Immigration and native health: Evidence from 1920s US immigration policy*

Patrick Szurkowski

University of Pittsburgh

August 13, 2024

Abstract

Multiple channels have been identified through which immigrant inflows produce spillover health effects on individuals in receiving locations. However, the net health impacts of immigration flows have yet to be identified. This project examines how an exogenous decline in immigration flows impacts the lifespans of native-born individuals. A set of 1920s immigration laws in the United States imposed quotas based on national origin. The effects of this policy change are identified in a difference-in-difference framework. Results indicate that average policy exposure was associated with significant declines in lifespan driven by job transition and relocation towards more deleterious health conditions.

JEL CODES:

Keywords:

*I am grateful to

1 Introduction

Health and immigration economists have documented many ways by which immigration inflows impact the health of residents. These effects can be roughly grouped into two categories; those affecting labor market conditions and those affecting local public health[1]. In labor markets, immigrants have been observed generating positive health spillovers by sorting into more hazardous positions in the place of native residents. On the other hand, negative public health effects have been observed as inflows increase congestion and strain local resources. The relative importance of each of these channels is unknown and residents have the option of relocating to avoid the effects of a shock entirely. These dynamics imply that the reduced-form health impacts of immigration might be positive, negative, or insignificant.

In the 1920s, the US implemented the first national-level policies aimed at completely overhauling immigration[2]. A series of two laws, the Emergency Quota Act (1921) and Immigration Act of 1924, implemented quotas which cut yearly inflows to around one-eighth of previous levels. This decline fundamentally changed the economy for residents, leading to transitions from rural agriculture to urban manufacturing. During this period, there was a mortality penalty associated with urban residence and the lack of workplace safety standards led to significant health risk in industry. It's reasonable to expect that these immigration-driven employment changes produced significant health costs but no estimates exist currently.

This paper examines changes in the lifespans of working age Americans after these immigration policies severely restricted immigration inflows to the United States. The historical setting here is ideal for examining the effects of an immigration shock on lifespan as an outcome because the lifespan of nearly all individuals in the sample population is observed[3]. This project draws on census and vital records from Ohio and North Carolina digitized as part of the LIFE-M project to measure long-term individual outcomes. This wide scope captures in health effects driven by changes by labor market and local conditions temporally distant from the timing of the shock. Within the health and immigration literature, these results represent the first treatment effect estimates cast in terms of a change in lifespan. The incidence of treatment effects across the population is identified and conditioning on

[1]This delineation is common in spatial equilibrium models of immigration examining changes in welfare subsequent to policy changes (Jia et al., 2023).

[2]Prior, state-level policies had been enacted sporadically and the Chinese Exclusion Act targeted one group in particular.

[3]In a setting with a more modern immigration shock, one would expect some portion of the sample population to not yet be deceased. This would generate sample selection bias along characteristics associated with shorter life expectancy. More troublesome, if the impact of these characteristics was not constant over time, it might be impossible to achieve unbiased estimation of the treatment effect.

relocation and occupation change status is used to identify the importance of changing local and labor market health conditions.

Results suggest that the average exposure to policy change across counties was associated with a roughly one year decline in lifespan among working age native-born individuals. The baseline analysis is extended to examine the incidence across subgroups within this population finding significant treatment effect heterogeneity. The effect is concentrated among younger cohorts ($<$ 35 years old) with this group experiencing a decline of over two years of life expectancy at average treatment levels compared with no significant effect among older cohorts. Gender and urban status played an important role in the magnitude of the treatment effect. Exposure is associated with significant declines for both genders and among urban and rural residents but effect sizes are larger among women and individuals living in urban areas.

To identify the importance of changes in labor market and local conditions, I test for heterogeneity among the subgroups conditioning on relocation and occupation change statuses. Conditioning on occupation change in the following decade, strong negative declines in lifespan are associated with exposure in the occupation switching subset while marginally negative declines are observed among non-switchers. These results suggest that labor market sorting drove the significant health declines while changes in local conditions were marginally significant or net zero. Both movers and stayers see negative treatment effects associated with exposure but the effect is marginally larger among movers (potentially picking up coterminous flows from rural to urban areas and from agriculture to manufacturing). These results suggest that the aggregate effect of changes induced in labor markets and local conditions led to health declines.

As a supplemental analysis, two occupation and industry-specific outcomes are used to identify changing job conditions. Occupation-specific average lifespans are calculated with the LIFE-M dataset and changes in the average lifespan of occupation of employment are studied in a national sample of working age men. These results suggest immigration shock exposure was associated with transitions into lower average lifespan occupations. Injury rate data collected by the Bureau of Labor Standards from 1940 allows a similar analysis to be conducted looking at cross-industry transitions with accident rates as a measure of workplace safety. These results indicate that greater exposure was associated with transitions to industries with higher rates of disabling injury.

To identify the impact of changes in immigration flows on the health of native residents, I use variation in areas' exposure to a drastic change in immigration policy. This paper utilizes on the empirical strategies of Abramitzky et al. (2019) and Ager et al. (2024) at the county-level. The Emergency Quota Act and Immigration Act of 1924 implemented

origin-based quotas on inflows from non-Western hemisphere countries. The quotas' effect was to impose binding constraints on immigrant flows from Southern and Eastern Europe while leaving Northern European immigration unconstrained. Local exposure varies due to differences in the composition of the foreign-born population in each area. Areas with larger portions of their immigrant community originating from constrained origins experienced sharper declines in immigrant inflows.[4] Identification uses this quasi-random "composition" component of the shock to observe the effect of limiting future chain migration to an area[5].

The identifying assumption required for the causal effect of the policy change to be consistently estimated is that areas that were more or less exposed were on similar health development trajectories. I verify this assumption with several pieces of evidence. To support the validity of the treatment's conditional independence, I show that the treatment effect estimates are consistent with the inclusion of covariates from a set of demographic, capital, and labor market controls chosen by LASSO procedure. Estimated treatment effects are consistent across alternative specifications. The LIFE-M dataset is combined with full count census data create a dataset composed of nearly the entire universe of individual lifespans in Ohio and North Carolina in 1900 to 1940. The outcome of interest is the change in average lifespan of working age (aged 15-55) individuals living in an area from one census to the next. The direct observation of lifespans within the treated population is a significant innovation. With other studies in this field focusing measured aged-adjusted mortality or other indicators of health, the outcome here is the first which captures treatment effects occurring over the entire future life-cycle of each treated individual.

This project contributes to our understanding of the impacts of immigration during the Age of Mass Migration and, more specifically, the impacts of the 1920s immigration policy changes. Results imply that the 1920s policy changes had a negative impact on the health of Americans. More proximally, they suggest that the large immigrant flows during this time period produced positive health spillovers to the native-born population, in addition to the labor market and demographic changes that have been identified. This study identifies the effects of the decreased immigration that occurred from 1925-1929 on individuals of working age during this decade; if these results are counter-factually extrapolated out to the larger immigration flows observed in the decades beforehand, generations of Americans prior to the policy change received health benefits from immigrant inflows to their local areas.

[4] Greenwood and Ward (2015) and Abramitzky et al. (2019) show how these quota policies impacted immigrant inflows at the national and local level.

[5] I control for pre-treatment foreign-born population share to account for the size component of the shock and focus on the coefficient from a shift-share-like treatment built off of pre-treatment origin shares and each origin countries exposure to the policy change to identify the impacts of an exogenous change in immigrant flows.

A long line of literature has examined the impact of immigrant inflows on conditions in receiving locations (Borjas, 1995; Card, 2001; Borjas, 2003; Ottaviano and Peri, 2012). Research has identified many significant effects of immigrant inflows in the US historical setting (Abramitzky et al., 2012, 2014; Tabellini, 2020; Sequeira et al., 2020; Eriksson and Ward, 2022). Additionally, dynamics around economic and demographic changes associated with the 1920s immigration policy changes have been studied extensively (Greenwood and Ward, 2015; Massey, 2016; Abramitzky et al., 2019). Alongside Ager et al. (2024), which examines changes in disease rates in cities driven by changes in urban congestion among immigrant communities, this project identifies the relevance of both local conditions and labor market conditions as relevant pathways through which the shock operated.

This project contributes to the literature examining the impacts of immigration on health conditions in destination locations by providing the first estimates of these impacts in terms of life expectancy. Research in this area has used a variety of proxies for health or health-related outcomes. A number of papers have used task and occupational risk data to identify how immigrants lessen the health burdens on native-born individuals in labor markets (Orrenius and Zavodny, 2009; Giuntella and Mazzonna, 2015; Giuntella et al., 2018, 2019; Dillender and McInerney, 2020). Another set of literature has examined the effects of immigration on health operating through changes in local conditions. Multiple papers have identified effects operating through increasing health care utilization, health expenditures, and infant/child & age-adjusted mortality rates (Gunadi, 2020; Rana et al., 2021; Aygün et al., 2021).

The rest of this paper is organized as follows: in the next section, 2, I describe the historical context surrounding immigration in this period, the political economy surrounding the policy changes, and health trends. Section 3 describes the data used for this analysis. Section 4 describes the empirical strategy used. Section 5 provides baseline results and investigates the incidence across the native-born population. Section 6 delves into labor market mechanisms. Section 7 provides closing remarks.

2 Historical context

2.1 Age of Mass Migration

Immigration in the late 19th and early 20th century United States was characterized by consistently large inflows and has been termed the Age of Mass Migration. In the late 19th century, these flows consisted primarily of individuals originating from Western and Northern Europe but the composition changed to be largely composed of individuals from Central, Eastern, and Southern Europe in the 20th century. Focusing on the setting of this study,

Figure 1 shows the overall foreign-born population share in Ohio and North Carolina while highlighting the portions consisting of immigrant from different national origins. Foreign-born shares peak in 1920 and decline afterward. Immediately noticeable is how this trend is driven by the expansion and contraction of the foreign-born population from Central and Eastern Europe.

For much of the 20th century, the conventional wisdom was that immigrants came to the US poorer than US-born individuals and worked hard to close this income gap over time. Compared to this story, Abramitzky et al. (2014) observe something different regarding the economic mobility possible for immigrants during this time period. In actuality, immigrants saw similar rates of occupational mobility as the US-born population. Those that came from poorer (richer) countries obtained lower (higher) income occupations in the US and, generally, remained occupied in these jobs. Figure 2 shows the distribution of foreign-born employment over time grouped by major 1950 basis occupation groups. The largest share of employment is composed of individuals employed as laborers but there is a declining trend in the size of this group across the entire time period. The decline here is driven by occupational upgrading, as the existing population of migrants ages, and changes in the occupational distribution of newly arrived immigrants post-quota.

Research examining the impacts of migration within the health literature set in the modern period grapple with the implications of the healthy migrant effect.[6] There is some evidence that this effect was relevant during the Age of Mass Migration. Spitzer and Zimran (2018) document positive selection into immigration on health characteristics within localities in Italy. Given similar health distributions, positive selection on health characteristics into migration would imply that immigrants had greater health capital than the US-born population.

In conjunction with World War I, the 1920s policies brought an end to the period of open migration to the United States. The disruption to international trade and relations essentially halted immigration during the war. After the war had ended, flows rebounded but were quickly brought to a halt by policy (the 1920s quotas). After the immigration policy changes, Massey (2016) finds that there was greater positive selection associated with exposure to the quota. The policies reduced the supply of low and unskilled labor through reduction in flow size and through reduction in immigrants working in these occupations among the remaining inflows. With immigrant concentration in urban areas, these areas with an existing hunger for low-skilled labor became even moreso.

[6]Immigrants to the United States are healthier than US-born individuals upon arrival but their health declines over time spent in the US to converge to natives.

2.2 Political economy of immigration policy

Concerns over the presence of immigrants and whether they were assimilating to US culture began before the policies implemented in the 1920s. In the 1800s, anti-immigrant sentiment on the east coast coalesced into an organized political party in the 1830s and 1840s, the Know-Nothings (Higham, 2002). Anti-European sentiment temporarily faded as a national concern with the dissolution of this party but immigrant from other areas continued to face legislative barriers to entry. The Chinese Exclusion Act passed in 1882 and placed a 10 year ban on all immigration from China. The Dillingham Commission began in 1906 with a focus on the assimilation of immigrant populations. Following the findings of the commission, support for anti-immigration policies galvanized and movement toward federally-implemented restrictions increased. In 1917, literacy tests were required for entry to the US but had little impact on the size of immigration flows due to widespread literacy. The end of the age of essentially open migration to the US was brought to a close with the 1920s immigration acts, the Emergency Quota Act in 1921 and the Immigration Act of 1924.

Increasing restrictions were in a significant way driven by stereotypes about the consequences of admitting immigrants. Markel and Stern (2002) describe how, across the 20th century, the perception of immigrants posing a significant public health risk was larger than the actual risk created. The early 20th century represented a time period in medical thought where germ and genetic theories were acknowledged but incomplete understandings led to the belief that genetics were at the root of observed differences in health and socioeconomic outcomes. In particular, these theories were used to justify stereotypes about the newer waves of immigrants being inherently more disease-prone or less economically successful. This set of beliefs justified an argument in anti-immigrant political thought that restricting immigration among these groups was the answer to infectious disease. As Ager et al. (2024) highlight, this was a false belief, immigrants concentrated in poorly maintained and crowded tenements in cities were more prone to disease because these local conditions would leave any population housed in these areas more exposed. As a self-fulfilling prophecy, reducing immigrant inflows did decrease infectious disease rates, not because immigrants brought disease with them but because there was no longer an population inflow being housed in unsafe local conditions.

The first act, the Emergency Quota Act, introduced quota policies which would stay in effect for more than 40 years until 1965. In the form introduced in 1921, these quota were based off of national origin with each country assigned a quota set at 3% off of the foreign-born population living in the US as of 1910 from that country. Immigration from the Western Hemisphere was left unconstrained. Three years later, the Immigration Act of 1924 revised these policies. First, the quota amounts were reduced from 3% to 2% of foreign-born

population. Second, this act moved the base year from which foreign-born populations were measured from 1910 to 1890. This revision had a significant effect on the composition of immigrants allowed entry to the US. From 1890 to 1910, shares of immigration flows from Northern and Western Europe decreased and those from Eastern, Central, and Southern Europe increased. This change had the effect of imposing binding constraints on the more recent waves of immigration while leaving Northern and Western European immigration essentially unconstrained.

2.3 Early 20th c. US health trends

The early 20th century United States is a setting with significantly higher mortality rates and lower life expectancies than studies examining health dynamics in the modern setting. Prior to the the establishment of OSHA or any form of federal workplace safety standards, the variation in on-the-job risk from white collar to unskilled and industrial positions is wider than the current time. There was some recognition of the risk workers in manufacturing faced, as the predecessor to OSHA, the Bureau of Labor Standards, began to collect workplace fatality data in the 1920s; this would culminate in the reporting of virtually all workplace fatalities by the end of the 1930s.

Among the distribution of occupations producing different effects, there is evidence that immigrants occupied those with worse health effects. Figure 3 reports the average lifespan of US-born individuals employed in occupations split by greater or less than average exposure to the quota policy shock and average foreign-born employment share. Moving from occupations with lower shares of foreign-born employment to higher shares, there is a decline in average lifespan, and, moving from low quota exposure to high quota exposure there is a decline in average lifespan. There is also evidence that immigrants resided in locations with worse public health conditions. Figure 4 reports average lifespans categorizing locations by high and low exposure and share. A similar pattern holds. Lifespans are lower in areas with higher exposure and foreign-born populations share. With the immigration shock driving shifts in labor supply of varying sizes, native-born population would be expected to transition to lower average lifespan locations and occupations.

This study takes place during a broad decreasing trend in mortality risk and increase in life expectancies in the early 20th century. From the mid-19th century to the 1940s, mortality rates in urban and rural areas consistently declined (Haines, 2001). Of interest, though crude mortality rates declined essentially continuously over the time period, there is some evidence of stagnation in this trend during the 1920s. Cutler and Meara (2001) examine age-adjusted rates by bin; the decline in the overall mortality rate in the 1920s was

driven almost entirely by decreasing infant mortality rates while mortality rates for cohorts aged 15 or older remained flat (compared to a declining trend in decades before and after).

Crude mortality rates for foreign- and US-born populations tells a similarly misleading story. From 1900 to 1960, crude mortality rates for foreign-born and US-born individuals diverge significantly with a rapid increase in mortality among the foreign-born. This effect is driven by the overall composition of the US-based foreign-born population shifting towards older ages. As the quotas cut off future flows, the existing population aged and mortality rates appear to increase. Breaking crude rates down to age cohorts by native-born status, Bakhtiari (2022) show a decline in mortality rates across all cohorts in both US-born and foreign-born cohorts.

A final important dynamic is the transition seen in urban and rural mortality rates. In both, a decline in mortality rates is observed across the sample period (Bakhtiari, 2022). During this time period, there is a significant urban mortality penalty. This was driven by population density, industrial concentration, and disease prevalence. As medical and public health developments were implemented across the US, this gap closed. Ager et al. (2024) study the impacts of the quota policies on conditions in urban areas; a decline in population density would be expected to decrease mortality rates due to infectious disease and they observe just this effect. Further, it appears that infectious disease's presence in US urban areas was not a product of immigrants 'importing' the illnesses to these areas.

In the modern health and immigration literature, there is a widely documented health gap between native-born and immigrant populations with immigrants arriving to their destination in better health. This effect is not observed in the US during this time period (despite evidence of positive selection on health characteristics) and immigrant populations appear to be less healthy than the US-born population. This gap appears to be significantly driven by the concentration of immigrant populations in urban areas. As the urban mortality penalty disappears so to does the gap between US-born and immigrant populations. By the end of the 20th century, the familiar immigrant mortality advantage is observed in the United States.

3 Data

This study combines several existing datasets to produce a county-level dataset covering Ohio and North Carolina at decade intervals from 1900 to 1930. The LIFE-M project, full count census, and quota exposure data calculated in Abramitzky et al. (2019) are used to produce this dataset. In addition to local demographic and economic characteristics, county-level quota exposure measures and average working-age individual lifespans are derived from

these combined data. The LIFE-M data allows for the direct observation of individual lifespans and increases the rates at which individuals may be linked across censuses by utilizing information from vital records.

The LIFE-M (Longitudinal, Intergenerational Family Electronic Microdata) data "provides a life-course and intergenerational perspective on the evolution of health and economic outcomes" (Bailey et al., 2022). The combination of vital records (birth, marriage, and death certificates) with census records provides a panel representation of individual demographic and economic characteristics, locations, and major life events and census and event intervals; this means that individual locations are noted at decade intervals and when major events such as birth, marriage, or death occur. These records span multiple generations in Ohio and North Carolina. Construction of these records began with collection of records for a 'core generation' born between 1900 and 1929. From there, two generations previous were identified from vital records as well as the children of the core generation for a total of four generations of families. This data uses information available in vital records and marriage certificates to supplement information used in standard linking methods, enhancing cross-dataset and cross-census matches[7].

This analysis focuses on the population of working age individuals in 1900, 1910, 1920, and 1930. This means that the majority of individuals included in the analysis are comprised of those from the core generation and their parents.[8] Of the close to 3 million working age individuals observed in Ohio and North Carolina in the pre-treatment period (1900), the LIFE-M dataset allows me to record the lifespan for 49.2% of this population. The average lifespan among the working age population in 1900 in my sample is 71.6 years. If these averages seem high, note that conditioning on reaching age 15 partials out the effects of exceptionally high infant mortality rates in this time period[9]. Focusing on the working age sample eliminates concerns over correlations with factors driving the decline in infant and childhood mortality that were ongoing.

Industry-level disabling injury rates per 100,000 work hours are collected from the Monthly Labor Review, a BLS publication (Kossoris and Kjaer, 1941). This publication presents results from a survey of over 20,000 business establishments in 1939 and 1940. The number of employees, work hours, number of disabling injuries, and total days lost due to disability are reported by industry (the industries in this report closely align with the 1950 census basis).

[7]See (Bailey et al., 2022) for more information about data collection and matching method used. Marriage forms' identification of parents and maiden names significantly improves match rates for women, both to census records and between them. These records also aid in the identification of household structure (which itself improves census matching) and the identification of broader cross-household family structures.

[8]This is relevant because authors note that the LIFE-M data suffers from two-sided censoring in early and late records due to incomplete digitization.

[9]Cutler and Meara (2004) find an infant mortality rate of $> 16,000$ per $100,000$ individuals in 1900

Injuries are reported for employees of the establishment and self-employed individuals working in these establishments. Notably absent are agriculture and mining industries (though manufacturing establishments serving both are included).

From full-count census data (Ruggles et al., 2024), several county-level statistics used in the construction of the outcomes, treatment variable, and covariates are derived. Population counts and the samples of individuals used to calculate foreign-born population shares of each national origin encompass all individuals who are between 15 and 55 years of age in 1900. Records from the LIFE-M data are linked to full count census data from the 1900 to 1930 censuses.[10] County-level lifespans are calculated by identifying the 15 to 55 aged population in each county fitting the relevant individual criteria and linking to LIFE-M birth and death records to obtain observed lifespans. Within the set of counties in Ohio and North Carolina, urban and rural subsets are identified by calculating the average portion of each county's population living in urban areas and defining urban areas as those with greater than or equal to the mean urban population share.

The treatment variable, quota exposure, is calculated by combining the initial population shares within the foreign-born community of each national origin with the nation-specific quota exposures calculated in Abramitzky et al. (2019). These nation quota exposure values are reported in Figure 5. In the simplest form, the nation-specific exposure is calculated as a dummy variable indicating whether immigration from a country was constrained by the 1924 quotas. The treatment here is then the sum of the share of the foreign-born population from each country times the dummy.

$$QE_c = \sum_o \frac{Foreign\ Pop._{1900,o}}{Total\ Foreign\ Pop._{1900}} \times \mathbf{1}\{Constrained_o\}$$

where c indexes the county, o indexes the origin country, and $\mathbf{1}\{Constrained_o\}$ is an indicator for whether immigration flows from country o faced a binding constraint after the 1920s policy changes. The quota exposure treatment of this form is referred to as $QE1$ hereafter. This quota exposure measure can be interpreted as indicating what percentage of the foreign-born community in an area were of national origins facing binding constraints.

Alternative treatment variables are based off of Abramitzky et al. (2023)'s definitions and are termed $QE2$ and $QE3$ moving forward. The preferred treatment measure is $QE3$.

[10]The LIFE-M data provide crosswalks to link these datasets by the histid field; this field is commonly used in cross-census and cross-dataset crosswalks. When creating a set of individuals linked across censuses (such as for identifying location, occupation, or industry change), Census Linking Project crosswalks are used; the LIFE-M data doesn't link to the 1930 census and individuals linked between the 1920 and 1930 censuses are a key population of interest in this empirical setup.

In these measures, the indicator variable is replaced with a value representing the percentage of immigration flows constrained under varying counterfactual assumptions, $QI_{o,2}$ and $QI_{o,3}$. $QE2$ assumes that immigration in the period from 1922 to 1929 would have been the same as 1902 to 1909. $QE3$ calculates counterfactual flows in the 1920s based off of separate time trend regression for each national origin. For a specific origin country, this is calculated as the difference between projected counterfactual flows and quota slots assigned in the 1920s, normalized by counterfactual flows; this sets quota intensity equal to zero if the quota slots given are greater than projected flows and the measure increases to one as greater shares of the projected flows are constrained from entry.

4 Empirical strategy

This project utilizes the identification strategy of Abramitzky et al. (2019) and Ager et al. (2024) and begins by estimating the effect of exposure to the quota policies on the working age population in Ohio and North Carolina. Afterward, I examine sub-populations and subsets of counties to more closely identify the incidence of these effects across the whole population. Additionally, several alternative outcomes are examined.

Results are estimated using a difference-in-difference framework with counties observed repeatedly each decennial census (ten year intervals). County and year fixed effects control for effects which are constant within counties over time or across locations in each year. The estimating equation is

$$Y_{ct} \quad = \quad \alpha_c + \delta_t + PT_t \times QE_c \; \beta + PT_t \times X'_{c,t=1900} \; \lambda + \epsilon_{ct}$$

where Y_{ct} is the dependent variable for county c in time period t, δ_c and σ_t represent county and year fixed effects, QE_c is the quota exposure calculated for county c, $X'_{c,t=1900}$ is a vector of pre-treatment controls from county c in 1900, PT_t is a dummy indicating whether the time period t is post-treatment, and $epsilon_{ct}$ are the residuals. The dependent variable, Y_{ct} is the differenced average lifespan of individuals aged 15-55 in each county as of the census. $X_{c,t=1900}$ always contains the foreign-born share of the population in 1900 and sometimes includes additional controls; these controls are interacted with the post-treatment dummy to allow for non-constant trends with respect to these pre-treatment controls over time. The results presented use heteroscedasticity robust standard errors clustered at the treatment (county) level. Observations are weighted by 1900 native-born working age population.

I study counties in Ohio and North Carolina from 1900-1910 and 1920-1930. Following Abramitzky et al. (2019) exclude 1910-1920 to avoid the disruptions to immigration and local

conditions caused by WWI. In the appendix, I include specifications with 1910-1920 included and show that my results do not significantly changed (in fact, estimated treatment effects are even larger in these specifications). Differencing the outcome has two advantages. First, it converts the conditional parallel trends assumption into a conditional mean independence assumption; so we can interpret the required condition as stating that the average change in lifespan is equal across treatment levels conditional on included controls. The coefficient on the base treatment term tells us what the correlation between the treatment and the change in lifespan from 1900 to 1910 is, and so, this coefficient is an indicator of whether this condition is violated. Second, this eliminates concerns over heterogeneous treatment effects making our estimates inconsistent (Roth et al., 2023). The change in treatment status being driven by a shock-based innovation (in this case, a policy change causing an exogenously varying shock across locations) reduces concerns about serial autocorrelation that might arise from use of a shift-share treatment (Abramitzky et al., 2023).

The specification used is as close as possible to the canonical DiD case (two periods, all treated in the second period, treated and untreated states) to avoid as many of the issues highlighted in the recent DiD literature as possible.[11] The main departure comes from the treatment variable; quota exposure is non-binary but always positive. This specification preserves the two-period setup with all units treated in the second period. This meets a sufficient condition to ensure weights on each comparison are always positive (de Chaisemartin and d'Haultfouielle, 2023).[12]

There are several assumptions required for consistent estimation of the average treatment effect in the difference-in-difference framework. I will briefly state these assumptions and then discuss the likelihood of each being met. First, SUTVA (Stable Unit Treatment Value Assumption) must hold; that is each location must experience the shock from the policy change only through the treatment value observed in that location. Secondly, as previously mentioned the conditional parallel trends assumption must be met. Third, the non-anticipation assumption must be met; treatment levels must not have been predictable prior to the treated period, if not, the outcome values prior to this period might be influenced by the treatment level. Lastly, an overlap condition must be met; for any set of potential observable characteristics, there must be a non-zero probability of being treated or untreated.

SUTVA may be violated if internal flows between counties or from other population flows respond to treatment (e.g. - individuals moving the during the Great Migration differentially

[11] See Roth et al. (2023) and De Chaisemartin and d'Haultfoeuille (2023) for reviews of recent developments in the difference-in-difference and TWFE literature.

[12] That condition is that the treatment during the treated period is higher than the average across both periods. Since none are treated in the first and all are treated in the second, this condition is met.

settle in more or less highly treated areas). I provide two pieces of evidence that this is not a significant concern; these results are provided in the appendix. First, I provide alternative specifications where I restrict the studied population to white individuals born in Ohio or North Carolina and see no significant change in the estimated treatment effect. Second, I test whether the proportion of individuals I link from one census to the next who move out of their initial county changes in response to treatment. Here, I observe no significant change in internal migration rates associated with quota exposure.

The conditional parallel trends assumption requires locations to be on similar health development trajectories conditional on controls. To test this assumption, I run a LASSO procedure of the counties' quota exposure levels on local demographic, labor, and capital characteristics in 1900 and provide alternative specifications with these controls included. Results here indicate that the gender ratio in 1900 may be a significant predictor of treatment level in addition to foreign-born population share. The alternative specifications do not significantly change the estimated treatment effect. Abramitzy et al (2023) run a similar LASSO procedure at the State Economic Area (SEA) level which indicates that percentage of workers employed in agriculture may be relevant; I also include alternative specifications with this control in the appendix.

The non-anticipation assumption may be violated if immigrants foresaw the impacts of the policy change ahead of its implementation and changed their settlement patterns or policymakers foresaw their areas quota exposure and shifted their behavior in a way that changed the trajectory of health development. Statistical evidence and contemporary commentary suggest the effects of the policy change were not predicted in this way. First, immigration flows were significantly decreased during the 1910s as World War One went on but they rebounded in the 1920s prior to the quotas implementation. Abramitzky et al. (2019) demonstrate that this treatment is uncorrelated with immigrant flows in prior decades.

5 Changes in resident health

5.1 Baseline

Examining trends in the outcome in levels from 1900 to 1930, average lifespans in the sample decline from 1900 to 1920 and then increase from 1920 to 1930. The average cohort lifespan across counties in 1920 is 66.8 years of age with a standard deviation of 3.36. The average treatment level across counties are 0.155 (for QE3) and 0.123 (for QE1) with standard deviations of 0.160 and 0.162 respectively.

In Table 1, the coefficients on the exogenous (quota exposure) and endogenous (foreign-

born pop. share) components of the immigration shock are reported for two quota exposure measures (QE1 and QE3). With both sets of geographic fixed effects (county and state), I find virtually identical significantly negative effects associated with post-treatment quota exposure; the coefficient for QE3 is −6.80 (sd of 1.32 for state FE and 1.53 for county FE) and, for QE1, −7.35 (sd of 1.33 for state FE and 1.51 for county FE) Given mean treatment levels, this implies that the ATE in a county with mean exposure is associated with a decline of 1.05 and 1.19 years of lifespan, respectively. My preferred exposure measure moving forward is QE3, the counterfactual estimates here take into account trends in migration waves from countries over time rather than assuming linear or constant immigrant trends. These results are driven by declines occurring in the later half of the 1920s; it's possible that even larger effects might be observed if the policies had alternatively been implemented at the beginning of the decade (and individuals saw five more years of exposure to conditions with decreased immigration flows).

It is also potentially interesting to note the result from the endogenous portion of the shock (represented in my regression model by foreign-born % in 1900); the coefficient here is significant and positive (7.17, sd = 1.26 and 7.40, sd = 1.30) implying that areas where large immigrant populations were settling remained on higher health development trajectories post-shock. These results align with observations from other research identifying the settlement patterns of immigrants in the early 20th century. Large immigrant communities developed in urban areas. At this time, there was still a significant urban mortality penalty but this gap was closing as public health investments and medical advances decreased the danger from infectious disease exposure in cities.

5.2 Labor market and location transitions

The analysis presented aggregates individuals to treatment-level (i.e. the county) and should be viewed as an intent-to-treat design for an important reason, the underlying population is mobile. Realized changes in the health conditions faced might come from changes within the location and labor sector individuals reside in at the time of treatment or as a result of individual transition to another sector with different conditions. This subsection identifies the relevance of the changes in local conditions and labor market conditions driven by sectoral transitions and within sector changes by observing the job and location change decisions of individuals in the following decade. Regressions accounting for these conditionals are presented in Table 2.

The first column of the table presents the preferred specification from the baseline results for comparison; this corresponds to the second column of the Table 1. These results suggest

there was a significant change in the average lifespan of individuals across the entire working age population. Column (2) presents results from difference-in-difference regressions comparing populations who responded to the shock by changing location to those who remained in the treated county. Among the non-moving population, there are significant declines in life expectancy, -4.95 (sd = 2.30). There's suggestive evidence that the population which moved after the shock experienced a larger shock but the interaction term here is not significantly different from the base term, -2.36 (sd = 3.09).

Column (3) reports results differentiating the working age population between those observed migrating to a different county in the following decade and those identified in the same county in both periods. I find that the treatment effect estimated among the stayer population is slightly smaller than the pooled treatment effect, -4.95 (sd = 2.30). The interaction term for the mover population suggests they experienced a larger shock but it is not statistically significant, -2.36 (sd = 3.09). The previous results suggest that a small portion of the aggregate effect operated through changes in local conditions. The difference in coefficients here suggest that changes in local health conditions experienced was driven by job change-associated relocation. Abramitzky et al. (2019) note that stayers saw a decline in occupational income score that movers were able to avoid. It is possible that movers traded health capital for income (in much the same way as immigrants have been observed doing in other settings).

5.3 Heterogeneous treatment effects

The significant negative effect observed in my baseline results is in essence a weighted average of potentially heterogeneous treatment effects across groups in the underlying population. To identify these effects and their incidence across the general population, this section reports results from regressions splitting the population of each county by group or examining subsets of counties fitting certain criteria. I look for heterogeneous treatment effects in sub-populations by gender, age, internal migrant status (post-treatment), and examine whether urban and rural counties experienced the shock differently. Results from these regressions are reported in Table 3.

Column (1) reports the same results as column (2) of Table 1; here, the third quota measure is used as the treatment variable with county fixed effects and 1900 foreign-born population share included as a saturated control. This setup is replicated across the remaining columns with a dummy variable indicating the sub-population interacted with the treatment and foreign-born share. If there are concerns that there are heterogeneous treatment effects with respect to the exogenous portion of the policy change, there might reasonably

be concerns about heterogeneous treatment effects with respect to the endogenous portion as well. Saturating the model in all time periods with pre-treatment foreign-born population share accounts for this variation. Additionally, all regressions (like the baseline estimates) are reported with robust standard errors clustered by county and weighted by working age population in 1900.

Column (2) reports results examining the impact of quota exposure on male and female cohorts' average lifespans across counties. Within this sample, both groups follow a similar pattern to the overall trend (declining from 1900 to 1920 then increasing from 1920 to 1930). The base term interacted with the post-treatment dummy indicates the effect on women while the effect on men is found by combining the first and second row coefficients. I find that the treatment effect on women is significantly larger than that on men. The coefficient for women is -10.62 (sd = 2.22) and quota exposure post-treatment for men is 6.71 (sd = 2.71). Both men and women see a decline in lifespan associated with exposure to the quota policy but the effect is larger for women than men.

These results may be driven by several factors. Women may see shocks to lifetime income driven by changes in marriage sorting, local economic condition decline, or spousal income changes. There may be more drastic changes in sorting in labor markets among women than men; women employed outside of household production were restricted to employment in manufacturing and low-skilled white collar jobs during this time. Thus, a change in sorting among women may lead to a drastic change in job conditions as they move from white collar to manufacturing positions. Men on the other hand have fewer restrictions on employment possibilities; those that change positions may move from some already dangerous job to one marginally worse.

Column (3) reports results splitting the sample into sub-populations of 15 to 34 years old and 35 to 55 years old. In the sample population, the 35 to 55 year old population lives longer than the 15 to 34 year old population in all years except for 1930. Results here indicate that the decline in lifespan caused by the policy change was concentrated among the younger working age cohorts. There's no statistically significant effect among the working age population aged 35 to 55. On the other hand, there is a strong statistically significant effect among those aged 15 to 34; the coefficient for younger cohorts is -14.70 (sd = 3.41). This effect implies that there was a decline in the lifespan of younger working age individuals of roughly two and a quarter years associated with the average policy exposure level. The stronger result here makes some sense. Younger cohorts are exposed for a larger share of their lifespan to the post-quota state. This is consistent with the labor market model suggested by Llull (2017); younger cohorts are more significantly impacted because they have accumulated less human capital and so are more likely to respond to changes in labor

supply across occupations with job change. Older cohorts have accumulated more human capital tailored to their existing profession and so are less likely to transition professions.

Column (4) compares urban and rural subsets. Urban and rural counties are divided between those with greater than the mean share of population living in urban areas and less than it. The base post-treatment coefficient represents the treatment effect in rural areas while the urban treatment effect is found by combining these coefficients. The results in rural areas produce an estimated treatment effect similar in magnitude to the overall treatment effect, -4.04 (sd = 1.04). The coefficient for urban areas implies that the treatment effect may be notably larger than that in rural areas, -5.15 (sd = 3.50).

Both urban and rural areas which are more exposed to the immigration shock would be predicted to see income declines which might drive health declines. The potentially larger effect in urban areas could be driven by the transitions into manufacturing and towards cities observed by Abramitzky et al. (2023). Manufacturing employment opportunities were concentrated in urban areas. If individuals transitioned from agriculture to manufacturing, they'd need to move from areas where the former were located to areas where the later were. This location changes is monetarily and socially costly. Individuals already located in urban areas avoided these costs and would be more likely to make this employment change.

5.4 Robustness checks

In the Appendix, I report results from multiple alternative specifications. Table A.1 displays results from baseline estimates with the second quota measure; while this treatment measure is dis-preferred because of the potentially unrealistic assumption that immigrant flows would have been constant across time periods in the counterfactual, the estimated treatment effect is significant, negative and of a larger magnitude than the other quota exposure measures. Table A.2 and Table A.3 reports baseline estimates without weighting observations by working age population in 1900; the estimated treatment effect is smaller in magnitude but still statistically significant, this is likely driven by larger treatment effects in urban areas. Table A.4 and A.5 report baseline results with observations for each county from 1900 to 1910 included; the estimates remain statistically significant and of the same magnitude as those excluding this time period. Table A.6 and A.7 report results from a two period regression comparing the outcomes in levels in 1920 and 1930; these estimates are statistically significant and negative but the estimated treatment effect is slightly smaller than the baseline specification with outcomes differenced. Table A.10 reports regression results including covariates from Abramitzky et al. (2023)'s LASSO procedure (1900's agricultural employment share). The inclusion of this additional control doesn't significantly change estimated

coefficients.

Varying the base year from which the share component of the treatment variable may suggest how remaining unobserved factors influence my estimates. In Table A.8 and Table A.9, the quota exposure variables are recalculated for each county setting the base year to 1920 instead of 1900. Selecting the earlier year is preferred because it is less likely to be correlated with unobserved endogenous factors (or, put in another way, one would expect the treatment with 1920s shares to covary more strongly with unobserved variables and bias my treatment effect estimates more). Results from these regressions follow the same significance and sign patterns as the earlier base year results. Since the change of base year's only effect should be in potentially strengthening the covariance between the treatment and unobserved bias, this suggests that the impact of unobserved variables is to downward bias my treatment effect estimates.

6 Labor market mechanisms

A large number of spatial equilibrium models of internal immigration break changes in utility (individual health capital) in a location after a shock into labor market changes (i.e., wages, employment health conditions) and local amenity values (i.e. public services, housing, environmental or health conditions) (Jia et al., 2023). Baseline results describe the change in health capital potentially operating across both pathways. Ager et al. (2024) present evidence that significant health effects of the policy change operated through local conditions. The following section presents results indicating that the immigration shock led to labor sorting into occupations and industries with more deleterious health conditions.

Cross-industry and cross-occupation variation in health conditions establish the relevance of changes in labor market health conditions in the net effect on health. The outcomes examined are observed at the occupation and industry level and so the relevant sample which treatment impacts are individuals who might be driven to switch occupations or industries by the shock. These results represent a lower bound on the effect operating through changes in labor market conditions as there might be within occupation (or even single employment) changes in work patterns unobservable without individual data on task assignment.

To examine how sorting across occupations impacted the health conditions faced, the outcome of interest is the average lifespan of working age individuals employed in that occupation in 1900 in the LIFE-M dataset. To examine cross-industry transitions, I use Department of Labor survey information on rates of disabling injuries from the June 1941 Monthly Labor Review. This survey reports information from over 20,000 business establishments spanning the majority of US industry (though notably absent are agriculture and mining

establishments. The entire working age population in the 1900 and 1920 full count censuses are linked to the following census with Census Linking Project crosswalks and matched to the industry/occupation-specific measures. The county-level variable used in regressions is the average difference between the beginning and end of decade outcomes in linked individual sample.

The identification strategy here is directly comparable to Abramitzky et al. (2019). Differences in employment health indicators are driven by observed transitions between industries or occupations among the linked sample of working age individuals. An alternative sample examined is the sample of individuals aged 15-35 at the time of the shock; one would expect to see stronger effects among this population since they've accumulated less human capital related to previous occupation choices and should more flexibly respond to the changes in labor supply. The empirical specification used is identical to the previous section with the additional inclusion of state-year FE. The controls included are 1900 foreign-born working age population percentage, state-year, and county FE. Standard errors are robustly clustered at the treatment-level (by county).

I focus on two sample definitions to examine estimate the average treatment effect and identify whether aggregate changes are driven by changes in transition rates or worsening job prospects. The full sample of linked individuals provides an estimate of the average treatment effect across the entire working age population in a location. The next subsample examined are the subset of linked individuals who changed occupation or industries over the decade; the coefficient from regressions with this sample identify whether switchers in more exposed locations transitioned to more or less deleterious job conditions.

6.1 Occupation-specific changes

Table 4 presents results from the county-level difference-in-difference regressions with the difference in average lifespan of occupation of employment as the outcome variable. The outcome uses average lifespans calculated by occupation with the LIFE-M data. These outcomes are matched to the occupations of employment for the entire United States male working age population who are linked to the next census with Census Linking Project crosswalks.

Columns 1 through 4 use the entire sample of working age men linked from one decade to the next in the 1900 to 1930 censuses (either 15 to 55 years old or 15 to 35 years old) while columns 5 through 8 focus on the subset of individuals who switched occupations during this period. The first set of results provides a coefficient which should approximate the average treatment effect on the entire male working age population. The second sample's

identifies whether treatment variation impacts occupation-associated health conditions on the intensive margin (so, conditional on switching occupations, does a larger shock lead to switching to even worse health conditions in the next occupation). Odd columns present results with just state-year FE and even include state-year and county FE. The preferred specification is that including county FE (the other is included to demonstrate the stability of coefficient estimates).

Among all working age men (aged 15-55), quota exposure is associated with a significant decline in the average lifespan of occupation of employment, -0.10 (sd = 0.05), and larger effects among the younger set of cohorts (aged 15-35), -0.13 (sd = 0.06). The larger magnitude among early career workers is consistent with the amount and form of human capital accumulated as a worker gains more experience inducing frictions in job transition. Results from the subset of individuals who changed occupations suggests there were significant effects on the intensive margin as well. Among all working age men, there is a significant decline in the average lifespan of the next occupation, -0.14 (sd = 0.07), and the magnitude is once again larger among early career workers, -0.17 (sd = 0.07).

These results imply that the shock from quotas led to significant occupational downgrading to worse health conditions. They align with theoretical expectations of a model where immigrants were employed in positions with larger health capital costs. This doesn't prove the existence of the healthy migrant effect in this time period but is consistent with the existence of this effect.

6.2 Industry-specific changes

Table 5 presents results from the county-level difference-in-difference regressions with the difference in the number of disabling injuries per 100,000 work hours as the outcome variable. The outcome uses injury rates collected by the BLS from over 20,000 business establishments in 1940 and reported in the June 1941 Monthly Labor Review. Injury rates are reported by industry[13]. These industry-specific rates are matched to the industries of employment for the entire United States male working age population who are linked to the next census with Census Linking Project crosswalks.

Columns 1 and 2 use the entire sample of working age men linked from one decade to the next in the 1900 to 1930 censuses (either 15 to 55 years old or 15 to 35 years old) while columns 3 and 4 focus on the subset of individuals who switched occupations during this period. The interpretation of the first set as the average treatment effect and the second set as a measure of the change in conditions on the intensive margin is identical

[13]These industries align significantly with the 1950 basis used in the census

to the occupation-specific analyses in the previous section. One difference here is that the identifying variation comes from the population who could be induced to change industries (rather than occupations).

Column 1 reports results from the regression with the full sample of adult working age men; quota exposure is associated with a significant increase in the injury rate per 100k hours, 0.13 (sd = 0.06). These per hour injury rates can be converted to estimates of the number of disabling injuries per 100k workers in a year. Assuming 50 hour work weeks, the mean exposure to the quota was associated with 47 more disabling injuries per 100k workers each year. Among early career workers (15-35 years old), these effects are slightly larger and significant, 0.15 (sd = 0.07). Columns 3 and 4 present results from the subset of the sample observed switching industry of employment in the following decade. The coefficients for the full sample and early career workers are positive and of similar magnitude, 0.22 (sd = 0.10) and 0.24 (sd = 0.10) respectively. Results indicate that industry switchers were likelier to enter even more dangerous industries as the treatment increased.

7 Conclusion

The restrictions enacted in the 1920s brought an end to a decades long open immigration regime in the US and represent, potentially, the starkest change of these policies in US history. A priori, it's unclear what the aggregate effect of health spillovers to the native-born community would be. There are many potential channels through which the shock might operate. Further, existing research shows shows that one shock might generate positive and negative effects operating through different channels. The labor shock might produce negative spillovers as natives sort into more dangerous jobs or otherwise via declines in income. On the other hand, these policy changes have been identified as producing positive spillovers from reduced overcrowding and infectious disease in cities Ager et al. (2024). Additionally, other mechanisms driving within occupation or location changes in health conditions have been identified as relevant in the modern setting.

Results from this analysis suggest there were large declines in the health capital of the resident population in the wake of quota implementation. This casts doubt on the long-running belief that immigrants negatively impact health conditions in the United States. The concentration of significant treatment effects among the population engaging in occupation changes suggest conditions in labor market opportunities drove a significant portion of the net effect. Individuals changing occupation and industry in response to the shock moved into positions with more deleterious effects on health capital. These results suggest some of the effect was driven by changes in on-the-job characteristics and also the coincident relocation

to areas with worse local health conditions.

A decline in lifespan among working age cohorts exposed to the shock in the 1920s has potentially far reaching consequences on economic conditions and health expenditures in local areas thereafter. Spillover effects of increased health utilization among these cohorts might lead to healthcare congestion in markets impacting later generations. The increased health expenses might also strain local public services requiring the diversion of resources from other productive means to the healthcare sector. Future research might focus on identifying within location, occupation, industry, or job changes in health conditions, use newer linking methods to extend the analysis to the female labor force, or consider policy responses which might've been enacted in response to the changes in health capital.

References

Abramitzky, R., Ager, P., Boustan, L. P., Cohen, E., and Hansen, C. W. (2019). The effects of immigration on the economy: Lessons from the 1920s border closure.

Abramitzky, R., Boustan, L., Eriksson, K., Rashid, M., and Pérez, S. (2022a). Census Linking Project: 1900-1910 Crosswalk.

Abramitzky, R., Boustan, L., Eriksson, K., Rashid, M., and Pérez, S. (2022b). Census Linking Project: 1920-1930 Crosswalk.

Abramitzky, R., Boustan, L. P., and Eriksson, K. (2012). Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. *American Economic Review*, 102(5):1832–1856.

Abramitzky, R., Boustan, L. P., and Eriksson, K. (2014). A nation of immigrants: Assimilation and economic outcomes in the age of mass migration. *Journal of Political Economy*, 122(3):467–506.

Ager, P., Feigenbaum, J. J., Hansen, C. W., and Tan, H. R. (2024). How the other half died: Immigration and mortality in us cities. *Review of Economic Studies*, 91(1):1–44.

Alacevich, C. and Nicodemo, C. (2024). The effect of immigration on occupational injuries: Evidence from administrative data. *Oxford Bulletin of Economics and Statistics*, 86(2):209–235.

Aygün, A., Kırdar, M. G., and Tuncay, B. (2021). The effect of hosting 3.4 million refugees on native population mortality. *Journal of health economics*, 80:102534.

Bailey, M. J., Lin, P. Z., Mohammed, A. S., Mohnen, P., Murray, J., Zhang, M., and Prettyman, A. (2022). Life-m: The longitudinal, intergenerational family electronic microdatabase.

Bakhtiari, E. (2022). The missing mortality advantage for european immigrants to the united states in the early twentieth century. *Demography*, 59(4):1517–1539.

Bellés-Obrero, C., Martin Bassols, N., and Vall Castello, J. (2021). Safety at work and immigration. *Journal of Population Economics*, 34(1):167–221.

Borjas, G. J. (1995). The economic benefits from immigration. *Journal of economic perspectives*, 9(2):3–22.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The quarterly journal of economics*, 118(4):1335–1374.

Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, 19(1):22–64.

Cutler, D. M. and Meara, E. (2001). Changes in the age distribution of mortality over the 20th century.

De Chaisemartin, C. and d'Haultfoeuille, X. (2023). Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey. *The Econometrics Journal*, 26(3):C1–C30.

Dillender, M. and McInerney, M. (2020). The role of mexican immigration to the united states in improved workplace safety for natives from 1980 to 2015. *Journal of health economics*, 70:102280.

Eriksson, K. and Ward, Z. (2022). Immigrants and cities during the age of mass migration. *Regional Science and Urban Economics*, 94:103593.

Giuntella, O. (2013). Why does the health of immigrants deteriorate? evidence from birth records. *IZA Discussion Paper No. 7588*.

Giuntella, O., Kone, Z. L., Ruiz, I., and Vargas-Silva, C. (2018). Reason for immigration and immigrants' health. *Public health*, 158:102–109.

Giuntella, O. and Mazzonna, F. (2015). Do immigrants improve the health of natives? *Journal of health economics*, 43:140–153.

Giuntella, O., Mazzonna, F., Nicodemo, C., and Vargas-Silva, C. (2019). Immigration and the reallocation of work health risks. *Journal of Population Economics*, 32:1009–1042.

Greenwood, M. J. and Ward, Z. (2015). Immigration quotas, world war i, and emigrant flows from the united states in the early 20th century. *Explorations in Economic History*, 55:76–96.

Gunadi, C. (2020). Immigration and the health of us natives. *Southern Economic Journal*, 86(4):1278–1306.

Haines, M. R. (2001). The urban mortality transition in the united states, 1800-1940* voici u, n test. In *Annales de démographie historique*, number 1, pages 33–64. Cairn/Softwin.

Higham, J. (2002). *Strangers in the land: Patterns of American nativism, 1860-1925*. Rutgers University Press.

Jia, N., Molloy, R., Smith, C., and Wozniak, A. (2023). The economics of internal migration: Advances and policy questions. *Journal of economic literature*, 61(1):144–180.

Kossoris, M. D. and Kjaer, S. (1941). Industrial injuries in the united states during 1940. *Monthly Labor Review*, 53(2):327–354.

Lebowitz, M. (2023). Trump sparks republican backlash after saying immigrants are 'poisoning the blood' of the u.s.

Markel, H. and Stern, A. M. (2002). The foreignness of germs: the persistent association of immigrants and disease in american society. *The Milbank Quarterly*, 80(4):757–788.

Massey, C. G. (2016). Immigration quotas and immigrant selection. *Explorations in Economic History*, 60:21–40.

Orrenius, P. M. and Zavodny, M. (2009). Do immigrants work in riskier jobs? *Demography*, 46:535–551.

Ottaviano, G. I. and Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European economic association*, 10(1):152–197.

Peri, G. and Sparber, C. (2009). Task specialization, immigration, and wages. *American Economic Journal: Applied Economics*, 1(3):135–169.

Rana, R. H., Alam, K., and Gow, J. (2021). Financial development and health expenditure nexus: A global perspective. *International Journal of Finance & Economics*, 26(1):1050–1063.

Roth, J., Sant'Anna, P. H., Bilinski, A., and Poe, J. (2023). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics*.

Ruggles, S., Flood, S., Sobek, M., Backman, D., Chen, A., Cooper, G., Richards, S., Rogers, R., and Schouweiler, M. (2024). Ipums usa: Version 15.0 [dataset].

Sequeira, S., Nunn, N., and Qian, N. (2020). Immigrants and the making of america. *The Review of Economic Studies*, 87(1):382–419.

Siegel, M. (2023). Migrants bussed from the border is not just an immigration crisis. think public health.

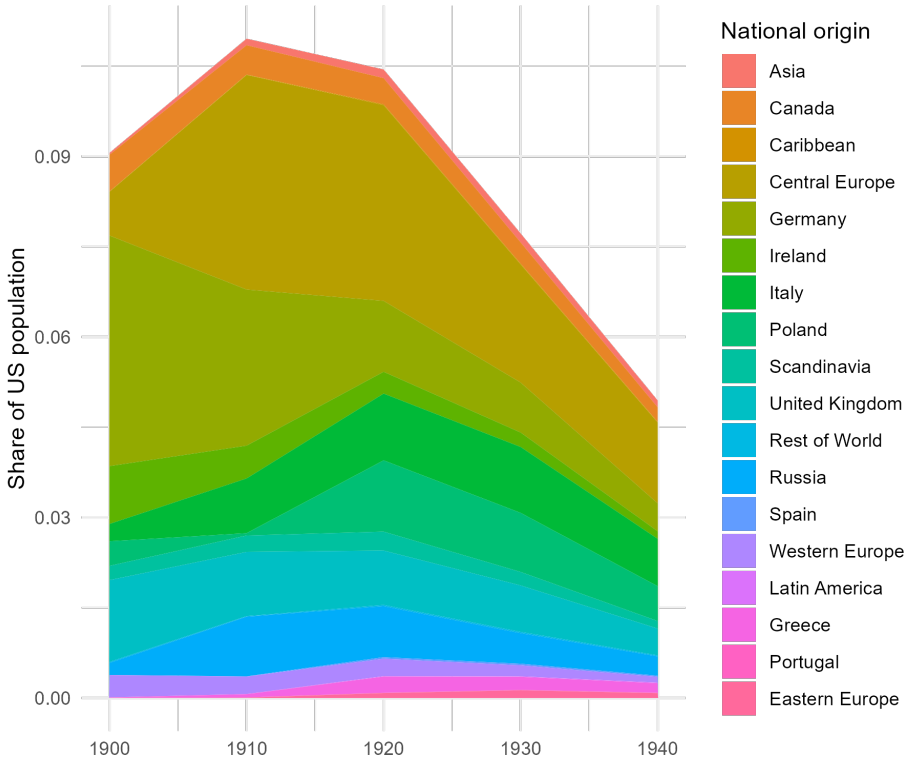
Sparber, C. and Zavodny, M. (2022). Immigration, working conditions, and compensating differentials. *ILR Review*, 75(4):1054–1081.

Spitzer, Y. and Zimran, A. (2018). Migrant self-selection: Anthropometric evidence from the mass migration of italians to the united states, 1907–1925. *Journal of Development Economics*, 134:226–247.

Tabellini, M. (2020). Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration. *The Review of Economic Studies*, 87(1):454–486.

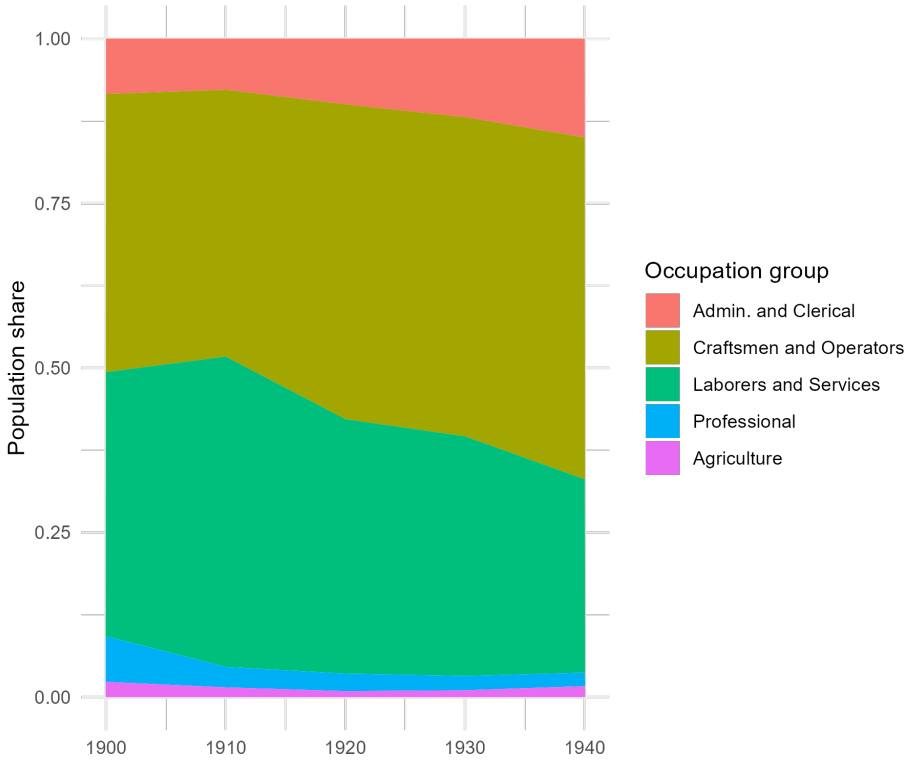
Ward, Z. (2017). Birds of passage: Return migration, self-selection and immigration quotas. *Explorations in Economic History*, 64:37–52.

Figure 1: National origin population shares over time - Ohio and North Carolina



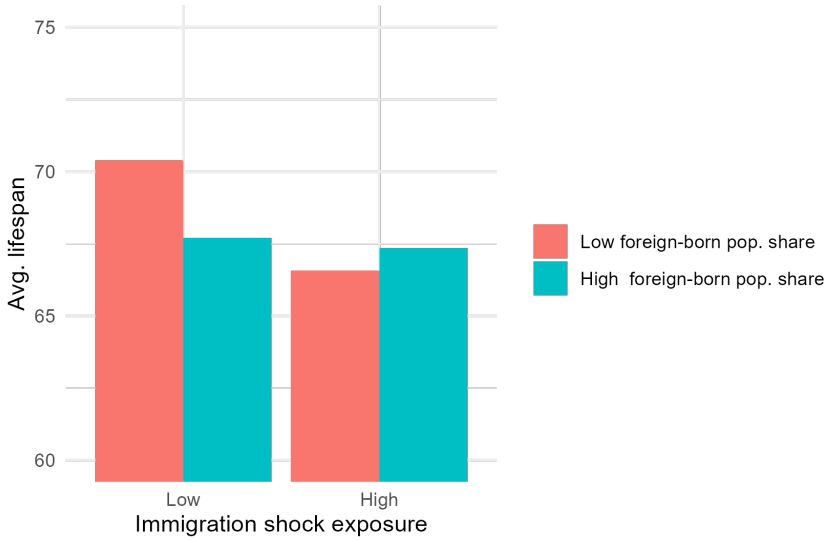
Notes: This figure plots the share of the working-age population of Ohio and North Carolina of different national origins over time. Full count census data is used to produce this figure.

Figure 2: Occupations shares in foreign-born pop. over time



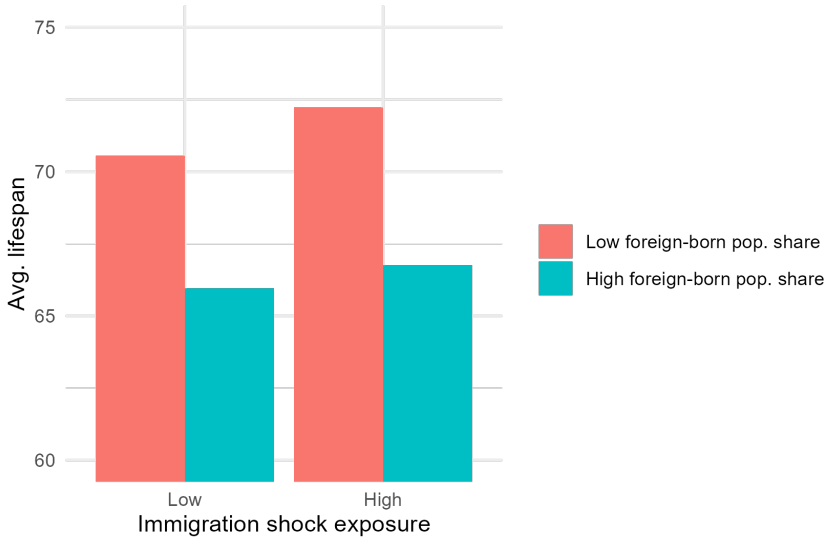
Notes: The figure reports the share of the foreign-born population employed in Ohio and North Carolina by 1-digit census occupation code. Full count census data is used to produce this figure.

Figure 3: Lifespans in exposed and unexposed occupations



Notes: The figure reports average lifespans of native-born working age individuals living in Ohio and North Carolina in 1920 binned along two axes related to individual occupation. The first axis indicates whether an individual's occupation was one with greater or less than median foreign-born employment share across occupations. The second axis indicates whether the occupation had greater or less than median exposure to the quota policies enacted in the 1920s (the specific measure used is the quota exposure measure used in the empirical analysis).

Figure 4: Lifespans in exposed and unexposed areas



Notes: The figure reports average lifespans of native-born working age individuals living in Ohio and North Carolina in 1920 binned along two axes related to their location in . The first axis indicates whether an individual's location was one with greater or less than median foreign-born population share across locations. The second axis indicates whether the occupation had greater or less than median exposure to the quota policies enacted in the 1920s (the specific measure used is the quota exposure measure used in the empirical analysis).

Figure 5: Quota exposures by national origin (Abramitzky et al., 2023)

Country Group	(1) Quota Intensity 1	(2) Quota Intensity 2	(3) Quota Intensity 3	(4) World War I Intensity
A. High-Restriction Countries				
Asia	1	0.912	0.947	0.496
Central Europe	1	0.935	0.968	0.978
Eastern Europe	1	0.830	0.935	0.957
Greece	1	0.935	0.965	0.502
Italy	1	0.921	0.962	0.887
Portugal	1	0.840	0.945	0.411
Rest of World	1	0.640	0.686	0.000
Russia	1	0.834	0.933	0.950
Spain	1	0.867	0.980	0.140
B. Low-Restriction Countries				
Germany	0	0.000	0	0.919
Ireland	0	0.133	0	0.789
Scandinavia	0	0.540	0.100	0.675
United Kingdom	0	0.293	0	0.795
Western Europe	0	0.497	0.559	0.716
C. Non-Restriction Countries				
Canada	0	0	0	0
Caribbean	0	0	0	0.112
Latin America	0	0	0	0
Mexico	0	0	0	0

Notes: Quota exposures from Abramitzky et al. (2023). Measure 1 is an indicator for whether immigrant flows were constrained by imposed quotas. Measures 2 and 3 capture the relative proportion of immigrant flows which were prevented from entry by each national quota; measure 2 assumes constant counterfactual flows in the absence of quota implementation while measure 3 assumes counterfactual flows are estimates from a time trend regression of each national origin.

Table 1: Baseline regressions

	Dep. Var.: Δ Avg. Lifespan			
	QE3		QE1	
	(1)	(2)	(3)	(4)
Policy Exposure	3.48***	−8.67	3.68***	−2.22
	(1.01)	(21.89)	(1.03)	(9.65)
Policy Exposure×(1920 − 1930)	−6.80***	−6.80***	−7.35***	−7.35***
	(1.32)	(1.53)	(1.33)	(1.51)
Foreign Pct×(1920 − 1930)	7.17***	7.17***	7.40***	7.40***
	(2.11)	(1.26)	(2.10)	(1.30)
Mean of Dep. Var.	69.5	69.5	69.5	69.5
Std. Dev. of Dep. Var.	2.6	2.6	2.6	2.6
County FE	No	Yes	No	Yes
State FE	Yes	No	Yes	No
N	364	364	364	364
Adjusted R^2	0.83	0.72	0.83	0.72

Notes: This table presents the coefficients from the DiD specification described in the empirical specification. The dependent variable in these specifications is defined as the change over each decade in the average lifespan of working age individuals living in the county as of the beginning of the decade. Columns 1 and 2 present results with QE3 as the treatment variable and columns 3 and 4 present results with QE1. Every county is observed from 1900 to 1910 and 1920 to 1930. Counties consist of those in Ohio or North Carolina with linked vital statistics records. Observations are weighted by total working age population in 1900. In parentheses, robust standard errors, clustered by county, are reported. Significance levels: 90%, ** 95%, *** 99%.

Table 2: Labor and local market transition

	All	Mover status	Job change status
	(1)	(2)	(3)
Policy Exposure	2.54	−10.38	20.92
	(11.28)	(14.53)	(21.41)
Policy Exposure×(1920 − 1930)	−6.80***	−4.95**	−0.83
	(1.53)	(2.30)	(2.86)
× Mover pop.		−2.36	
		(3.09)	
× Occ. switching pop.			−11.79**
			(4.69)
Mean of Dep. Var.	69.5	71.3 (Stayer) 67.6 (Mover)	72.4 (Non-switcher) 71.1 (Switcher)
Std. Dev. of Dep. Var.	2.6	2.89 (Stayer) 4.23 (Mover)	3.46 (Non-switcher) 2.97 (Switcher)
County FE	Yes	Yes	Yes
N	364	718	716
Adjusted R^2	0.72	0.69	0.21

Notes: This table presents the coefficients from stacked DiD regressions. The dependent variable in these specifications is defined as the change over each decade in the average lifespan of working age individuals living in the county as of the beginning of the decade. Every county is observed from 1900 to 1910 and 1920 to 1930 and across an axis of potential heterogeneous treatment effects. Column 1 corresponds to results from column 2 of the baseline regression results table. Counties consist of those in Ohio or North Carolina with linked vital statistics records. In parentheses, robust standard errors, clustered by county, are reported. Significance levels: * 90%, ** 95%, *** 99%.

Table 3: Heterogeneity analysis

	All	Gender	Age	Urban status
	(1)	(2)	(3)	(4)
Policy Exposure	2.54	7.16	1.02	1.84
	(11.28)	(7.15)	(13.11)	(10.56)
Policy Exposure×(1920 − 1930)	−6.80***	−10.62***	−1.75	−4.04***
	(1.53)	(2.22)	(1.53)	(1.04)
× Male		6.71**		
		(2.71)		
× Age < 35			−14.70***	
			(3.41)	
× Urban area				−5.15
				(3.50)
Mean of Dep. Var.	69.5	70.2 (Women) 71.0 (Men)	71.0 (≥ 35) 70.3 (< 35)	71.4 (Rural) 68.8 (Urban)
Std. Dev. of Dep. Var.	2.6	4.19 (Women) 2.22 (Men)	2.58 (≥ 35) 4.43 (< 35)	2.38 (Rural) 2.35 (Urban)
N	364	728	728	364
Adjusted R^2	0.72	0.74	0.70	0.78

Notes: This table presents the coefficients from stacked DiD regressions. The dependent variable in these specifications is defined as the change over each decade in the average lifespan of working age individuals living in the county as of the beginning of the decade. Every county is observed from 1900 to 1910 and 1920 to 1930 and across an axis of potential heterogeneous treatment effects. Column 1 corresponds to results from column 2 of the baseline regression results table. Counties consist of those in Ohio or North Carolina with linked vital statistics records. In parentheses, robust standard errors, clustered by county, are reported. Significance levels: * 90%, ** 95%, *** 99%.

Table 4: Labor market mechanism (occ. lifespan) regressions

	Dep. Var.: Mean change in average lifespan of employed occupation							
	All		Age ≤ 35		All switchers		Age ≤ 35 switchers	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Policy Exposure	0.09**	2.02	0.10*	3.13	0.13*	2.32	0.13*	3.57
	(0.05)	(1.71)	(0.05)	(2.69)	(0.07)	(1.99)	(0.07)	(3.07)
Policy Exposure$\times(1920-1930)$	−0.08	−0.10**	−0.10	−0.13**	−0.12	−0.14**	−0.14	−0.17**
	(0.07)	(0.05)	(0.07)	(0.06)	(0.11)	(0.07)	(0.10)	(0.07)
Foreign Pct$\times(1920-1930)$	−0.01	0.06	−0.17	−0.10	−0.17	−0.07	−0.34**	−0.25**
	(0.14)	(0.07)	(0.13)	(0.08)	(0.19)	(0.09)	(0.17)	(0.10)
Mean of Dep. Var.	−0.30	−0.30	−0.42	−0.42	−0.39	−0.39	−0.51	−0.51
Std. Dev. of Dep. Var.	0.34	0.34	0.48	0.48	0.41	0.41	0.53	0.53
County FE	No	Yes	No	Yes	No	Yes	No	Yes
State FE	Yes	No	Yes	No	Yes	No	Yes	No
N	5,288	5,288	5,288	5,288	5,284	5,284	5,282	5,282
Adjusted R^2	0.55	0.71	0.65	0.73	0.55	0.71	0.65	0.73

Notes: This table presents the coefficients from the DiD specification described in the empirical specification. Counties consist of those in Ohio or North Carolina with linked vital statistics records. Observations are weighted by total working age population in 1900. In parentheses, robust standard errors, clustered by county, are reported. Significance levels: 90%, ** 95%, *** 99%.

Table 5: Injury rate regressions

	Dep. Var.: Change in injury rate (per 100k hours)			
	All	Age $\leq$ 35	All switchers	Age $\leq$ 35 switchers
	(1)	(2)	(3)	(4)
Policy Exposure	−0.92	−1.23	−1.15	−1.47
	(1.04)	(0.75)	(2.02)	(1.24)
Policy Exposure×(1920 − 1930)	0.13***	0.15**	0.22**	0.24**
	(0.06)	(0.07)	(0.10)	(0.10)
Foreign Pct×(1920 − 1930)	−0.14	−0.13*	−0.19*	−0.15
	(0.06)	(0.07)	(0.10)	(0.10)
Mean of Dep. Var.	−0.001	0.001	0.005	0.007
Std. Dev. of Dep. Var.	0.37	0.42	0.63	0.67
N	5,248	5,224	5,214	5,192
Adjusted R^2	0.61	0.59	0.62	0.60

Notes: This table presents the coefficients from DiD regressions. The dependent variable in these specifications is defined as the change in number of disabling injuries per 100k work hours for the industry of employment of working age individuals living in each county as of the beginning of the decade. Every county is observed from 1900 to 1910 and 1920 to 1930. In parentheses, robust standard errors, clustered by county, are reported. Significance levels: * 90%, ** 95%, *** 99%.

Appendix

A. Figures and tables

Table A.1: QE2 as treatment variable

	Dep. Var.: Δ Avg. Lifespan	
	(1)	(2)
Policy Exposure	3.89^{***}	-3.64
	(1.06)	(13.11)
Policy Exposure $\times(1920-1930)$	-7.98^{***}	-7.98^{***}
	(1.37)	(1.65)
Foreign Pct$\times(1920-1930)$	7.34^{***}	7.34^{***}
	(2.27)	(1.33)
Mean of Dep. Var.	69.5	69.5
Std. Dev. of Dep. Var.	2.6	2.6
County FE	No	Yes
State FE	Yes	No
Observations	364	364
Adjusted R^2	0.83	0.73

Notes: This table presents the coefficients from the DiD specification described in the empirical specification. The dependent variable in these specifications is defined as the change over each decade in the average lifespan of working age individuals living in the county as of the beginning of the decade. Every county is observed from 1900 to 1910 and 1920 to 1930. Counties consist of those in Ohio or North Carolina with linked vital statistics records. In parentheses, robust standard errors, clustered by county, are reported. Significance levels: * 90%, ** 95%, *** 99%.

Table A.2: QE3 - No population weights

	Dep. Var.: Δ Avg. Lifespan	
	(1)	(2)
Policy Exposure	2.40^{***}	-9.76
	(0.82)	(21.57)
Policy Exposure$\times(1920-1930)$	-4.62^{***}	-4.62^{***}
	(1.04)	(1.12)
Foreign Pct$\times(1920-1930)$	9.97^{***}	9.97^{***}
	(2.82)	(2.70)
Mean of Dep. Var.	69.5	69.5
Std. Dev. of Dep. Var.	2.6	2.6
County FE	No	Yes
State FE	Yes	No
N	364	364
Adjusted R^2	0.63	0.45

Notes: This table presents the coefficients from the DiD specification described in the empirical specification. The dependent variable in these specifications is defined as the change over each decade in the average lifespan of working age individuals living in the county as of the beginning of the decade. Every county is observed from 1900 to 1910 and 1920 to 1930. Counties consist of those in Ohio or North Carolina with linked vital statistics records. In parentheses, robust standard errors, clustered by county, are reported. Significance levels: * 90%, ** 95%, *** 99%.

Table A.3: QE1 - No population weights

	Dep. Var.: Δ Avg. Lifespan	
	(1)	(2)
Policy Exposure	2.49^{***}	-3.29
	(0.83)	(9.54)
Policy Exposure$\times(1920-1930)$	-5.20^{***}	-5.20^{***}
	(1.01)	(1.11)
Foreign Pct$\times(1920-1930)$	10.14^{***}	10.14^{***}
	(2.91)	(2.77)
Mean of Dep. Var.	69.5	69.5
Std. Dev. of Dep. Var.	2.6	2.6
County FE	No	Yes
State FE	Yes	No
N	364	364
Adjusted R^2	0.64	0.46

Notes: This table presents the coefficients from the DiD specification described in the empirical specification. The dependent variable in these specifications is defined as the change over each decade in the average lifespan of working age individuals living in the county as of the beginning of the decade. Every county is observed from 1900 to 1910 and 1920 to 1930. Counties consist of those in Ohio or North Carolina with linked vital statistics records. In parentheses, robust standard errors, clustered by county, are reported. Significance levels: * 90%, ** 95%, *** 99%.

Table A.4: QE1 - Incl. 1910 to 1920 obs.

	Dep. Var.: Δ Avg. Lifespan	
	(1)	(2)
Policy Exposure	2.20^{***}	1.00
	(0.65)	(6.50)
Policy Exposure$\times(1920-1930)$	-7.03^{***}	-7.03^{***}
	(1.20)	(1.36)
Foreign Pct$\times(1920-1930)$	7.89^{***}	7.89^{***}
	(1.36)	(1.29)
Mean of Dep. Var.	69.5	69.5
Std. Dev. of Dep. Var.	2.6	2.6
County FE	No	Yes
State FE	Yes	No
N	546	546
Adjusted R^2	0.79	0.70

Notes: This table presents the coefficients from the DiD specification described in the empirical specification. The dependent variable in these specifications is defined as the change over each decade in the average lifespan of working age individuals living in the county as of the beginning of the decade. Every county is observed from 1900 to 1910, 1910 to 1920, and 1920 to 1930. Counties consist of those in Ohio or North Carolina with linked vital statistics records. In parentheses, robust standard errors, clustered by county, are reported. Significance levels: * 90%, ** 95%, *** 99%.

Table A.5: QE3 - Incl. 1910 to 1920 obs.

	Dep. Var.: Δ Avg. Lifespan	
	(1)	(2)
Policy Exposure	2.18^{***}	-0.57
	(0.67)	(14.87)
Policy Exposure$\times(1920-1930)$	-6.54^{***}	-6.54^{***}
	(1.22)	(1.43)
Foreign Pct$\times(1920-1930)$	7.67^{***}	7.67^{***}
	(1.23)	(1.22)
Mean of Dep. Var.	69.5	69.5
Std. Dev. of Dep. Var.	2.6	2.6
County FE	No	Yes
State FE	Yes	No
N	546	546
Adjusted R^2	0.78	0.70

Notes: This table presents the coefficients from the DiD specification described in the empirical specification. The dependent variable in these specifications is defined as the change over each decade in the average lifespan of working age individuals living in the county as of the beginning of the decade. Every county is observed from 1900 to 1910, 1910 to 1920, and 1920 to 1930. Counties consist of those in Ohio or North Carolina with linked vital statistics records. In parentheses, robust standard errors, clustered by county, are reported. Significance levels: * 90%, ** 95%, *** 99%.

Table A.6: QE3 - Outcomes in levels

	Dep. Var.: Avg. Lifespan	
	(1)	(2)
Policy Exposure	2.39*	17.28***
	(1.31)	(5.71)
Policy Exposure×(1930)	−3.85**	−3.85***
	(1.73)	(0.87)
County FE	No	Yes
State FE	Yes	No
Mean of Dep. Var.	69.5	69.5
Std. Dev. of Dep. Var.	2.6	2.6
N	364	364
Adjusted R^2	0.78	0.89

Notes:

Notes: This table presents the coefficients from a DiD specification comparing outcomes in levels. The dependent variable in these specifications is defined as the average lifespan of working age individuals living in the county as of the beginning of the decade. Every county is observed in 1920 and 1930. Counties consist of those in Ohio or North Carolina with linked vital statistics records. In parentheses, robust standard errors, clustered by county, are reported. Significance levels: * 90%, ** 95%, *** 99%.

Table A.7: QE1 - Outcomes in levels

	Dep. Var.: Avg. Lifespan	
	(1)	(2)
Policy Exposure	1.78	16.38***
	(1.21)	(4.93)
Policy Exposure×(1930)	−4.25***	−4.25***
	(1.57)	(0.84)
County FE	No	Yes
State FE	Yes	No
Mean of Dep. Var.	69.5	69.5
Std. Dev. of Dep. Var.	2.6	2.6
N	364	364
Adjusted R^2	0.78	0.89

Notes:

Notes: This table presents the coefficients from a DiD specification comparing outcomes in levels. The dependent variable in these specifications is defined as the average lifespan of working age individuals living in the county as of the beginning of the decade. Every county is observed in 1920 and 1930. Counties consist of those in Ohio or North Carolina with linked vital statistics records. In parentheses, robust standard errors, clustered by county, are reported. Significance levels: * 90%, ** 95%, *** 99%.

Table A.8: Baseline regressions - 1920 shares

	Dep. Var.: Δ Avg. Lifespan			
	QE3		QE1	
	(1)	(2)	(3)	(4)
Policy Exposure	1.07	−2.48	1.07	−1.44
	(0.75)	(7.26)	(0.67)	(5.07)
Policy Exposure×(1920 − 1930)	−2.80***	−2.80**	−2.68***	−2.68***
	(0.98)	(1.17)	(0.85)	(1.04)
Foreign Pct×(1920 − 1930)	7.41***	7.41***	7.51***	7.51***
	(2.37)	(1.43)	(2.35)	(1.43)
Mean of Dep. Var.	69.5	69.5	69.5	69.5
Std. Dev. of Dep. Var.	2.6	2.6	2.6	2.6
County FE	No	Yes	No	Yes
State FE	Yes	No	Yes	No
N	368	368	368	368
Adjusted R^2	0.81	0.69	0.81	0.69

Notes: This table presents the coefficients from the DiD specification described in the empirical specification. The dependent variable in these specifications is defined as the change over each decade in the average lifespan of working age individuals living in the county as of the beginning of the decade. Columns 1 and 2 present results with QE3 as the treatment variable and columns 3 and 4 present results with QE1; for both treatment variables, the share component has been calculated with the base year set to 1920 rather than 1900. Every county is observed from 1900 to 1910 and 1920 to 1930. Counties consist of those in Ohio or North Carolina with linked vital statistics records. Observations are weighted by total working age population in 1900. In parentheses, robust standard errors, clustered by county, are reported. Significance levels: * 90%, ** 95%, *** 99%.

Table A.9: Stacked regressions - 1920 shares

	All	Gender	Age	Urban status
	(1)	(2)	(3)	(4)
Policy Exposure	0.68	3.40	−1.31	1.29
−10.20				
	(10.24)	(6.41)	(12.01)	(9.07)
(12.21)				
Policy Exposure×(1920 − 1930)	−2.80	−4.32***	−0.29	−2.82**
	(1.17)	(1.51)	(1.19)	(1.11)
× Male		3.08*		
		(1.84)		
× ¡ 35 yo			−6.28**	
			(2.93)	
× Urban area				−1.94
				(2.10)
Mean of Dep. Var.	69.5	70.2 (Women) 71.0 (Men)	71.0 (≥ 35) 70.3 (< 35)	71.4 (Rural) 68.8 (Urban)
Std. Dev. of Dep. Var.	2.6	4.19 (Women) 2.22 (Men)	2.58 (≥ 35) 4.43 (< 35)	2.38 (Rural) 2.35 (Urban)
N	368	736	736	368
726				
Adjusted R^2	0.69	0.72	0.67	0.77
0.68				

Notes: This table presents the coefficients from the DiD specification described in the empirical specification. The dependent variable in these specifications is defined as the change over each decade in the average lifespan of working age individuals living in the county as of the beginning of the decade. Columns 1 and 2 present results with QE3 as the treatment variable and columns 3 and 4 present results with QE1; for both treatment variables, the share component has been calculated with the base year set to 1920 rather than 1900. Every county is observed from 1900 to 1910 and 1920 to 1930. Counties consist of those in Ohio or North Carolina with linked vital statistics records. Observations are weighted by total working age population in 1900. In parentheses, robust standard errors, clustered by county, are reported. Significance levels: * 90%, ** 95%, *** 99%.

Table A.10: Baseline regressions with Abramitzky et al. (2023) LASSO covariates

	Dep. Var.: Δ Avg. Lifespan			
	QE3		QE1	
	(1)	(2)	(3)	(4)
Policy Exposure×(1920 − 1930)	−4.35***	−4.35***	−5.41***	−5.41***
	(1.15)	(1.31)	(1.19)	(1.35)
Foreign Pct×(1920 − 1930)	−13.24***	−13.24***	−13.21***	−13.21***
	(1.28)	(1.29)	(1.26)	(1.25)
Mean of Dep. Var.	69.5	69.5	69.5	69.5
Std. Dev. of Dep. Var.	2.6	2.6	2.6	2.6
County FE	No	Yes	No	Yes
State FE	Yes	No	Yes	No
N	364	364	364	364
Adjusted R^2	0.87	0.81	0.87	0.81

Notes: This table presents the coefficients from the DiD specification described in the empirical specification. The dependent variable in these specifications is defined as the change over each decade in the average lifespan of working age individuals living in the county as of the beginning of the decade. Columns 1 and 2 present results with QE3 as the treatment variable and columns 3 and 4 present results with QE1. In addition to pre-treatment foreign-born population share, covariates also include pre-treatment agricultural employment share. Every county is observed from 1900 to 1910 and 1920 to 1930. Counties consist of those in Ohio or North Carolina with linked vital statistics records. Observations are weighted by total working age population in 1900. In parentheses, robust standard errors, clustered by county, are reported. Significance levels: * 90%, ** 95%, *** 99%.

Table A.11: Baseline regressions with LASSO covariates

	Dep. Var.: Δ Avg. Lifespan			
	QE3		QE1	
	(1)	(2)	(3)	(4)
Policy Exposure$\times(1920-1930)$	-6.46^{**}	-6.46^{***}	-6.87^{***}	-6.87^{***}
	(2.53)	(2.10)	(2.34)	(2.00)
Foreign Pct$\times(1920-1930)$	13.83	13.83	12.92	12.92
	(20.32)	(15.07)	(20.33)	(15.13)
Mean of Dep. Var.	69.5	69.5	69.5	69.5
Std. Dev. of Dep. Var.	2.6	2.6	2.6	2.6
County FE	No	Yes	No	Yes
State FE	Yes	No	Yes	No
N	364	364	364	364
Adjusted R^2	0.83	0.73	0.84	0.74

Notes: This table presents the coefficients from the DiD specification described in the empirical specification. The dependent variable in these specifications is defined as the change over each decade in the average lifespan of working age individuals living in the county as of the beginning of the decade. Columns 1 and 2 present results with QE3 as the treatment variable and columns 3 and 4 present results with QE1. In addition to pre-treatment foreign-born population share, covariates also include pre-treatment gender ratio. Every county is observed from 1900 to 1910 and 1920 to 1930. Counties consist of those in Ohio or North Carolina with linked vital statistics records. Observations are weighted by total working age population in 1900. In parentheses, robust standard errors, clustered by county, are reported. Significance levels: * 90%, ** 95%, *** 99%.