The Hidden Costs of Mass Deportations: Early-Life Mortality and the 1930s Mexican Repatriation

Sergio Eduardo López Araiza Bejar* Jesse McDevitt-Irwin†

October 2026

PLEASE CLICK FOR AN UPDATED VERSION

Abstract

Nativism and anti-immigrant sentiment have been a common thread throughout US history, leading to exclusionary policies and deportation campaigns designed to expel immigrant groups. From 1928 to 1935, the US government implemented a deportation campaign based on fear and pushed between 330,000 and one million Mexicans to leave the country. We study the effect of this campaign on early-life mortality of children born to Mexican-immigrant families that remained in the US. In order to study the impact of this repatriation campaign on health, we overcome the lack of disaggregated stillbirth and infant mortality data by leveraging the fact that, in utero and in infancy, males are more frail than females. We proxy early-life mortality using childhood sex ratios and leverage the rollout of the campaign, which started in Texas in 1928 and didn't move to the rest of the country until 1931. Using a triple-differences estimator, we find a large increase in the childhood sex ratio for Mexican-immigrant households in Texas in 1930, equivalent to moving from the early-life mortality rates of the white population in 1930 to those of the Black population. Finally, we find evidence that the deportation campaign led to lower healthcare utilization and public services utilization more broadly.

1 Introduction

Nativism and anti-immigrant sentiment have been a common thread throughout US history (Abramitzky and Boustan, 2017), leading to exclusionary policies and deportation campaigns designed to deport and expel immigrant groups. There is still limited evidence on the impact that

*Corresponding author. Northwestern University Kellogg: sergio.lopezaraiza@kellogg.northwestern.edu
†University of Chicago: jrmcirwin@uchicago.edu.

these policies have on the health of the children of the targeted population. In this paper, we provide new evidence on the health consequences of one of the largest removal campaigns in US history.

Prior to and during the Great Depression, the US government pushed between 330,000 and one million Mexicans to leave the country. Lack of state capacity led the immigration authorities to design a 'repatriation' campaign based on creating widespread terror among the Mexican community. Beginning in 1928 in Texas and spreading to the rest of the country by 1931, the campaign had enormous impacts on the lives of Mexican immigrants. The threat of immigration raids that ended up with Mexican immigrants being herded into trains was constant and a "climate of fear" descended over the Mexican community (Goodman, 2020, p. 172). We focus on the start of the campaign between 1928 and 1930, which took place in Texas before moving to the rest of the country in 1931.

We study the effect of this campaign on the early-life mortality—which consists of miscarriages, stillbirths, and infant deaths—of children of Mexican-immigrant families that remained in the US. The primary challenge is the lack of detailed and trustworthy statistics for the Mexican-immigrant population. To overcome this challenge, we proxy early-life mortality using childhood sex ratios. Childhood sex ratios leverage the biological frailty of males during both gestation and infancy and thus reflect the combined effect of miscarriages, stillbirths, and infant mortality. We calculate childhood sex ratios using census data, which also allows us to identify the immigration status of the household.

The use of childhood sex ratios as a proxy for early-life mortality is increasingly recognized in economics and related fields (Sanders and Stoecker, 2015; Valente, 2015; Guimbeau et al., 2022; McDevitt-Irwin and Irwin, 2025; Castillo and Dribe, 2026) and is well supported by the data in our setting. Infant mortality statistics from the 1920s reflect a propensity of stillbirth around 20% higher for males than for females. The same is true for infant mortality: males are around 20% more likely to die than females before age 1. These empirical patterns are consistent with the medical and demographic literature that document male frailty in early life (Waldron, 1998; Catalano et al., 2005; Loomans et al., 2013).

To estimate causal effects, we use the fact that the campaign targeted Mexicans and compare outcomes of Mexican-immigrant households to those of US-born white households, before and after the campaign started. This comparison allows us to study whether the health outcomes for children of Mexican-immigrant households changed trajectories after the campaign started. Importantly, because the health of Mexican immigrants could have been affected by other factors after 1920, we incorporate a third difference to the research design. We consider Mexican-immigrant households outside of Texas, who were not affected by the campaign until 1931. Including this group in the empirical strategy allows us to account for other sources of difference in health between Mexican-immigrant and white US-born households. The identification assumption in this triple-differences model is that, in the absence of the onset of the Mexican Repatriation Campaign, the difference in

health outcomes between children of Mexican-immigrant households and white US-born households would have evolved similarly for families in Texas than in the other US-Mexico border states. Our results are robust to using difference-in-differences designs, leveraging within-Mexican, cross-state variation as well as within-Texas, cross-ethnicity variation.

We find a large increase in the childhood sex ratio for Mexican-immigrant households in Texas in 1930. Our logistic regression estimation renders a $\sim$5 point shift in the log under-3 sex ratio, which is roughly equivalent to a 5% shift towards females. These are large effects, as the biological healthy sex ratio at birth is typically around 5% more male than female. This difference is close in magnitude to the sex ratio gap between Black and white American families in 1930. Moreover, a back of the envelope calculation suggests an increase in the combined stillbirth and infant mortality rates[1] of roughly 13.8 percentage points.

We find similar patterns in the 1940 census. The campaign ended in 1935, so we should find a similar shift toward females in the cohorts who are aged 5 to 9 in 1940, but not those who are under age 5 in 1940. However, since the campaign moved to the rest of the country in 1931, a cross-state comparison is no longer feasible to extract causal effects. Therefore, we run a within-Texas differences-in-difference design to compare how the childhood sex ratios changed for cohorts age 0 to 3 and cohorts age 5 to 9 between children of Mexican-immigrant families and white US-born families. We find that in the 1930 census, ages 3 and under see a shift towards females in their sex ratio, but ages 5-9 do not. On the other hand, children aged 5-9 see an increase in the childhood sex ratio of a similar magnitude in 1940, while those aged 0-3 do not. This pattern is consistent with the campaign increasing early-life mortality for children in Mexican-immigrant families.

We test the robustness of our results in multiple ways. First, we make sure that the detrimental health effects are concentrated among the cohort exposed to the campaign while in gestation or in infancy and not in older cohorts. Even if the campaign increased mortality for older kids, we wouldn't expect to find a change in the childhood sex ratios given that the male frailty channel is mainly present before age one. As expected, we don't find any changes in the sex ratios for older cohorts.

Next, we make sure that the increases that we find in the childhood sex ratios of Mexican-immigrant households are not driven by a broader change affecting minorities. We do so by running placebo regressions for households where the household head is either a non-Mexican immigrant or a Black American. Contrary to what we find for children in Mexican-immigrant households, we estimate a negative triple interaction coefficient for all these groups, significant at the 10% level for children of foreign-born household heads.

To check that our results reflect a rise in early-life mortality and not just selective out-migration

[1] That is , early-life mortality, net of miscarriages

patterns[2], we look at the number of children in the household. This is not our primary outcome of interest because it compounds early-life mortality with fertility decisions, to which the childhood sex ratio is agnostic to. We find that the number of children decreases among Mexican-immigrant households in Texas in 1930. The effect is smaller than what our sex ratio results imply in terms of early-life mortality, likely reflecting differential changes in fertility decisions by Mexican and non-Mexican households across the period[3].

Finally, we look at household characteristics to provide evidence against our results being driven by selective out-migration. Given that the campaign was designed to drive Mexican immigrants out of the United States, it could be the case that better-off Mexican-immigrant households moved out of Texas with a higher frequency. If this were the case, our results could be driven by sample composition rather than by the campaign directly affecting the targeted population's health. We find that Mexican households in Texas in 1930 actually tend to have higher occupational scores and are less crowded. Thus, we don't find any evidence of selective out-migration driving our results.

After looking at the health impacts of the Mexican Repatriation Campaign, we provide evidence consistent with Mexican-immigrant families retreating from public places. This would have been a side effect of the campaign, which was designed to create a 'climate of fear' among Mexican immigrants to induce self-deportation (Goodman, 2020). Aside from convincing Mexican immigrants to leave the country, this would have also led Mexican immigrants to avoid public spaces. To look into this mechanism, we develop a new dataset of birth registration completeness in Texas. We show that birth registration declined among Texas Mexicans upon the onset of the deportation campaign. This suggests a decline in healthcare utilization, as doctors and nurses were required to register births, which made medical assistance at birth an important correlate of registration (Eriksson et al., 2018). Prior literature has found that a decrease in the supply of healthcare in a similar context increased infant mortality (Liebert and Mäder, 2026), making this a likely driver of the increase in early-life mortality that we document. More broadly, this is indicative of a withdrawal from public services.

This paper contributes to several strands of literature. First, it contributes to the broad literature on the effects of deportations and exclusion campaigns in the US. These studies have looked at both contemporary (Amuedo-Dorantes et al., 2022; Vu, 2024) and historical (Long et al., 2026; Chen and Xie, 2024) settings, including this specific deportation campaign (Lee et al., 2022; Cortes and Sant'Anna, 2024). Overall, these papers have found adverse effects for the labor market outcomes

[2]Even if the historical literature doesn't mention it, one could be worried that families with Mexican children could be more likely to either move out of the country, move their male children out of the household, or hide them from the enumerator as a result of the campaign.

[3]It is well documented that the 1920s and 1930s are decades in which fertility rates are in decline for white households. Also, an increase in early-life mortality could lead to a compensating increase in fertility among Mexican households

of natives or backlash from the targeted population.

Ours is the first paper looking at early-life mortality among the targeted population of a fear based deportation campaign. The closest literature are previous papers that studied the effects of modern immigration enforcement on birth-weight (Amuedo-Dorantes et al., 2022; Vu, 2024). However, their setting is quite different from ours in terms of how immigration enforcement looked like. While they are looking at increased immigration enforcement, we focus on the onset of a campaign designed to induce fear in the targeted community to induce self deportation. Thus, the chilling effects found when examining these modern policies could potentially be even larger in a setting similar to ours.

We also provide new evidence of Mexican population health in the United States, which is largely unknown prior to 1950[4], and show how it was shaped by immigration policy. While previous work on racial health gaps has focused on outcomes for Black and white Americans (Eriksson et al., 2018; Feigenbaum et al., 2019), the health gaps for Mexican immigrants are less well known due to data constraints: Mexicans were classified as 'whites' and were therefore not separated out in published vital statistics reports. Because our proxy measure for health comes directly from Census data, we can identify Mexican immigrants whereas previous research could not.

Finally, we also make a contribution to the historical literature looking at the Mexican Repatriation Campaign. The distinction between voluntary and coerced repatriation of this campaign has long been debated in the historical literature. While historians agree that the bulk of Mexicans and Mexican-Americans returned to Mexico somewhat voluntarily, most authors argue that the repatriation campaign was so violent and widely publicized that they were coerced into leaving the US to avoid going through the official deportation procedure (Alanís Enciso, 2017; Balderrama and Rodríguez, 2006; Goodman, 2020; Lytle-Hernández, 2010). Some historians even go as far as calling it a policy intended to banish the Mexican and Mexican-American population from the United States (Ramírez, 2025). However, others have viewed the massive Mexican and Mexican-American exodus of the 1930s as largely voluntary and not driven by coercion or intimidation (Gratton and Merchant, 2013). Our results document the detrimental health effects endured by the Mexican population that stayed in the US, supporting the coercion interpretation.

The rest of the paper proceeds as follows. Section 2 describes the historical context, from the trends in Mexican immigration prior to the Mexican Repatriation Campaign. Section 3 details the relationship of childhood sex ratios with fetal-infant loss, and why they make a good proxy. Section 4 describes our data sources. Section 5 shows our empirical approach and results, while section 6 shows our robustness tests. Section 7 explores the healthcare and public space avoidance

[4]While labor assimilation outcomes and selection for Mexican immigrants have been previously studied (e.g. Escamilla-Guerrero et al., 2021; Escamilla-Guerrero and López-Alonso, 2023), literature on health outcomes has been scarcer due to lack of data availability

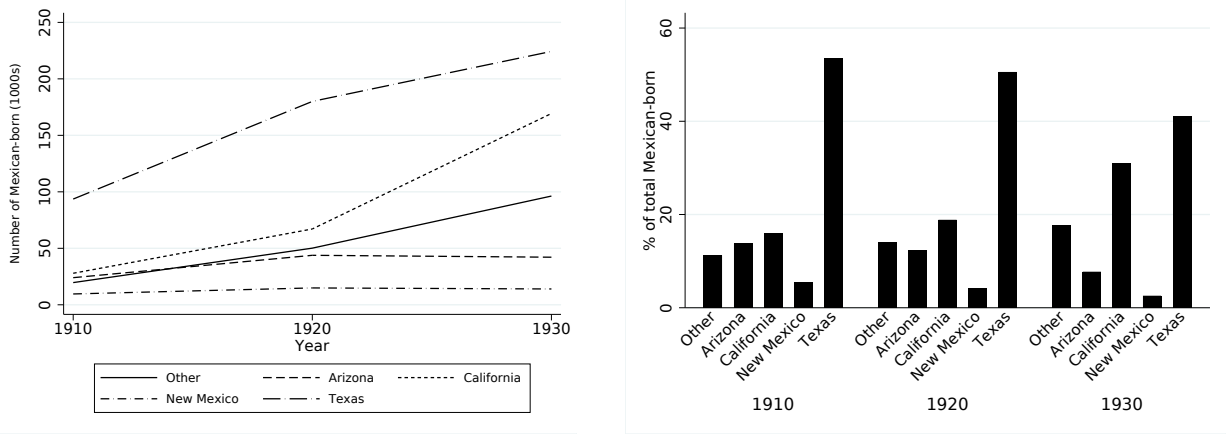
mechanism, and section 8 concludes.

2 Historical Context

2.1 Mexican Immigration to the US and anti-Mexican Sentiment

Mexican immigration to the US was relatively low in the 19th century. The vast majority of Mexican-Americans during this period were the around 100 000 Mexicans who lived in the US Southwest when it still belonged to Mexico (Lopez Araiza, 2026). Immigration to the US during the 19th century mostly came from Europe and Asia (Abramitzky and Boustan, 2017).

Mexican immigration to the US took off after 1900: between 1900 and 1910 the Mexican-immigrant population in the US almost doubled. Mexican immigration continued to increase into the 1920s, driven by the destruction of the Mexican Revolution (1910-1920) and the demand for agricultural laborers in the Southwest. Mexican immigration continued until the start of the Great Depression. By 1930, Mexicans represented 4.5% of the US foreign born population, and were overwhelmingly concentrated in the US-Mexico border states. In Figure 1, we plot the distribution of Mexican immigrants in the US, seeing that Texas was the largest population, especially up until 1930.



(a) Total Mexican-born population

(b) Percentage of Mexican born population in state

Figure 1: Distribution of Mexicans in the US over time

Notes: This figure depicts the Mexican-born presence in the US throughout our period. Panel A shows the total number of Mexican born individuals over 18 years of age in each state. We present the states in the US-Mexico border, which is where our analysis focuses on, and the rest of the US. Panel B shows the percentage of the total Mexican-born individuals in the US over the age of 18 were enumerated in each state.

Texas was one of the main destinations for Mexican migrants for multiple reasons: it was fairly well connected to Mexico via the railroad, there was an agricultural boom, and Mexican presence

had been established for centuries. The Mexican population in Texas greatly increased within a few decades: in 1890 there were roughly 50,000 Mexican immigrants living in Texas, by 1920 there were over 250,000, representing over half of all the Mexican immigrants in the US (McKay, 1985, table 2).

This sudden inflow of Mexican migrants created a growing anti-Mexican sentiment in the Southwest. Signs stating "No dogs, negroes, or Mexicans allowed" were common, particularly in Texas (Balderrama and Rodríguez, 2006, p. 42). Moreover, mainstream eugenics views thought of Mexicans as inferior to whites, blaming them for poor county level health indicators and accusing them of disproportionately making use of health resources (Molina, 2006). Mexicans had generally been classified as 'white' by the US government up until this point, when a rise in anti-Mexican sentiment spurred the creation of a new 'Mexican' race category in the 1930 census[5].

2.2 Immigration Laws at the Turn of the Century

The massive wave of Mexican immigration to the US coincided with a broader rise in anti-immigrant sentiment and policy. The 1920s was a decade of radical change in US immigration policy and enforcement. In the previous decades, there were very few restrictions on immigration to the US from Europe and Latin America. European immigration was severely limited by the 1921 immigration act (amended in 1924), which both fixed the total number of immigrants and established quotas in an effort to reduce immigration from Southern and Eastern Europe (Abramitzky and Boustan, 2017, 2022).

Immigration from Mexico, however, was largely unaffected by these laws and continued to grow, eventually leading to political backlash. A series of bills were introduced to congress in the 1920s which would have restricted Mexican immigration. The most stark example of these bills was sponsored by the Texas representative John Box in 1926—popularly known as the "Box Bill"—and it intended to set quotas specifically for Mexicans, as the 1921 and 1924 acts did for Europeans. After multiple debates, these initiatives failed due to the interest in agribusiness owners to keep their access to low-wage Mexican labor (Alanís Enciso, 2017).

In 1929, the first major law affecting Mexican immigration was passed: the Undesirable Aliens Act (Ramírez, 2025; Weber, 2013; Alanís Enciso, 2017). This law made it a deportable offense to enter the United States through any means that was not an official port of entry. While not retroactive in principle, *de facto* immigrants needed to be able to show that they either entered the US prior to 1929 or, if they did so after, that they came through an official port of entry (Ramírez, 2025; Goodman, 2020). Moreover, even though the law didn't mention Mexicans specifically as a targeted group, they were the only immigrants that were able to enter the US by crossing the almost

[5]The "Mexican" race category disappeared after the 1930 census.

2000-mile long border at any point, which allowed them to avoid going through official entry points and paying the entry taxes. The fact that this Act disproportionately affected Mexican immigrants was not lost to congressmen and senators, which is evidenced by the fact that "no other race or nationality was mentioned nearly as often" in the debates (Fish, 2022, p. 1088)[6].

The fact that Mexican immigrants were particularly vulnerable to this new law wasn't lost to them either. Newspapers in Mexico and the United States that catered to the Mexican community were very alarmed by the passage of the Undesirable Aliens Act and the discussion of the "Box Bill", sometimes conflating both (Alanís Enciso, 2020). A contemporary immigration official noted that these policy changes "put the fear of God in their [Mexican immigrants'] hearts." (cited in Goodman, 2020, p. 41).

2.3 The Mexican Repatriation Campaign

The US government started a deportation campaign targeting Mexicans in 1928. It justified the removal campaign on the grounds that it would improve the lives of Americans (Balderrama and Rodríguez, 2006). This rationale is made explicit in the Department of Labor's 1931 *Annual Report*, where they say that "the unlimited immigration of the past is in a large measure responsible for the oversupply of some classes of labor". As unemployment soared during the Great Depression, government officials argued that the removal of Mexican immigrants and Mexican-Americans would free up jobs for US-born whites. The government claimed that removing 'aliens' would leave more welfare resources for the rest of the population.

However, the federal government lacked the state capacity to actually arrest and deport hundreds of thousands of Mexicans. Instead, they created a hostile climate for Mexican immigrants to pressure them to return to Mexico: immigration raids were highly publicized and featured in newspapers in order to maximize fear (Goodman, 2020). Contemporary public officials who promoted the campaign noted that a well-publicized immigration raid could "scare" many immigrants into fleeing. Goodman (2020) cites a contemporary commentator saying that "alien hunting[...] became a gladiatorial spectacle in which agents raided homes, churches, picket lines, public spaces, bars, dance halls, and pool halls, sometimes without a warrant". Emory Bogardus, a contemporary sociologist, stated that "it takes only an insinuation from a welfare official in the United States to create widespread fear among Mexican immigrants" (Bogardus, 1933).[7]

Prior to eventually covering the entire country, the repatriation campaign began in Texas. Texas was the home of both the largest Mexican population in the United States and of multiple politicians

[6]For example, Congressman Thomas Blanton of Texas said "There were thousands that came across the border unlawfully. You can go up and down the Mexican border in Texas for 10 and 15 miles and you will not see a house, and sometimes you will not see an immigration agent for 40 miles. They can wade across the river in lots of places. They come into Texas by hordes all of the time" when arguing in favor of the act (cited in Fish, 2022).

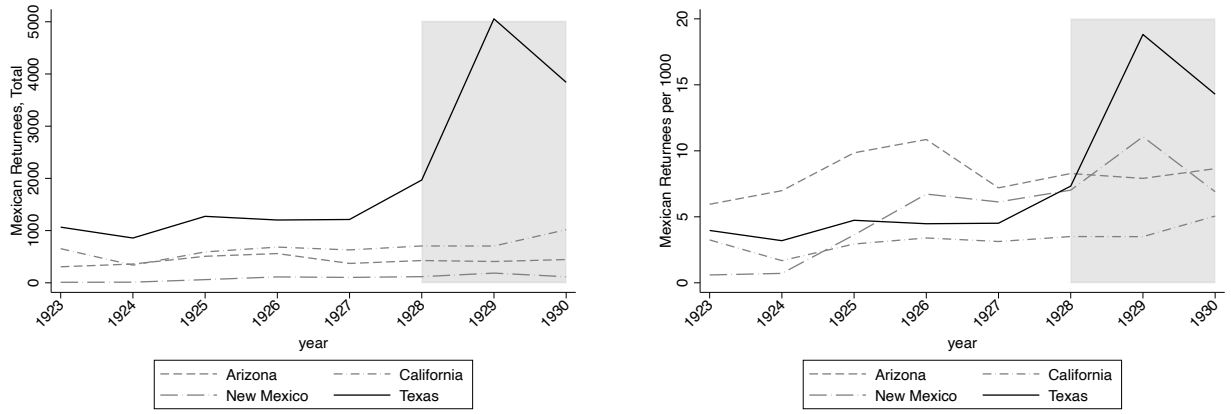
[7]We found this quote through Goodman (2020, p. 45).

that pushed for the passage of the "Box Bill". Both factors came together to create a coalition of federal and local politicians willing to increase the pressure on Mexicans.

In the summer of 1928 the Department of Labor's Bureau of Immigration started an intense deportation campaign targeting Mexicans in the Lower Rio Grande Valley. By 1930 it had extended across the state. Many thousands of Mexican immigrants had been formally deported and many more had chosen to repatriate (McKay, 1985; Weber, 2013). Contemporary newspapers reported widespread fear among Texas Mexicans, including those who had resided in the state for decades (McKay, 1985).

Using data from the annual reports of the Commissioner General of Immigration to the Secretary of Labor from 1923 to 1930, figure 2 shows how the onset of the campaign was concentrated in Texas. While these numbers represent registered departures—which more than 80% of the time meant deportations—and are unable to capture 'voluntary repatriations', they reflect the intensity of federal immigration raids and enforcement. Panel A of figure 2 shows the total number of Mexican returnees in the four states of the US-Mexico border. There is a clear jump in the number of returnees from 1927 to 1928 and then from 1928 to 1929, when the deportation campaign accelerated thanks to the passage of the Undesirable Aliens Act. Importantly, no other state displays this pattern prior to 1931. Panel B takes into account that the baseline number of Mexicans in each state is different, so it shows the number of Mexican returnees per 1000 Mexicans in the state—according to the 1930 census.[8] While Arizona had a higher returnee propensity prior to 1928, we see this shift in 1929 given that Mexicans in Texas see a peak in their propensity to return to Mexico.

[8]We use the number of Mexicans in the 1930 census because we are particularly interested in the years of the deportation campaign (1928-1930). The patterns look very similar if we instead use the Mexican-immigrant population in the 1920 census.



(a) Total returnees

(b) Returnees per 1,000 Mexicans

Figure 2: Mexican returnees by state 1923-1930

Notes: This figure presents the official number of Mexican departures, as reported in the annual reports of the Commissioner General of Immigration to the Secretary of Labor from 1923 to 1930. While some of these departures may have been voluntary, for this time period around 80% of them represented deportations. Panel A shows the total number of returnees for the four states in the US-Mexico border. Panel B calculates the total number of returnees per 1000 Mexicans in the state using the 1930 census. The shaded area from 1928 to 1930 represents the period of the Mexican repatriation campaign in Texas.

In 1931 the repatriation campaign spread beyond Texas to the rest of the country. Most studies of the campaign (e.g. Balderrama and Rodríguez, 2006) tend to focus in California, where the most famous raids took place (e.g. 'La Placita' raid in February of 1931, the first one reported outside of Texas). However, accounts of Mexicans being loaded into trains to be taken back to Mexico can be found across the country, from New York and Boston to Detroit and Chicago.

Most of the so-called 'voluntary repatriation' therefore reflected the coercive pressure of the federal, state, and local governments working together to push Mexican immigrants to leave the country. Estimates of the number of Mexican immigrants and Mexican-Americans pushed to leave the US during this time period range anywhere from 330,000 to one million (Gratton and Merchant, 2013). Even the lower bound of these estimates represents around the third of the Mexican and Mexican-American population in the US at the time (Lee et al., 2022)[9]. Today, the campaign is widely considered a black mark on US history, with the California government issuing an official apology in 2006 (NPR, 2006).

[9]To put these numbers in perspective, around 410,000 people were deported in 2012.

3 Childhood Sex Ratios (CSR) and Fetal-Infant Health

Individual vital statistics are largely unavailable for the US prior to the mid-20th century, making it difficult to measure the effect of the deportation campaign on health. Census microdata contain key variables like ethnicity, nativity, location and occupation, but lack health outcomes. To solve this problem, we turn to childhood sex ratios—-easily available from census microdata—as indicators of infant mortality and fetal loss.[10]

In early life, males are more frail than females.[11] This includes gestation—where male fetuses are more likely to be miscarried—, birth—where males are more likely to be stillborn—and in infancy, when males are more susceptible to common infant causes of death (Waldron, 1998).

Several factors contribute to excess male stillbirth and infant mortality. One is the lack of a redundant X chromosome, which makes males more likely to suffer congenital birth defects. Males also grow at a faster rate *in utero* and during infancy, placing higher metabolic demands on the mother. Male fetuses are therefore more susceptible to under-nutrition. Males are also larger, making them more susceptible to birth trauma and anoxia. Males are also more likely to suffer preterm birth and low-weight-for-gestational age. In infancy, males are also more susceptible to infectious disease, having a weaker immune response than females. In the US between 1910 and 1930, male infants experienced more than 20% higher rates of stillbirth and infant mortality than females (U.S. Department of Commerce, Bureau of the Census, 1934).[12]

Taking these findings, we follow McDevitt-Irwin and Irwin (2025) and define under-3 childhood sex ratios as $CSR = log\left(\frac{female}{male}\right)$. Given that males are more prone to die than females, a higher sex ratio—that is, a sex ratio more female skewed—should signal worse health outcomes overall. Other work in economics have followed a similar path and have used sex ratios as indicators of health (Sanders and Stoecker, 2015; Valente, 2015; Guimbeau et al., 2022).

We verify the empirical relationship of sex ratios and infant mortality in our setting using US data for 1910–1940. We pair infant mortality rates from vital statistics to sex ratios from the decennial census. This does not capture the effect of stillbirths, but allows us to look at a much wider range of mortality rates, as stillbirths are generally not available in the US data prior to 1922. We draw on McDevitt-Irwin and Irwin (2025), plotting the under-5 sex ratio against the average of infant mortality in the prior 5 years. We do this for a number of US states and cities which had data from 1900 to 1940. We see in Figure 3 that infant mortality is clearly closely related to childhood

[10] For the most similar application of sex ratios to our own, see McDevitt-Irwin and Irwin (2025) and Castillo and Dribe (2026).

[11] For a review, see (Waldron, 1998, 64-83). This has been known for centuries; see, for example, Struyck (1740), Wargentin (1755), and Clarke (1786). For discussion see Théré and Rohrbasser (2006).

[12] This is similar across years and across races. On the national level, from 1918-1931 excess male mortality across stillbirths and infancy ranged from 21% to 23%. In 1931, the excess male mortality among racial groups ranged from 18% for 'other colored' to 21% for Blacks and 23% for whites.

sex ratios in the census. Moreover, the slope of the line is .254, very similar to what we get from the theoretical predictions of the vital statistics for the contemporary US. We model this linear relationship in appendix A1.

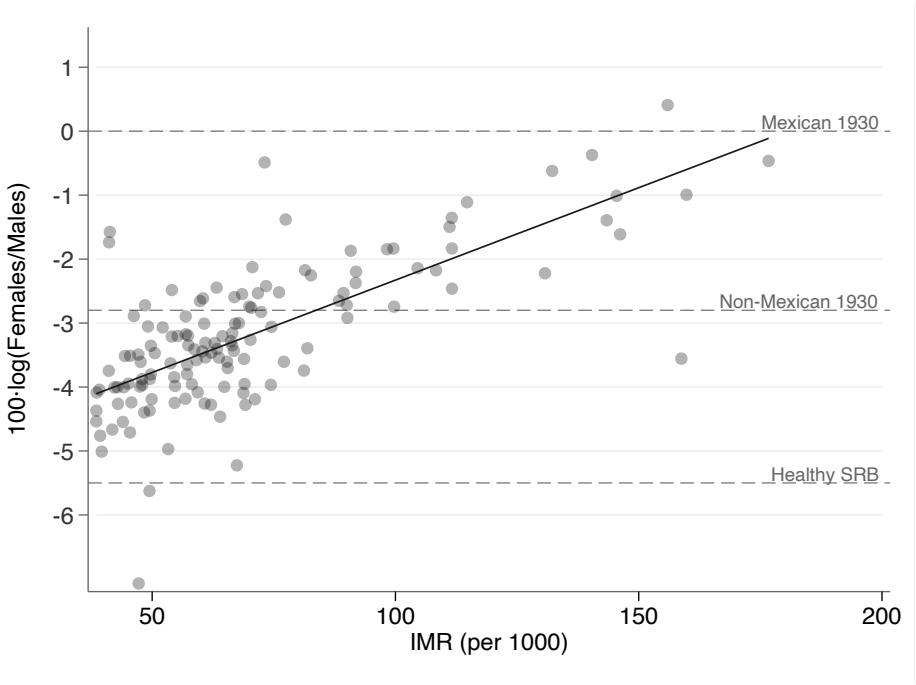


Figure 3: State- and city-level log childhood sex ratios (under-5) against infant mortality, United States, 1900-1940

Notes: State- and city-level log childhood sex ratios (100 × log(females/males), under-5, decennial census) against infant mortality (per 1,000, vital statistics), United States, 1900-1940. Solid line is the OLS fit (slope = 0.254, R^2 = 0.592); dashed lines mark the 1930 Texas Mexican-origin and non-Mexican under-5 sex ratios. Data from McDevitt-Irwin and Irwin (2025), who get the sex ratios from census data and the infant mortality rates from vital statistics reports like U.S. Department of Commerce, Bureau of the Census (1934). We mark 5.5 male as a 'healthy' sex ratio at birth, reflecting the biological norm in healthy populations.

Regarding the Mexican Repatriation Campaign, we see two major channels through which it could have increased miscarriages, stillbirths and infant mortality, leading to female skewed sex ratios. These are access to medical care and maternal stress.

On the former, the deportation campaign likely discouraged Mexican mothers from accessing medical care. Doctors and nurses played a crucial role by providing important information about infant feeding practices and hygiene, both of which were vital to reducing infectious disease risk (Wolf, 2001). The importance of doctors is hard to overstate: Liebert and Mäder (2026) show that a decline in the supply of medical care led to an increase in infant mortality during the 1930s in Germany. Liebert and Mäder (2026) find that these increases in the infant mortality rate were driven by inflammatory bowel disease (colitis, enteritis, diarrhea), stillbirths, and respiratory infections and

other infectious diseases (influenza, pneumonia, measles). All of these are male-biased forms of infant mortality, so their increase would leave relatively more females in the surviving population, shifting the childhood sex ratio.

Second, negative shocks to maternal health, especially stress, lead to a female-skewed sex ratio at birth due to an increase in miscarriages, which are disproportionately male (Catalano et al., 2005). Maternal stress also leads to worsened birth outcomes, especially preterm birth, which would have led to higher rates of stillbirth and infant mortality (Loomans et al., 2013). Therefore, increased psychological stress induced by the campaign would have led directly to a female-skewed sex ratio at birth, as well as through increasing infant mortality.

In sum, males experience higher rates of fetal loss, stillbirths, and infant deaths. The corollary which we exploit is that higher rates of fetal and infant loss skew the surviving population towards females. This pattern is clearly visible in high-mortality historical populations, where females often outnumbered males, compared to today, where there are typically 5-6% more males than females in the childhood population (McDevitt-Irwin and Irwin, 2025). This pattern is also clear from the sex ratio at birth.

4 Data

4.1 Full Count Census

For our main analysis, we use the United States Full Count Census microdata for 1910, 1920, and 1930[13] from IPUMS (Ruggles et al., 2024). These data allow us to identify children of household heads that are 3 years old or younger for every census year in the states of the US-Mexico border: Arizona (AZ), California (CA), New Mexico (NM), and Texas (TX). For each of these children, we identify whether the household head was born in Mexico and, therefore, was targeted by the deportation campaign. Given that most of the children in our sample were born in the United States—over 90% of children of Mexican born household heads were born in the US—we use the household head place of birth as our main variable to identify children of Mexican immigrants.[14]

We focus on households headed by Mexican immigrants because they would have been the most vulnerable to the deportation campaign, and can be defined unambiguously. We therefore exclude

[13]We don't use the 1940 census in the main analysis for two main reasons. The first one is that, prior to the 1930s, the rolout of the campaign was clearly delimited to Texas. Thus, focusing on variation up until 1930 allows us to leverage the rollout to get cross-state variation. Second, the children that would have been affected by the campaign in the 1940 census would be between 5 and 9 years old, which raises some issues of potential selective migration. Nevertheless, we explore the effect of the repatriation campaign in the 1940 Census in section 5.4.

[14]We focus on Mexican immigrants as opposed to the larger Mexican-American community because they were the main target of the deportation campaign in Texas. Mexican-Americans started being targeted by immigration authorities when the campaign extended to the rest of the country.

Mexican-American households—that is, households for which the household head is of Mexican ancestry but was born in the US. The control group is composed of US-born white households.

The full count census also allows us to know other household characteristics, such as its location, the occupation, literacy, and age of the household head and spouse, and the number of families living in the household. We use some of these characteristics—mainly the household head and spouse literacy and age—to create a more suitable comparison, as both parental age and literacy tend to be correlated with children's health outcomes.

4.2 Other Data Sources

We complement the full count census data with other data sources. The first one is the Texas birth indexes from 1926 to 1930. These indexes were prepared by the Texas Human Health Services for genealogical purposes and contain, for each registered birth, the name and last name of the child and parents, as well as the birth date and birth county. We use surname classification to determine whether a should be considered Mexican, as these data do not include ethnicity or parents' birthplace.

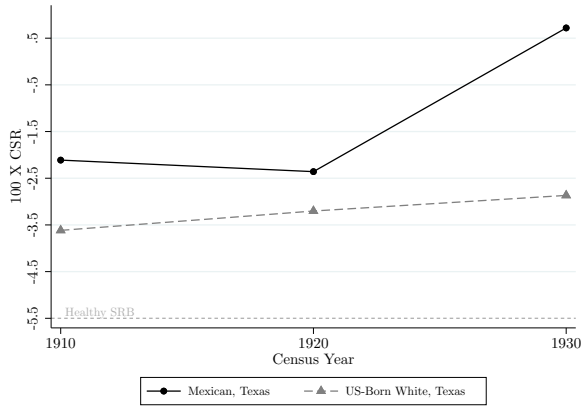
A second additional data source is the Census Place Project (Berkes et al., 2023), which allows us to place individuals in historical census places into stable geographies. This is particularly relevant given that our specification will be controlling for county fixed effects, so we want counties to remain fixed across time.

We also use contemporary vital statistics reports, such as U.S. Department of Commerce, Bureau of the Census (1934). These reports include births, deaths, and stillbirths for the Registration Area of the US. They are very limited, as Texas was not included until 1933, and they do not separate out Mexicans from other ethnicities. Still, they provide us with general levels of fetal-infant loss during the period, and allow us to validate childhood sex ratios as indicators of infant mortality. We also use the infant mortality rates from these data to estimate birth registration completeness.

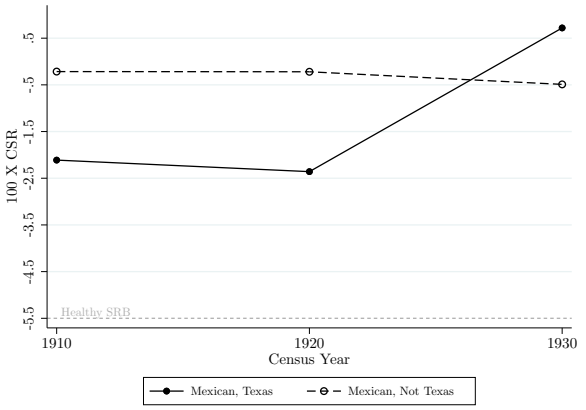
5 Main Analysis: Health Consequences

5.1 Descriptive Patterns

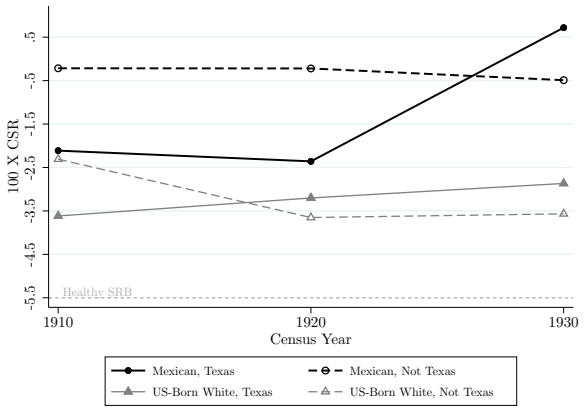
In figure 4, we plot age 3 and under sex ratios across 1910 to 1930, for states in the US-Mexico border to show the sources of variation that we leverage. We focus on ages 3 and under because these ages for the 1930 census had been exposed to the deportation campaign during infancy and/or gestation. We restrict the sample to Mexican-immigrant households and US-born white households.



(a) Within Texas, Across Households



(b) Within Mexican Households, Across States



(c) Across Households and Across States

Figure 4: Under-3 sex ratios for children in Mexican-immigrant and US-born white households

Notes: This figure presents the under-3 sex ratio for Mexican-immigrant households and for US-born white households from 1910 to 1930. A household is defined as Mexican if the household head was born in Mexico. Panel A restricts the sample to families in Texas and shows the time series by distinguishing by household type (either Mexican-immigrant or US-born white households). Panel B presents the time series for Mexican-immigrant families and distinguishes them by their state of residence: either Texas or on the other US-Mexico border states (California, Arizona, and New Mexico). Panel C shows all the different time series together, presenting them differentially for Mexican-immigrant and US-born households that live either in Texas or on the other US-Mexico border states (California, Arizona, and New Mexico).

Panel A restricts the sample to families located in Texas and it shows the evolution of the under-3 childhood sex ratio for Mexican and US-born white households[15]. For US-born whites we can see the childhood sex ratio slowly increasing throughout this period. We see a different behavior from the under-3 sex ratio for children in Mexican-immigrant households in Texas. In 1910, the under-3 sex ratio was already more female skewed for Mexican-immigrant households than for US-born

[15]A household is considered "Mexican" if the household head was born in Mexico.

white households. This is consistent with the historical public health literature that documents worse health outcomes for Mexican immigrants during this time period (Preston and Haines, 1991; Molina, 2006). From 1910 to 1920 the under-3 sex ratio slowly decreased for Mexicans. Thus, prior to the deportation campaign, fetal-infant health outcomes seem to have evolved more favorably for Mexicans than for US-born white households in Texas. However, from 1920 to 1930 we see a break in this trend for Mexican-immigrant households: the under-3 sex ratio shifted towards females by around 3 points.

Panel B shows the trends for Mexican-immigant households across the US-Mexico border, displaying the data for Texas separately. In ligth grey we display the trend for Mexican-immigrant households in California, Arizona, and New Mexico. In 1910 we see Mexican-immigrant households in Texas have better health outcomes than Mexican-immigrant households in the other US-Mexico border states. However, this changes after 1920, when Mexican-immigrant households in Texas present a change in the trend of their under-3 childhood sex ratio that is not present for Mexican-immigrant households in the rest of the border states.

Finally, panel C displays all the sources of variation we exploit in our empirical specification. It adds to the series displayed in panels A and B the evolution of the under-3 childhood sex ratio for US-born white households outside of Texas. For this group, the under-3 childhood sex ratio shifted towards males, which is consistent with improved health for the general population during this period, as previously documented in the literature Cutler and Miller, 2005; Tapia Granados and Diez Roux, 2009.

5.2 Empirical Strategy

We focus on the start of the repatriation campaign that took place from 1928 to 1930 in Texas. We do so for two reasons. First, the fact that the campaign was first deployed only in Texas allows us to use cross-state variation with the Mexican-immigrant population, which is useful to alleviate concerns that there may be something else happening specifically to Mexicans close to the 1930 census driving the results. Second, by looking at infants in the 1930 census, we alleviate concerns that sex ratios became female skewed due to non-health factors like sex-selective migration.

We formalize the descriptive results in Figure 4 with a triple-differences specification. First, we use time variation, as the campaign began in 1928. Thus, we expect worse fetal-infant health outcomes in 1930 than in previous years. Our second source of variation is based on ethnicity: we compare outcomes for children of Mexican-immigrant households to those of children in US-born white households. The campaign targeted Mexican immigrants, so we expect more female skewed sex ratios for children in households where the household head was born in Mexico. Our third source of variation is geographic. By 1930, the campaign had begun in Texas, but had not yet

extended to the rest of the country, particularly to the other US-Mexico border states. Therefore, we compare Mexican households in Texas to those in the other US-Mexico border states.

We run the following regression:

$$female_{ict} = \beta Mexican_i \times After_t \times Texas_c + \Gamma X_i + \gamma_t + \omega_c + \epsilon_{ict} \tag{1}$$

where $female_{ict}$ is a dummy variable that indicates if the individual *i* in the regression is a female, $Mexican_i$ is a dummy variable that indicates if the individual *i* lives in a household with a Mexican-born household head, $After_t$ is a dummy that indicates if the time *t* in which we are observing individual *i* is in 1930—after the deportation campaign started—, and $Texas_c$ is a dummy that indicates whether county *c* is in Texas. γ_t are census year fixed effects, ω_c are county fixed effects. Finally, X_i is a vector of controls at the household level which, aside from the $Mexican_i$ dummy, includes four variables: age of the household head and household spouse, and dummy variables that indicate whether the household head and household spouse are literate. All lower term interactions of our dummies of interest are absorbed via fixed effects as well.

Given that we expect the campaign to worsen fetal-infant health specifically for Mexicans living in Texas in 1930, we expect $\beta > 0$.

We estimate this regression via a logistic regression because it provides us with changes to the childhood sex ratio directly. Given that $CSR = \log\left(\frac{girls}{boys}\right)$ and that $\log\left(\frac{girls}{boys}\right) \rightarrow \log\left(\frac{P[female]}{1-P[female]}\right)$, the childhood sex ratio is the sample analogue of the log-odds of being female. If we use a logistic regression, then, we can estimate directly how the childhood sex ratio changes with our RHS variables when estimating equation (1).

We also estimate the regression detailed in equation (1) via OLS. While the coefficient of interest is not directly related to the childhood sex ratio as in the case of the logistic regression, it tells us the marginal changes in the propensity of females as our right hand side variables change. This is of course relevant and has direct implications for interpretation: with the female propensity rising, the childhood sex ratio will unambiguously increase, as both the numerator will rise and the denominator will fall. Therefore, the simple linear probability model is itself informative on how the campaign is changing the fetal-infant health outcomes of Mexicans.

5.3 Main Results

The results are shown in Table 1 below. Panel A presents the results when estimating the equation via a logistic regression and Panel B presents the results when estimating it using OLS. The coefficients for Panel A directly translate into changes in the childhood sex ratio, while the coefficients in Panel B tell us how the propensity of being female is changing. Both sets of coefficients go in the same

direction given that an increase in the female propensity monotonically translates into an increase in the childhood sex ratio.

Column 1 presents the least saturated version of the model, in which the lower term interactions are not collinear with the fixed effects. We include all needed lower term interactions to make our coefficient of interest interpretable. Columns 2 and 3 present a similar regression with different levels of saturation. In column 2 all lower level interaction terms are interacted with the time, Texas, and Mexican fixed effects. In column 3 we also interact the household level controls—household head and spouse literacy and age—with these fixed effects.

Regardless of the level of saturation and the estimation method, we find an increase in the childhood sex ratio—or in the female propensity for the LMP—for children of Mexican-immigrant households in Texas in the 1930 Census, right after the campaign started. The increase in fetal-infant losses caused by the repatriation campaign are robust to both the estimation method and the level of saturation. Moreover, increasing the level of saturation of the regression leads to higher estimates of β. This means that allowing for more flexibility for how we model the demographics, time, and location controls actually renders a marginally higher coefficient.

For interpretation, we focus on Panel A (logistic regression) because it yields an estimate of the change in log sex ratios, and Column 2 because is uses the canonical level of saturation. As can be seen, the childhood sex ratio increased by 0.047 points for Mexicans in Texas in 1930, a jump from the pre-period that is consistent with Figure 4.

Table 1: Changes in the Childhood Sex Ratio

	(1)	(2)	(3)
Dependent variable:	*Female*	*Female*	*Female*
Panel A: Logit			
HHH Mexican × After × Texas	0.0442**	0.0470**	0.0505***
	(0.0195)	(0.0196)	(0.0195)
Panel B: Linear Probability Model			
HHH Mexican × After × Texas	0.0110**	0.0117**	0.0126***
	(0.0049)	(0.0049)	(0.0049)
R-squared	0.0003	0.0003	0.0003
Observations	1,415,347	1,415,347	1,415,347
Fixed effects			
County	Yes	Yes	Yes
Year	Yes	—	—
Year × HHH Mexican	—	Yes	Yes
Texas × HHH Mexican	—	Yes	Yes
Texas × Year	—	Yes	Yes
Demographics × (Mexican, Year, Texas)	—	—	Yes
Controls			
HHH & HHS age, HHH & HHS literacy	Yes	Yes	Yes
SE clustered by county	Yes	Yes	Yes

Notes: This table presents the coefficient β from equation 1 under different specifications and estimation methods. Column 1 has the least saturated specification, column 2 interacts all lower terms with the fixed effects, and column 3 adds the household level controls to the interactions. Panel A shows the coefficients when estimating the equation with a logistic regression, which translates directly into changes in the childhood sex ratio. On the other hand, Panel B shows the coefficients when estimated using OLS, and, therefore, the coefficients it shows represent marginal changes in the probability that the average kid in the sample is a female. Specifically, the coefficients shown in this table tell us how the childhood sex ratio is different for children of Mexican households, in Texas, in 1930. */**/*** denote significance at the 10/5/1 percent level.

To interpret the magnitude of this coefficient, we can first turn to the sex ratios we see in Figure 4. A healthy sex ratio at birth is in the range of 5–6 log points male (-0.05 to -0.06 in terms of our estimate coefficients), as established by a large body of work in human biology Orzack et al., 2015. In 1920, the white native-born US 3-and-under sex ratio was -2.6, roughly 3 points away from healthy. The Texas Mexican 3-and-under sex ratio was about -2.3, some 3.3 points from healthy. In 1930, the white native-born sex ratio remained roughly constant. Our coefficient tells us that the Mexican sex ratio shifted around 5 points toward females. This means that the distance from a healthy sex ratio more than doubled, suggesting that fetal-infant loss more than doubled.

We can also benchmark this increase using the 1930 levels in the childhood sex ratios for the white and Black population in our sample. This gap is of roughly 3 log-points, 0.03 in terms of our

coefficient. Given that we estimate an increase in the childhood sex ratio for Mexican-immigrant families exposed to the campaign, this would be roughly equivalent to taking the Mexican-immigrant population from health levels held by the US-born white population and taking them to outcomes that are 50% worse than those of the Black population, which we know was historically undeserved by medical care.

With more assumptions, discussed at length in Appendix A1, we can do a back-of-the-envelope calculation for the implied change in stillbirths and infant mortality as a result of the campaign. First, we need to assume a change in the sex ratio at birth, which reflects miscarriages. We assume that the effect of the campaign on the sex ratio at birth was comparable to the baseline difference between whites and non-whites in 1931: 1.67 points. Subtracting this from our 4.7 points, we are left with a 3.03 point shift.

As explained in detail in Appendix A1, we can model the childhood sex ratio as:

$$CSR = SR_B + \tau(sb + IMR)$$

where SR_B is the sex ratio at birth, including live and non-live births, τ is excess male mortality, sb is the stillbirth rate and IMR is the infant mortality rate. Using CDC reports from the 1920s, we find an average τ of 0.22.

Thus, we can then divide the 3.03 point shift (0.0303) over τ, which yields a roughly 13.8 percentage point increase in combined stillbirths and infant mortality:

$$\Delta(sb + q) = \frac{\beta - \Delta SR_B}{\tau} = \frac{.0303}{.22} = 0.138$$

To compare this increase to baseline levels, we take IMR and stillbirths available for the 1920s from the 1931 vital statistics report. Combined stillbirths and infant mortality rates were about 0.16 for the non-white population, and about 0.09 for the white population. So a plausible interpretation of our coefficient is that for the Texas Mexican population combined stillbirths and infant mortality went from around the same level as whites outside of Texas—as shown in figure 4—, to worse than that of the contemporary Black population.

5.4 Childhood Sex Ratio and the Broader Campaign

As discussed in Section 2, the repatriation campaign accelerated in 1931 and lasted until 1935. Thus, we would expect to find a shift towards females in the childhood sex ratios of Mexican families in the 1940 census as well. These changes in the childhood sex ratio should be concentrated in the cohorts born while the campaign was active: those who would be 5 years or older in 1940. In this section, we look at whether we can find this effect on the broader campaign.

The main challenge to do this is the fact that, since the campaign spread to the rest of the country in 1931, we can't use the geographic variation that we used in the previous section. Therefore, we limit our attention to Mexican-immigrant and US-born white families in Texas from 1910 to 1940. We take the universe of children in Texas between 0 and 9 years old in the 1910, 1920, 1930, and 1940 censuses and estimate the following regression via logit:

$$\begin{aligned} female_{ict} = \beta_0 + \beta_1 Mexican_i + \beta_2 I(year_t = 1930) + \beta_3 Mexican_i \times I(year_t = 1930) \\ + \beta_4 I(year_t = 1940) + \beta_5 Mexican_i \times I(year_t = 1940) + \Gamma X_i + \gamma_c + \omega_t + \epsilon_{ict} \end{aligned} \tag{2}$$

where once again $Mexican_i$ is a dummy that indicates whether the household head was born in Mexico, X_i are the same household level controls—household head and spouse age and literacy—, and γ_c and ω_t are county and census year fixed effects, respectively. In this regression, β_3 will tell us if the propensity of being female differentially changes for Mexicans in 1930 and β_5 will tell us if it does in 1940.

We show the estimated coefficients of interest in table 2 below. Column 1 presents the coefficients when estimating the regression with the full sample. We find a shift toward females for both the 1930 and 1940 censuses. Columns 2 and 3 separate out these results into sub-samples in which we expect to find different changes: children 3 years of age or younger and 5 years of age or older. In column 2 we find that the under-3 sex ratio shifted toward females in 1930 but not 1940. This is precisely what we expected given that children 3 years of age or younger in 1940 would have been born after 1936 and, therefore, would not have been affected by the campaign.

In column 3, where we look at children 5 years old or older, we find the opposite pattern. For 1930, we don't find a statistically significant increase in the childhood sex ratio of this age group. In 1940, there was a large statistically significant shift toward females in the age 5-9 sex ratio. The 1940 coefficient is what we would expect to see for cohorts exposed to the campaign in utero and infancy: children that were born during the campaign see a female-skewed sex ratio, likely reflecting worse fetal-infant health.

Table 2: Under-10 Sex Ratios

	(1) Female	(2) Female	(3) Female
Mexican × 1930	0.0231*	0.0408***	0.0084
	(0.0127)	(0.0136)	(0.0196)
Mexican × 1940	0.0294***	0.0161	0.0348**
	(0.0097)	(0.0139)	(0.0157)
Observations	2,408,754	957,700	1,209,136
FE	County Year	County Year	County Year
SE Cluster	County	County	County
Sample	All	Under 3	Above 5

Notes: This table presents the coefficients from running the regression specified in equation 2 using a logistic regression on a repeated cross section of children 9 years old or younger in Texas in the 1910, 1920, and 1930 censuses. Mexican is a dummy variable that identifies whether the household head was born in Mexico. Column 1 runs the regression with the full sample, column 2 runs the regression with children 3 years old or younger, and column 3 with children 5 years old or older. */**/*** denote significance at the 10/5/1 percent level.

6 Robustness

We have shown that there was a large shift towards females among children of Mexican-immigrant households exposed to the deportation campaign during infancy and/or gestation. We have argued that this reflects an increase in early-life mortality (miscarriages, stillbirths, and infant mortality). In this section, we address alternative explanations for this shift in sex ratios. First, we show that the effect is only found in cohorts who were affected during gestation and/or infancy, not older ages. Then we show that the effect is only present for Mexican-immigrant households, not other minorities. Then we show that alternative specifications, using only within-Texas variation and using only within-Mexican variation, give us similar results to our triple-differences approach. Then, we show that the number of children in Mexican-immigrant households decreases, just as we would expect if fetal-infant loss increased. Finally, we provide evidence against selective migration as driving our results.

6.1 Childhood Sex Ratio for Cohorts Exposed to the Campaign at Older Ages

To show that the estimated increase in the childhood sex ratio is indeed related to fetal-infant loss, we estimate the triple-differences interaction for different age partitions of children of household

heads. Given that the male frailty is mainly during gestation and infancy—not later childhood (Hill and Upchurch, 1995)—we would not expect to find any significant increase in the female propensity for cohorts of children born before 1927, who would have been older than infants by the beginning of the campaign in the 1930 census.

We show the estimated triple-differences coefficients in figure 5 below. We once again do so using a logistic regression for ease of interpretation, but the results remain unchanged in terms of direction and statistical precision when using OLS. On the left, we display the baseline coefficient analogous to the one displayed in table 1, column 2, panel A. As discussed above, we find an increase in the log-odds of being female of around 0.05.

Moving to the right, we show the coefficients for two different age groups: between 4 and 6 years old and between 7 and 9 years old. We find a negative coefficient for Mexican children in Texas in 1930—that is, the childhood sex ratio is becoming more male biased—that is not statistically significant. Thus, in accordance to what we would expect if the childhood sex ratio changes are reflecting changes in health rather than other issues, we don't see a more female biased childhood sex ratio for cohorts that were older than infants by the time the campaign started.

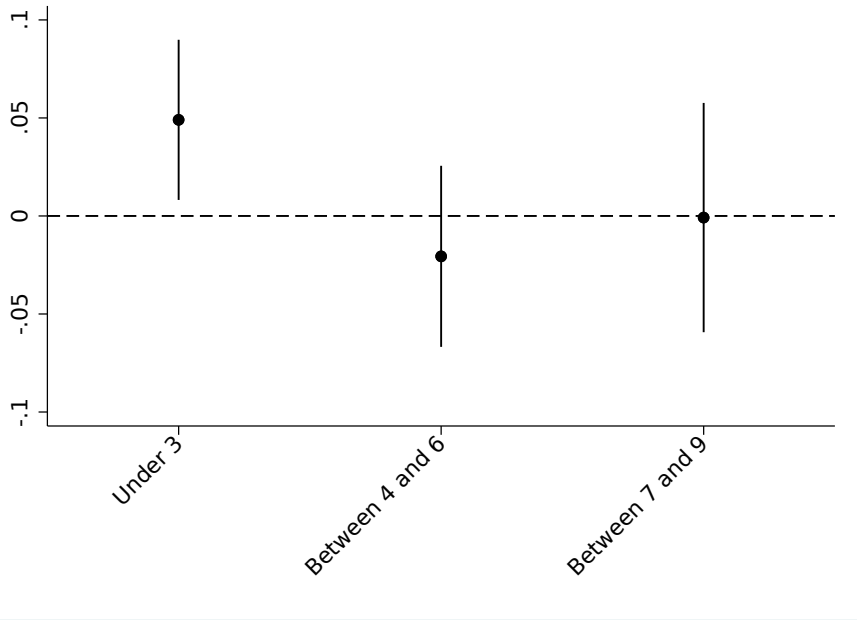


Figure 5: Coefficient Plots For Different Cohorts

Notes: This figure presents the coefficients when running the regression jointly for different age cohorts. The coefficient on the left shows our baseline estimated coefficient with children 3 years of age or younger. To the right, we estimate the same triple-differences regression with two different age groups: children between 4 and 6 years of age and children between 7 and 9 years of age.

6.2 Childhood Sex Ratio for Other Minority Households

Another potential concern is whether the results are specific to Mexican-immigrant households, as opposed to being driven by changes happening to minorities more generally in our context. This could happen if discrimination got differentially worse in Texas in 1930 or if local policies worsened the conditions of other minorities. If this were the case, we would expect to find a similar increase in the childhood sex ratio for other minorities when compared to US-born whites as we found with Mexican-immigrant households.

To explore this, in this section we replicate the results of table 1, panel A, column 2, but look at other households aside from Mexican-immigrant households. If the results in the previous section were driven by changes affecting other minorities, we would expect to find a similar coefficient for them. We use the same specification as in column 2—with enough saturation so that lower-term interactions are absorved by the fixed effects—and estimate it via logistic regression. We present the results in table 3.

Table 3: CSR Changes for Other Minorities

	(1) Female	(2) Female	(3) Female
Group × After × Texas	0.0470** (0.0196)	−0.0519* (0.0273)	−0.0171 (0.0323)
Observations	1,415,347	1,292,358	1,199,695
Placebo group	HHH Mexican	HHH Foreign Born (Excluding Mexicans)	HHH Black
Fixed effects			
County	Yes	Yes	Yes
Year × Group	Yes	Yes	Yes
Texas × Group	Yes	Yes	Yes
Texas × Year	Yes	Yes	Yes
Controls			
HHH & HHS age, HHH & HHS literacy	Yes	Yes	Yes
SE Cluster	County	County	County

Notes: This table presents the coefficient β_7 from equation 1 for different groups estimated via a logistic regression. Column 1 presents the triple interaction for Mexican migrant households and corresponds to the coefficient displayed in Table 1, column 2, Panel A. Column 2 uses as a group of interest households where the household head is foreign born and column 3 where household heads are black. The omitted category are households where the household head is a native born white. */**/*** denote significance at the 10/5/1 percent level.

Column 1 of table 3 presents the coefficient for Mexican-immigrant households, and therefore

corresponds to the coefficient presented in table 1, column 2, panel A. Column 2 focuses on foreign-born households—excluding Mexican households—and column 3 focuses on US-born Black households. The comparison group for each of these is households with a US-born white household head.

As can be seen, the coefficients from columns 2 and 3 are negative rather than positive. Moreover, the coefficient for foreign-born households is statistically significant. This brings us reassurance that our main results are indeed driven by something affecting specifically Mexicans in Texas in 1930 rather than a more general change. Thus, we are confident that the worsened fetal-infant health outcomes were caused by the start of the Mexican Repatriation campaign.

6.3 Within Texas and Within Mexican Differences

Our preferred specification is the triple-differences approach described above because it leverages all the levels of variation available to us between 1928 and 1930. However, we could alternatively choose to either focus on the Mexican population and look at changes across states and time or on the Texas population and look at the changes for Mexicans in 1930 instead. That is, you could look at either of the second differences to measure the fetal-infant health consequences of the campaign. In this section, we do both to show that our result does not depend on a specific empirical approach.

First, we focus on Texas given that the campaign started there and leverage the fact that the campaign affected only Mexican households and only by the 1930 census. We restrict our observations to families in Texas and use the differences-in-difference approach described in equation (3) below. As before, households where the household head is a US born white are the reference group.

$$female_{ict} = \delta_0 + \delta_1 Mexican_i + \delta_2 After_t + \delta_3 Mexican_i \times After_t + \Gamma X_i + \gamma_t + \omega_c + \epsilon_{ict} \quad (3)$$

Alternatively, we can also focus on the Mexican-immigrant households and leverage the fact that the campaign started and remained in Texas from 1928 through 1930. Thus, we'd expect the childhood sex ratio to increase for Mexican households in Texas in 1930 while remaining unchanged in the other border states. This approach is laid out in equation (4) below, which we estimate by restricting our sample to Mexican-immigrant households. Here, the omitted category are Mexican-immigrant households in California, Arizona, and New Mexico.

$$female_{ict} = \alpha_0 + \alpha_1 After_t + \alpha_2 Texas_c + \alpha_3 After_t \times Texas_c + \Gamma X_i + \gamma_t + \omega_c + \epsilon_{ict} \quad (4)$$

The estimated coefficients can be found in table 4. Panel A presents the results for the within-Texas DiD detailed in equation (3), while Panel B the within-Mexican DiD detailed in equation (4).

Column 1 presents the coefficients obtained via OLS estimation, while column 2 the ones obtained when using a logistic regression.

Table 4: Differences in Differences coefficients

	(1) Female	(2) Female
Panel A: Within Texas Variation		
HHH Mexican X After	0.0103***	0.0413***
	(0.0035)	(0.0138)
Observations	710,520	710,520
R-squared	0.0004	
Panel B: Within Mexican Variation		
After X Texas	0.0102**	0.0407**
	(0.0047)	(0.0189)
Observations	147,857	147,836
R-squared	0.0025	
FE	County Year	County Year
SE Cluster	County	County
Estimation Method	OLS	Logit

Notes: This table presents the estimated coefficients of equation 3 –within Texas variation– in Panel A and of equation 4 –within Mexican variation– in Panel B. Column 1 presents the coefficients when estimated via OLS and column 2 when estimated via logistic regression. Standard errors are clustered at the county level. */**/*** denote significance at the 10/5/1 percent level.

The coefficients are remarkably similar to those presented in table 1, where we use a triple-differences approach to leverage all the sources of variation. Finding very similar estimates when focusing on two of these sources of variation brings us reassurance that our findings are not driven by the empirical approach selected but rather remain similar when using other reasonable strategies.

6.4 Number of children in Mexican households

In this section, we estimate our triple-differences specification, but take the number of children in the household as the outcome variable. A decrease in the number of children in Mexican-immigrant households after the onset of the campaign would be consistent with our interpretation of the changes in childhood sex ratios being driven by fetal-infant loss. Rather than using our individual level data, we instead construct a new dataset from the full count US censuses between 1910 and 1930 at the household level. We keep households in the US-Mexico border with a female spouse between 15 and 35 years old for which the household head is either a Mexican immigrant or a US

born white. The number of observations increases here, because the regression is at the household level, and there are many households without any children age 3 and under.

The results can be found in Table 5. Column 1 focuses on the number of children 3 years old or younger, while columns 2 and 3 focus on girls and boys 3 years old or younger, respectively. Column 1 shows that Mexican households in Texas in 1930 saw a decrease of almost 0.05 children under 3 years of age or younger, on average, significant at the 5% level. This would imply having 5 less children for every 100 in 1910 and 1920. This rate is smaller than the degree of fetal-infant loss which we estimate above, but that is likely due to changes in fertility choices. Mexican households that lose a child could try and succeed in having another, and white households—our comparison group—saw fertility rates decline from 1920 to 1930 (Coale and Zelnik, 1963).

Columns 2 and 3 show a similar and consistent story as well. In both columns the estimated coefficient of interest is negative. However, the coefficient in column 2 is very noisy and statistically insignificant, so we only document a decrease in the number of male children 3 years of age or younger. This is consistent with our previous results and with the male frailty literature described in section 3.

Table 5: Number of children in Mexican households

	(1)	(2)	(3)
	Under 3	Under 3	Under 3
	All	Female	Male
HHH Mexican X Year 1930 X Texas	-0.0485**	-0.0090	-0.0394***
	(0.0245)	(0.0132)	(0.0130)
Observations	1,874,388	1,874,388	1,874,388
R-squared	0.0601	0.0288	0.0280
Fixed effects			
County	Yes	Yes	Yes
Year × HHH Mexican	Yes	Yes	Yes
Texas × HHH Mexican	Yes	Yes	Yes
Texas × Year	Yes	Yes	Yes
Controls			
HHH & HHS age, HHH & HHS literacy	Yes	Yes	Yes
SE Cluster	County	County	County
Mean	.68	.33	.34
Mean for Mexicans	.82	.41	.42

Notes: This table presents the estimated coefficients of equation 1 but having the number of children in the household as a dependent variable. The coefficients are estimated via OLS. The sample is composed of households with the spouse present and between the ages of 15 and 35. Column 1 uses as a dependent variable the total number of children in the household 3 years old or younger, column 2 the total number of girls, and column 3 the total number of boys. Standard errors are clustered at the county level. */**/*** denote significance at the 10/5/1 percent level.

6.5 Selective Migration

One final concern is that the increase in fetal-infant loss is true, but that it reflects compositional change from selective migration. If richer, healthier Mexican immigrants were the first to leave, then the population left would have worse health outcomes than before. This would not reflect an actual increase in fetal-infant loss, but just a shift in the composition of Mexican-immigrant households.

The historical literature tends to emphasize that returnees were negatively selected in terms of occupations (McKay, 1985; Weber, 2013). If this were the case and given the fact that richer families tend to have better health outcomes, selective out-migration would actually work against our findings. We once again make use of the census data to test this claim made in the historical literature.

We collapse the data at the household level as we did in section 6.4 and run the same triple-differences design to see whether household observable characteristics change with the campaign. Our results can be found in table 6 below.

Table 6: Household Characteristics in 1930

	(1) *HH Occscore*	(2) *HHH Occscore*	(3) *Number of Families*	(4) *People in Household*
Mexican × Year 1930 × Texas	−0.0208	0.7760	−0.0640**	−0.2070***
	(0.9798)	(0.5551)	(0.0311)	(0.0608)
Observations	1,874,388	1,874,388	1,874,388	1,874,388
R-squared	0.0435	0.1079	0.0090	0.1266
Fixed effects				
County	Yes	Yes	Yes	Yes
Year × HHH Mexican	Yes	Yes	Yes	Yes
Texas × HHH Mexican	Yes	Yes	Yes	Yes
Texas × Year	Yes	Yes	Yes	Yes
Controls				
HHH & HHS age, HHH & HHS literacy	Yes	Yes	Yes	Yes
SE Cluster	County	County	County	County
Mean	30.995	22.816	1.248	4.292

Notes: This table presents the results of running a triple interaction regression such as the one specified in equation 1 but with a different set of outcome variables. The sample is collapsed at the household level and it includes all households regardless of their fertility outcomes. Column 1 has as a LHS variable the sum of all occupational scores of people in the household. The Census didn't ask for income until 1940, so we use the median income of each occupation in the 1950 census –when enumerators also tracked self employed individuals– as an occupational score. These occupational scores should reflect each occupation's permanent income, net of short term noise, as argued by Abramitzky et al. (2012, 2014). Column 2 focuses on the household head's occupational score. Columns 3 and 4 look into the household composition. Column 3 looks into differences in the number of families within each household and column 4 looks at the number of people within the household. */**/*** denote significance at the 10/5/1 percent level.

Columns 1 and 2 proxy the household income using occupational scores[16]. Column 1 adds up all the positive occupational scores in the households to account for the fact that, to compensate occupational downgrades, other members of the household could start working. We find no significant difference for Mexicans in Texas in 1930. Column 2 measures the occupational score of the household head and, doesn't find a significant change either.

Columns 3 and 4 take a different approach and instead seek to measure the material conditions of Mexican households. If families face worse material conditions, they might end up sharing a living space by either mixing families within the household or by having more distant members of the families living in the household. Thus, column 3 looks at the number of families within the household, while column 4 at the number of people. As with the occupational score, we don't find Mexican families in Texas to be living more crowded households. Instead, we find a decrease in the number of people in the household, which is consistent with our early-life loss results.

Overall, we view this as evidence that out-migration was not positively selected, which means that our health results can't be explained by the composition of Mexican households. Moreover, these results also point to the fact that material conditions didn't worsen for Mexicans with the onset of the campaign, which could have been a plausible mechanism.

Finally, to address any remaining concerns of selective out-migration driving our results, we do a bounding exercise to calculate how large this outmigration should be to drive our results. Using census data, previous literature has found roughly 20-25% of US Mexicans left during the broad repatriation campaign (Gratton and Merchant, 2013). The sex ratio is a function of infant mortality, which has a floor of 0. The childhood sex ratio then has a floor (for large populations) of a healthy sex ratio at birth, which is 5–6 log points male (Orzack et al., 2015). Suppose that 20% of Mexican immigrants had left Texas by 1930, a very conservative assumption since departures peaked after 1930. Suppose that these return migrants were positively selected, and had very low rates of IMR. Their offspring sex ratio would then be about -5.5 log points. We can easily calculate the effect of this out-migration on the population sex ratio. In 1920, Mexican immigrants in Texas had sex ratios at -2.4 in 1920. Under these assumptions, selective outmigration could, at most, shift the sex ratio from -2.4 to -1.625 log points. This is a shift of 0.775 points, when our estimated coefficient was 4.7 points. Therefore, selective outmigration could at most only explain around 17% of our estimated increase in the childhood sex ratio.

[16]The census didn't ask for income until 1940 so we have to proxy household income using occupational scores. These are provided by IPUMS and impute the median income reported at the occupation level in the 1950 census. While an imperfect measure, they provide a reasonably good measure of the household's permanent income. For an in-depth discussion about using occupational scores in this time period look at Abramitzky et al. (2012, 2014).

7 Mechanisms

7.1 Withdrawal from Public Life

So far we have shown that the Mexican Repatriation Campaign increased early-life mortality for Mexican households. We've done so by proxying fetal-infant loss using the childhood sex ratio, which leverages the fact that in infancy and in utero males are more frail than females. We now turn to explain why this happened.

We focus on withdrawal from public services and healthcare[17]. Anecdotal evidence suggests that during the campaign Mexicans withdrew from public life. They avoided relief stations and medical facilities. We provide novel quantitative evidence for this retreat from public services using Texas birth registration data. Matching the data to the census, we make new estimates of birth registration. We find a decline in birth registration.

Moreover, as we discuss further below, the decline in birth registration which we find among Texas Mexicans suggests a reduction in the frequency of births assisted by medical professionals (doctors and nurses), who were required by law to record births. Therefore, the decline in birth registration is not only an indicator of a withdrawal from public life, but also a sign of reduced healthcare utilization. In an era when doctors and nurses provided important care and information to mothers, this is likely an important driver of the increase in fetal-infant loss we document in section 5.3.

7.2 Decrease in Healthcare Utilization

Fear, intimidation, and coercion were the driving force behind the Mexican Repatriation Campaign, as detailed in section 2. The lack of state capacity and the ambition to deport as many Mexicans as possible led the federal immigration authorities to design a campaign intended to terrorize Mexicans and get them to leave on their own. As a result, Mexicans withdrew from public life and public places (Goodman, 2020), which includes medical visits. Similar changes in health seeking behavior have been observed among migrants targeted by deportation campaigns today (New York Times, 2026).

We lack the data to directly measure healthcare utilization in the 1920s and 1930s. Aside from a few questions regarding extreme medical conditions—such as blindness and deafness—the census does not record medical care. To proxy for changes in healthcare utilization, we estimate birth

[17] An obvious channel is maternal stress, which would directly impact the sex ratio through increased miscarriages, stillbirths, and infant mortality, as discussed in section 3. However, we cannot measure maternal stress. Biomarkers, like blood cortisol levels, are completely unavailable for our period. Anecdotal evidence like newspapers and interviews give examples of the stress experienced by Mexican-immigrant women, but they do not provide enough data to run quantitative tests.

registration completeness in Texas.

US birth registration systems were managed by the states, and remained very incomplete well into the 1930s. Moreover, registration systems did not cover the population equally. Southern states, including Texas, had worse coverage than most Northern states. Minorities, like the Black population, had lower rates of birth registration than whites Eriksson et al. (2018). This was also the case for Mexican immigrants and Mexican-Americans, as we show below.

One of the key determinants of registration was having a medically assisted birth. Doctors, nurses, and midwives were required by law to register births, and could be subject to fines if they did not report births to state authorities (Eriksson et al., 2018).[18] We use a measure of birth registration completeness at the county-Mexican level to proxy changes in fetal-infant healthcare.

To construct this measure, we follow Eriksson et al. (2018), who in turn drew on older work by Coale and Zelnik (1963). The underlying assumption is that the census had better coverage of the population than birth registration systems. We take the census single-age cohorts in 1930 as the true number of surviving infants. We infer year of birth from child age, which is imperfect as the census asked 'age at last birthday' in 1930. Taking the infant mortality rates from the Vital Statistics reports (e.g., U.S. Department of Commerce, Bureau of the Census, 1934), we then infer how many children would have been born to end up with the observed cohort size in 1930.

There are two main differences between our approach and the one in Eriksson et al. (2018). First, we differentiate between children born to Mexican parents and other children. The census can identify individuals born in Mexico because it directly asks for the place of birth, but the birth registration data doesn't have this information. Thus, we infer Mexican origin using last names: we calculate for each last name the proportion of individuals that claim they were born in Mexico in the 1930 census or have Mexican heritage[19]. We then classify a last name as Mexican if more than 80% of individuals in the 1930 census in Texas with that last name claim to have been born in Mexico. To keep things consistent, we then use this Mexican label for both the census and birth data. More details about this construction can be found in Appendix A2.

A second challenge when calculating birth registration completeness for Mexicans is the fact that we don't have infant mortality rates for this specific group that we can use to infer the number of births. Moreover, this paper is documenting an increase in fetal-infant losses, which should be partially explained by an increase in infant mortality rates for children born to Mexican families. We opt to take a conservative route and use the infant mortality rate at the county level reported in

[18] In the case of deliveries performed in hospitals, registration would be even more likely given the fact that a paper trail was produced and, therefore, lack of registration could be more easily traced.

[19] Mexican last names don't differ enough between first and later generation immigrants for us to be able to distinguish them. Consequently, this section deviates from the main analysis because we lump Mexican immigrants and Mexican-Americans under the same "Mexican" category. Given that the onset of the campaign did not target Mexican-Americans as intensely as it did Mexican immigrants, this data issue would bias our results downwards.

the 1933 Vital Statistics Report (U.S. Department of Commerce, Bureau of the Census, 1936)—the first one for which we have data from Texas—and use it to infer births for both Mexicans and non-Mexicans.

Given the fact that—according to both our childhood sex ratio measures and the historical medical literature (e.g. Molina, 2006)—-baseline infant mortality was higher for Mexicans than for non-Mexicans, this assumption should bias our calculation of birth registration for Mexicans upwards.[20] Moreover, if the infant mortality rate is differentially increasing for Mexicans, this bias would be higher in magnitude by 1930.[21] Therefore, what we estimate in terms of the percentage of births registered for Mexicans should be interpreted as an upper threshold.

Figure 6 presents the birth registration completeness patterns for Mexicans and non-Mexicans in Texas between 1926 and 1930. Overall, we see lower registration birth completeness rates for Mexicans throughout the period, but, between 1927 and 1929, we see the rate increase at a similar pace for both groups. In 1930, however, we see a substantial drop in the Mexican registration rate, while registration completeness for non-Mexicans continued to grow. This break in trend implies that babies born in Mexican families were less likely to be registered after the start of the campaign. This could be due to a reduction in the number of physician-assisted births, as Mexicans reduced healthcare utilization. More broadly, it speaks to their withdrawal for public services, of which healthcare is an important part.

[20] If the infant mortality rate for Mexican children was indeed higher, we would infer a higher number of births, which should imply a higher number of inferred births.

[21] A similar argument can be made for the families that went back to Mexico as a result of the campaign.

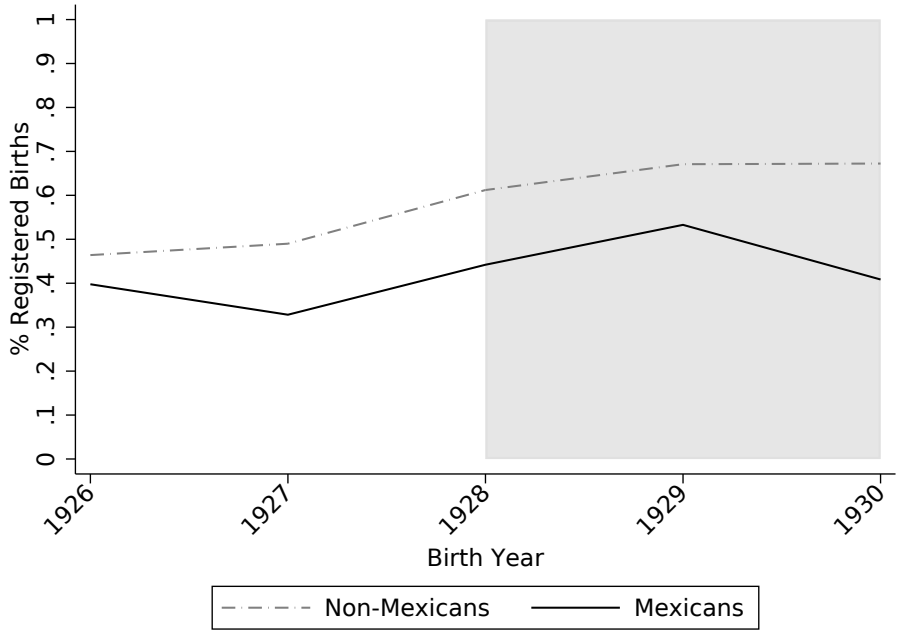


Figure 6: Birth Registration Completeness

Notes: This figure presents the percentage of registered births for Mexicans and non-Mexicans in Texas between 1926 and 1930. Mexican households are inferred using last names. The number of births is calculated using the 1930 census and the CDC IMR at the county level in 1933, the first year for which we have data from Texas. More details about the construction of these variables are provided in Appendix A2).

The second difference between our exercise and Eriksson et al. (2018) is that they provide estimates at the state level, while we do so at the county level. Given that the census doesn't capture the county of birth, our underlying assumption is that, for children three years of age or younger, the county of residence should roughly coincide with the county of birth registration.

To formally test the drop in the Mexican registration after the campaign started, we run a differences in differences regression to leverage the variation that the onset of the campaign brought to Mexicans. We run this regression within Texas, which is what our birth registration data covers.

$$registration_{mct} = \beta_0 + \beta_1 After_t + \beta_2 Mexican_m + \beta_3 After_t \times Mexican_m + \gamma_t + \omega_c + \epsilon_{mct} \quad (5)$$

where $After_t$ is a dummy that indicates if year t is bigger than 1927 and $Mexican_m$ is the dummy that we construct to identify Mexican births. We include time (γ_t) and county (ω_c) fixed effects. The results can be found in table 7.

Table 7: Birth Registration Completeness

	(1) % Births Registered	(2) % Births Registered
After X Mexican	-0.0822**	-0.1222***
	(0.0376)	(0.0357)
Observations	1,724	1,695
R-squared	0.5015	0.7368
non-Mexican mean	.48	.48
Mexican mean	.36	.36
FE	County Year	CountyXYear
SE Cluster	County	County

Notes: This table presents the DiD estimated coefficient for birth registration completeness described in equation 5. Column 1 has year and county fixed effects, while column 2 uses countyXyear fixed effects. Standard errors are clustered at the county level. */**/*** denote significance at the 10/5/1 percent level.

Column 1 has year and county fixed effects, while column 2 uses county and year interaction fixed effects to account for any changes that could have happened at the county level. We cluster the standard errors at the county level. In both specifications we find a large and statistically significant decline in the registration completeness of Mexicans: 8 and 12 percentage points respectively. Given the strong link between birth registration and doctor assisted births, we view this as evidence that Mexicans avoided medical care as a result of the campaign. In an era when doctors and nurses provided important care and information to mothers, this is likely a driver of the increase in fetal-infant loss we document in section 5.3.

8 Conclusion

In this paper we analyze the effect of the Mexican Repatriation Campaign (1928-1935) on early-life outcomes for children in Mexican-immigrant households in the US-Mexico border states. We focus on early-life mortality, composed of miscarriages, stillbirths, and infant mortality.

Measuring early-life mortality outcomes directly is not feasible due to lack of trustworthy, disaggregated enough data. To overcome the lack of high-quality, granular vital statistics from this period, we use childhood sex ratios as our outcome of interest. Childhood sex ratios reflect maternal health (through miscarriages) and fetal-infant loss because, biologically, males are more frail than females in gestation and infancy.

Texas was the only state where the repatriation campaign had begun by 1930, giving us a natural identification strategy. Our main empirical strategy uses a triple-differences approach that leverages

three sources of variation: spatial, given that the campaign started and remained in Texas; household, given that it targeted Mexican-immigrant households; and time, given that we expect to see the effects of the campaign in the 1930 census.

We find that the repatriation campaign coincided with a large shift towards females in the under-3 sex ratio in Mexican-immigrant households. We confirm this shift using the 1940 census, finding that the cohorts exposed to campaign during gestation and infancy, born 1930-1935, also experienced this female shift. This effect reflects early-life mortality: a combination of increases in infant mortality, stillbirths, and miscarriages. We calculate that this shift in sex ratios likely reflected a 13.8 percentage point increase in combined stillbirths and infant mortality.

The female-shift in childhood sex ratios seen in Mexican-immigrant households is robust to a number of tests. Our results are similar across various levels of saturation of fixed effects, with more stringent specifications making the coefficient increase in magnitude. Our results are not present in placebo groups, such as other minorities or in older cohorts. We find an accompanying decrease in the number of children in Mexican-immigrant households, as we would expect. Finally, we provide evidence against selective migration explaining our results.

To explain these results, we use the Texas birth registration index for 1926-1930 and leverage two facts: (1) birth registration in Texas was grossly incomplete and (2) doctor assisted births tended to be registered. We construct a novel dataset of birth registration completeness at the county-race-year level as a proxy measure for having doctor assisted births. We find that, after the campaign started, Mexicans saw a substantial (8 to 12 percentage point) drop in their birth registration completeness. This decline is birth registration reflects a withdrawal from public life indicative of the stress and fear present in the Mexican community. It points towards a reduction in healthcare utilization, as doctors and nurses were required to register births.

Our findings add to the growing literature about the negative health consequences of violent immigration enforcement in the US. Our setting is unique in its scale and intensity relative to previous research, as the 1928-1935 campaign was far larger than any in the 21st century. Previous research has demonstrated that deportation campaigns fail to benefit non-immigrants in terms of labor outcomes. Our research goes a step further to show just how destructive these campaigns are for the targeted community and their mostly US-born children.

References

Abramitzky, R. and Boustan, L. (2017). Immigration in American economic history. *Journal of Economic Literature*, 55(4):1311–1345.

Abramitzky, R. and Boustan, L. P. (2022). *Streets of Gold: America's Untold Story of Immigrant Success*. PublicAffairs, New York.

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.

Alanís Enciso, F. S. (2017). *They Should Stay There: The Story of Mexican Migration and Repatriation during the Great Depression*. University of North Carolina Press, Chapel Hill. Translated by Russ Davidson.

Alanís Enciso, F. S. (2020). La 'Ley Box' y la prensa mexicana (1929-1930). *Historia Mexicana*, pages 1681–1727.

Amuedo-Dorantes, C., Churchill, B., and Song, Y. (2022). Immigration enforcement and infant health. *American Journal of Health Economics*, 8(3):323–358.

Balderrama, F. E. and Rodríguez, R. (2006). *Decade of Betrayal: Mexican Repatriation in the 1930s*. UNM Press.

Berkes, E., Karger, E., and Nencka, P. (2023). The Census Place Project: A method for geolocating unstructured place names. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

Bogardus, E. S. (1933). Mexican repatriates. *Sociology and Social Research*, pages 169–176.

Castillo, M. and Dribe, M. (2026). The mortality gradient in town and country: Social class and childhood mortality in Sweden at the turn of the twentieth century. *The Economic History Review*.

Catalano, R., Bruckner, T., Anderson, E., and Gould, J. B. (2005). Fetal death sex ratios: A test of the economic stress hypothesis. *International Journal of Epidemiology*, 34(4):944–948.

Chen, S. and Xie, B. (2024). Institutional discrimination and assimilation: Evidence from the Chinese Exclusion Act of 1882. *Explorations in Economic History*, 94:101615.

Clarke, J. (1786). Observations on some causes of the excess of the mortality of males above that of females. *The London Medical Journal*, 9(Pt 2):179.

Coale, A. J. and Zelnik, M. (1963). *New Estimates of Fertility and Population in the United States: A Study of Annual White Births from 1855 to 1960 and of Completeness of Enumeration in the Censuses from 1880 to 1960*. Princeton University Press, Princeton.

Cortes, G. and Sant'Anna, V. (2024). Send them back? the real estate consequences of repatriations. *Journal of Political Economy: Macroeconomics*. Accepted, Journal of Political Economy: Macroeconomics.

Cutler, D. and Miller, G. (2005). The role of public health improvements in health advances: The twentieth-century United States. *Demography*, 42(1):1–22.

Duncan, B. and Trejo, S. J. (2025). Which mexicans are white? enumerator-assigned race in the 1930 census and the socioeconomic integration of mexican americans. *ILR Review*, 78(1):62–85.

Eriksson, K., Niemesh, G. T., and Thomasson, M. (2018). Revising infant mortality rates for the early twentieth century United States. *Demography*, 55(6):2001–2024.

Escamilla-Guerrero, D., Kosack, E., and Ward, Z. (2021). Life after crossing the border: Assimilation during the first Mexican mass migration. *Explorations in Economic History*, 82:101403.

Escamilla-Guerrero, D. and López-Alonso, M. (2023). Migrant Self-Selection and Random Shocks: Evidence from the Panic of 1907. *The Journal of Economic History*, 83(1):45–85.

Feigenbaum, J. J., Muller, C., and Wrigley-Field, E. (2019). Regional and racial inequality in infectious disease mortality in U.S. cities, 1900–1948. *Demography*, 56(4):1371–1388.

Fish, E. S. (2022). Race, history, and immigration crimes. *Iowa Law Review*, 107(3):1051–1105.

Goodman, A. (2020). *The Deportation Machine: America's Long History of Expelling Immigrants*. Princeton University Press.

Gratton, B. and Merchant, E. (2013). Immigration, repatriation, and deportation: The Mexican-origin population in the United States, 1920–1950. *The International Migration Review*, 47(4):944–975.

Guimbeau, A., Menon, N., and Musacchio, A. (2022). Short- and medium-run health and literacy impacts of the 1918 Spanish flu pandemic in Brazil. *The Economic History Review*, 75(4):997–1025.

Hill, K. and Upchurch, D. M. (1995). Gender differences in child health: Evidence from the demographic and health surveys. *Population and Development Review*, pages 127–151.

Klein, J. P. and Moeschberger, M. L. (2003). *Survival Analysis: Techniques for Censored and Truncated Data*. Statistics for Biology and Health. Springer, New York, 2nd edition.

Lee, J., Peri, G., and Yasenov, V. (2022). The labor market effects of Mexican repatriations: Longitudinal evidence from the 1930s. *Journal of Public Economics*, 205:104558.

Liebert, H. and Mäder, B. (2026). Physicians and the production of health: Returns to health care during the mortality transition. *Review of Economics and Statistics*, 108(2):452–469.

Long, J., Medici, C., Qian, N., and Tabellini, M. (2026). The impact of the Chinese Exclusion Act on the economic development of the Western US. Working paper.

Loomans, E. M., van Dijk, A. E., Vrijkotte, T. G. M., et al. (2013). Psychosocial stress during pregnancy is related to adverse birth outcomes: Results from a large multi-ethnic community-based birth cohort. *European Journal of Public Health*, 23(3):485–491.

Lopez Araiza, S. (2026). Overnight americans: Economic assimilation of mexican crossed by the border. *Working Paper*.

Lytle-Hernández, K. (2010). *Migra!: A History of the U.S. Border Patrol*. University of California Press, Berkeley.

McDevitt-Irwin, J. and Irwin, J. R. (2025). Infant mortality among US whites in the 19th century: New evidence from childhood sex ratios. *Demographic Research*, 52:303–350.

McKay, R. R. (1985). *Texas Mexican Repatriation During the Great Depression*. University of Oklahoma at Norman.

Molina, N. (2006). *Fit to Be Citizens? Public Health and Race in Los Angeles, 1879–1939*. University of California Press, Berkeley.

New York Times (2026). Migrants are skipping medical care, fearing ICE, doctors say. https://www.nytimes.com/2025/05/08/health/migrants-health-care-trump.html. Accessed August 5, 2026.

NPR (2006). Remembering California's 'Repatriation Program'. https://www.npr.org/2006/01/02/5079627/remembering-californias-repatriation-program. Produced by Melissa Block, January 2, 2006.

Orzack, S. H., Stubblefield, J. W., Akmaev, V. R., Colls, P., Munné, S., Scholl, T., Steinsaltz, D., and Zuckerman, J. E. (2015). The human sex ratio from conception to birth. *Proceedings of the National Academy of Sciences*, 112(16):E2102–E2111.

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

Ramírez, M. A. (2025). *Banished Citizens: A History of the Mexican American Women Who Endured Repatriation*. Harvard University Press, Cambridge, MA.

Ruggles, S., Nelson, M. A., Sobek, M., Fitch, C. A., Goeken, R., Hacker, J. D., Roberts, E., and Warren, J. R. (2024). IPUMS ancestry full count data: Version 4.0 1850–1940. Minneapolis: IPUMS.

Sanders, N. J. and Stoecker, C. (2015). Where have all the young men gone? Using sex ratios to measure fetal death rates. *Journal of Health Economics*, 41:30–45.

Struyck, N. (1740). *Inleiding Tot de Algemeene Geographie, Benevens Eenige Sterrekundige En Andere Verhandelingen*. Issak Tirion, Amsterdam.

Tapia Granados, J. A. and Diez Roux, A. V. (2009). Life and death during the great depression. *Proceedings of the National Academy of Sciences*, 106(41):17290–17295.

Théré, C. and Rohrbasser, J.-M. (2006). Facing death in the early days of life: Inequality between the sexes in Enlightenment demographic thought. *The History of the Family*, 11(4):199–210.

U.S. Department of Commerce, Bureau of the Census (1934). Birth, stillbirth, and infant mortality statistics for the birth registration area of the United States, 1931. Technical Report 17th Annual Report, U.S. Government Printing Office, Washington, D.C.

U.S. Department of Commerce, Bureau of the Census (1936). Birth, stillbirth, and infant mortality statistics for the continental United States, the territory of Hawaii, the Virgin Islands: 1933. Technical Report Nineteenth Annual Report, Government Printing Office, Washington, DC.

Valente, C. (2015). Civil conflict, gender-specific fetal loss, and selection: A new test of the Trivers–Willard hypothesis. *Journal of Health Economics*, 39:31–50.

Vu, H. (2024). I wish I were born in another time: Unintended consequences of immigration enforcement on birth outcomes. *Health Economics*, 33(2):345–362.

Waldron, I. (1998). Sex differences in infant and early childhood mortality: Major causes of death and possible biological causes. In Lopez, A. D. and Ruzicka, L. T., editors, *Too Young to Die: Genes or Gender?*, number 155 in Population Studies, pages 64–83. United Nations, New York.

Wargentin, P. (1755). *Anmärkningar Om Nyttan Af Årliga Förtekningar På Födda Och Döda i et Land*. Kongl. Vetenskaps Academiens Handlingar För År 1755. Sjette Stycket, 16: 241–253.

Weber, J. (2013). Homing pigeons, cheap labor, and frustrated nativists: Immigration reform and the deportation of Mexicans from South Texas in the 1920s. *Western Historical Quarterly*, 44(2):167–186.

Wolf, J. H. (2001). *Don't Kill Your Baby: Public Health and the Decline of Breastfeeding in the Nineteenth and Twentieth Centuries*. Ohio State University Press, Columbus.

A1 Appendix A: Modeling Childhood Sex Ratios

The sex ratio can be modeled as the ratio of two Kaplan-Meier processes, with a certain number of males and females born, who then experience sex-specific death probabilities.[22] Consider the sex ratio, taken throughout as females per male, among survivors to age X, $SR_X = \frac{F_X}{M_X}$, where F_X and M_X are the female and male population counts. Let SR_B be the sex ratio of total births (including stillbirths and live births), sb_m and sb_f the male and female stillbirth probabilities, and q_m and q_f the male and female infant mortality probabilities. Because miscarriages cannot be measured empirically, we collapse them into the sex ratio of total births SR_B.[23]

The survival chain from conception to age X is multiplicative:

$$M_X = M_B\,(1 - sb_m)\,(1 - q_m)\,, \qquad F_X = F_B\,(1 - sb_f)\,(1 - q_f)$$

so the sex ratio among survivors factors as a product:

$$SR_X = \frac{F_X}{M_X} = SR_B \cdot \frac{1 - sb_f}{1 - sb_m} \cdot \frac{1 - q_f}{1 - q_m}$$

Taking the natural logarithm turns the product into a sum:

$$\ln SR_X = \ln SR_B + [\ln(1 - sb_f) - \ln(1 - sb_m)] + [\ln(1 - q_f) - \ln(1 - q_m)]$$

At the mortality levels of the 1920s United States the stage-specific losses are small enough that $\ln(1 - x) \approx -x$, which gives the interpretively transparent form:

$$\ln SR_X \approx \ln SR_B + (sb_m - sb_f) + (q_m - q_f) \tag{A1}$$

Because male losses exceed female at both stages, each parenthesized difference is positive and adds to $\ln SR_B$, raising the survivor ratio toward parity.

We then express this male excess mortality in relative terms. Let sb and q denote the total (both-sex) stillbirth and infant-mortality rates, and define excess male mortality as the male-female gap divided by the total rate, $\tau_{sb} = \frac{sb_m - sb_f}{sb}$ and $\tau_q = \frac{q_m - q_f}{q}$, so that $sb_m - sb_f = \tau_{sb} \cdot sb$ and $q_m - q_f = \tau_q \cdot q$. Substituting into equation (A1):

$$\ln SR_X \approx \ln SR_B + \tau_{sb} sb + \tau_q q$$

[22] See Klein and Moeschberger (2003, ch. 4) for a textbook explanation of this model.

[23] Miscarriages could be easily modeled as a third term in this decomposition, as could mortality at older ages. However, without data on miscarriages, we cannot test key assumptions, like the degree of excess male miscarriages relative to excess male stillbirths and infant mortality.

CDC vital statistics for the 1920s show these two penalties to be approximately equal, which motivates the explicit assumption $\tau_{sb} = \tau_q = \tau$. Under this restriction the two stages combine and the decomposition collapses to a single mortality term in the total rates:

$$\ln SR_X \approx \ln SR_B + \tau(sb + q) \tag{A2}$$

The female skew is then simply the common penalty τ scaled by the combined stillbirth and infant mortality probabilities, $sb + q$.

We verify our model using the 1931 Vital Statistics report, which gives the number of total live births by sex, number of stillbirths, and number of infant deaths for the USA Registration Area (U.S. Department of Commerce, Bureau of the Census, 1934). In Figure 7, we calculate the sex ratio of all births (adding live births and stillbirths, SRB), the sex ratio of live births (SRLB), and the implied sex ratio of survivors to age 1 (SR1). We see a clear pattern, where fetal-infant loss skews the survivor sex ratio towards females. Moreover, the slope from the SRB to the SRLB and that from SRLB to SR1 are roughly the same, supporting our assumption above of a constant excess male mortality. The slopes also appear similar across time, with different years having very similar slopes, again supporting our assumption of τ being constant. The non-white population appear to have female-skewed sex ratios regardless of fetal-infant loss. This is likely due to miscarriages, which are not captured in these data.

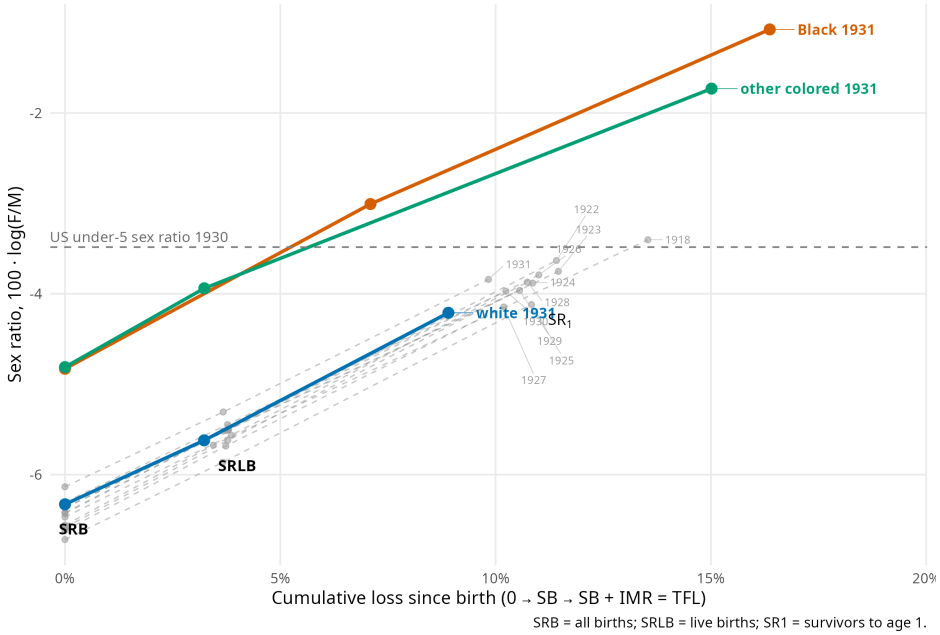


Figure 7: Attrition of the sex ratio from all births (SRB) to live births (SRLB) to survivors to age 1 (SR_1), United States, 1918-1931

Notes: The vertical axis is 100·log(F/M); horizontal position is cumulative loss since conception, the first segment giving the stillbirth (SB) and the second the infant-mortality (IMR) component, with the slopes the associated rise in the sex ratio. The dashed line marks the 1930 under-five sex ratio. Source: 1931 Vital Statistics Report, Table B (U.S. Department of Commerce, Bureau of the Census, 1934). 1930 sex ratio from IPUMS full count data, restricted to the same area as the vital statistics coverage.

A2 Estimating Birth-Registration Completeness by County, Ethnicity, and Birth Year, Texas 1926–1930

Here we explain how we estimate the completeness of birth registration in Texas at the birth-year $\times$ county $\times$ ethnicity level, where ethnicity is *Mexican* or *non-Mexican*, for birth cohorts 1926–1930. We combine data from the 1930 census with individual level birth registrations from Texas to create a novel dataset. Here describe the data sources, assumptions on mortality, estimator, and the identification of the Mexican-origin population. We follow the work of Eriksson et al. (2018).

A2.1 Motivation

Birth registration in the United States was incomplete well into the 1930s. The degree of under-registration varied systematically across states, urban-rural, and race. (Eriksson et al., 2018) document that births to African Americans in the rural South were unregistered at rates of roughly 15–25%. Texas suffered from poor registration, and it was the last state to the the Federal Birth Registration Area, which required a 90% coverage. Its large Mexican-origin population was precisely the kind of group most likely to be missed by local registrars.

We are interested in the degree of registration completeness itself, rather than just estimating a statewide adjustment factor. The reverse-survival estimator delivers exactly this: an independent estimate of the true number of births in each county–ethnicity–cohort cell, against which registered births can be compared.

A2.2 The reverse-survival estimator

We use the 1930 census as our 'ground truth' for the number of children born. The children of a given single-year age group enumerated in a census are the survivors of the births that occurred in the corresponding period. We then divide this count with the probaility of survival, to account for mortality. This gives an estimate of the actual number of births. We divide the number of births actually recorded by this estimate of births, yielding a completeness ratio. This follows a long line of work in demography and economic history (United Nations 1983; Preston, Heuveline, and Guillot 2001; Eriksson, Niemesh, and Thomasson 2018).

Let a cohort be indexed by county c, race r, and birth year t. Let N_{crt} be the number of children of that cohort enumerated in the 1930 census, let B^{reg}_{crt} be the number of births registered for that cohort, and let S_{crt} be the probability that a kid born into the cohort survives from birth to the census enumeration date. The census birth count is then

$$\hat{B}_{crt} = \frac{N_{crt}}{S_{crt}}, \tag{B3}$$

and registration completeness is

$$C_{crt} = \frac{B^{\text{reg}}_{crt}}{\hat{B}_{crt}} = \frac{B^{\text{reg}}_{crt}\, S_{crt}}{N_{crt}}. \tag{B4}$$

A value $C_{crt} \approx 1$ indicates effectively complete registration; $C_{crt} < 1$ indicates net under-registration; and the implied count of missed births is $\hat{B}_{crt} - B^{\text{reg}}_{crt}$. Values above one can arise from sampling noise in small cells, from over-registration (e.g., delayed registrations attributed to the cohort), or from errors in the survival or age inputs; however, this doesn't happen in our data.

A2.3 Data sources

A2.3.1 Registered births (numerator)

We get registered births by county, birth year, and ethnicity for 1926–1930 from Texas vital-registration records. They are organized as one observation per county × birth year × Mexican/non-Mexican cell (`births_n`). These counts are the numerator B^{reg}_{crt} in our completeness ratio. We use 3-digit FIPS codes to match counties to the census data. We classify births as Mexican or non-Mexican by the parental surname.

A2.3.2 Census child counts (denominator input)

We use the 1930 Census to get counts of children by birth cohort, N_{crt}. We get the Census microdata from IPUMS (Ruggles et al., 2024). We focus on Texas residents and children of completed ages 0–4. Then we aggregate person records to county × single-year-of-age × ethnicity cells. We use the birthplace to identify Mexicans in the 1930 census. We convert county ICPSR codes into FIPS codes, and construct year of birth from the age variable.

A2.3.3 Infant mortality (survival input)

We get infant mortality data from contemporary vital statistics. We have county-level infant mortality rates (IMR) for 1933 from the U.S. Census Bureau's *Birth, Stillbirth, and Infant Mortality Statistics: 1933* (published 1936). We build county-level figures as the sum of the county remainder (Section D, excluding incorporated places of 10,000+) and the corresponding large cities (Sections B and C). The resulting $\text{IMR}_c = 1000 \times (\text{infant deaths}_c)/(\text{registered births}_c)$ reflects the entire county. We attach county FIPS codes through the county-name crosswalk used in the 1939–40 vital-statistics file. Infant mortality is not available at the ethnicity-county level. We discuss the implications of this below.

A2.4 Mapping census age to birth cohorts

The 1930 census began enumeration on April 1, 1930 and was over before the end of the month. A child recorded at age a was therefore born, on average, a years earlier, spanning the interval from about April $(1930 - a - 1)$ to April $(1930 - a)$. We assign each age group to a nominal calendar birth year,

$$t = 1930 - a, \qquad a \in \{0, 1, 2, 3, 4\} \;\Rightarrow\; t \in \{1930, 1929, 1928, 1927, 1926\}, \tag{B5}$$

recognizing that each single-year age bin straddles two calendar years relative to a calendar-year registration count. This introduces a modest cohort-edge mismatch that is largest for the youngest (age 0, only part of calendar 1930 is complete at enumeration) and oldest (age 4, part of the cohort was born in 1925) groups. Because the mismatch is a fixed feature of the April enumeration date and is common across counties and races, it is unlikely to drive cross-sectional differences in estimated completeness; as robustness, the birth-year labeling can be shifted by one year or the cohorts interpolated to calendar years without changing our key results.

A2.5 Survival and the mortality schedule

A2.5.1 Survival from birth to enumeration

Birth cohorts are defined by age in April 1930, not exact age. We draw on Preston, Heuveline, and Guillot (2001) to model these cohorts and their survival. Let q_x denote the probability that a child alive at exact age x dies before exact age $x + 1$, for $x = 0, \ldots, 4$, and let

$$\ell_a = \prod_{j=0}^{a-1} \big(1 - q_j\big), \qquad \ell_0 = 1, \tag{B6}$$

be cumulative survival to exact age a. A child enumerated at completed age a has survived all full prior years of age and is, on average, partway through the current year a when they are enumerated. Our key probability is the survival from birth to census, which we write as

$$S_a = \ell_a \big(1 - \kappa_a \, q_a\big), \tag{B7}$$

where $\kappa_a \in (0, 1)$ is a separation factor capturing the fraction of within-year-a mortality already experienced by the average enumerated child of that age. For age 0 this reduces to $S_0 = 1 - \kappa_0 q_0$.

For most ages, the separation factor is simply one half, but not for the first year. The separation factor for infancy is greater than one-half because deaths are heavily concentrated in the first weeks of life, so that a child observed at age 0 in the census has, on average, already passed through

most of the first-year mortality risk (Preston, Heuveline, and Guillot 2001, ch. 3, discuss this at length). We set $\kappa_0 = 0.75$, consistent with the front-loaded infant-death schedule assumed in the county-level analysis and with the adjustment logic in Eriksson, Niemesh, and Thomasson (2018); for ages 1–4 mortality is spread more evenly within the year and we set the separation factor to one half: $\kappa_{1\text{–}4} = 0.50$. Both of these parameters can be adjusted. We do robustness checks for $\kappa_0 \in \{0.70, 0.85\}$ and $\kappa_{1\text{–}4} \in \{0.40, 0.60\}$. Because infant mortality is much greater than mortality at ages 1–4, the estimates are only weakly sensitive to these choices. ———Jesse stopped here 9/26

A2.5.2 Infant mortality q_0 from the observed IMR

We take the infant-mortality probability directly from the observed rate,

$$q_{0,cr} = \frac{\text{IMR}_c}{1000} \times m_r, \tag{B8}$$

where IMR_c is the 1933 county-total infant-mortality rate and m_r is a race multiplier discussed in Section A2.5.4. We use the 1933 IMR as a stand-in for the 1926–1930 rates. It is the earliest and only county-level infant-mortality series available for Texas. Holding q_0 fixed across 1926–1930 assumes no strong trend in infant mortality over the immediately preceding years.

A2.5.3 Deriving child mortality q_1–q_4

We do not have data on mortality between ages 1 and 5, so we derive it directly from infant mortality using a fixed reference schedule of child-to-infant mortality ratios:

$$q_{a,cr} = k_a \, q_{0,cr}, \qquad (k_1, k_2, k_3, k_4) = (0.160,\ 0.080,\ 0.055,\ 0.040). \tag{B9}$$

Our variation in mortality, therefore, comes from the observed q_0. We draw on model life tables for the ratio of child to infant mortality (Coale and Demeny 1983; United Nations 1983). Mortality falls steeply after infancy and continues to decline across early childhood. This constant ratio is a simplification. Typically, as mortality declines, infant mortality takes up a larger proportion of total mortality, meaning the childhood-to-infant mortality ratio decreases. Given the limited importance of post-infant mortality to our setting, we stay with this simplifying assumption, rather than using a complete model life table for each county-year-race cell.

A2.5.4 Race differential in mortality

Our IMR data are for the county total, and are not broken up by race. Mexican-origin infant mortality was likely higher than white infant mortality during this period. We assume equal survival,

which likely understates Mexican mortality, overstates Mexican census-implied births, and thereby biases estimated Mexican completeness upward. Thus, our estimated birth registration completeness for Mexicans is an upper bound.

A2.6 Identifying the Mexican and non-Mexican populations

We identify Mexican immigrant households in the 1930 census based on the place of birth of the household head. We then construct a list of surnames from these data, and calculate the probability of being Mexican origin for each name. For both the census counts and the birth registration data, we classify a child as 'Mexican' if their surname has a probability of at least 80% of being Mexican. This ensures that we have a consistent definition across the two data sets. The 1930 census did record "Mexican" as a separate race category, but there was systematic misclassification both regionally and by household characteristics Duncan and Trejo, 2025. Moreover, the birth registration data does not have a race, ethnicity, or parental birthplace field, leaving parental surname as the only useful information on whether a child is of Mexican origin.

A2.7 Estimation and aggregation

In order to calculate our estimator of birth registration completeness, we merge three datasets. First, the 1933 county-total IMR is converted to q_0 and joined to the census cells via FIPS. Within these county $\times$ race $\times$ age cell we use IMR to calculate the full survival schedule $\{q_0, \ldots, q_4\}$, the cumulative survivorship ℓ_a, the birth-to-census survival S_a, and the implied births $\hat{B} = N/S_a$. We merge these implied births with a dataset of registered births on county $\times$ race $\times$ birth year to form C_{crt}. We flag any cells which are missing one of these data sources. A census cell without registered births is an extreme case of under-registration, while registered births without any matched census implied births reflects coding gaps.

We report completeness at the cell level and also aggregate. Statewide (or subgroup) completeness is computed as a ratio of sums,

$$\bar{C}_{rt} = \frac{\sum_c B^{\text{reg}}_{crt}}{\sum_c \hat{B}_{crt}}, \tag{B10}$$

rather than as an unweighted mean of cell ratios, so that large counties receive appropriate weight and small-cell noise is not amplified. A binary small-cell flag (census count or registered births below 25) marks observations where sampling and rounding noise dominate.

References

Coale, Ansley J., and Paul Demeny. 1983. *Regional Model Life Tables and Stable Populations.* 2nd ed. New York: Academic Press.

Duncan, Brian, and Stephen J. Trejo. 2025. "Which Mexicans Are White? Enumerator-Assigned Race in the 1930 Census and the Socioeconomic Integration of Mexican Americans." *ILR Review.* (Earlier version: NBER Working Paper 31623, 2023.)

Eriksson, Katherine, Gregory T. Niemesh, and Melissa Thomasson. 2018. "Revising Infant Mortality Rates for the Early Twentieth Century United States." *Demography* 55(6): 2001–2024. (Earlier version: NBER Working Paper 23263, 2017.)

Gratton, Brian, and Myron P. Gutmann. 2000. "Hispanics in the United States, 1850–1990: Estimates of Population Size and National Origin." *Historical Methods* 33(3): 137–153.

Preston, Samuel H., Patrick Heuveline, and Michel Guillot. 2001. *Demography: Measuring and Modeling Population Processes.* Oxford: Blackwell.

Ruggles, Steven, et al. *IPUMS USA* [complete-count 1930 census microdata]. Minneapolis, MN: IPUMS. (Cite the version used.)

United Nations, Department of International Economic and Social Affairs. 1983. *Manual X: Indirect Techniques for Demographic Estimation.* Population Studies No. 81. New York: United Nations.

U.S. Bureau of the Census. 1936. *Birth, Stillbirth, and Infant Mortality Statistics: 1933.* Washington, DC: U.S. Government Printing Office.