omitted year. Recall that in 1900, the statistically significant determinants of city location were elevation, distance to anchorage area, and tributary density.

For cities that emerged during the 1910-1940 period – not a time of great urban growth –

---

[12] We have also estimated these interactions with decadal effects, and find them consistent with the results here. We prefer the era effects because they match more closely with key policy adjustments.

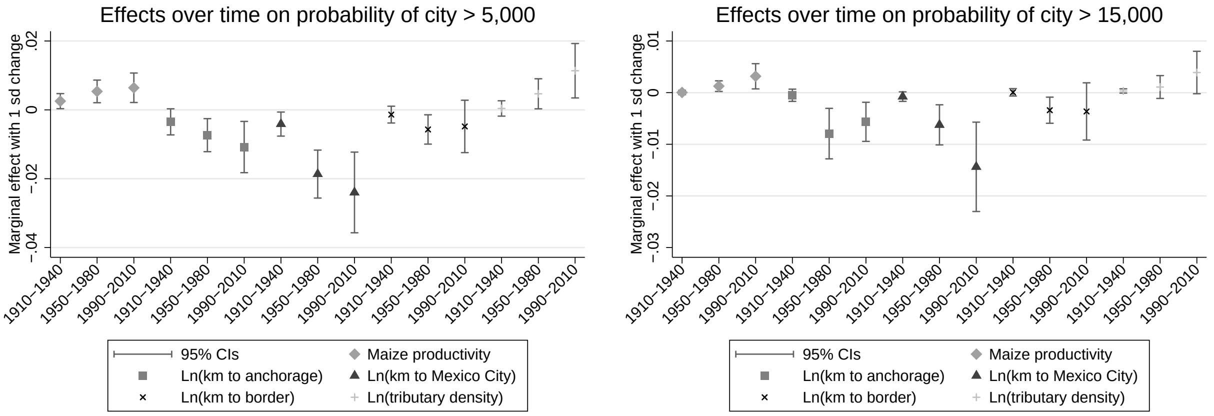
[13] We also tested for spatial dependence within a 25 km radius, and found no interesting differences in results.

determinants were largely that same as in 1900. This is illustrated by the lack of statistically significant effects of any of the covariates for this first period. However, starting in 1950, with the beginning of the ISI period and Mexico's era of urban growth, determinants of city emergence begin to change. In particular, distances begin to have stronger effect: cities of all sizes are less likely to emerge far from anchorage areas and Mexico City. It is unclear whether the latter effect is driven by the fact that Mexico City is by far the largest internal market or because of political pull, but this distance is clearly a dominant factor for the development of cities from 1950 onwards. Distance from borders only has a statistically significant impact on city development for cities larger than 50,000 inhabitants, and only relatively large in magnitude during the free trade era. Interestingly, while cities were more likely to cross the 50,000 threshold near the border during this period, overall population densities increased farther from the border, highlighting the importance of urban activity in trade.

Maize productivity has statistically significant and positive impacts on small city emergence, but these effects are relatively small in magnitude compared to the distance impacts. They do not drive in any way the emergence of large cities, probably because these cities are not constrained by local agricultural production. Historical population density significantly increases the probability of the emergence of a small city in the free trade era, and although the direction of the effects is similar for medium and large cities, they are not statistically significant. These results are the same whether or not we control for the role of neighboring cities.

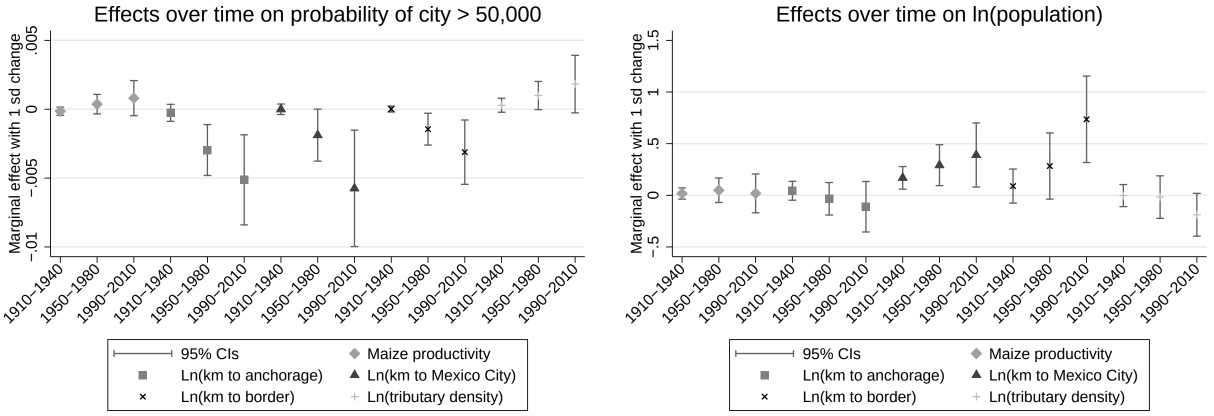
This section unpacks the timing of shift in the determinants of city location that was obscured by the cross-sectional regressions. The factors determining new city emergence during the ISI period were different from those in the 1900 in the sense that distance to Mexico City, the main market and political center of the country, became increasingly important. As trade opened, US border proximity began to pull towns over the highest urban threshold. Historical persistence, as measured by tributary density in 1570, gave certain

Figure 7: Evolving marginal effects on city emergence



(a) > 5,000

(b) > 15,000



(c) > 50,000

(d) Ln(AHL density)

Figures show the marginal effects on outcomes of a one standard deviation change in covariates. Estimates come from tables A5-A7, column (6). These are partial results.

areas an advantage in the development of more recent cities, though its impact was smaller in magnitude than the forces of trade-related geography.

6 Conclusion

In this paper we have tracked the emergence of Mexico's city system from 1900 to 2010. We began by looking spatially at the correlation between locational fundamentals, historical populations, and proxies for trade in cross-sections during the colonial period, and at the beginning and end of our data on the modern era: 1900 and 2010. Colonial era population show surprising persistence, even in the face of a population collapse estimated at 90%; by far the most substantial determinant of density in the mid and late colonial periods was population in 1570. The key determinants of population density in 1570 were those associated with proximity to the political center of the Aztec Empire, agricultural productivity, disease environment, and ability to defend against aggressors (variation in elevation). Even in the colonial era, however, trade began to influence population densities – by the eve of independence, higher populations were located in areas closer to anchorage points.

Historical populations sowed the seeds for urbanization: historical population density was deterministic of the first large cities in Mexico in 1900, and continued to exert pull as the century progressed. However, the forces of maritime trade – distance to anchorage – were even more important, as were the physical qualities of the location (elevation). By 2010, large cities were most likely to be found in agriculturally productive areas near Mexico City, close to anchorage areas and in flatter locations. However, the nature of the relationships between covariates and city location is quite different between the two periods. Geographic variables (elevation, maize productivity, distance to coast, and distance to fresh water) favored the center of the country in 1900. By 2010, these same variables predicted large increases in city presence along the coasts and in the south. By contrast, variables directly associated with trade (distance to the US border, Mexico City, and anchorage areas), predict cities around anchorage areas and in stripes across the middle of Mexico in 1900, whereas these same covariates predict significant increases in city presence in 2010 in the northeast. Furthermore, the power of these combinations of covariates also changes over time: the strictly geophysical variables have greater predictive power in 1900, while the distance variables are more predictive power in 2010. Our panel analysis highlights the timing of this shift. These estimations show that cities emerging in the revolutionary and post-revolutionary period (1910-1940) did so largely in similar places to where they existed in 1900. During the ISI period, the role of proximity Mexico City increased, while in the most recent liberalization era, proximity to US border began to pull cities over the largest urban threshold.

The story of Mexico's grand urbanization highlights the ever-shifting influence of history, geographic fundamentals, and trade. It encompasses the entire range of urbanization scenarios, and illustrates how the role of geography changes over time as policies shift. The evidence here also sheds light on the discussion about whether populations persist in an area because that place is always inherently geographically preferable (Davis & Weinstein 2002), or because once a population is settled, it helps resolve ambiguity about where future investment will locate (Krugman 1991, Bleakley & Lin 2012). The emergence of cities along the northern border and in the arid regions of the northeast provides important evidence for latter. It is notable that the collapse of the colonial population was not sufficient to disrupt the centripetal force of Tenochitlan. What was required was the transformation induced by improvements in transport technology combined with the opening up of trade with the north, thus decoupling urbanization from geographic destiny.

References

Ades, A. F. & Glaeser, E. L. (1995), ‘Trade and Circuses: Explaining Urban Giants’, *Quarterly Journal of Economics* **110**(1), 195–227.

Anas, A., Arnott, R. & Small, K. A. (1998), ‘Urban spatial structure’, *Journal of economic literature* **36**(3), 1426–1464.

Bairoch, P. (1988), *Cities and Economic Development. From the Dawn of History to the Present*, University of Chicago Press, Chicago, IL.

Bank, W. (1970), ‘Mexico - fourth road project (english)’, *Transportation projects series; no. PTR 47. Washington, D.C.* .

Bates, R. H. (1981), *Markets and States in Tropical Africa*, University of California Press, Berkeley, CA.

Beatty, E. (2001), *Institutions and Investment: the political basis of industrialization in Mexico before 1911*, Vol. 42, Stanford University Press Stanford, CA.

Behrens, K. (2007), ‘On the location and lock-in of cities: Geography vs transportation technology’, *Regional Science and Urban Economics* **37**(1), 22–45.

Bleakley, H. & Lin, J. (2012), ‘Portage and Path Dependence’, *Quarterly Journal of Economics* **127**(2), 587–644.

Bosker, M. & Buringh, E. (2015), ‘City seeds: geography and the origins of the European city system’, *Journal of Urban Economics* .

Bosker, M., Buringh, E. & van Zanden, J. L. (2013), ‘From Baghdad to London: Unraveling Urban Development in Europe, the Middle East, and North Africa, 800–1800’, *Review of Economics and Statistics* **95**(4), 1418–1437.

Coatsworth, J. H. (1978), ‘Obstacles to economic growth in nineteenth-century Mexico’, *The American Historical Review* **83**(1), 80–100.

Coatsworth, J. H. (1981), *Growth Against Development: The Economic Impact of Railroads in Porfirian Mexico*, Northern Illinois University Press, DeKalb, IL.

Coatsworth, J. H. & Williamson, J. G. (2004), ‘Always protectionist? Latin American tariffs from independence to Great Depression’, *Journal of Latin American Studies* **36**(02), 205–232.

Cook, S. F. & Borah, W. W. (1960), *The Indian population of central Mexico, 1531-1610*, Vol. 44, University of California Press.

Cook, S. F. & Borah, W. W. (1971), *Essays in population history: Mexico and the Caribbean*, Vol. 1, Univ of California Press.

Davis, D. R. & Weinstein, D. E. (2002), 'Bones, Bombs, and Break Points: The Geography of Economic Activity', *American Economic Review* **92**, 1269–1289.

De Vries, J. (1984), *European Urbanization, 1500-1800*, Vol. 4, Routledge.

Dincecco, M. & Onorato, M. G. (2016), 'Military conflict and the rise of urban Europe', *Journal of Economic Growth* **21**(3), 259–282.

Duranton, G. (1999), 'Distance, land, and proximity: economic analysis and the evolution of cities', *Environment and Planning A* **31**(12), 2169–2188.

Ficker, S. K. (2004), 'The export boom of the Mexican revolution: Characteristics and contributing factors', *Journal of Latin American Studies* **36**(02), 267–296.

Fujita, M. & Mori, T. (1996), 'The role of ports in the making of major cities: self-agglomeration and hub-effect', *Journal of Development Economics* **49**(1), 93–120.

Fujita, M. & Mori, T. (1997), 'Structural stability and evolution of urban systems', *Regional science and urban economics* **27**(4), 399–442.

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

Gerhard, P. (1993*b*), *The North Frontier of New Spain*, University of Oklahoma Press, Norman, OK.

Gerhard, P. (1993*c*), *The Southeast Frontier of New Spain*, University of Oklahoma Press, Norman, OK.

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

Glaeser, E. L. (2007), The economics approach to cities. NBER Working Paper 13696.

Glaeser, E. L. (2014), 'A World of Cities: The Causes and Consequences of Urbanization in Poorer Countries', *Journal of the European Economic Association* **12**(5), 1154–1199.

González Navarro, M. (1974), *Población y sociedad en México (1900–1970)*, Universidad Nacional Autónoma de México, México, D.F.

Haber, S. H. (1992), 'Assessing the Obstacles to Industrialisation: The Mexican Economy, 1830–1940', *Journal of Latin American Studies* **24**(01), 1–32.

Haber, S., Klein, H. S., Maurer, N. & Middlebrook, K. J. (2008), *Mexico Since 1980*, Cambridge Univrsity Press, New York.

Haber, S., Razo, A. & Maurer, N. (2003), *The Politics of Property Rights: Political Instability, Credible Commitments, and Economic Growth in Mexico, 1876–1929*, Cambridge University Press, New York, NY.

Harris, J. R. & Todaro, M. P. (1970), 'Migration, Unemployment, and Development: A Two-Sector Analysis', *American Economic Review* **60**(1), 126–142.

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

Henderson, J. V. (2005), 'Urbanization and growth', *Handbook of economic growth* **1**, 1543–1591.

Henderson, J. V., Squires, T. L., Storeygard, A. & Weil, D. N. (2016), The Global Spatial Distribution of Economic Activity: Nature, History, and the Role of Trade. National Bureau of Economic Reserarch Working Paper 22145.

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

Jedwab, R., Christiaensen, L., Gindelsky, M. et al. (2014), 'Rural Push, Urban Pull and... Urban Push? New Historical Evidence from Developing Countries', *The George Washington University, Institute for International Economic Policy Working Papers* **4**.

Jedwab, R., Kerby, E. & Moradi, A. (2015), 'History, path dependence and development: Evidence from colonial railroads, settlers and cities in Kenya', *The Economic Journal* .

Knight, A. (1991), 'Popular Culture and the Revolutionary State in Mexico', *The Hispanic American Historical Review* **74**(3), 393–444.

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

Konishi, H. (2000), 'Formation of hub cities: transportation cost advantage and population agglomeration', *Journal of Urban Economics* **48**(1), 1–28.

Krugman, P. (1991), 'History Verus Expectations', *Quarterly Journal of Economics* **106**(2), 651–667.

Krugman, P. (1993), 'On the number and location of cities', *European Economic Review* **37**(2-3), 293–298.

Krugman, P. & Elizondo, R. L. (1996), 'Trade policy and the third world metropolis', *Journal of Development Economics* **49**(1), 137–150.

Mellinger, A., Sachs, J. D. & Gallup, J. L. (2000), Climate, coastal proximity, and development, *in* 'The Oxford Handbook of Economic Geography,', Oxford University Press, New York, NY.

Moreno-Bird, J. C., Nápoles, P. R. & Valdivia, J. C. (2004), 'NAFTA and the Mexican Economy: A Look Back on a Ten-Year Relationship', *NCJ Int'l L. & Com. Reg.* **30**, 997.

Moreno-Brid, J. C., Santamaria, J. & Rivas Valdivia, J. C. (2005), 'Industrialization and economic growth in Mexico after NAFTA: the road travelled', *Development and Change* **36**(6), 1095–1119.

NOAA (2014), 'Anchorage areas'. NOAA Office for Coastal Management, US Government Publishing Office.

Nunn, N. & Puga, D. (2012), 'Ruggedness: The Blessing of Bad Geography in Africa', *The Review of Economics and Statistics* **94**(1), 20–36.

Overman, H. G. & Venables, A. J. (2005), *Cities in the developing world*, Centre for Economic Performance, London School of Economics and Political Science.

Pirenne, H. (1925), *Medieval cities*, Doubleday.

Redding, S. J. & Sturm, D. M. (2008), 'The costs of remoteness: Evidence from German division and reunification', *The American Economic Review* **98**(5), 1766–1797.

Redding, S. J., Sturm, D. M. & Wolf, N. (2011), 'History and industry location: Evidence from German airports', *Review of Economics and Statistics* **93**(3), 814–831.

Reynolds, C. W. (1970), *The Mexican economy: twentieth-century structure and growth*, Yale University Press.

Sellars, E. & Alix-Garcia, J. (2017), 'Labor Scarcity, Land Tenure, and Historical Legacy: Evidence from Mexico'. Working paper, Oregon State University and Texas A&M.

Sobrino, J. (2010), *Migración interna en México durante el siglo XX*, Consejo Nacional de Población.

Tilly, C. (1989), 'Cities and States in Europe, 1000-1800', *Theory and Society* **18**(5), 563–584.

Unikel, L. (1976), *El desarrollo urbano de México: diagnóstico e implicaciones futuras*, El Colegio de México, México, DF.

United States Central Intelligence Agency (1990), *The World Factbook (1990)/Mexico*, CIA.

Warman, A. (2001), *El campo mexicano en el siglo XX*, Fondo de Cultura Económica.

Weber, M., Martindale, D. & Neuwirth, G. (1958), *The City*, Free Press New York.

World Bank (2016), 'World development indicators'.

Appendix A Further description of AHL data: for online publication only

The AHL omits some small localities recorded in earlier censuses that had ceased to exist prior to the development of the geographical code system in 1978. In some states (such as Guerrero, México, Tamaulipas, and Yucatán), all localities that shift municipal boundaries prior to 1978 are omitted in the early years of the dataset, though nearly all of the omitted communities are rural. The central *delegaciones* of Mexico City are also omitted from the AHL prior to the administrative reform of 1970.

We edit and supplement the AHL data in several ways. First, we supplement missing population data using information from Unikel (1976), which contains lists of urban centers over time and complete population counts for central Mexico City, and information from the *División Territorial* which contains a list of localities in the country as of each census. Data on city placement in 1878 are also taken from Unikel. Where geographic data are missing (about 15% of localities in the dataset), we use information on the history of a locality from the AHL's "Movements" file to link it to the location of a contemporary settlement. We do this primarily by using text analysis to identify alternative geographic identifiers that have been connected with the community and connecting these identifiers to a geographic location today. For example, a locality's identifier changes when it switches between municipalities, is absorbed into a larger locality, or when it breaks off from a parent locality. For the small number of historical localities for which we were unable to identify a location, we place the locality at the position of the *cabecera municipal*, the seat of local government. These localities are small (usually less than 100 in population) rural areas that had been depopulated prior to 1990, and are not used in the majority of our empirical analysis.

Appendix B Additional tables: for online publication only

Table B1: Key features of eras in Mexican history

Year	Era	Population[a]	Trade	Transport
1520	Colonial era	20 m falls to 1.5 m by 1650 2.5 m	Crown centralizes control of trade	*Tlameme*[b]
1821	Independence; Wars with Spain (1829), US (1846–48), France (1862-67)	6 m	First tariff law: 25 ad valorem tax on all imports; ratio of tariffs to imports: 36% [c]	Single road capable of supported wheeled traffic (Mexico City to Chihuahua, via Zacatecas and Durango)[d]
1876			Ratio of tariffs to imports: 46%	1873: First railroad: Mexico City to Veracruz; < 5 km of road passable by four wheeled carts per 10,000 inhabitants[e]
1877	Porfiriato	10.4 (1884); 12.6 m (1895)	2% commodity exports/GDP[f] 1900: 3% commodity exports/GDP	1877: 640 km rr (mule-driven);
1911		11 cities over 50,000 in 1900	3% commity exports/GDP	> 5,000 km rr, freight transport costs fall by 2/3rds
1910	Revolution & restructuring	15.2 m	4% commodity exports/GDP	Railroad destruction
			% exports/GDP: 11.0 in 1930[g]	Total roads: 91,609 km; "good" roads: 17,670 in 1930
1940		19.7 m	% exports/GDP: 6.7	Additional road construction 1935-1940[h]: 3,694 km
1950	ISI	25.8 m	% exports/GDP: 9.9	13,600 km paved road
1970		48.2 m	% exports/GDP: 7.7	1968: 67,000 km roads[i]
1980	Economic restructuring	66.8 m	% exports/GDP: 10.7	14,225 km rr; 1990: 210,000 km roads; 65,000 km paved[j]
2010		112.3 m	% exports/GDP: 29.9	26,704 km rr[k]; 366,095 km roads, 72,577 km paved

---

[a] Population sources: colonial: Gerhard (1993*a*); 1821: Coatsworth (1978); 1910-2010: INEGI

[b] Source: Hassig (1985)

[c] Source: Coatsworth & Williamson (2004)

[d] Source: Haber (1992)

[e] Source: Beatty (2001)

[f] These calculations use value of commodity exports from Ficker (2004) and GDP values from Mitchell. I am having trouble reconciling exchange rates for this period, so these are likely to be wrong.

[g] Source for 1930-2010: World Bank (2016)

[h] Source: Reynolds (1970)

[i] Source: Bank (1970)

[j] Source: United States Central Intelligence Agency (1990)

[k] Source: World Bank (2016)

Table B2: Summary Statistics

	mean	sd	max	min
Population density, 1900	125.02	839.78	38505.43	0.00
% mun pop in cities > 5000	0.16	0.28	1.00	0.00
% mun pop in cities > 2500 & < 5000	0.11	0.21	1.00	0.00
% mun pop in cities < 2500	0.73	0.32	1.00	0.00
Total population 1900	8796.50	14572.10	408962.00	181.00
Km to US border	682.23	270.37	1329.89	0.00
Km to DF	424.91	386.85	1758.38	0.00
Mean maize	0.81	0.65	3.56	0.00
Km2, 1900 municipalities	1016.53	2411.73	69083.00	3.00
Observations	18588			

Table B3: Correlates of urban grid cells in 1900, urban cutoff of 5,000

	Dep. var: Has locality of $>$ 15,000 population = 1								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Elevation $<$ 1500 m	-0.016***	-0.018***	-0.014**	-0.012**	-0.021**	-0.019***	-0.014**	-0.018**	-0.019**
	(0.006)	(0.007)	(0.006)	(0.006)	(0.008)	(0.007)	(0.007)	(0.008)	(0.008)
Std grid cell elevation	-0.000*	-0.000**	-0.000	-0.000	-0.000*	-0.000	-0.000	-0.000	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(km coast)	0.007***	0.010***	0.008***	0.008***	0.013***	0.010***	0.010***	0.009***	0.012***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)
Ln(km water)	-0.000	-0.000	0.001	0.001	-0.000	0.001	0.001	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
Ln(maize productivity)	0.008**	0.005	0.008	0.005	0.009**	0.014***	0.008*	0.015***	0.008**
	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Ln(Km to border)	-0.001	0.006	0.005	0.005	0.012**	0.012**	0.010**	0.004	0.011**
	(0.003)	(0.005)	(0.004)	(0.004)	(0.006)	(0.006)	(0.004)	(0.003)	(0.005)
Ln(Km to DF)	-0.013***	-0.010***	-0.010***	-0.008***					
	(0.003)	(0.003)	(0.003)	(0.002)					
Ln(km anchorage)	-0.013***	-0.015***	-0.014***	-0.014***	-0.018***	-0.015***	-0.016***	-0.013***	-0.018***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Ln(Tributary density, 1570)		0.003		0.005***	0.007***		0.009***		0.007***
		(0.002)		(0.002)	(0.002)		(0.003)		(0.002)
Ln(Tributary density, 1650)			0.001	-0.001		0.004	-0.002		
			(0.002)	(0.002)		(0.003)	(0.003)		
City of 15000, t-1, 50 km								0.010*	0.007
								(0.006)	(0.006)
Observations	9342	6491	5842	5736	6491	5842	5736	9342	6491
Adjusted R^2									

Unit of observation is the 15 x 15 km grid cell. Clustered standard errors are in parentheses. * $p< 0.10$, ** $p<0.05$, *** $p < 0.01$. Coefficients robust to excluding Mexico City area. Reported coefficients are marginal effects from a probit.

Table B4: Correlates of urban grid cells in 2010, urban cutoff of 50,000

	Dep. var: Has locality of $>$ 15,000 population = 1								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Elevation $<$ 1500 m	-0.018***	-0.017***	-0.016***	-0.015**	-0.020**	-0.021**	-0.018**	-0.023***	-0.021***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.008)	(0.009)	(0.008)	(0.008)	(0.007)
Std grid cell elevation	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000**	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(km coast)	0.004*	0.005*	0.004*	0.004*	0.007**	0.005*	0.006*	0.005**	0.006*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)
Ln(km water)	-0.003**	-0.004**	-0.003**	-0.003**	-0.003**	-0.003*	-0.003*	-0.003**	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Ln(maize productivity)	0.010**	0.007*	0.006	0.005	0.011**	0.014***	0.010**	0.018***	0.010**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)
Ln(Km to border)	-0.011***	-0.006***	-0.006***	-0.006***	-0.000	0.001	-0.001	-0.007***	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.004)	(0.003)	(0.002)	(0.002)
Ln(Km to DF)	-0.017***	-0.014***	-0.015***	-0.014***					
	(0.002)	(0.003)	(0.003)	(0.003)					
Ln(km anchorage)	-0.015***	-0.015***	-0.014***	-0.013***	-0.017***	-0.016***	-0.016***	-0.014***	-0.016***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.002)	(0.003)
Ln(Tributary density, 1570)		0.003*		0.001	0.009***		0.008**		0.008***
		(0.002)		(0.003)	(0.003)		(0.004)		(0.003)
Ln(Tributary density, 1650)			0.002	0.002		0.006*	0.001		
			(0.002)	(0.003)		(0.004)	(0.003)		
City of 15000, t-1, 50 km								0.021***	0.015**
								(0.005)	(0.006)
Observations	9342	6491	5842	5736	6491	5842	5736	9342	6491
Adjusted R^2									

Unit of observation is the 15 x 15 km grid cell. Clustered standard errors are in parentheses. * p< 0.10, ** p<0.05, *** p < 0.01. Coefficients robust to excluding Mexico City area. Reported coefficients are marginal effects from a probit.

Table B5: Correlates of cities: population > 50,000

	(1)	(2)	(3)	(4)	(5)	(6)
City of 15000, t-1, 50 km			-0.002***		-0.002**	
			(0.001)		(0.001)	
City of 50000, t-1, 50 km				-0.001		
				(0.001)		
Ln(km to border), 1910-1940	-0.000	0.000	-0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(km to border), 1950-1970	-0.002***	-0.001**	-0.002***	-0.002***	-0.001**	-0.001**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(km to border), 1980-2010	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(km to Mex City), 1910-1940	-0.000	0.000	-0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(km to Mex City), 1950-1970	-0.003***	-0.002**	-0.003***	-0.003***	-0.002**	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(km to Mex City), 1980-2010	-0.008***	-0.008***	-0.008***	-0.008***	-0.008***	-0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)
Ln(km to anchorage), 1910-1940	-0.000	0.000	-0.000	-0.000	0.000	-0.000
	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
Ln(km to anchorage), 1950-1970	-0.003***	-0.003**	-0.003***	-0.003***	-0.003**	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(km to anchorage), 1980-2010	-0.005***	-0.004***	-0.005***	-0.005***	-0.005***	-0.005***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Maize productivity, 1910-1940	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Maize productivity, 1950-1970	0.001	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Maize productivity, 1980-2010	0.002**	0.001	0.002**	0.002**	0.001	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(km to coast), 1910-1940	0.000	-0.000	0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(km to coast), 1950-1970	0.001*	0.001**	0.001*	0.001*	0.001**	0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(km to coast), 1980-2010	0.001*	0.002**	0.001*	0.001*	0.002**	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(km to water), 1910-1940	0.000	0.000	0.000	0.000	0.000	0.000

	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(km to water), 1950-1970	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(km to water), 1980-2010	-0.000*	-0.000	-0.001**	-0.000*	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
SD elevation, 1910-1940	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
SD elevation, 1950-1970	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
SD elevation, 1980-2010	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1570 trib density), 1910-1940		0.000			0.000	0.000
		(0.000)			(0.000)	(0.000)
Ln(1570 trib density), 1950-1970		0.001*			0.001**	0.001*
		(0.000)			(0.000)	(0.000)
Ln(1570 trib density), 1980-2010		0.001**			0.001**	0.001*
		(0.001)			(0.001)	(0.001)
Observations	111499	77525	111499	111499	77525	64592
Adjusted R^2	0.008	0.010	0.008	0.008	0.010	0.009

Unit of observation is the grid cell, 1900-2010. Robust standard errors clustered at state level are in parentheses. Outcome is a binary variable equal to one when the size of the city exceeds 50,000. Column (6) excludes 1921 and 1980 censuses. Marginal effects should be interpreted as differences from 1900, the excluded year.

Table B6: Correlates of cities: population > 15,000

	(1)	(2)	(3)	(4)	(5)	(6)
City of 15000, t-1, 50 km			-0.002		-0.002	
			(0.002)		(0.003)	
City of 50000, t-1, 50 km				0.001		
				(0.002)		
Ln(km to border), 1910-1940	-0.000	-0.000	-0.001	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(km to border), 1950-1970	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(km to border), 1980-2010	-0.004*	-0.004*	-0.004*	-0.004*	-0.004*	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Ln(km to Mex City), 1910-1940	-0.001	-0.001	-0.001	-0.001	-0.000	-0.001*
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(km to Mex City), 1950-1970	-0.007***	-0.007***	-0.007***	-0.007***	-0.007***	-0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Ln(km to Mex City), 1980-2010	-0.017***	-0.019***	-0.017***	-0.017***	-0.019***	-0.018***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Ln(km to anchorage), 1910-1940	-0.000	-0.001	-0.000	-0.000	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(km to anchorage), 1950-1970	-0.006***	-0.008***	-0.006***	-0.006***	-0.008***	-0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Ln(km to anchorage), 1980-2010	-0.007***	-0.007***	-0.007***	-0.006***	-0.007***	-0.006***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Maize productivity, 1910-1940	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Maize productivity, 1950-1970	0.003***	0.003**	0.003***	0.003***	0.003**	0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Maize productivity, 1980-2010	0.007***	0.006**	0.007***	0.007***	0.006**	0.007**
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)
Ln(km to coast), 1910-1940	0.000	0.000	0.000	0.000	0.000	0.001*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(km to coast), 1950-1970	0.002**	0.003**	0.002**	0.002**	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(km to coast), 1980-2010	0.001	0.003**	0.001	0.001	0.003**	0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(km to water), 1910-1940	0.000	0.000	-0.000	0.000	0.000	-0.000

	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(km to water), 1950-1970	-0.001**	-0.001	-0.001**	-0.001**	-0.001	-0.001
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
Ln(km to water), 1980-2010	-0.001	-0.000	-0.001	-0.001	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
SD elevation, 1910-1940	-0.000	0.000	-0.000	-0.000	0.000	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
SD elevation, 1950-1970	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
SD elevation, 1980-2010	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1570 trib density), 1910-1940		0.000			0.000	0.000
		(0.000)			(0.000)	(0.000)
Ln(1570 trib density), 1950-1970		0.001			0.001	0.001
		(0.001)			(0.001)	(0.001)
Ln(1570 trib density), 1980-2010		0.003**			0.003**	0.003*
		(0.001)			(0.001)	(0.002)
Observations	110396	76590	110396	110396	76590	63777
Adjusted R^2	0.018	0.023	0.018	0.018	0.023	0.020

Unit of observation is the grid cell, 1900-2010. Robust standard errors clustered at state level are in parentheses. Outcome is a binary variable equal to one when the size of the city exceeds 15,000. Column (6) excludes 1921 and 1980 censuses. Marginal effects should be interpreted as differences from 1900, the excluded year.

Table B7: Correlates of cities: population > 5,000

}	(1)	(2)	(3)	(4)	(5)	(6)
City of 15000, t-1, 50 km			0.004**		0.004*	
			(0.002)		(0.002)	
City of 50000, t-1, 50 km				0.012***		
				(0.003)		
Ln(km to border), 1910-1940	-0.000	-0.000	-0.000	-0.000	-0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(km to border), 1950-1970	-0.004**	-0.005**	-0.004**	-0.004**	-0.004**	-0.005**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Ln(km to border), 1980-2010	-0.004	-0.005	-0.004	-0.003	-0.005	-0.005
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)
Ln(km to Mex City), 1910-1940	-0.003**	-0.003**	-0.003**	-0.003**	-0.003**	-0.005**
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Ln(km to Mex City), 1950-1970	-0.018***	-0.021***	-0.018***	-0.018***	-0.021***	-0.023***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Ln(km to Mex City), 1980-2010	-0.029***	-0.030***	-0.029***	-0.028***	-0.030***	-0.030***
	(0.006)	(0.007)	(0.006)	(0.006)	(0.007)	(0.007)
Ln(km to anchorage), 1910-1940	-0.001	-0.001	-0.001	-0.001	-0.001	-0.004*
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Ln(km to anchorage), 1950-1970	-0.006***	-0.006**	-0.006***	-0.006***	-0.005**	-0.008***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Ln(km to anchorage), 1980-2010	-0.013***	-0.013***	-0.013***	-0.012***	-0.013***	-0.011***
	(0.002)	(0.004)	(0.002)	(0.002)	(0.004)	(0.004)
Maize productivity, 1910-1940	0.003*	0.003*	0.003**	0.003**	0.003*	0.006**
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Maize productivity, 1950-1970	0.011***	0.009***	0.011***	0.011***	0.009***	0.012***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Maize productivity, 1980-2010	0.015***	0.013***	0.015***	0.015***	0.013***	0.014***
	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)
Ln(km to coast), 1910-1940	0.001	0.001	0.001	0.001	0.001	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(km to coast), 1950-1970	0.002	0.003*	0.002	0.002	0.003*	0.004**
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Ln(km to coast), 1980-2010	0.002	0.003	0.002	0.002	0.003	0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Ln(km to water), 1910-1940	-0.001	-0.000	-0.001	-0.001	-0.000	-0.000

	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
Ln(km to water), 1950-1970	-0.002	-0.001	-0.001	-0.001	-0.001	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(km to water), 1980-2010	-0.001	-0.000	-0.001	-0.001	-0.000	-0.000
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
SD elevation, 1910-1940	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
SD elevation, 1950-1970	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
SD elevation, 1980-2010	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1570 trib density), 1910-1940		0.000			0.000	0.000
		(0.001)			(0.001)	(0.001)
Ln(1570 trib density), 1950-1970		0.004**			0.003**	0.004**
		(0.001)			(0.002)	(0.002)
Ln(1570 trib density), 1980-2010		0.009***			0.009***	0.009***
		(0.003)			(0.003)	(0.003)
Observations	106723	73650	106723	106723	73650	61286
Adjusted R^2	0.029	0.035	0.029	0.030	0.035	0.032

Unit of observation is the grid cell, 1900-2010. Robust standard errors clustered at state level are in parentheses. Outcome is a binary variable equal to one when the size of the city exceeds 5,000. Column (6) excludes 1921 and 1980 censuses. Marginal effects should be interpreted as differences from 1900, the excluded year.