Disease and Urban Development

(Preliminary and Incomplete)

by

Douglas Hanley, Economics, University of Pittsburgh

Shawn McCoy, Economics, UNLV

and

Werner Troesken, Economics, University of Pittsburgh

1. Introduction

Arguably the single greatest change in urban life over the past two hundred years has been a near doubling in life expectancy in cities throughout the Western World and the elimination of the urban mortality penalty. As late as 1900, overall mortality in American cities was 20 to 60 percent than in non-urban places, and it was not until 1940 that mortality rates in urban and non-urban places were equalized. It is well known improvements in public health and disease control drove these reductions in urban mortality ( ). Although eighteenth and nineteenth-century observers believed that high disease rates and poor sanitation deterred urban in-migration, there is very little direct evidence on the role public health improvements played in stimulating urban growth and the evidence that does exist seems to suggest that the effects of such improvements on growth could not have been very large. First, the economic history literature indicates that workers in cities were fully compensated for the disamenities associated with increased mortality risk ( ). If so, the eradication disease might not have resulted in sharp increases in population levels or growth rates. Second, and along these lines, if one looks at population growth in cities in the era before effective public health measures, casual empiricism would seem to suggest disease was of second-order importance. London and New York, for example, were havens of high mortality and poor health, and yet experienced rapid rates of population growth during the pre-public-health era.

Accordingly, in this paper, we formally explore the relationship between disease and urban development. The analysis begins with a model locational choice that highlights how disease inhibits urban growth in a world without effective public health measures: places with high productivity attract more people, but in a world without effective disease control, in-migration to high productivity places also generates higher disease rates, which discourages further in-migration. More precisely, the utility flow generated by any given locality is a positive function of productivity (a constant) and a negative function of disease costs. Disease costs are the result of a congestion externality so that the cost of disease rises with the population of the locality. The costs of disease, however, can be lessened by a disease mitigation technology or any other shock (such as a disruption in trade or a change in climate) that alters the disease propensity of the locality. Agents also pay a cost to relocate to a new place, and that cost is an increasing function of the number of people migrating out of a locality. In the model's steady state, utility flows are equalized across localities and places with higher productivity levels will have higher populations and larger realized congestion effects in the form of higher population-induced disease costs.

The model yields three sets of testable predictions. First, in a world where it is not possible to mitigate disease through technology, high productivity places would have higher population levels and higher disease rates than low productivity places. Second, the introduction of an effective disease mitigation technology in all cities, would induce movement away from previously low-disease/low-productivity places toward high-productivity places. To the extent that such movements are not instantaneous, this would result in both higher population levels and high population growth rates in places with high productivity. Put more simply, one expects population growth rates across places to diverge (at least temporarily) after

the discovery and implementation of effective disease mitigation technologies, with populations relocating to high-productivity/high-population places because the population-induced costs of disease have fallen. Third, the general introduction of a disease mitigation technology would have a relatively small impact on cities that were, for whatever reason, relatively immune to disease.

We test these predictions by assembling a panel data set of American counties with time-consistent borders and exploring how population growth was related to the rise and fall yellow fever. Poorly understood and controlled for much of the nineteenth century, yellow fever (which was spread by a mosquito) would erupt suddenly after years of lying dormant, often killing 5 to 25 percent of a local population within a few month's time. Although the sporadic and intermittent nature of yellow fever rendered it a trivial cause of death in overall death counts for the nineteenth century, historians have long claimed that its frightening Ebola-like symptoms and its potential to cause sudden widespread death discouraged in-migration and undermined urban growth. Qualitative historical accounts suggest high yellow fever places were also high productivity places. In particular, yellow fever tended to strike cities with geographic features, such as a coastal location, that rendered those cities focal points for trade, migration, and economic activity, which in turn, brought yellow fever.

Consistent with the predictions of the model, the formal econometric evidence shows that before the onset of effective strategies to control and prevent the disease, yellow fever afflicted high-population and high-productivity counties more so than small ones. Population growth was also slower in high-population counties than in low-population ones so that counties were converging in their population and density levels. Moreover, once exogenous shocks in trade and scientific knowledge made it possible for local governments to control and prevent yellow fever, the disease subsided and population levels in large counties began to diverge from those in small counties. Finally, the general introduction of technologies mitigating yellow fever had no impact on populations in places that were, because of their geography and climate, invulnerable to yellow fever. These patterns are robust to a variety of potential confounders, including the arrival of the railroad, unobserved time-varying shifts in the value of fixed county-level characteristics, and regional variation in the timing of effective yellow fever controls.

The model and results presented here contribute to three literatures. First, over the past half century, economists have grown increasingly interested in how disease influences long term economic growth and development. The relevant literature is vast and evolving. Initial studies suggest that diseases like malaria undermine economic development and explain the slow growth rates observed in equatorial places (Sachs, etc). Later studies, however, present strong evidence that, at least at a macro level, disease has little or no effect on long-term macroeconomic growth. For example, . . . More recent studies go even further and suggest that the Black Death might have stimulated long term growth. One puzzle in all of this is that there is a large body of well-identified work in applied microeconomics showing that at least at the individual level, exposure to disease in early life can have serious adverse effects on long-term economic outcomes (Almond, Bleakley). Given the current state of the literature, it is not

immediately obvious how to reconcile this micro-level evidence with the prevailing macroeconometric evidence.

Second, there is strong evidence that investments in public health and disease control played a central role in explaining the urban mortality transition the large increases in life expectancy over the course of the late-nineteenth and early-twentieth century. Most of this work focuses on public water and sewer systems, and suggests a large rate of social return on these investments. Despite this evidence, we know almost nothing about how improved sanitation and public health altered the locational choices of individuals and overall urban population growth. The results here indicate that, in at least one setting, the eradication of disease had large effects on locational choices and population growth.

Third, the American South has long lagged behind the North in economic performance, and only after World War II did incomes begin to converge. Until 1940, income per capita in the South was 45 to 60 percent of the U.S. average; by 1980, that deficit had been reduced to 80 to 95 percent (Wright 1987). Standard explanations for these patterns fall into one of three categories. Institutional explanations consider national labor standards (Wright 1987); Civil Rights legislation (Wright 2013, Collins 2003); and the decline of paternalism and other institutions hostile to black economic progress (Alston and Ferrie 1993, 1999). Technological explanations focus on air-conditioning (Biddle 2008, 2012), electrification (Downs 2014), and the mechanization of agriculture (Alston and Ferrie 1993, 1999). Disease-based explanations consider the eradication of hookworm and malaria (Bleakley 2007, 2010; Kitchens, 2013). The results provide further evidence on the importance of disease, and implicate yet another disease (i.e., yellow fever) in slowing Southern economic development. Although historians have long claimed that yellow fever disrupted trade and discouraged migration to Southern cities, ours is the first paper to formalize and test such beliefs.

2. Model

See model attached to end of paper

3. Yellow Fever: An Overview

Although largely unheard of today, on a city-by-city level, yellow fever was responsible for the worst epidemics in American history. In 1878, a yellow fever epidemic in Memphis killed 1 of every 8 city residents. The death toll in Memphis would have been even higher had two-thirds of the white population not fled the city (Humphreys 1992, p. 5; Wrenn 1987). In 1853, a severe yellow fever epidemic affected cities and towns throughout Louisiana. In Alexandria, the pestilence killed 16 to 20 percent of the population; in Baton Rouge, it killed 5 percent; and in Shreveport, it killed close to a quarter of the population (Keating 1879, p. 89). Although it was most pronounced in coastal cities in the American South, it struck northern port cities as well. In Philadelphia, for example, outbreaks in 1699 and 1792 killed nearly 1 of every 10 residents (Duffy 1953). To put these numbers in context, consider the experience of Camden, New Jersey during the great influenza pandemic of 1918. At this time, Camden had a higher death rate from influenza than any other major American city (U.S. *Mortality Statistics* 1918). Influenza killed 1 of every 80 Camden residents, one-tenth the death rate observed in

Memphis a half-century earlier. Along the same lines, more people died in the Memphis yellow-fever epidemic than in the Chicago fire, the San Francisco earthquake, and the Johnstown flood combined (Bloom 1993, pp. 1-4).

Not until early 1900s did scientists establish that yellow fever was spread by a mosquito, the *Aedes aegypti*. The peculiar characteristics of this mosquito help explain why yellow fever existed almost exclusively in cities and towns, and was largely unknown in rural areas. Sometimes referred to as a "cistern mosquito," *A. aegypti* is a small and gray-backed insect common throughout the American South and the Carribean. It breeds in fresh water sources that are clear and relatively free of organic activity such as cisterns, metal gutters, and buckets. Such man-made containers are attractive, in part, because the mosquitos can cement their eggs to a stable and flat surface. By the same token, the *A. aegypti* avoids marshes, swamps, lakes, and water sources otherwise polluted with mud, urine, and feces. *A. aegypti* flourishes in temperatures between 70˚ and 90˚ Fahrenheit; its activities begin to slow when temperatures dip below 70˚ and it will not feed when temperatures are below 60˚. It becomes inert at temperatures below 50˚. Nevertheless, the eggs of *A. aegypti* are robust, and can survive relatively mild winters (Bloom 1993, pp. 22-28; Carter 1914, pp. 4-10).

As its name implies, the disease adversely affected liver function (resulting in jaundice) and caused a high fever. Other symptoms included headache, restlessness, chills, and nausea. For those who survived, the disease reached its peak three or four days after the onset of symptoms. For those who did not, death usually came after a week of suffering. One or two days before death, the patient's kidneys would shut down and urine output would cease. Profuse internal hemorrhaging resulted in blackened vomit, and bleeding from the gums, nose, mouth, and even old bruises (Cooper and Kiple 1993; Humpheys 1992, p. 6).

Vulnerability to yellow fever differed by age, race, and immigrant status. In contrast to most other prominent nineteenth century diseases, yellow fever bore disproportionately on older children and adults; it typically resulted in fairly mild cases in young children. Also, after thousands of years of repeated exposure, Africans developed a genetic resistance to the disease which persons of European extraction did not possess. Along the same lines, yellow fever was known as Stanger's Disease because in cities that suffered repeated exposure from the disease, newly-arrived migrants were by the most vulnerable segment of the population because those populations had yet to experience any selection based on natural immunity status. (Carrigan 1970; Kiple and Kiple 1977; Tunali ??).

In an era of limited medical understanding, medical treatments for the disease ranged from the benign to the malignant. On the latter end of the spectrum, some physicians used poisons such as antimony and mercury, as well as bleeding and ice-cold baths to treat the sick (Hogg 1840). On the former, homeopaths used bed rest, cool baths, various herbal mixtures, and diet in an effort to check the disease. Mainstream and non-homeopathic treatments for the disease were so poor that life insurance companies were said to offer discounts to city-dwellers who committed to pursue homeopathic treatments for yellow fever (Carrigan 1970). In part because of the inability to effectively treat and care for sick patients, the afflicted faced a significant risk of death. In severe epidemics, case fatality rates were said to be as high as 70 to

90 percent, though other reports suggest rates as low as 10 and 25 percent (e.g., Keating 1879, pp. 77-98; and Spinzig 1880, pp. 155-74).

There is also some debate among medical historians if the high case fatality rates associated with yellow fever historically reflected poor diagnostic skills and under-reporting of the true incidence of the disease (Humphreys 1992, pp. 5-6). According to others, recent laboratory experiments indicate that yellow fever has many strains, and that the pathogenic properties of these strains differ markedly (Bloom 1993, pp. 10-11). This suggests it might have been possible to have case fatality rates as 90 percent in some outbreaks and as low as 5 or 10 percent in others.

Yellow fever would lay dormant for years, and even decades, and then erupt suddenly, killing hundreds or thousands of people within a few months. In Memphis, for the half-century preceding 1878, in only four years (1853, 1855, 1867, and 1873) did anyone in the city die of yellow fever. Similar patterns can be observed in other cities in the American South. Yellow fever epidemics struck Savannah, Georgia in 1820, 1854, and 1876 while in the intervening years, no one in the city perished from the disease. Rivaling the Memphis epidemic of 1878, the 1876-Savannah outbreak killed as many as 1 of every 13 residents who remained in the city during the epidemic. Similarly, in Charleston, South Carolina, yellow fever epidemics struck the city on and off during the eighteenth and nineteenth centuries. In 1871 and 1872, 113 Charleston residents were killed in a yellow fever outbreak, but over the next three decades, no one else was stricken (Toner 1873; Keating 1879, pp. 80-98). That yellow fever would vanish for years and then erupt with sudden ferocity suggests that the disease not endemic to the United States.

In this regard, yellow fever was highly correlated with trade, particularly the slave trade. In large port cities, it would strike during years of unusually high trading activity, and nineteenth-century public health experts noticed that epidemics typically started with the arrival of ships from places where yellow fever was rife and endemic, such as West Africa, the West Indies, and South America (Beeson and Troesken 20xx, Keating 1879, pp.). For example, yellow fever was unknown in Europe until Columbus made contact with the West Indies, after which the disease began striking ports in Spain, Portugal, and the Mediterranean shoreline (Hand 1879). Along the same lines, a Congressional inquiry in 1878 concluded "that in all countries outside of the West Indies . . . yellow fever is an exotic disease; and in all [such countries] its introduction can be traced, either directly or indirectly, to the West Indies (Spinzig 1880, pp. 111-12)." During the early 1900s, after it was well known that the yellow fever was spread by mosquitoes, public health officials in the United States began investigating ships arriving from foreign ports for evidence that they were carrying the *Aedes agypti*, the mosquito that transmits yellow fever. Their research indicated that mosquitoes were not uncommon, and that they could indeed survive the voyage from ports in the Carribean and South America.[1]

[1] In one example, the two-masted sailing ship *John H. Crandon* was inspected at a Gulf Coast quarantine station after a twenty-two day journey from Vera Cruz, Mexico, a port where yellow fever was endemic. Larvae were found in the ships ballast tanks, and all during the voyage to the United States mosquitos were in abundance. There was "a constant buzz" in the ships forecastle, and anyone entering was "sure to be attacked by several mosquitos." By the time boat arrived at the American gulf, "a veritable plague" of mosquitoes inhabited the boat (Grubbs 1903, p. 27; Carter 1902, pp. 7-15).

By the same token, when foreign trade stopped, so too did the yellow fever epidemics. The clearest example of this come from New Orleans during the Civil War and Union occupation after the war. With the start of the Civil War, total trade (imports plus exports) passing through the Port of New Orleans fell from 202 million (constant 1860 dollars) in 1860 to zero by 1862 and remained at 4 to 7 million dollars (2 to 3 percent of its pre-war level) until the end of the war in 1865. As late as 1870, trade in New Orleans had only recovered to 60 percent of its pre-war levels. This reduction in trade was historically unprecedented in terms of its depth and length, and with it came an historically unprecedented disruption in yellow fever epidemics in the city. For the first time in the (post-1810) history of New Orleans, there was an unprecedented 7-year interval without a single reported case of yellow fever in the city, and it was only after trade returned that yellow fever also returned in 1867 (Troesken 2015, pp. aa-bb). The same pattern was also observed during the Revolutionary War, the War of 1812, and the embargo leading up to the War of 1812: in each of these conflicts, trade ceased, and so too did yellow fever; the disease only returned with peace (Monette 1842, pp. 139-42).

Geographic Distribution of Yellow Fever

Before 1875, there were at least 723 epidemics of yellow fever in 229 cities and towns across the United States (Toner 1873). With few exceptions, these towns were below 500 feet above sea level, and were coastal, bordering rivers, large bayous, the Atlantic Ocean, or the Gulf of Mexico. Although yellow fever bore disproportionately on places in the American South, the available data indicate that northern cities were not exempt from the disease. More than half of all yellow fever epidemics took place in four states: Louisiana (23 percent of all epidemics); New York (12 percent); Texas (10 percent); and South Carolina (9 percent). Moreover, within in these states, epidemics were concentrated in large cities that served as major seaports. New Orleans accounted for 66 percent of all the epidemics in Louisiana, New York City for 74 percent of all epidemics in New York state, and Charleston for 81 percent of all epidemics in South Carolina. It is notable that aside from New Orleans, the American city most often afflicted with yellow fever was New York, the largest and fasting growing city in the country. Equally notable, however, is the absence of yellow fever in places like Portland, Maine; Newport, Rhode Island; and cities located on and around the Great Lakes region and inland rivers far removed from the Mississippi.

Aside from New York, the diffusion of yellow fever was greatest in the Southern states. In Alabama, 11 different cities were afflicted, at one point or another, by the disease, though more than half of the epidemics in the state occurred in Mobile. In Florida, thirteen cities were struck, with three cities bearing the brunt of the outbreaks, Saint Augustine, Key West, and Pensacola. In Louisiana, yellow fever erupted in 43 different cities and towns. Aside from New Orleans, the state's most vulnerable towns were Alexandria, Baton Rouge, Shreveport, Thibodeaux, and Washington. In Mississippi, yellow fever was also widespread, affecting 20 different localities. Natchez, where yellow fever struck most often, accounted for only 24 percent of the outbreaks observed in the state. Yellow fever was most diffuse, however, in Texas, where outbreaks occurred in at least 41 places. The two Texas cities with the most frequent visitations from yellow fever were Galveston and Houston, which together account for only 28 percent of all outbreaks in the state.

The Decline of Yellow Fever: Three Important Moments

Plotting decadal totals of yellow fever epidemics in major American cities, Figure 1 shows three sharp drops in the incidence of the disease over time. These drops play a key role in our estimation.

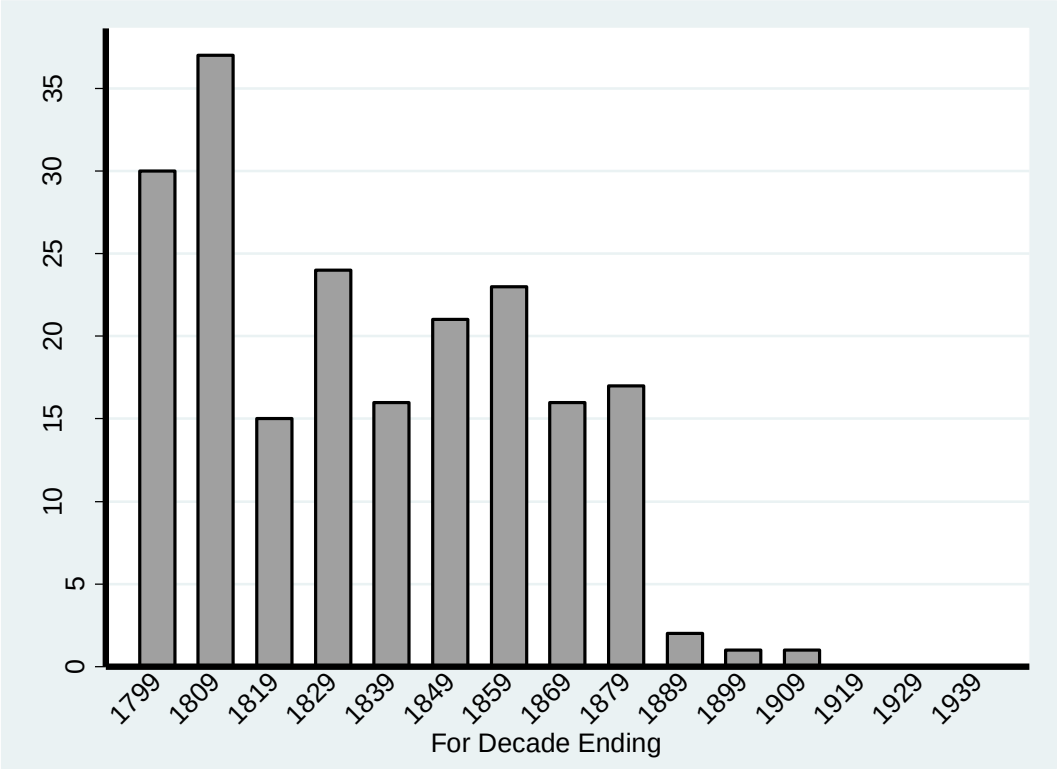


Figure 1. Total Number of Yellow Fever Epidemics in Major American Cities by Decade

The first drop occurs in the wake of the abolition of the slave trade in 1807. With the abolition of the slave trade, trade with the West Indies, Africa, and other places where yellow fever was endemic slowed. This slowdown in trade affected northern ports much more heavily than Southern ports which retained much of their trade with the West Indies and South America. For example, shortly after the slave trade was abolished, yellow fever largely disappeared from both New York City and Philadelphia, though it continued to strike in the American South with nearly the same frequency and severity as it had previously. Because Northern ports bore the loss of the slave trade more heavily than ports in the South, common sense seems to suggest that population growth and economic activity would have slowed more in Northern port cities than those in South. However, the model developed in section 2 predicts just the opposite: with the drop in the propensity to get hit with yellow fever, the congestion

costs associated with the disease would have fallen, and induced in-migration to Northern cities previously afflicted with yellow fever.

The second drop occurs around 1880, and the third around 1900. Aside from a few relatively small epidemics in Jacksonville (Florida), New Orleans, and elsewhere, after 1880 yellow fever struck with much less frequency and severity in the American South. This shift came about quite by accident. After 1880, cities throughout the U.S. began to rapidly expand their public water systems. Although cities extended public water systems largely to combat typhoid fever and diarrheal diseases, the effect was to bring piped water to populations that had previously relied on cisterns for water. In turn, urban dwellers began abandoning their cisterns, a favorite breeding ground of the *A. aegypti*. If yellow fever were an important hindrance to urban development, cities that previously had been vulnerable to severe and repeated epidemics should have started to grow faster once those epidemics began to subside (Bloom 1993, pp. 24, 226-28; Carter 1914, pp. 16-23).

After Walter Reed demonstrated that yellow fever was spread by mosquitos around 1900, preventing outbreaks of the disease was a comparatively easy task and would have promoted urban growth in cities and places hitherto most vulnerable to the disease if yellow fever represented a significant barrier to growth. Note that by 1900, New Orleans was the only city in the United States that was still experiencing yellow fever on a regular basis, and the benefits of Reed's discovery were concentrated mainly in that city which quickly adopted procedures to destroy mosquitoes and prevent reproduction.

Yellow Fever as a Congestion Externality: Five Case Studies

The model developed in section 2 predicts that in a world without effective disease mitigation technologies, disease and productivity are correlated, and that the mechanism driving this correlation is a congestion externality resulting from the increased population of high productivity places. More simply, the model reflects the following intuition: high productivity places attract more people, and because people carry disease, increased productivity is associated with more disease. While we test this logic formally in a later section, here we present case studies of five of the largest cities in early America: New York City, Philadelphia, Boston, Charleston, and New Orleans. The central message of these case studies is twofold. First, yellow fever was pronounced in these large and fast growing cities, though it was less serious in Boston because of its northern and relatively inhospitable climate for mosquitoes. Second, it was pronounced in these cities because they had natural features (notably easy access to water transportation) that attracted people and trade, and with the people and trade, came the mosquitoes that carried yellow fever.

Case studies to follow…

4. DATA

In this section, we describe the data used in our formal econometric work.

Population. Total population data (county-level) was retrieved from NHGIS for each decennial census from 1840 to 2010. Geographic data delineating county boundaries as they existed in each census year was also retrieved. We use these data to construct a balanced of total population counts at the county level.

Methods for linking census count variables through time include: 1) nominal linking; 2) areal interpolation; and 3) the method of least common bounds. Substantial changes in both the number of counties, the shape of county boundaries, and changes in county names, made linking county-level population counts nominally, infeasible. For instance, 1840 census includes 1285 counties; 2010 census includes 3221 counties. The method of areal interpolation would allow one to re-aggregate population counts from one census to another based on the proportion of each source zone assigned to each target zone. More specifically, this method would assign the total value of the attribute of interest (total population) from each source zone (for example, 2010, county-level, population counts) to target zones (for example, 1840 county polygons) according to the areal proportion of each source zone. However, while feasible, this approach assumes uniformly distributed population density. The extent to which both the number of counties as well as the shape of counties has changed over the last 170 years both elevate our concerns regarding the validity of this assumption.

The method of least common bounds is attractive because it makes no assumptions on the distribution of population counts within a county. This method, which we implement in GIS (code provided by Randy Walsh), groups counties between censuses together based on the smallest set of polygons that completely contain all counties between all censuses. Applying this procedure to our data resulted in 693 LCBs. We then aggregate total population counts based on each counties LCB group identifier for each decennial census. This approach allows us to document changes in population counts across time across a geographically stable unit of observation. 5 of the 693 LCBs contained counties that did not have population information recorded in the 1840 – 2010 censuses, these were dropped. We base our analysis on the remaining, 688 LCBs for which we can completely aggregate population counts to for the entire period of our study (1840 – 2010) resulting in a balanced panel of 12,384 observations. In more robust specifications, we drop LCBs that lie above the 99th percentile with respect to the land area they subsume; this panel includes 12,276 observations (682 LCBs x 18 decennial census years).

Figures (2a) and (2b) plot total population growth in our study area over time, as well as average county / LCB population growth.

Figure 2.a.: Total Study Area Log-Population

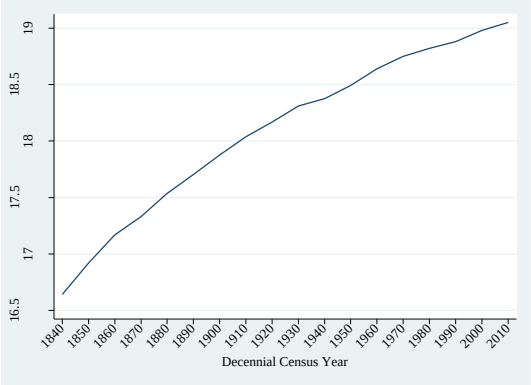
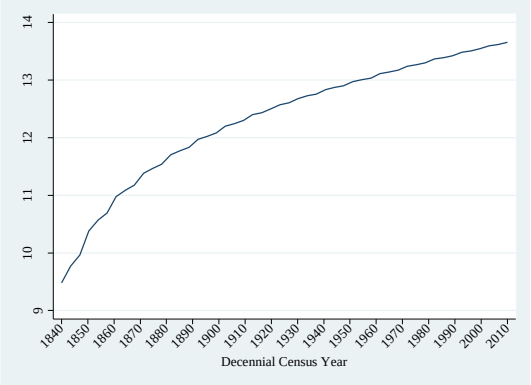


Figure 2.b.: Average County (LCB) Log-Population



Yellow Fever. Data regarding the distribution of yellow fever was compiled from J.M. Toner's 1873 publication, "The Distribution and Natural History of Yellow Fever as it has Occurred at Different Times in the United States." As shown in Figure (3), this work documents the names of localities where yellow fever has appeared in the US.

Figure 3. Excerpt from Toner Document

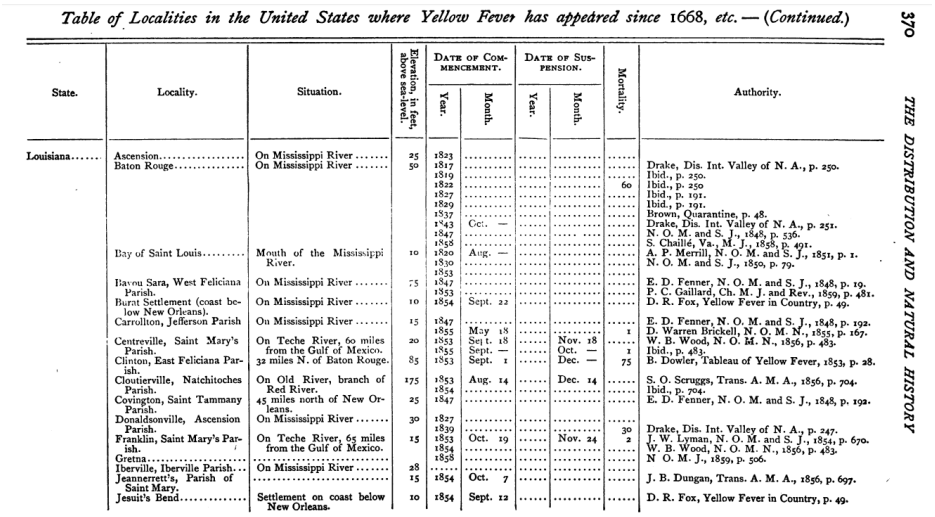


Table of Localities in the United States where Yellow Fever has appeared since 1668, etc. — (Continued.)

370 THE DISTRIBUTION AND NATURAL HISTORY

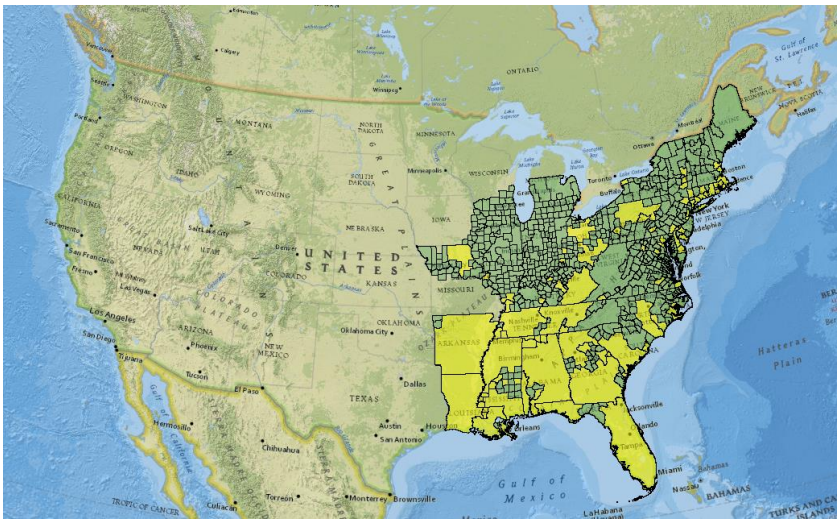
State.	Locality.	Situation.	Elevation, in feet, above sea-level.	Date of Commencement. Year.	Date of Commencement. Month.	Date of Suspension. Year.	Date of Suspension. Month.	Mortality.	Authority.
Louisiana.....	Ascension.....	On Mississippi River	25	1823	.....	.....	.....	.....	
	Baton Rouge.....	On Mississippi River	50	1817	.....	.....	.....	.....	Drake, Dis. Int. Valley of N. A., p. 250.
				1819	.....	.....	.....	.....	Ibid., p. 250.
				1822	.....	.....	.....	60	Ibid., p. 250
				1827	.....	.....	.....	.....	Ibid., p. 191.
				1829	.....	.....	.....	.....	Ibid., p. 191.
				1837	.....	.....	.....	.....	Brown, Quarantine, p. 48.
				1843	Oct. —	.....	.....	.....	Drake, Dis. Int. Valley of N. A., p. 251.
				1847	.....	.....	.....	.....	N. O. M. and S. J., 1848, p. 536.
				1858	.....	.....	.....	.....	S. Chaillé, Va., M. J., 1858, p. 491.
	Bay of Saint Louis.....	Mouth of the Mississippi River.	10	1820	Aug. —	.....	.....	.....	A. P. Merrill, N. O. M. and S. J., 1851, p. 1.
				1830	.....	.....	.....	.....	N. O. M. and S. J., 1850, p. 79.
				1853	.....	.....	.....	.....	
	Bayou Sara, West Feliciana Parish.	On Mississippi River	75	1847	.....	.....	.....	.....	E. D. Fenner, N. O. M. and S. J., 1848, p. 19.
				1853	.....	.....	.....	.....	P. C. Gaillard, Ch. M. J. and Rev., 1859, p. 481.
	Burat Settlement (coast below New Orleans).	On Mississippi River	10	1854	Sept. 22	.....	.....	.....	D. R. Fox, Yellow Fever in Country, p. 49.
	Carrollton, Jefferson Parish	On Mississippi River	15	1847	.....	.....	.....	.....	E. D. Fenner, N. O. M. and S. J., 1848, p. 192.
				1855	May 18	.....	.....	1	D. Warren Brickell, N. O. M. N., 1855, p. 167.
	Centreville, Saint Mary's Parish.	On Teche River, 60 miles from the Gulf of Mexico.	20	1853	Sept. 18	.....	Nov. 18	.....	W. B. Wood, N. O. M. N., 1856, p. 483.
				1855	Sept. —	.....	Oct. —	1	Ibid., p. 483.
	Clinton, East Feliciana Parish.	32 miles N. of Baton Rouge.	85	1853	Sept. 1	.....	Dec. —	75	B. Dowler, Tableau of Yellow Fever, 1853, p. 28.
	Cloutierville, Natchitoches Parish.	On Old River, branch of Red River.	175	1853	Aug. 14	.....	Dec. 14	.....	S. O. Scruggs, Trans. A. M. A., 1856, p. 704.
				1854	.....	.....	.....	.....	Ibid., p. 704.
	Covington, Saint Tammany Parish.	45 miles north of New Orleans.	25	1847	.....	.....	.....	.....	E. D. Fenner, N. O. M. and S. J., 1848, p. 192.
	Donaldsonville, Ascension Parish.	On Mississippi River	30	1827	.....	.....	.....	.....	
				1839	.....	.....	.....	30	Drake, Dis. Int. Valley of N. A., p. 247.
	Franklin, Saint Mary's Parish.	On Teche River, 65 miles from the Gulf of Mexico.	15	1853	Oct. 19	.....	Nov. 24	2	J. W. Lyman, N. O. M. and S. J., 1854, p. 670.
				1854	.....	.....	.....	.....	W. B. Wood, N. O. M. N., 1856, p. 483.
	Gretna.....	.....		1858	.....	.....	.....	.....	N. O. M. J., 1859, p. 506.
	Iberville, Iberville Parish...	On Mississippi River	28	.....	.....	.....	.....	.....	
	Jeannerrett's, Parish of Saint Mary.	.....	15	1854	Oct. 7	.....	.....	.....	J. B. Dungan, Trans. A. M. A., 1856, p. 697.
	Jesuit's Bend.....	Settlement on coast below New Orleans.	10	1854	Sept. 12	.....	.....	.....	D. R. Fox, Yellow Fever in Country, p. 49.

Source: Toner (1873).

Of interest to us is the list of locality names indicating regions of the U.S. that experienced a yellow fever epidemic. In order to identify these localities in our population panel, we nominally link each record in Toner's (1873) table to county area names and state names

recorded in the 1870 census. If a given locality name appeared in Toner's (1873) manual, but not in the 1870 census, we researched the history of said locality to determine changes in its name and then linked back to the 1870 census. Finally, we flagged LCBs that were determined to have a history of yellow fever. Figure (4) plots our study area and each of the 688 LCBs included in our analysis. LCBs with a history disease are shown in yellow.

Figure 4. Map of County Boundaries and Study Area



Figures 4 and 5 plot annual totals, and cumulative mortality in the U.S., due to Yellow Fever. These graphs are based on an estimate of mortality due to each yellow fever epidemic recorded in Toner's manual. However, as shown in Figure (3), there are many instances with missing mortality information; in these cases, we assign a mortality value of one.

Figure 4. Annual Mortality due to Yellow Fever

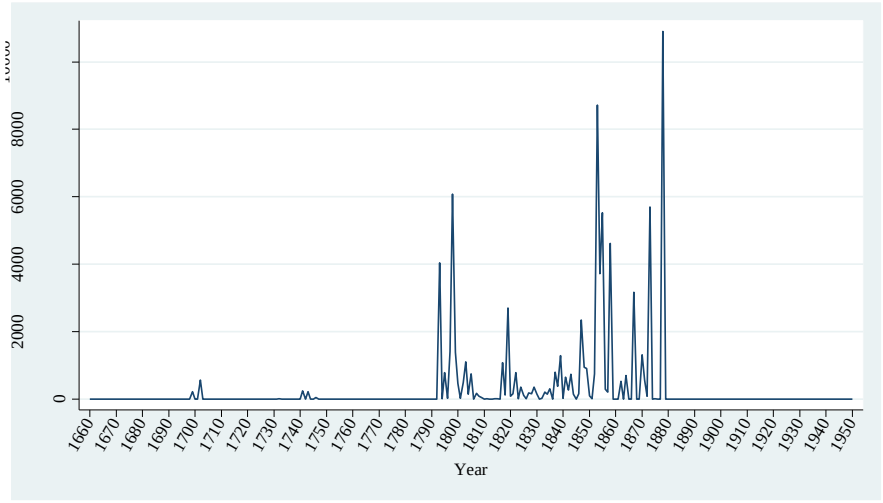
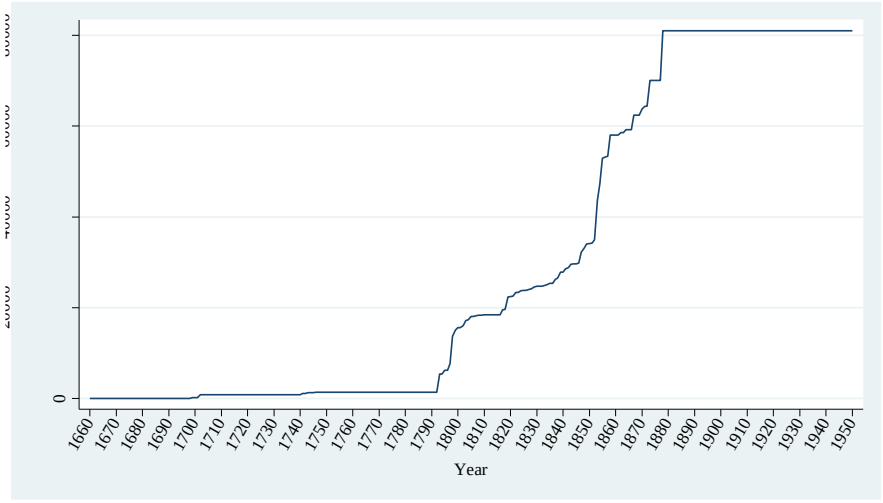


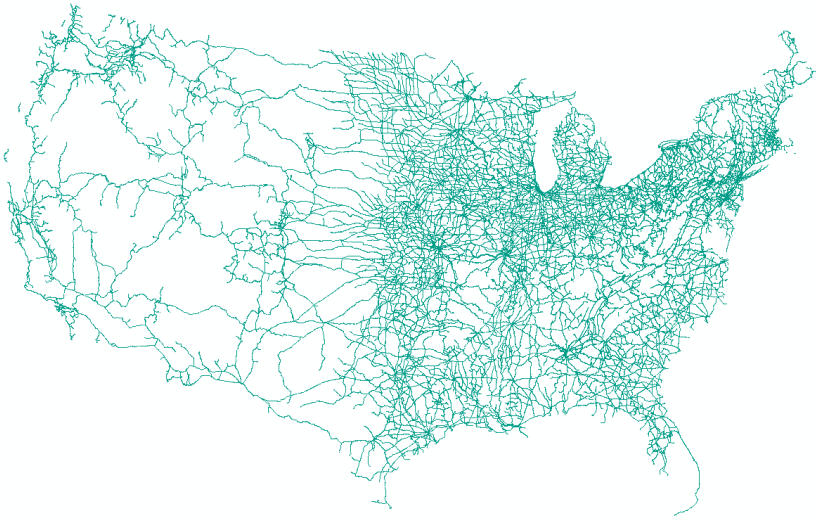
Figure 5. Cumulative Mortality due to Yellow Fever



One potential confounder in our empirical analysis is the development and extension of the American rail system, and we needed to develop data to control for that.

Rail Road. Information detailing the construction of the railroad is compiled using Jeremy Atack's (2015) "Historical Geographic Information Systems (GIS) database of U.S. Railroads for years 1830 to 1972." The extent of the railroad is illustrated in Figure (6) below.

Figure 6. American Railroad Network.



Source Data: Atack (2015)

This spatial data set indicates the rough dating (field = "InOpBy") that each stretch of line was built. We intersect Atack's (2015) railroad data with the LCB polygons in our data to determine which segments of the railroad lie within each LCB. We then compute the cumulative length (in meters) of rail line built in each LCB for each decennial census year. We denote this variable by: $Railroad_{it}$. Figure (7) plots total annual rail line built in our study area. Figure (8) plots cumulative rail line built.

Figure 7. Annual Rail Line Built (meters)

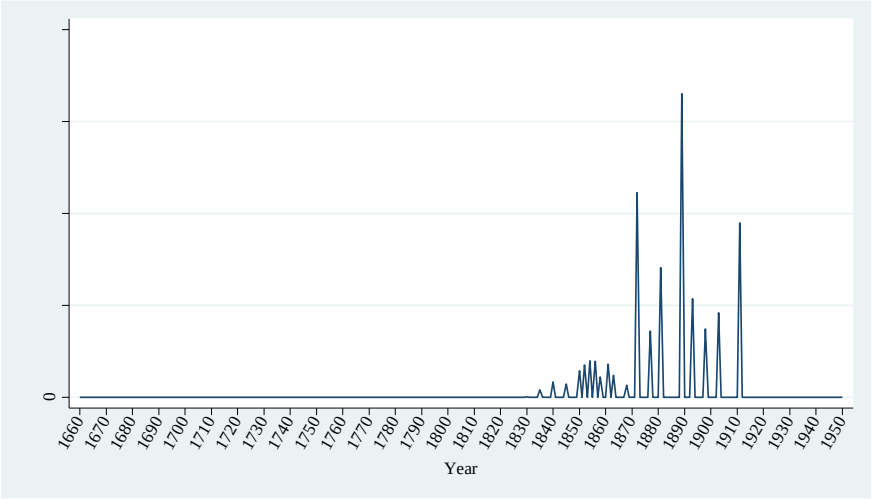
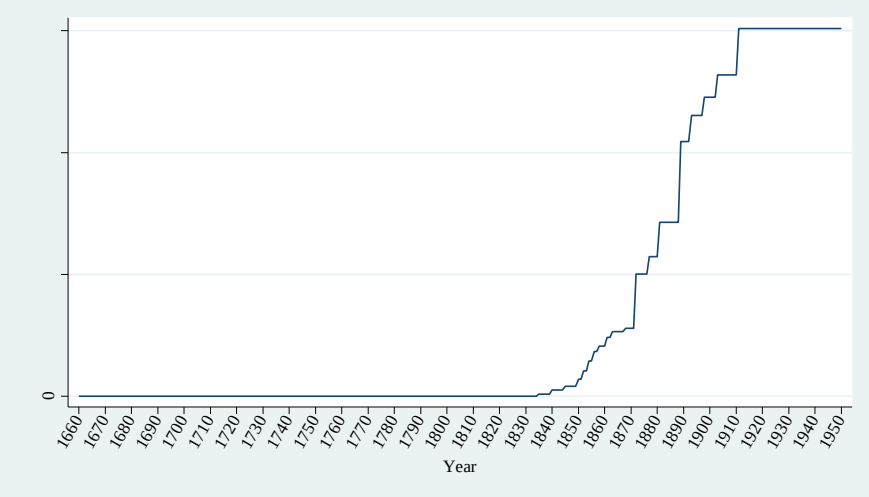


Figure 8. Cumulative Rail Line Built (meters)



5. METHODS

***Baseline Population Differences*.** Our baseline empirical models document population differences between counties with and without a history of disease by comparing changes in population with and without a history of disease. Letting i denote a county (LCB) and t each decennial census year, we estimate variants of the linear regression model,

$$\ln(pop_{it}) = \alpha + \beta \cdot I[Yellow\ Fever]_i + \epsilon_{it} \quad (1)$$

which we estimate separately for the years 1840 and 2010. $\ln(pop_{it})$ denotes the log-population level of county i in time t and $I[Yellow\ Fever]_i$ an indicator variable set equal to one for any county with a history of disease. Of interest to us are coefficient estimates of β which indicate the mean difference in log-population levels of counties with, and without, a history of disease.

Beta-Convergence. We study the degree to which counties with a history of disease converged or diverged between 1840 and 2010 by estimating variants of a beta-convergence model. These models take the form,

$$\ln(pop_{i,2010}) - \ln(pop_{i,1840}) = \alpha + \beta \cdot I[Yellow\ Fever]_i + \lambda \cdot \ln(pop_{i,1840}) + \epsilon_{it} \quad (2)$$

We also considering variants of this regression replacing the dependent variable with the log of population density.

Difference-in-Differences Framework. Our motivation for considering estimates of the beta-convergence model is its prevalence in the extant literature. However, this approach ignores population dynamics between 1840 and 2010. To characterize these properties, we study changes in the differences of population between counties with and without a history of yellow fever by estimating variants of the linear regression model,

$$\ln(pop_{it}) = \alpha + \sum_{k} (\beta_k \cdot I[Yellow\ Fever]_i \times I[Year\ k]_{it}) + \sum_{k} \gamma_k \cdot I[Year\ k]_{it} + \sum_{i} \rho_i \cdot County_i + \epsilon_{it} \quad (3)$$

$I[Year\ k]_{it}$ is an indicator variable set equal to one for decennial census year k, with $k \in \{1850, \ldots, 2010\}$. $County_i$ is a complete set of county (LCB) fixed effects. Estimates of β_k indicate mean differences in log-population levels between decennial census years k and 1840 in counties with a history of disease, relative to changes in population levels in counties without a history of disease over the same time frame, controlling for time-invariant regional effects captured by $County_i$. This model allows us to investigate the rate at which population levels between counties with and without a history of disease converged or diverged through the evolution of coefficient estimates of β_k.

6. RESULTS

Baseline Population Differences. We present estimates of equation (1) in Table (1). Columns (1) and (2) report estimates of equation (1) using log population and log population density, respectively. Columns (3) and (4) replicate Columns (1) and (2) restricting attention to LCBs that lie below the 99th percentile with respect the land area they subsume. These models show that in 1840, YF counties were 135% larger and 54% more dense than non YF counties. However, YF counties are 186% larger and 105% more dense in terms of 2010 population levels.

Table (1): Population Differences

2010 Population Differences

	(1)	(2)	(3)	(4)
Dep. Var:	ln(Pop 2010)	ln(Pop Dens 2010)	ln(Pop 2010)	ln(Pop Dens 2010)
Sample:			Area < 99th	Area < 99th
1.I_Yellow_Fever	1.869***	1.049***	1.746***	1.118***
	(0.163)	(0.168)	(0.160)	(0.173)
Observations	688	688	682	682

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

1840 Population Differences

	(1)	(2)	(3)	(4)
Dep. Var:	ln(Pop 1840)	ln(Pop Dens 1840)	ln(Pop 1840)	ln(Pop Dens 1840)
Sample:			Area < 99th	Area < 99th
1.l_Yellow_Fever	1.358***	0.539***	1.306***	0.678***
	(0.122)	(0.140)	(0.121)	(0.129)
Observations	688	688	682	682

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Beta-Convergence. We present model estimates of equation 2 in Table (2). Columns (1) and (2) estimate equation (2) using our complete sample, Columns (3) and (4) consider estimates on our more restrictive sample. Each regression relates differences in 2010 and 1840 population levels to a YF indicator variable and baseline, 1840 population levels. Coefficient estimates of the YF indicator are positive and significant while coefficient estimates of baseline, 1840 population are negative and significant. These results show that between 1840 and 2010, YF counties diverged in population from non YF counties. In contrast, counties that were initially large or more dense appeared to converge to smaller or less dense counties.

Table (2): Beta-Convergence

Beta Convergence: 1840 to 2010

	(1)	(2)	(3)	(4)
Dep. Var.:	ln(Pop 2010) - ln(Pop 1840)	ln(Pop Dens 2010) - ln(Pop 1840)	ln(Pop 2010) - ln(Pop 1840)	ln(Pop Dens 2010) - ln(Pop Dens 1840)
Sample:			Area <99th	Area <99th
1.l_Yellow_Fever	1.078***	0.804***	0.992***	0.805***
	(0.150)	(0.132)	(0.141)	(0.135)
ln_pop_1840	-0.417***		-0.423***	
	(0.0535)		(0.0545)	
ln_pop_1840_dens		-0.545***		-0.539***
		(0.0540)		(0.0548)
Observations	688	688	682	682

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

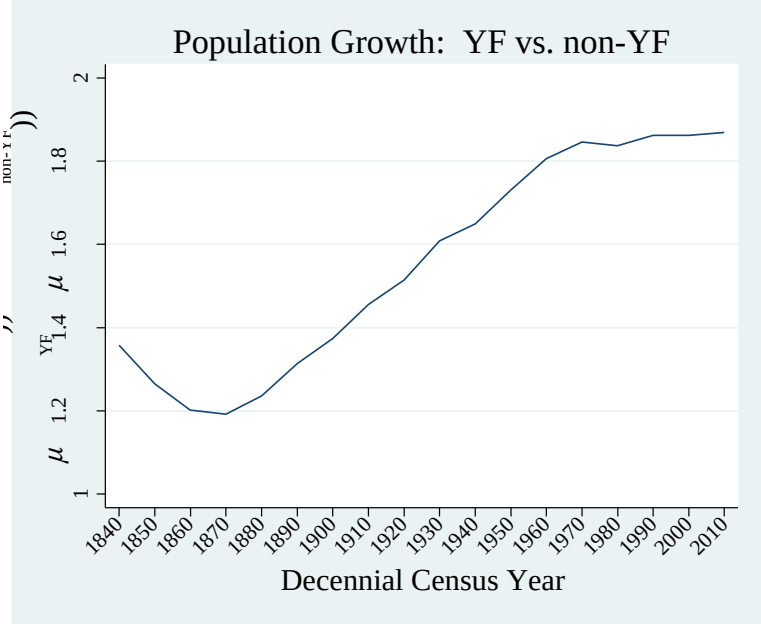
Difference-in-Differences Estimates.

Estimates of equation (1) and (2) shown in Tables (1) and (2) are useful for contrasting our results with the results in previous works, but ignore the transitionary population dynamics between YF and nonYF counties in the years between 1840 and 2010. To obtain a general sense of the these transitionary dynamics, we plot compute mean differences in log population levels between YF and nonYF counties for each decennial census year,

$$\mu[\ln(pop_{it})]_{YF} - \mu[\ln(pop_{it})]_{non-YF}$$

Each difference of means is plotted in Figure (9). Note that the end points in 1840 and 2010 correspond to coefficient estimates reported in Table (1): YF counties were 135% larger in 1840, and 180% larger in 2010; suggestive evidence of population divergence over the study period. More generally, notice from the figure that while YF counties diverged from nonYF counties in the long-term, in the years prior to 1870, our data show that these were converging to nonYF counties.

Figure (9).



We explore these dynamics more formally by estimating variants of equation (3); these results are shown in Table (3), which report coefficient estimates of each YF x decennial census year interaction term. Model estimates indicate that YF counties experienced a 9% decline in population relative to nonYF counties in 1850, relative to 1840 which is indicative of population convergence. This effect further increases in magnitude towards -15.6% in 1860 and -16.6% in 1870 which indicates that YF continued to converge in population to non YF counties. Estimates decrease in magnitude in 1880, but remain negative showing that while YF counties began to diverge after 1880, their population levels were not restored (in relative terms) to 1840 levels. Comparing the relative sign and magnitude of each coefficient estimate shows that population levels of YF counties were not restored to 1840 levels until 1900.

Table (3)

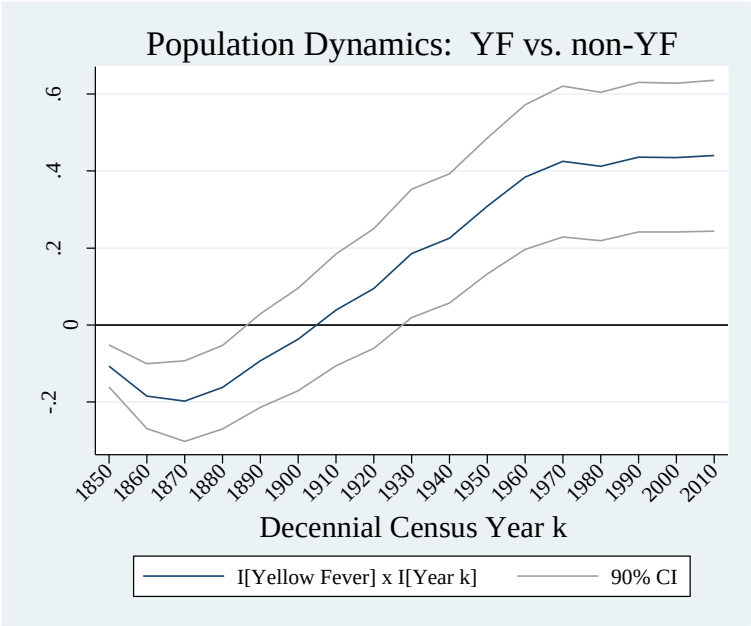
Population Dynamics: YF vs. Non-YF Counties 1840 to 2010

	(1)	(2)
Dep. Var.:	ln(Pop)	ln(Pop)
Sample:		Area < 99th
1.I_Yellow_Fever#1850.year	-0.0932***	-0.107***
	(0.0330)	(0.0331)
1.I_Yellow_Fever#1860.year	-0.156***	-0.185***
	(0.0522)	(0.0512)
1.I_Yellow_Fever#1870.year	-0.166***	-0.198***
	(0.0639)	(0.0633)
1.I_Yellow_Fever#1880.year	-0.122*	-0.162**
	(0.0676)	(0.0659)
1.I_Yellow_Fever#1890.year	-0.0438	-0.0929
	(0.0762)	(0.0738)
1.I_Yellow_Fever#1900.year	0.0162	-0.0374
	(0.0831)	(0.0808)
1.I_Yellow_Fever#1910.year	0.0982	0.0390
	(0.0903)	(0.0879)
1.I_Yellow_Fever#1920.year	0.158	0.0952
	(0.0968)	(0.0943)
1.I_Yellow_Fever#1930.year	0.250**	0.186*
	(0.104)	(0.101)
1.I_Yellow_Fever#1940.year	0.291***	0.225**
	(0.105)	(0.102)
1.I_Yellow_Fever#1950.year	0.373***	0.309***
	(0.111)	(0.107)
1.I_Yellow_Fever#1960.year	0.448***	0.384***
	(0.118)	(0.114)
1.I_Yellow_Fever#1970.year	0.488***	0.425***
	(0.124)	(0.119)
1.I_Yellow_Fever#1980.year	0.479***	0.412***
	(0.123)	(0.117)
1.I_Yellow_Fever#1990.year	0.504***	0.436***
	(0.124)	(0.118)
1.I_Yellow_Fever#2000.year	0.504***	0.435***
	(0.125)	(0.117)
1.I_Yellow_Fever#2010.year	0.511***	0.440***
	(0.127)	(0.119)
Observations	12,384	12,276
County FE	YES	YES
Year FE	YES	YES
No. Clusters	688	682

***p<.01 **p<.05 *p<.1. Standard errors (shown in parenthesis are clustered at the county level

To illustrate the transitionry dynamics between YF and nonYF over time, Figure (10) plots coefficient estimates corresponding to Column (1) of Table (3); together with 90% confidence intervals. With 1840 as the base year, decreases in these estimates indicate relative convergence between YF and nonYF counties, increases in these estimates indicate relative divergence.

Figure (10)



Sensitivity: Controlling for the impact of the railroad. Estimates of equation (3) shown in Table (3) and Figure () ignore the impact of the railroad. We test the sensitivity of these results to flexible controls for the rail road in Table (4). Column (1) replicates Column (2) of Table (3). Column (2) tests the sensitivity of the results in Column (1) to controlling for cumulative amount of rail line built in each decennial census year; Column (3) allows the effect of the railroad to vary by decennial census year, Column (4) allows the effect of the railroad to vary by decennial census year, fit separately for YF and nonYF counties.

Table (4)

Population Dynamics: YF vs. Non-YF Counties 1840 to 2010
Sensitivity to the Arrival of the Railroad

VARIABLES	(1) ln_pop	(2) ln_pop	(3) ln_pop	(4) ln_pop
1.I_Yellow_Fever#1850.year	-0.107***	-0.114***	-0.0999***	-0.111**
	(0.0331)	(0.0337)	(0.0376)	(0.0459)
1.I_Yellow_Fever#1860.year	-0.185***	-0.204***	-0.155***	-0.206***
	(0.0512)	(0.0523)	(0.0566)	(0.0704)
1.I_Yellow_Fever#1870.year	-0.198***	-0.228***	-0.190***	-0.217**
	(0.0633)	(0.0653)	(0.0695)	(0.0853)
1.I_Yellow_Fever#1880.year	-0.162**	-0.206***	-0.179**	-0.174*
	(0.0659)	(0.0685)	(0.0711)	(0.0887)
1.I_Yellow_Fever#1890.year	-0.0929	-0.169**	-0.144*	-0.0717
	(0.0738)	(0.0786)	(0.0799)	(0.0991)
1.I_Yellow_Fever#1900.year	-0.0374	-0.133	-0.0978	-0.00215
	(0.0808)	(0.0875)	(0.0880)	(0.108)
1.I_Yellow_Fever#1910.year	0.0390	-0.0637	-0.0358	0.0932
	(0.0879)	(0.0953)	(0.0959)	(0.118)
1.I_Yellow_Fever#1920.year	0.0952	-0.0240	0.00432	0.166
	(0.0943)	(0.104)	(0.104)	(0.128)
1.I_Yellow_Fever#1930.year	0.186*	0.0663	0.0827	0.293**
	(0.101)	(0.111)	(0.112)	(0.138)
1.I_Yellow_Fever#1940.year	0.225**	0.106	0.122	0.336**
	(0.102)	(0.111)	(0.113)	(0.139)
1.I_Yellow_Fever#1950.year	0.309***	0.190	0.205*	0.447***
	(0.107)	(0.117)	(0.119)	(0.148)
1.I_Yellow_Fever#1960.year	0.384***	0.265**	0.280**	0.549***
	(0.114)	(0.124)	(0.127)	(0.159)
1.I_Yellow_Fever#1970.year	0.425***	0.306**	0.322**	0.599***
	(0.119)	(0.129)	(0.133)	(0.168)
1.I_Yellow_Fever#1980.year	0.412***	0.292**	0.315**	0.569***
	(0.117)	(0.126)	(0.130)	(0.164)
1.I_Yellow_Fever#1990.year	0.436***	0.316**	0.347***	0.588***
	(0.118)	(0.127)	(0.131)	(0.164)
1.I_Yellow_Fever#2000.year	0.435***	0.315**	0.352***	0.571***
	(0.117)	(0.126)	(0.130)	(0.163)
1.I_Yellow_Fever#2010.year	0.440***	0.321**	0.360***	0.565***
	(0.119)	(0.127)	(0.132)	(0.165)
Observations	12,276	12,276	12,276	12,276
County FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Railroad	NO	YES	YES	YES
Railroad x Year	NO	NO	YES	YES
Railroad x Year x I[Yellow_Fever]	NO	NO	NO	YES
No. Clusters	682	682	682	682

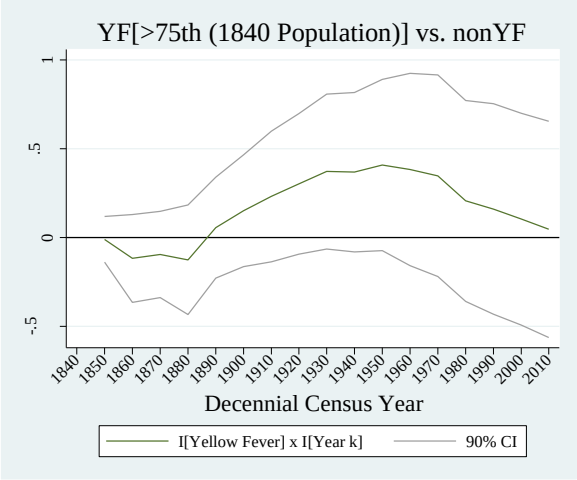
***p<.01 **p<.05 *p<.1. Standard errors (shown in parenthesis
are clustered at the county level. Models omit LCBs lying above the 99th percentile with respect to area.

The Role of Initial Population. Our descriptive results show that even at early points of American history, YF counties were larger and more dense. This motivates us to think about the role that initial population levels and initial population density levels played in American urban development. We take two complementary approaches to analyze these dynamics.

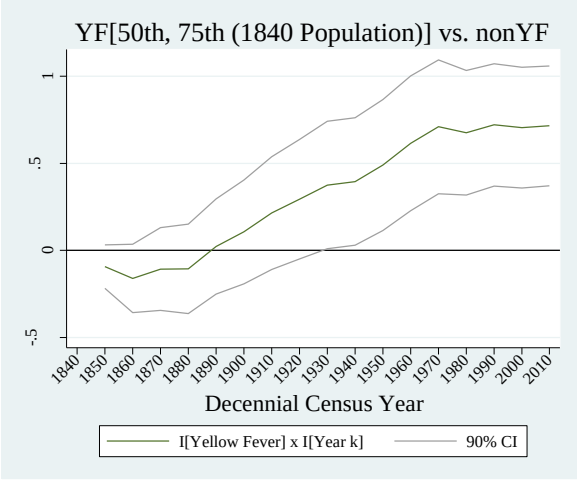
First, we look at the distribution of initial population as well as initial population density levels of YF counties. Based on these distributions, we split YF counties into quartiles, and compare population dynamics between YF and nonYF counties over time within each quartile. These results of partitioning YF counties based on 1840 population levels are illustrated in Figure (11). Each model is based on our most robust empirical specification which controls for the effect of the railroad through time separately for YF and nonYF counties.

Figure (11)

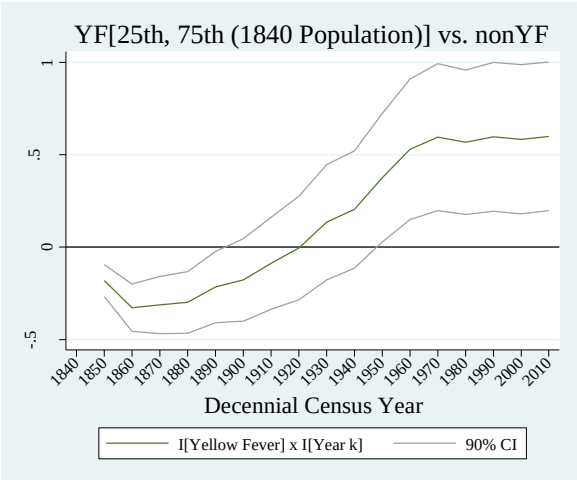
(a)



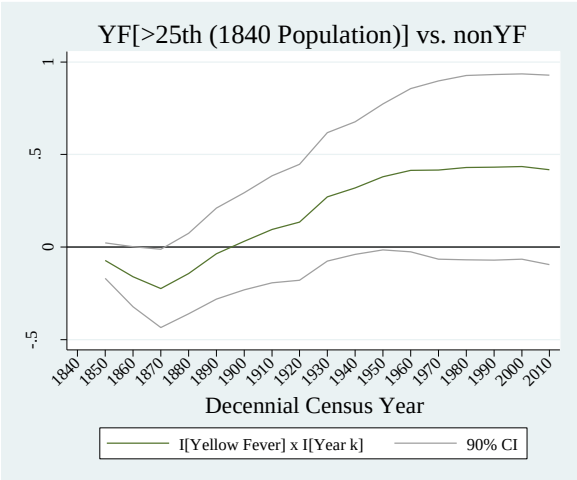
(b)



(c)



(d)

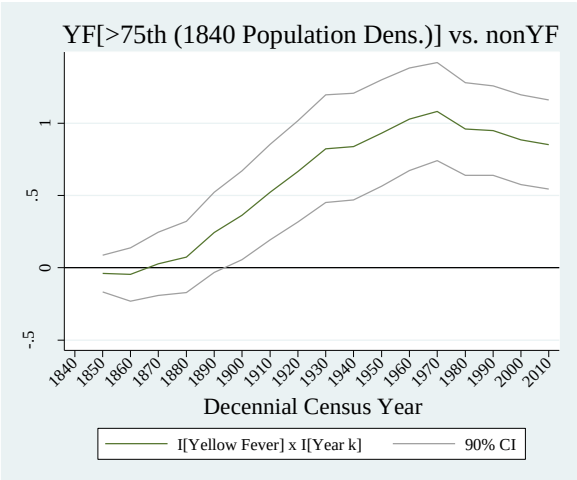


Panel (a) studies population dynamics between YF and non YF counties, restricting attention to YF counties that lie above the 75th percentile with respect to 1840 population counts. Panels (b) – (d) restriction attention to YF counties that lie between the 50th and 75th percentile (panel b), between the 25th and 50th percentile (panel c) and below the 25th percentile (panel d), respectively.

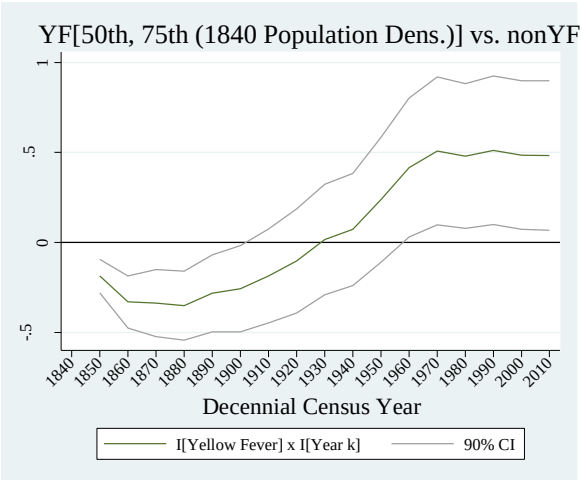
Next, we focus on the distribution of initial population density levels of YF counties. Based on these distributions. These results of partitioning YF counties based on 1840 population density levels are illustrated in Figure (12).

Figure (12)

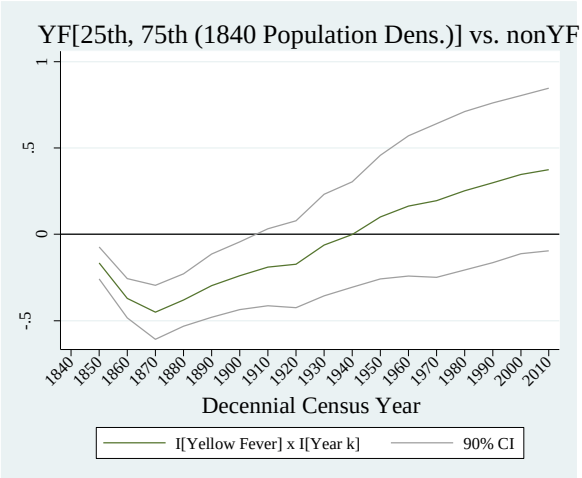
(a)



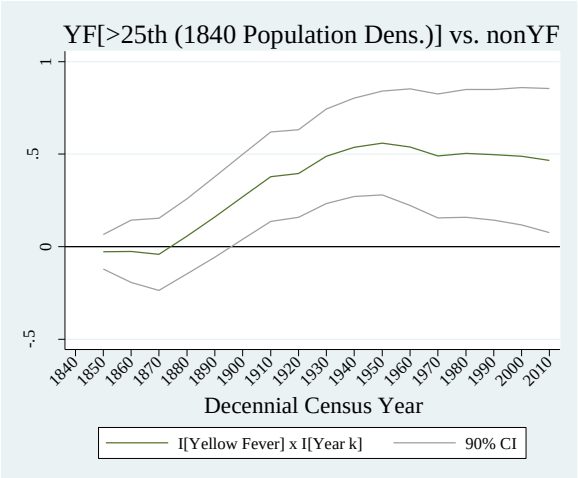
(b)



(c)

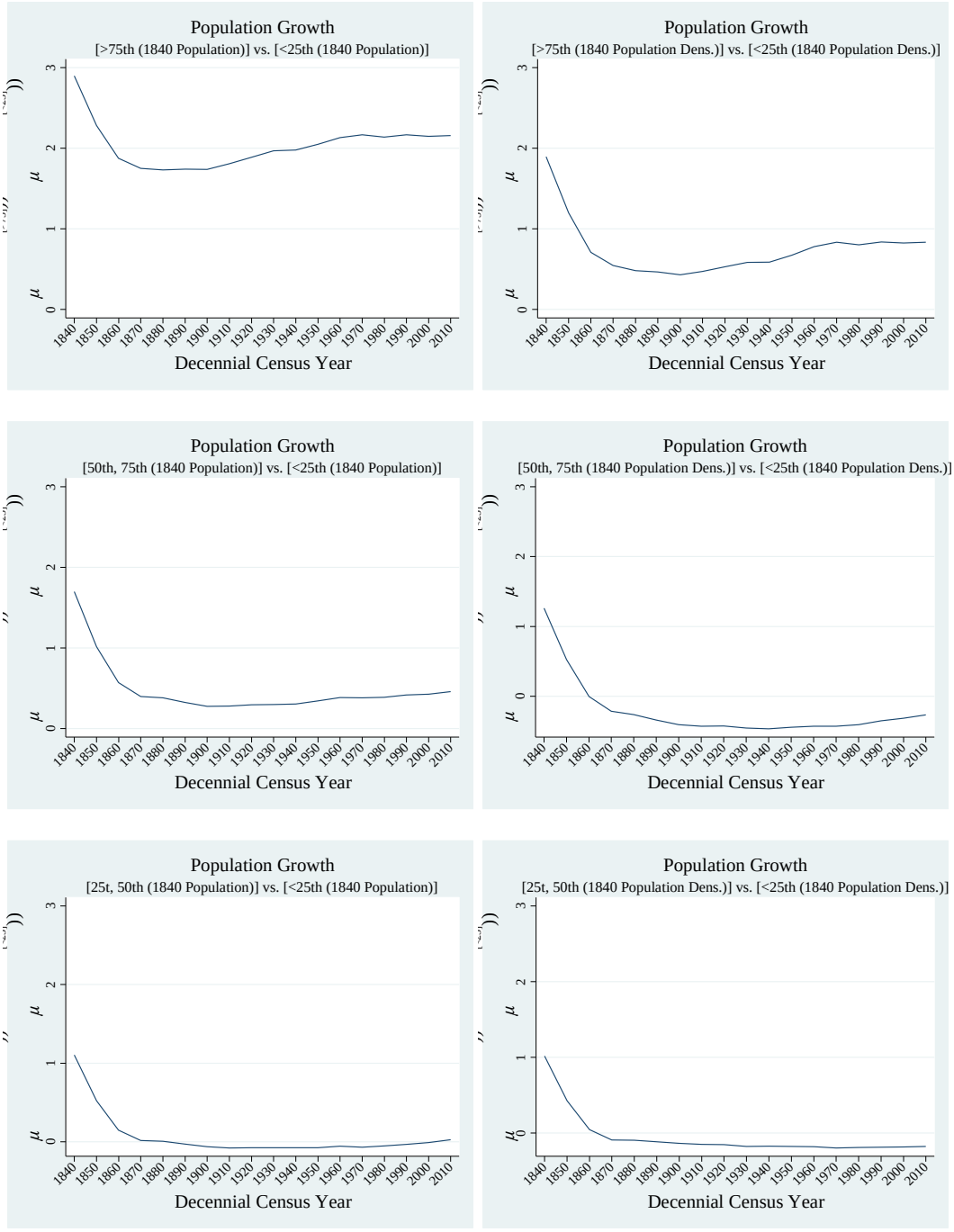


(d)



Convergence based on Initial Population Levels. Next, we study the role population played, independent of the role of YF. To do this, we look at the overall distribution of 1840 population levels and 1840 population density levels. For each distribution, we split our sample of LCBs into 4 quartiles. Based on these quartiles, we compare mean differences over time between LCBs. These results are illustrated in Figure (13). The first column partitions counties based on initial population levels, the second column partitions counties based on population density levels.

Figure 13.


Population Growth
[>75th (1840 Population)] vs. [<25th (1840 Population)]
Population Growth
[>75th (1840 Population Dens.)] vs. [<25th (1840 Population Dens.)]
Population Growth
[50th, 75th (1840 Population)] vs. [<25th (1840 Population)]
Population Growth
[50th, 75th (1840 Population Dens.)] vs. [<25th (1840 Population Dens.)]
Population Growth
[25t, 50th (1840 Population)] vs. [<25th (1840 Population)]
Population Growth
[25t, 50th (1840 Population Dens.)] vs. [<25th (1840 Population Dens.)]
Decennial Census Year

Final Thoughts.

Comparing population dynamics across counties based solely on initial, 1840 population levels indicates that initially larger counties generally converged in population to initially smaller counties over the 1840 to 2010 study period. This may partially explain why YF counties converged to nonYF counties at early points of American history to the extent that YF counties had initially larger, baseline population levels compared to nonYF counties. However, baseline population levels only appear to partially explain our findings; initially large places tended to converge to initially small places across the full duration of our sample; a result more consistent with the predictions of Solow growth theory. In contrast, YF counties rapidly diverged from nonYF counties in a period of time where YF epidemics began to disappear. In fact, by the late 1990s, YF counties diverge so much, population levels exceeded baseline 1840 levels; a result that is largely inconsistent with the predictions of Solow growth theory. Interpreted through the lense of our theoretical model, this finding seems to suggest that disease (or the dis-utility associated with the perception of disease) played an important role in inhibiting growth in YF counties early on. We draw a number of key insights from this anaylsis.

First, these results show that YF is a predictor of long-term growth potential.

However, long-term growth potential may induce disease in the absence of public health. In this sense, economic growth may be self-inhibiting if public health infrastructure isn't adequately addressed. As a result, one might regard the role of public health in promoting sustainable growth as something even more fundamental than we previously thought.

Disease and Cities

Let localities be indexed by i. At any given time, let the flow utility of a particular locality be

$$u_i = z_i - (1 + y_i)c(p_i)$$

where z_i is the productivity in said locality, y_i is the cost of disease (either psychic or physical), and $c(\cdot)$ maps from population p_i into a cost of congestion.

1 Steady State

In this case, any locality with $p_i > 0$ has $u_i = \bar{u}$ for some $\bar{u} > 0$. This value is determined by conservation of people

$$\sum_i p_i = P$$

Suppose that c is increasing, then we will have

$$p_i = c^{-1}\left(\max\left\{0, \frac{z_i - \bar{u}}{1 + y_i}\right\}\right)$$

and so $\bar{u}$ will satisfy

$$\sum_i c^{-1}\left(\max\left\{0, \frac{z_i - \bar{u}}{1 + y_i}\right\}\right) = P$$

2 Forward Looking

The previous analysis assumed that agents make location decisions myopically. In reality they will be forward looking and will incur relocation costs. Suppose that relocation costs are *Oregon Trail* like and depend on the region of origin

$$r_i = \phi(x_i)$$

where x_i is the number of people leaving a locality. Their present value of being in a locality will be

$$\delta v_i - \dot{v}_i = u_i + \max\{0, \bar{v} - v_i - r_i\}$$

where $\bar{v} = \max_i v_i$. In this case, we will then arrive at

$$r_i = \bar{v} - v_i$$

Hence we will also have

$$\delta v_i - \dot{v}_i = u_i$$

Thus we arrive at

$$x_i = \phi^{-1}(r_i) = \phi^{-1}(\bar{v} - v_i)$$

Localities with $v_i = \bar{v}$ will experience entry. Assume a uniform rationing rule, so that

$$e_i = \frac{\sum_i x_i}{|\{i|v_i = \bar{v}\}|}$$

This can be formulated as a coupled system of differential equations in p_i

$$\dot{p}_i = e_i - x_i$$

The steady state of which corresponds to the results in the previous section, since $\dot{v}_i = \dot{p}_i = 0$.

3 Intuition

People want to live *la dolce vita*, and in this case that means getting paid your marginal product in high productivty (z_i) localities and not dying of or having to worry about dying of disease (y_i). Disease here is modeled as a congestion externality, in the sense that it gets worse with higher population density.

Places with high flow utility will in general have high present valuations, and you can always move on to greener pastures if need be. These value gradients will cause people to pay a moving cost (r_i) to migrate. Moving costs are determined by the number of people exiting a particular locality. One could think of this as reflecting some scarce resource that must be purchased locally.

In steady state, flow utilities (u_i) will be equalized. Hence, localities with high productivity and/or low propensity for disease will have higher populations and larger realized congestion effects. When the disease mitigation technology improves ($y_i \to 0$), congestion effects will be dampened in cities that had endemic disease. As a result, people will migrate to those cities and they will reach their full potential.

This alone should be enough to generate the observed differences between the growth rates of localities with and without endemic disease. One interesting question is whether one would expect a increase or decrease in the variance of city sizes overall. This would depend on the correlation between productivity (that is, before the effects of disease) and disease prevalence. Generally, one would expect an increase in variance, but if cities with endemic disease are also low productivity, the opposite could potentially arise.

Side note: I'm working a neat variance decomposition along these lines. We could also potentially get some nice closed form results on this and on aggregated productivity gains from reallocation if we assume linear congestion costs.