Accepted Manuscript

Title: THE ENGINE AND THE REAPER: INDUSTRIALIZATION AND MORTALITY IN LATE NINETEENTH CENTURY JAPAN

Author: John P. Tang

PII: S0167-6296(17)30830-5
DOI: http://dx.doi.org/10.1016/j.jhealeco.2017.09.004
Reference: JHE 2056

To appear in: *Journal of Health Economics*

Received date: 29-2-2016
Revised date: 31-7-2017
Accepted date: 6-9-2017

Please cite this article as: {http://dx.doi.org/

THE ENGINE AND THE REAPER: INDUSTRIALIZATION AND MORTALITY IN LATE NINETEENTH CENTURY JAPAN

By JOHN P. TANG*

*Australian National University, Research School of Economics, 26 LF Crisp Building, Acton, ACT 2601, Australia (john.tang@anu.edu.au). The author would like to thank the two anonymous referees and editor in the preparation of this manuscript. Useful comments were also received from Jan Stuhler, Tue Gorgens, Timothy Hatton, Gaurab Aryal, Osamu Saito, Kyoji Fukao, Masato Shizume, Chiaki Moriguchi, Chika Yamauchi, Patrick Nolen, Alfonso Herranz-Loncan, Laura Panza, Latika Chaudhary, Remi Jedwab, Walker Hanlon, Shari Eli, Kris Inwood, Price Fishback, Dan Bogart and participants at the Australasian Cliometrics Workshop, Western Economic Association International conference, ANU Centre for Economic History transport workshop, World Economic History Congress, the Canadian Network of Economic History conference, and various departmental seminars. Dek Joe Sum, Amanda Maclean, Shigeo Morita, and Mitchell Lee provided excellent research assistance. This research is supported by funding from the Australian Research Council (DE120101426) and the Japan Society for the Promotion of Science.

Economic development improves long-run health outcomes through access to medical treatment, sanitation, and higher income. Short run impacts, however, may be ambiguous given disease exposure from market integration. Using a panel dataset of Japanese vital statistics and multiple estimation methods, I find that railroad network expansion is associated with a six percent increase in gross mortality rates among newly integrated regions. Communicable diseases accounted for most of the rail-associated mortality, which indicate railways behaved as transmission vectors. At the same time, market integration facilitated by railways corresponded with a twenty percent increase in total capital investment nationwide over ten years.

Keywords: difference-in-differences, matching estimators, market integration, mortality Kuznets curve, transportation infrastructure
JEL Codes: I15, J61, N75, O24

In the year 1886, a major wave of cholera swept through Japan and killed 108,405 people by the year's end.[1] This single disease accounted for more than 1 out of 9 deaths that year, compared to the previous year's toll of 9,310 deaths and the following year's total of 654. While most of the country experienced high rates of mortality, areas with railroad access had a higher incidence, 336 deaths per 100,000 compared to the 245 in prefectures without rail access.[2] Mortality rates for all areas were similar in the years immediately before and after the epidemic, around 1 death per 100,000. Another outbreak of cholera four years later claimed only one third the number of casualties, but again prefectures with rail access had higher mortality rates, 81 per 100,000 compared to 66 in areas without rail. As the international spread of the Zika virus in recent years has also demonstrated, increased transport access may facilitate disease transmission as well as commerce and economic development (World Health Organization 2016b).[3]

The relationship between transport infrastructure and disease transmission qualifies a well-established stylized fact that industrialization improves public health. Extensive scholarship finds improvement in longevity, child mortality, and adult heights as economies develop (McKeown 1976; Fogel 1986; Hatton 2014). These gains in turn foster increases to productivity and promote continued

[1] Mortality rates based on resident population figures; these statistics are from Japan Statistical Association (1967) and are further described in the section on data.

[2] Rosenberg (1987) also finds a relationship between the spread of various cholera epidemics and water transport in New York City during the mid-nineteenth century.

[3] The Zika virus, which causes birth defects and auto-immune disorders, was initially identified in human patients in Uganda in 1952, spread to French Polynesia starting in 2007, and currently has active local transmission in 34 countries and territories (Centers for Disease Control and Prevention 2016).

development as part of a virtuous health-wealth cycle (Costa 2013). Health improvements can arise from two forces: macroeconomic growth, which allow governments to invest in sanitation, medical facilities, and other public health measures; and higher living standards, including greater access to and consumption of food as well as the acquisition of human capital.

Less obvious is how the timing and spatial variation inherent to the transition from a traditional economy to an industrialized one affects mortality, particularly if labor mobility increases the frequency of exposure to infectious diseases. In other words, the health benefits of industrialization can be absent in its initial stages, vary between regions depending on their degree of integration, and differ by cause of death. While the health risks of industrial activity are illustrated by studies of unsafe work conditions, unhealthy urban environments, exposure to foreign diseases, and industrial pollution, there are few studies that specifically identify market integration via improved transport as a causal factor in increased mortality (Johnston 1995; Walker 2009; Beach and Hanlon, forthcoming).[4]

The contribution of this paper is to estimate the impact of transport access on mortality, using the introduction of the railroad in Japan at the end of the nineteenth century as a quasi-natural experiment. Railways, powered by the steam engine, were one of the first modern technologies diffused around the world and credited with leading industrialization through domestic and international market integration, cheaper and faster shipping, and intersectoral linkages (Fogel 1964; Summerhill 2005; Donaldson, forthcoming). In Japan, their adoption and use led to increased firm activity and industrial agglomeration, and between the 1870s, when railroads were introduced, and the eve of the First World War, the economy was transformed from an isolated, agrarian society to an emerging industrial

[4] Notable exceptions to this include Haines et al (2003) and Rosenberg (1987) for the United States.

power with higher income levels, burgeoning trade, and a nascent colonial empire (Tang 2014; Perkins and Tang 2017).[5] The country's rapid transformation, however, was remarkable as well for its persistently high and rising mortality rates and the differential trends between regions, with rural death rates steadily converging upward toward urban ones. Compared to the experiences of other countries during the 1800s, increasing mortality in rural areas is striking given that they had net outward migration, less industrial activity, and healthier environmental and nutritional conditions (Taeuber 1958; Johansson and Mosk 1987; Honda 1997).[6]

What can explain these unusual mortality patterns during the early phases of Japanese industrialization at a time when mortality rates in western economies were declining? One hypothesis is that the railroad not only helped to integrate the domestic market in goods and labor, but also to spread communicable diseases like tuberculosis and influenza. While transport has been implicated in disease transmission since the Black Death and the discovery of the New World, a sustained increase in mortality due to disease exposure may not be as obvious in the modern period or within a homogeneous population that had stable food supplies, improving living standards, and similar immunity (Komlos 2000; Morse 2001).[7] To test whether rail access increased domestic mortality through disease transmission, one can compare regions before and after they receive railways to other areas with similar initial levels of development but lacking rail access. Separating mortality by disease group sharpens the transmission mechanism

[5] During the Meiji Period (1868-1912), Japan acquired the Ryukyu islands (1879, now known as Okinawa), Taiwan (1895) and Korea (1910).

[6] During Britain's industrialization (1776-1841), while urban mortality rates were consistently higher than rural ones, their ratio was relatively stable (Williamson 1990, p. 54). In contrast, Sweden's rapid industrialization was not associated with negative health outcomes (Sandberg and Steckel 1997). Szreter (2004, p. 81) suggests that industrialization "exerts intrinsically negative population health effects among those communities most directly involved in the transformations which it entails."

[7] While food production and access to foreign supplies may have mitigated mortality from nutrition and disease, improved transport also increased exposure to contagion (Taeuber 1958, p. 50).

played by the railroad since increased exposure and mobility would likely have a stronger impact on food-borne and communicable diseases and not on other causes of death.

Using both matching estimation and a difference-in-differences linear regression framework and a panel dataset disaggregated by region and disease group, I find rail access accounts for 5.5 percent of average annual mortality across prefectures integrated into the national rail system between 1884 and 1893. Put differently, rail-associated mortality represents about 66 percent, or 111 of the 169 additional deaths, of the increase in total mortality per post-rail year in each of the affected prefectures compared to pre-rail years.[8] These figures account for measures of per capita income, urbanization, industrial activity, and medical treatment and are robust to both estimation methodologies. Furthermore, the difference-in-differences specifications indicate that 75 percent of rail-associated deaths is due to communicable diseases, which is consistent with the hypothesis that improved transport facilitated the transmission of contagion, but had a negligible impact on non-communicable causes of death.[9] Corroborating these results is the finding that lower population density is associated with disproportionately higher mortality among regions following rail access. As suggested in existing work on Japan and other countries, the integration of smaller, isolated areas with the national market may raise the frequency of exposure since the former lack both sufficient urban infrastructure and immunity

[8] The difference in average annual mortality rates by prefecture before and after rail access period is 169 additional deaths per 100,000; of which 111 deaths are associated with rail. The 111 deaths annual figure is calculated using the per disease group average treatment effect on areas gaining rail access during the period 1884 to 1893; see Table 4, column C for results. As discussed later, the remaining increase in mortality not explained by rail may be due to better reporting of all types of mortality from increased political centralization or increased incidence across all disease categories from market integration unrelated to railways.

[9] The three disease groups classified as communicable are acute infectious, digestive, and respiratory. These categories include diseases listed as infectious (chronic and acute) by the World Health Organization (2016a); see appendix A for a list of diseases within each category. The paper uses the terms communicable, contagious, and infectious interchangeably, with the main differences being whether transmission requires direct person-to-person contact or indirect contact via discharges (communicable); are due to biological contaminants like bacteria or viruses (infectious); or encompasses aspects of both other terms (contagious) (Merriam-Webster 2016).

within the population from outbreaks of communicable disease (Taeuber 1858; Cain and Hong 2009).

While the arrival of the steam engine may have heralded the grim reaper in late nineteenth century Japan, the economy also benefited from increased capital investment due to railways and their associated agglomeration and scale economies (Tang 2014). In the counterfactual scenario with no additional inter-regional railways between 1884 and 1893, Japanese capital investment would have been 17.8 percent lower nationwide. This is equivalent to a 15.8 percent per capita decrease in the period immediately preceding the country's industrial takeoff at the turn of the century, which relied on domestic market integration and export growth (Perkins and Tang 2017). One interpretation of these findings is that in this early stage of the country's development, there may have existed a tradeoff between health and wealth.

I. Background and literature review

The relationship between economic development and health outcomes generally finds a positive influence of the former on the latter. Laborers earning higher wages are able to consume more and better food while governments can provide public health services and clean water (Costa and Steckel 1997; Ferrie and Troesken 2008). The availability of medical treatment also delivered improved outcomes, with the development of vaccines, medications, and disease prevention techniques leading to substantial reductions in mortality across a range of diseases (Cutler et al 2006).[10]

Notwithstanding health gains in the long term, these contributions in the short run may vary in importance or differ between regions and thus be absent or

[10] Smallpox vaccination reduced the severity of epidemics in Japan between 1886 and 1908, while cholera and bubonic plague saw substantial reductions in mortality by the early 1900s (Taeuber 1958, p. 51).

even negative on average. For example, during Britain's high period of industrial activity in the mid-nineteenth century, the decrease in mortality rates may have owed more to improved nutrition than to public health measures or medical treatment (McKeown 1976; Cutler et al 2006; Fogel 1986).[11] For the United States in the early 1800s, soldiers experienced declines in heights and life expectancy before recovering later in the century even as industrial activity and average incomes grew steadily over the period (Haines et al 2003; Margo 2000; Margo and Steckel 1983; Komlos 1998, 2012). This "ante-bellum puzzle" finds that food availability and diet, particularly early age nutrition, were positively correlated with health outcomes while deteriorating external conditions and stagnant wage growth had negatively associations.[12]

Another typical feature of industrialization is increased urbanization, which historically increases mortality rates (Cutler et al 2006; Atack et al 2010). Contributing factors include crowded living and unhealthy work environments, which aid the spread of communicable diseases, as well as limited access to fresh food, clean water, and clean air, although this may be outweighed by higher wages (Johnston 1995; Williamson 1982). Economic development may also have a more direct effect in the creation of industrial hazards like the transmission of respiratory diseases in confined spaces, toxic pollution and effluent, and longer work hours (Lewchuk 1991; Macintyre 1997; Ferrie 2003; Gagnon et al 2011; Tang 2015). In the United States during the 1800s, increased disease exposure had a major impact on life expectancy, with mortality rates corresponding to

[11] McKeown (1976) differentiates between airborne versus water- or food-borne communicable diseases, with the former unaffected by sanitation measures. Improved food hygiene practices like pasteurization started in the late 1800s, while effective medical intervention, such as the development of penicillin (1928), streptomycin for tuberculosis (1946), and cardiovascular treatment were developed after the nineteenth century. Before the twentieth century, medical treatment of infectious diseases was largely palliative given the lack of knowledge about germ theory and disease transmission.

[12] Countering these studies is research that finds no causal link between income and life expectancy starting in the nineteenth century (Lindert 1983), and the interplay between nutritional intake and diseases regardless of how the latter are carried (Preston and van de Walle 1978). Around the turn of the twentieth century, despite a relative undersupply of water treatment and physicians, American mortality rates in rural areas declined (Komlos 1998; Higgs 1973)

population and housing density (Meeker 1972; Preston and van de Walle 1978; Clay and Troesken 2006). Healthier living environments via improved hygienic practices in Europe over the same period contributed more strongly to adult heights compared with increased income and education (Szreter and Mooney 1998; Millward and Baten 2010; Hatton 2014).

The difference in the timing of these factors may give rise to what can be considered a mortality Kuznets curve for industrialization, where the impacts of industrial activity worsen health outcomes initially but improve later as rising living standards, investments in public health measures, and medical treatment exert their influences (Johansson and Mosk 1987; Honda 1997; Szreter 2004).[13] While this relationship is well-documented for heights, it is less clear for mortality and in countries industrializing over a shorter time period (Costa and Steckel 1997). One channel for the delay in health gains may be access to modern transportation, which speeds up the process of industrialization by increasing production efficiency, the movement of goods and people, and urbanization while generalized economic improvements lag behind (Haines et al 2003; Atack et al 2009). In the United States, for example, access to transport via rail or water was associated with an increase in mortality and a decrease in height during the mid 1800s (Haines et al 2003).

Compared to other historical transitions, Japan's industrialization may offer greater insight to the health-wealth relationship given its rapidity, availability of data, and fewer confounding factors. Japanese growth was aided by its adoption of foreign technologies and institutions, the activities of its

[13] Cutler et al (2006) credit improvements to public health measures after 1870, with the "acceptance of the germ theory of disease in the 1880s and 1890s, which led to a wave of new public health initiatives and the conveyance of safe health practices to individuals" (ibid, p. 102). The spread of germ theory coincides with this paper's period of analysis but may not have had much of an effect given the country's reliance on eastern medical practices and limited public health spending. Cutler et al also identify three stages in the historical change in mortality: the first phase, from the mid 1700s to the mid 1800s, due to increased nutrition and economic growth; the second phase up to the 1930s, where public health mattered most; and the third phase thereafter, with reductions due to medical treatment and advances (ibid, p. 106).

entrepreneurs, and the favorable conditions in which it integrated with the world economy (Sussman and Yafeh 2000; Mitchener et al 2010; Tang 2011, 2013, 2016). Based on industrial growth rates and the share of manufacturing value to total output, Japan's arrival as an industrialized economy can be dated as early as the mid 1890s (Benetrix et al 2015; Perkins and Tang 2017). A major innovation catalyzing this transition was the railroad, introduced in 1872 and reaching most regions throughout the country by the early twentieth century. In practical terms, rail transport allowed a day's journey to be shortened to an hour, bringing mineral ores and perishable silk in neighboring prefectures within easy reach of major cities and the international markets.[14] Recent work shows that railroad expansion led to increased firm activity and investment in areas that received the transportation infrastructure earlier (Tang 2014).

At the same time, better transportation links, coupled with unfettered migration, urbanization, and population growth, may have increased the transmission of communicable diseases like tuberculosis and cholera, leading to a net increase in mortality (Johnston 1995).[15] This rise in mortality rates occurred despite increased food spending per capita, income from trade, and industrial activity (Japan Statistical Association 1962; Yamazawa and Yamamoto 1979, table 3; Hayami and Ruttan 1985).[16] Moreover, mortality rates exceeded population growth rates during the 1880s; stayed relatively constant at 21 deaths

[14] Tokyo, the capital and largest city, was connected to Yokohama, in neighboring Kanagawa prefecture, in 1872 and brought deep-sea access to the capital. Osaka, the second largest city, was connected to the port of Kobe in Hyogo prefecture in 1874.

[15] Japanese labor markets were deregulated with the ending of feudal land ownership and occupational restrictions in the 1870s, allowing migration within and between regions (Flath 2014). The six largest cities grew from 2.4 million inhabitants in 1888 to 6.1 million in 1918, or 11 percent of the total population, while urban areas exceeding 50,000 inhabitants increased their population share from 7 to 17 percent over the same period (Johnston 1995, p. 64). Mortality from respiratory illnesses, which included tuberculosis, bronchitis, and pneumonia, rose from 17.8 percent of total mortality in 1888 to 29.2 percent in 1898 (ibid, p. 60).

[16] Food spending increased by about two percent each year in real terms between 1885 and 1900 and exports increased by nearly nine percent per annum. Between the years 1886 and 1899, silk reeling factories increased from 411 to 2,217 and those for cotton spinning and weaving increased from 89 to 1,370 (Johnston 1995, pp. 75-77). Silk textiles used domestically sourced cocoons, which competed with land used for rice cultivation; cotton textiles relied primarily on imported raw cotton.

per thousand in the next decade; and slowly declined to 19 deaths per thousand by the late 1930s. [17] Industrialized prefectures experiencing higher rates throughout the end of the nineteenth century, and then had a marked decrease afterward (Honda 1997, pp. 276-77). Mortality also varied significantly between disease groups and was higher across all diseases in the treatment prefectures in the 1880s to early 1890s compared those in the control group, as shown in Table I.
[Table I]

Starting in the early 1870s, government concern about the social impact of communicable diseases led to policies targeting cholera, dysentery, typhoid fever, smallpox, diphtheria, and typhus, which were legally designated as acute infectious diseases.[18] Despite regulations to control their spread (including revocation of medical licenses for physicians caught concealing diagnosed cases), epidemics of cholera continued over the next few decades, notably in 1886, albeit with decreased mortality over time (Johnston 1995, p. 62). In contrast, cases of tuberculosis, a chronic respiratory disease, increased dramatically in the late 1800s through the early twentieth century, reaching an all time high in 1918 (ibid, p. 87). Compared to countries in western Europe, Japan had initially similar tuberculosis mortality rates in the late nineteenth century, although average income levels and public health spending were much lower (Honda 1997, p. 265). Government spending on public health was mostly limited to smallpox vaccination and sanitary education, with considerably more money devoted to military and industrial activities (ibid, p. 267).[19] This may explain why the disease's incidence remained high well into the 1900s even as it fell in other

[17] The probability of death for males improved significantly in the 1890s, especially for those under 15 years and between 25 and 49 (Taeuber 1958, p. 51).

[18] Most contagious diseases were already present in Japan before the onset of industrialization, with the exception of plague and typhus (ibid, p. 257).

[19] Central government spending on sanitation services was initially rather low, averaging 10,000 yen in real terms per year between 1885 and 1894 before increasing significantly in the next decade to reach 220,000 yen in 1904, with a similar pattern for water treatment and provision (Emi and Shionoya 1966, table 13).

industrial economies. It was during the late nineteenth century, however, with the large mortality differential between industrial and agricultural regions, discrete changes in railroad access, and the relative neglect in public health spending and increasing industrial activity, that one can more clearly identify the relationship between transport and disease mortality.[20]

II. Research design

This paper tests the hypothesis that improved transport access corresponds with higher mortality rates in early stages of industrialization, ceteris paribus.[21] This may be surprising since railroad construction itself has been part of catch-up historical economic development: on the one hand, industrialization leads to higher average incomes, allowing for better public health, medical treatment, and nutrition; on the other, urbanization and workplace hazards may have deleterious impacts, especially for certain disease types. However, these factors neglect the role of transport in raising the frequency of disease exposure and transmission to areas that may have been naturally protected due to isolation (Johansson and Mosk 1987; Taeuber 1958).[22]

[20] Honda (1997) finds higher mortality rates in industrialized prefectures compared to agricultural ones before 1900, but does not differentiate between those with and without rail access. Modern medical treatment, following the discovery of germ theory, led to the development of vaccines and antibiotics in the first half of the twentieth century, which postdates the period of analysis in this paper.

[21] There are a few important qualifications to this statement. First, this study analyzes a period during which the germ theory of disease was not widely accepted; this theory was embraced following Robert Koch's discovery of the bacteria that caused tuberculosis in 1882 (Johnston 1987, p. 11). Second, the government was unable or unprepared to deal with public health issues due to lack of enforcement of sanitary practices and insufficient equipment (e.g., microscopes for diagnosis were not available in all prefectures before 1891) (ibid, pp. 241, 275). Third, the Japanese public was relatively uninformed about modern sanitation practices and medicine that were available in western countries; for example, tuberculosis was considered heritable and incurable in the pre-war period (ibid, pp. 32 and 259). Even with increased government awareness, medical school trained physicians did not outnumber traditional practitioners until 1908 and the gap in knowledge between the populace and government persisted throughout the pre-war era (ibid, pp. 231 and 249).

[22] An alternative explanation is that railroads allowed for selection on population subgroups, with less healthy (e.g., older) individuals remaining in rural areas and thus accounting for the increase in mortality in those regions. Unfortunately, there are no age-disaggregated mortality statistics by region for this period to test this explanation, although the communicable disease groups encompass ailments that target different age groups or are age indiscriminant.

Fortunately, the historical context of Japan's industrial transition and potential impact on health provide a suitable framework for analysis, and sharpens the focus on the roles of transport and market integration. In the late nineteenth century Japan, explanatory factors associated with health improvements, such as government spending on public health and medical advances like antibiotics and vaccines based on germ theory, were modest or missing. Another potential confounding factor, industrial pollution, was also important before the turn of the twentieth century compared to industrialized economies like the United Kingdom since there were fewer factories and concentrated in a handful of urban areas.[23] Reforms in land and tax policy also encouraged labor mobility as workers were no longer tied to the land, who could then take advantage of the newly introduced railway and thus acted as vectors for disease transmission (Flath 2014). This was particularly pronounced with tuberculosis, which increased in geographic spread due to the emergence of modern textile factories in the late nineteenth and early twentieth centuries and the railroad network constructed at the outset of Japanese industrialization (Johnston 1987).[24]

Since railways expanded over time and across regions, one can identify its effect as the network grew extensively as well as account for intensity of use.[25] This can be approached by comparing mortality rates in regions that gained rail access to those that did not, and further contrasting differences in trends before an

[23] Beach and Hanlon (forthcoming) note the role of coal smoke exposure on increased mortality in mid-nineteenth century Great Britain. This factor is less relevant to late nineteenth century Japan, which did not rely as heavily on fossil fuel use in industries (i.e., labor-intensive textile manufacture) and most of these factories were concentrated in the large urban areas of Osaka and Tokyo, which are excluded from the main analytical results.

[24] "Had it not been for improved transportation and the rapid growth of the textile industry, tuberculosis might have stayed in the cities. But the living and working conditions in the textile industry made factories hothouses of disease. As soon as tuberculosis impaired the ability of worker's ability to perform their duties, companies invariably fired them and sent them back to their home villages. There, they often infected family members and neighbors, and in villages that had had little prior exposure to tuberculosis, mortality soared. As villages became increasingly afflicted, labor recruiters went to more distant villages to find wiling workers, and as a result, the disease spread from one village to the next in concentric waves" (ibid, pp. 415-16).

[25] Railway construction began in the 1870s, but it was not until the early 1880s that significant lengths of track were laid across the country, coinciding with the collection of vital statistics across regions.

area receives rail access to the years after. This change in transport infrastructure allows for a treatment-control research design and a causal interpretation of the impact of railways on mortality, where prefectures integrated into the national rail system can be considered a treatment group while those without access during the same period comprise a control group. To further improve identification of transport access as the mechanism for transmitting contagion, disease groups are also grouped by communicability. This latter source of variation highlights the mechanism of transport itself as opposed to any detrimental effects from induced industrial activity: deaths due to communicable diseases (or any contagion carried by biological hosts) should be affected, but other disease types should not.

A. Data sources

The newly assembled data used in the analysis come from multiple sources, and are novel for both their level of disaggregation by disease group and region and in the matching of regional mortality rates with economic and industrial indicators at the same geographic level. Moreover, these data are comprehensive and cover the entire population and economy, providing an unbiased and detailed view of changes occurring within the country. For vital statistics, the government's Cabinet Bureau of Statistics compiled annual prefecture-level mortality data on twelve different disease categories starting in 1883 (Japan Statistical Association 1962).[26] These include: acute infectious, blood, bone and joint, developmental and nutritional, digestive, external injury, nervous, poisoning, respiratory, skin and muscle, urogenital, and unclear groups.[27]

[26] Aside from the poisoning and unclear groups, each disease category comprises multiple diseases ranging from 8 (skin and muscular) to 29 (digestive); see appendix A for details.

[27] The unclear category is particularly useful in assessing the consistency of disease classification over time and between the treatment and control groups. Controlling for prefecture-level characteristics in the regression model, the shares of unclassified mortality remain statistically unchanged for both treatment and control groups throughout the period, indicating that precision in medical diagnosis did not improve in the country among either group of prefectures or disproportionately among treatment group prefectures (which arguably could have attracted better qualified physicians due

The statistics, which apply to the resident population in a prefecture (i.e., inclusive of migrants), were recorded under the central government's classification system by local administrative offices (Johnston 1995, pp. 57-58).[28] In 1901, the government revised this system with different, more clinically more precise categories, but there is no direct mapping between the two classifications.[29] Besides mortality data, the government also collected data on the number of public and private hospitals in operation, pharmacies, and health professionals for each prefecture annually.[30]

For industrial and demographic data, the Cabinet Bureau yearbooks also provide prefectural figures of firms, capital investment across industry types, population, and different measures of land area and value. The first two economic series are used to measure industrial development in each region, while per capita income can be proxied by privately owned land value divided by the resident population as well as invested capital per capita. These nominal series are converted into constant 1990 yen using the national consumer price index compiled by Jorda et al (2017). Unfortunately, no prefecture level price indices are available on an annual basis for this period.

Urbanization, which was inconsistently documented before the 1920 population census, is approximated with population density based on low gradient or habitable land area. This adjustment is preferable to using total land area since

---

to rail access). While disease diagnosis may have been inaccurate during this period, these results suggest that it was uniformly applied and thus do not threaten the identification strategy.

[28] These categories are not precise in that digestive illnesses may have included gastrointestinal illnesses due to unclean water as well as stomach cancer while nervous conditions may include both cerebral hemorrhages and tertiary syphilis. Developmental and nutritional diseases included different forms of cancer, senility, and old age (Johnston 1995, p. 323). Local physicians made the diagnoses, which may also be imprecise given that the majority was not trained at medical schools and used traditional homeopathic techniques (ibid, p. 58). Deaths for the registered population (*honseki*) were separately recorded after adjustment for the migrant population, but do not have disease disaggregation.

[29] Internationally, there was no agreed system available until the 1893 publication of Jacques Bertillon's classification *International Classification of Causes of Death* (Moriyama et al 2011, p. 12).

[30] While rudimentary, these data are far more comprehensive and available at an earlier period than most contemporary countries, including the United States (Moriyama et al 2011). There are also data on disease mortality by age group, although these are not disaggregated by prefecture and are considered unreliable for younger age groups (Taeuber 1958, pp. 42-43). Hospital data are missing for 1884.

only 16 percent of Japanese land is arable and most of the population resided in these low-lying coastal areas and alluvial plains (Trewartha 1945; Taeuber 1958). In particular, land area is subdivided by gradient level as well as by type (i.e, fields, forests, mountainous areas); both classification systems are used in the analysis. To measure railroad network expansion, I use the *Rail Stations of Japan* handbook, which includes all rail stations built in the country by date and location starting in 1872 (Chuo Shoin 1995; Tang 2014). Together, I use these data to construct a balanced panel dataset of mortality rates by disease group, prefecture, and year and include controls for economic activity and other determinants of health.[31]

B. Empirical framework

This paper uses two empirical strategies, matching estimation on observables in areas that gain rail access to those that did not as well as a difference-in-differences regression model that compares mortality rates for each disease group before and after a prefecture receives rail access (i.e., treatment) against regions that do not receive access over the same years. The period of analysis covers the years 1883 to 1893, which spans the first major expansion of railroad construction across regions in Japan as well as the economy's industrial takeoff as measured by the share of industrial output to national production.[32] Over this period, main trunk lines increased from 202 to 897 kilometers, while local railways grew from 101 to 2,223 kilometers (Japan Statistical Association 2007, table 12-7).[33] Annual passengers carried increased six-fold, from 5.2

[31] Appendix B lists the prefectures with year of rail access, rail stations per 100,000 residents in year 1893, and mortality per 100,000 residents in year 1893, and number of observations based on data availability for the full set of covariates.

[32] Tang (2014) uses the same treatment period to analyze regional firm activity patterns, while Perkins and Tang (2017) estimate Japan's entry to industrialized status as occurring between 1897 and 1907, based on five-year moving averages.

[33] The number of public and private rail stations grew from 53 to 394 between 1883 and 1893.

million in 1883 to 32 million in 1893 (ibid, table 12-8a). Since consistent vital statistics on disease groups do not begin until 1883, that year is used to start the period of analysis.[34] For the end year, 1893 precedes the First Sino-Japanese War (1894-95), which corresponded with increased use of railways to transport wounded soldiers and thus may potentially confound the results for disease transmission and location of death registration. Government spending on public health measures also quadrupled in 1895, from 76 thousand yen the previous year to 307 thousand yen (Emi and Shionoya 1966, table 13). Thus, the effect of missing data for prefecture level government spending would be magnified starting in the mid 1890s, potentially biasing the results.[35]

In the matching estimation analysis, all prefectures are included to calculate average treatment effects from rail access on mortality, with breakdowns by disease categories.[36] The prefectures are then separated into groups depending on period of access to highlight differences that may exist depending on length of treatment. In the difference-in-differences estimates, the eleven prefectures that gained rail access on or before 1883 are excluded to test for pre-treatment comparability and obtain a clean treatment effect; the omitted areas include the major urban regions of Tokyo and Osaka.[37] In both estimation methods, Okinawa prefecture is also omitted since it does not have industrial data during this period. For the difference-in-differences model, Kagawa is excluded as well since it

[34] Mortality statistics are available in earlier years for some prefectures, but used a different disease group classification system that continued to change until 1883 (Japan Statistical Association 1967).

[35] There was consolidation of public health initiatives starting in the 1890s, such as the establishment of the Greater Japan Medical Association in 1893 and modern treatment for injuries and illnesses from the Russo-Japanese war in 1904-05 (Lock 1980, pp. 62 and 232). The government also commissioned studies on workplace health between 1903 and 1913, with particular attention paid to the increase in the spread of tuberculosis in rural areas due to industrialization and military activities (Johnston 1987, pp. 103-7, 122).

[36] Disease category mortality statistics are complete across prefectures by year, but not all prefectures have complete series for the full period between 1883 and 1893 due to data availability for covariates. See appendix B for the number of observations by prefecture. Missing years of prefectural data are: Aichi (1891, 1892); Aomori (1885); Chiba (1885); Gifu (1891, 1892), Hokkaido (1885, 1892); Kagawa (1883-1888); Mie (1891); Nagano (1885); Nagasaki (1885); Nara (1883-1887); Saga (1883); Saitama (1887); Shizuoka (1885); Toyama (1883); Yamagata (1883); and Yamanashi (1883, 1885).

[37] The eleven prefectures excluded because of early rail access (pre-1884) include Gifu, Gunma, Fukui, Hokkaido, Hyogo, Kanagawa, Kyoto, Osaka, Saitama, Shiga, and Tokyo.

receives rail access the same year that mortality statistics are reported, meaning it does not have a pre-treatment comparison. These restrictions bring the number of prefectures down to 34 out of 47 in the latter estimation framework, but can be compared with some of the results from the matching estimation analysis.

Figure I shows the prefectures grouped by period of rail access on a map of Japan both at the beginning and end of the treatment period.[38] The first group (i.e., always treated) includes 11 prefectures that gained rail access before 1884 and includes the major cities of Tokyo and Osaka. The second group (i.e., new treatment) comprises 17 prefectures that gained rail access between 1884 and 1893. Since mortality data are available starting in 1883, prefectures in this group have at least one year of pre-treatment data to establish pre-treatment comparability in the difference-in-differences analysis. The remaining group (i.e., control) includes the 17 prefectures did not gain rail access until after 1893.[39]
[Figure I]

For the matching estimation analyses, average treatment effects of railways for all prefectures (ATE) and those with treatment (ATET) are estimated on annual mortality rates. The variable $Death_{ijt}$ is measured in the annual number of deaths per 100,000 inhabitants for each disease group i or combination of disease groups in a prefecture j in a given year t. Treatment is defined as an indicator variable for rail access, $Rail_{jt}$ with zero for no rail access or one for access beginning in the year when the prefecture first has a rail station built until the end of the period.

Prefectures are matched on the following observables: the value of privately owned land divided by the resident prefectural population, $LandVal_{jt}$,

[38] The excluded prefecture Okinawa, comprising small islands south of the main four islands, is not shown and would be included in the post-1893 rail access group if mortality and industrial data were available.

[39] Treatment group prefectures include Aichi, Aomori, Fukuoka, Fukushima, Hiroshima, Ibaraki, Iwate, Kumamoto, Mie, Miyagi, Nagano, Nara, Niigata, Okayama, Saga, Shizuoka, and Tochigi. The control group includes Akita, Chiba, Ehime, Ishikawa, Kagoshima, Kochi, Miyazaki, Nagasaki, Oita, Shimane, Tokushima, Tottori, Toyama, Wakayama, Yamagata, Yamaguchi, and Yamanashi.

which proxies for individual wealth since prefectural output statistics are unavailable; the population density of each prefecture using low gradient land (i.e., less than 3 degrees in gradient), $PopDen_{jt}$, as a proxy for urbanization; the number of industrial firms per 100,000 prefectural inhabitants, $IndFirms_{jt}$, as a measure of industrial activity; and the number of public and private hospitals per 100,000 prefectural inhabitants, $Hosp_{jt}$, as a measure of medical service provision. Note that the prefectural population figures, like those for the mortality statistics, are based on the resident population, not only those officially registered, and thus account for inter-prefectural migration flows.[40] For nearest neighbor matching (NNM), individual matches within disease groups are made using Mahalanobis distance on the above four covariates. Since these are continuous measures, I make a bias correction for the reported average treatment effects (Abadie and Imbens 2011). For propensity score matching (PSM), the same four covariates and disease group are used in a logistic model to find the closest match.

For the difference-in-differences reduced form linear regression model, I use the same covariates as in the matching estimation model, but specify the functional form and include additional covariates. The basic specification is:

(1) $Death_{ijt} = \beta_0 + \beta_1 \cdot Rail_{jt} + \beta_2 \cdot PrefVar_{jt} + \beta_3 \cdot Rail_{jt} \bullet PrefVar_{jt} + \beta_4 \cdot Disease_i + \beta_5 \cdot Pref_j + \beta_6 \cdot Year_t + \varepsilon_{ijt}$, where

$Death_{ijt}$ = mortality rate for disease group i by prefecture j in year t

$Rail_{jt}$ = rail access variable for prefecture j in year t

$PrefVar_{jt}$ = explanatory variables for the proxies of per capita income, urbanization, industrial activity, and medical services for prefecture j in year t

$Rail_{jt} \bullet PrefVar_{jt}$ = interaction of rail access and explanatory variables for prefecture j in year t

[40] Prefectural migration data are not available until the twentieth century, starting with the 1920 population census.

$Disease_i$ = disease group fixed effect

$Pref_j$ = prefecture fixed effect

$Year_t$ = year fixed effect

ε_{ijt} = error term

Note that one salient difference between the matching estimation and difference-in-differences linear regression models is that the latter allows the treatment variable $Rail_{jt}$ to be a continuous measure.[41] By using the number of rail stations per 100,000 prefectural inhabitants in a given year, this treatment variable more precisely accounts for the density of transportation access and intensity of use compared with an indicator.[42] The difference-in-differences model also allows the inclusion of prefecture fixed effects to remove unobserved heterogeneity or idiosyncratic shocks, which would be more difficult to capture in a matching framework.

To further improve identification of rail access as a vector of disease transmission, I include an indicator variable for communicable disease, which applies to the three categories of acute infectious (e.g., cholera, typhoid), digestive (e.g., hepatitis, food-borne illnesses), and respiratory (e.g., pneumonia, tuberculosis) ailments.[43] In the matching estimation model, this variable is used to differentiate the effect of rail access on the mortality series between communicable and non-communicable disease groups. For the difference-in-differences model, this variable is interacted with rail access (dummy and continuous) and reported in separate specifications to account for a heterogeneous treatment effect. This variable also allows use of the full sample of prefectural and

[41] In matching estimation, treatment does not strictly have to be binary, but can be aggregated to discrete groupings. This, however, still creates imperfect matches and thus are biased (Angrist and Pischke 2009, p. 77). Matching on a single continuous covariate can be accomplished with minimal efficiency loss, but requires a large number of matches; matching on multiple continuous variables (as performed in this paper) is generally inconsistent and inefficient (Abadie and Imbens 2006, p. 236).

[42] Note that direct measures of prefecture level freight and passengers by rail station or line are unavailable.

[43] See footnote 9 for disease definitions and classification.

disease group data to compare mortality trends, increasing the power of the regression estimates. In all double- and triple-differences specifications, I include fixed effects for disease group, prefecture, and year and use robust standard errors clustered by disease group and prefecture.

To check for robustness, I use alternative measures for the four time-varying prefecture level control variables: $CapInvest_{jt}$, which is total capital investment across industries divided by the resident population; $PopDenAlt_{jt}$, which is population density based on habitable land (i.e., rice paddies, fields, and residential land); $IndustShr_{jt}$, the share of capital invested in industrial firms (excluding transport) to total capital; and $Doct_{jt}$, the number of doctors per 100,000 inhabitants, respectively. In the difference-in-differences specifications, rail access is also interacted with these four control variables individually and all together since railways have been shown to facilitate economic activity as well as urbanization and land value (Atack et al 2010; Tang 2014; Donaldson and Hornbeck, forthcoming). These in turn may have provided government and private resources for public health services (Allen 1946; Onji and Tang, 2017).[44]

There are some advantages to choosing one estimation method over another with consequences for the interpretation of results (Angrist and Pischke 2009). Matching estimators are more flexible in that misspecified functional form is less problematic; matches between groups depend on non- or semi-parametric calculations of similarity in observables and are "weighted together to produce a single overall average treatment effect" (ibid, p. 69). Different matching methods can also be used to check the consistency of the estimates. Difference-in-differences models, on the other hand, allow inclusion of fixed effects to account for unobserved heterogeneity and are more consistent and efficient in the use of continuously measured covariates. Specifications with interaction terms are also

[44] The Japanese government relied on the 1875 land tax for the bulk of its revenues in the early Meiji Period before the implementation of personal and corporate income taxes later in the century.

easier to interpret and have individual coefficient estimates to differentiate their effects.

To have a meaningful interpretation of the estimates from a difference-in-differences model, a number of issues need to be addressed. The first is in the selection of appropriate treatment and control groups, which should not differ in the trends of observables before the onset of the treatment. To check that this condition is satisfied, I exclude prefectures that gained rail access before 1884 to have data to test for pre-treatment comparability. I then run the full specification of the reduced form model only on the pre-treatment years with the panel of diseases by prefecture and substitute an indicator of future rail access (i.e., intention to treat) for the rail access variable. When interacted with the other covariates, these interaction terms identify whether the two groups differ in the trends of observables before access. As shown in the tables, virtually all of the pre-treatment results, given in the tables, have statistically insignificant results for rail access or its interactions with the other explanatory variables before treatment. The one exception, which I discuess in the robustness section, indicates a reversal of trend, which corroborates the main results.

Figure I provides a visual comparison of the geography among prefecture groups. Prefectures in the newly treated and untreated groups are similar in their dispersion across the country, size, and proximity to the major cities of Tokyo and Osaka. Prefectures in the control group were also not disproportionately disadvantaged at the start of the treatment period and include wealthy agricultural and commercial regions like Toyama and Nagasaki, respectively. The newly treated and untreated groups of prefectures also begin the treatment period with similar mortality rates among individual disease groups. However, since newly treated prefectures gain rail access over different years, their length of treatment and a possible lag in mortality may obscure direct year comparisons or period end comparisons. Besides the time-varying treatment variable and year fixed effects in

the main regression specifications, I also account for access year differences by using balanced treatment duration and period averages in the robustness checks.

A second consideration is whether the dependent variable of mortality rates influenced the decision to construct railways in particular places or times, i.e., endogeneity or reverse causality. Based on government documents from the late nineteenth century, this appears unlikely since the primary motives for building the railways were national security, market integration, and the promotion of strategic industries; there were no references to public health concerns (Japan Railway Bureau 1887, p. 44; Free 2008). Furthermore, in line with its security and economic objectives, the central government did not allow for local lobbying of placement for railways until 1892 after the passage of the Railway Construction Law (Tang 2014). Additional research using instrumental variables corroborates this interpretation and indicates that railway expansion was causal for industrialization and urbanization (Yamazaki 2017). If regions were concerned about potential health impacts or were interested in promoting economic activity that may indirectly affect public health, these changes in construction would postdate the period of analysis in this paper. Moreover, the central government also did not consider health outcomes before the 1894 Sino-Japanese War, which also postdates the treatment period.

The placement and timing of construction and routes were also subject to geographic conditions (i.e., landslides, mountain passes) and thus do not strictly adhere to shortest distances or the government's anticipated timeframe (Japan Railway Bureau 1886, p. 18).[45] As shown in Figure 1, railways radiated northward and westward from the main cities of Tokyo and Osaka, respectively,

[45] While the treatment window already minimizes potential endogeneity between rail construction and economic activity, there is no evidence that public health concerns motivated railway placement in the annual reports issued by the Japanese Railway Bureau. This is underscored by the results in the robustness checks indicating that there was a delay in observing increased mortality among rail-accessible prefectures, which officials may not have attributed to rail network integration.

along coastal areas due to the costs of construction and technical difficulty in crossing uneven terrain, thus providing an exogenous source of spatial variation. Numerous changes to route placement due to failed attempts at construction progress were made by the central government without external or local consultation during the period of analysis in this paper, and indicate that the variation in timing of railway expansion was also reasonably exogenous (Japan Railway Bureau 1887, p. 6). To corroborate this circumstantial evidence and whether one can observe differential health outcomes before actual railway construction, I employ pre-treatment specifications and placebo tests.

A final related issue is that both railroad construction and health outcomes may have been consequences of industrialization, and thus do not have a meaningful relationship with each other. Unlike the developmental experiences of earlier developing countries like the United Kingdom, Japanese industrialization occurred after railroads were introduced and triggered increased industrial agglomeration and firm scale (Tang 2014, Perkins and Tang 2017). Thus, railways may have contributed both directly to changes in mortality through increased labor mobility as well as indirectly through increased exposure to industrial activity (e.g., pollution) or urbanization (e.g., crowding). Since the latter mechanism can still be interpreted as a consequence of railway expansion and is included as a separate control variable in the analysis, the indirect effects on mortality from industrialization do not threaten the identification strategy outlined in this paper. That said, the results from the analysis suggest a more direct effect on mortality given the type of diseases experiencing increased mortality and the unusual pattern of between regions (i.e., increasing mortality in rural areas with less industrial activity), which will be further discussed in the following sections.

III. Results

Matching Estimation

The results for the matching estimation analysis are shown in Tables II and III. In the first table, I estimate average treatment effects (ATE) and average treatment effects on the treated (ATET) prefectures for annual mortality rates by disease group over the period 1883 to 1893, and include all prefectures in Japan that have mortality and industrial data.[46] These include both the groups of prefectures that gained rail access before 1884 (i.e., always treated) and those that gained access between 1884 and 1893 (i.e., newly treated) as well as the group that remained without rail before 1894 (i.e., control or untreated). I use two types of matching estimators: nearest neighbor matching (NNM) based on the Mahalanobis distance in observable variables between treated and untreated prefectures, and propensity score matching (PSM) with the same set of observables. The variables are time-varying and at the prefecture level and include land value per capita, low gradient population density, industrial firms per capita, and hospitals per capita as well as disease group. Matches using NNM are adjusted for bias given the continuously measured covariates, and the estimates are for one match between the treated and untreated prefecture groups. I report robust standard errors for all estimates.

[Table II]

The top panel in Table II uses mortality across all disease groups aggregated into a single series, while the middle two panels split the sample into those designated as communicable or not. The bottom panel has separate series for the twelve disease categories. Most ATE and all ATET estimates are positive

[46] As mentioned earlier, Okinawa prefecture is omitted from all the analyses due to missing industrial data during the period and hospital data are unavailable for year 1884. This reduces the potential number of observations from 6,204 (i.e., 47 prefectures for 12 disease groups across 11 years) to 5,520 (46 prefectures, 12 disease groups, 10 years). Individual years of covariate data (e.g., land value, firm counts) are also missing for various prefectures (see appendix B for prefectural observation totals), further reducing the sample to 5,172 observations.

and significant and indicate that rail access is associated with higher levels of mortality per capita across the country and treated prefectures. The estimates in the top panel for aggregate mortality suggest that even without specifying cause of death, which may be inaccurate given the state of medical knowledge, railways have a pronounced association with the increasing mortality over this period.

Since a rail effect can be observed in total mortality, it may be possible to observe heterogeneous effects depending on the cause of death. This can be done by separating the effects on mortality statistics by communicable and non-communicable disease groups. The magnitudes are much higher for those in the communicable disease groups, which suggest that while railways as a vector of disease transmission do not account for all of the mortality difference, they are associated with the bulks of the deaths. Rising mortality in general may be due to better reporting in government records or increased prevalence of disease transmission outside of railways, but these cannot be accounted for in the model or with the observables. The NNM and PSM estimates vary in magnitudes within each of the four panels, but are not systematically larger for one estimation method. Note that the coefficients in the lower panels sum up to those reported in the aggregate mortality total; e.g., the 22.7 NNM ATET deaths for the twelve disease group panel is approximately one-twelfth the 272.8 NNM ATET deaths in aggregate.

[Table III]

Table III further refines this analysis by separating the sample of prefectures into two groups. This allows for direct comparisons to assess whether there may have been heterogeneity among prefectures that depend on the length of treatment, which may bias the results that combine them as a single treatment group. The treatment effects are reported for mortality aggregated across all disease groups as well as in a panel of twelve individual groups. In the top panel, I compare early rail access prefectures (i.e., always treated) to those that did not

receive access before 1894 (i.e., control or untreated). The middle panel compares the always treated group to prefectures that gain rail access between 1884 and 1893, and the bottom panel compares the newly treated group to those that were untreated. These results are largely consistent with those in Table II in that most ATE and ATET estimates are positive and significant regardless of group comparison.

Magnitudes, however, are noticeably larger in the panels that include the always treated prefectures (i.e., rail access before 1884), while those comparing the newly treated to the untreated prefectures are smaller and not consistently significant for the ATE estimates. This suggests that areas that gained rail access early (e.g., Tokyo and Osaka) have a generally higher level of mortality compared to the rest of the country, and this is anecdotally supported in the population health literature (Taeuber 1858). Furthermore, in the middle panel the ATET is larger than the ATE, which may be due to a lack of endemic disease populations in the newly treated prefectures and thus a short term increase in mortality incidence from exposure. In the bottom panel, the difference in significance between the ATE and ATET estimates also suggest that the impact of rail access disproportionately raised mortality only in the rail accessible prefectures, but not when averaged across both groups. To check whether there were distributional differences between prefectures within the treatment group, in the next subsection I use the difference-in-differences framework and report the ATET estimates.

Difference-in-Differences Estimation

To provide a more nuanced picture of the impact of rail access on mortality, I use the difference-in-differences framework and specify a functional form. This also allows me to observe a clean treatment effect by comparing mortality trends in areas before they gain rail access to years following access alongside a comparison to areas that were untreated throughout the same years.

Since mortality data are available only starting in 1883, this means early rail access prefectures need to be excluded since there are no pre-treatment data to verify the comparability of the treatment and control groups. These early access prefectures also contain outliers in economic development and population (e.g., Tokyo, Osaka, Hokkaido), which may have other unobserved heterogeneity not picked up with matching estimation or available observables.[47] Thus, for most of the remaining analysis, the specifications will focus on the subset of 34 prefectures, half that gained access between 1884 and 1893 (treatment) and half without rail access throughout the period (control).[48]

[Table IV]

The regression results shown in Table IV give the difference-in-differences estimates for mortality rates in the panel of disease groups using rail access as a dummy variable. To show comparability with the matching estimation methods, both the full sample of prefectures (Columns A and B) and the subset of prefectures that excludes early access areas (Columns C and D) are used. Neither the full sample nor the subsample specifications in columns A and C have a statistically significant coefficient on rail access, in contrast to those in the matching estimates. This may be due to the inclusion of prefecture and year fixed effects, which absorb idiosyncratic shocks unaccounted for with matching estimation. Adding the interaction term with the communicable disease indicator, however, changes the findings slightly: the coefficient of the interaction is positive and weakly significant and the ATET for the sample is as well.

To check whether the treatment group of prefectures in the subset is comparable in observables to the control group, I run a pre-treatment specification

[47] These major metropolitan areas had more foreign residents or commerce, which may have introduced different strains of disease, workers commuting between regions, or possibly larger endemic disease populations, all of which are unobserved and can affect reported measures and account for the relatively higher group mortality rates estimated in Table III.

[48] In terms of observations, this reduces the possible number from 5,172 (46 prefectures) to 3,852 (34 prefectures) across twelve disease categories and complete covariate data as shown in Tables IV and V.

(Column E) that includes interactions of those observables with an intention to treat indicator in all years before the treatment group gains rail access. None of the interactions is statistically significant, suggesting that the two groups are comparable in the years before the treatment group receives rail access.

Besides rail access and disease communicability, the explanatory variable of population density (Row 3) is also consistently negative and statistically significant across most specifications. The interpretation of this finding is that less densely populated areas had disproportionately higher mortality rates with or without railways. This is consistent with earlier research that documents this unusual convergence toward higher mortality rates in urban areas (Taeuber 1958, Johansson and Mosk 1987).[49]

[Table V]

To assess whether variation in the level of treatment mattered for mortality, I substitute rail station density for the rail access indicator; results are shown in Table V. Unlike the estimates from the previous table, the coefficient on rail station density is statistically significant in the subsample, and the ATET estimates for both the baseline specification (Column C) and the one with the communicable disease interaction (Column D) are positive and highly significant. This suggests that the intensity of use more precisely reflects the impact of rail access on mortality, which corroborates the interpretation that frequency of exposure may be the mechanism driving increased mortality. Note that the ATET estimate for the baseline specification, 9.3 annual deaths per disease group in each prefecture, is roughly similar to that for the matching estimate of 9.1 in Table III. The difference-in-differences results also provide more information in that the bulk of the increase in mortality is found in communicable disease categories, 16.4 out of the 22.5 additional deaths.

[49] Taeuber (1958, p. 293) notes that villages in Japan had lower income, more ignorance, and fewer adequate medical facilities, all of which may have counteracted the potential gains from industrialization.

[Table VI]

Expanding on the baseline specifications, I include additional interaction terms between rail station density and each of the four covariates in separate and combined specifications, shown in Table VI. Of these four, both land value per capita and population density are statistically significant in their separate specifications (Columns A and B) and in combination with other interactions (Column E). The positive coefficients on land value per capita and its interaction indicate that while wealth is associated with higher mortality, this was more pronounced with rail access. This finding is unsurprising given other studies that indicate a rational tradeoff between health outcomes and increased earnings for migrants choosing to move to less healthy industrializing cities (Williamson 1982, 1990). Somewhat complicating this interpretation, however, are the negative signs on population density and its interaction with rail access, which I interpret as higher rural mortality further rising once integrated with the national rail network. While the absence of age disaggregated mortality data by region and disease group make it difficult to identify whether this was due to negative selection in health among migrants, the interactions indicate that railways catalyzed these developments.

Neither of the other two control variables, industrial firms per capita and hospitals per capita, has a statistically significant interaction with railways. This may reflect the still nascent level of industrial activity in these areas and Japan as a whole during its pre-industrialization drive as well as the unsophisticated level of medical treatment and technology. It may appear peculiar that the availability of hospitals is weakly correlated with increased mortality in all prefectures regardless of rail access. This may be due to either increased personnel to document cases of mortality as well as the facilities being used for quarantines and thus allowing disease transmission within them.

The interaction term with communicable disease is positive and weakly significant throughout all specifications, and the ATET estimates are all positive and mostly significant as well. The specifications in which there is no discrete rail impact on mortality are those with the interaction between population density and rail station density. This indicates that the observed increase in mortality among prefectures gaining rail access was unevenly distributed and largely borne among more rural areas. This disproportionate increase in rural mortality from rail access may be due to increased exposure to urban-oriented contagions before the implementation of quarantines and investment in sanitation measures like water treatment.[50] In rural areas without rail access, relative isolation provided a naturally protective environment against these ailments and thus these were not yet endemic leading to the higher average mortality observed in bigger or better connected regions (Johansson and Mosk 1987, p. 213; Taueber 1958, p. 51).

The negative relationship between mortality and the interaction of rail access and population density in the short run, while unusual compared with the experiences of other countries like the United Kingdom, has some support in the early development of the Japanese industrial economy. Since the difference-in-differences analysis excludes the largest cities of Tokyo and Osaka, those areas that remain in the sample contain mostly mid-sized municipalities and prefectural capitals. These second-tier prefectures had the widest variation in mortality rates and may have lacked of other forms of infrastructure to cope with the growth facilitated by railways (Taeuber 1958, p. 296; Cain and Hong 2009). Alongside these cities and their inadequate sources of water and sewage disposal were farms and paddies that continued to fertilize the land with human excrement, which

[50] In the case of tuberculosis, increased mobility between cities and the countryside was due to railways: "[h]ad it not been for improved transportation and the rapid growth of the textile industry, tuberculosis might have stayed in the cities... As soon as tuberculosis impaired the ability of workers' ability to perform their duties, companies invariably fired them and sent them back to their home villages. There, they often infected family members and neighbors, and in villages that had had little prior exposure to tuberculosis, mortality soared" (Johnston 1987, pp 415-16).

posed health risks with exposure to food-borne illnesses (Taeuber 1958, p. 51).[51] Gastrointestinal diseases were one of the leading killers across the country during the prewar period, along with respiratory diseases like tuberculosis and pneumonia (Honda 1997, p. 265).

IV. Robustness checks

To check the robustness of the different control variables associated with mortality, I substitute alternative measures for each of the four non-rail covariates used in the earlier difference-in-differences specifications. In Table VII, the four proxy variables for per capita income, urbanization, industrialization, and medical services are substituted with per capita capital investment; population density in non-mountainous or forested areas; the share of capital invested in industrial sectors to total capital; and the number of doctors per 100,000 inhabitants, respectively. The last column uses all four alternative measures together. These alternative measures are arguably less precise than those used in earlier specifications, even if data availability is more complete.[52] That is because capital investment excludes residential property value; low gradient land is a subset of non-mountainous or forested areas; industrial capital share is more concentrated in highly urban areas; and doctors at the time include those practicing traditional medicine.

Across all specifications, the coefficient on the interaction of rail access and communicable disease (Row 10) is positive, statistically significant, and

[51] Unlike cotton textile factories, which were based mainly in urban areas, silk filatures were found in rural locations close to silkworm producers. Factory workers in general "were subject to high rates of morbidity from tuberculosis" and that this disease "presented the largest single threat to the health of the women who worked in the textile industry" (Johnston 1995, pp. 81, 84). Many textile workers were working class, from rural areas, and recruited on short-term contracts who then returned home to convalesce or die, acting as vectors of disease in their communities.

[52] The number of observations increases in Column A since there are more data available for capitalization than land value; similarly in Column D, there are data for doctors in year 1884, unlike for hospitals.

similar in magnitude to the estimates from the previous tables. There is more variation in the average treatment effects on the treated prefectures (Row 11), which are mostly positive and significant. Population density based on habitable land type, which is a more expansive definition, has larger coefficients on its individual and interaction variables. The other two explanatory variables to proxy for income, industrial development, and medical services remain mostly insignificant as in the earlier tables.

[Table VII]

Since the panel regressions in earlier tables give estimates averaged across all years between 1883 and 1893, they may obscure the effect of rail access depending on the length of treatment. Table VIII presents results that balance the number of treatment years in the treatment group, starting with one year of rail access (Column A) and progressing up to the average annual effect of five cumulative years of treatment (Column E). These specifications indicate that the impact of rail access is progressive, with no significant effect in the first two years, but gradually increasing starting in the third. Epidemiological evidence supports this interpretation, for example tuberculosis that has a prognosis of three years to fatality if left untreated (Tiemersma et al 2011). However, since the disease groups do not differentiate mortality by all constituent diseases, it is not possible to isolate incubation and duration to fatality with the current data.[53]

[Table VIII]

To test whether the relationship between rail access and mortality is spurious, I use a placebo specification in column F that includes a balanced panel of the last year before rail access (i.e., pre-treatment). Unlike the pre-treatment

[53] The vital statistics data separately report mortality for internal pulmonary disease (in Japanese, *uchi haibyo*), but this category is an imprecise measure for pulmonary tuberculosis as it includes all diseases with lung inflammation and was not based on medical testing. The first year of mortality statistics for pulmonary tuberculosis (in Japanese, *haikekkaku*) was in 1899, post-dating the first wave of railroad expansion and the period of analysis in this paper (Japanese Statistical Association 1962, volume 21).

specification in Table IV, the interaction between future rail access and population density is weakly statistically significant. However, the sign on the coefficient is positive, which indicates that prior to rail access, future treatment prefectures had a positive association between population density and mortality. The sign reversal in years following rail access shown in earlier tables indicate that railways may have had an even larger distributional effect and worsened rural mortality. The remaining explanatory variables are statistically insignificant and there is no average treatment effect.

Finally, to verify whether the results may be driven by serial correlation in time trends, which may bias the standard errors in the estimates, I collapse the panel data of twelve diseases into two period observations, pre- and post-treatment (Bertrand et al 2004). For the pre-treatment period, this is limited to one year of data to allow all 34 prefectures in the earlier analysis to be included; for the post-treatment period, I use the average of the dependent variable and covariates for the years of 1891 to 1893. These results are shown in Table IX, which has rail access measured both as an indicator (columns A and B) and in station density (columns C and D), along with a separate specification for pre-treatment period comparison (column E). In both sets of rail access estimates, the coefficient on the interaction of railways and communicable diseases is positive and statistically significant. The ATET estimates are also all positive and mostly significant, with similar magnitudes as those in earlier tables. The pre-treatment comparison has no significant interaction terms or ATET estimate, indicating the two groups of prefectures were comparable before railways were introduced in half of them.

[Table IX]

Taken together, the results from the matching estimation and difference-in-differences regression models indicate that the introduction of railways is consistently associated with increased disease mortality in the short run,

particularly those that are communicable in nature. This relationship is robust to specifications that aggregate across disease categories, separated by communicability, or in a panel of twelve groups, thus mitigating concern about misclassified diseases, improved diagnosis, or multiple causes of deaths. The results are also consistent whether using the full sample of Japanese prefectures or the subsample that provides a clean treatment effect as well as interaction with the covariates, substitution of similar variables, length of treatment period, and averages between pre- and post-treatment years.

What may be somewhat surprising is that most of the other explanatory variables have a modest or no relationship with mortality rates and the average treatment effect, with the main exception being population density. A possible explanation for this is the short period of analysis, with the effects of economic growth not diffused as quickly as the discrete change in transport access. The period of analysis also precedes the growth of the cotton textile industry, which began in the late 1880s and was a major driver of Japanese industrialization at the turn of the twentieth century until the second world war (Braguinsky et al 2014). This may explain why the variables for industrial activity and per capita income have muted significance. At the same time, since the expansion of the railway preceded or coincided with the drive toward industrialization, this reinforces the interpretation that it was transport access as opposed to changes in the composition of industry that was largely responsible for observed mortality differences.

The weak significance of both proxies for per capita income may be due to the imprecision in using wealth measures like land value and capital investment. Wealth was also highly unevenly distributed and may overstate the actual dispersion of wages earned by the bulk of the population living outside the major metropolitan areas of Tokyo and Osaka, which are excluded from the difference-in-differences analysis (Moriguchi and Saez 2008). Regarding medical services,

government investment in public welfare and sanitation was minimal before the 1890s and both the measures of doctors and hospitals overstate their efficacy in treatment (Eli 2015). The muted changes in these measures all serve to amplify the effect of rail access, which had greater variation over this period.

Only the proxy for urbanization has some influence on mortality, but with an unexpected negative sign. This may be due to urbanization being imperfectly captured by the modified population density measures and, for the difference-in-differences results, the exclusion of prefectures with the highest urbanization since the latter had access to railways before the period of analysis and received much higher inflows of migrants. Their exclusion also may also affect the measures of types of industrial activity, since much of the investment in heavier industry firms and capital was located in these two metropolitan areas, which are not captured in the variables for industrial firms per capita or capital share.

Notwithstanding the negative short-term impact on mortality, the broader economic impact from railway expansion can be inferred from the regression estimates. Extrapolating demographic and economic changes observed in control group prefectures to the country as a whole is tenuous given unobserved externalities from railways on other sectors. Nevertheless, one can consider how the national economy would have behaved without rail network expansion in the treatment prefectures given their pre-treatment levels of activity. In the absence of rail expansion in the 1880s to early 1890s, capital investment nationwide would have been lower by approximately 17.8 percent, with a decrease of 15.8 in per capita investment (adjusting for 2.4 percent lower population growth in the counterfactual treatment prefectures), by the end of the treatment period in 1893. This makes sense since the period of analysis immediately preceded Japan's industrial takeoff in the late 1890s and saw rapid growth in manufacturing and resource extraction, with railways credited with catalyzing industrial agglomeration and scale economies (Tang 2014).

V. Conclusions

Studies of the role of transport infrastructure usually focus on changes in industrial activity, market access, and urbanization, but with less attention paid to their consequences on public health. This research identifies another channel through which railways had an aggregate and distributional impact on the economy, using detailed mortality statistics and analytical frameworks that allow for a causal interpretation. The results from this paper's analysis support the hypothesis that railway expansion, with its variation across time, space, and intensity of use, contributed to higher mortality rates. Specifically, they are associated with two-thirds of the increase in observed mortality rates among prefectures gaining access between 1884 and 1893, which is equivalent to nearly six percent of total mortality in post-rail access years. While not all of the change in mortality can be explained by railroad expansion, which is reasonable given the existence of other forms of transport and the contributions of industrial activity and urbanization, the finding that it represents such a large share is both surprising and disturbing.[54]

Disaggregation by cause of death indicates that the higher mortality incidence is due mainly to communicable diseases that could take advantage of increased labor mobility and penetration to previously sheltered rural areas. The use of rail station density as an intensive measure of utilization also adds more precision to the mortality estimates, in contrast to a simple indicator found in other studies. While data limitations prevent more detailed exploration of specific

[54] Other possible explanations for non-communicable disease group mortality also increasing with railways include the imprecision of disease classification among the twelve groups of diseases in appendix A; multiple causes of death with only one assigned as primary; and improved reporting of mortality across all disease categories with increased political centralization with railways. The data used in the analysis preclude testing these hypotheses, but may be identified with more detailed data possibly available at the city or district level.

diseases, municipalities, and the direction of transmission between regions, the preponderance of circumstantial evidence points toward railways as responsible for spreading mortality with large and quantifiable short-run effects. These findings also corroborate earlier research and case studies on why mortality rates increased during the early years of Japanese industrialization, why health outcomes varied by region, and why proximity to different forms of transport mattered (Haines et al 2003).

In a more general context, the Japanese health experience may appear unusual compared with contemporaneous Europe and North America, but not when viewed as an example of how health outcomes may lag industrial development like the United States in the ante-bellum period. Insights from Japan's relatively fast industrial transition may also be applied to non-western and peripheral economies seeking to follow its pattern (Williamson 1990, p. 52). Despite improvements in medical technology and public health measures, heterogeneity in their provision among developing countries is common. Moreover, transport infrastructure remains a primary target of public and private investment in these economies, even if the immediate implications remain unclear. In retrospect, these impacts are clear in late nineteenth century Japan, which both saw increases in military and capital investment not matched with public health spending or notable changes in income levels or medical advances (Honda 1997, p. 253). In an increasingly globalized economy, whether in historical Japan or the present day, both domestic and international markets are integrating much faster than government spending on services and public awareness. Whether on balance the long-term benefits of industrial development outweigh the short-run health hazards occurred in the process is a tradeoff that policymakers need to anticipate in their pursuit of economic growth.

REFERENCES

Abadie, A and Guido Imbens. 2006. "Large sample properties of matching estimators for average treatment effects." *Econometrica*, 74(1): 235-267.

--- 2011. "Bias-corrected matching estimators for average treatment effects." *Journal of Business and Economic Statistics*, 29: 1-11.

Allen, G.C. 1946. *A Short Economic History of Modern Japan, 1867-1937*. London: George Allen and Unwin, Ltd.

Angrist, Joshua A. and Jorn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Atack, Jeremy, Fred Bateman, Michael Haines, and Robert Margo. 2010. "Did Railroads Induce or Follow Economic Growth? Urbanization and Population Growth in the American Midwest, 1850-1860." *Social Science History*, 34(2): 171-197.

Beach, Brian and W. Walker Hanlon. Forthcoming. "Coal Smoke and Mortality in an Early Industrial Economy." *Economic Journal*.

Benetrix, Agustin, Kevin O'Rourke, and Jeffrey Williamson. 2015. "The Spread of Manufacturing to the Periphery, 1870 to 2007." *Open Economies Review* 26(1): 1-37.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics*, 119(1): 249-275, February.

Cain, Louis and Sok Chul Hong. 2009. "Survival in 19th Century Cities: The Larger the City, the Smaller Your Chances." *Explorations in Economic History*, 46(4): 450-463, October.

Centers for Disease Control and Prevention. 2016. "Zika Virus: Areas with Zika." URL: http://www.cdc.gov/zika/geo/active-countries.html, accessed 29 February 2016.

Chuo Shoin. 1995. *Ekimei Jiten* [Rail Stations in Japan], 5th edition. Japanese language text. Tokyo: Chuo Shoin Henshubu.

Clay, Karen and Werner Troesken. 2006. "Deprivation and Disease in Early Twentieth-Century America." *NBER Working Paper Series*, Number 12111. Cambridge: National Bureau of Economic Research.

Costa, Dora. 2013. "Health and the Economy in the United States, from 1750 to the Present." *NBER Working Paper Series*, Number 19685. Cambridge: National Bureau of Economic Research, November.

Costa, Dora and Richard Steckel. 1997. "Long-Term Trends in Health, Welfare, and Economic Growth in the United States." *Health and Welfare during Industrialization*. Eds. Richard Steckel and Roderick Floud. Chicago: University of Chicago Press.

Cutler, David, Angus Deaton, and Adriana Llera-Muney. 2006. "The Determinants of Mortality." *Journal of Economic Perspectives*, 20(3): 97-120, Summer.

Donaldson, David. Forthcoming. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review*.

Donaldson, David and Richard Hornbeck. Forthcoming. "Railroads and American Economic Growth: A `Market Access' Approach." *Quarterly Journal of Economics.*

Eli, Shari. 2015. "Income Effects on Health: Evidence from Union Army Pensions." *Journal of Economic History*, 75(2): 448-478, June.

Emi, Koichi and Yuichi Shionoya. 1966. "Government Expenditure." *Estimates of Long-Term Economic Statistics of Japan Since 1868*, volume 7. Eds.

Kazushi Ohkawa, Miyohei Shinohara, and Mataji Umemura. Tokyo: Toyo Keizai Shinposha.

Ferrie, Joseph. 2003. "The Rich and the Dead: Socioeconomic Status and Mortality in the United States, 1850-1860." *Health and Labor Force Participation over the Life Cycle: Evidence from the Past.* Ed. Dora Costa. Chicago: University of Chicago Press.

Ferrie, Joseph and Werner Troesken. 2008. "Water and Chicago's Mortality Transition, 1850-1925." *Explorations in Economic History*, 45: 1-16.

Flath, David. 2014. *The Japanese Economy*, third edition. Oxford: Oxford University Press.

Fogel, Robert. 1964. *Railroads and American Economic Growth: Essays in Econometric History*. Baltimore: Johns Hopkins Press.

--- 1986. "Nutrition and the Decline in Mortality since 1700: Some Preliminary Findings." *Long-Term Factors in American Economic Growth*. Eds. Stanley Engermann and Robert Gallmans. Chicago: University of Chicago Press.

Free, Dan. 2008. *Early Japanese Railways, 1853-1914.* Tokyo: Tuttle Publishing.

Gagnon, Alain, Marc Tremblay, Helene Vezina, and Jamie Seabrook. 2011. "Once Were Farmers: Occupation, Social Mobility, and Mortality During Industrialization in Saguenay-Lac-Saint-Jean, Quebec 1840-1971." *Explorations in Economic History*, 48: 429-440.

Haines, Michael, Lee Craig, and Thomas Weiss. 2003. "The Short and the Dead: Nutrition, Mortality, and the `Antebellum Puzzle' in the United States." *Journal of Economic History*, 63(2): 382-413, June.

Hatton, Timothy. 2014. "How Have Europeans Grown So Tall?" *Oxford Economic Papers*, 66: 349-372.

Hayami, Yujiro and Vernon Ruttan. 1970. “Korean Rice, Taiwan Rice, and Japanese Agricultural Stagnation: An Economic Consequence of Colonialism.” *Quarterly Journal of Economics* 84(4): 562-589.

Higgs, Robert. 1973. "Mortality in Rural America, 1870-1920: Estimates and Conjectures." *Explorations in Economic History*, 10: 177-195.

Honda, Gail. 1997. "Differential Structure, Differential Health: Industrialization in Japan, 1868-1940." *Health and Welfare during Industrialization*. Eds. Richard Steckel and Roderick Floud. Chicago: University of Chicago Press.

Japan Railway Bureau. 1886. *Meiji Juukudo Tetsudou Kyoku Nenpo* [Meiji Year 19 Railway Bureau Yearbook]. Tokyo: Tetsudou Kyoku.

--- 1887. *Meiji Nijuudo Tetsudou Kyoku Nenpo* [Meiji Year 20 Railway Bureau Yearbook]. Tokyo: Tetsudou Kyoku.

Japan Statistical Association. 1962. *Nihon Teikoku Tokei Nenkan* [Statistical Yearbook of the Japanese Empire], volumes 1-33. Tokyo: Tokyo Ripurinto Shuppansha.

--- 2007. *Historical Statistics of Japan: New Edition*. Tokyo: Japan Statistical Association.

Johansson, S. Ryan and Carl Mosk. 1987. "Exposure, Resistance and Life Expectancy: Disease and Death during the Economic Development of Japan, 1900-1960." *Population Studies*, 41: 207-235.

Johnston, William. 1987. *Disease, medicine, and the state: a social history of tuberculosis in Japan, 1850-1950*. Doctoral dissertation. Cambridge: Harvard University.

--- 1995. *The Modern Epidemic: A History of Tuberculosis in Japan*. Cambridge: Harvard University Press.

Jorda, Oscar, Moritz Schularick, and Alan M. Taylor. 2017. "Macrofinancial History and the New Business Cycle Facts." *NBER Macroeconomics*

Annual 2016, volume 31. Eds. Martin Eichenbaum and Jonathan A. Parker. Chicago: University of Chicago Press.

Komlos, John. 2012. "A Three-Decade History of the Ante-Bellum Puzzle: Explaining the Shrinking of the U.S. Population at the Onset of Modern Economic Growth." *Journal of the Historical Society*, 12(4): 395-445.

---. 1987. "Shrinking in a Growing Economy? The Mystery of Physical Stature during the Industrial Revolution." *Journal of Economic History,* 58(3): 779-802.

Lewchuk, Wayne. 1991. "Industrialization and Occupational Mortality in France Prior to 1914." *Explorations in Economic History*, 28: 344-366.

Lindert, Peter. 1983. "English Living Standards, Population Growth, and Wrigley-Schofield." *Explorations in Economic History*, 20: 131-155.

Macintyre, Sally. 1997. "The Black Report and Beyond: What are the Issues?" *Social Science and Medicine*, 44(6): 723-745.

Margo, Robert. 2000. *Wages and Labor Markets in the United States, 1820-1860.* Chicago: University of Chicago Press.

Margo, Robert and Richard Steckel. 1983. "Heights of Native Born Northern Whites during the Antebellum Period." *Journal of Economic History*, 43(1): 167-174.

McKeown, Thomas. 1976. *The Modern Rise of Population*. New York, New York: Academic Press.

Meeker, Edward. 1972. "The Improving Health of the United States, 1850-1915." *Explorations in Economic History*, 9: 353-373.

Merriam-Webster. 2016. Medical definitions of "communicable", "contagious", and "infectious" disease. URL: http://www.merriam-webster.com/medical. Accessed 1 September 2016.

Millward, Robert and Joerg Baten. 2010. "Population and Living Standards, 1914-1945." *The Cambridge Economic History of Modern Europe*, volume 2. Cambridge: Cambridge University Press.

Mitchener, Kris, Masato Shizume, and Marc Weidenmier. 2010. "Why Did Countries Adopt the Gold Standard? Lessons from Japan." *Journal of Economic History*, 70(1): 27-56.

Morse, Stephen. 2001. "Factors in the Emergence of Infectious Diseases." *Plagues and Politics*. Ed. Andrew Price-Smith. London: Palgrave Macmillan.

Moriguchi, Chiaki and Emmanuel Saez. 2008. "The Evolution of Income Concentration in Japan, 1886-2005: Evidence from Income Tax Statistics." *Review of Economics and Statistics*, 90(4): 713-734.

Moriyama, Iwao, Ruth Loy, and Alastair Robb-Smith. 2011. *History of the Statistical Classification of Diseases and Causes of Death*. Eds. Harry Rosenberg and Donna Hoyert. Washington: U.S. Government Printing Office.

Onji, Kazuki and John P. Tang. 2017. "A Nation without a Corporate Income Tax: Evidence from Nineteenth Century Japan." *Journal of Economic History*, 77(2): 440-472.

Perkins, Dwight and John P. Tang. 2017. "East Asian Industrial Pioneers: Japan, Korea, and Taiwan." *The Spread of Modern Industry to the Periphery since 1871*. Eds. Kevin O'Rourke and Jeffrey Williamson. Oxford: Oxford University Press.

Preston, Samuel. 2007. "The Changing Relation Between Mortality and Level of Economic Development." *International Journal of Epidemiology*, 36: 484-490.

Preston, Samuel and Etienne van de Walle. 1978. "Urban French Mortality in the Nineteenth Century." *Population Studies*, 32: 275-297.

Rosenberg, Charles. 1987. *The Cholera Years: The United States in 1832, 1849, and 1866*. Chicago: University of Chicago Press.

Sandberg, Lars and Richard Steckel. 1997. "Was Industrialization Hazardous to Your Health? Not in Sweden!" *Health and Welfare During Industrialization.* Eds. Richard Steckel and Roderick Floud. Chicago: University of Chicago Press.

Summerhill, William. 2005. "Big Social Savings in a Small Laggard Economy: Railroad-Led Growth in Brazil." *Journal of Economic History*, 65(1): 72-102.

Sussman, Nathan and Yishay Yafeh. 2000. "Institutions, Reforms, and Country Risk: Lessons from Japanese Government Debt in the Meiji Era." *Journal of Economic History*, 60(2): 442-267.

Szreter, Simon. 2004. "Industrialization and Health." *British Medical Bulletin*, 69: 75-86.

Szreter, Simon and G. Mooney. 1998. "Urbanisation, Mortality and the Standard of Living Debate: New Estimates of the Expectation of Life at Birth in 19th Century British Cities." *Economic History Review*, 50: 84-112.

Taeuber, Irene. 1958. *The Population of Japan*. Princeton, New Jersey: Princeton University Press.

Tang, John P. 2011. "Technological Leadership and Late Development: Evidence from Meiji Japan, 1868 to 1912." *Economic History Review*, 64(S1): 99-116.

--- 2013. "Financial Intermediation and Late Development: Evidence from Meiji Japan." *Financial History Review*, 20(2): 111-135.

--- 2014. "Railroad Expansion and Industrialization: Evidence from Meiji Japan." *Journal of Economic History*, 74(30): 863-886.

--- 2015. "Pollution Havens and the Trade in Toxic Chemicals: Evidence from U.S. Trade Flows." *Ecological Economics*, 112: 150-160.

--- 2016. "A Tale of Two SICs: Industrial Development in Japan and the United States in the Late Nineteenth Century." *Australian Economic History Review*, 56(2): 174-197.

Tiemersma, Edine W., Marieke J. van der Werf, Martien W. Borgdorff, Brian G. Williams, and Nico J. D. Nagelkerke. 2011. "Natural History of Tuberculosis: Duration and Fatility of Untreated Pulmonary Tuberculosis in HIV Negative Patients: A Systematic Review." *PLoS ONE* 6(4): e17601.

Walker, Brett. 2009. *Toxic Archipelago: A History of Industrial Disease in Japan*. Seattle: University of Washington Press.

Williamson, Jeffrey. 1982. "Was the Industrial Revolution Worth It? Disamenities and Death in 19th Century British Towns." *Explorations in Economic History* 19: 221-245.

--- 1990. *Coping with City Growth During the British Industrial Revolution*. Cambridge: Cambridge University Press.

World Health Organization. 2016a. "Fact sheets: infectious diseases." URL: http://www.who.int/topics/infectious_diseases/factsheets/en/, accessed 1 September 2016.

--- 2016b. "Zika virus fact sheet." URL: http://www.who.int/mediacentre/factsheets/zika/en/, accessed 29 February 2016.

Yamazaki, Junichi. 2017. "Railroads, technology adoption, and modern economic development: evidence from Japan." Institute of Social and Economic Research Discussion Paper 1000. Osaka: Institute of Social and Economic Research.

Yamazawa, Ippei and Yuzo Yamamoto. 1979. "Foreign Trade and Balance of Payments." *Estimates of Long-Term Economic Statistics of Japan Since*

1868, volume 14. Eds. Kazushi Ohkawa, Miyohei Shinohara, and Mataji Umemura. Tokyo: Toyo Keizai Shinposha.

APPENDIX A: DISEASE CATEGORIES

The following breakdown of included diseases for each disease group in the analysis comes from the Cabinet Bureau of Statistics, although the two disease categories of poisoning and unclear illness are not further disaggregated (Japan Statistical Association 1962, volume 9, table 243).

blood (9): heart inflammation, cardiac hypertrophy, intracardial meningitis, valvular disease, inner heart inflammation, heart attack, artherosclerosis, arterial disease, venous disease

bone and joint (9): bone and joint inflammation, bone inflammation, periostitis, osteomyelitis, bone ulcer, bone necrosis, joint abcess, bone injury, osteomalacia

developmental and nutritional (18): structural abnormality, hypoplasia and stunting, dental illness, adenopathy and swollen lymph nodes, rickets, gout, leprosy, edema, diabetes, gangrene, cancer and tumors, nutritional deficiency, *shikoujibaigutsu*, anemia, chlorosis, leukemia, senility, goiter

digestive (29): mouth and tongue disorder, parotid gland disorder, stomach disorder, peritonitis, abdominal abcess, intestinal parasite, ascites, hernia, intestinal blockage, abdominal catarrh, gastric ulcer, stomach stenosis, gastric distension, stomach pain, hematemesis, intestinal bleeding, diarrhea, intestinal catarrh, acute gastroenteritis, chronic gastroenteritis, appendicitis, intestinal disorder, mesentery disorder, spleen disorder, cholelithiasis and gallstones, jaundice, acute hepatitis, chronic hepatitis

external injury (14): burns, frostbite, electric shock, crushing, shooting, incision, stabbing, bites, bruising, suffocation, hanging, strangling, drowning, suicide

infectious (27): typhoid, typhus, dysentery, Asia cholera, diphtheria, croup, smallpox, measles, scarlet fever, beriberi, undulant fever, Ross fever, sepsis, septicemia and blood poisoning, hospital gangrene, whooping cough, puerperal fever, contagious parotitis and mumps, malignant salivary gland inflammation, cerebrospinal meningitis, rheumatism, gonorrhea, syphilis, herpes, chancre, rabies, anthrax

nervous (21): meningitis, brain edema, encephalitis, stroke, cerebral embolism, cerebral palsy, cerebral anemia, cerebral hyperemia, insanity, spinal myelitis, spinal meningitis, spinal exhaustion, spinal paralysis, hypochondria, eclampsia, pediatric epilepsy, hysteria, epilepsy, chorea, tetanus, ear disease

poisoning (1)

respiratory (16): laryngitis, tracheal tuberculosis, bronchodialation, pneumonia, tuberculosis, hemoptysis and lung hemorrhage, emphysema, asthma, lung inflation failure, lung gangrene, lung paralysis, pulmonary edema, pleurisy, pleural effusion

skin and muscular (8): bleeding ulcers, subcutaneous ligament inflammation, carbuncle, mange, myositis and muscle inflammation, phlebitis and vein inflammation, umbilical disorder, muscular atrophy

urogenital (12): urinary cystitis, bladder stones, urethral stricture, uremic, orchitis, Bright's disease, vaginal catarrh, endometritis, uterine prolapse, uterine bleeding, uterine cramps, ovarian cyst

unclear (1)

APPENDIX B: PREFECTURAL DATA DESCRIPTION

Prefectures	Rail Access Year	Stations per 100k (1893)	Mortality per 100k (1893)	Observations (1883-1893)
Aichi	1886	1.0	2,318	96
Akita	1899	0	2,064	120
Aomori	1891	1.6	1,999	108
Chiba	1894	0	1,933	108
Ehime	1914	0	2,287	120
Fukui	1882	0.2	2,195	120
Fukuoka	1889	2.1	2,226	120
Fukushima	1887	0.9	2,096	120
Gifu	1883	0.4	2,282	96
Gunma	1883	1.8	2,217	120
Hiroshima	1891	0.3	2,233	120
Hokkaido	1880	5.1	2,235	96
Hyogo	1874	1.3	2,566	120
Ibaraki	1885	1.0	1,985	120
Ishikawa	1897	0	2,257	120
Iwate	1890	2.2	2,028	108
Kagawa	1889	0.6	2,253	60
Kagoshima	1900	0	1,604	120
Kanagawa	1872	1.5	1,733	120
Kochi	1924	0	2,126	120
Kumamoto	1891	0.6	2,126	120
Kyoto	1876	0.4	2,220	120
Mie	1890	1.8	2,106	108
Miyagi	1887	1.5	2,225	120
Miyazaki	1911	0	2,080	108
Nagano	1888	1.0	2,206	108
Nagasaki	1897	0	2,009	108
Nara	1890	1.6	2,473	72
Niigata	1886	0.3	2,464	120
Oita	1897	0	2,463	120
Okayama	1890	1.0	2,183	120
Okinawa	2003	0	2,473	0

Source: Author's calculations based on Chuo Shoin (1995) and Cabinet Bureau of Statistics, Japan Statistical Association (1962). The number of observations for each prefecture is for twelve disease categories for each year of available mortality and covariate data ; see footnote 36 for individual years of missing data by prefecture.

APPENDIX B: PREFECTURAL DATA DESCRIPTION

Prefectures	Rail Access Year	Stations per 100k (1893)	Mortality per 100k (1893)	Observations (1883-1893)
Osaka	1874	0.8	3,092	120
Saga	1889	0.9	2,345	108
Saitama	1883	1.2	2,117	108
Shiga	1880	2.1	2,333	120
Shimane	1908	0	2,215	120
Shizuoka	1888	1.9	1,802	108
Tochigi	1885	2.9	2,209	120
Tokushima	1899	0	2,421	120
Tokyo	1872	1.2	2,439	120
Tottori	1902	0	2,286	120
Toyama	1897	0	2,365	108
Wakayama	1898	0	2,199	120
Yamagata	1901	0	2,401	108
Yamaguchi	1899	0	2,133	120
Yamanashi	1897	0	2,290	96

Source: Author's calculations based on Chuo Shoin (1995) and Cabinet Bureau of Statistics, Japan Statistical Association (1962). The number of observations for each prefecture is for twelve disease categories for each year of available mortality and covariate data; see footnote 36 for individual years of missing data by prefecture.

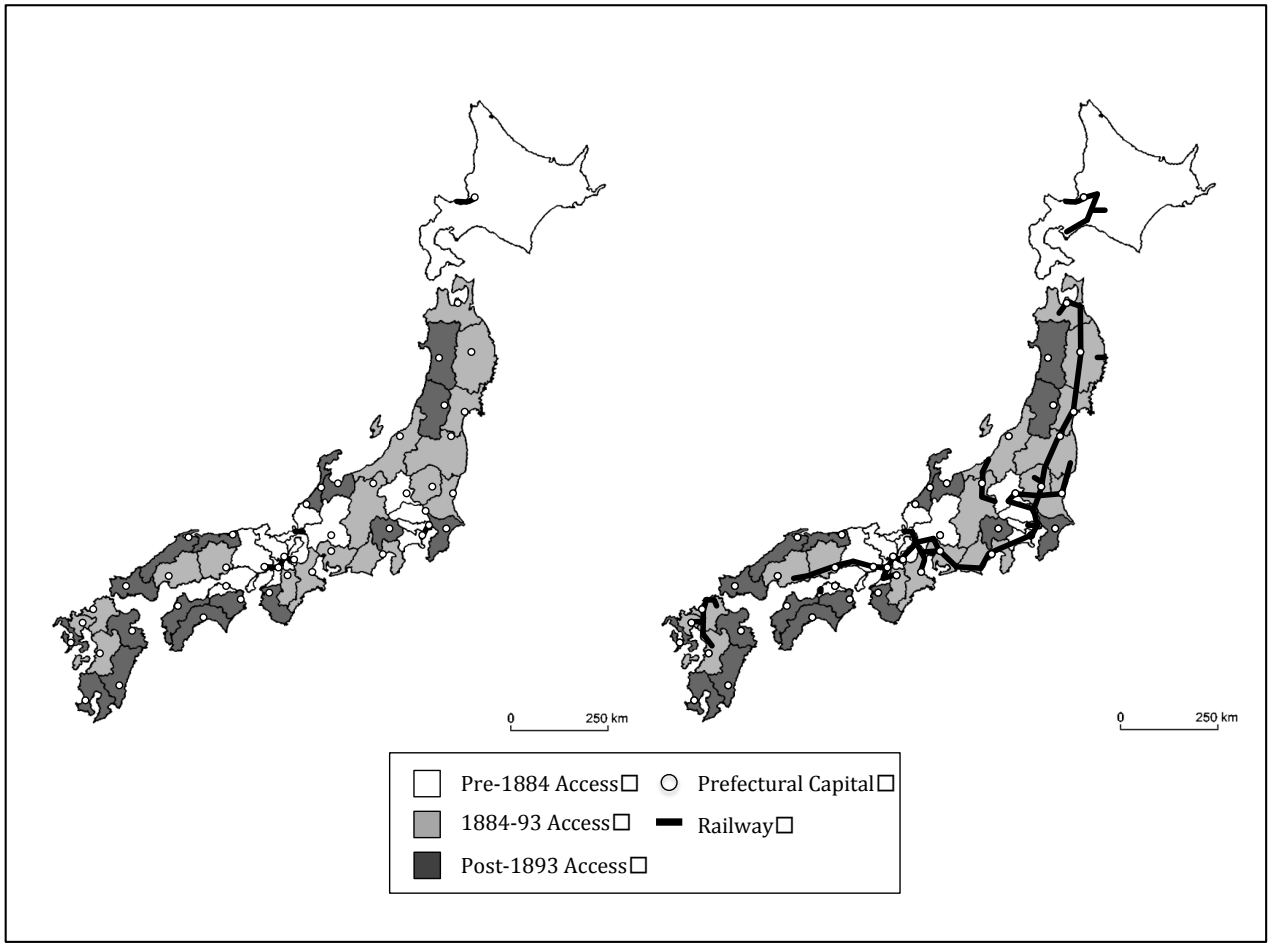

Figure I: Japanese Rail Network, 1883 and 1893

Source: see text. Okinawa prefecture is not shown.

Table I: Mortality Rates, Prefectural Averages

Year	1883	1887	1891	1895	1899
Population	881,364[a]	862,965[b]	871,677	902,787	966,022
Annual percent change		0.1	0.8	0.8	1.6
Prefectures with rail access		2.5	1.2	1.0	1.7
Prefectures without rail access		-1.0	0.3	0.5	1.1
Mortality per 100k	1,661.3	1,934.9	2,059.2	2,001.8	2,025.5
Annual percent change		4.4	0.7	-1.3	-0.7
Prefectures with rail access		0.1	1.8	-1.7	-0.9
Prefectures without rail access		6.4	-0.5	-0.6	-0.3
Disease Groups					
Communicable, Total	785.2	920.3	1,012.3	986.2	978.6
Acute Infectious	111.3	149.8	143.0	206.5	149.8
Digestive	446.5	448.8	497.7	423.9	430.6
Respiratory	227.4	321.7	371.6	355.8	398.2
Non-Communicable, Total	876.1	1,014.5	1,046.8	1,015.5	1,046.8
Blood Related	38.8	62.9	76.7	86.6	107.6
Bone and Joint	9.2	16.1	19.8	20.2	23.9
Developmental	298.7	319.1	327.9	307.4	315.1
External Injury	38.1	42.4	61.4	46.4	50.3
Nervous	368.2	418.7	419.1	408.6	417.3
Poisoning	1.7	1.0	1.4	1.5	2.6
Skin and Muscle	25.7	31.6	30.5	26.9	31.3
Urogenital	51.5	64.5	66.0	70.0	77.4
Unclear	44.2	58.2	44.0	47.9	21.3

Notes: [a]In 1883, two prefectures did not have established borders (Kagawa and Nara) and thus are missing data. Also missing data are Miyazaki, Saga, and Toyama. [b]In 1887, Kagawa prefecture remained without borders and Nara prefecture is missing data. Population based on resident statistics. Annual percentage change for population and mortality rates calculated using previous four years except where noted.

Source: Author's calculations based on Cabinet Bureau of Statistics, Japan Statistical Association (1962).

Table II: Matching Estimation Results, All Prefectures

DV: Annual Prefecture Mortality per 100k	ATE	ATE on Treated	Observations	Method
Aggregated Total (1 series)	242.540***	272.817***	431	NNM

	(46.507)	(53.163)		
	243.786***	186.994***	431	PSM
	(59.600)	(52.794)		
Communicable Disease Group (3 series)	59.003***	65.639***	1,293	NNM
	(8.368)	(9.201)		
	34.082**	62.076***	1,293	PSM
	(17.402)	(19.436)		
Non-communicable Disease Group (9 series)	7.281***	8.433***	3,879	NNM
	(1.882)	(2.529)		
	7.264	22.984**	3,879	PSM
	(12.544)	(10.999)		
All Disease Groups (12 series)	20.212***	22.735***	5,172	NNM
	(2.518)	(2.961)		
	14.462	65.085***	5,172	PSM
	(23.782)	(17.051)		

Notes: All average treatment effects are for the treatment of rail access as an indicator variable with one match per series between treated and control groups. Nearest neighbor matching (NNM) uses exact matching on disease group and Mahalanobis distance on observable covariates, which are corrected for large sample bias. Propensity score matching (PSM) estimates are calculated using a logistic model on the same observables as in the nearest neighbor matching model. Okinawa prefecture is missing covariate data and thus dropped from the analysis. Robust standard errors reported.

Significance levels: *10%, **5%, ***1%

Table III: Matching Estimation Results, Prefecture Group Comparison

DV: Annual Prefecture Mortality per 100k	ATE	ATE on Treated	Observations	Method
Pre-1884 Rail Access (Always Treated) and Post-1893 Rail Access (Never Treated)				
Aggregated Total (1 series)	370.715***	335.213***	273	NNM
	(73.842)	(82.733)		
	382.751***	277.694***	271	PSM
	(104.945)	(68.956)		
All Disease Groups (12 series)	30.893***	27.934***	3,276	NNM
	(3.955)	(4.409)		
	31.237	71.559***	3,252	PSM
	(20.250)	(22.193)		
Pre-1884 Rail Access (Always Treated) and 1884-1893 Rail Access (New Treatment)				
Aggregated Total (1 series)	360.211***	438.212***	268	NNM
	(66.207)	(72.797)		
	280.081***	351.853***	268	PSM
	(51.604)	(55.322)		
All Disease Groups (12 series)	30.018***	36.518***	3,216	NNM
	(3.441)	(3.787)		
	46.037***	72.547***	3,216	PSM
	(16.591)	(21.362)		
1884-1893 Rail Access (New Treatment) and Post-1893 Rail Access (Never Treated)				
Aggregated Total (1 series)	61.565	109.497***	321	NNM
	(43.476)	(41.952)		
	-19.896	65.374	321	PSM
	(46.086)	(58.074)		
All Disease Groups (12 series)	5.130**	9.125***	3,852	NNM
	(2.452)	(2.855)		
	-14.066	90.007***	3,852	PSM
	(18.950)	(18.235)		

Notes: All average treatment effects are for the treatment of rail access as an indicator variable with one match per series between treated and control groups. Nearest neighbor matching (NNM) uses exact matching on disease group and Mahalanobis distance on observable covariates, which are corrected for large sample bias. Propensity score matching (PSM) estimates are calculated using a logistic model on the same observables as in the nearest neighbor matching model. Okinawa prefecture is missing covariate data and thus dropped from the analysis. Robust standard errors reported.

Significance levels: *10%, **5%, ***1%

Table IV: Disease Panel Regression Results, Rail Indicator

DV: Mortality per 100k	A	B	C	D	E
					Pre-treatment
Rail Access indicator	4.318	-0.236	2.919	-1.636	dropped
	(3.498)	(2.867)	(3.657)	(3.152)	
Land Value p.c., 1990	27.741**	27.741**	18.923*	18.923*	12.203
constant yen	(11.654)	(11.625)	(10.987)	(10.953)	(11.425)
Pop. Density, low gradient	-11.528**	-11.528**	-10.362*	-10.362*	-10.932**
	(5.462)	(5.462)	(5.538)	(5.538)	(5.505)
Industrial Firms per 100k	-0.317	-0.317	-0.373	-0.373	-0.485
	(0.405)	(0.406)	(0.443)	(0.446)	(0.534)
Hospitals per 100k	2.703*	2.703*	6.491	6.491	8.488
	(1.487)	(1.487)	(4.068)	(4.069)	(6.346)
Interaction with Rail					
Communicable indicator		18.218*		18.218*	
		(10.375)		(10.372)	
Land Value p.c., 1990					19.795
constant yen					(25.367)
Pop. Density, low gradient					-19.492
					(59.847)
Industrial Firms per 100k					0.382
					(1.628)
Hospitals per 100k					-2.789
					(9.379)
ATE on Treated	4.318	17.982*	2.919	16.582	-40.079
	(3.498)	(10.206)	(3.657)	(10.184)	(139.714)
Year Coverage	1883-93	1883-93	1883-93	1883-93	1883-93
Disease Groups	12	12	12	12	12
Prefectures	46	46	34	34	34
Observations	5,172	5,172	3,852	3,852	2,748
F-statistic	17.55***	17.00***	15.07***	14.87***	10.39***
Within R-squared	0.083	0.085	0.083	0.085	0.103

Notes: All specifications include fixed effects for disease group, prefecture, and year. Robust standard errors are clustered by disease group and prefecture. The average treatment effect of rail access on treated prefectures is calculated for rail access and its interactions over the treatment period. The pre-treatment specification includes all interactions with prefecture-level control variables and rail access. Okinawa prefecture is missing land value data and thus dropped from all specifications. Kagawa prefecture is missing pre-treatment data and thus omitted from restricted sample specifications.

Significance levels: *10%, **5%, ***1%

Table V: Disease Panel Regression Results, Rail Station Density

DV: Mortality per 100k	A	B	C	D	E
					Pre-treatment
Rail Stations per 100k	4.507	1.392	8.879***	4.670	dropped
	(3.067)	(2.598)	(3.376)	(2.849)	
Land Value p.c., 1990	27.997**	27.997**	17.353	17.353	12.203
constant yen	(11.720)	(11.694)	(11.005)	(10.964)	(11.425)
Pop. Density, low gradient	-11.747**	-11.747**	-11.212**	-11.212**	-10.932**
	(5.469)	(5.476)	(5.491)	(5.491)	(5.505)
Industrial Firms per 100k	-0.290	-0.290	-0.267	-0.267	-0.485
	(0.413)	(0.412)	(0.446)	(0.449)	(0.534)
Hospitals per 100k	2.423	2.423	7.376*	7.376*	8.488
	(1.533)	(1.502)	(4.104)	(4.103)	(6.346)
Interaction with Rail					
Communicable indicator		12.461		16.838*	
		(7.810)		(8.943)	
Land Value p.c., 1990					19.795
constant yen					(25.367)
Pop. Density, low gradient					-19.492
					(59.847)
Industrial Firms per 100k					0.382
					(1.628)
Hospitals per 100k					-2.789
					(9.379)
ATE on Treated	4.621	14.202*	9.292***	22.507**	-40.079
	(3.144)	(8.129)	(3.533)	(9.298)	(139.714)
Year Coverage	1883-93	1883-93	1883-93	1883-93	1883-93
Disease Groups	12	12	12	12	12
Prefectures	46	46	34	34	34
Observations	5,172	5,172	3,852	3,852	2,748
F-statistic	18.25***	18.09***	15.22***	15.36***	10.39***
Within R-squared	0.084	0.085	0.086	0.089	0.103

Notes: All specifications include fixed effects for disease group, prefecture, and year. Robust standard errors are clustered by disease group and prefecture. The average treatment effect of rail access on treated prefectures is calculated for rail station density and its interactions over the treatment period. The pre-treatment specification includes all interactions with prefecture-level control variables and rail access. Okinawa prefecture is missing land value data and thus dropped from all specifications. Kagawa prefecture is missing pre-treatment data and thus omitted from restricted sample specifications.

Significance levels: *10%, **5%, ***1%

Table VI: Disease Panel Regression Results, Interaction Effects

DV: Mortality per 100k	A	B	C	D	E
Rail Stations per 100k	-15.048	12.754***	2.667	5.365	-17.250
	(12.217)	(4.786)	(4.286)	(5.642)	(17.054)
Land Value p.c., 1990	18.703*	16.376	17.056	17.383	17.782
constant yen	(11.138)	(10.906)	(10.927)	(10.977)	(11.018)
Pop. Density, low gradient	-11.308**	-10.768*	-11.326**	-11.228**	-10.918**
	(5.493)	(5.494)	(5.501)	(5.488)	(5.515)
Industrial Firms per 100k	-0.289	-0.264	-0.356	-0.271	-0.416
	(0.448)	(0.447)	(0.463)	(0.447)	(0.462)
Hospitals per 100k	6.310	7.941*	7.135*	7.336*	6.111
	(4.125)	(4.119)	(4.157)	(4.071)	(4.209)
Interaction with Rail					
Communicable indicator	16.838*	16.838*	16.838*	16.838*	16.838*
	(8.788)	(8.863)	(8.906)	(8.929)	(8.516)
Land Value p.c., 1990	18.803*				44.368**
constant yen	(11.138)				(19.093)
Pop. Density, low gradient		-8.860**			-12.264***
		(3.623)			(3.764)
Industrial Firms per 100k			0.483		0.635
			(0.833)		(0.786)
Hospitals per 100k				-0.494	-0.526
				(3.337)	(3.708)
ATE on Treated	37.073***	7.657	20.993**	30.975***	26.277
	(12.419)	(10.891)	(9.571)	(10.175)	(16.348)
Year Coverage	1883-93	1883-93	1883-93	1883-93	1883-93
Disease Groups	12	12	12	12	12
Prefectures	34	34	34	34	34
Observations	3,852	3,852	3,852	3,852	3,852
F-statistic	14.44***	14.32***	14.40***	14.52***	12.46***
Within R-squared	0.090	0.091	0.089	0.089	0.093

Notes: All specifications include fixed effects for disease group, prefecture, and year. Robust standard errors are clustered by disease group and prefecture. The average treatment effect of rail access on treated prefectures is calculated for rail station density and its interactions over the treatment period. Okinawa prefecture is missing land value data and thus dropped from the analysis. Kagawa prefecture is missing pre-treatment data and thus omitted from specifications.

Significance levels: *10%, **5%, ***1%

Table VII: Disease Panel Regression Results, Alternative Control Variables

DV: Mortality per 100k	A	B	C	D	E
Rail Stations per 100k	3.835	5.172*	4.076	0.543	-2.026
	(3.077)	(2.872)	(2.910)	(2.635)	(10.810)
Land Value p.c., 1990		17.003*	17.687	5.876	
constant yen		(10.193)	(10.988)	(8.611)	
Pop. Density, low gradient	-12.954**		-11.284**	-18.491***	
	(5.089)		(5.501)	(4.914)	
Industrial Firms per 100k	-0.367	-0.248		-0.292	
	(0.436)	(0.448)		(0.477)	
Hospitals per 100k	6.848	7.674*	6.550		
	(4.234)	(4.129)	(4.221)		
Alternative Measures					
Capitalization p.c., 1990	37.673				47.324
constant yen	(48.933)				(46.344)
Pop. Density, arable land		-23.307**			-40.314***
		(11.527)			(10.149)
Industrial Capital Share			-0.080*		-0.009
			(0.045)		(0.051)
Doctors per 100k				0.116	0.081
				(0.149)	(0.137)
Interaction with Rail					
Communicable indicator	16.999*	16.838*	16.838*	21.659***	21.774***
	(8.818)	(8.968)	(8.880)	(7.885)	(7.838)
ATE on Treated	21.772**	23.032**	21.886**	23.232***	-3.750
	(8.924)	(9.323)	(9.230)	(8.236)	(18.152)
Year Coverage	1883-93	1883-93	1883-93	1883-93	1883-93
Disease Groups	12	12	12	12	12
Prefectures	34	34	34	34	34
Observations	3,888	3,852	3,852	4,224	4,260
F-statistic	15.32***	15.41***	15.33***	16.26***	13.15***
Within R-squared	0.089	0.089	0.090	0.108	0.109

Notes: All specifications include fixed effects for disease group, prefecture, and year. Robust standard errors are clustered by disease group and prefecture. The average treatment effect of rail access is calculated for station density and its interactions over the treatment period. Okinawa prefecture is missing land value data and thus omitted from some specifications. Kagawa prefecture is missing pre-treatment data and thus is omitted from all specifications.

Significance levels: *10%, **5%, ***1%

Table VIII: Disease Panel Regression Results, Balanced Treatment Periods

DV: Mortality per 100k	A	B	C	D	E	F
Rail Stations per 100k	0.060	3.101	4.209*	2.672	2.720	dropped
	(3.982)	(2.669)	(2.385)	(2.351)	(2.487)	
Land Value p.c., 1990	17.460	20.544*	18.044	17.646	17.667	14.001
constant yen	(11.094)	(11.448)	(11.459)	(11.374)	(11.316)	(11.515)
Pop. Density, low	-12.092**	-11.624**	-11.217**	-10.933**	-10.833**	-10.699*
gradient	(5.574)	(5.570)	(5.534)	(5.519)	(5.508)	(5.457)
Industrial Firms per 100k	-0.244	-0.207	-0.311	-0.309	-0.313	-0.503
	(0.502)	(0.502)	(0.497)	(0.491)	(0.488)	(0.537)
Hospitals per 100k	5.055	5.194	6.331	6.134	6.430	8.587
	(5.276)	(5.040)	(4.724)	(4.560)	(4.502)	(6.279)
Interaction with Rail						
Communicable indicator	-13.706	7.339	10.326	13.225	15.079*	dropped
	(12.747)	(11.394)	(10.398)	(8.884)	(8.550)	
Land Value p.c., 1990						-18.027
constant yen						(19.197)
Pop. Density, low						120.428*
gradient						(71.481)
Industrial Firms per 100k						1.762
						(1.203)
Hospitals per 100k						2.397
						(8.645)
ATE on Treated	-2.395	3.449	6.656	8.874*	10.921**	266.509
	(2.174)	(3.712)	(4.699)	(4.921)	(5.245)	(162.492)
Treatment Duration	1 year	2 years	3 years	4 years	5 years	-1 year
						(Placebo)
Disease Groups	12	12	12	12	12	12
Prefectures	34	34	34	34	34	34
Observations	2,640	2,832	3,036	3,204	3,324	2,436
F-statistic	8.40***	9.60***	10.66***	10.76***	12.11***	8.12***
Within R-squared	0.079	0.080	0.079	0.077	0.078	0.097

Notes: All specifications include fixed effects for disease group, prefecture, and year. Robust standard errors are clustered by disease group and prefecture. The average treatment effect of rail access is calculated for station density and its interactions over the treatment period. Okinawa prefecture is missing land value data and Kagawa prefecture is missing pre-treatment data, so both are omitted from specifications.

Significance levels: *10%, **5%, ***1%

Table IX: Disease Panel Regression Results, Grouped Periods

DV: Mortality (per 100k)	A	B	C	D	E
	Rail indicator	Rail indicator	Station density	Station density	Pre-treatment
Rail Access	5.244	-1.982	11.284**	3.598	0.591
	(4.771)	(4.297)	(4.493)	(3.755)	(22.567)
Land Value p.c., 1990	32.386	32.386	26.661	26.661	dropped
constant yen	(28.109)	(27.586)	(28.335)	(27.791)	
Pop. Density, low	-8.101	-8.101	-9.158*	-9.158*	-3.739
gradient	(5.255)	(5.259)	(5.185)	(5.186)	(3.435)
Industrial Firms per 100k	-0.042	-0.042	-0.314	-0.314	-2.306
	(0.807)	(0.805)	(0.825)	(0.824)	(2.768)
Hospitals per 100k	8.245	8.245	11.131	11.131	-8.590
	(8.405)	(8.351)	(8.227)	(8.252)	(9.899)
Interaction with Rail					
Communicable indicator		28.906***		30.745***	-5.226
		(8.994)		(11.058)	(15.949)
Land Value p.c., 1990					14.064
constant yen					(46.662)
Pop. Density, low					5.296
gradient					(4.110)
Industrial Firms per 100k					1.434
					(3.767)
Hospitals per 100k					3.154
					(13.188)
ATE on Treated	5.244	26.924***	6.200**	18.869***	5.761
	(4.771)	(9.625)	(2.468)	(6.112)	(43.121)
Year Coverage	1883, 1891-93	1883, 1891-93	1883, 1891-93	1883, 1891-93	1883
Disease Groups	12	12	12	12	12
Prefectures	34	34	34	34	34
Observations	816	816	816	816	408
F-statistic	14.80***	13.94***	15.32***	14.32***	91.54***
Within R-squared	0.141	0.177	0.153	0.186	0.910

Notes: All specifications include fixed effects for disease group, prefecture, and period. Robust standard errors are clustered by disease group and prefecture. The average treatment effect of rail access is calculated for rail access or average station density and their interactions over the treatment period for treated prefectures. Okinawa prefecture is missing land value data and Kagawa prefecture is missing pre-treatment data, so both are omitted from specifications.

Significance levels: *10%, **5%, ***1.