The Birth of Mass Unauthorized Immigration in the United States[*]

Elizabeth U. Cascio

Ethan G. Lewis

Aryen Shrestha

December 13, 2025

Abstract

The Immigration and Nationality Act of 1965 imposed, for the first time in U.S. history, annual restrictions on lawful immigration from the Western Hemisphere. This paper introduces and implements a method for estimating highly granular local unauthorized rates in the wake of the surge in Latin American immigration that ensued. Using within-state, cross-county variation in these estimated unauthorized rates for foreign-born Mexican women and multiple comparison groups, we then estimate how maternal legal status affects the health of U.S. born – and thus U.S. citizen – children. We find that maternal authorization increases birth weight in the second generation.

I. Introduction

The long-term unauthorized population in the United States faces restrictions on jobs, travel, and program access that may inhibit not just their own economic mobility, but that of their U.S. born children. The issue is salient: in 2023, an estimated 4.6 million U.S. born children had unauthorized parents – nearly the highest number on record (Passel and Krogstad, 2025).[1] The intergenerational impacts of immigrant legal status are nevertheless understudied. In part, this may reflect the fact that legal status is not reported in U.S. administrative data, and self-reports of or proxies for status in U.S. survey data are measured with error. Even if observable, true status may also be correlated with unobserved determinants of child well-being.

This paper addresses both challenges, focusing on the decade that marked the rise of mass unauthorized immigration in the U.S. – the 1970s. The Immigration and Nationality Act of 1965 (INA) set the first-ever limits on lawful entry from the Western Hemisphere, opening the door for unauthorized immigration from Mexico and other Latin American countries.[2] Amendments to the INA in 1976 exacerbated this effect by extending country-specific annual quotas to the Western Hemisphere. This is illustrated in Figure 1, which plots lawful and total Mexican immigrants among U.S. residents in early 1980, by year of entry. A measure of unauthorized arrivals (Cascio and Lewis, 2025), the gap between total and lawful Mexican inflows began an explosive expansion in the late 1960s. Mexican arrivals to the U.S. in the late 1970s were more likely to be unauthorized than not.

Building on this logic, we introduce and implement a method for estimating highly granular local unauthorized rates in early 1980, in the wake of the post-INA surge in Mexican

---

[1] The unauthorized immigrant population in the United States is estimated to have reached a record 14 million in 2023, accounting for 27% of the foreign-born and 4.1% of the U.S. population overall (Passal and Krogstad, 2025).
[2] Mexico dominated immigration from Latin America during this period, accounting for half of immigrants from Latin America and 61 percent Latin American parents of U.S. born children in early 1980 (Ruggles et al., 2025).

unauthorized immigration. Using within-state, cross-county variation in these estimated

unauthorized rates for foreign-born Mexican women and multiple comparison groups (U.S. born

women, foreign-born Mexican women in 1970), we then estimate how maternal legal status

affects the health of U.S. born – and thus U.S. citizen – children. We find that maternal

authorization increases birth weight in the second generation.

The rise of mass unauthorized immigration thus left a mark on births in the U.S. – not

only increasing the foreign-born share but also lowering birth weight relative to a counterfactual

where these new arrivals were lawful residents. Data limitations prevent us from distinguishing

whether effects are driven by the unobservables of new Mexican arrivals or the conditions they

faced after entry. However, controls for maternal observables increase our estimates, suggesting

they reflect treatment more than selection.[3] The correlation between birth weight and school

achievement (Figlio et al., 2014) and educational attainment (Black, Devereux, and Salvanes,

2007, Royer, 2009) moreover suggests that the negative impacts of having unauthorized parents

are not confined to worse health at birth, an important consideration for future research.

II.   Background

The Mexican government long discouraged emigration to the U.S., fearing that

population loss would endanger its economic development. However, the state changed its

position and policies in the 1970s due to a combination of poor local economic conditions and

high population growth (Minian, 2018). Mexican inflows to the U.S. skyrocketed. Of the 2.74

million foreign-born Mexicans residing in the U.S. in 1980, more than half – 1.59 million – had

arrived in the 1970s, and the pace of arrival accelerated as the decade wore on.[4]

---

[3] This finding also argues against the idea that unauthorized status differentially induces childbearing among women who tend to have lower birth weight babies, i.e., increases negative selection into motherhood.
[4] These and the Census-based figures in Figure 1 are the authors' calculations from the 1980 Census 5% public-use microdata sample (Ruggles et al., 2025), assuming a 25% Mexican undercount (Borjas, Freeman, and Lang, 1991).

The legal environment in which Mexicans entered the U.S. in the 1970s was very different than it had been through the mid-1960s. Due to a combination of pressure from southwestern agricultural interests and foreign diplomacy (Yang, 2020), immigration from Mexico – and indeed the entirety of the Western Hemisphere – was not quota-restricted under the 1924 Immigration Act, unlike inflows from Europe and the rest of the world. Mexican arrivals were, however, not always welcomed or lawful after 1924: due to concerns over competition with native workers, Mexicans were repatriated in large numbers during the Great Depression, and from 1953 to 1954, "Operation Wetback" deported an estimated one million Mexicans, some with U.S. citizenship. Still, starting during World War II, Mexican men routinely came to the U.S. both temporarily and lawfully as agricultural contract labor ("Braceros"), establishing a pattern of circular migration for men from rural areas in central Mexico (Minian, 2018).

The situation changed starkly in the mid-1960s. Motivated again by concerns over domestic employment, President Johnson ended the Bracero program in 1964.[5] In 1965, the INA then established the first-ever quota (of 120,000 entrants annually) on immigration from the Western Hemisphere. The law also abolished the national-origins quota system imposed on the Eastern Hemisphere in 1924, replacing it with a higher worldwide limit on lawful immigration, a preference system favoring family reunification and skill, and 20,000-person annual country-specific limits. The preference system and the 20,000 annual limits were extended to Mexico and the rest of the Western Hemisphere in the 1976 amendments to the INA. The 1976 amendments also ended the labor certification exemption for parents of U.S. born children that had long applied to the Western Hemisphere (Legomsky and Gravo, 1976).

---

[5] Existing evidence suggests that these concerns were misplaced: neither the Bracero exclusion (Clemens, Lewis, and Postel, 2018) nor Mexican repatriations during the Great Depression (Lee, Peri, and Yasenov, 2022) increased the employment of U.S. born workers.

The combination of increasing Mexican inflows and more restrictions on lawful pathways to entry meant that most of the new Mexican arrivals by the late 1970s did not have authorization. Indeed, taking the gap between total and lawful arrivals as an estimate of unauthorized flows, the counts plotted in Figure 1 suggest that 61% of 1975-79 Mexican arrivals residing in the U.S. in 1980 were unauthorized,[6] compared to 45% for 1970-74 arrivals, 30% for 1965-69 arrivals, and 10% for 1960-64 arrivals. Overall, 37% of foreign-born Mexican U.S. residents in 1980 were in the country unlawfully. While there were not great differences in Mexican unauthorized rates across gender (Appendix Figure 1), some areas had higher unauthorized rates than others. The next section formalizes how we detect this variation.

III. Data and Sample

A. Unauthorized rates

Formally, we define the unauthorized rate as one minus the authorized rate, where the authorized rate is the share of the total foreign-born population that is naturalized or otherwise lawfully resident in the U.S., either permanently (via a Green Card) or temporarily (via a visa). For example, for members of foreign-born group j born in year y residing in county c in year t, the unauthorized rate is given by:

$$(1) \qquad u_{ct}^{jy} = 1 - nat_{ct}^{jy} - olr_{ct}^{jy},$$

where nat_{ct}^{jy} is the naturalization rate and olr_{ct}^{jy} is the (non-citizen) lawful resident rate. In principle, unauthorized rates can thus be estimated at a highly granular level.

We estimate the components of equation (1) for Mexican women ($j = MEXF$) in $t = 1979$ using several data sources. The numerator of the lawful resident rate comes from the *Alien*

---

[6] The "other lawful residents" figure for 1975-79 Mexican arrivals may exceed expectations based on a 20,000 annual quota on Mexico since some categories of non-citizen lawful entrants are not quota restricted. For example, naturalized citizens can sponsor certain categories of migrants (like spouses and parents) in unlimited numbers.

Address Reports 1980 Public Use File (U.S. Department of Justice, 1992), which gives country

of origin, sex, age, year of entry, and zip code of residence (which we map to counties) for lawful

residents in the U.S. in 1979.[7] Estimates of the denominator of the lawful resident rate, as well

as estimates of the 1979 naturalization rate, derive from published tabulations and public-use

microdata from the 1980 Census (Manson et al., 2024, Ruggles et al., 2025).[8]

Figure 2 shows that 1979 unauthorized rates were generally rising across cohorts of

foreign-born Mexican women of childbearing age. The peak (of approximately 45%) occurred

among women born in the late 1950s, the most represented cohorts among foreign-born Mexican

women residing in the U.S. as of early 1980 (Appendix Figure 2). These patterns reflect the

increase in unauthorized rates across entry cohorts into the 1970s (Figure 1) and the relative

youth of new arrivals.

To estimate the unauthorized rate for Mexican mothers giving birth in county c in 1979,

we aggregate estimates of $u_{c,1979}^{MEXF,y}$ from equation (1) across cohorts to the county level using

weights constructed from 1979 birth records (described in Section II.B). That is,

$$
(2) \qquad\qquad u_{c,1979}^{MEXF} = \sum_y \omega_{c,1979}^{MEXF,y} u_{c,1979}^{MEXF,y},
$$

where $\omega_{c,1979}^{MEXF,y}$ is the share of births in county c in 1979 to Mexican mothers born in year y.

B.   Birth outcomes

Outcomes come from Vital Statistics Natality Detail Data, which contain individual birth

records for either the universe or a 50% random sample of births, depending on the state of

occurrence and year. We focus on birth weight (in grams), the natural log of birth weight, and

---

[7] While the data release is technically for 1980, very few visa holders report entering in 1980 relative to 1979 or earlier: in this subsample, only 0.27% entered the U.S. in 1980, compared to 3.52% in 1979. In the subsample of interest – Mexican women born 1935 to 1965 – 97.3% were lawful permanent residents or held Green Cards.
[8] See the appendix for details on these data and our calculations. Note that the 1980 Census was taken on April 1, 1980, making it comparable to the *Alien Address Reports 1980 Public Use Data* in terms of timing.

indicators for different parts of the birth weight distribution.[9] We collapse these outcomes to the

county level, separately by child year of birth for foreign-born Mexican and U.S. born mothers.[10]

In addition to county of residence at the time of birth and maternal country of origin, the birth

records include maternal age, which we use as a control variable and to estimate maternal year of

birth. They also include child sex and parity, which we also use to construct control variables.[11]

We restrict attention to Natality Detail Data from two years: 1979 and 1970. Outcomes

from 1979 match the estimated unauthorized rates in equation (2). Outcomes from 1970 – the

first year in which foreign-born mothers are identified in Natality Detail Data – are used in a

specification check, then as an additional difference in the research design (see Section III).

C.   *Sample and descriptive statistics*

To define the analysis sample, we start with the top 100 counties as ranked by the number

of births to foreign-born Mexican women in 1970. From there, we first limit attention to states

where there are at least four counties in this group, to facilitate within-state estimation. We then

drop two counties that have inadequate data and a third county – Los Angeles (LA) – whose size

and high unauthorized rate make it exceedingly influential in weighted estimation. Our final

analysis sample includes 79 counties across 5 states: California, Texas, Illinois, Arizona, and

New Mexico. The included counties represent nearly 84% of 1970 births to Mexican mothers in

the U.S., exclusive of LA county. More details, including a county list, are in the appendix.

Appendix Table 1 gives descriptive statistics for our sample.[12] Mexican mothers gave

birth to heavier babies than U.S. born mothers in both 1970 and 1979, but the gap in birth weight

---

[9] Other characteristics commonly explored in studies using Natality Detail Data, like frequency and timing of prenatal care visits and maternal educational attainment, are not reported for all states during the 1970s.
[10] Three foreign countries of origin were identified in Natality Detail Data in the 1970s: Canada, Cuba, and Mexico.
[11] We accessed these data through the National Vital Statistics System repository at the National Bureau of Economic Research (https://www.nber.org/research/data/vital-statistics-natality-birth-data).
[12] Throughout, we weight by the group-specific county share of 1970 births in the sample. Within each group (Mexican mothers, U.S. born mothers), we thus give more weight to counties with more representation in 1970. In

narrowed over the decade, as unauthorized immigration expanded. The average Mexican woman giving birth in 1979 resided in a county where 31% of Mexican women of the same age were unauthorized, but there was substantial variation in unauthorized rates, even within state.[13]

III. Results

County-level unauthorized rates in 1979 were not randomly assigned (Appendix Table 2). Our empirical approach must therefore address potential bias from correlation between 1979 unauthorized rates and unobserved determinants of contemporaneous birth outcomes for Mexican mothers. We attempt to do this by incorporating data on birth outcomes for U.S. born mothers in the same counties at the same time, as well as data from 1970.

We begin with the following difference-in-differences specification for 1979 births:

$$(3) \qquad bw_{jc(s)} = \theta^{DD} u_{c,1979}^{MEXF} \times D_j^{MEXF} + \delta_c + \gamma_{js} + \varepsilon_{jc(s)},$$

where $bw_{jc(s)}$ represents an average birth weight outcome for children born in 1979 to mothers of group j residing in county c (in state s); $u_{c,1979}^{MEXF}$ is the 1979 unauthorized rate, defined above; and D_j^{MEXF} is a dummy variable set to one if group j consists of births to Mexican mothers, zero if births to U.S. born mothers. Pooling data for the two groups, we can include county fixed effects, δ_c, which remove bias from shared county-specific shocks to birth weight. To account for local but group-specific shocks, the model also includes group-by-state fixed effects, γ_{js}.

The first column of Table 1 reports the least squares estimate of θ^{DD}, the parameter of interest in equation (3), from a model where birth weight (in grams) is the outcome variable.

---

any combined sample, however, each group overall gets equal weight. We use 1970 data to construct weights, rather than 1979 data, because fertility could be endogenous to unauthorized immigration.

[13] A regression of county unauthorized rates on state indicators (weighted by 1970 births to foreign-born Mexican women) yields an R^2 of 0.5276. Thus, about half of the variation in county-level 1979 unauthorized rates is within state. Counties in California and Illinois on average have significantly higher unauthorized rates: 44.4% and 48.4%, respectively, compared to 21.5%, 23.9%, and 16% in Arizona, New Mexico, and Texas.

Mechanically, this estimate captures how much more, on average, the 1979 unauthorized rate affected 1979 birth outcomes for Mexican mothers than U.S. born mothers in the same state. In the absence of spillovers, this estimate may alternatively be interpreted as the effect of maternal unauthorized status on birth weight at the level of the individual child. Among Mexican mothers, having authorization to be in the U.S. increases birth weight by 115 grams on average.

For this estimate to have a causal interpretation, unobservables cannot vary differently with $u_{c,1979}^{MEXF}$ across groups. Column 2 of Table 1 re-estimates model (3) for births in 1970, before unauthorized immigration from Mexico had significantly grown (Figure 1), but when unobserved local drivers of differential outcomes across Mexican and U.S. born mothers may have already been present. The estimate of θ^{DD} is much smaller, at an insignificant 8 grams. Columns 3 and 4 of Table 1 show that the difference in the column 1 and 2 coefficients – a triple difference estimate – is not just statistically significant, but also largely insensitive to controls for maternal age and indicators for child sex and parity – strong predictors of birthweight.[14] To the extent that maternal unobservables and observables are positively correlated, this finding suggests that selection on unobservables is not biasing our estimates upward in magnitude.[15]

Table 2 Panel A repeats the specification in the last column of Table 1 for the county-level average of the natural log of birth weight (column 2) and low birth weight (<2500 grams) and high birth weight (>4000 gram) shares (columns 3 and 4). Though robust to functional form, the estimates imply that effects are driven by reductions in birth weight in the upper end of the

---

[14] We allow the coefficients on these controls to vary both over time and across groups. The data strongly reject the restriction of equality of coefficients ($p = 0.0034$), mostly due to significant heterogeneity in coefficient controls across groups in 1970 ($p = 0.0065$).

[15] We also estimate model (3) using predicted birth weight as the outcome, where the prediction is based on group-specific relationships in 1970 between birth weight and a cubic in maternal age, indicators for second, third, and fourth or higher parity, and child sex. A benefit of this approach is that it does not rely on possibly treatment-induced changes in the relationship between birth weight and maternal and child observables. Estimates of θ^{DD} (and θ^{DDD} in model (4)) for predicted birth weight are positive (and statistically significant in the case of θ^{DDD}), suggesting that higher unauthorized shares shift birth composition to favor heavier infants (Appendix Table 3).

distribution. This conclusion is echoed in Figure 3, which presents estimates for the share of 1979 births in each quartile of the 1970 Mexican birth weight distribution. The hollow dots, corresponding to the Panel A specification, imply that some of upper-tail reductions in birth weight occurred below the high birth weight threshold.

The remaining panels of Table 2 present estimates from alternative triple-difference specifications, based on the model (additional controls suppressed):

$$(4) \qquad bw_{jc(s)t} = \theta^{DDD} u_{c,1979}^{MEXF} \times D_j^{MEXF} \times T_t^{1979} + \delta_{ct} + \gamma_{jst} + \gamma_{jc} + \varepsilon_{jc(s)t},$$

where T_t^{1979} is a dummy equal to one if $t = 1979$, zero if $t = 1970$, and δ_{ct}, γ_{jst}, and γ_{jc} are vectors of county-by-year, group-by-state-by-year, and group-by-county fixed effects, respectively. The γ_{jc} replace the direct effect of $u_{c,1979}^{MXF} \times D_j^{MXF}$ in the otherwise equivalent model that is presented in Panel A. A benefit of the γ_{jc} is that they control much more flexibly for sources of bias at the county and group level that are fixed over time.

Estimates of θ^{DDD} from model (4), shown in Panel B and as the solid dots of Figure 3, are mostly larger in magnitude, suggesting some downward bias in the baseline specification. For example, being authorized to be in the U.S. increases birth weight for the children of Mexican mothers by 177 grams (column 1) and 5.3% (column 2), up from the 134 grams and 4% in Panel A. Adding group- and time-varying effects of 1970 county characteristics to model (4), as shown in Panel C and the solid triangles in Figure 3, raises the coefficients even further, though the additional controls do not add significantly to the explanatory power of the model.

IV.   Conclusion

Using cross-county variation in unauthorized inflows from Mexico in the years following the INA, this paper has presented novel estimates of the impact of parental legal status on child health. The magnitude of our findings is significant, more than large enough to explain the

narrowing gap in birth weight between the children of Mexican and U.S. born women during the 1970s.[16] Whether effects of this magnitude continue to the present day is an open question.

While our research design aimed to uncover a causal effect, limitations on the available data meant that we could not exploit variation from before versus after the INA. Unlike the Immigration Reform and Control Act of 1986, studied in Cascio, Cornell, and Lewis (2024), the INA also did not change the status of immigrants already resident in the U.S., but rather altered the legal environment in which they would have entered. We therefore cannot rule out that our estimates reflect a change in the characteristics of migrants at the time of entry, though the (lack of) sensitivity to controls for maternal characteristics suggests an interpretation based more on treatment than selection. To the extent that conditions in the U.S. contributed to poor birth outcomes in the second generation, there are also many potential mechanisms – limited job opportunities, stress, low take-up of public benefits, among others – that require additional data and analysis to explore.[17] We leave that exploration to future research.

---

[16] Combined with the average unauthorized rate in the sample (31%), the estimates in Panel A suggest that, on average, unauthorized immigration lowered the birth weight for Mexican mothers by 41.5 grams (134 x 0.31). By contrast, the gap in Mexican-U.S. born mother gap in birth weight fell by about 25 grams between 1970 and 1979 (Appendix Table 1).

[17] Aggressive immigration enforcement has been shown to lower both benefits take-up in mixed status households (Watson, 2014, Alsan and Yang, 2024) and birth weight (Strully et al., 2020, Tome et al., 2021).

References

Alsan, Marcella and Crystal Yang. 2024. "Fear and the Safety Net: Evidence from Secure Communities." *Review of Economics and Statistics* 106(6): 1427–1441.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2007. "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes." *Quarterly Journal of Economics* 122(1): 409-439.

Borjas, George J., Richard B. Freeman, and Kevin Lang. 1991. "Undocumented Mexican-Born Workers in the United States: How Many, How Permanent?" In John M. Abowd and Richard B. Freeman (eds.) *Immigration, Trade, and the Labor Market*. Chicago: University of Chicago Press.

Cascio, Elizabeth U., Paul Cornell, and Ethan G. Lewis. 2024. "The Intergenerational Effects of Permanent Legal Status." NBER Working Paper 32635. Cambridge, MA: National Bureau of Economic Research.

Cascio, Elizabeth and Ethan Lewis. 2025. "Opening the Door: Immigrant Legalization and Family Reunification in the United States." *Journal of Labor Economics* 43(4).

Clemens, Michael A., Ethan G. Lewis, and Hannah M. Postel. 2018. "Immigration Restrictions as Active Labor Market Policy: Evidence from the Bracero Exclusion." *American Economic Review* 108(6): 1468-87.

Figlio, David, Jonathan Guryan, Krzysztof Karbownik, and Jeffrey Roth. 2014. "The Effects of Poor Neonatal Health on Children's Cognitive Development." *American Economic Review* 104(12): 3921-3955.

Lee, Jongkwan, Giovanni Peri, and Vasil Yasenov. 2022. "The Labor Market Effects of Mexican Repatriations: Longitudinal Evidence from the 1930s" *Journal of Public Economics* 205, 104558.

Legomsky, Stephen H. and Mitchell D. Gravo. 1976. "Afterword: The Immigration and Nationality Act Amendments of 1976." *San Diego Law Review* 14: 326-335.

Manson, Steven, Jonathan Schroeder, David Van Riper, Katherine Knowles, Tracy Kugler, Finn Roberts, and Steven Ruggles. 2024. IPUMS National Historical Geographic Information System: Version 19.0 [dataset]. Minneapolis, MN: IPUMS.

http://doi.org/10.18128/D050.V19.0

Minian, Ana Raquel. 2018. *Undocumented Lives: The Untold Story of Mexican Migration.* Cambridge: Harvard University Press.

Passel, Jeffrey S. and Jens Manuel Krogstad. 2025. "U.S. Unauthorized Immigrant Population Reached a Record 14 Million in 2023." https://www.pewresearch.org/race-and-ethnicity/2025/08/21/u-s-unauthorized-immigrant-population-reached-a-record-14-million-in-2023/ (Accessed September 4, 2025).

Royer, Heather. 2009. "Separated at Girth: U.S. Twin Estimates of the Effects of Birth Weight." *American Economic Journal: Applied Economics* 1(1): 49-85.

Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Grace Cooper, Julia A. Rivera Drew, Stephanie Richards, Renae Rodgers, Jonathan Schroeder, and Kari C.W. Williams. 2025. *IPUMS USA: Version 16.0* [dataset]. Minneapolis, MN: IPUMS.

https://doi.org/10.18128/D010.V16.0

Strully, Kate W., Robert Bozick, Ying Huang, and Lane F. Burgette. 2020. "Employer Verification Mandates and Infant Health." *Population Research and Policy Review* 39: 1143-1184.

Tome, R., M.A. Rangel, C.M. Davis, and L. Bellows. 2021. "Heightened Immigration Enforcement Impacts U.S. Citizens' Birth Outcomes: Evidence from Early ICE Interventions in North Carolina." *PLoS ONE* 16(2): e0245020.

United States Department of Justice. Immigration and Naturalization Service. Alien Address Reports, [United States]: 1980 Public Use File. Inter-university Consortium for Political and Social Research [distributor], 1992-02-16. https://doi.org/10.3886/ICPSR07998.v1

Watson, Tara. 2014. "Inside the Refrigerator: Immigration Enforcement and Chilling Effects in Medicaid Participation." *American Economic Journal: Economic Policy* 6(3): 313-338.

Yang, Jia Lynn. 2020. *One Mighty and Irresistible Tide: The Epic Struggle over American Immigration, 1924-1965*. New York, NY: W.W. Norton & Company.
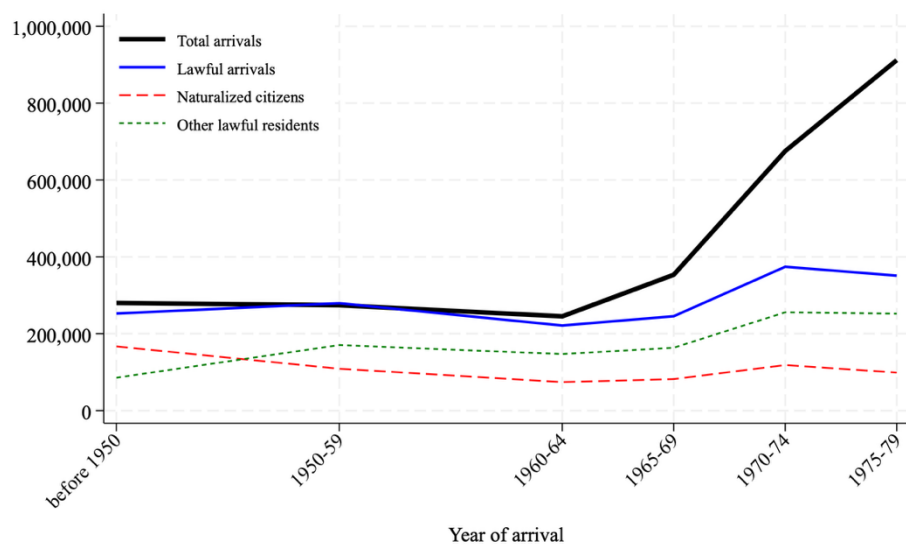
FIGURE 1. TRENDS IN IMMIGRATION FLOWS, BY LEGAL STATUS: 1980 U.S. RESIDENTS BORN IN MEXICO

Note: Total arrivals and naturalized citizens are calculated from 1980 Census 5% public use microdata (Ruggles et al., 2025). Other lawful residents (Green Card and temporary visa holders) are drawn from the *Alien Address Reports 1980 Public Use File* (U.S. Dept. of Justice, 1992). Total arrivals are the sum of naturalized citizens and other lawful residents. Data are restricted to U.S. residents born in Mexico.
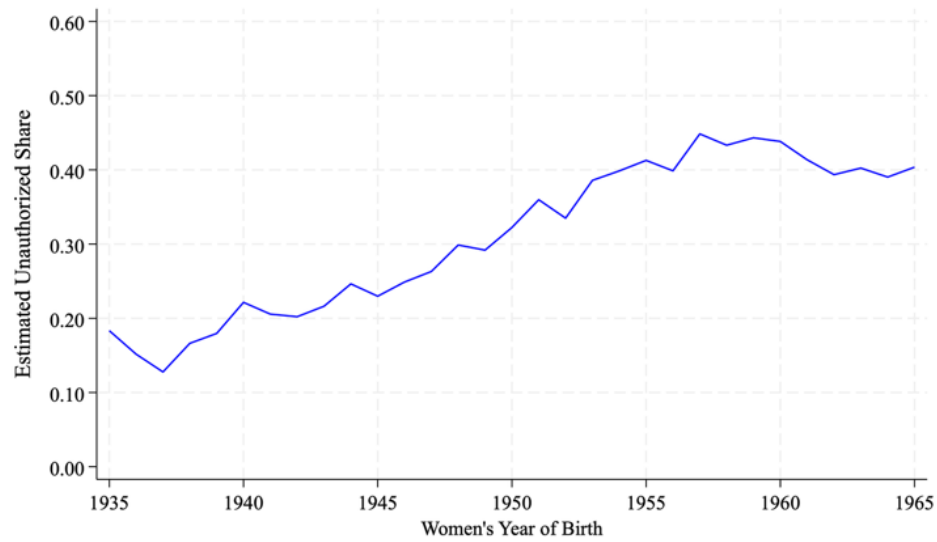
FIGURE 2. TRENDS IN THE ESTIMATED UNAUTHORIZED RATE, BY BIRTH COHORT: 1980 U.S. RESIDENT WOMEN BORN IN MEXICO

Note: Unauthorized rates calculated based on equations (1) and (2). Data are from 5% public use microdata (Ruggles et al., 2025) and published tabulations (Manson et al., 2024) of the 1980 Census and the *Alien Address Reports 1980 Public Use File* (U.S. Department of Justice, 1992). See appendix for further details.
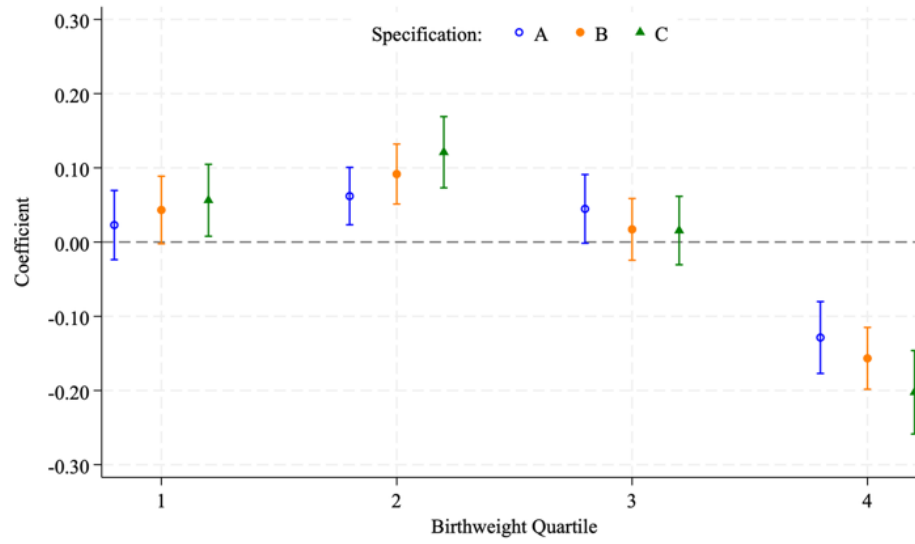
FIGURE 3. TRIPLE DIFFERENCE ESTIMATES BY QUARTILE OF THE 1970 MATERNAL MEXICAN BIRTHWEIGHT DISTRIBUTION

Note: Figure plots triple-difference coefficients that align with the specifications in Panels A-C of Table 2. See Table 2 notes for details. Outcomes are indicators for quartiles of the 1970 distribution of birthweight for Mexican mothers in the counties of our analysis sample. The 2nd, 3rd, and 4th quartiles begin at 3062, 3374, and 3686 grams, respectively.

TABLE 1 – DIFFERENCE-IN-DIFFERENCES AND TRIPLE-DIFFERENCE ESTIMATES OF THE EFFECT OF MATERNAL
UNAUTHORIZED STATUS ON BIRTH WEIGHT (IN GRAMS)

| | DD: 1979 | DD: 1970 | DDD: 1979 - 1970 | |
	(1)	(2)	(3)	(4)
Coefficient	-114.9	8.18	-123.0	-134.0
	(45.7)	(50.3)	(46.8)	(53.1)
Observations	158	158	316	316
R-squared	0.880	0.881	0.892	0.927
County fixed effects	Y	Y	Y	Y
State x *MEXF* fixed effects	Y	Y	Y	Y
State x *MEXF* x 1979 fixed effects	N	N	Y	Y
County x 1979 fixed effects	N	N	Y	Y
birth characteristics x *MEXF* x 1979	N	N	N	Y
p: birth characteristics interactions				0.0034

Note: Columns (1) and (2) present least squares estimates of θ^{DD} from model (3) using Natality Detail Data from 1979 and 1970, respectively. Column (3) presents the difference between the column (1) and column (2) estimates, and column (4) regression adjusts that difference for year- by group-specific effects of birth characteristics (maternal age and indicators for child sex and parity). *MEXF*=1 for children born to Mexican women, and *MEXF*=0 for children born to native women. Regression estimates are weighted by the group-specific county share of 1970 births in the sample, and standard errors (in parentheses) are clustered on county. Sample consists of births to Mexican and U.S. born mothers residing in one of 79 counties across five states: California, Texas, Illinois, Arizona, and New Mexico. See text and appendix for more details. DD=difference-in-differences coefficient, DDD=triple differences coefficient.

Dependent variable:	Birth weight (in grams) (1)	ln (Birth weight) (2)	Low birth weight (3)	High birth weight (4)
Panel A. County and state-by-year fixed effects				
$u_{c,1979}$ x *MEXF* x 1979	-134.0	-0.040	-0.0045	-0.069
	(53.1)	(0.016)	(0.0227)	(0.022)
R-squared	0.927	0.928	0.866	0.892
p: birth characteristics interactions	0.003	0.004	0.083	0.000
Panel B. Add county-by-group fixed effects				
$u_{c,1979}$ x *MEXF* x 1979	-176.7	-0.053	0.0088	-0.079
	(44.1)	(0.013)	(0.0191)	(0.018)
R-squared	0.985	0.985	0.954	0.975
p: birth characteristics interactions	0.003	0.004	0.004	0.013
Panel C. Add 1970 county characteristics x MEXF x 1979				
$u_{c,1979}$ x *MEXF* x 1979	-209.7	-0.063	-0.0158	-0.098
	(52.2)	(0.016)	(0.0191)	(0.026)
R-squared	0.987	0.987	0.962	0.977
p: birth characteristics interactions	0.000	0.000	0.013	0.250
p: 1970 county chars x *MEXF* x 1979	0.423	0.436	0.163	0.345

Note: Panel A repeats the specification in Table 1 column (4). Panel B substitutes county-by-group fixed effects for the direct effect of $u_{c,1979}$ x *MEXF* (mode1 (4)), and Panel C adds year and group-varying effects of 1970 county characteristics. Year corresponds to child's year of birth. See Table 1 notes for further details.

Online Appendix for

The Birth of Mass Unauthorized Immigration in the United States

by

Elizabeth U. Cascio, Ethan G. Lewis, and Aryen Shrestha

December 13, 2025

A1. Treatment Definition

A1.1 Unauthorized Share of Foreign-Born Mexican Mothers

Our treatment variable is $u_{c(s),1979}^{MEXF}$ – the unauthorized share among foreign-born Mexican women residing in U.S. county c in state s and giving birth in 1979.[1] We estimate this as:

$$(A1) \qquad u_{c(s),1979}^{MEXF} = \sum_y \omega_{c,1979}^{MEXF,y} u_{c(s),1979}^{MEXF,y},$$

where $u_{c(s),1979}^{MEXF,y}$ is the 1979 unauthorized share among foreign-born Mexican women born in year y and residing in county c in state s. We restrict y to be between 1935 and 1965 (inclusive), which keeps only women between the ages of 15 and 45 in our sample.[2] Construction of this variable is described in Appendix Section A1.2.

$\omega_{c,1979}^{MEXF,y}$ is then the share of U.S. born children in 1979 who have foreign-born Mexican mothers, in county c and birth cohort y, relative to the total number of such children in county c across the 1935 to 1965 maternal birth cohorts of foreign-born Mexicans.[3] That is, $\omega_{c,1979}^{MEXF,y} = N_{c,1979}^{MEXF,y} / \sum_{y \in [35,65]} N_{c,1979}^{MEXF,y}$, where $N_{c,1979}^{MEXF,y}$ is the count of children born in 1979 to foreign-born Mexican women (themselves born in year y) residing in county c. Counts of U.S. born children come from Natality Detail Data, which contain individual birth records for either the universe or a 50% random sample of births, depending on state and year.[4] Equation (A1) thus estimates the unauthorized rate across all foreign-born Mexican women giving birth in 1979 in county c by taking a weighted average of county- and cohort-specific unauthorized rates.

---

[1] For expository ease, we include the "s" subscript in this appendix only on terms where state-level data are relevant in their estimation.

[2] These ages represent, respectively, the 1st and 99th percentiles of maternal age for foreign-born Mexican women giving birth in the U.S. in 1979.

[3] Thus, by construction, $\sum_y \omega_{c,1979}^{MEXF,y} = 1$ for all c.

[4] All states in our sample have 50% samples in 1970. In 1979, IL and TX data are the universe of births, but data for the remaining three states (AZ, CA, NM) remain 50% random samples.

A1.2 Estimating County- and Cohort-Specific Unauthorized Rates

The 1979 unauthorized share of foreign-born Mexican women born in year *y* residing in county *c* in state *s*, $u_{c(s),1979}^{MEXF,y}$, is defined as a residual, i.e., as one minus the authorized share for this group. We estimate the authorized share for this group as the sum of two quantities: (i) the share of this group who are naturalized citizens ($nat_{c(s),1979}^{MEXF,y}$); and (ii) the share of this group who are otherwise lawful residents ($olr_{c(s),1979}^{MEXF,y}$):

(A2)
$$u_{c(s),1979}^{MEXF,y} \approx 1 - \left(nat_{c(s),1979}^{MEXF,y} + olr_{c(s),1979}^{MEXF,y} \right).$$

Other Lawful Residents

The lawful resident share of this group is given by

(A3)
$$olr_{c(s),1979}^{MEXF,y} \equiv \frac{OLR_{c,1979}^{MEXF,y}}{POP_{c(s),1979}^{MEXF,y}},$$

where $OLR_{c,1979}^{MEXF,y}$ is the <u>number</u> of female Mexican Green Card and temporary visa holders (and $olr_{c(s),1979}^{MEXF,y}$ is the rate) born in year *y* and residing in county *c* in 1979, and $POP_{c(s),1979}^{MEXF,y}$ is the <u>number</u> of foreign-born Mexican females in the same cohort and county, also in 1979. We obtain $OLR_{c,1979}^{MEXF,y}$ using the *Alien Address Reports 1980 Public Use File* (U.S. Department of Justice, 1992), which record the sex, country of origin, year of birth, and zip code of residence (which we code to counties) for all foreign-born migrants lawfully resident in the U.S. in early 1980.

We obtain $POP_{c(s),1979}^{MEXF,y}$ via approximation from the Census taken April 1, 1980:[5]

(A4)
$$POP_{c(s),1979}^{MEXF,y} \approx POP_{c,1980}^{HISPF,y} \left(\frac{POP_{s,1980}^{MEX,y}}{POP_{s,1980}^{HISP,y}} \right) \left(\frac{\frac{POP_{c,1980}^{MEX}}{POP_{c,1980}^{HISP}}}{\frac{POP_{s,1980}^{MEX}}{POP_{s,1980}^{HISP}}} \right),$$

---

[5] Throughout, we also adjust for a projected 25% Census undercount of Mexicans in the 1980 Census (Borjas, Freeman, and Lang, 1991).

where $POP_{c,1980}^{HISPF,y}$ is the number of Hispanic women in county c born in year y in 1980, from Census tabulations distributed by NHGIS (Manson, et al., 2024). The second term represents the Mexican share among Hispanics in state s in birth cohort y, which we estimate from the 5% public-use microdata sample (PUMS) of the 1980 Census (Ruggles et al., 2025). This term thus scales down Hispanic female population in the county and cohort by the Mexican share among Hispanics in that cohort in the state as a whole. To account for the fact that some counties have higher foreign-born Mexican population shares than others, the final term then scales by the share of foreign-born Mexicans among Hispanics at the county level relative to the state level, using county-level figures reported by NHGIS (Manson, et al., 2024).[6]

Naturalized Citizens

Lastly, the naturalized share of this group is defined as:

$$(A5) \qquad nat_{c(s),1979}^{MEXF,y} \equiv \frac{NAT_{c(s),1979}^{MEXF,y}}{POP_{c(s),1979}^{MEXF,y}}$$

where $NAT_{c(s),1979}^{MEXF,y}$ is the number of female Mexican naturalized citizens born in year y and residing in county c. Unlike $OLR_{c(s),1979}^{MEXF,y}$, $NAT_{c(s),1979}^{MEXF,y}$ is not observable in administrative data. In this case, we estimate the rate directly rather than separately estimating its components:

$$(A6) \qquad nat_{c(s),1979}^{MEXF,y} \approx nat_{s,1980}^{MEXF,y} \left(\frac{\frac{NAT_{c,1980}^{HISP,y}}{POP_{c,1980}^{HISP,y}}}{\frac{NAT_{s,1980}^{HISP,y}}{POP_{s,1980}^{HISP,y}}} \right)$$

The first term is the naturalized share among foreign-born Mexican women born in year y residing in early 1980 in state s, which we estimate from the 5% 1980 Census PUMS (Ruggles, et al., 2025). The second term accounts for the fact that naturalization rates vary across counties;

---

[6] Census tabulations are based on larger samples than public-use microdata samples, so are preferable to use where available.

because we cannot observe county-level naturalization rates for Mexicans (even overall), we scale by the Hispanic naturalization rate at the county relative to the state level, using county-level naturalized citizen and population counts reported by NHGIS (Manson, et al., 2024).

A2.   Sample Definition

We first restrict our sample to the top 100 U.S. counties ranked by the number of births to foreign-born Mexican mothers in 1970. Out of the total 45,060 such births, these counties account for 92.4%. From this group, we drop Los Angeles County because it is an extreme outlier: in 1970, it alone had 15,212 births to foreign-born Mexican mothers (33.8% of such births in the top-100 sample),[7] and an estimated unauthorized rate nearly twice the average as in our ultimate estimation sample.

We then drop states with fewer than four counties in the sample. This yields 81 counties across five states: California, Texas, Arizona, New Mexico, and Illinois. Together, these represent 85.1% of births to foreign-born Mexican mothers in 1970 (excluding LA County). Finally, we drop two counties for which we cannot compute $u_{c(s),1979}^{MEXF,y}$ for *all* maternal birth cohorts.[8] This reduces the sample to 79 counties, accounting for 84.4% of births to foreign-born Mexican mothers in 1970 (excluding LA). Our final sample consists of 79 counties across 5 states.[9]

---

[7] The next largest county (Harris County, IL) accounted for only about 6%.

[8] The two dropped counties are both in New Mexico. De Baca county reports that $POP_{c,1980}^{HISPF,y} = 0$ for three maternal birth cohorts (y), and Los Alamos county reports that $POP_{c,1980}^{MEX} = 0$, which affects the construction of the estimated unauthorized rate across all maternal birth cohorts y (equation (A4)).

[9] They are, in **Arizona** (6): Cochise, Maricopa, Pima, Pinal, Santa Cruz, Yuma; **California** (32): Alameda, Colusa, Contra Costa, Fresno, Imperial, Kern, Kings, Madera, Merced, Monterey, Orange, Placer, Riverside, Sacramento, San Benito, San Bernardino, San Diego, San Francisco, San Joaquin, San Luis Obispo, San Mateo, Santa Barbara, Santa Clara, Santa Cruz, Solano, Sonoma, Stanislaus, Sutter, Tulare, Ventura, Yolo, Yuba; **Illinois** (6): Cook, Du Page, Kane, Lake, Rock Island, Will; **New Mexico** (5): Bernalillo, Chaves, Dona Ana, Hidalgo, Mora; **Texas** (30): Bell, Bexar, Brazoria, Cameron, Culberson, Dallas, Deaf Smith, Dimmit, Ector, El Paso, Fort Bend, Galveston, Guadalupe, Harris, Hidalgo, Lubbock, Maverick, Nueces, Pecos, Presidio, Reeves, Starr, Tarrant, Tom Green, Travis, Uvalde, Val Verde, Webb, Willacy, Zavala.

A3. References

Borjas, George J., Richard B. Freeman, and Kevin Lang. 1991. "Undocumented Mexican-Born Workers in the United States: How Many, How Permanent?" In John M. Abowd and Richard B. Freeman (eds.) *Immigration, Trade, and the Labor Market*. Chicago: University of Chicago Press.

Manson, Steven, Jonathan Schroeder, David Van Riper, Katherine Knowles, Tracy Kugler, Finn Roberts, and Steven Ruggles. 2024. IPUMS National Historical Geographic Information System: Version 19.0 [dataset]. Minneapolis, MN: IPUMS. http://doi.org/10.18128/D050.V19.0
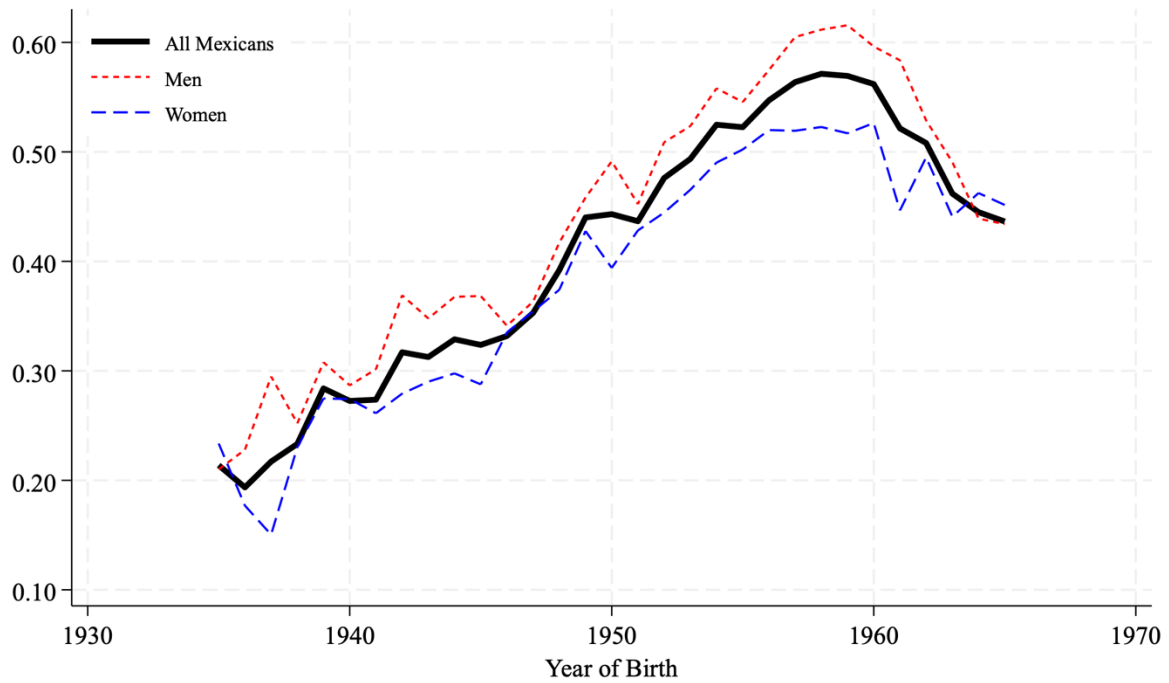
Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Grace Cooper, Julia A. Rivera Drew, Stephanie Richards, Renae Rodgers, Jonathan Schroeder, and Kari C.W. Williams. 2025. *IPUMS USA: Version 16.0* [dataset]. Minneapolis, MN: IPUMS. https://doi.org/10.18128/D010.V16.0

United States Bureau of the Census. *County and City Data Book [United States] Consolidated File: County Data, 1947-1977*. Inter-university Consortium for Political and Social Research [distributor], 2012-09-18. https://doi.org/10.3886/ICPSR07736.v2
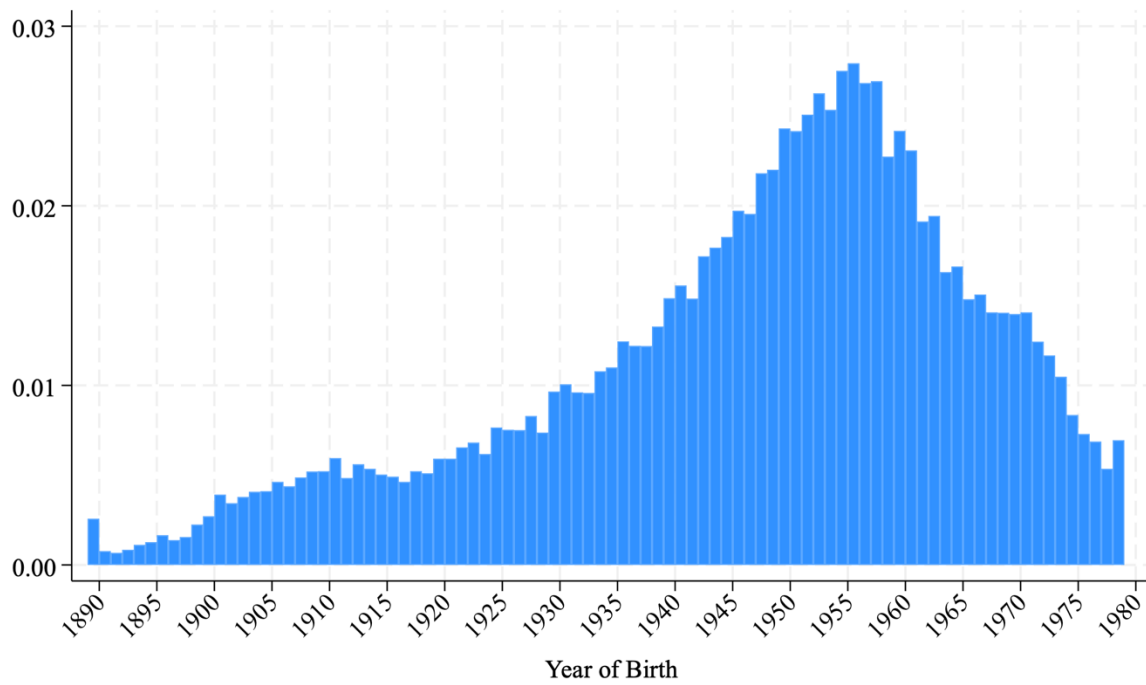
United States Department of Justice. Immigration and Naturalization Service. *Alien Address Reports, [United States]: 1980 Public Use File*. Inter-university Consortium for Political and Social Research [distributor], 1992-02-16. https://doi.org/10.3886/ICPSR07998.v1

A4. Tables and Figures

APPENDIX FIGURE 1. 1980 MEXICAN UNAUTHORIZED SHARE BY GENDER AND YEAR OF BIRTH

Note: Unauthorized shares calculated as the gap in the share of total and lawful arrivals of Mexican migrants in 1980, by gender and year of birth. Total arrivals and naturalized citizens are calculated from 1980 Census 5% public use microdata (Ruggles et al., 2025). Other lawful residents (Green Card and temporary visa holders) are drawn from the *Alien Address Reports 1980 Public Use* File (U.S. Dept. of Justice, 1992). Total arrivals are the sum of naturalized citizens and other lawful residents. Details on calculation of these figures are given in the appendix.

APPENDIX FIGURE 2. BIRTH YEAR FOR 1980 U.S. RESIDENT WOMEN BORN IN MEXICO.

Note: Histogram of the 1980 share of foreign-born Mexican women residing in the U.S., by year of birth. Sample restricted to those with non-U.S. born parents. Calculated from the 1980 Census 5% public-use microdata (Ruggles et al., 2025).

Mother's ethnicity:	1979		1970	
	MEXF	*NATIVEF*	*MEXF*	*NATIVEF*
Birth weight	3,389	3,327	3,357	3,275
	[45.4]	[62.1]	[55.3]	[48.9]
Ln of birth weight	8.13	8.11	8.12	8.09
	[0.01]	[0.02]	[0.02]	[0.02]
Low birth weight share	0.05	0.07	0.06	0.08
	[0.01]	[0.02]	[0.02]	[0.01]
High birth weight share	0.11	0.10	0.10	0.08
	[0.02]	[0.02]	[0.03]	[0.01]
Mother's age	25.63	24.76	26.81	24.35
	[0.62]	[0.62]	[0.65]	[0.56]
Male share	0.51	0.51	0.52	0.51
	[0.02]	[0.01]	[0.04]	[0.01]
Second child share	0.26	0.32	0.21	0.28
	[0.03]	[0.01]	[0.03]	[0.02]
Third child share	0.17	0.15	0.20	0.20
	[0.02]	[0.01]	[0.03]	[0.01]
Fourth or more child share	0.24	0.10	0.36	0.17
	[0.05]	[0.03]	[0.06]	[0.04]
Number of births	2,481	25,266	1,258	29,742
	[2,178]	[24,597]	[986]	[33,245]
Observations	79	79	79	79

Note: Means are reported with standard deviations in brackets and are sourced from Natality Detail Data, which contain individual birth records for either the universe or a 50% random sample of births, depending on state and year. Data are aggregated to group-by-county-by-child birth year averages, and statistics are weighted by the group-specific county share of 1970 births in the sample. Birth outcomes include average birth weight (in grams) and natural log of birth weight and shares low birth weight (<2500 grams) or high birth weight (>4000 grams). Birth characteristics include maternal average age at childbirth and the share of children who were male at birth. Parity variables indicate the shares of children who were second, third, or fourth (or higher) birth order.

	Unauthorized Rate (x100)
Hispanic share (1970)	-0.645
	(0.239)
Unemployed share (1970)	-01.011
	(1.615)
Population per square mile (1970)	- -0.000236
	(0.000963)
Urban population share (1970)	0.0190
	(0.125)
Population percentage change (1960-1970)	0.153
	(0.120)
Population percentage, 18 or older (1970)	0.126
	(1.291)
Population percentage, 25 or older with high school (1970)	-0.385
	(0.466)
Services employment share (1970)	0.558
	(1.563)
Construction employment share (1970)	-1.603
	(2.103)
Families with income < $3000 share (1970)	0.341
	(0.652)
Constant	50.90
	(74.21)
Observations	79
R-squared	0.750
State fixed effects	Yes
F: 1970 county chars.	3.588
p: 1970 county chars.	0.000594

Note: Table reports the coefficients from a regression of the estimated 1980 unauthorized share (multiplied by 100) on 1970 county characteristics and state fixed effects. Robust standard errors are in parentheses. Regression is weighted by the number of births to foreign-born Mexican women in 1970. Characteristics are sourced from the County and City Data Book (United States Bureau of the Census, 2012).

	DD: 1979	DD: 1970	DDD: 1979 - 1970	
	(1)	(2)	(3)	(4)
Coefficient	2.6	-13.4	16.0	17.3
	(7.9)	(9.0)	(6.6)	(7.9)
Observations	158	158	316	316
R-squared	0.983	0.989	0.987	0.997
County fixed effects	Y	Y	Y	Y
State x *MEXF* fixed effects	Y	Y	Y	Y
State x *MEXF* x 1979 fixed effects	N	N	Y	Y
County x 1979 fixed effects	N	N	Y	Y
County x *MEXF* fixed effects	N	N	N	Y
1970 county chars x *MEXF* x 1979	N	N	N	Y
p: 1970 county chars x *MEXF* x 1979				0.373

Note: Outcome is predicted birth weight (in grams), where the prediction is based on the group-specific 1970 relationship between birth weight and a cubic in maternal age, indicators for second, third, and fourth or higher parity, and an indicator for child sex. Columns (1) and (2) present least squares estimates of θ^{DD} from model (3) using Natality Detail Data from 1979 and 1970, respectively. Column (3) presents the difference between the column (1) and column (2) estimates, and column (4) regression adjusts that difference for county-by-group fixed effects and year-by group-specific effects of the 1970 county characteristics listed in Appendix Table 2. *MEXF*=1 for children born to Mexican women, and *MEXF*=0 for children born to native women. Regression estimates are weighted by the group-specific county share of 1970 births in the sample, and standard errors (in parentheses) are clustered on county. Sample consists of births to Mexican and U.S. born mothers residing in one of 79 counties across five states: California, Texas, Illinois, Arizona, and New Mexico. DD=difference-in-differences coefficient, DDD=triple differences coefficient.