European Settlements, Settler Mortality, and the Uneven Development of the Americas

Javier A. Birchenall*
University of California at Santa Barbara

November 12, 2024

Abstract

I analyze early European settlement patterns and the mortality of Protestant and Catholic European clergy arriving in the New World from the 16th to the 20th centuries. Settler mortality estimates reveal significant temporal and spatial heterogeneity, with favorable disease environments in Spanish America early on giving way to favorable disease environments in some areas of British North America by the 18th century. I argue that the economic reversal between once-rich tropical and densely populated areas to now-rich temperate areas partly reflects Europe's transformation of the New World's first-nature geography, including a worsening of disease ecologies and a reversal in factor endowments.

Keywords: Disease, Economic Development, Settler Mortality
JEL classification: N10, O11, P51
Communications to: Javier A. Birchenall
Department of Economics, 2127 North Hall
University of California, Santa Barbara CA 93106
Javier.Birchenall@ucsb.edu

*I am grateful to Andres Alvarez and the economic history group at Universidad de los Andes (Bogotá); Martin Fiszbein, Felipe Valencia-Caicedo, and to the participants at the RIDGE Economic History meeting; and to Greg Casey, Juan Cobo Betancourt, David Cuberes, Jesus Fernandez-Villaverde, Maria del Carmen Pérez Artés, and Rodrigo Soares for helpful comments. I am indebted to numerous Gretler fellows for their help with this project.

1 Introduction

This paper studies the settlement patterns of European colonies in the New World, and presents new estimates of settler mortality rates drawn from the life histories of European missionaries and clergy arriving in the New World between the 16th and the 20th centuries. I employ these estimates to shed light on the potential of historical factors in explaining the uneven development paths in the Americas. Work on the historical origins of economic development has long recognized the significance of the location, objectives, and scale of European settlements in shaping New World development trajectories. The pioneer work of Acemoglu et al. [1] highlighted the role of disease environments in determining settlements patterns and the incentives of European colonizers toward "extractive" or "inclusive" institutions. The objective of this paper is to better understand the various mechanisms that connect settler mortality rates and settlement patterns.

The settlement of Spanish America and British North America differed in their timing, location, and scale. Instead of focusing solely on the role of settler mortality as a determinant of settlement patterns, I examine how European settlements themselves shaped disease ecologies and factor endowments, often regarded as invariant *first-nature geography*. I argue that it is unreasonable to treat the geography of ex-colonies as outside the realm of European influence. Europeans forced two major and related transformations upon the disease ecology of the New World. The first transformation began with Columbus' arrival and the introduction of European diseases to previously unexposed (e.g., virgin) populations. The second transformation began with the systematic arrival of African slaves and the introduction of diseases from the African pool that thrived in the tropical and semitropical climates of the Americas, but could not have thrived in temperate climates. These transformations significantly impacted native and settler mortality. Because settler mortality estimates at a specific time and place may provide a partial view of Europe's experience, thinking about the difficulties of establishing European settlements requires thinking about disease ecologies and changes in first-nature geography over time and space.[1]

[1] The timing and identity of different colonizers are closely related. There are no early British settlements in the New World or late Spanish settlements in areas comparable to those settled by the British. Only the Caribbean islands come close to providing a case study for the influence of Spanish, English, French, Dutch, and Danish settlements in similar geographic settings and times (Feyrer and Sacerdote [29]).

There are no comprehensive sources to examine early settlements patterns and settler mortality across regions. Acemoglu et al. [1] relied on mortality estimates from the life histories of Catholic bishops (Gutierrez [40]), and the mortality of American, British, and French soldiers across the globe in the 19th century (Curtin [17]). I mainly examine the life histories of Protestant and Catholic missionaries and bishops serving across the world from the 16th century to the 20th centuries. Because the mortality estimates rely on several secondary sources, I will focus on the main implications of the new settler mortality estimates rather than on the details about the sources, which are given in the Appendix.

The first lesson of the paper is that the high settler mortality rates of tropical America in the 19th century were a colonial legacy rather than an invariant feature of the geography of the New World. In the early stages of colonization, Spain faced a disease ecology that favored European settlements: Spanish settlers had immunity to European diseases and strategically favored areas with large native populations, often located in healthier regions. Although data about early settlements is limited (Hardoy and Aranovich [42]), the information provided by Spanish chroniclers in the 1580s (López de Velasco [75]) and the 1630s (Vasquez de Espinosa [27]) does not suggests that *vecinos* (i.e., Spanish settlers) were discouraged by mortality rates. Spanish settlers followed an "invasive species" pattern, dispersing randomly over time and space. Eventually, settlements reached a size distribution consistent with Zipf's Law, where settlers attached to areas with large native populations and mineral deposits. In the most populous cities in Spanish America, the number of European settlers doubled every 25 years. This growth rate is comparable to the growth rate of American cities during the age of mass migration.

Europe's encounter with the African disease pool did result in high settler mortality (McNeill [54]), but these encounters occurred when Spanish America was already settled. African diseases mostly affected the settlement patterns of late European comers.[2] Yellow fever and falciparum malaria, the diseases that defeated the British invasion of Cartagena

[2] Conflict was also a cause of mortality for European settlers, but Europeans had a military advantage over native populations. The war with the *Araucanians*, the indigenous Mapuches of Chile and Argentina, was the only long-lasting conflict in Spanish America. In North America, the *American Indian Wars* spanned over several centuries and were more deadly for European settlers than some conflicts in Spanish America partly because native North American tribes had acquired horses and firearms as early as the 16th century from strategic alliances with European countries (Steele [71]).

in 1741 and the French construction of the Panama Canal in 1881-1888, did not originate in the tropics of the New World and were late additions to America's disease pool (Diamond and Panosian [19]). The late confluence of disease pools created unfavorable disease environments in areas where diseases of African origin would become endemic, and even in some areas of North America where these diseases visited in epidemic cycles until the 18th century (Duffie [22]). In British North America, I find that missionaries faced unfavorable conditions in virtually all colonies, with the exception of New England, which attracted very few European settlers after 1650 and had steady out-migration for decades. Protestant missionaries in the Middle and Southern colonies, for example, experienced a four-fold higher risk of death per year of service relative to missionaries in New England.

Over time, there was a complete reversal in the mortality rates of Protestant missionaries in British North America and in the tropical regions of the New World. Early arrivals in British North America experienced unfavorable mortality conditions similar to those faced by late arrivals in Central America, while early arrivals in Central America enjoyed favorable mortality conditions comparable to those of late arrivals in British North America. This reversal is also evident in the mortality rates of Catholic bishops. Catholic bishops arrived late to British North America, however. In the early years, the mortality of Catholic bishops in Spanish America was indistinguishable from that of bishops in other parts of the world, except in Africa, where the risk of death per year of service was twice the risk in Europe.

The relatively high mortality rates faced by early settlers in British North America did not disappear entirely. Like Portugal's settlement of Brazil, British settlements in North America were overwhelmingly coastal. After the American War of Independence and the Louisiana purchase, settlements moved to the interior of the country. The disease environment in the frontier areas of settlement during the Western Expansion of the early 19th century was also unfavorable. In the Southern US, soldiers faced disease and mortality conditions comparable to those in the west coast of Africa due to constant presence of yellow fever (Lawson [47]). For instance, the latitude-mortality gradient *within* US counties in the early 19th century was larger than the latitude gradient of mortality *across* countries today (Acemoglu et al. [1]). Using population density from the US decennial censuses since 1790 and the mortality rates from frontier military forts in Lawson [47], the source used

by Curtin [17], suggests a positive association between settler mortality rates and frontier settlements along the Western Expansion. The second lesson of the paper is that neither early nor late frontier settlements in British North America actively sought to avoid areas with unfavorable disease environments.

These lessons suggest that not only were economic fortunes reversed during colonial times, but disease environments also experienced drastic transformations. To account for the reversal of fortunes among countries colonized by European powers, I consider a *sophisticated geography hypothesis*. Attributing economic reversals to first-nature geographical factors in the form of disease environments and factor endowments is difficult if these factors are seen as invariant or unchangeable (Acemoglu et al. [2]). Disease environments and factor endowments, however, underwent significant changes in the Americas as a consequence of colonial activity. European diseases devastated native populations leading to declines of up to 75-90 percent in some areas. Native populations in the Caribbean, the first area settled by Europeans, were exterminated within a few years and unable to recover at all (Newson [58]). In extractive economies, depopulation had contractionary effects with labor shortages incentivizing the forced migration of African workers and their diseases. This global confluence of disease pools further worsened disease ecologies in tropical America and limited population growth relative to the temperate areas of North America where tropical diseases could not become endemic.

Related literature. Settler mortality rates from Gutierrez [40] and Curtin [17] are central to the *colonial origin* hypothesis proposed by Acemoglu et al. [1]. They have shown that settler mortality and the feasibility of settlements shaped colonial strategies and current economic performance. Their work has had a transformative effect on the study of institutional and historical legacies in economic development. It is impossible to synthesize the many emerging literatures that study European settlements and settler mortality.[3] I build upon Acemoglu et al. [1] by focusing on two aspects: the settlements patterns at inception and the changes in settler mortality over time.

The new settler mortality estimates are very different from the estimates in Gutierrez

[3] A very selected list includes Acemoglu et al. [3], Droller [21], Easterly and Levine ([23], [24]), Glaeser et al. [35], Gorodnichenko et al. [37], Hall and Jones [41], Maloney and Valencia-Caicedo [51], Putterman and Weil [63], Sequeira et al. [65], and Sokoloff and Engerman [70].

[40] and Curtin [17].[4] Gutierrez [40] considered Catholic bishops but only in Spanish America, while Curtin [17] focused on military mortality but only in the late colonial era, the 19th century. Longitudinal data and a wider regional coverage are needed for a comparative perspective. I place special emphasis on religious organizations with a long-standing presence and a broad geographic reach in the New World. The new estimates imply a different view of disease environments and settlement patterns in colonial times. They show that Spain succeeded in establishing settlements throughout wide geographic areas, although typically concentrated in dense enclaves with favorable disease environments, large native populations, and mineral deposits. British settlements, in contrast, were confined to coastal areas. In British North America, while New England offered a favorable disease environments, this region attracted fewer settlers than the Middle and Southern colonies, whose settler mortality rates far exceeded mortality in New England.

While I derive new settler mortality estimates, I will not use the new estimates to examine the role of historical legacies on long-run development. This decision is rooted in the belief that the economic possibilities of former colonies should not be studied solely through the lens of European agency. There is an inherit Eurocentrism in some of the existing literature that emphasizes the tragic history of European migration to the tropical world, while elevating the institutions, human capital, ideas, and culture that Europeans brought or failed to bring to their colonies at the expense of the impact of colonization on native people and their economies. Most of the literature that studies historical legacies maintains a passive view of native populations. Assessing causality in historical episodes also requires information that is often unavailable and exclusion restrictions that are difficult to verify. Instrumental variable strategies, for instance, perform poorly in small samples and cross-country data allows for no more than a handful of ex-colonies to be considered at a point in time.

[4] Existing settler mortality data mixes nonhomogeneous populations and uses data imputations due to limited coverage. Some of these limitations have been pointed out by Albouy [5], and considered by Acemoglu et al. [4], but data improvements have mainly amounted to a reshuffling of existing sources. Subramanian [72], following a suggestion by Simon Johnson, used mortality estimates from the British insurance industry. Mortality estimates from insurance data are heavily concentrated in the late 19th century. There are still some imputations but, critically, a substantial gap in the coverage of Latin America, as "the life insurance industry was only later to develop in Spain and Portugal, and both the Spanish and Portuguese militaries did not keep very strong mortality records."

2 Early European Settlements

I first characterize the early settlement patterns of Spanish America. I only provide general remarks about British settlements, but I will study the settlement of the western frontier of the United States after the American War of Independence in a subsequent section.[5]

2.1 A Primer

The European settlement of the New World was staggered. Spain and Portugal had a century-long head start over England and France, whose late arrival led them to settle in areas already claimed by Spain or considered to be of limited value. By the time English settlements began, Spanish America was already settled.

There are no reliable estimates of the number of European settlers in the New World, but Altman and Horn ([7], Table 1.1) and Engerman and Sokoloff ([26], Table 1) provide a sense of the relative magnitudes and timing of settlements by country of origin. In both accounts, Spain was the primary source of European settlers for almost two centuries. In Altman and Horn [7], about 437,000 settlers left Spain between 1500 and 1650. They provide no further estimates for Spain, though the war of succession in Spain (1701-14) led to a second wave of significant migration (Pérez-Artés [62]). In the 17th century, about 400,000 settlers left England to areas that included the West Indies, New England, and the Chesapeake. In the 17th and 18th centuries, 321,000 settlers left Britain between 1700 and 1780, including 190,000 settlers from Scotland and 250,000 from Ireland.

Table 1, taken from Butzer [10], highlights the importance of Spanish and British settlements relative to other European countries. It also showcases the success of Spain in America: Spanish America was the only area where the European population grew after a century of settlement. (European descendants in Table 1 do not include *mestizos*.) British

[5] I will only discuss Portuguese settlements in passing, and will not consider French settlements. Despite its size, France did not contribute nearly as many settlers as Spain and England partly because "New France, did not pass for a country of rapid fortunes" (Altman and Horn [7], p. 131). As noted in Altman and Horn ([7], pp. 162-163), "the French were not, however, reluctant to emigrate to America in general but merely to Canada in particular. The climate, with its harsh winter, mosquito-laden summer, and radically shortened growing season, and the continual wars with the Iroquois did little to boost the reputation of the colony." Voltaire's famous plea is well-known: "You know that these two nations [England and France] are at war about a few acres of snow in the neighborhood of Canada, and that they have expended much greater sums in the contest than all Canada is worth."

and French colonies experienced negative population growth, as settlers were initially directed to the Caribbean where the warmer climate increased the virulence and transmission of many diseases. Table 1 also suggest a worsening of the disease environments in the Caribbean in the 17th century, as attrition in British and French settlements was many times higher than in Spanish settlements. Within North America, settlements were directed toward mid-Atlantic and Southern colonies, whose populations grew at lower rates than in New England.

Table 1. Demography of the New World in the first century after settlement.

	Number of European	
	Settlers	Descendants
Spanish colonies to 1600		
Mexico	90	92
Central America		20
South America	63	116
Caribbean	22	12
Total	175	240
British colonies to 1700		
Newfoundland	39	3
New England		93
Mid-Atlantic	116	54
Southern		104
Caribbean	222	36
Total	377	290
French colonies to 1700		
Quebec	7	13
Acadia		2
Caribbean	38	6
Total	45	21

Notes: Data in thousands from Butzer ([10], Table 1).

The following features of the European settlement patterns of the New World will be important in the paper:

(i) The Spanish settlement of the New World was fast and wide in scope and scale. It exhibited a *spatial distribution* of settlements consistent with a pattern of diffusion of "invasive species," where settlers followed random dispersals in space, coupled with exponential growth; and a *size distribution* of settlements consistent with Zipf's Law attributed to the highly structured distribution of native populations and mineral deposits.

(ii) The initial English settlement of the New World favored equally the West Indies and North America, where settlements followed a narrow *North-South axis* and concentrated in scattered coastal areas with relatively unfavorable disease environments. In the Western Expansion after independence, US settlements also faced areas with relatively high settler mortality rates and unfavorable disease environments.

I will not discuss the forced transportation of African slaves, even though the Atlantic slave trade played a large demographic role in the settlement of the New World, especially in Brazil and the Caribbean. In British North America, slaves experienced high rates of natural *increase*, while slaves in the Caribbean experienced high rates of natural *decrease*. As Fogel and Engerman note ([30], p. 25), "the rate of natural decrease in the West Indies varied from 5 to 2 percent per annum during most of the 18th century. Over time, the rate of decrease moved toward zero. But it was not until well into the 19th century that the black population of the West Indies became demographically self-sustaining."

2.2 Spanish America

The European settlement of the Americas began with Christopher Columbus's misguided attempt to find a route to Asia. The Taíno cacique Guacanagarí allowed Columbus to leave 39 of his crewmen from the *Santa Maria*, which had wrecked on Christmas Day in 1492 off the coast of *Española*, originally "Ayti" to the local inhabitants. Upon his return in 1493, Columbus found all settlers had died. Spanish settlers reportedly provoked the Caonabo and Mayreni Indians as each settler seized "five women to minister to his pleasure;" Traboulay ([73], p. 25). Columbus relocated his base to the current Dominican Republic and established *La Isabella*, the first European settlement in the Americas.

The mortality rate among natives and the first European settlers in the Caribbean was high, likely due to typhus or influenza (Guerra [39], Cook [15]).[6] *Española* was quickly overrun, and settlers moved to Puerto Rico (1508), Jamaica (1509), and Cuba (1511) decimating the native populations in the Caribbean. Settlements on the mainland of Central and South America began in 1508 following two simultaneous thrusts from *Española* that

[6] Smallpox would not arrive to the Caribbean until 1518. Yellow fever arrived with the systematic introduction of African slaves. The Philippines saw no mortality increase after Spain's arrival (Cook [15]).

would eventually lead to the Mexicas (1519) and the Incas (1532).

Spatial distribution of settlements. Spanish settlements varied over space and time. A way to represent settlement patterns is using invasive species models (Shigesada et al. [67]). Assume an initial number of settlers released at the origin, and let $t \geq 0$ denote time since arrival. Suppose also that an inflow of settlers takes place exponentially at rate $\mu > 0$ and that settlements diffuse through space following a random motion with a diffusion coefficient $D > 0$. The density of settlers is Gaussian and widens over time, with a *settlement frontier* expanding at a rate

$$r(t) \simeq 2\sqrt{\mu D t}, \tag{1}$$

where $r(t)$ is the (radial) distance of the new settlements, taken with respect the origin of the invasion. Simply put, a random diffusion of settlers across space implies linear growth in the radius of the settlement frontier as a function of time.

Hardoy and Aranovich [42] compiled (and corroborated) data on the Spanish settlements circa 1580 and 1630. Their data relies on the reports received by Juan López de Velasco after his appointment as cosmografo-cronista en el Consejo de Indias in 1571 and by the direct accounts from Fray Antonio Vasquez de Espinosa, partly published in 1629. Both sources describe the demography and geography of the Americas, navigational currents, and customs and beliefs of indigenous peoples. Their accounts report the date of "fundacion" of some of the cities, which I associated with t in expression (1). I also consider additional sources, such as Lockhart and Schwartz [49].

Demographic data includes the number of *vecinos* (neighbors), which I take as the number of Spanish settlers. *Vecinos* typically referred to heads of households with the right to participate in government and hold public office.[7] As noted by Hardoy and Aranovich ([43], p.66) "vecinos seem to have been those adult males eligible for citizenship in a Spanish city, and that definition did not include the residents of indigenous cities." The data is not comprehensive of all the inhabited cities at the time or the actual size of the cities.

The data is drawn from separate sources, with slightly different spatial coverage, but with

[7] Vecinos is the equivalent of the English term "burghers," which represents individuals with certain rights and privileges within their cities (Borah [8]).

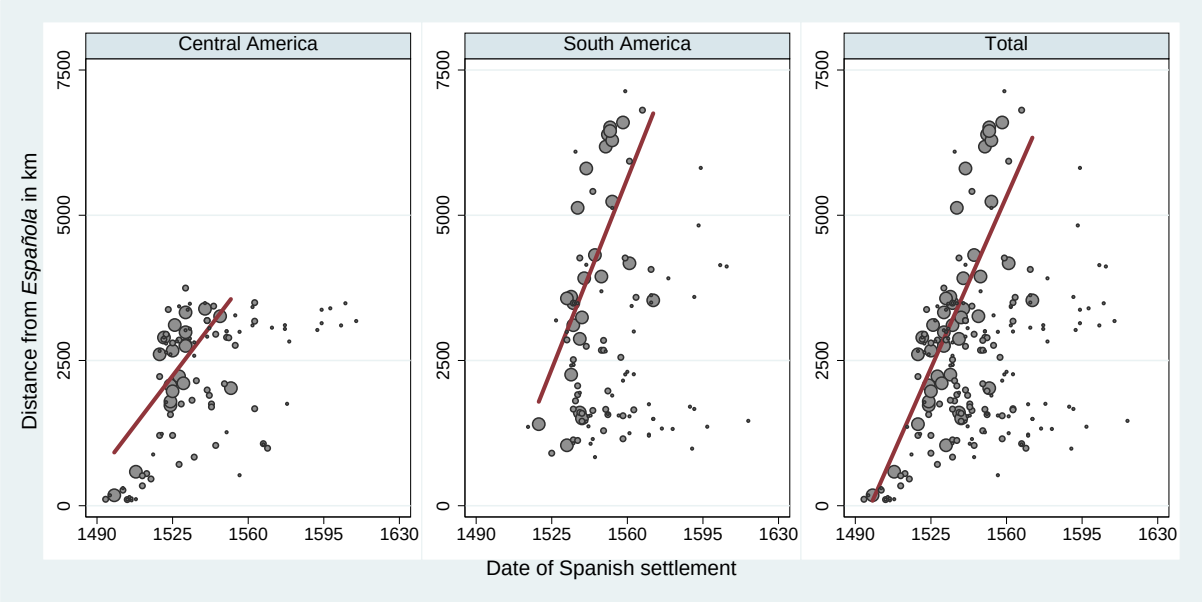


Figure 1: Radial distance (in km.) from *Española* and date of settlement. Line fitted to settlements with more than 100 vecinos in 1580. Small markers represent expanded settlement dates beyond those listed in Hardoy and Aranovich [42].

high compatibility for medium to large settlements. Data for the 1580s contains information for 189 cities, out of a total registry of 225. Data for the 1630s contains 165 cities, out of a registry of 331 (including abandoned settlements).[8]

The early colonial system of Spanish America consisted of the viceroyalties of New Spain (Mexico, Central America, the United States, and Venezuela) and Peru (Panama and South America with the exception of Venezuela). In line with the colonial system, I consider Spanish settlements in *Central America*, which roughly coincide with the viceroyalty of New Spain, and *South America*, which roughly coincide with the viceroyalty of Peru. Although Mexico is geographically part of North America, for the purposes of this paper, I reserve *North America* to cover only the present-day United States and Canada.

The Caribbean was the beachhead for the invasion of the rest of the New World, so I consider *Española* as the center of the invasion. Figure 1 plots the radial distance $r(t)$ to

[8]The year 1580 is significant because the initial phase of Spanish exploration ended with the conquest of Peru in the 1540s, leading to a phase of consolidation and colonization starting around 1580. Although areas such as California, Venezuela, and the Southern Cone remained sparsely populated and peripheral until the Bourbon reforms, virtually all the colonial cities that would serve as administrative capitals for Spain were established before 1580. The year 1630 is significant because it provides Spanish settlement information *before* the systematic introduction of the African disease pool, the Bourbon reforms, and the gradual decline of the empire.

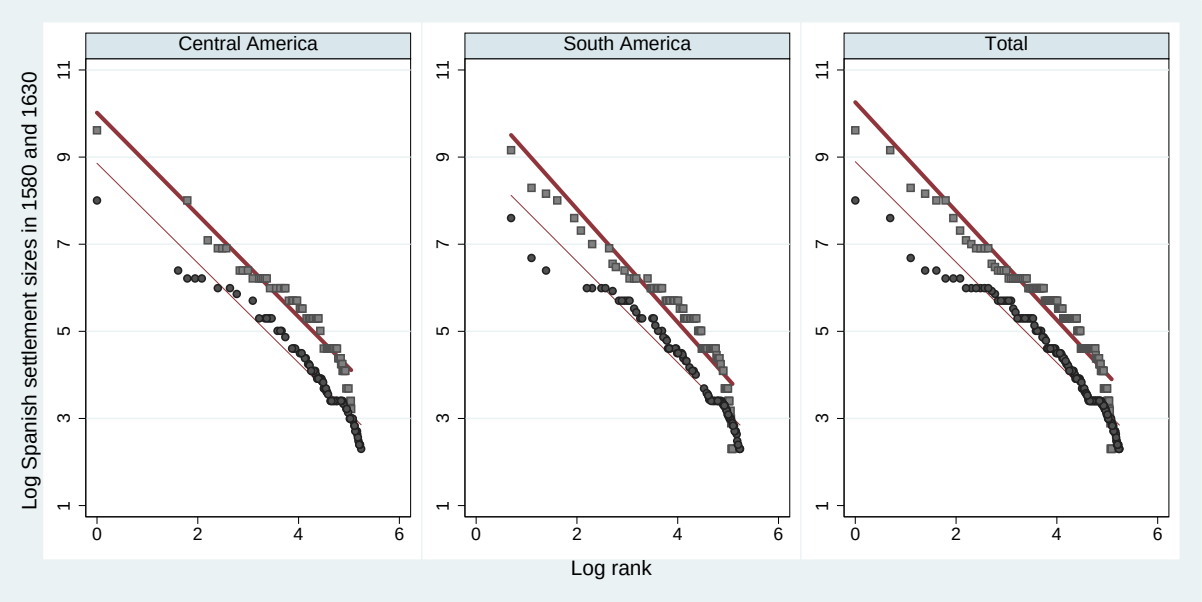


Figure 2: Double log plot for settlement size, 1580 and 1630, and rank.

Española as a function of the date of settlement, t. Expression (1) suggests that $r(t)$ should increase linearly over time. A linear representation fits the data well. In settlements with more than 100 *vecinos* in 1580, a linear relationship accounts for more than one-third of the variation in the data. The radius of the Spanish settlement frontier increased by about 45km per year in Central America and by about 90km per year in South America. (Considering settlements of any size lowers the distance covered per year, as shown in the Appendix.) Including cities not dated in Hardoy and Aranovich [42] induces minimal changes in Central America, but it leads to insignificant results in South America. South American settlements were limited by the Andes, which is not accounted for in expression (1).

Size distribution of settlements. The number of *vecinos* varied substantially across settlements and over time. I represent the size distribution of European settlements relating size and rank in a double log plot. Under power law distributions, a relationship between log-size and log-rank should be described by a straight line. Figure 2 plots such relationship for the 1580 and 1630 data, across the two main areas of Spanish geographic expansion. Spanish settlements in either year or geographic location are characterized by a double-log relationship. The size distributions are parallel with a slope numerically and statistically equal to one, as in *Zipf's law* representation in urban economics (Table A7).

Table 2. Determinants of European settlement sizes and growth in Spanish America.

	A. Log-*vecinos*, 1580				B. Log-*vecinos*, 1630				C. Growth	
	Central	South	All	All	Central	South	All	All	1630-1580	
Log-N_j	*0.07*	0.05	*0.06*	0.04	*0.11*	*0.13*	*0.12*	*0.10*	*0.045*	*0.051*
	(0.02)	(0.05)	(0.02)	(0.02)	(0.04)	(0.06)	(0.03)	(0.03)	(0.02)	(0.02)
Distance				*0.57*				*0.43*		−0.10
				(0.10)				(0.16)		(0.16)
Date				*−0.02*				*−0.01*		0.01
				(0.00)				(0.00)		(0.00)
Silver				−0.10				*0.32*		*0.34*
				(0.13)				(0.19)		(0.20)
Gold				0.16				−0.04		−0.20
				(0.18)				(0.09)		(0.21)
Port				*1.65*				*1.90*		0.61
				(0.47)				(0.51)		(0.44)
Mean	3.94	4.01	3.98	3.98	5.24	5.03	5.12	5.12	0.91	0.91
R^2	0.08	0.01	0.04	0.24	0.20	0.05	0.11	0.30	0.03	0.09
N Obs	96	93	189	167	70	92	162	134	93	85

Notes: Robust standard errors in parentheses. Italics denote statistical significance at the 90 percent or more. Log-*vecinos* represents settlement sizes in logs; Log-$N_j$is log-population estimates of natives within 100km, from the cities in Chandler [13] in 1500; *Distance* is relative to Española; *Date* is the date of the settlement; *Silver* and *Gold* are the number of silver and gold deposits within 100km of the European settlement; *Port* is an indicator for the cities listed in the text; and *Mean* is the average value of the dependent variable.

What accounts for the European settlement patterns in Figure 2? The coexistence of a few large settlements and several small settlements is consistent with dominant areas attracting disproportionate shares of settlers. As in agglomeration models in urban economics such as *preferential attachment* models (Simon [69]), new settlers likely attached to areas with an already large native population. As these areas become more attractive with the arrival of European settlers, dominance became self-reinforcing leading to a skewed size distribution of settlers. To study the importance of native populations, Table 2 examines the relationship between the number of native inhabitants and the number of *vecinos* in 1580 and 1630. I consider native cities and their population in Chandler [13], estimated prior to the arrival of Europeans in 1500. For each city, I measure its distance to the cities in Chandler [13], and count the size of the native population within a 100 kilometers radius.[9]

[9]There is a large degree of uncertainty about the actual pre-contact native population numbers (Maloney and Valencia Caicedo [51], Newson [59]), but large cities are a useful benchmark as there are perhaps less

Table 2 controls for the presence of silver and gold deposits within 100km of the settlement.[10] I also include an indicator for whether the Spanish settlement was a central port for Spain, i.e., La Habana, Veracruz, Cartagena, Santo Domingo, San Juan, Lima, or Acapulco (the main departure port towards the Philippines). Peripheral ports such as Montevideo, Buenos Aires, and Caracas would become important only after the Bourbon reforms.

Central American settlements in 1580 were directed towards areas with large native urban populations, but the association was not robust. In South America, there was no association between native populations and European settlers in 1580, possibly because this was still a period of exploration. South America is farther away from *Española* and settled at later times. By 1630, however, European settlements in both Central and South America were directed towards areas with large native populations, in ways consistent across regions. Point estimates in Panel B are equal across regions and robust to the many controls.

Table 2, Panel C, examines the growth rate of the number of *vecinos* between 1580 and 1630, which differences out fixed effects associated with each location. The sample size is slightly reduced, but the size of the native population and the presence of silver deposits are significant determinants of settlement growth. The average annual growth rate in Panel C is $0.91/50 = 0.0182$, which implies a doubling time of 38 years. (Panels A and B imply faster growth of $(5.12 - 3.98)/50 = 0.023$, or a 30 year doubling time.) To better understand the estimates in Panel C, it is helpful to consider the largest cities in each region. In Central America, the largest cities were Tenochtitlan (80,000), Texcoco (60,000), and Tlaxcala (40,000). In South America, they were Cuzco (50,000), Quito (30,000), and Bogotá (20,000).

The initial size of urban populations and the availability of mineral deposits are causally exogenous to the settlement decisions, although omitted variables may operate as a common cause of native and European settlements. To get a sense of the growth rate of European

difficulties with determining population size than with smaller areas. There is nothing special about 100km. I examined distances from 50km to 500km, and results are virtually unchanged.

[10] I rely on the the Mineral Resource Data System (MRDS) maintained by the United States Geological Survey (USGS), available in https://mrdata.usgs.gov/. The MRDS provides a general geologic setting of deposits for major nonfuel minerals. There is no indication of the size of the deposits, but there are records of the amout of silver and gold exported from the main mines in the New World; see, e.g., Palma [60]. As the size of the deposits was unknown and actual extraction depended on a number of factors, the presence of the deposits is more informative about settlements than the actual extraction rate.

settlements, consider a point estimate of 0.051 and a measure for city size of 20,000. Urban areas should see an annual inflow of settlers at a rate of

$$\text{Settlement growth rate in large native city: } 0.0182 + \frac{0.051 \times \log(20{,}000)}{50} = 0.028,$$

which yields a doubling time of 25 years. The average number of silver mines within 100kms is 0.29. As a comparison, a silver mining city should see a growth rate of

$$\text{Settlement growth rate in mining city: in } 0.0182 + \frac{0.34 \times 1}{50} = 0.025.$$

Because large cities were geographically concentrated, large urban populations may not properly capture the influence of native populations in the pattern of Spanish settlement. López de Velasco [75] reported the number of tributary indians and the number of *encomenderos* in some cities in the 1580s. I show in the Appendix (Table B1) that *vecinos* were driven to areas with tributary indians. There is a positive influence for both *encomenderos* and non-*encomenderos*, but tributary indians were more important for non-*encomenderos*, likely because land was more critical than labor for encomenderos.

2.3 British America

The English settlement of North America mirrored the settlement of Portuguese America instead of Spain's territorial strategy likely because there were fewer and less densely concentrated native populations, and limited mineral deposits (Sokoloff and Engerman [70]).[11] As in Brazil, settlements were coastal and followed a narrow North-South axis, instead of the East-West direction of the early 19th century's Western Expansion.[12]

[11] Portuguese settlements were overwhelmingly scattered coastal ports. As noted by Lockhart and Schwartz ([49], p. 229), Brazil's population and its cities remained "like crabs on the beach" until the discovery of gold, and later diamonds, in *Minas Gerais* in 1690s. Throughout the 17th century, Portuguese settlements occurred at a slower pace and with much larger number of African slaves than Portuguese settlers; Carrara ([11] , Table 1). Sugar cane and gold played a key role in the development of Brazil's economy and demography during the colonial period (Naritomi et al. [57]).

[12] Unlike Portuguese America, the interior of the US did not have rich mineral deposits. During the 1830s, gold was discovered in the Appalachian Mountains. The land where these deposits were discovered belonged to the Cherokee Nation. The land was soon expropriated, and the Cherokee tribes were forcibly removed. In the 1840s, small amounts of mineral deposits, mainly gold, were discovered in California, prompting the arrival of Chinese immigrants as the main source of labor in mining.

North America was also settled much later than Spanish America. Initial settlement attempts by Spain, France, and England failed.[13] The first permanent English settlement was established in Virginia in 1607, half a century after Spain's settlements in Florida. Mortality rates in Jamestown were catastrophic. At the end of its first year, 1607-1608, only 38 of the original 104 settlers were alive (Vaughan [74]). A decade after the initial settlement, and a migration of more than 1,700 settlers, Virginia counted only 351 European settlers. The Virginia Company sent 3,570 people between 1619 and 1622 and at least 6,000 people went to Virginia between 1607 and 1624. In 1625 the population was near 1,200 (Kupperman [46], p. 24).

Mortality conditions improved over time, especially in New England (Gallman [34], Merrens and Terry [55]).[14] Estimates of death rates in New England during the 17th and 18th centuries are as low as 15 per thousand, while settler mortality rates in Virginia and Maryland were as high as 250 and 140 per thousand (Fogel et al. ([31], p. 75), Gemery ([38], Table 5.6)). These mortality rates are comparable to the mortality rates of European soldiers in the west coast of Africa during the 19th century (Curtin [17]). It is not possible to determine when North America became more attractive to settlers than the Caribbean, but the Puritan colony in the island of *Providence* (off the coast of Nicaragua but under Colombian jurisdiction) ended in 1641 due to Spanish hostilities rather than disease (Burns [9]). Likewise, Barbados was the main colonial destination for indenture servants until the 1650s (Galenson [33]). From the 1650s to the 1770s, with the exception of the 1730s, Maryland and Virginia dominated the market for servants. New England, instead, experienced out-migration for more than a century since the 1680s (Smith [68], Gemery ([38], Table 5.8)).

Settlements in the West Indies concentrated on the Lesser Antilles (including Barbados) and the Orinoco delta in Guyana where the boundary between Spanish and Portuguese

[13] Spain attempted multiple settlements in six expeditions to Florida from 1513 to 1563. Ponce de Leon's attempts (with over 200 men) in 1513 and 1521 failed. He died in the 1521 campaign. In 1526 and 1527 more than 500 men under Panfilo Narvaez were also defeated by natives. Hernando de Soto's campaign in 1539 reached the lower Mississippi but he also died. Only in 1565 under Don Pedro Menendez de Aviles (with 600 men), a lasting base at St. Augustine was established, Steele [71].

[14] Out of the 102 passengers and crew of *the Mayflower,* 44 died within departure from England and the first four months after landfall on November 11, 1620 in Plymouth (Deetz and Scrott-Deetz [20]). During the first winter, 20 out of the 28 men died in the Quebec post founded by Samuel de Champlain in 1608.

possessions was blurry (Burns [9]). The British expansion into the Caribbean encountered native and Spanish opposition, and a disease environment altered by European influences. Settler mortality was high due yellow fever and malaria, which only occur in climates with mean temperatures over 20 degrees Celsius (McNeill [54]). Yellow fever first stroke Barbados in 1647 before spreading through the Caribbean (Carter [12]). When Britain seized Jamaica from Spain (1655), British soldiers experienced high mortality rates due to disease. British expeditions to Guadalupe (1689), Martinique (1692), St. Dominigue (1695), Darién (1698), and the French expedition to Cartagena (1697), were unsuccessful due to yellow fever (Burns [9], McNeill [54]). Falciparum malaria and hookworm also arrived with infected Europeans and African slaves in their passage through the Caribbean. In North America, malaria became hyperendemic in all the colonies, except for New England. Yellow fever epidemics frequented Southern ports, but receded from Northern and mid-Atlantic cities after the late 18th century (Duffie [22], Kunitz [45]).

2.4 Overview

The settlement patterns of Spanish and British America differed markedly. Spain establish settlements across almost all areas of the New World in a short period of time. Santiago de Chile, more than 6,000kms away from *Española*, was settled in 1541, less than 50 years from Columbus' arrival. In contrast, North America west of the Appalachian mountains remained mostly unsettled until the early 19th century. Had British settlers advanced at the rate of advance of Spanish settlers, they would have reached the Pacific coast from the interior within a half-century from the settlement of Jamestown. The scale of the Spanish and British settlements also differed considerably. For close to 150 years, Spain had the largest presence in the New World, with the number of settlers in urban areas growing at rates comparable to growth rates during the age of mass migration.

Because the settlement of Spanish America began when tropical diseases from the African pool had not yet visited the New World, Spanish settlers had an epidemiological advantage over native populations. Settlers also targeted areas with a high concentration of natives, which were typically at high altitudes where mineral deposits can be found and where disease environments are less hostile. Altitude acts as a protective factor against disease, as many

tropical vectors and bacteria are sensitive to altitude and UV radiation (Gade [32]). British settlers saw as much potential in the West Indies as in North America, but, by their late arrival, settlers in tropical areas faced a disease environment that included deadly diseases from the African pool. In North America, settlers were scattered along a narrow coastal corridor. New England, whose disease environment is similar to that of northwest Europe, had the lowest mortality rates in the New World, although New England experienced no significant immigration after the 1650s.

3 New Settler Mortality Estimates

I turn now to estimates of settler mortality rates to illustrate the temporal and spatial differences in the disease environments settlers faced in the New World. I rely on the life histories of missionaries documented from missionary archives and compiled by several sources discussed in detail in the Appendix.

The object of analysis is a missionary's survival *spell* measured from the first arrival or appointment to a particular location. I compare the mortality experiences of different populations using Kaplan-Meier estimates and Cox proportional hazard models of the following form:

$$\lambda(t) = \lambda_0(t) \exp\left\{\sum\nolimits_i \beta_i \mathbb{I}_i + \mathbf{x}_i \phi\right\}, \tag{2}$$

where t represents the time since appointment or arrival (i.e., analysis time), and $\mathbf{x}_i$ are controls that may alter the chances of a missionary's survival relative to a baseline hazard $\lambda_0(t)$. The indicator $\mathbb{I}_i$ differentiates locations or time periods. Because the hazard ratio $\exp\{\beta_i\}$ provides a more intuitive and direct interpretation of spatial and temporal differences on the risk of death per year of service, I report the hazard ratios rather than the coefficients themselves. A coefficient $\beta_i > 0$ indicates an increased hazard $\exp\{\beta_i\} > 1$ or an increased risk of death for missionaries in location/period i relative to the baseline represented by $\lambda_0(t)$.

To maximize comparability, I treat life events uniformly across sources, but focus on comparisons *within* a specific data source rather than on comparisons between sources. Catholic bishops were older and had more religious authority and economic power than

missionaries. Data for Catholic bishops is also more complete for birth dates and age at arrival than some of the Protestant sources. As it is common with survival data, most of the censoring takes place because missionaries resigned or abandoned their positions prior to their death. In some instances, the mode of exit is known. I use information about exits due to illnesses in the Appendix. Both the Kaplan-Meir survival plots and Cox hazard estimates can handle right-censoring data. Missionaries sometimes transferred between stations, so I focus on the first spell in the New World.

As a caveat, it is not possible to determine if the health of European missionaries and bishops is representative of the overall health in the colonies or the health of potential settlers. Bishops, for example, had higher social status and resources than missionaries. Missionary encounters also differed from military encounters. Missionary work focused on evangelical, medical, and educational activities. Opportunities for diseases to spread among missionaries were hence fewer than among stationed soldiers or soldiers in campaigns.[15] Missionary efforts also focused on areas with large native populations, which were attractive to actual European settlers in Spanish America. While the location of mission stations was strategic, the assignment of missionaries and bishops to a location was more or less random. A strong sense of service and martyrdom often encouraged missionaries to serve in unhealthy areas.

Based on a later period, Lennox [48] studied personnel records of American missionaries worldwide between 1812 and 1920. He focused on missionary attrition, particularly on life expectancy on the field. Overall, he found no significant selectivity of missionaries based on age or health conditions relative to the population at large, no systematic difference in the duration of service relative to Catholic missionaries, and a decreasing trend in missionary mortality starting in 1820. (This trend was also documented by Curtin [17] for military mortality.) Pérez-Artés [62] has found that Spanish settlers were positively selected from the population in terms of numeracy skills. When possible, I include controls for education and social rank, and for age at arrival if birth dates are available.

[15] As Fogel et al. ([31], p. 82) noted early on, "[m]ortality and morbidity rates of soldiers, even when not engaged in combat, were somewhat higher than for civilians, because barracks-style living tended to promote the spread of infection. [...] One should, therefore, be extremely cautious about using military data as a source of information on the *absolute* level of mortality for civilian populations."

3.1 Protestant Missionary Data

Different Protestant missions (i.e., Puritans, Anglicans, Baptists, Methodists, and Lutherans) settled in North America, some as early as the mid-17th century. Table 3 provides a summary of the data sources for Protestant missionaries and bishops.

Table 3. Selected Protestant Missionary Societies in the New World.

Region	Arrival/Appointment			Number of		
	Earliest	Latest	Median	Arrivals	Deaths	Spells
	A. New England Company (NEC, Kellaway [44])					
US: New England	1649	1785	1709	406	508	263
	B. Society for the Propagation of the Gospel (SPG, Pascoe [61])					
US	1702	1780	1744	294	155	153
Canada	1703	1900	1858	1464	202	201
Central America (a)	1708	1900	1851	429	80	80
Africa (b)	1765	1900	1879	583	73	73
Australia	1810	1900	1866	528	29	29
Asia (c)	1821	1900	1876	686	146	146
Europe	1849	1900	1887	154	11	11
	C. Baptist Missionary Society (BMS, Myers [56])					
India	1793	1891	1853	269	202	202
Africa	1795	1892	1882	88	56	56
Central America (d)	1813	1881	1838	94	91	91
Europe	1844	1889	1873	14	7	7
China (e)	1845	1892	1884	35	11	11
	D. London Missionary Society (LMS, Lovett [50])					
South Africa (f)	1797	1894	1841	281	71	71
India	1798	1893	1840	247	68	68
South Pacific	1797	1894	1836	244	43	43
Asia	1814	1895	1861	162	25	25
Caribbean (g)	1806	1853	1838	35	9	9
	E. Bishops from Foreign Missionary Society (FMS, Emery [25])					
US: Northeast	1784	1921	1870	72	50	50
US: South	1790	1921	1887	96	57	57
US: Midwest	1819	1921	1890	74	42	42
US: West	1853	1921	1902	43	20	20
China	1844	1921	1888	8	5	5
Africa	1851	1921	1877	5	4	4

Notes: (a) Antigua and Barbuda, Bahamas, Dominica, and Panama; (b) Congo, Mauritius, Saint Helena, Sierra Leone, South Africa, and Zimbabwe; (c) Bangladesh, China, India, Indonesia, Japan, Korea, Myanmar, and Sri Lanka; (d) Jamaica, West Indies, and Belize; (e) including Japan; (f) mostly South Africa with a few missionaries in Sierra Leone (2), Central Africa (21) and Madagascar (58), Mauritius (4); (g) Jamaica, Trinidad (4), and Tobago (2).

Protestant missionary data covers the earliest and largest Protestant denominations in Britain's colonies in North America, the West Indies, and Central America. Founded in 1649, the New England Company (NEC) was the oldest permanent Protestant (Puritan) missionary society in the New World. The NEC operated until the American War of Independence in Massachusetts, Connecticut, Rhode Island, New Hampshire, and coastal Maine (Kellaway [44]). The Society for the Propagation of the Gospel (SPG) was established in 1701, to promote the Anglican Christian faith in the American colonies. The SPG was one of the largest and most influential missionary organizations in North America, so it plays a central role in this section. Over time, the SPG expanded to the Caribbean, Africa, India, the Pacific (Pascoe [61]). After the American War of Independence, the SPG continued missionary work in Canada, which remained under British control. Canada was the main destination of SPG missionaries (Table 3).

Both the NEC and the SPG were active in North America during colonial times. The Baptist Missionary Society (BMS), founded in 1792, was part of a broader movement among American and British Protestant churches to engage in overseas missionary work. Most missionaries in the BMS served in India, under British East India Company rule. While there are no BMS missionaries in North America, missionaries served in Jamaica, the West Indies, and Belize. The London Missionary Society (LMS), founded on 1795, was interdenominational and also operated in multiple parts of the world including Jamaica, but with a very limited presence in North or South America (Lovett [50]). Finally, while the Domestic and Foreign Missionary Society of the Protestant Episcopal Church of the United States (FMS) was an *American* missionary society, demographic information on its Bishops covers frontier areas of the western United States where the FMS engaged with native American communities. The FMS had a presence in Africa and China, but with an insignificant number of missionaries (Table 3).

3.2 Catholic Missionary Data

Catholic missionaries accompanied Spanish settlers in their exploration and colonization efforts beginning with Columbus's second voyage. In North America, Franciscan missionary efforts began in the 16th century in the Southwest, but this region was peripheral in the

long-term Spanish and British settlement of North America. Table 4 provides a summary of the data sources for Catholic missionaries and bishops. The first Catholic Missions (CM) to British America and New France, early in the 17th century, were mainly located in New England and the Great Lakes region where the Abnaki, Huron, Iroquois, Illinois, and Ottawa resided. Catholic missions, however, did not progressed in North America partly because the introduction of European diseases led to political instability and social unrest among the many indigenous tribes.[16]

Table 4. Catholic Missionaries in the New World.

Region	Arrival/Appointment			Number of		
	Earliest	Latest	Median	Arrivals	Deaths	Spells
	A. Catholic Missions to North America (CM, Shea [66])					
US: New England (a)	1613	1832	1687	61	39	38
US: Midwest (b)	1615	1764	1678	93	59	53
	B. Catholic Bishops (CB, Cheney [14])					
Europe	1080	1921	1741	9314	8742	8742
Central America	1511	1921	1773	173	168	168
South America	1536	1921	1820	309	275	275
North America	1730	1921	1887	400	397	397
Africa	1468	1921	1895	148	143	143
Asia	1523	1921	1884	576	512	512
Australia (c)	1842	1921	1887	91	90	90

Notes: (a) Maine and New York; (b) Huron, Ottawa, Illinois, and Louisiana; (c) includes New Zealand, the Solomon islands, Togo, and smaller islands collectively known as Pacific Oceania.

The Catholic Church's centralized structure provides a more systematic and comprehensive data for bishops appointed in Rome. Vatican records in the *Hierarchia Catholica* contain a comprehensive list of high-ranked officials with information on their birth, ordinations, appointments, and death since 1080. The mortality estimates in Gutierrez [40], used as the main input for settler mortality in Latin American by Acemoglu et al. [1], relied on Vatican records from *Hierarchia Catholica* compiled by Ritzler and Sefrin [64].

[16] Between 1610 and 1640, European diseases killed about half of the Huron population. The weakened villages were then taken by the Iroquois, who further decimated the Huron nation. Several French Jesuits priest were charged with witchcraft, tortured, and killed eventually being canonized in 1930 as the *North American Martyrs*. As noted by Shea ([66], p. 177) "to dash still more the prospects of Christianity on the Huron Lakes, the summer of 1637 witnessed the pestilence return with renewed fury in their fated country. The calumnies against the missionaries daily increased: not only the medicine-men and the common people, but even the chiefs openly charged the missionaries with destroying the land by witchcraft."

Cheney [14] has digitized and uploaded historical information from *Hierarchia Catholica* and supplemented it with additional sources. His work is not officially sanctioned by the church, but it seeks to offer a comprehensive and current database of information regarding Catholic bishops, cardinals, and other prominent figures, up until present times. In order to keep the sources comparable, I limit the data in Cheney [14] to appointments taking place before 1920.

While the overwhelming majority of bishops were located in Europe, Table 4 shows extensive coverage in the different world regions. It also shows a timing consistent with the staggered pattern of European settlement. The many Catholic Bishops appointed to various dioceses in Europe before and during the European settlement of the New World are one potential reference group for survival estimates.

3.3 The North American Experience

I first provide some graphical evidence that illustrates the nature of settler mortality in British North America. The findings in this section speak to whether the mortality rates of soldiers in temperate areas of the New World in the mid-19th century (Curtin [17]) are representative of settler mortality conditions over time and space within North America. A comparative analysis of settler mortality in the New World is presented below.

Figures 3, 4, and 5 plot survival rates as a function of the number of years since arrival to the New World for Catholic Missionaries (CM) in North America, and for Protestant missionaries from the New England Company (NEC) and the Society for the Propagation of the Gospel (SPG). I report the Kaplan-Meier estimate, whose jagged steps depend on the number of missionaries at risk at each point in time, and the smooth baseline Cox hazard estimate. As it will be the case with all survival estimates in the paper, the Cox proportional hazard model fits the data well. In the figures, the *median survival* is the time where survival reaches the dashed line representing 50 percent, or the "half-life."

Figure 3 plots survival curves for the earliest missionaries in the sample, the CM sample. This includes the survival of missionaries in New England (and New France) and in the Midwest, which includes the Great Lakes region an a few missionaries in Louisiana. Relative to the Midwest, the figure suggest a significantly lower settler mortality in early colonial New

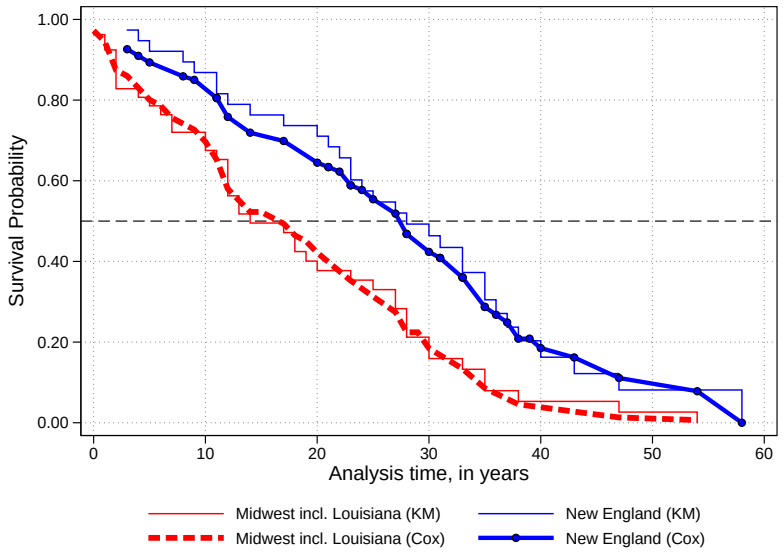


Figure 3: Kaplan-Meier and baseline Cox hazard survival curves for Catholic Missionaries (CM) to North America, 1613-1832, for missionaries in New England (and New France) and missionaries in the Great Lakes region (and Louisiana).

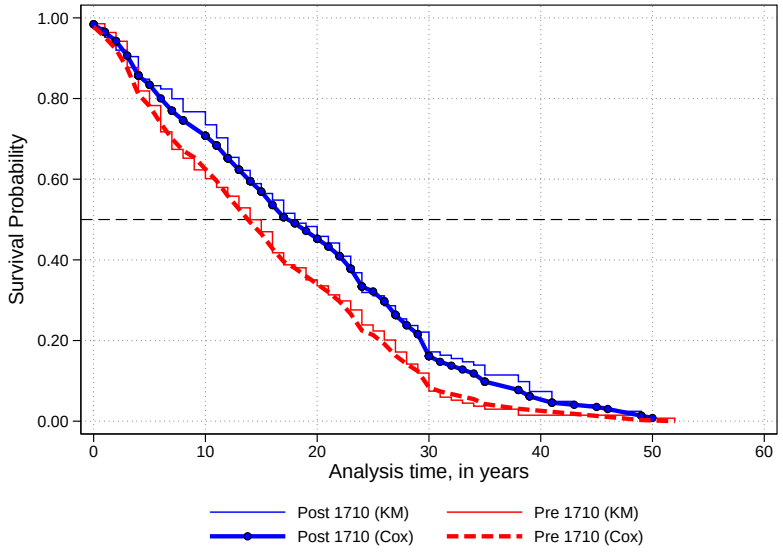


Figure 4: Kaplan-Meier and baseline Cox hazard survival curves for Protestant Missionaries from the New England Company (NEC), 1647-1785, for missionaries who arrived before and after 1710, the median arrival time of the NEC.

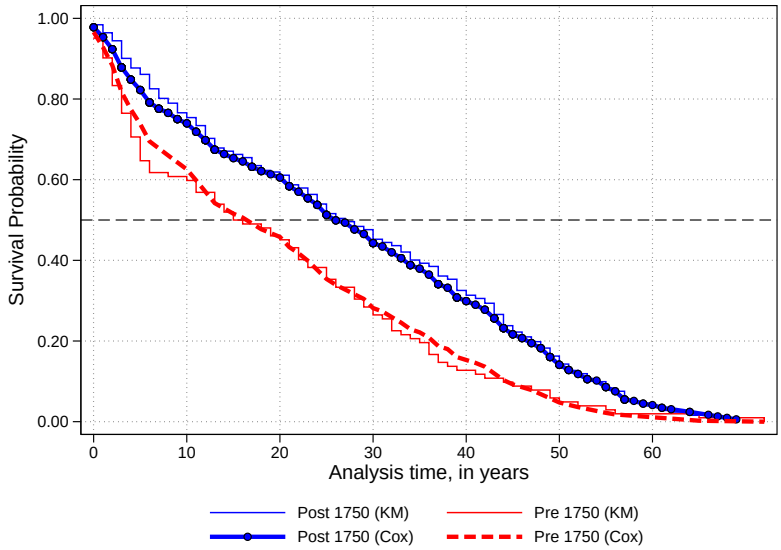


Figure 5: Kaplan-Meier and baseline Cox hazard survival curves for Protestant Missionaries from the Society for the Propagation of the Gospel (SPG), 1702-1900, for missionaries who arrived to North America before and after 1750, the median arrival time to the US in the SPG.

England. Figure 4 examines Protestant missionaries from the NEC, who worked exclusively in New England. Instead of comparing regions, this figure plots the survival curves for missionaries who arrived to New England before and after the median arrival year in the sample or 1710 (Table 3). Consistent with the historical accounts in Section 2, this figure suggests that survival in New England improved over time.

To assess the robustness of the survival curves, the Appendix considers Cox hazard estimates with added controls, including controls for the year of arrival and an indicator for missionary killings as the cause of death. Table A1 shows that, as in Figure 3, Catholic missionaries in the Midwest were about twice as likely to die per year of service compared to Catholic missionaries in New England. It also shows that the higher prevalence of missionary killings does not fully account for the increased mortality risk in the Midwest. I also find that the risk of death in New England was 36 percent higher for early arrivals relative to late arrivals, as in Figure 4. Commissioners experienced a reduced risk of death, but controlling for Commissioner status does not alter the estimate of early arrivals. The downward trend in mortality in New England is also robust.

Table 5. Cox model survival estimates within North America.

	A. Differential survival in North America: SPG						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
North America <1750	*1.55*		*1.36*	*1.45*	*1.48*		
	(0.21)		(0.25)	(0.20)	(0.29)		
Arrival year						*0.98*	*1.05*
						(0.01)	(0.02)
Southern colonies		*2.95*	*2.24*		*3.84*		*7.34*
		(0.68)	(0.61)		(2.12)		(3.74)
Middle colonies		*1.38*	1.16		1.03		1.77
		(0.20)	(0.22)		(0.32)		(0.62)
New England		*0.74*	*0.63*		*0.53*		0.91
		(0.12)	(0.11)		(0.14)		(0.27)
Educated				0.89	0.81		0.77
				(0.12)	(0.11)		(0.10)
Priest				*0.65*	*0.50*		*0.49*
				(0.07)	(0.06)		(0.06)
Latitude					1.05		1.03
					(0.04)		(0.04)
Longitude					*0.98*		*0.98*
					(0.00)		(0.00)
	B. Differential survival in the United States: SPG						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
United States<1750	*1.51*		1.27	*1.41*	*1.39*		
	(0.20)		(0.23)	(0.19)	(0.27)		
N. Spells	354	354	354	354	354	354	354

Notes: In Panel A, the variable *North America<1750* is an indicator for arrival to North America prior to that year; *Arrival year* represents the year of arrival of the missionary relative to 1492, in decadal units (i.e., divided by 10); *Educated* and *Priest* are indicators for these categories in the SPG; *Southern Colonies* include Maryland, South and North Carolina, and Georgia; *Middle Colonies* includes New Jersey, New York, Pennsylvania, and Delaware; *New England* includes stations in Massachusetts, Connecticut, Rhode Island, New Hampshire, Maine, Vermont, and Narragansett; *Latitude* and *Longitude* are based on the location of the missionary station. Panel B reports estimates for arrivals to the US before 1750, under the specifications of Panel A.

Both the CM and NEC samples have limited geographic and temporal coverage. Figure 5 examines temporal differences in survival for missionaries from the SPG in North America, which contains wider geographic and temporal coverage than the CM and NEC samples. I use 1750 as the cutoff for distinguishing between early and late arrivals to North America because this is the median arrival year in the United States. Therefore, the mortality risk for early missionaries is compared not only to late arrivals to the US, but to a baseline that

also includes missionaries arriving to Canada after the American War of Independence.

Table 5 reports Cox hazard estimates that examine differential mortality risks in North America. Panel A, Column (1), suggest a 55 percent higher risk of death per year of service for missionaries arriving before 1750, or $\exp\{\beta_i\} = 1.55$, as in Figure 5. Column (2) adds regional indicators for the American colonies, excluding temporal indicators. The most notable point of Column (2) is the substantially higher risk of death in Southern Colonies relative to New England; a ratio of 4 to 1, or $2.95/0.74 = 3.98$. These large regional differences in mortality remain unchanged once the timing of arrivals is included in Column (3). Being an ordained priest significantly reduced the risk of death in Column (4), but it does so without altering the higher mortality risk for early arrivals. Column (5) introduces a number of additional controls including temporal and regional indicators, controls for whether the missionary was educated or an ordained priest, as well as the latitude and longitude of the missionary stations. Latitude is not a significant factor, which is expected given its correlation with the Southern Colonies indicator. Longitude is a negative and significant risk factor, likely due to the reduced health risks and warfare with native populations when SPG missionaries moved to the interior of Canada. Because the arrival year, the composition of the missionary pool, and the locations of the missions are correlated with each other, the many controls in Columns (5) likely overstate the regional differences in mortality. The indicator for early arrivals in Column (5), however, is not statistically different from that in Column (1) suggesting that the graphical pattern in Figure 5 is robust.

Column (6) includes the arrival year instead of a discrete control and reports a decreasing trend in mortality over time. Column (7) includes the same set of controls as specification (5), and reports an increasing trend in mortality. The observed increasing trend in mortality in Column (7) is likely the result of confounding between the arrival of ordained priests and missionaries to southern stations. Mortality changes over time were also not smooth enough to be properly captured by the constant linear trend assumed in Columns (6) and (7). Panel A looks at North American arrivals. Table 5, Panel B, reports estimates for early missionary arrivals to the US under the same specifications as in Panel A (only the temporal indicator is reported for convenience). For the most part, both sets of estimates are statistically equal because the early mortality conditions in Canada and the United States were quite similar.

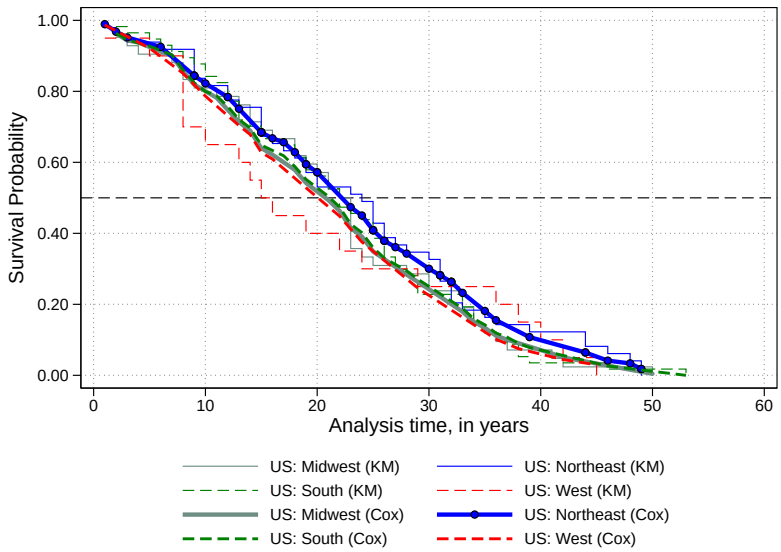


Figure 6: Kaplan-Meier and baseline Cox hazard survival curves for Bishops from the Foreign Missionary Society (FMS), 1784-1921, for American Bishops in different US regions.

The previous figures and tables focused on settler mortality during colonial times. To examine regional differences at later dates, I use information from bishops in the United States between 1784 and 1920 appointed by the Foreign Missionary Society (FMS), Figure 6. FMS bishops reached Western states in 1853, during the Western Expansion. The figure shows no significant difference in mortality across US regions from the mid-19th century on. This finding is also evident in the Cox hazard estimates available in the Appendix (Table A2). The temporal and spatial differences in survival rates during colonial times are not visible in the FMS data. This is possibly because mortality disparities declined over time, but more likely because the population of Bishops is more selective than the missionary population. As noted before, bishops are older and had higher religious and economic status.

The key takeaway from Figures 3 to 6 and Tables 5 and 6 is the large degree of variation over time and space in missionary mortality in North America. This suggests that mortality in military forts in the Northern US during the mid-19th century (Curtin [17]) understated the mortality experience of European settlers in colonial British America by focusing on later periods and in the only area where survival improved over time. New England reached an exceptional disease environment relative to the mid-Atlantic and Southern colonies, which

were the main settlement areas of British North America.

3.4 Comparative Mortality: Protestant Missionaries

I move now to examine differences in settler mortality between North America and other regions in the New World in the SPG data. The first part of this section speaks to the differences in the mortality of Protestant missionaries in North and Central America, over time. An advantage of the comparative mortality estimates here is that they rely on a uniform source. An advantage of looking at Central America is that this was the main area where British, French, Dutch, and Spanish settlers overlapped in the New World. To validate some of the mortality differences, I also consider Baptist missionaries, but at later times. The Central American stations provide information about the mortality experience of British settlers in Antigua and Barbuda, Bahamas, Belize, Dominica, Jamaica, and Panama. There are limited mortality estimates for protestant missionaries in South America.

Table 6 reports Cox hazard estimates that examine differential mortality risks between Central and North America, during early and late arrival times. As 1744 was the median arrival year to the US (Table 3), I continue using 1750 as the cut-off for early arrivals to North America and the US. For Central America, the median arrival year was 1851 (Table 3), so I consider 1850 as the cut-off for early arrivals. Yellow fever became widespread throughout the Greater Caribbean in the 18th century (McNeill [54]), but the early sample sizes are too small for a balanced comparison. To the extent that knowledge about tropical disease increased in later periods, a late cut-off would bias mortality estimates toward finding high mortality early on and a decline in mortality for late arrivals (which, incidentally, is the opposite of my findings). A cut-off in 1850 also aligns with the period of the settler mortality estimates in Curtin [17]. To determine baseline mortality, I consider all world regions with SPG missionaries listed in Table 3, but my focus is on differences within the New World.

Column (1) includes indicators for North American and Central American missionaries, regardless of the arrival time. Column (1) suggest a reduced risk of death per year of service for missionaries in North America, with $\exp\{\beta_i\} = 0.59$. As the estimate for Central America is not significant, Column (1) also suggests that the mortality experience of SPG

missionaries in Central America was not different from than of missionaries in other parts of the world including Asia, Africa, Europe, and Australia. Column (2), as in Table 5, shows that mortality conditions in North America significantly improved over time. The mortality risk for missionaries arriving after 1750 was one-third of the risk faced by early North American arrivals (i.e., $0.52/1.52 = 0.35$).

Table 6. Cox model survival estimates for SPG across regions.

	A. Differential Survival for North America				
	(1)	(2)	(3)	(4)	(5)
North America<1750		*1.52*	*1.53*	*1.42*	*1.46*
		(0.19)	(0.20)	(0.20)	(0.21)
North America	*0.59*	*0.52*	*0.52*	*0.35*	*0.38*
	(0.04)	(0.09)	(0.09)	(0.09)	(0.10)
Central America<1850			*0.49*	*0.43*	*0.42*
			(0.12)	(0.10)	(0.10)
Central America	1.23	1.23	*1.82*	1.32	*1.63*
	(0.18)	(0.18)	(0.28)	(0.34)	(0.47)
Australia					*0.65*
					(0.16)
Added controls	No	No	No	Yes	Yes
	B. Differential Survival for United States				
	(1)	(2)	(3)	(4)	(5)
United States<1750		*1.50*	*1.51*	*1.63*	*1.66*
		(0.25)	(0.26)	(0.31)	(0.32)
United States	0.92	*0.72*	*0.71*	*0.70*	*0.76*
	(0.08)	(0.09)	(0.09)	(0.11)	(0.12)
Central America<1850			*0.51*	*0.46*	*0.44*
			(0.12)	(0.11)	(0.10)
Central America	*1.67*	*1.67*	*2.42*	*2.12*	*2.97*
	(0.23)	(0.23)	(0.36)	(0.45)	(0.82)
Australia					*0.52*
					(0.16)
Added controls	No	No	No	Yes	Yes
N. Spells	693	693	693	692	692

Notes: In Panel A, *North America<1750* is an indicator for arrivals prior to that year; *Central America<1850* is also an indicator for arrivals prior to that year. Controls include the latitude and longitude of the missionary station, and whether the missionary was educated or a priest.

Column (3) adds the early arrival indicator for SPG missionaries to Central America. Instead of an improvement in survival, Column (3) shows that SPG missionaries arriving

to Central America after 1850 faced a three-times higher risk of death per year of service relative to missionaries arriving early on. While the first arrival to Central America took place in 1708 (soon after the foundation of the SPG), after the introduction of diseases from the African pool, Column (3) suggests that disease environments improved over time in North America, but worsened over time in the tropics. In fact, because the mortality estimates are based on comparisons *within* the SPG, the point estimates suggest that early mortality in North America was not statistically different from late mortality in Central America. In other words, the mortality conditions British settlers to North America faced was similar to those faced by European settlers in tropical America after the mid-19th century. This is a complete reversal in the healthiness between tropical and temperate areas of the New World.

Column (4) adds to specification (3) controls for whether the missionary was educated or a priest, as well as the missionary station's latitude and longitude. The results are mostly unchanged relative to the reversal of mortality risks in Column (3). Column (5) adds an indicator for missionaries arriving in Australia, whose disease environment has also been viewed as favorable to Europeans settlers. As Column (5) shows, missionaries in Australia did encountered healthier environments, but their mortality risks were not different from those faced by missionaries in Central America early on. Finally, instead of examining North America as a single unit, Table 6, Panel B, examines early arrivals to the US. The specific magnitudes differ slightly, but the main message remains valid: while mortality conditions improved over time in North America, and the US in particular, they worsened over time in the tropics of the New World.

European settlers in tropical areas faced unfavorable disease environments during the 19th century (Curtin [17]). These unfavorable conditions were evident in the SPG missionary data (Table 6), but also visible in the mortality of missionaries from other denominations, such as the Baptist Missionary society and the London Missionary society described in the Appendix (Tables A3 and A4). While patterns in the SPG are clear, they may be the result of missing controls. The SPG data contains information on birth dates for a few missionaries. Birth information is important to control for age at arrival, a natural determinant of mortality. Although sample sizes are reduced by about two-thirds, the main

implications of Tables 5 and 6 remain unchanged after the inclusion of age at arrival as a control, as shown in Table A5. As expected, older missionaries in the SPG had a significantly higher risk of death relative to younger missionaries arriving at the same time. The age-at-arrival effects in the SPG are identical to age effects in the sample of Catholic bishops (to be discussed in the next subsection), which is reassuring for validity. I also include a control for "smallpox" and "drowning" when reported as the causes of death, although not all drownings happened in the crossing of the Atlantic. Information about causes of death is limited, but the controls do not alter the general findings.

3.5 Comparative Mortality: Catholic Bishops

I now examine the comparative mortality of Catholic Bishops (CB) across New World regions and time periods. The CB contains information on bishops in South and Central America dating back to the 16th century, but only after 1730 for North America (Table 3). As I have taken 1920 to be the end point in the samples examined in the paper, I present Cox hazard estimates from specification (2) for three data subgroups that vary depending on arrival times. The first group includes arrivals prior to 1720, the second prior to 1820, and the last includes arrivals prior to 1920. Differences in the regional indicators across these samples reflect how mortality changed over time in relation to its baseline. As in the comparative estimates in the SPG data, I use as a baseline the mortality conditions in world regions other than the New World. The CB data, in particular, has a large number of European bishops, which is useful for comparing mortality rates across the New and Old Worlds at a time when Catholic bishops were not yet present in North America.

Figure 7 plots Kaplan-Meier and Cox survival estimates for bishops appointed by the Vatican between 1080 and 1720. As no Catholic bishop was yet appointed to North America, the figure only reports survival estimates for bishops in Spanish America. That is, for Central and South America. Figure 7 shows that bishops in the New World experienced an increased risk of death relative to European bishops, regardless of whether the bishops were in South or Central America. Figure 8 extends the time range of the previous figure to 1820, hence adding Catholic bishops in North America. As the figure shows, the Cox survival estimates in North America are not different from those in Europe. Catholic bishops

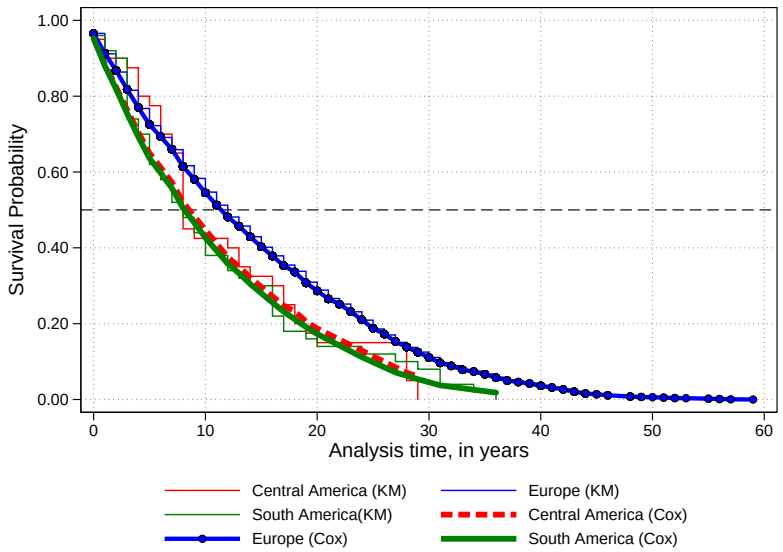


Figure 7: Kaplan-Meier and baseline Cox hazard survival curves for Catholic Bishops (CB) appointed between 1080-1720, in different regions.

in North America had higher survival than bishops in Central and South America. (Survival curves until 1920 are not very different from those until 1820, as seen in the Appendix.)

Table 7 reports Cox hazard estimates that examine differential mortality risks between Central and South America, with different cutoffs for the appointment date. Panel A reports estimates for Catholic Bishops (CB) appointed up until 1720, as shown in Figure 7. Column (1) suggests that all non-European regions had somewhat uniformly higher mortality than Europe, with Africa having the highest relative mortality risk (a doubling of the risk of death relative to the baseline European bishops). Column (2) adds age at appointment, which is available in a large number of observations for the CB data (I only considered the sample where age is available). Age is a robust predictor of survival, and its point estimate is indistinguishable from the estimate of the missionaries of the SPG reported in Table A3. Age at appointment renders most of the spatial differences in mortality insignificant, with the exception of CB in Africa, which still has a high and significant mortality risk for bishops. This means that at the time of settlement, or prior to 1720, settler mortality rates in Spanish America were not statistically different from mortality conditions in Europe other than by a compositional effect associated with the appointment of older bishops to the New World.

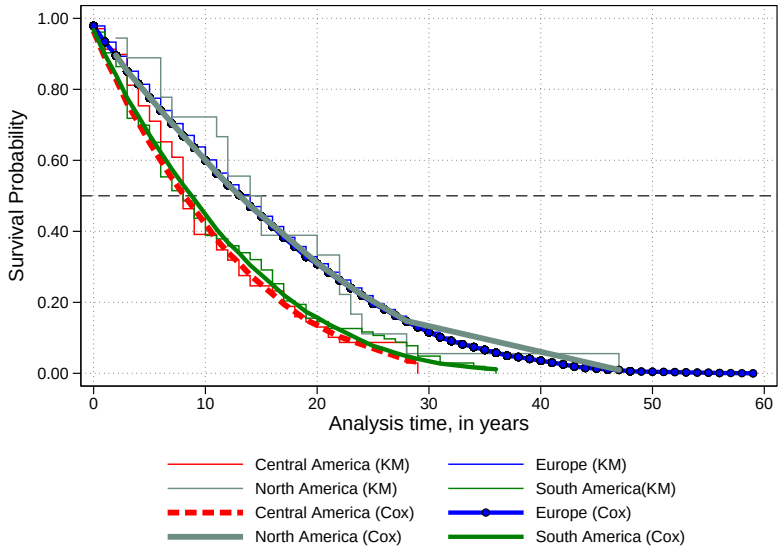


Figure 8: Kaplan-Meier and baseline Cox hazard survival curves for Catholic Bishops (CB) appointed between 1080-1820, in different regions.

Table 7. Cox model survival estimates for CB across regions.

	A. Appointments <1720			B. Appointments <1820		
	(1)	(2)	(3)	(4)	(5)	(6)
Central America	*1.34*	1.04	0.97	*1.70*	*1.34*	1.19
	(0.17)	(0.14)	(0.14)	(0.18)	(0.16)	(0.16)
South America	*1.40*	1.18	1.19	*1.57*	*1.41*	*1.38*
	(0.19)	(0.13)	(0.13)	(0.15)	(0.12)	(0.12)
Africa	*2.32*	*2.01*	*2.02*	*1.80*	1.09	1.06
	(0.60)	(0.57)	(0.57)	(0.54)	(0.60)	(0.57)
Asia	*1.21*	1.08	1.14	*1.15*	*1.23*	*1.29*
	(0.13)	(0.12)	(0.12)	(0.08)	(0.13)	(0.10)
North America				0.99	0.83	0.94
				(0.22)	(0.18)	(0.21)
Age		*1.04*	*1.05*		*1.05*	*1.05*
		(0.00)	(0.00)		(0.00)	(0.00)
Appointment year			*0.98*			*0.98*
			(0.00)			(0.00)
N. Spells	1918	1918	1918	4420	4420	4420

Notes: The sample in Panel A covers Catholic bishops appointed up until 1720, and Panel B, appointed up until 1820. Catholic bishops appointed up until 1920 are studied in the Appendix.

Column (3) adds a control for the appointment year. The hazard estimate of 0.98 indicates a decline in the mortality of bishops as time goes on. Adding this control does

no alter the main findings in Column (2). Panel B expands the sample to include bishops appointed up until 1820, hence including North American bishops, as in Figure 8. Because the sample is expanded and it continues to include the early appointments, an increase in the hazard estimate relative to Panel A indicates a worsening in survival relative to early appointments. Indeed, Column (4) suggests that bishops appointed to Central and South America after 1820 faced more unfavorable disease environments than those appointed to the same regions but in the earlier period. As in Figure 8, there are no mortality differences between Catholic bishops in North America and Europe.

Column (5) adds a control for the age at appointment to the 1820 sample. While the hazard estimates in Column (5) decline relative to those in Column (4), the indicators for Central and South America remain significant and different from the estimate for North America, which is not statistically different from 1. In other words, mortality conditions in Spanish America worsened between 1720 and 1820, relative to conditions up until 1720 when no differences between these areas and Europe existed. Column (6) adds a control for the appointment year. In Column (6), South American bishops still have higher mortality than bishops appointed to North America and Europe.

3.6 Overview

The main takeaway of this section is a reversal in settler mortality rates between British North America and Spanish America. The mortality of Protestant missionaries varied over time and among regions within North America, but there was downward trend that gave way to improved survival after the 18th century, in contrast to an upward trend in tropical regions. Similarly, the mortality rates among Catholic Bishops in Spanish America was relatively low during earlier times, and comparable to the mortality Catholic bishops in Europe. Mortality, however, deteriorated after the 18th century. The reversal in the healthiness of the different areas of the New World implies that settler mortality rates in the mid-19th century overstated the mortality experience of European settlers in Spanish America while understating the mortality experience in British North America.

4 The US Western frontier

At the time of the *Declaration of Independence*, the United States covered a narrow area along the Atlantic coast. In the first decennial census, 1790, the population of the US approached 4 million, with Virginia as its most populous state with 750,000 people. Only six towns in the US had a population of more than 10,000 people. New York, the largest city at that time, had only around 33,000 residents, about one-third the population of Mexico City. Central Mexico, estimated to have had about 11 million people in 1519, also approached 4 million in 1793, after reaching a population of 1.5 million in 1650 (Borah [8]). Through conquest and land purchases, the area of the US increased since the early 18th century. The *1763 Proclamation Line* that prohibited Anglo-American settlements west of the Appalachian Mountains was eventually abandoned. From occupying 867 thousand square miles in 1790, the US territory doubled in size by 1820, and it double once more by the 1840s. (Current area size is 3.8 million square miles.)

Table 8. Military mortality in US and settler mortality across countries.

	Mean	Std. Dev.	Min.	Max.	Obs.
	US military forts, Lawson [47]				
Military mortality Rate, 1822-1825	38.45	39.57	6.23	242	47
Mortality Rate (military returns), 1819-1838	35.45	26.57	3.48	128	45
Mortality Rate (medical returns), 1819-1838	24.04	18.37	2.31	96	45
Disease prevalence rate, 1819-1838	2,634	893	820	4,817	45
Latitude	0.41	0.06	0.27	0.51	57
Longitude	0.91	0.08	0.74	1.06	57
Altitude (in feet)	249.4	288.3	0	896	54

Notes: Mortality rate and Disease prevalence per 1,000.

The territorial expansion beyond the Appalachian mountains and into the Louisiana territory was accompanied by the deployment of US army troops to secure the territory.[17] As military forts were part of the Medical Department established in 1818, each fort reported

[17] As the British did with Jamaica, Napoleon sought to use St. Dominigue to "re-launch France on the path of an American empire," and to invade North America through New Orleans and the Mississippi valley. Out of a military force of over 65,000 European soldiers dispatched, about 80 percent died of "yellow fever in the ports, in garrisons, in hospitals, or onboard ships." Furthermore, "after the St. Dominigue's debacle, Napoleon soured the prospect of a restored French Caribbean empire and sold Louisiana and the heartland of North America to the United States," as noted by McNeill ([54], pp. 260-267).

detailed information on the health and mortality of stationed soldiers. Curtin [17] relied on the reports from Lawson [47] to characterize the health conditions in the northern and southern US in the early 19th century. Lawson ([47] pp. 52-53) reported the number of deaths and the strength of each military fort during 1822-1825, with relatively wide geographic coverage. Data for 1819-1838 contains detailed information by fort, but some temporary forts available in 1822-1825 were not listed. The 1819-1838 contains mortality statistics based on medical and military regimental returns.[18]

Table 8 summarizes mortality rates per 1,000, relative to the strength of the fort (i.e., the number of stationed soldiers). Mortality rates within the US exhibited high dispersion. Low mortality was prevalent in New England and the Midwest, while high mortality was prevalent in the South. The highest mortality in Table 8 comes from Baton Rouge, in a fort with a strength of 1,053 men (in 1822-1825) and a mortality rate of 242.20/1,000. Military returns yield a mortality rate for New Orleans (1819-1838) of 128.20/1,000 and a mortality in Key West of 115.38/1,000.[19] These mortality rates are comparable to the mortality of European troops in the west coast of Africa, and much larger than the aggregate of 34 deaths per 1,000 reported in Curtin ([17], Table 1) for the Southern US or the 15 deaths per 1,000 of the Northern US, the value used by Acemoglu et al. [1] for the US.

Disease mortality was largely environmentally determined. Latitude accounts for a considerable fraction of the variation in mortality within the US. Figure 9 plots the log-military mortality rate per 1,000 during 1822-1825, as a function of the latitude of the fort. Mortality from military and medical returns in 1819-1838 yield virtually identical estimates. (The Appendix reports estimates that also include longitude as a control.) The figure indicates a strong latitude gradient in mortality. This gradient is similar for frontier and no frontier

[18] Medical returns often associated mortality to the "agency of climate." Although, the majority of native tribes had aligned with the British in order to protect their tribal lands during the War of 1812, mortality caused by conflict with native Americans was relatively rare in Lawson [47]. Lawson [47] also reports "epidemic deaths" as a separate category, but the reports are very scattered.

[19] Regarding Baton Rouge, Lawson ([47], pp. 28-30) notes that "[t]he causes of death are not mentioned" but that "[t]he reports of last year, [ending 31st March,] show more sickness and mortality at Baton Rouge than at any other post." Further on, he notes that "Baton Rouge, or more properly speaking, the swamps of the Mississippi, proved literally the grave of the regiment." As noted by Lawson ([47], pp. 290, 303), "at Key West, as in the islands generally of the West Indies, yellow fever has prevailed with much malignity." For instance, "[y]ellow fever prevailed among the naval forces stationed at this island in 1824." Indeed, "[T]he mortality of this station is extraordinarily high [as] every man, on an average, has been reported sick once in every two months and a half."

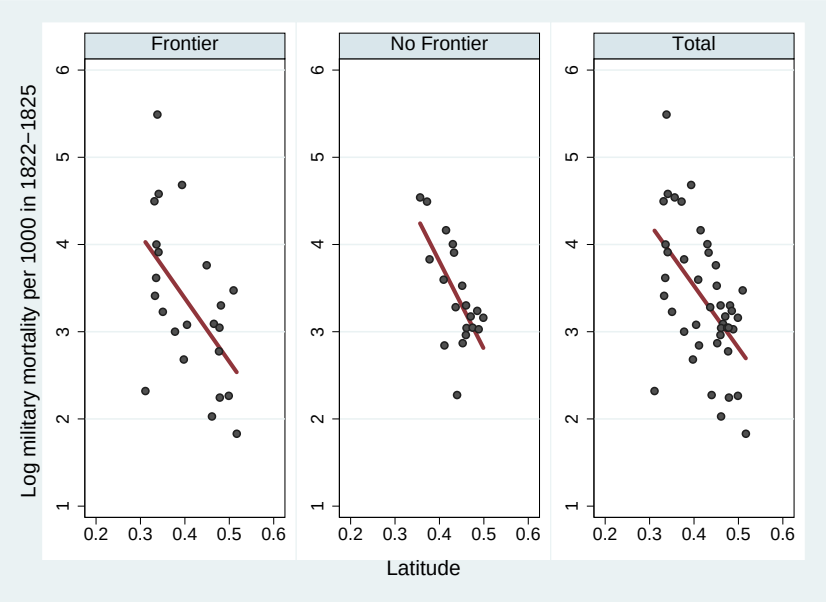


Figure 9: Military mortality rates in US forts, 1822-1825.

forts, with two important differences. First, there is a positive influence of longitude on frontier forts, but not in non-frontier forts, which are almost invariably located near coastal areas. This indicates that forts further away from settled coastal areas experienced higher mortality than non-frontier forts or forts in established settlements. The influence of longitude does not undo the impact of latitude. Second, latitude and longitude account for a larger fraction of the variance in frontier forts than in non-frontier forts. As shown in the Appendix, latitude accounts for as much as 65 percent of the variance of mortality in frontier forts relative to 34 percent of the variance in non-frontier forts.

How large is the latitude gradient *within the US*? Consider the relationship between latitude and the *cross-country* estimates of settler mortality in Acemoglu et al. [1]. (Their Table 2, in the Appendix, appears to contain several typos, which are also registered in the online data files provided by the authors. The typos are unlikely to alter any of their results. The mortality estimates for Mali, Nigeria, and Gambia exceed 1,000, an impossibility. I have replaced their estimates by shifting one decimal point for their reported rates to 294.0, 200.4, and 147.0 respectively.) A log-linear relationship of the form

$$\log(\text{Military Mortality Rate}_i) = \alpha + \beta|\text{Latitude}_i| + \varepsilon_i$$

yields cross-country estimates of $\beta = -4.76$ (s.e. 0.87, $R^2 = 0.31$, Obs.= 87) for their entire sample and $\beta = -2.03$ (s.e. 0.97, $R^2 = 0.12$, Obs.= 40) for ex-colonies. The point estimate for frontier forts within the US in Figure 9 is $\beta = -7.25$ (s.e. 2.59, $R^2 = 0.30$, Obs.= 23), which is larger than the cross-country estimates. Relative to cross-country data, the large gradient within the US is a consequence of the high mortality rates in Southern forts, which are located at latitudes that are relatively far from the Equator.

While the placement of military forts is rarely random, US soldiers were more or less randomly assigned across forts. Suppose that the military mortality rates in Lawson [47] are time invariant and representative of the disease conditions of the county where these forts were located. Accompanying each fort, there is geographic and meteorological information to map each of the permanent forts to current US counties, whose population sizes are available from decennial censuses since 1790. In what follows, I define "frontier" forts as those located in counties with no population in the 1790 census, while "non-frontier" represents forts in areas with existing settlements in 1790. In frontier areas, a long-difference involving the population density of the area where these forts were located should reveal the long-run relationship between settler (i.e., military) mortality and population settlements.[20] In non-frontier areas, long-differences should reveal the relationship between settler mortality and demographic changes driven by migration and natural population growth.

Consider a split sample between frontier and non-frontier areas and a linear projection

$$\log S_{j,t} - \ln S_{j,1790} = \phi_t \log(\text{Military Mortality Rate}_j) + \xi_{j,t}, \tag{3}$$

indexed over a time horizon t, where $S_{j,t}$ denotes the population density of county j at horizon t. (Using population levels rather than population densities provides stronger results. Results are also unchanged if land size is used as a separate control, as shown in the Appendix.) The coefficients ϕ_t trace out the impulse-response to differences in mortality rates at different horizons $t = 1800, 1810, ..., 2010$. The sample of counties in the frontier

[20] It is generally not possible to discriminate between European settlers and settlers of European descent. Many of the settlers to the South and frontier regions were from New England. Indeed, to explain the relatively high mortality in southern posts, Lawson ([47], p. 54) noted that "let it be borne in mind that the troops operating in our southern regions consist mostly of recruits enlisted in the north, with constitutions generally predisposed by intemperance to those diseases which are peculiar to the unacclimated."

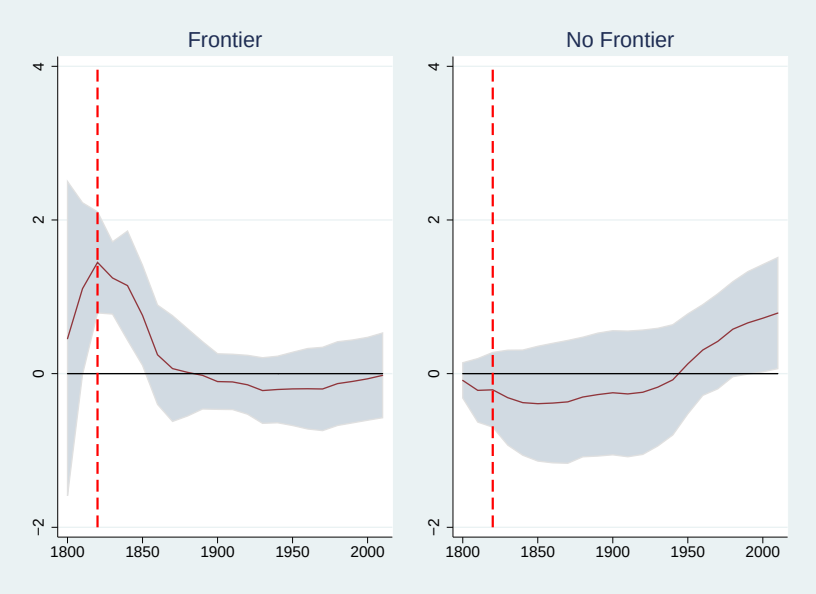


Figure 10: Military mortality, 1822-1825, and log-population density in frontier and non-frontier US settlements.

regressions is unbalanced, as it changes by the settlements themselves.[21] For frontier areas, as $S_{j,1790} = 0$, I consider a dependent variable based only on $\ln S_{j,t}$, which is equivalent to assuming that $S_{j,1790} = 1$ in (3).[22]

Figure 10 plots the coefficients ϕ_t across decennial US censuses for frontier and non-frontier counties using mortality measures for 1822-1825. (Mortality rates according to military and medical returns for 1819-1838 yield even stronger positive effects, as shown in the Appendix.) The dashed line represents the date of measurement of military mortality. Mortality rates were unrelated to population growth in non-frontier areas. In frontier areas *higher* military mortality rates translated into *higher* population densities between 1800 and 1850. The positive influence of military mortality was short-lived. By 1860, mortality rates were unrelated to population density across the once "empty" areas of the US.

As Lawson [47] also provides information about the number of cases of disease, it is possible to construct a disease prevalence measure relative to the number of stationed soldiers in each fort. Figure 11 plots the relationship between the prevalence of disease (per 1,000)

[21] From no counties settled by 1790, the sample includes 12, 15, 17, 22, and 24 counties until 1840. Sample size remain unchanged until reaching 27 frontier counties by 1900.

[22] An alternative definition of "frontier" based on less than two people per square mile in the 1800 census, yields a balanced panel but the same qualitative response of mortality on frontier settlements as in the current definition (see Appendix).

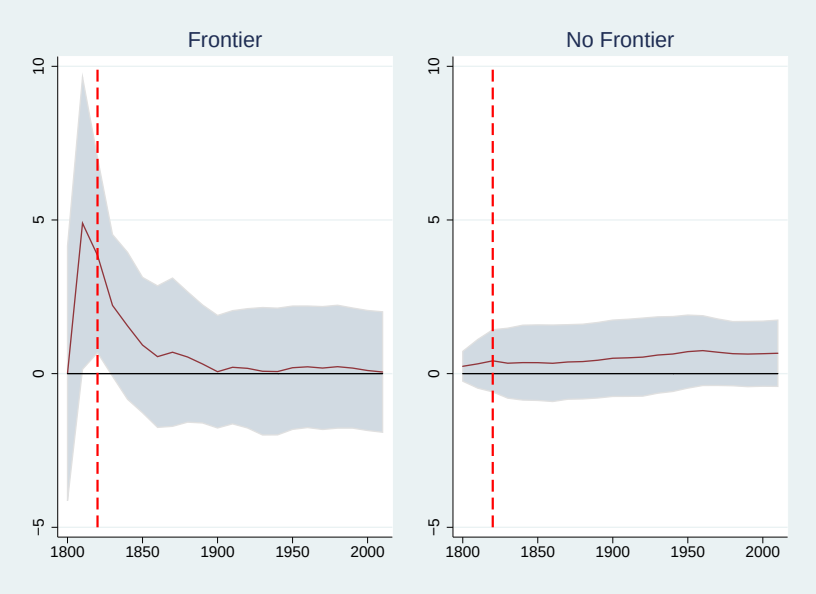


Figure 11: Log-disease prevalence, 1819-1838, and log-population density in frontier and non-frontier US settlements.

and the density of population settlements in frontier and non-frontier areas. As in Figure 10, unhealthy frontier locations were more attractive for settlement than healthy locations.

Overall, the initial British settlement of North America was limited to coastal areas where disease environments varied greatly and were not uniformly favorable to European settlers. New England did reach a favorable disease environment due in part to its cooler climate, but low settler mortality translated into steady out-migration rather than into higher European settlements. European settlers were attracted to the unfavorable Middle and Southern colonies due in part to their greater economic potential. As in the North-South settlement of the Atlantic coast, the positive relationship between settler (i.e., military) mortality and the East-West settlements along the Western Expansion was likely due to economic opportunities that made unhealthy frontier areas desirable for settlement.

5 A Reversal of Geographic Fortunes

Spain settled primarily in healthy areas with large native populations (Table 2). Britain, after Caribbean settlements met with little success, settled along a coastal corridor in North America, as previously done by the Portuguese in Brazil. While health conditions in North

America were not uniformly favorable (apart from New England, Table 3), unhealthiness did not limited frontier settlements in the US (Figures 10 and 11). By the late 18th and early 19th centuries, there was a reversal in economic and demographic conditions between Spanish America and British North America

To gain some perspective on this reversal of fortunes, consider a simple colonial economy. Native population size is N and land endowments are L. The number of settlers is S, and their labor is perfectly substitutable with native labor N. Total population is $P = N + S$. Output is produced by $F(P, L)$, with constant returns to scale. Colonizers extract rents from native workers at a rate τ^N, and from native land owners at a rate τ^L. The return to a colonial enterprise with S settlers and factor endowments (N, L) is

$$\Pi(S|N, L) = F(P, L) - (1 - \tau^N)F_P(P, L)N - (1 - \tau^L)F_L(P, L)L,$$

where $\tau^N = \tau^L = 1$ gives full extraction, so that $\Pi(S|N, L) = F(P, L)$, while $\tau^N = \tau^L = 0$ gives colonizers a competitive return, so that $\Pi(S|N, L) = F_P(P, L)S$. Then,

$$\Pi_N(S|N, L) = \left[\tau^N - \frac{\alpha(1 - \tau^L)}{\sigma}\right] F_P(P, L) - (1 - \tau^N)F_{PP}(P, L)N, \text{ and} \tag{4}$$

$$\Pi_L(S|N, L) = \left[\tau^L - \frac{(1 - \alpha)(1 - \tau^N)}{\sigma}\frac{N}{P}\right] F_L(P, L) - (1 - \tau^L)F_{LL}(P, L)L, \tag{5}$$

where $\alpha < 1$ is the output elasticity of land and $\sigma > 0$ is the elasticity of substitution.

Consider expression (4). The main implications of (5) are symmetric hence I omit them. Under full extraction $\tau^N = \tau^L = 1$, native populations are unambiguously beneficial, as $\Pi_N(S|N, L) = F_P(P, L) > 0$. If no rents can be extracted, i.e., $\tau^N = \tau^L = 0$, native populations are unambiguously detrimental, as $\Pi_N(S|N, L) = F_{PP}(P, L)S < 0$. Even under complete labor extraction $\tau^N = 1$, some rents are lost to land owners, $\alpha(1 - \tau^L)/\sigma$. These losses depend on the land extraction rate τ^L and the production function. More rents are lost when land is a larger input in production, i.e., a higher α, or when land is not easily substitutable for labor, i.e., lower σ.

The previous expressions are consistent with Spain's settlement strategy directed toward areas with native populations and rich land endowments, and with the need for "extractive

institutions" with $\tau^N > 0$ and $\tau^L > 0$ to ensure that $\Pi_N(S|N,L) > 0$ and $\Pi_L(S|N,L) > 0$.

Demographic collapse. Spanish America faced a demographic collapse with large and permanent contractionary consequences unparalleled in British North America. A demographic collapse reduced extractive and competitive rents alike. Native populations faced contractionary forces through the collapse of intensive production techniques such as irrigation canals, terracing, and raised fields, and a muted Malthusian response to labor scarcity. The "wage" bill of native populations, given by $(1-\tau^N)F_P(P,L)N$, depends on τ^N. The higher the degree of extraction, the more muted the wage response and hence the weaker the ability of the native economy to restore population growth.

Native populations declined due to disease, forced labor, and violent conflict and conquest. The spread and impact of European diseases depended on environmental factors, but "it was not uncommon for a single epidemic to carry off one-third, or even one-half, of the population of an area, and in the early colonial period native populations were ravaged by epidemics at almost regular ten-year intervals;" Newson ([58], p. 256).[23] Smallpox, measles, plague, chickenpox, influenza, and other respiratory conditions are highly contagious and particularly damaging in urban settings, but the less densely settled lower elevations areas would eventually be the ones that experienced a permanent collapse of native populations. In the Caribbean, native populations became virtually extinct by the mid-16th century. In Mexico and Central America, the initial demographic collapse was followed by a slow recovery during the 17th century, while in the Andes, native populations continued to decline throughout the colonial period (Newson [58] and [59]).

Depopulation also prompted the use of African slaves as factor substitutes leading to a further deterioration of the disease environment by the introduction of the African disease pool. This additional disease exchange compounded the initial impact of European settlements on native economies, particularly in tropical areas where the changed disease

[23] In addition, instead of a single shock, Spanish America endured a relentless sequence of epidemics. Whitmore [76] simulated the population decline in the Basin of Mexico from a series of epidemics from European diseases between 1520-1521 and 1615-1616. In total, there were 20 epidemics of smallpox, measles, plague, typhus, mumps, influenza, cocoliztili (a generic Nahuatl word for epidemic diseases) in the lapse of less than 100 years. The arrival of diseases of African origin meant an additional worsening of disease environments that further reduced the potential for a demographic and economic recovery. Tropical diseases had disproportionate effects on ports and coastal areas, where yellow fever and malaria became permanently fixtures of Spanish America.

environment limited population growth for African slaves and European settlers, and made the recovery of native populations impossible.

Land expansions. Disease was a concern for settlers in British North America, but marked seasonality meant that tropical diseases manifested through seasonal *epidemic* visits rather than as *endemic* diseases (Kunitz [45]).[24] Favorable disease ecologies eventually made new settlements attractive, but also ensured natural population growth. Positive population growth has been the driving force of American economic growth since colonial times. Even the demographic experience of US slaves was comparatively more favorable than that of slaves in the West Indies. If the US had replicated the demographic experience of the West Indies, instead of a population of one million African Americans in 1800, African Americans would have been only 186,000; Fogel and Engerman ([30], Figure 7).

When $N = 0$, settler population growth, represented by a change in S, is a special case of expression (4),

$$\Pi_S(S|0, L) = \left[\tau^N - \frac{\alpha(1-\tau^L)}{\sigma}\right] F_P(S, L), \tag{6}$$

with L also representing available land. The special case of expression (5) under $N = 0$, is

$$\Pi_L(S|0, L') = \tau^L F_L(S, L) - (1-\tau^L) F_{LL}(S, L)L. \tag{7}$$

Expression (6) says that population growth is beneficial if the substitution possibilities between labor and land are large, or if there is a high degree of rent extraction from land owners. If $\tau^L = 1$ natural population growth would be unambiguously profitable, as $\Pi_S(S|0, L) = \tau^N F_P(S, L)$. Land expansions in (7) are unambiguously beneficial for a colonial enterprise with no native populations, as $\Pi_L(S|0, L) > 0$ regardless of the values of τ^L or σ. Simply put, land and rent extraction from land owners are central when $N = 0$.

Independence from Britain created an opportunity to settle a frontier with abundant and productive land in the interior of the US. Native populations in Spanish America decreased to $N' < N$, while natural population growth and land expansions led to $S' > S$ and $L' > L$ in British North America. In sum, the initial advantage $\Pi(S|N, L) > \Pi(S|0, L)$ of Spanish America, was reversed to $\Pi(S'|0, L') > \Pi(S|N', L)$ in favor of British North America.

[24] Smallpox was present in the colonies, but its demographic effects were small. Except for recurrent visits to Charleston, yellow fever was never a serious problem in British North America (Duffie [22]).

6 Concluding Remarks

The economic fortunes of Spanish America and British North America were reversed during the 18th and 19th centuries, resulting in persistent income disparities that remain to this day. Institutional explanations for their uneven development paths typically rely on divergent institutional factors driven by differences in first-nature geographic aspects such as disease environments and factor endowments. In contrast to second-nature geography, which recognizes political institutions and other human-made influences, first-nature geographic characteristics are often seen as invariant. However, the confluence of European and African disease pools brought by European settlers significantly worsened disease environments in Spanish America relative to British North America. The demographic collapse of native populations in Spanish America and the opening of indigenous lands in British North America suggest that disease environments and factor endowments were in the realm of European colonial influences, rather than purely exogenous and invariant factors.

Because native populations were smaller, "extractive" institutions in British North America operated mainly on land markets rather than on labor markets, as typically done in Spanish America. It is common to view British North America as having "inclusive" institutions, even though settlers in Spanish America had a higher degree of social integration with native populations than in British North America, where the social order has always been highly segregated. In Spanish America, several native populations have retained their ancestral lands, while in British North America, land expansions took place at the expense of the exclusion and displacement of native populations into segregated enclaves or confined reservations. Today, the largest native American populations in the United States are concentrated in areas outside the core of British colonial settlement. The largest fractions of native American populations continue to reside in or near their ancestral lands only in Alaska, where settlers from European descent have always been a minority for natural reasons; New Mexico, where initial European settlements followed a Spanish patterns of assimilation and colonial rule; and Oklahoma, where displaced tribes were relocated under the *Indian Removal Act* of 1830.

References

[1] Acemoglu, D., Johnson, S., and Robinson, J.A. (2001) "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review* 91, 1369-401

[2] Acemoglu, D. Johnson, S., and Robinson, J.A. (2002) "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics* 117, 1231-94

[3] Acemoglu, D., Robinson, J.A., and Johnson, S. (2003) "Disease and Development in Historical Perspective," *Journal of the European Economic Association* 2, 397-405

[4] Acemoglu, D., Johnson, S., and Robinson, J.A. (2012) "The Colonial Origins of Comparative Development: An Empirical Investigation: Reply," *American Economic Review* 102, 3077-110

[5] Albouy, D.Y. (2012) "The Colonial Origins of Comparative Development: An Empirical Investigation: Comment.," *American Economic Review* 102, 3059-76

[6] Allen, W.O.B., and McClure, E. (1898) **Two Hundred Years: The History of the Society for Promoting Christian Knowledge 1698-1898**, Kessinger Publishing

[7] Altman, I. and Horn, J.P. (1991) **"To make America:" European Emigration in the Early Modern Period**, University of California Press

[8] Borah, W.W. (1977) **New Spain's Century of Depression**, University of California Press

[9] Burns, A. (1954) **History of the British West Indies**, Allen & Unwin

[10] Butzer, K.W. (1992) "The Americas Before and After 1492: An Introduction to Current Geographical Research," *Annals of the Association of American Geographers* 82, 345-68

[11] Carrara, A. A. (2014) "The Population of Brazil, 1570-1700: A Historiographical Review," *Tempo* 20, 1-21

[12] Carter, H.R. (1931) **Yellow Fever: An Epidemiological and Historical Study of its Place of Origin**, Williams & Wilkins Company

[13] Chandler, T. (1987) **Four Thousand Years of Urban Growth: An Historical Census**, Lewiston

[14] Cheney, D.L. (2013, November 23) **The Hierarchy of the Catholic Church: Current and Historical Information about its Bishops and Dioceses**, https://www.catholic-hierarchy.org/

[15] Cook, N. D. (2002) "Sickness, Starvation, and Death in Early Hispaniola," *Journal of Interdisciplinary History* 32, 349-86

[16] Croil, J. (1894) **The Noble Army of Martyrs & Roll of Protestant Missionary Martyrs from A.D. 1661 to 1891**, Presbiterian Board of Publications

[17] Curtin, P.D. (1989) **Death by Migration: Europe's Encounter with the Tropical World in the Nineteenth Century**, Cambridge University Press

[18] Davis, K.G. (1975) "The Living and the Dead: White Mortality in West Africa, 1684-1732," in Engerman S.L. and Genovese, E.D. (Eds.) **Race and Slavery in the Western Hemisphere: Quantitative Studies**, Princeton University Press

[19] Diamond, J. and Panosian, C. (2006) "Continental Asymmetry of the Origins of Human Infectious Diseases," in Hämäläinen, P. (Ed.) **When Disease Makes History: Epidemics and Great Historical Turning Points**, Helsinki University Press, 17-44

[20] Deetz, J. and Scott-Deetz, P. (2000) **The Times of Their Lives: Life, Love and Death in Plymouth Colony**, W.H. Freeman and Company

[21] Droller, F. (2018) "Migration, Population Composition and Long Run Economic Development: Evidence from Settlements in the Pampas," *Economic Journal* 128, 2321-52

[22] Duffie, J. (1953) **Epidemics in Colonial America**, Louisiana State University

[23] Easterly, W. and Levine, W. (2003) "Tropics, Germs, and Crops: How Endowments Influence Economic Development," *Journal of Monetary Economics* 50, 3-39

[24] Easterly, W. and Levine, R. (2016) "The European Origins of Economic Development," *Journal of Economic Growth* 21, 225-57

[25] Emery, J.C. (1921) **A Century of Endeavor, 1821-1921: A Record of the First Hundred Years of the Domestic and Foreign Missionary Society of the Protestant Episcopal Church in the United States of America**, Department of Missions New York

[26] Engerman, S.L., and Sokoloff, K.L. (2002) "Factor Endowments, Inequality, and Paths of Development among New World Economies," *Economía* 3, 41-109

[27] Vasquez de Espinosa, A. (1629 [1948]) **Compendio y Descripción de las Indias Occidentales,** Smithsonian Institution

[28] Eubel, (1913, 1914, 1923) **Hierarchia Catholica Medii Aevi: I (1198-1431), II (1431-1503), III (1503-1592)**, Monasterii Sumptibus et typis librariae Regensbergianae

[29] Feyrer, J., and Sacerdote, B. (2009) "Colonialism and Modern Income: Islands as Natural Experiments," *Review of Economics and Statistics* 91, 245-62

[30] Fogel, R. and Engerman, S. (1974) **Time on the Cross**, Little Brown

[31] Fogel, R., Engerman, S., Trussell, J., Floud, R., and Pope, C. (1978) "The Economics of Mortality in North America, 1650-1910: A Description of a Research Project," *Historical Methods* 11, 75-108

[32] Gade, D.W. (1979) "Inca and Colonial Settlement, Coca Cultivation and Endemic Disease in the Tropical Forest," *Journal of Historical Geography* 5, 263-79

[33] Galenson, D. (1981) **White Servitude in Colonial America: An Economic Analysis**, Cambridge University Press

[34] Gallman, J.M. (1980) "Mortality among White Males: Colonial North Carolina," *Social Science History* 4, 295-316

[35] Glaeser, E. L., La Porta, R., Lopez-De-Silanes, F., and Shleifer, A. (2004) "Do Institutions Cause Growth?" *Journal of Economic Growth* 9, 271-303

[36] Goodall, N. (1954) **A History of the London Missionary Society, 1895-1945**, Oxford University Press

[37] Gorodnichenko, Y. and Roland, G. (2017) "Culture, Institutions, and the Wealth of Nations," *Review of Economics and Statistics* 99, 402-16

[38] Gemery, H.A. (2000) "The White Population of the Colonial United States, 1607-1790," in Haines, M. and Steckel, R. (Eds.) **A Population History of North America**, Cambridge University Press

[39] Guerra, F. (1988) "The Earliest American Epidemic: The Influenza of 1493" *Social Science History* 12, 305-25

[40] Gutierrez, H. (1986) "La Mortalité des Evêques Latino-Américains aux XVIIe et XVIIIe Siècles," *Annales de Démographie Historique*, 29-39

[41] Hall, R.E., and C.I. Jones (1999) "Why Do Some Count Others?" *Quarterly Journal of Economics* 114, 83-116

[42] Hardoy, J.E. and Aranovich, C. (1969) "Urbanización en América Hispánica entre 1580 y 1630," Boletin del Centro de Investigaciones Historicas y Esteticas 11, 9-89

[43] Hardoy, J.E. and Aranovich, C. (1970) "Urban Scales and Functions in Spanish America Toward the Year 1600: First Conclusions," *Latin American Research Review* 5, 57-91

[44] Kellaway, W. (1962) **The New England Company, 1649-1776: Missionary Society to the American Indians**, Barned & Noble

[45] Kunitz, S.J. (1984) "Mortality Change in America, 1620-1920," *Human Biology* 56, 559-582

[46] Kuperman, K.O. (1979) "Apathy and Death in Early Jamestown," *Journal of American History* 66, 24-40

[47] Lawson, T. (1840) **Statistical Report on the Sickness and Mortality in the Army of the United States, Jan. 1819 to Jan. 1839**, Jacob Gideon Jr

[48] Lennox, W.G. (1933) **The Health and Turnover of Missionaries**, Foreign Missions Conference

[49] Lockhart, J. and Schwartz, S.B. (1985) **Early Latin America: A History of Colonial Spanish America and Brazil**, Cambridge University Press

[50] Lovett, R. (1899) **The History of the London Missionary Society, 1795-1895**, Oxford University Press

[51] Maloney, W.F. and Valencia Caicedo, F. (2016) "The Persistence of (Subnational) Fortune," *Economic Journal* 126, 2363-401

[52] McDaniel, A. (1995) **Swing Low, Sweet Chariot: The Mortality Cost of Colonizing Liberia in the Nineteenth-Century**, University of Chicago Press

[53] McNeill, W.H. (1976) **Plagues and People**, Anchor Books

[54] McNeill, J.R. (2010) **Mosquito Empires: Ecology and War in the Greater Caribbean, 1620-1914**, Cambridge University Press

[55] Merrens, H.R. and Terry, G.D. (1984) "Dying in Paradise: Malaria, Mortality, and the Perceptual Environment in Colonial South Carolina," *Journal of Southern History* 50, 533-50

[56] Myers, J.B (1892) **The Centenary Volume of the Baptist Missionary Society 1792-1892**, BMS press

[57] Naritomi, J., Soares, R.R., and Assunção, J.J. (2012) "Institutional Development and Colonial Heritage within Brazil," *Journal of Economic History*72, 393-422

[58] Newson, L.A. (1993) "The Demographic Collapse of Native Peoples of the Americas, 1492-1650," *Proceedings of the British Academy* 81, 247-288

[59] Newson, L.A. (2006) "The Demographic Impact of Colonization," in **The Cambridge Economic History of Latin America, Vol 1: The Colonial Era and the Short Nineteenth Century**, Bulmer-Thomas, V., Coatsworth, J.H, Cortés Conde, R. (eds), 143-184, Cambridge University Press

[60] Palma, N. (2022) "The Real Effects of Monetary Expansions: Evidence from a Large-scale Historical Experiment," *Review of Economic Studies* 89, 1593-627

[61] Pascoe, C.F. (1901) **Two Hundred Years of the S. P. G.: An Historical Account of the Society for the Propagation of the Gospel in Foreign Parts, 1701-1900 (Based on a Digest of the Society's Records)**, Volume 2, Published at the Society's Office

[62] Pérez-Artés, M.C. (2023) "Numeracy Selectivity of Spanish Migrants in Colonial America (Sixteenth-Eighteenth Centuries)," *Economic History Review* 77, 503-22

[63] Putterman, L. and Weil, D.N. (2010) "Post-1500 Population Flows and the Long Run Determinants of Economic Growth and Inequality," *Quarterly Journal of Economics* 125, 1627-82

[64] Ritzler R. and Seerin, P. (1952, 1958, 1968), **Hierarchia Catholica Medii et Recentioris Aevi: V (1667-1730), VI (1730-1799), VII (1800-1846)**, Monasterii Sumptibus et typis librariae Regensbergianae

[65] Sequeira S, Nunn N, and Qian N. (2020) "Immigrants and the Making of America," *Review of Economic Studies* 87, 382-419

[66] Shea, J.G. (1854) **History of the Catholic Missions Among the Indian Tribes of the United States, 1529-1854**, Arno Press

[67] Shigesada, N., Kawasaki, K., and Takeda, Y. (1995) "Modeling Stratified Diffusion in Biological Invasions," *The American Naturalist* 146, 229-51

[68] Smith, D.S. (1972) "The Demographic History of Colonial New England," *Journal of Economic History* 32, 165-183

[69] Simon, H.A. (1955) " On a Class of Skew Distribution Functions," *Biometrica* 42, 425-40

[70] Sokoloff, K.L. and Engerman, S.L. (2000) "History Lessons: Institutions, Factors Endowments, and Paths of Development in the New World," *Journal of Economic Perspectives* 14, 217-32

[71] Steele, I.K. (1994) **Warpaths: Invasions of North America**, Oxford University Press

[72] Subramanian, K. (2007) "The Colonial Origins of Comparative Development: New Settler Mortality Data," Undergraduate Honor Thesis, Department of Economics, Amherst College, https://www.amherst.edu/media/view/18871/original/Subramanian.pdf

[73] Traboulay, D.M (1994) **Columbus and Las Casas: The Conquest and Christianization of America, 1492-1566**, University Press of America

[74] Vaughan, A.T. (1975) **American Genesis: Captian John Smith and the Founding of Virginia**, Boston

[75] López de Velasco, J. (1574 [1894]) **Geografía y Descripción Universal de las Indias**, Fortanet

[76] Whitmore, T.M. (1991) "A Simulation of the Sixteenth-Century Population Collapse in the Basin of Mexico," *Annals of the Association of American Geographers* 81, 464-87

7 Appendix A: Settler Mortality Rates

Settler mortality estimates. I construct settler mortality estimates from several historical secondary sources with enough information to carry out demographic analyses. Most of the sources for North America are based on *missionary rolls* for Protestant missions, although I consider Catholic Missions (CM) founded in 1632, primarily from Jesuit priests in New England and Canada. As reported by Croil ([16], p. 167), regarding Protestant missions, the "principal [British] foreign mission societies" and their date of foundation are: New England Company (NEC), 1649; Society for the Propagation of the Gospel in Foreign Parts (SPG), 1701; Baptist Missionary Society (BMS), 1792; London Missionary Society (LMS), 1795; Church Missionary Society, 1799; Methodist Missionary Society, 1817; Church of Scotland, 1824; Free Church of Scotland, 1843; Presbyterian Church of Ireland, 1840; Presbyterian Church of England, 1844; United Presbyterian Church, 1847; China Inland Mission, 1862. I consider the first fourth missions, which are not only the earliest, but also the largest in North America, and the only ones active before the American War of Independence.[25] I also consider Bishops from the Foreign Missionary Society of the Protestant Episcopal Church of the United States (FMS), which is an American society founded after independence. I use the FMS for comparability with the Catholic Bishop data. The details of the relevant sources follow.

◊ *Protestant missionary rolls*: The Protestant missionary denominations studied here are the earliest and largest missionary denominations in North America. They include:

NEC: Established in 1649, The New England Company (NEC) was the oldest English Protestant missionary society dedicated to the conversion of Native Americans in New England (Kellaway [44]). The NEC operated until the American War of Independence, but its operations were relatively small in scale and very geographically concentrated. Kellaway ([44], pp. 289-303) reports the "names of all members of the New England Company elected before 1776 and of all Commissioners appointed before that date" including the entry date to the society as well as the death year.

SPG: The Society for the Propagation of the Gospel in Foreign Parts (SPG) was founded by the Church of England, London 1701, with the purpose of spreading Christianity to British colonies overseas. The SPG was the main promoter of the cause and interests of the Church of England in the American colonies. It was also one of the largest and most influential missionary organizations, as missionaries sometimes served as colonial administrators. In its early years, the SPG sought to teach the gospel to native Americans and European settlers in North America, but over time it expanded to the Caribbean, Africa, India, the Pacific, and other areas leading to over 5,000 mission stations at the close of

[25] As with many other societies, there is a historical account of the Society for Promoting Christian Knowledge (SPCK), in Allen and McClure [6], but there is no missionary roll or information about particular missionaries.

the 19th century. Each station, typically staffed by Anglican priests and missionaries, included a church or chapel, as well as a school and a hospital. To celebrate the two hundred years of the SPG, C.F. Pascoe, the Secretary of the Society from 1905 to 1933, complied two volumes listing comprehensive historical accounts from the founding of the society in 1701 to 1900, including information on the society's leadership, finances, and administration (Pascoe [61]). Of particular interest for this paper, is the *missionary roll* based on "every effort to secure an accurate and complete list" of missionaries obtained from "the careful scanning of hundreds of volumes of printed matter" including personal journals and letters. The missionary roll includes detailed historical records for over 4,000 missionaries. Although not all records are complete, they aimed to provide "full particulars as to birth, education, ordination, location, and death" for all missionaries.

BMS: Records from the Baptist Missionary Society (BMS), founded in 1792 by a group of (Calvinistic) Baptists led by William Carey, are available as part of the celebration of the first 100 years of the society. The BMS was a product of the Second Evangelical Awakening The BMS sent its first missionaries (William Carey and John Tomas) to North India (1793) "to proclaim the message of the redeeming love of Christ to the Hindus." Later work began in Sri Lanka (1812), the Caribbean (1814), Cameroon (1841), China (1859), and the Congo (1879). The missionary roll in Myers ([56], pp. 313-325) contains demographic details about the missionaries sent abroad by the BMS "since the formation of the society" in 1792 until 1892 to India, the Orissa Mission, the Java and Sumatra Mission, Ceylon, China, Japan, Africa, Jamaica, Other West Indian Islands, Honduras, Belize, and Europe.

LMS: The London Missionary Society was an Protestant mission formed "from the union of various denominations" and founded in England in 1795, with the aim of spreading Christianity to regions with little or no Christian presence, Goodall ([36], p. 3). To honor the centenary of its foundation, the two volume work of Lovett ([50], p. vi) provided an "accurate and complete picture of the origin and the administration of the Society, and the great results which have been achieved [...] in so many of the great mission-fields of the world." This included a "complete list of the missionaries who laboured in" many different areas of the world where the LMS was present, including a list "intended to show the year of arrival of missionaries and the length of service;" see Lovett ([50], pp. 794-802, Volume I, and pp. 737-747 in Volume II). Goodall [36] provided an account of the succeeding half century of the LMS, including a list of the missionaries in service since 1895, "based on the most accurate records available," up to 1954.

FMS: The Domestic and Foreign Missionary Society of the Protestant Episcopal Church of the United States (FMS) was an *American* missionary society founded in 1820, driven by growing interest in domestic and foreign missionary work. Most efforts were concentrated among Native American tribes in the western United States, but missionaries reached countries as far as Liberia and China. Julia Emery, Secretary of the Women's Auxiliary to the Board of Missions, provided a comprehensive account of the society's history as part of the centennial celebration of the FMS. In Emery ([25] pp. 399-439), she listed detailed information on the "dates of organization of the different States and Territories," including the names of the Bishops, their dates of consecration as Bishops within the church, and the date of their death. The Bishops in the FMS in Emery [25] were descendants of European settlers instead of European settlers themselves, but the geographic coverage within the US

includes frontier areas. It is also possible to compare the mortality of Protestant Bishops with that of Catholic Bishops discussed below, because both are more selected populations than the populations of missionaries.

◊ *Catholic Missionaries and Bishops*

CM: Leaving aside Franciscans who accompanied Spanish explorers to North America in the 16th century, Jesuit missionaries initiated work in North America in 1632. Shea [66] contains detailed information about Catholic Missionaries (CM) in New England and the Great Lakes region. This includes the Abnaki mission (Maine), the Huron mission (upper Canada, Michigan, and Ohio), the Iroquois mission (New York), the Ottawa mission (Wisconsin and Michigan), the Illinois mission, and the Louisiana mission.

CB: Vatican records compiled as part of the *Hierarchia Catholica* collection contain biographical information for a comprehensive list of bishops, archbishops, cardinals, popes, and other prelates of the Roman Catholic church listed alphabetically by diocese, along with information on their birth, ordinations, appointments, and death. The text is in Latin. Early work by the Franciscan friar and historian P. Conrad Eubel, published in 1898, contained data from 1198 to 1592 (in three volumes), and was based on extensive research of historical documents and other sources (Eubel [28]).[26] Mortality estimates in Gutierrez [40], used as the key input for settler mortality in Latin American in Acemoglu et al. [1], relied on data from Volumes V-VII (Ritzler and Sefrin [64]).

David L. Cheney [14] has digitized and uploaded historical information from the *Hierarchia Catholica* and supplemented it with multiple additional sources such as the *Annuario Pontificio*, an annual directory of the Catholic Church. The digitized data is publicly available at https://www.catholic-hierarchy.org/. His work is not officially sanctioned or approved by any Catholic Church authority, but it aims to provide accurate and up-to-date information on the various bishops, cardinals, and other leaders of the Catholic Church around the world, up until today. Before the 20th century, the data covers over 10,000 European-born bishops appointed to numerous dioceses around the world.

Robustness of survival curves. To assess the robustness of the survival curves, I present Cox hazard estimates of the figures reported in the text, as well as supplementary findings of auxiliary interest.

Table A1 reports Cox hazard estimates associated with Figures 3 and 4. For the CM sample, Panel A, controls include an indicator for the Midwest, a control for the year of arrival (normalized for convenience), and an indicator for missionary killings as the cause of death. In the absence of controls (as Figure 3), Table 5, Column (1) suggests that Catholic missionaries in the Midwest were about twice as likely to die per year of service compared to Catholic missionaries in New England, i.e., $\exp\{\beta_i\} = 1.96$. Controlling for arrival year also

[26] The work was originally intended to comprise the period of 1198 to 1431 alone, with a chronology of Cardinals and Popes only, but there are now several volumes edited by different authors and periods of time with a much broader coverage that includes bishops since the second volume. The volumes are based on the following chronology (I) 1198-1431, (II) 1431-1503, (III) 1503-1592, (IV) 1592-1667, (V) 1667-1730, (VI) 1730-1799, (VII) 1800-1846, (VIII) 1846-1903. (The origin date is 1198 because the regular and uninterrumpted collection of Vatican records began under the Pope Innocent III, 1198-1216.)

suggests a doubling in the risk of death in the Midwest, or $\exp\{\beta_i\} = 2.24$ in Column (2), although the increased value of the Midwest indicator and the positive trend in mortality are likely confounded by fact that late arrivals took place in the Midwest rather than in New England. For instance, there is no observable trend in mortality in Column (3). Column (4) suggests that the higher prevalence of missionary killings, some of which were associated with disruptions caused by the introduction of European diseases, does not fully account for the increased mortality risk in the Midwest.

For the NEC sample, Panel B of Table A1, considers an indicator for early arrivals (as in Figure 4), a continuous control for the arrival year, and an indicator for being a *Commissioner* of Indian Affairs, associated with higher social status and economic conditions. Table 5, Column (5), suggests that early arrivals faced a higher risk of death in New England of about 36 percent, or $\exp\{\beta_i\} = 1.36$, relative to late arrivals. Commissioners experienced a reduced risk of death, but controlling for Commissioner status in Column (6) does not alter the estimate of early arrivals. Using a continuous year of arrival as control also suggests a downward trend and a reduced morality risk for Commissioners, Columns (7) and (8).

Table A2 reports Cox hazard estimates for differences in mortality across bishops in the FMS, relative to New England and the Mid-Atlantic colonies (the omitted categories). Column (1) reports estimates consistent with Figure 6. While Column (1) suggests a higher relative risk of death in each of these regions, none of the estimates are statistically significant. The latitude and longitude of the State where the bishops resided (not of the regions), added as controls in Column (3), also lack significance. Only appointment year is statistically different from zero. This estimate suggests an increasing risk of death for bishops appointed at later dates. As discussed in the analysis of Table 6, this reflects the fact that appointment year and region are highly correlated. Bishops moved to areas with higher mortality risks as time went on.

Table A1. Cox model estimates across regions and time in North America.

	A. Survival in CM					B. Survival in NEC			
	(1)	(2)	(3)	(4)		(5)	(6)	(7)	(8)
Midwest	*1.96*	*2.24*		*1.78*	<1710	*1.36*	*1.32*		
	(0.44)	(0.54)		(0.45)		(0.16)	(0.16)		
Arrival year		*1.03*	1.00	*1.04*	Arrival year			*0.95*	*0.95*
		(0.02)	(0.02)	(0.02)				(0.01)	(0.01)
Killed				*2.91*	Commissioner		*0.76*		*0.78*
				(0.91)			(0.10)		(0.11)
N. Spells	91	91	91	91	N. Spells	263	263	263	263

Notes: Robust standard errors in parentheses. The variable *<1710* is an indicator for arrival to North America prior to that year; *Midwest* is an indicator for Catholic Missions (CM) in the Great Lakes region (including Louisiana); *Arrival year* represents the year of arrival of the missionary relative to 1492, in decadal units (i.e., divided by 10); *Killed* is an indicator for the cause of death in the NEC; and *Commissioner* is an indicator for being a Commissioner for Indian Affairs in the NEC.

Table A2. Cox model estimates across regions and time in US, FMS.

	(1)	(2)	(3)	(4)
Midwest	1.17			0.99
	(0.24)			(0.30)
South	1.14			0.89
	(0.21)			(0.24)
West	1.22			0.82
	(0.34)			(0.47)
Appointment year		*1.15*		*1.16*
		(0.04)		(0.04)
Latitude			0.56	0.37
			(0.39)	(0.38)
Longitude			1.59	0.98
	(0.34)		(0.88)	(1.15)
N. Spells	168	168	168	168

Notes: Robust standard errors in parentheses. *Appointment year* represents the year of appointment as bishop relative to 1492, in decadal units (i.e., divided by 10); *Latitude* and *Longitude* are based on the State where the bishops resided, not the regions used as indicators.

Tropical areas exhibit higher settler mortality, relative to temperate areas in the 19th century. For example, Figure 12 plots the survival curves for Baptist missionaries in Central America (i.e., Jamaica, West Indies, and Belize), and other areas such as sub-Saharan Africa, India, and China (including Japan, but with very few observations, as shown in Table 3). The figure shows that by the mid- to late-19th century, mortality conditions in Central America were not different from those in other tropical places, with the exception of sub-Saharan Africa where missionaries faced a considerably higher risk of death than anywhere else in the world.

Table A3 reports Cox hazard estimates for differences in mortality across bishops in the BMS, relative to Europe (the omitted category). It considers mostly tropical locations in Central America, India, and Africa. There are very few observations for China (which includes Japan). Column (1) reports hazard estimates consistent with Figure 12, and suggest a four-fold higher mortality risks in sub-Saharan Africa. Column (2) shows an increasing trend in mortality, but a positive trend is not significant once regional indicators are included. Columns (3) to (5) add controls that include a priest status, and an indicator for women. While in both cases there is a reduced risk of death, these effects are not significant. The indicator for sub-Saharan Africa is stable across specifications and suggests a considerably higher risk of death relative to any other world region.

Figure 13 plots the survival curves for missionaries from the London Missionary Society in the Caribbean (mostly Jamaica), South Africa, India, China and a collection of South Pacific Islands (i.e., Tahiti, Bora Bora, the Marquesas). While survival curves in the Caribbean are below the other regions, there was no statistical difference with the South Pacific Islands. Table A4 reports Cox hazard estimates for differences in mortality across missionaries of the LMS, relative to South Pacific islands (the omitted category). Column (1) includes only the regional controls, while Column (2) includes only a control for the date of acceptance. Column (3) includes both sets of controls.

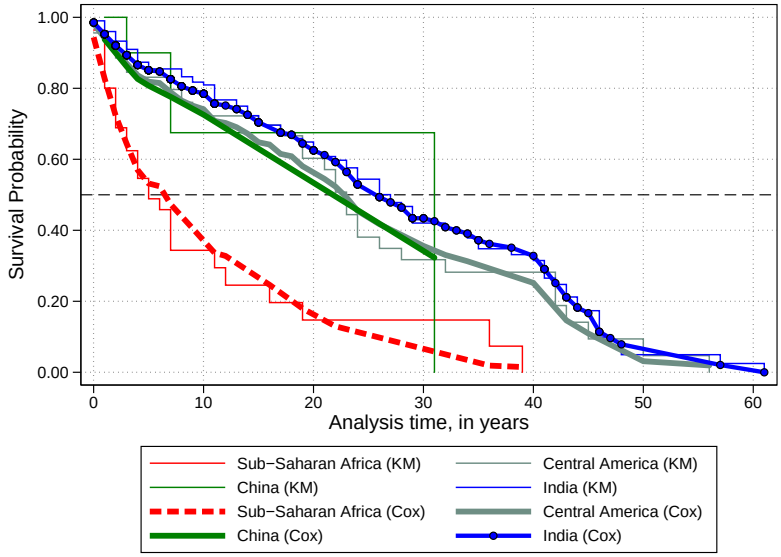


Figure 12: Kaplan-Meier and baseline Cox hazard survival curves for missionaries from the Baptist Missionary Society (BMS), 1793-1894, in different tropical regions.

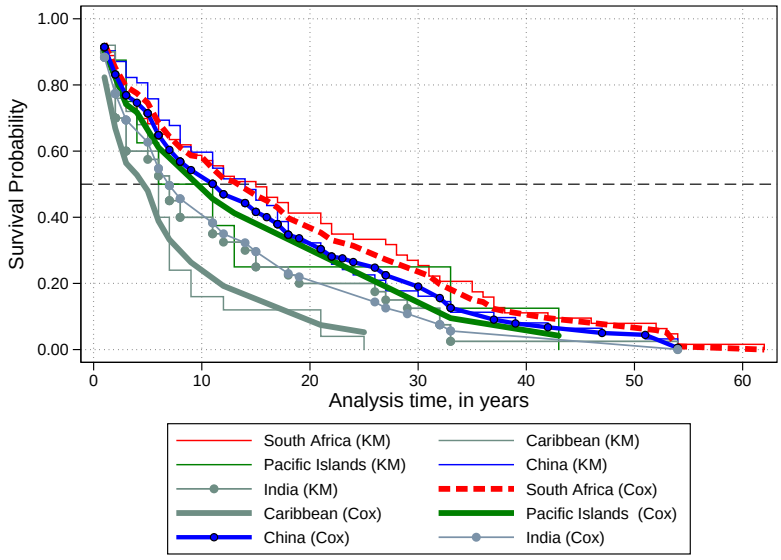


Figure 13: Kaplan-Meier and baseline Cox hazard survival curves for missionaries from the London Missionary Society (LMS), 1795-1895, in different tropical regions.

Table A3. Cox model estimates across world regions, BMS.

	(1)	(2)	(3)	(4)	(5)
Central America	1.41		1.45	1.46	1.46
	(0.87)		(0.90)	(0.90)	(0.90)
India	1.14		1.18	1.19	1.19
	(0.68)		(0.71)	(0.72)	(0.72)
China	1.50		1.50	1.50	1.50
	(1.10)		(1.11)	(1.11)	(1.10)
Africa	*4.46*		*4.43*	*4.42*	*4.47*
	(2.80)		(2.80)	(2.79)	(2.83)
Acceptance year		*1.10*	1.01	1.01	1.01
		(0.04)	(0.04)	(0.04)	(0.04)
Pastor				0.76	0.76
				(0.76)	(0.76)
Woman					0.70
					(0.79)
N. Spells	367	367	367	367	367

Notes: Robust standard errors in parentheses. *Acceptance year* represents the year of acceptance for service relative to 1492, in decadal units (i.e., divided by 10); *Pastor* and *Woman* are indicators variables.

Missionaries in South Africa had a robust statistically significant reduced risk of death per year of service. Missionaries in China had an increased risk, but the increased hazard becomes insignificant once acceptance date is added as control. Adding acceptance date suggest a reduced risk for Indian missionaries, which was insignificant but on the margin for significance in Column (1). An advantage of the LMS data is that it reports close to 50 women missionaries. An indicator for these missionaries is included in Column (4). It shows an increased three-fold risk of death for women Column (5) includes the acceptance date and the regional controls. While the magnitude of the estimated coefficient declines, there is a still a considerably higher statistical risk of death per year of service for women in the LMS (Table A4).

Table A5 reports estimates that rely on the same specifications of Table 6, Panel A, but with the addition of *age at arrival* as a control. Given the limited availability of birth year information in the SPG, sample sizes for North American missionaries drop considerably, from 354 to 113. Column (1) includes the indicator for early arrivals in the new sample, without any other covariate. The estimated coefficient β_i is not statistically different from zero, which is expected because of the smaller sample variation. For instance, because birth year information is less common in early records. Column (2) shows that age has a strong influence on the risk of death per year of service. An extra year of age at arrival increase the risk of death by about 6 percent. This specific value will later be compared with the age at arrival estimate for Catholic Bishops. Although regional differences in Column (3) are smaller than in Table 6, the indicator for New England still suggests a more favorable disease environment. The higher favorability of New England, the positive influence of age at arrival, and the reduced risk associated with being a Priest are still evident in Columns (4) and (5).

Table A4. Cox model estimates across world regions, LMS.

	(1)	(2)	(3)	(4)	(5)
Caribbean	0.81		0.72		0.72
	(0.32)		(0.27)		(0.27)
South Africa	*0.62*		*0.60*		*0.60*
	(0.13)		(0.12)		(0.12)
China	*1.56*		1.14		1.14
	(0.34)		(0.25)		(0.25)
India	0.71		*0.65*		*0.64*
	(0.14)		(0.12)		(0.12)
Acceptance year		*1.01*	*1.01*		*1.01*
		(0.00)	(0.00)		(0.00)
Woman				*3.34*	*2.19*
				(1.68)	(0.90)
N. Spells	198	198	198	198	198

Notes: Robust standard errors in parentheses. *Acceptance year* represents the year of acceptance for service relative to 1492, in decadal units (i.e., divided by 10); *Woman* is an indicator variable.

Table A5. Age effects in Cox model estimates for SPG in North America.

	A. Differential survival in North America: SPG				
	(1)	(2)	(3)	(4)	(5)
Age at arrival		*1.06*		*1.06*	*1.05*
		(0.01)		(0.01)	(0.01)
North America <1750	1.09	0.95			1.68
	(0.26)	(0.23)			(0.54)
Southern colonies			1.47	0.88	0.39
			(0.98)	(0.47)	(0.31)
Middle colonies			1.00	1.09	0.69
			(0.26)	(0.30)	(0.26)
New England			*0.53*	*0.54*	*0.30*
			(0.12)	(0.10)	(0.11)
Educated					1.36
					(0.33)
Priest					*0.56*
					(0.13)
Latitude					0.98
					(0.04)
Longitude					0.98
					(0.01)
N. Spells	113	113	113	113	113

Notes: As in Table 6.

Table A6 includes age at arrival in the same specifications as Table 7, Panel A. Again, as birth information is limited in the SPG data, sample sizes are considerably reduced, from

693 to 207. Column (1) shows that, overall, North America had a more favorable disease environment than Central America. It also shows that mortality conditions in Central America were not statistically different from mortality in other areas. Column (2) maintains these findings unaltered, and shows that the age of a missionary at arrival increased the mortality risk by about 4 percent, as in Table A3. Columns (3) and (4) examine temporal and spatial differences in mortality. As in Table 7, they show that while mortality conditions improved in North America over time, settler mortality worsened over time in Central America. This is the case even after controlling for the latitude and longitude of the missionary station, as well as for whether the missionary was educated or a priest. Finally, as in Table 7, Column (5) shows that early disease environments in Central America were as favorable to European settlers as the disease environment of Australia.

Table A6. Age effects in Cox model estimates for SPG across regions.

	A. Differential Survival for North America				
	(1)	(2)	(3)	(4)	(5)
Age at arrival		*1.04*	*1.04*	*1.05*	*1.05*
		(0.00)	(0.00)	(0.00)	(0.00)
North America<1750			0.97	1.06	1.17
			(0.22)	(0.26)	(0.30)
North America	*0.41*	*0.41*	*0.40*	*0.43*	0.73
	(0.06)	(0.06)	(0.06)	(0.18)	(0.34)
Central America<1850			*0.26*	*0.26*	*0.28*
			(0.08)	(0.08)	(0.10)
Central America	0.88	1.03	*2.00*	*2.88*	*7.37*
	(0.23)	(0.25)	(0.60)	(1.38)	(4.84)
Australia					*0.24*
					(0.17)
Added controls	No	No	No	Yes	Yes
N. Spells	207	207	207	206	206

Notes: As in Table 7.

The SPG data report some causes of death for a few missionaries. I added a control for the listing of "drowning" and "smallpox" as causes of death, but do not report the results here, as they are virtually unchanged with the baseline estimation. The hazard coefficients on drowning is 3.06, and that on smallpox is 7.08. Both are statistically significant effects, but none of the alters the estimates presented in the text or the Appendix.

Figure 14 plot the survival estimates for Catholic bishops up until 1920, by region. As noted in the text, while there are minor differences between Central and South America, survival in North America was closer to the survival in Europe, and higher than in the other regions of the New World.

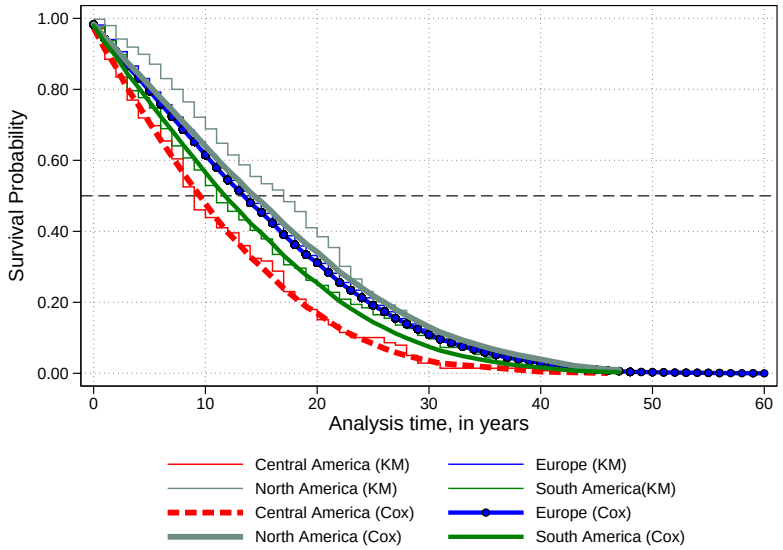


Figure 14: Kaplan-Meier and baseline Cox hazard survival curves for Catholic Bishops (CB) appointed between 1080-1920, in different regions.

Table A7. Cox model survival estimates for Catholic Bishops.

	Appointment <1920		
	(1)	(2)	(3)
Central America	*1.53*	*1.50*	*1.46*
	(0.13)	(0.13)	(0.14)
South America	*1.18*	*1.16*	*1.23*
	(0.08)	(0.07)	(0.07)
Africa	0.86	1.00	1.17
	(0.07)	(0.10)	(0.13)
Asia	*0.91*	*1.09*	*1.23*
	(0.04)	(0.05)	(0.05)
North America	*0.92*	1.02	*1.20*
	(0.03)	(0.04)	(0.05)
Age		*1.05*	*1.05*
		(0.00)	(0.00)
Appointment year			*0.98*
			(0.00)
N. Spells	7757	7757	7757

Notes: As in Table 8.

Diffusion and settlement. Consider the derivation of the range expansion and time relationship in the text. Let $S(x, y, t)$ denote the density of settlers as a function of time and space (x, y), and S_0 the initial number of settlers. Under a short-distance dispersal, the diffusion equation from a spatial random walk yields a fundamental solution of the form $S^*(x, y, t) = S_0 \exp\{\mu t - (x^2 + y^2)/4Dt\}/4\pi Dt$. The settlement frontier is the contour

$S^*(r,t) = [S_0/\sqrt{4\pi Dt}]^{-1} \exp\{\mu t - r^2/4Dt\}$, where r is the radial distance from the origin. For a small value $\bar{S}$ of settlers, $r(t)^2/4Dt = \mu t + \ln[S_0/\bar{S}4\pi Dt]$. After a certain period, as t increases, a linear range-versus-time relationship, such as (1) emerges. Additional terms beyond simple exponential growth that induce saturation or faster-than-linear expansion are not relevant in the early settlements, but may be incorporated as discussed by Shigesada et al. ([67], eqs. 11a-11-c).

Table A8. Relationship between distance (in kilometers) and settlement date.

	A. More than 100 *vecinos*			B. Any settlement size		
	Central America	South America	All areas	Central America	South America	All areas
Settlement date	*43.3*	*91.2*	*78.3*	*27.4*	*62.3*	*47.0*
	(15.3)	(32.0)	(14.8)	(8.9)	(17.1)	(8.3)
R^2	0.39	0.31	0.41	0.19	0.16	0.23
N Obs	23	30	53	59	58	117
Settlement date (expanded)	*16.9*	−3.67	*9.7*	*23.3*	8.3	*20.4*
	(4.13)	(8.74)	(5.13)	(3.3)	(7.4)	(3.8)
R^2	0.32	0.01	0.03	0.28	0.01	0.11
N Obs	44	60	104	107	112	219

Notes: Robust standard errors in parentheses. Values in italics denote statistical significance at the 90 percent or more. Settlement dates from Hardoy and Aranovich [42]. The expanded settlement date is based on dating of settlement (i.e., Spanish foundation) in settlements listed but not dated by Hardoy and Aranovich [42].

Table A8 estimates a linear range-versus-time relationship (1) across several specifications. The table presents separate estimates for settlements in Central America, which includes Mexico and the Caribbean islands, and South America (which excludes Panama). It considers the dates reported by Hardoy and Aranovich [42], but also expanded dates based on settlement or foundation dates for undated cities listed in Hardoy and Aranovich [42]. The expanded set of settlement dates includes smaller settlements compared to those in Hardoy and Aranovich [42], but also settlements at relatively later dates. The expanded sample essentially doubles the sample size of Hardoy and Aranovich [42].

The first panel, Panel A, considers Spanish settlements with more than 100 vecinos in 1580. The fitted line in Figure 1 is based on the Hardoy and Aranovich [42] sample in this panel. Overall, Table A8 shows that a random diffusion is able to account well for the settlement patterns in Central America, but not for South America. The narrow corridor in the Andes in South America is less likely to agree with the assumption of a homogeneous environment, as considered by the invasive species models.

Settlement size. A second important relationship is that between the size distribution of settlements and their rank. Table A9 reports estimates of Zipf's Law that relate log-population size with log-rank. It considers several specifications for the different areas and data sources. In all estimates, this relationship is negative and statistically significant. In the 1580s, and for Hardoy and Aranovich [42] data, the point estimate of the relationship equals one. The strong negative relationship is visible even in the original data reported by López de Velasco [75] for the 1580s and by Vasquez de Espinosa [27] for the 1630s.

Table A9. Relationship between log settlement size and rank.

	A. Hardoy and Aranovich [42]			B. Expanded data			C. Original data
	Central America	South America	All areas	Central America	South America	All areas	All areas
Log rank, 1580	*−1.00*	*−1.05*	*−1.03*	*−1.13*	*−1.17*	*−1.15*	*−1.25*
	(0.05)	(0.05)	(0.03)	(0.06)	(0.05)	(0.04)	(0.05)
R^2	0.95	0.95	0.95	0.94	0.95	0.94	0.93
N Obs	59	58	117	96	93	189	178
Log rank, 1630	*−1.07*	*−1.13*	*−1.10*	*−1.17*	*−1.30*	*−1.24*	*−1.31*
	(0.02)	(0.04)	(0.02)	(0.04)	(0.05)	(0.04)	(0.04)
R^2	0.97	0.95	0.96	0.94	0.87	0.89	0.90
N Obs	34	41	75	70	92	162	146

Notes: Robust standard errors in parentheses. Values in italics denote statistical significance at the 90 percent or more. Hardoy and Aranovich [42] includes data with known settlement dates. The expanded date Panel B does not consider the date of the settlement. Panel C uses the original data reported by López de Velasco [75] for 1580 and Vasquez de Espinosa [27] for 1630.

8 Appendix B: Frontier Settlements

Frontier settlements in Spanish America. Hardoy and Aranovich [42] compiled information from multiple colonial sources, including data from Juan López de Velasco, who gathered detailed information about the Spanish provinces in the New World and the Philippines, and compiled it in his *Geografía y Descripción Universal de las Indias* (López de Velasco [75]) officially published in 1894. The information reported by López de Velasco was obtained from detailed questionnaires sent to the viceroyalties regarding physical and political geography, native governments, land productivity, and names of the different settlements. All the information was unified in his report.

López de Velasco established the geographic limits of the Spanish domains, provided a description of native populations and basic facts about the "discovery" of the New World, reported the flora and fauna of the different regions, and discussed administrative aspects of the different settlements including the number of *vecinos*, which is the key object of study in Hardoy and Aranovich [42]. The book also provided the navigational currents that ships leaving Spain must follow to reach the different ports and details on the geographic locations of Spanish settlements based on longitude calculations from lunar eclipses. Although finished and presented to the King of Spain in 1574, his work remained confidential and was not published until 1894. Among the many details reported, López de Velasco ([75], pp. 11-12) noted the overall healthiness of the New World, attributing it to the absence of European diseases: "adonde los hombres de ordinario viven más sanos y libres de enfermedades que en estotro Nuevo Mundo antiguo." He also credited healthy diets associated with abundant food: "la abundancia de frutas y otros alimentos que se han hallado en la tierra." Finally, he observed deteriorating health conditions in some tropical areas due to worsening diseased environments: "aquellas partes hay algunas que no son ya tan sanas para los españoles ni

los indios como al principio." López de Velasco [75] also reported information about the number of *vecinos* who were *encomenderos*, the number of indian towns, and the number of tributary indians. The number of tributary indians was often preceded by caveats about the number of indians in hiding, hence his accounts are likely to underestimate the size of native populations.

Table B1 reports estimates of the relationship between the numbers of *vecinos*, *encomenderos*, no *encomenderos*, and the number of tributary indians. Information about encomenderos and tributary indians is incomplete, but López de Velasco [75] often reports the absence of *encomenderos* or indians. As there are considerable zeroes in his report, I provide estimates on levels rather than in logs. I also scale the number of tributary indians per 1000.

Table B1. Determinants of Spanish settlement types in López de Velasco [75].

	Vecinos		*Encomenderos*	Non-*encomenderos*	
	(1)	(2)	(3)	(4)	(5)
Tributary indians	*11.04*	*8.06*	0.68	*7.37*	*6.08*
	(3.14)	(1.87)	(0.41)	(1.56)	(1.19)
Encomenderos					*1.89*
					(0.63)
R^2	0.20	0.45	0.09	0.59	0.65
N Obs	96	55	55	55	55

Notes: Robust standard errors in parentheses. Values in italics denote statistical significance at the 90 percent or more. Regressions in levels, with tributary indians scaled on per 1000 basis.

The main message of Table B1, Column (1), is that Spanish settlers were driven to areas with large native populations, as it was also shown in the text for different specifications and large native cities. In Column (2), I restricted the sample to areas with information about the number of *encomenderos*, to make comparisons with additional specifications possible. Column (2) also agrees with the importance of native populations as determinants of Spanish settlements. Column (3) shows that the number of *encomenderos* and the number of tributary indians were weakly associated. The reason is that the *encomienda* was largely a rural institution, and most native populations were concentrated in medium to large cities. For instance, Column (4) shows that the association between the number of *vecinos* and tributary indians was essentially driven by non-*encomenderos*. In other words, Spanish settlers located in areas with large native population, but not to directly benefit from receiving *encomiendas*. In fact, Column (5) shows that the presence of *encomenderos* made non-*encomenderos* more likely without drastically diminishing the quantitative significance of the size of native populations.

Hardoy and Aranovich [42] also rely on reports from Fray Antonio Vasquez de Espinosa based on personal accounts as he worked and traveled across the different *audiencias* of Central and South America between 1612 and 1621. His work, *Compendio y Descripción de las Indias Occidentales* was rediscovered in the 1930s (Vasquez de Espinosa [27]). The book was found in the 1930s in the Barberini library in Rome while marked as "anonymous," but its authorship was verified; see the prologue of Vasquez de Espinosa [27] by Charles Upson Clark (who provided the English translation of the book published in 1942). The accounts

are mostly organized geographically, and consist of detailed descriptions of the geography, commercial, industrial, political, and social movements of the different *audiencias*. Vasquez de Espinosa [27] also provided rich accounts of life in the New World including a long list of medicinal plants. There are no references to high settler mortality rates other than particular instances such as the military expedition to Florida in 1539, the expedition to the Marañon river in Peru, and a mortality crisis in the island of Trinidad and in Guyana in 1595; Vasquez de Espinosa ([27] p. 52). Vasquez de Espinosa ([27] p. 108) actually praised the healthiness of Jamaica, a Spanish possession until in 1655.

Frontier mortality in the US. For context, Table B2 reports summary statistics for military mortality rates in Acemoglu et al. [1] from their original files.

Table B2. Settler mortality across countries.

	Mean	Std. Dev.	Min.	Max.	Obs.
Mortality Rate (all)	156.12	155.26	2.55	668	87
Mortality Rate (ex-colonies)	76.77	45.39	8.54	170	40
Latitude (all)	0.18	0.13	0.00	0.66	87
Latitude (ex-colonies)	0.21	0.13	0.01	0.66	40

Notes: The data from Acemoglu et al. [1] is taken from the authors' replication files with corrections for typos associated with mortality rates that exceed 1,000.

Table B3 provides OLS estimates of the influence of latitude and longitude on military mortality in US forts. I do not include altitude because there are a few forts with missing data, thus lowering sample sizes. For those with data, altitude had no influence in frontier or non-frontier fort mortality. The most important observation of the table is that latitude had a strong negative effect on mortality, regardless of the mortality estimate. Longitude does matter for military returns, indicating that forts further into the frontier had higher mortality than forts closer to the Atlantic coast.

Table B3. Relationship between military mortality and geographic characteristics.

	Log-military mortality rate 1822-1825			Log-mortality rate, 1819-1839					
				Military returns			Medical returns		
	FR	No FR	All	FR	No FR	All	FR	No FR	All
Latitude	*−7.29*	*−7.71*	*−7.99*	*−7.82*	*−11.94*	*−7.37*	*−9.08*	−11.42	*−8.79*
	(2.68)	(2.94)	(2.31)	(1.33)	(5.55)	(1.24)	(1.56)	(6.83)	(1.43)
Longitude	−1.48	2.66	−1.70	*4.30*	−4.01	0.50	3.29	−1.93	−0.45
	(2.87)	(3.42)	(1.13)	(1.83)	(6.01)	(0.91)	(2.19)	(8.14)	(1.06)
R^2	0.31	0.46	0.32	0.65	0.34	0.48	0.64	0.34	0.48
N Obs	23	20	43	25	20	45	25	20	45

Notes: Robust standard errors in parentheses. FR is short for frontier. Values in italics denote statistical significance at the 90 percent or more. The different mortality rates have been referenced in the paper.

Projections. I also construct alternative estimates for differences in fort mortality across US counties. The text relies on military mortality from 1822-1825 published as a

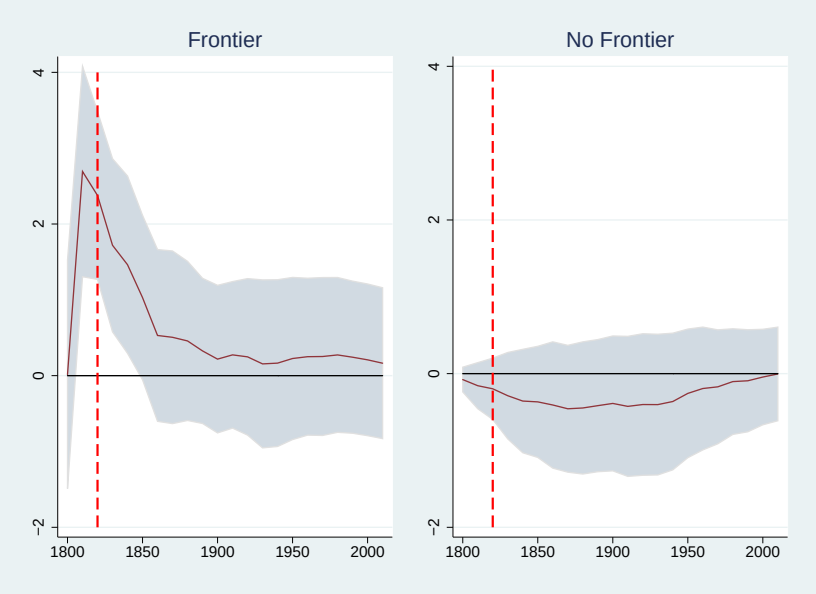


Figure 15: Mortality from military returns, 1819-1838, and log-population density in frontier and non-frontier settlements.

single list in Lawson [47], but it is possible to compile the medical and military returns for 1819-1838 by going over each individual fort. Figures 15 and 16 plot the same impulse-response as in the text, but with the alternative military and medical returns statistics. As in the text, mortality had a positive and short-lived association with settlements in the American frontier.

The 1790's US Census had some limitations associated with coverage and data incompleteness. I consider an alternative definition of "frontier" areas by looking at US counties with less than 2 people per square mile in the 1800 US Census. I consider local projections

$$\ln S_{j,t} - \ln S_{j,1800} = \phi_s[\text{Frontier}_j \times \ln(\text{Fort Mortality Rate}_j)] + \gamma_t X_j + \xi_{j,s},$$

where Frontier_j and $\ln(\text{Fort Mortality Rate}_j)$ are part of the controls and the coefficients ϕ_t have the same interpretation as in the text. They reflect the long-term response to mortality differences in the population settlement of frontier areas. Although I considered population densities, I also added the county's land area as a control in the previous regressions. Figure 17 plots the impulse-response estimates to differences in military mortality in frontier areas only. As in the previous figures, military mortality exerted a positive influence in the settlement of the American frontier.

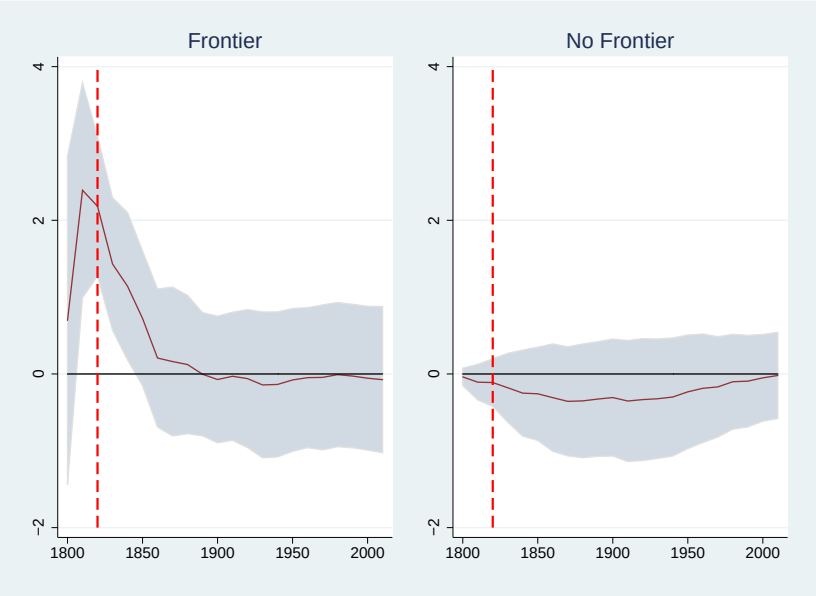


Figure 16: Mortality from medical returns, 1819-1838, and log-population density in frontier and non-frontier settlements.

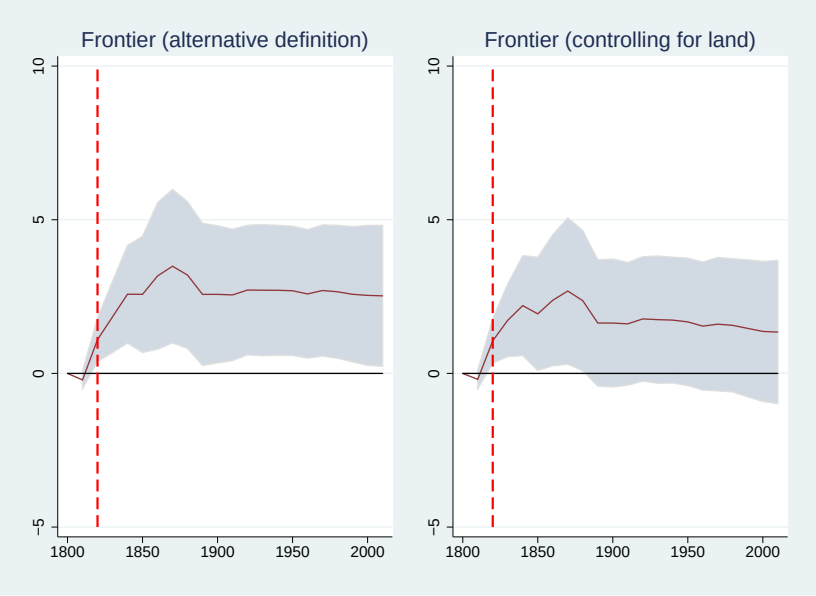


Figure 17: Frontier settlements and military mortality, 1819-1838, under alternative definition of frontier areas, and additional control for land.