Immigration and Infant Mortality in Massachusetts: Evidence from the Age of Mass Migration

Katherine Eriksson — *UC Davis and NBER*

Peter Z. Lin — *UC Davis*

Gregory Niemesh — *Miami University and NBER*

August 4, 2020

Abstract

This paper estimates the effect of immigration on infant mortality rate during the Age of Mass Migration. Specifically, we use a shift-share instrument and town-level panel data from Massachusetts between 1860 and 1915 to estimate the impact over a long period. We find a significant positive effect of immigrant inflows on native infant mortality before 1900, with this effect diminishing after 1900. We also find suggestive evidence that this effect is due to communicable diseases and over-crowding. And the public health investment helped mitigate the negative effect.

Keywords: Immigration; Infant Mortality; Age of Mass Migration.

JEL Classification: Q12; C22; D81.

1 Introduction

The Age of Mass Migration, 1850 to 1915, began with the Irish Potato Famine in 1845 and German Revolutions of 1848. About 30 million European immigrants came to the United States in the subsequent decade, many settling in the Northeast (Abramitzky, Boustan and Eriksson, 2014). Towards the end of the period, public health experts and policy makers worried that the tendency of immigrants to settle in overpopulated neighborhoods within cities meant a negative impact on residents' health in cities (Kraut, 1988).

In this paper, we estimate the effect of immigration on infant and child mortality in Massachusetts throughout this period from 1860 to 1915.

We focus on Massachusetts for two main reasons. First, during the early part of the Age of Mass Migration, Massachusetts was one of the places absorbing largest amount of immigrants. absorbed XX% of immigrants to the United States. The political backlash in Massachusetts against low-income Irish immigrants in the 1850's was one of the largest in the nation, partly based on the argument that the immigrants were negatively impacting native health (Alsan, Eriksson and Niemesh, 2018). Second, more practically, Massachusetts is the first state to institute vital registration (starting in 1842) and keeps high-quality death registration data after 1860. (Preston and Haines, 1991) We construct town-level infant mortality rates for each five year period using individual-level birth and death certificates, and to this match immigrant population shares by town taken from the Federal Census of Population or Massachusetts State Censuses. To estimate the causality, we adopt an shift-share instrumental variable strategy based on ethnic settlement patterns in 1855, 1880, and 1900 Card (2001); Altonji and Card (1991).

We find that immigration positively affected infant mortality before 1900, with one percentage point increase in the population share of immigrants raising the infant mortality rate by 0.22 percentage points. However, after 1900 this effect decreased to 0.16 percentage points. We posit that this diminished effect is due to public health investments such as clean water and sewer (Alsan and Goldin, 2019; Cutler and Miller, 2005). The improved city sanitation mitigated the negative impact of large population inflow on infant and children's health. To verify this posit, we show that higher immigrant share was associated with higher population density and that population density positively predicts infant mortality rates, but only before 1900. Therefore we conclude that the effect is likely driven by crowding in an era without clean water or sewers. We also argue that the effect is not only driven by composition whereby children born to immigrants have higher infant mortality rates, but that there are spillovers to the native population. Our results are robust to a range of checks.

Our finding contributes a specific historical literature on mortality transition happening in the late 19th and early 20th century, as well as the role of immigration in this transition. One recent study addressing a similar question to ours is by Ager et al. (2020). They leverage the quota act of 1920s and find cities with more missing immigrants because of the immigration restriction experienced sharper decline in mortality rates. They argue this effect is due to less crowded living conditions in populated cities. Our study differs with their work in at least several ways. First, our focus is primarily on a long period from mid-19th century to early 20th century, when city sanitation was still in a poor condition. Second, our primary focus is put on infants and early-age children, who were much more vulnerable to adverse city health conditions and less directly affected via labor market channel.

Our paper also adds knowledge to a broader literature evaluating impacts of immigration on native-born population. In addition to rich studies focusing on labor market outcomes Card (2001); Borjas (1995); Ottaviano and Peri (2012); Glitz (2012),a fast growing literature explores how immigration will affect the native's health in contemporary context . Existing studies find positive or at least no negative effects of immigration on native's health outcomes (Giuntella and Mazzonna, 2015; Giuntella et al., 2019; Gunadi, 2020). Studies also find that positive effects mostly concentrate on low-skilled native-born workers, who can possibly shift to less physically burdened jobs because immigrants take over this jobs. While the effects on either infants or early-age children can be very different, very limited research has been done to address this. Another special contribution of historical studies can make is that back to periods of early industrialization, we don't need to worry about the potential crowding out effect of health care resources.

The rest of this paper is organized as follows. In Section 2, we describe the empirical strategy and data construction. Section 3 presents main results, discusses the robustness of our results, and heterogeneous effects. Section 4 discusses suggestive evidence to channels, and Section 5 concludes.

2 Methodology and Data

2.1 Empirical Strategy

We estimate the effect of immigration on infant mortality rates in Massachusetts towns between 1860 and 1915 with the following model:

$$IMR_{it} = \alpha + \beta_0 imgrt_{it} + \beta_1 imgrt_{it} \times Post1900_t + \eta_i + \sigma_t + \epsilon_{it} \quad (1)$$

where dependent variable, IMR_{it}, is infant mortality rate in town i and census year t. The explanatory variable, $imgrt_{it}$, is immigration rate and measured by the percentage of population that were foreign-born. $Post1900_t$ is a dummy variable indicating for periods after the census year 1900 (including the year 1900). Inclusion of this interaction term allows immigration has a different effect on the outcome variable before and after the eve of twentieth century. We also include town fixed effects (η_i) and census year fixed effect (σ_t) to control for all time-invariant characteristics at towns and a statewide trend. ϵ_{it} is an error term.

The coefficients of our primary interest are β_0 and β_1. The former gives a marginal effect of immigration on infant mortality rate before 1900, and the later gives an estimated change of this effect in the post-1900 periods, when the marginal effect becomes $\beta_0 + \beta_1$. However, a major concern about this regression model is that an immigration rate could be highly endogenous even conditional on town and census year fixed effects. That is, the geographic distribution of immigrants was correlated with some other time-varying local conditions that could influence the outcome. If this was the case, the OLS estimates of β_0 and β_1 will be biased.

Therefore, to address this concern, we take an instrumental variable strategy that leverages a shift-share instrument (or past settlement instrument). A shift-share instrument is commonly used in literature to evaluate the effects of immigration shocks on labor market or other broader outcomes Altonji and Card (1991); Card (2001). The instrument is based on the fact that newly arriving immigrants tend to locate near previous immigrants coming from the same origin. Therefore, the lagged geographic distribution of early settlers can well predict the distribution of followers in the same origin group. In this spirit, we construct an instrumental variable for the actual immigration by interacting a lagged distribution with national changes (in this paper, statewide changes) of immigration in each group. The instrument is valid if the aggregate immigration flows were orthogonal to local conditions and the lagged distribution is fully controlled by town fixed effects.

In practice, we first construct the following shift-share instrument for $imgrt_{it}$:

$$imgrtIV_{it} = \frac{F\hat{B}_{it}}{F\hat{B}_{it} + N\hat{B}_{it}} \tag{2}$$

where $F\hat{B}_{it}$ and $N\hat{B}_{it}$ is predicted population that were foreign born and native born, respectively. And they were attained in the following way:

$$\begin{aligned} F\hat{B}_{it} &= \sum_{o} (\frac{F_{io,1855}}{F_{o,1855}}) \times F_{ot} \\ N\hat{B}_{it} &= \frac{N_{i,1855}}{N_{1855}} \times N_t \end{aligned} \tag{3}$$

where $F_{io,1855}/F_{o,1855}$ is the share of all immigrants from country 0 that resided in town i by the year of 1855. F_{ot} is the number of all immigrants from country o in Massachusetts by year t . The native-born population is also predicted in the same way, and $N_{i,1855}/N_{1855}$ is the share of all native-born population in town i by 1855.

We choose 1855 as a base year for lagged distribution because the age of mass migration from European countries started just after 1850. We cannot attain informative distribution to predict following migration until a sufficient number of immigrants arrived and established an initial pattern. In addition, detailed population data by nativity and origin countries at towns was mostly available after 1850. Regarding the origin countries, we categorize immigrants into eleven major groups, which are Ireland, Canada, England, Scotland, Germany, France, Italy, Portugal, Denmark, Norway and Sweden, and all other countries.[1] Immigrants from these ten specified regions accounted for 99.5% of all foreign-born population in Massachusetts by 1855 and still accounted for 77.5% of all immigrants by 1915.

Figure 4 shows the origin composition of immigrants in Massachusetts from 1855 to 1915. By 1855, the largest group of immigrants came from Ireland and they accounted for more than 70% of all foreign-born population in the state. The second and third largest group were from England

[1]The group of Canada includes the British Canada, French Canada, New Brunswick, Newfoundland, Nova Scotia, and Prince Edward Island; the group of England include Wales; and the group of Germany includes Austria, Switzerland, Belgium, and Netherlands.

and Canada. The three dominant groups accounted for 91% of all immigrants. The origin composition remained relatively stable until the mid-1880s, when an increasing number of immigrants arrived from Norway, Sweden, Italy, and Portugal. Afterwards, the share of immigrants from new countries keep increasing. It is also notable that the share of immigrants from all other countries (including some Eastern European countries such as Russia and Poland) quickly increased in the decade of 1900. By 1915, the share of immigrants from Ireland decreased below 25% and the share of immigrants from all other countries reached about 20%.

This significant composition change raises a technical issue of the above defined shift-share instrument. With immigration from new countries rising over time, the instrument merely based on 1855 distribution provides no or very limited prediction for geographic distribution of immigrants from new origins. Therefore, the instrument gets weaker as the lagged distribution becomes less predictive over time. In figure 5, we plot the correlation between an actual immigration rate and the shift-share instrument. Sub-figure (B) reveals that conditional on town and year fixed effects, the instrument only weakly predicts actual immigration rates ($coef$=0.08, t -value=1.51), though the unconditional correlation is a bit larger ($coef$=0.67, t-value=25.72, see sub-figure A).

A straight solution to this problem is updating the lagged distribution just before the origin composition was about to significantly change. The updated reference distribution will be more predictive than 1855 distribution as it is closer to following migration flows in the sense of timing. With this idea, we modify the instrument by updating the reference distribution in 1880 and 1900, respectively. Then the instrument with updated distribution has a formula as below:

$$imgrtIV'_{it} = \frac{FB'_{it}}{NB'_{it}} = \frac{\sum_o \theta_{io,t} \times F_{ot}}{\phi_{i,t} \times N_t}$$
$$\theta_{io,t} = \frac{F_{io,1855}}{F_{o,1855}} \times 1(t \leq 1880) + \frac{F_{io,1880}}{F_{o,1880}} \times 1(1885 \leq t \leq 1900) + \frac{F_{io,1900}}{F_{o,1900}} \times 1(t \geq 1905) \quad (4)$$
$$\phi_{i,t} = \frac{N_{i,1855}}{N_{1855}} \times 1(t \leq 1880) + \frac{N_{i,1880}}{N_{1880}} \times 1(1885 \leq t \leq 1900) + \frac{N_{i,1900}}{N_{1900}} \times 1(t \geq 1905)$$

where, $\theta_{io,t}$ and $\phi_{i,t}$ are the updated shift-share of population by birth countries. The function of $1(.)$ is an indicator for specified periods. The updated instrument now predicts the migration flow between 1885 and 1900 with the 1880 distribution. The migration flows between 1905 and 1915 is then predicted by the lagged distribution in 1900. The sub-figure (C) and (D) in figure 5 plots the correlation between actual immigration rate and the modified instrument. Conditional on town and year fixed effects, the updated instrument better predicts an immigration rate than before ($coef$=0.42, t-value=12.35). With this updated shift-share instrument, we estimate the equation (1) with 2SLS and we instrument the interaction term with $imgrtIV'_{it} \times Post1900_t$.

2.2 Data

We calculate infant mortality rates in Massachusetts towns between 1860 and 1915 based on registered death and birth certificates, which are provided by FamilySearch.org. Death and birth certificates typically report age, gender, event date and event place. With this information, we calculate the infant mortality rate in town i and census year t as below:

$$IMR_{it} = \frac{\sum_{j=t-2}^{j=t+2} Deaths_{ij}^{1}}{\sum_{j=t-2}^{j=t+2} Births_{ij}} \tag{5}$$

where $Deaths_{ij}^{1}$ is the total number of deaths under age 1 in town i and year j; and $Births_{ij}$ is the number of total births in the town and year. It is noteworthy that in regression, we only measure a mortality rate associated with a census year, which is actually a five-year moving average centering the census year. We take the five-year average of infant mortality rates instead of the original annual rates mostly because the annual data can bear greater measurement errors when deaths or births are misreported. This is an significant issue particularly for small places. Furthermore, we also exclude annual observations with total infant deaths exceeding total births counts from a moving average to make sure our calculation are not affected by obvious errors.[2]

To examine a broader effect of immigration on early-age children, we also calculate a mortality rate of children aged between 1 and 4 years. In a similar way, we first count the total deaths aged 1-4 in a specified towns and year. Then, we measure the mortality rate as the total deaths divided by an estimated population aged 1-4. We follow the same approach by Alsan and Goldin (2019) and calculate the population aged 1-4 as the difference between cumulative births and cumulative deaths for corresponding cohorts in previous years. Details on this calculation is provided in appendix.

[2] Based on our calculation, only 2.76% of all annual observations (509 out of 18,444) show obvious measurement errors.

An immigration rate in this paper is measured by the percentage of population that were foreign-born. Population data by nativity and birth countries comes from the Federal population censuses every decade since 1860. We add data from Massachusetts state population censuses every ten years since 1855. Finally, we compile a panel data of infant mortality rates and immigration rates in 315 Massachusetts towns (or cities) every five years between 1860 and 1915 (a total of twelve waves). [3] It is also noted that we adjust our data to reflect changes in town boundaries in the study period. The 315 towns are based on a set of consistent boundaries over time.

Figure 1 shows the statewide trends in our calculated infant mortality rates (left axis) and the percentage of immigrant in the population (right axis) from 1860 to 1915. Infant mortality reached a peak higher than 18% around 1875, levelling off from 1880 to 1895, and dramatically fell thereafter. Meanwhile, the population share of immigrants grew steadily from 20 percent in 1860 to over 30 percent after 1900. This overall trend hint a structural change in the correlation between immigration and infant mortality on the eve of 20th century.

Figure 2 maps the cross-town variation in calculated infant mortality in the census year 1860, 1880, 1900, and 1915. We find that before 1900, the counties located closer to Boston metropolitan area had higher infant mortality than the interior and rural counties. While comparing the map of 1900 with that of 1915, we find a larger decline of infant mortality in Boston metropolitan area. This pattern is possibly due to provision of clean water and sewerage services just introduced in the 1900s (Alsan and Goldin, 2019). Figure 3 maps the geographic variation in immigration rate across MA towns. It is easy to find that for most neighborhoods, immigration rate steadily increased and Boston area was continuously the place with highest immigration rate.

Table 1 reports summary statistics of the main variables. For the full sample including all towns and census periods, infant mortality rate has a mean of 0.123 (or 12.3%), with a standard deviation of 0.06. We also calculate the infant mortality rates by gender and seasons. Mortality rates of male infants were a bit higher than that of the female (0.133 vs. 0.11), but the difference is not statistically significant. Also, mortality rates in warmer months were a bit higher than that in colder months (0.07 vs. 0.05). The mortality rate of children aged 1-4 has a mean of 0.018 (or 1.8%), which was significantly lower than the infant mortality rate in our sample. Lastly, the immigration

[3] Due to the incomplete data in 1890 census, we only have 134 towns in that census year. Population by nativity are only available for towns with more than 2,500 residents in 1890 census.

rate has a mean of 0.174 (or 17.4%), with a standard deviation of 0.101.

3 Results

3.1 Main results

Table 2 shows main results of the estimated effect of immigration on infant mortality rate in Massachusetts between 1860 and 1915. Column (4) shows the 2SLS estimates of our preferred specification. The estimates show that immigration had a positive and significant effect on infant mortality rate in the periods before 1900. The point estimate suggests that 1 percentage point increase in immigration rate would raise an infant mortality rate by about 0.22 percentage point. This marginal effect also suggests that one standard deviation increase in immigration rate would increase an infant mortality rate by about 0.4 standard deviation in our sample.

However, we also find a negative and significant coefficient on the interaction term between immigration rate and post-1900 indicator, and this suggests the positive effect significantly shrank since the early twentieth century. We also calculate the linear combination $\beta_0 + \beta_1$, which gives the marginal marginal effect post 1900. Column (3) reports an estimated overall effect for the entire period. The OLS estimates are also reported in column (1) and (2) for comparison. Though the point estimates by 2SLS are a bit larger, endogeneity tests suggest they don't significantly differ. In short, our baseline results reveal that though immigration could increase infant mortality rate in Massachusetts between 1860 and 1915, this affect got significantly smaller after 1900.

We also check the robustness of our finding and table 3 reports these results. First, we weight the baseline regression by town population to allow for a greater variance of error term in a more populated place. Column (1) shows similar results with the unweighted regression. Second, we add county by census year fixed effects in regression and we find no significant change in our results (col. 2). Third, it might be a concern that our finding is only driven by some very large cities, such as Boston. Therefore, we exclude the city of Boston from the sample. Column (3) reveals the result and the estimates don't change a lot. Then, we further exclude all observations with more than 10,000 population from regression and again the results remain mostly unchanged (col. 4).

It might also be a concern that measurement errors of infant mortality rates will bias our estimates. It is more likely to observe inaccurate infant mortality rates at small places, where a small error in death or birth counts can cause a large error in the mortality rate. Therefore, we drop all observations with fewer than 50 births from regression and check if our results significantly change. Column (5) reports no significant change in our results. Lastly, we also take the logarithm of infant mortality rate as an outcome variable and we find similar qualitative results.

We also examine a broader effect of immigration on the mortality rate of children aged 1 to 5 in the same period. Table 4 reports the results. We find a positive and significant effect of immigration on the children's mortality rate at early ages before 1900. The point estimate in column (2), the preferred specification, suggests that one percentage point increase in an immigration rate would increase the children mortality rate by about 0.04 percentage point. Meanwhile, the negative and statistically significant coefficients on the interaction term suggest that the positive effect greatly reduced and fell into insignificance after 1900. In short, we find a much smaller effect of immigration on the mortality rate of children aged 1 to 5 than infants.

3.2 Heterogeneous Effects

Given that we find a positive effect of immigration on infant mortality and it significantly shrank after 1900, we further explore if this effect vary for different demographic groups or in different seasons. First, we estimate the effect of immigration on infant mortality rates by gender. Column (1) and (2) in Table 5 shows the results. Before the 1900, we find positive and significant effect for both gender. The positive effect for both gender all decreased after 1900, though the decline is greater and more significant for male.

Then, we check if the effect of immigration differ in different seasons. Some studies find evidence that infants and early-age children are more vulnerable to gastrointestinal diseases that had more prevalence in warmer months. Column (3) and (4) report the estimated effect of immigration on infant mortality in warmer months (April to September) and colder months (October to March). Results show that immigration had a greater positive effect in warmer months before the 1900s. One percentage point increase in immigration rate would increase the infant mortality in warmer months by about 0.16 percentage point, while only increase the effect on fall-winter mortality by 0.06 percentage point. In addition, the positive effect on Spring-Summer mortality significantly shrank after 1900, while the effect on fall-winter mortality remained quite stable after 1900.

3.3 Suggestive Channel

Why were infants and early-age children affected by immigration, and what explains the diminishing effect after 1900. One potential explanation is that the arrival of immigrants increased population and more crowed places had faster spread of some transmitted diseases. The recent study by Alsan and Goldin (2019) finds the introduction of safe water and sewerage services in Boston metropolitan area between 1895 and 1900 significantly reduced the children mortality caused by gastrointestinal and respiratory diseases. Such interventions and improved environment can mitigate the negative effect of increasing population on survival of infants.

We find supportive evidence that immigration did increase overall population in the study period. Column (1) in table 6 show that for the entire period, one percentage point increase in immigration rate increase the overall population by about 2.45%. Results in column (2) suggest that the crowding effect was even greater after 1900 than before.

However, the larger crowding effect of immigration did not cause greater effect on infant mortality. In fact, we find smaller effect after 1900. Results in column (3) and (4) might provide some clues. As column (4) reveals, though higher population could increase infant mortality before the 1900, this positive correlation fell into zero after 1900, possibly due to improved sanitation conditions in cities.

4 Conclusion Remarks

In this paper, we explore the impact of immigration on infant and early-age child mortality in Massachusetts between 1860 and 1915, a period in the Age of Mass Migration. With a panel data of calculated infant mortality rate at town level and a shift-share instrument strategy, we find a positive effect of immigration rate on infant mortality. More specifically, we find that one percentage point increase in immigration rate would increase an infant mortality rate by about 0.22 percentage point before 1900. This effect significantly decreases to about 0.16 percentage point after 1900, when city sanitation conditions began to be improved. We also find suggestive evidence for our posit that the adverse health effect was due to communicable diseases and overpopulated living environment.

Our findings help better understand the health impact of immigration in a historical context, where industrialization accelerated, city sanitation was not well developed, and public health care was not yet available. Though we find a positive effect of immigration on infant mortality, our findings don't support the nativist argument that immigrants brought in disease. Instead, the adverse effect was mainly due to the poor city living conditions associated with population growth. Fortunately, we see the adverse effect was significantly mitigated after 1900, when public health infrastructures were introduced.

References

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson (2014) "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration," *Journal of Political Economy*, Vol. 122, No. 3, pp. 467–506.

Ager, Philipp, James J Feigenbaum, Casper Worm Hansen, and Hui Ren Tan (2020) "How the Other Half Died: Immigration and Mortality in US Cities," Working Paper 27480, National Bureau of Economic Research.

Alsan, Marcella, Katherine Eriksson, and Gregory Niemesh (2018) "The Rise and Fall of the Know-Nothing Party," working paper.

Alsan, Marcella and Claudia Goldin (2019) "Watersheds in Child Mortality: The Role of Effective Water and Sewerage Infrastructure, 1880–1920," *Journal of Political Economy*, Vol. 127, No. 2, pp. 586–638.

Altonji, Joseph G. and David Card (1991) "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives," in *Immigration, Trade, and the Labor Market*: National Bureau of Economic Research, Inc, pp. 201–234.

Borjas, George J. (1995) "The Economic Benefits from Immigration," *Journal of Economic Perspectives*, Vol. 9, No. 2, pp. 3–22.

Card, David (2001) "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, Vol. 19, No. 1, pp. 22–64.

Cutler, David and Grant Miller (2005) "The Role of Public Health Improvements in Health Advances: The Twentieth-Century United States," *Demography*, Vol. 42, No. 1, pp. 1–22.

Giuntella, O., F. Mazzonna, C. Nicodemo, and C. Vargas-Silva (2019) "Immigration and the reallocation of work health risks," *Journal of Population Economics*, Vol. 32, No. 3, pp. 1009–1042.

Giuntella, Osea and Fabrizio Mazzonna (2015) "Do immigrants improve the health of natives?" *Journal of Health Economics*, Vol. 43, pp. 140 – 153.

Glitz, Albrecht (2012) "The Labor Market Impact of Immigration: A Quasi-Experiment Exploiting Immigrant Location Rules in Germany," *Journal of Labor Economics*, Vol. 30, No. 1, pp. 175–213.

Gunadi, C. (2020) "Immigration and the Health of U.S. Natives," *Southern Economic Journal*, Vol. 86, No. 4, pp. 1278–1306.

Haines, Michael R. (1979) "The Use of Model Life Tables to Estimate Mortality for the United States in the Late Nineteenth Century," *Demography*, Vol. 16, No. 2, pp. 289–312.

——— (1998) "Estimated Life Tables for the United States, 1850-1910," *Historical Methods*, Vol. 31, No. 4, p. 149, Last updated - 2013-02-23.

Kraut, Alan M. (1988) "Silent Travelers: Germs, Genes, and American Efficiency, 1890-1924," *Social Science History*, Vol. 12, No. 4, pp. 377–394.

Ottaviano, Gianmarco I. P. and Giovanni Peri (2012) "Rethinking the Effect of Immigration on Wages," *Journal of the European Economic Association*, Vol. 10, No. 1, pp. 152–197.

Preston, Samuel H. and Michael R. Haines (1991) *Fatal Years: Child Mortality in Late Nineteenth-Century America*: Princeton University Press.

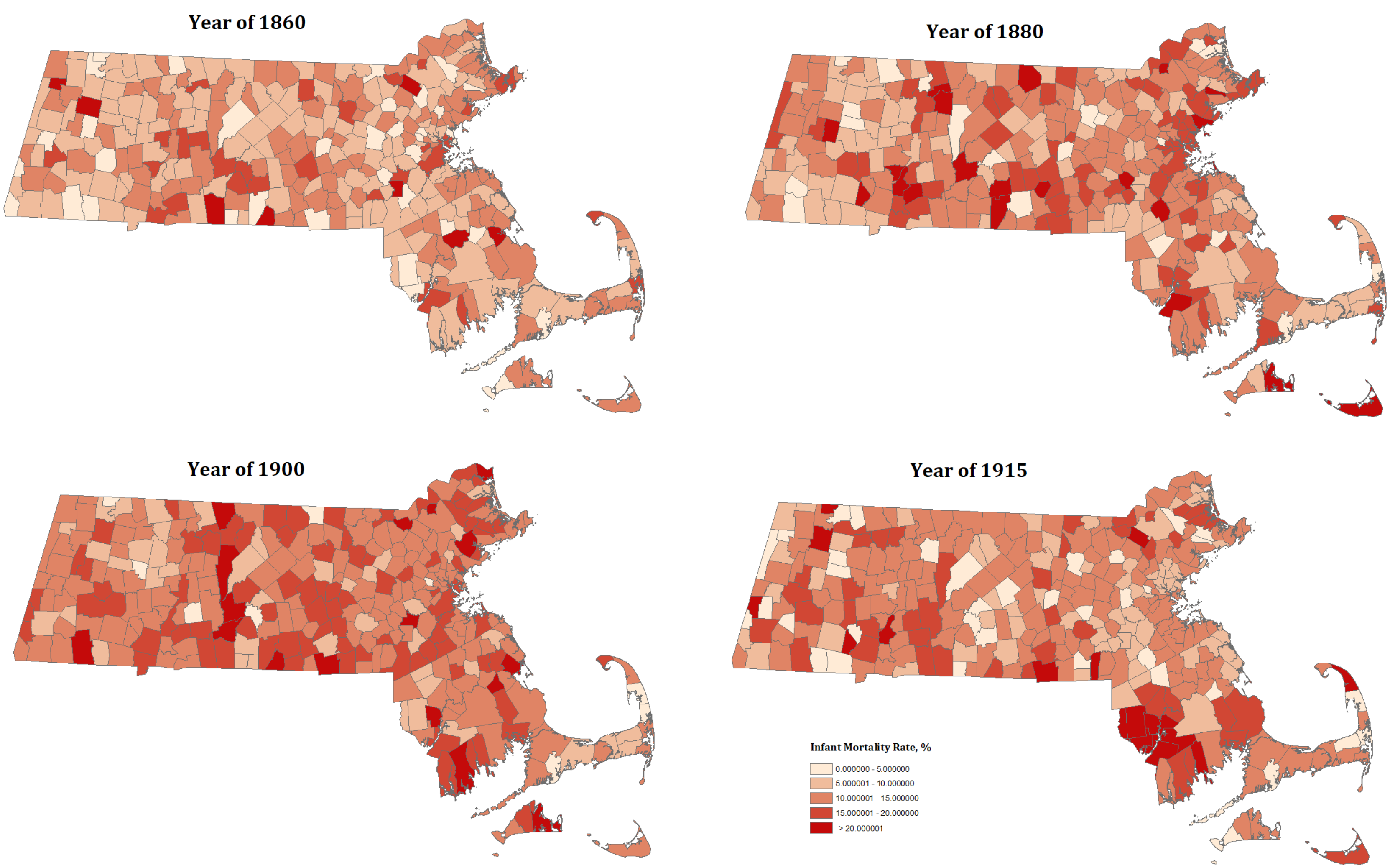
Wooldridge, Jeffrey M (1995) "Score diagnostics for linear models estimated by two stage least squares," *Advances in econometrics and quantitative economics: Essays in honor of Professor CR Rao*, pp. 66–87.

Figures and Tables

Figure 1: Infant Mortality and Immigrant Share in Massachusetts: 1860-1915

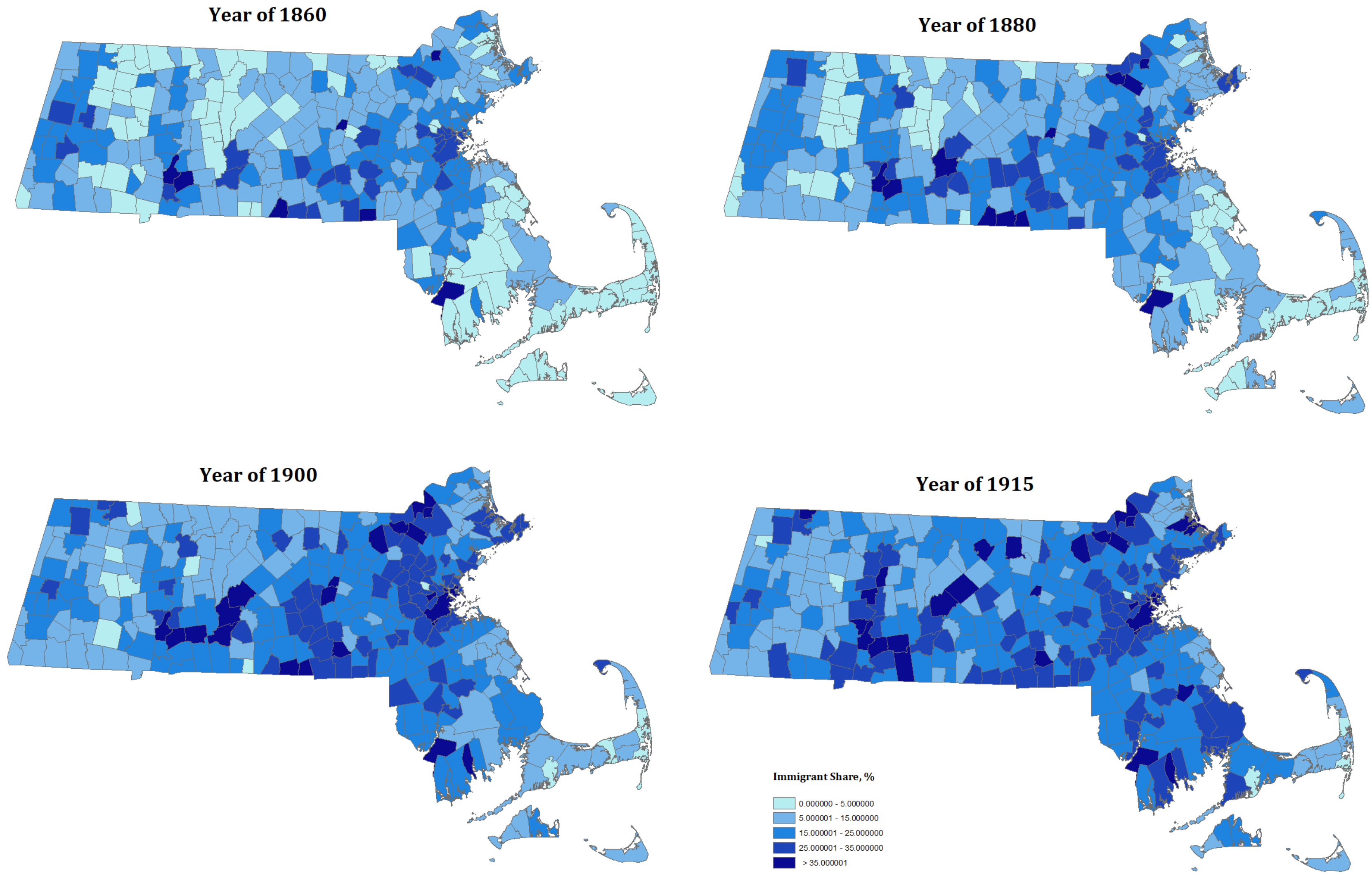
Note: The state-wide infant mortality is attained by dividing total deaths under age 1 by total births in each year between 1860 and 1915. The time series plotted is a 5-year moving average (2 years lag and 2 years lead) of the original annual series. Immigrant share is the percentage of foreign-born population and from the federal and state censuses in Massachusetts between 1860 and 1915.

Figure 2: Infant Mortality Rates in Massachusetts Towns



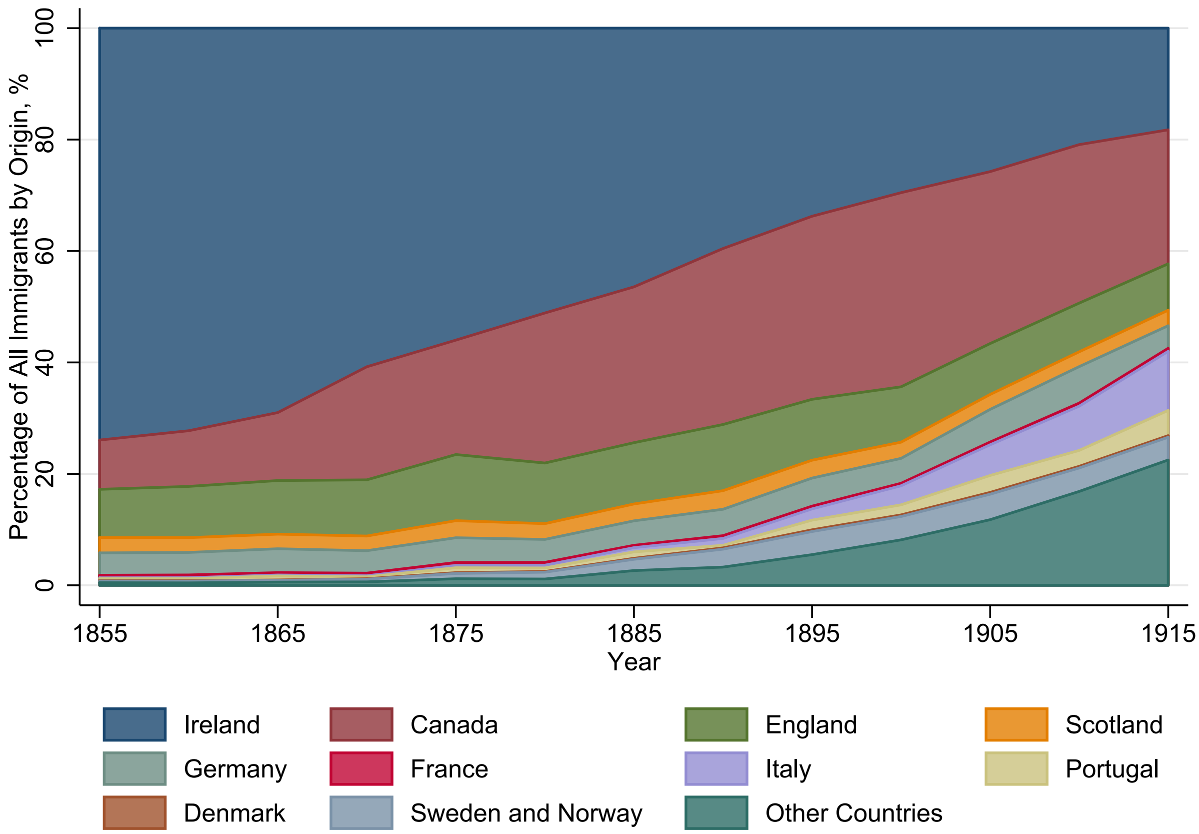
Note: For each specified census year t, we calculate the infant mortality rates as the total number of deaths under age 1 between year $t-2$ and $t+2$, divided by the total number of births in the same period. Data comes from the birth and death certificates registered in Massachusetts and is provided by FamilySearch.org. The boundaries have been adjusted to satisfy overtime comparison.

Figure 3: Immigration Rates in Massachusetts Towns



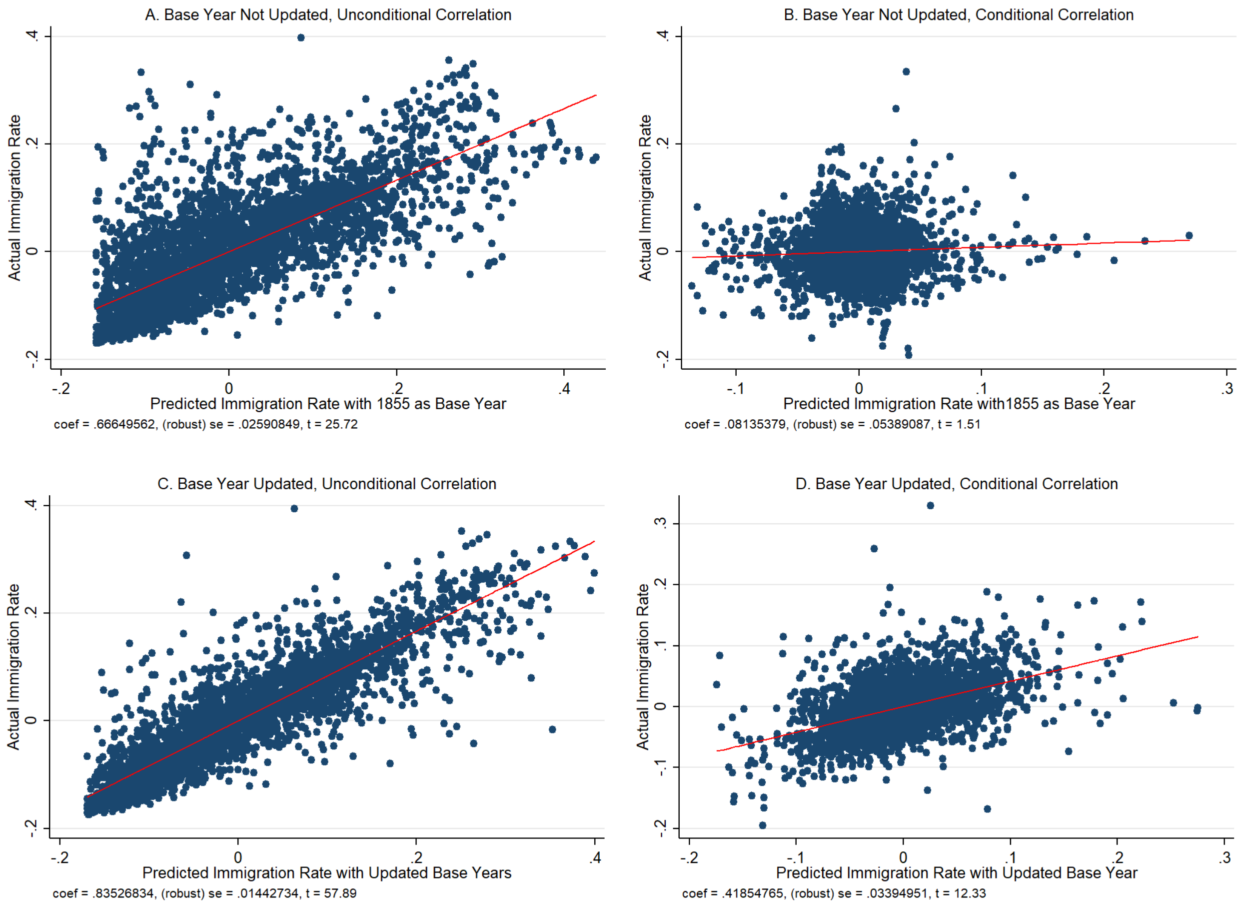
Note: Immigration rates are calculated as the percentage of population that were foreign-born in the specified census year. Population by nativity in the years 1860, 1880, and 1900 comes from the eighth, tenth, and twelfth population census. Data in the year 1915 comes from the population census of Massachusetts in 1915.

Figure 4: Origin Composition of Immigrants in Massachusetts: 1860-1915



Note: Data come from the federal and state censuses. The group of Canada includes the British Canada, French Canada, New Brunswick, Newfoundland, Nova Scotia, and Prince Edward Island. The category of England include England and Wales; the category Germany also include Austria, Belgium, Netherlands, and Switzerland.

Figure 5: Immigration Rate and Shift-share Instrument



Note: Figure (A) and (C) plot actual immigration rate against shift-share instrument without controlling for any co-variates. Figure (B) and (D) plot the two variables controlling for town fixed effects and census year fixed effects. We update the reference distribution in 1880 and 1900 for the instrument presented in Figure (C) and (D). The reported coefficients, standard errors, and t-value come from simple regressions. All standard errors are clustered at towns.

Table 1: Summary Statistics: Infant Mortality Rate and Immigration Share: 1855-1905

	Mean	Stand. Dev	Min	Max	Obs.
Infant Mortality Rate (IMR)	.123	.055	0	.930	3595
IMR, Male	.133	.068	0	.95	3595
IMR, Female	.110	.061	0	.903	3595
IMR, April - September	.070	.037	0	.549	3595
IMR, October - March	.053	.028	0	.463	3595
Mortality Rate, Age 1-4 (CMR)	.018	.019	0	.8	3595
Immigration Rate	.174	.101	0	.568	3595
Census Year	1887	17.67	1860	1915	3595

Note: The summary statistics presented in this table are unweighted across all towns and census years between 1860 and 1915.Gender-specific mortality rate is the ratio of deaths to births for male or female. Summer mortality rate is the ratio of deaths in April through September to year-round births; Winter mortality rate is the ratio of deaths in October through March to year-round births.

Table 2: Immigration and Infant Mortality Rates: 1860 - 1915

	IMR OLS (1)	IMR OLS (2)	IMR 2SLS (3)	IMR 2SLS (4)
Immigration Rate (β_0)	0.126***	0.149***	0.191***	0.219***
	(0.027)	(0.028)	(0.065)	(0.061)
Immigration Rate $\times$ Post 1900 (β_1)		-0.062***		-0.056**
		(0.021)		(0.026)
$\beta_0 + \beta_1$		0.087***		0.163**
		(0.030)		(0.071)
Town F.E.	Yes	Yes	Yes	Yes
Census Year F.E.	Yes	Yes	Yes	Yes
Kleibergen-Paap F-Statistic			166.06	78.14
F-stat of Endogeneity Test			1.51	1.11
N	3595	3595	3595	3595

Note: Dependent variables are infant mortality rates in all columns. Immigration rate is the percentage of population that were foreign-born. Regressions are conducted at the level of town by census years. The F-statistics of endogeneity test has a null hypothesis that the explanatory variables are exogenous and is based on the Wooldridge (1995) robust score test. We cannot reject the null hypothesis in our specifications. Standard errors are clustered at towns and reported in parentheses.

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

Table 3: Robustness: Immigration and Infant Mortality Rates: 1860 - 1915

	Weighted by population	Control for county-year F.E.	Drop Boston	Drop $Population \geq 10,000$	Drop $Births \leq 50$	$log(IMR)$
	(1)	(2)	(3)	(4)	(5)	(6)
Immigration Rate	0.199*	0.209***	0.214***	0.211***	0.200***	1.917***
	(0.115)	(0.065)	(0.061)	(0.061)	(0.054)	(0.468)
Immigration Rate× Post 1900	-0.145***	-0.053*	-0.052**	-0.045*	-0.076**	-0.401*
	(0.047)	(0.031)	(0.026)	(0.026)	(0.031)	(0.220)
Town F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Census Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap F-Statistic	26.35	70.42	78.14	78.04	84.11	84.83
N	3595	3583	3583	3568	3162	3543

Note: Coefficients are all estimated by 2SLS with an updated shift-share instrument. Dependent variable is the infant mortality rate for all columns, except for (6). Standard errors are clustered at towns and reported in the parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Immigration and Children Mortality Aged 1 to 5 in Massachusetts: 1860-1915

	CMR All Sample (1)	CMR All Sample (2)	CMR All Sample (3)	CMR Population Weighted (4)	CMR Drop Boston (5)
Immigration Rate	0.016	0.039***	0.040***	0.032*	0.039***
	(0.015)	(0.013)	(0.015)	(0.017)	(0.013)
Immigration Rate × Post 1900		-0.046***	-0.050***	-0.055***	-0.046***
		(0.011)	(0.012)	(0.008)	(0.011)
Town F.E.	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes
County Year F.E.	No	No	Yes	No	No
Kleibergen-Paap F-Statistic	166.06	78.14	70.42	26.35	78.14
N	3595	3595	3583	3595	3583

Note: All specifications are estimated with 2SLS and dependent variables are mortality of Children aged over 1 and under 5. Standard errors are reported in the parentheses and clustered at towns.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Heterogenous Effects of Immigration on Infant Mortality

	IMR, Male (1)	IMR, Female (2)	IMR, Apr. - Sept. (3)	IMR, Oct. - Mar. (4)
Immigration Rate	0.187***	0.260***	0.160***	0.058**
	(0.062)	(0.084)	(0.039)	(0.029)
Immigration Rate × Post 1900	-0.085***	-0.042	-0.065***	0.009
	(0.031)	(0.029)	(0.017)	(0.013)
Town F.E.	Yes	Yes	Yes	Yes
Census Year F.E.	Yes	Yes	Yes	Yes
Kleibergen-Paap F-Statistic	78.14	78.14	78.14	78.14
N	3595	3595	3595	3595

Note: All specifications are estimated by 2SLS. Infant mortality rates by gender are based on gender-specific death and birth counts. The mortality rate for specified months is measured by the death count in specified months divided by the year-round birth count. Standard errors are clustered at towns and reported in the parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Infant Mortality, Immigration, and Population: 1860-1915

	Log Total Population (1)	Log Total Population (2)	IMR (3)	IMR (4)
Immigration Rate	2.450***	0.688		
	(0.516)	(0.444)		
Immigration Rate × Post 1900		3.529***		
		(0.192)		
Log Total Population			-0.001	0.014***
			(0.003)	(0.004)
Log Total Population × Post 1900				-0.010***
				(0.002)
Town F.E.	Yes	Yes	Yes	Yes
Census Year F.E.	Yes	Yes	Yes	Yes
KP F-Statistic	166.06	78.14		
N	3595	3595	3595	3595

Dependent variables in column (1) and (2) are log of total town population, and the dependent variables in column (3) and (4) are infant mortality rates. Coefficients are estimated with 2SLS in column (1) and (2), while OLS estimates are given in column (3) and (4). Standard errors in all regressions are clustered at towns and reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix

Mortality Rate of Children Aged 1-4

We calculate a mortality rate of children aged between 1 and 4 years as below:

$$CMR_{it} = \frac{\sum_{j=t-2}^{t+2} Deaths_{i,j}^{1-4}}{\sum_{j=t-2}^{t+2} POP_{i,j}^{1-4}}$$

where CMR_{it}^{1-4} is the children mortality rate aged 1-4, $Deaths_{i,j}^{1-4}$ is total deaths aged 1-4 years in town i and year j, and $POP_{i,j}^{1-4}$ is the population aged 1-4 years in town i and year j. We adopt the same approach with Alsan and Goldin (2019) to estimate the population aged between 1 and 4 years by taking difference between the cumulative births and deaths for specified cohorts in previous years. Specifically,

$$POP_{i,j}^{1-4} = \sum_{a=1}^{a=4}[Births_{i,t-a} - \sum_{k=1}^{k=a} Deaths_{i,t-k}^{a-k}]$$

where $Births_{i,t-a}$ is the births in year $(t-a)$, and $Deaths_{i,t-k}^{a-k}$ is the deaths aged $(a-k)$ years in town i and year $(t-k)$.