ICE at the Door, Tests on the Floor: Student Achievement and Local Immigration Enforcement

CORA BENNETT, VIRGINIA GRAVES, AND BENJAMIN MEADOWS

Federal immigration statutes are enforced by Immigration and Customs Enforcement (ICE). However, ICE enforcement does not occur in a vacuum; it has a well-documented legacy of spillovers. Understanding the actual behaviors of immigration enforcement is exceedingly difficult owing to opaque or unavailable data. In this article, we are able to match the level of local immigration enforcement in local areas to school-district-level elementary and middle school achievement scores. Specifically, we merge ten years of individual-arrest ICE records from the New Orleans district field office obtained from a Freedom of Information Act request to district-level achievement scores to understand how immigration enforcement affects the human capital of children. We find that overall responses to enforcement as measured by district-level achievement scores are sharp and large, with achievement scores dropping anywhere between 0.27 and 0.52 of a standard deviation.

Keywords: immigration, economics of education, economics of immigration, immigration enforcement

In recent years, the documented effects of immigration enforcement on vulnerable populations have increased tremendously. For example, Secure Communities, a Department of Homeland enforcement program, directly reduced Hispanic citizen participation in Federal safety net programs (Alsan and Yang 2024) and both citizen and noncitizen labor hours (East et al. 2023). Opposing policies that reduce immigration enforcement have been shown to potentially increase an immigrant's willingness to report crime (Jacome 2022). Within the United States, scarce resources must be devoted toward immigration enforcement, and therefore, the efficacy and spillovers of those resources should be justly scrutinized, as documented by Caitlin Patler and Bradford Jones (2025, this issue). The tremendous scale at which enforcement has grown in the past twenty years makes this a highly pertinent question. Yet, most studies that document the actions of immigrants and immigration en-

Cora Bennett is assistant professor in the Department of Economics at the University of Notre Dame, United States. **Virginia Graves** earned a BS in economics, *summa cum laude*, from the University of Alabama, Birmingham, United States. **Benjamin Meadows** is assistant professor in the Department of Economics at the University of Alabama, Birmingham, United States.

 Bennett, Cora, Virginia Graves, and Benjamin Meadows. 2025. "ICE at the Door, Tests on the Floor: Student Achievement and Local Immigration Enforcement." *RSF: The Russell Sage Foundation Journal of the Social Sciences* 11(4): 104–22. https://doi.org/10.7758/RSF.2025.11.4.05. Direct correspondence to: Benjamin Meadows, at bgmeadow@uab.edu, 710 13th Street S, Birmingham, AL 35233, United States.

forcement either use rough proxies to identify the immigration status of an individual, such as race, or use broad immigration policy changes, such as the county-wide Secure Communities program. While using these broad individual characteristics or changes in immigration policy offers insight into how immigration policies affect individuals, it narrows the precision of research on such a significant topic.

In this study, we use GPS-coded or landmark-location, individual arrest, Immigration and Customs Enforcement (ICE) data. This data focuses on non-border apprehensions, which is a growing segment of ICE activity. With this data, we can gain a better understanding of the effect of ICE enforcement on the achievement scores of US schoolchildren because, rather than assessing the rollout of a policy geared to increase immigration enforcement, we assess the enforcement itself. When an immigration enforcement policy is rolled out, it could differ by its level of enforcement across treated areas, have delays in implementation, or have heterogeneous effects based on the treatment population. By using apprehension data, we are better able to identify the specific treatment that was implemented by these immigration enforcement policies.

We match this ICE arrest data to school-district-level achievement data from the Stanford Education Data Archive (SEDA). This affords us the ability to estimate the effect of local ICE enforcement on student achievement, measured at the school-district level. The arrest data specify the type, location, and date of each ICE arrest, which also affords a level of heterogeneity in analysis. We estimate a difference-in-differences model, following Brantly Callaway and Pedro H. C. Sant'Anna (2021), to compare the Hispanic-White test score gap between school districts with high levels of immigration enforcement activity and those without.

Two key sources of variation allow us to identify the effects. First, our outcome variable captures the difference between Hispanic and White students within a school district. We chose to focus on the Hispanic-White test score gap, as there is extensive literature on the racialized consequences of immigration enforcement (Flores and Schachter 2018; Asad and Clair 2018). Thus, not only are undocumented Hispanic students going to be affected by the shocks to their community caused by ICE apprehensions at higher rates than their White peers, but additionally, both undocumented and documented students might be directly affected if someone in their family is apprehended. Regardless, as Frank D. Bean and colleagues (2015) examine—legal status is a central axis of inequality for Hispanic Americans, and our study highlights one of these areas.

Our second source of variation is the differing levels of ICE enforcement across school districts. Although this variation may appear endogenous—since school districts with higher proportions of Hispanic families, regardless of immigration status, are likely to have a greater number of enforcement events—we do not believe it affects our outcome variable, which is the difference between Hispanic and White test scores.

The findings indicate that ICE enforcement within a school district significantly impacts Hispanic children's test scores, particularly when enforcement is substantial (right-tail events) and occurs directly in communities. Specifically, such enforcement actions increased the White-Hispanic test score gap by between 0.27 and 0.52 of a standard deviation in math and between 0.28 and 0.52 of a standard deviation in reading language arts.

This analysis also reveals a statistically precise negative effect on student performance on standardized test scores when there is large-scale ICE enforcement. An unintended consequence of immigration enforcement policies in the US is their effect on young, school-age children. Other literature by Benjamin Meadows (2021) points to the fact that these negative effects are not held by immigrants alone—young US citizens are affected as well.

We contribute to the broader literature on the unintended effects of immigration enforcement, and more specifically, the effects on children. In general, children of immigrants are worse off across many dimensions with more intense immigration enforcement. After increases in immigration enforcement, children live in an increased level of poverty (Amuedo-Dorantes et al. 2018) and can have blunted early childhood development (Gonzalez and Patler 2021).

Literature has developed around the effects of increased immigration enforcement on young children's educational achievement and attendance. J. Jacob Kirksey and Carolyn Sattin-Bajaj (2021) find that California children's academic performance was hindered by local immigration enforcement. Additionally, Kirksey and Sattin-Bajaj (2025, this issue) show that an ICE raid on Load Trail LLC in 2018 not only decreased four-year college enrollment among Hispanic and English-learner students (a common proxy for students lacking documentation) but also increased high school employment in those populations, likely stemming from students choosing to enter the workforce rather than four-year colleges. Laura Bellows (2019) also finds that activation of Secure Communities decreased achievement scores in Hispanic and Black students. Finally, while Thomas Dee and Mark Murphy (2018) document large disenrollments among Hispanic students in the face of Section 287(g) partnerships, Benjamin Meadows and Griffin Edwards (2023) find that most students in Southeastern states remained in the classroom during the brunt of harsh local immigration enforcement laws. The research presented in this article underscores the effects using more detailed data, confirming what was already understood from assessments of large-scale policies.

Biases against certain ethnicities or nationalities can translate into discriminatory practices in hiring, wage setting, and access to public services, contributing to economic inequality. These biases likely also translate to the classroom. For example, René D. Flores and Ariela Schachter (2018) demonstrate that factors such as national origin and social status cues (like occupation, education, and language fluency), as well as other characteristics or police records, influence whether someone is perceived as undocumented. Asad L. Asad and Matthew Clair (2018) expand this with the idea of racialized legal status (RLS), defining it as a seemingly race-neutral legal category that disproportionately affects racial/ethnic minorities. They argue that RLS functions as a social determinant of health, contributing to disparities through direct effects on individuals who hold the status and indirect spillover effects on their racial/ethnic group members. Cecilia Menjívar (2021) also highlights the extensive research exploring the criminalization and racialization of immigrants through legal frameworks, media portrayals, and enforcement action. This work underscores how legal statuses, enforcement programs such as 287(g), media stereotypes, and local contexts contribute to discrimination, disparate impacts, and negative experiences for various immigrant groups, often along racial and ethnic lines. Thus, if ICE enforcement programs target majority-Hispanic communities, there are multiplicative effects expected, and these ripple out throughout the community, regardless of enforcement purposes.

We also add to the literature that looks specifically at the effects of immigration enforcement actions, not just the enforcement policies. Catalina Amuedo-Dorantes and Francisca M. Antman (2022) find that ICE deportations reduce labor force participation in likely undocumented workers. Conversely, Daniel Corral (2021) finds that sanctuary policies may have an insulating effect on these typically deleterious consequences. At the same time, the interaction of natural disasters and visibility and detection by ICE has been investigated by Agustina Laurito and Ashley N. Muchow (2025, this issue), who show increased labor force participation in the face of home-country natural disasters, leading to greater detection by ICE.

Also, local policing practices and perceptions of legal status further complicate the education of immigrant communities. Joscha Legewie and Jeffrey Fagan (2019) demonstrate that aggressive policing tactics can negatively impact the educational outcomes of minority youth, potentially limiting their future economic opportunities. This is reinforced by Legewie and colleagues (2022), who find that increased police presence around college campuses can reduce the academic performance of undocumented male students. Using student data from the City University of New York, they observe particularly sharp effects in Black and South Asian men who face heightened surveillance, demonstrating significant consequences of local policing for undocumented individuals. The intersection of legal status and local law enforcement creates significant barriers to human capital development for this population, as our article reinstates.

DATA

To understand the effect of local immigration enforcement upon local student populations, we merge two unique data sets. First, we use ICE apprehension data from the New Orleans field office. Second, we merge this into the SEDA data archive, which includes the universe of student achievement scores.

Enforcement Data

To measure the effects of immigration enforcement on children's human capital accumulation, it is important to focus on ICE's on-the-ground activity. Specifically, we focus here on apprehensions by ICE. Typically, deportations are the most-reported and displayed events from ICE, yet they also represent the last link in a lengthy chain of administrative events (arrests, court dates, transfers to other facilities, final orders, and so on) (Moinester 2024). Alternatively, ICE apprehensions represent the genesis of that administrative cascade of events, and thus may still be highly influential in communities. There are several different types of ICE apprehensions, but in general, one can define an ICE apprehension as the point in time when a person is transferred to ICE custody, whether by jail transfer or in-person arrests. The data for daily ICE enforcement operations come from a Freedom of Information Act (FOIA) request we filed with ICE. These data are the records of every individual apprehended by ICE's Enforcement and Removal Operations (ERO) from 2008 to 2018. Each data point is an arrest at the individual level from the New Orleans field office, which covers Louisiana, Arkansas, Mississippi, Tennessee, and Alabama. Thus, our research and results are pinned to the Southeastern context. Each observation includes the individual's country of origin, apprehension location, arrest method, arrest program, and date. Arrest program typifies the program used to apprehend the individual under ERO, which can be the Criminal Alien Program (CAP), Section 287(g) of the Immigration and Nationality Act, or a targeted apprehension program that includes the arrest methods Non-Custodial Arrests and Located Arrests. Arrest method is defined as the type of apprehension ICE executes within the aforementioned programs and includes specific indicators such as local, state, or federal arrests under CAP.

In addition to location, we are able to isolate different types of operations undertaken by ICE.

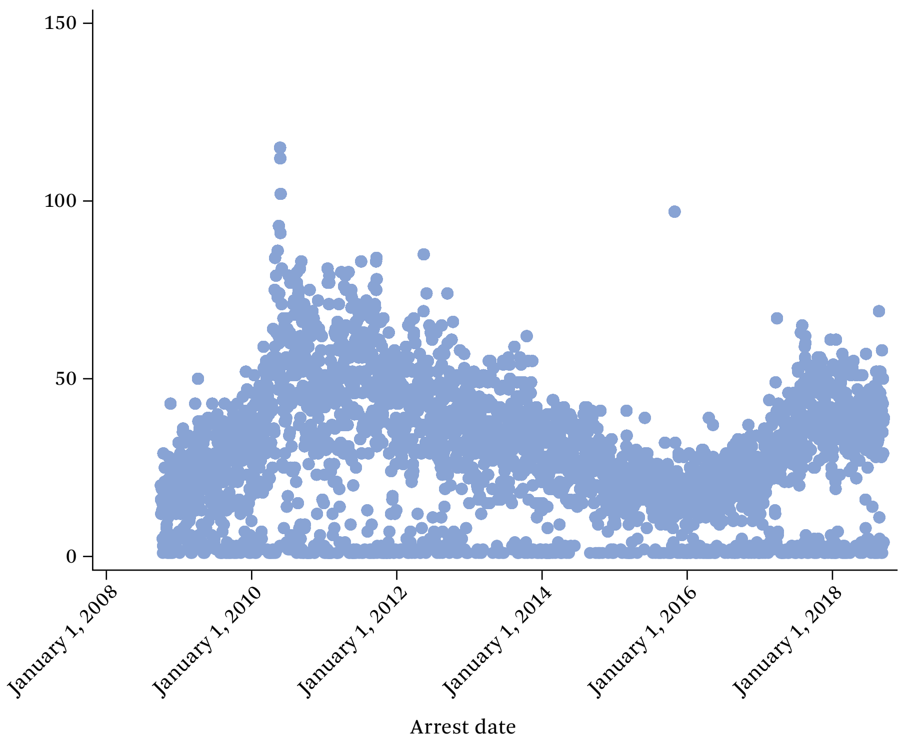
In our data, there are five common arrest methods outlined by ICE: Federal Criminal Alien Program, Local Criminal Alien Program, State Criminal Alien Program, Located Arrests, and Non-Custodial Arrests. For instance, ICE can detain or request to detain undocumented individuals arrested by other agencies for non-immigration offenses (drunk driving, theft, and so on) under CAP. The majority of the arrests in our data are these types of detainers, in which custody is transferred from the specific authority to ICE. Specifically, these are coded as CAP Local (78 percent), CAP State (3 percent), and CAP Federal (7.8 percent). For the remainder of the article, all CAP programs are treated as one.

The other arrest types are more direct apprehensions, with Located Arrests (4.6 percent), Non-Custodial Arrests (2 percent), and 287(g) arrests (1 percent). While these types of arrests are smaller in number, they are important to include as they codify the moments when ICE removes an individual from the public specifically for immigration violations. Located Arrests are where an individual is directly located and arrested by ICE. Since Located and Non-Custodial Arrests occur more visibly than jail transfers, it stands to reason that these arrest types could translate into larger community reactions, which may affect the well-being of students. It follows that the estimates of interest in our model rely on Located and Non-Custodial apprehensions rather than the total of ICE activity. Lastly, 287(g) arrests stem from the section of the Immigration and Nationality Act that authorizes local and state enforcement officials to perform immigration enforcement operations (Peacock 2025, this issue). In our data, these arrests typically occur only in the counties or districts that signed agreements with ICE to do so.

Overall, these ICE records obtained through FOIA create a picture of enforcement levels and totals on any given day in this study area. We display the yearly arrest totals for all types of ICE enforcement in figures 2–6. The general shape of ICE enforcement activities in figure 1—with peaks in the early 2010s and again in 2016–2018—matches different executive branch's administrative priorities.

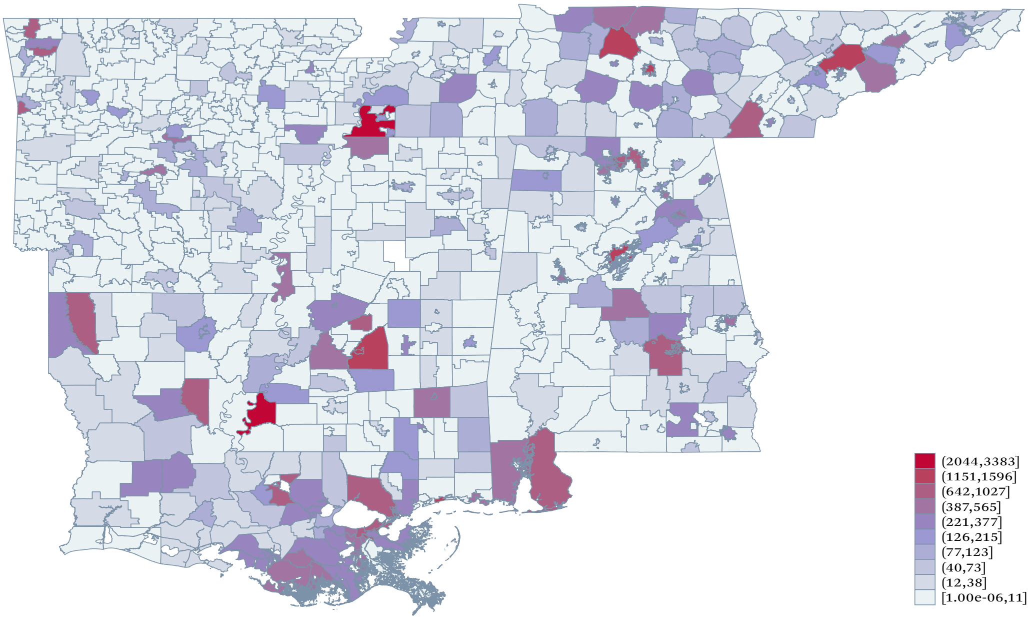
To merge this data with SEDA achievement outcomes, we leverage the shapefiles for each school district provided by SEDA—placing the GPS coordinates from the ICE data within the school districts. From there, we aggregate the number of arrests within our time period, which corresponds to the academic school year. When we aggregate all arrest types across all years, we see that the majority of the arrests occur in the largest cities—and therefore school districts—within the states: Fayetteville, Arkansas; Nashville, and Knoxville, Tennessee; Jackson, Mississippi; Birmingham, Alabama; and New Orleans, Louisiana (figures 2, A.1, and A.2). Additionally, figure 2 shows that port areas, such as New Orleans, Louisiana, and Mobile, Alabama, see increased enforcement numbers. When we aggregate only Located Arrests (arrests where ICE directly arrests an individual), we see a different pattern (figure 5). Further, when we incorporate Non-Custodial Arrests with Located Arrests areas such as Memphis, Tennessee; coastal Louisiana; and northern Alabama all show prominent numbers of these arrest types. As previously mentioned, 287(g) arrests only occur in a few areas due to agreements struck between ICE and local municipalities or counties (Fayetteville, Arkansas; Nashville and Knoxville, Tennessee; and Baton Rouge, Louisiana), as seen in figure 3. Lastly, in figure 4, CAP (jail transfer) arrests are highly

Figure 1. Total ICE Apprehension, by Day



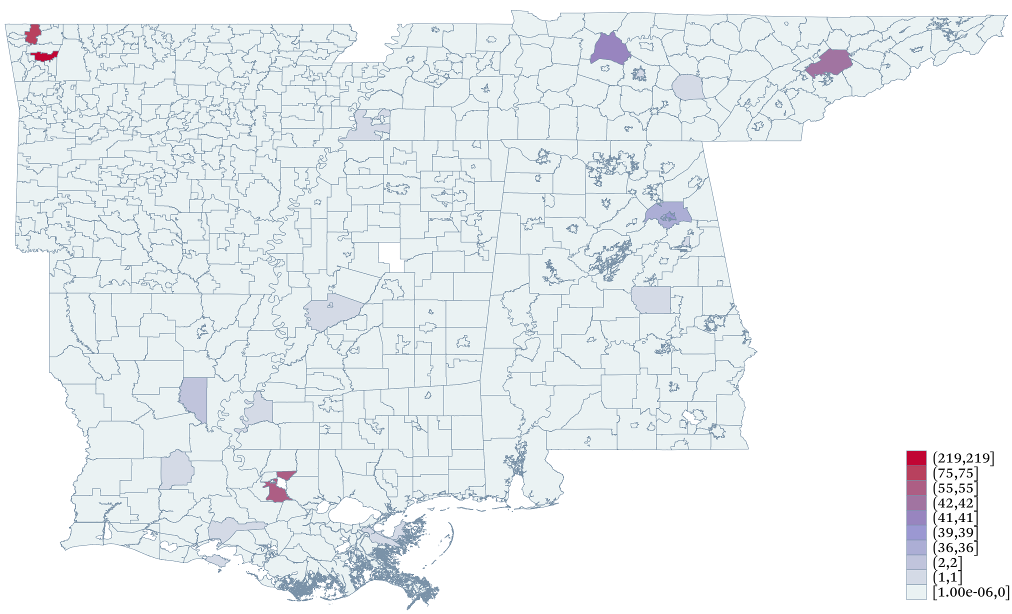
Source: Merged ICE x SEDA datafile.
Note: The total number of ICE apprehensions across all arrest types is aggregated by day and plotted for the entire panel.

Figure 2. All ICE Arrests, 2008–2018, in Southeastern School Districts



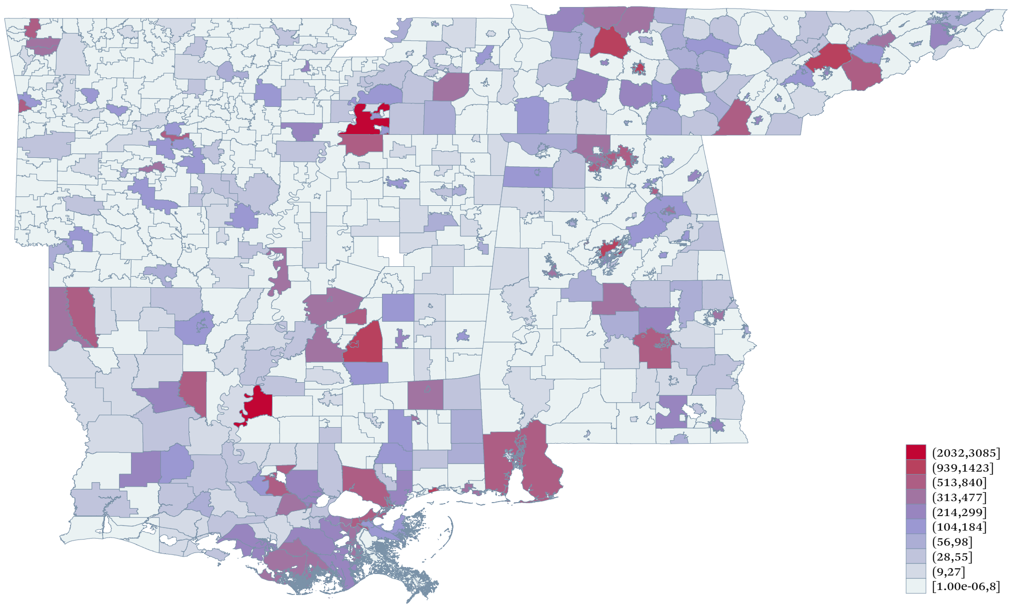
Source: Merged ICE x SEDA datafile.

Figure 3. All Section 287(g) ICE Arrests, 2008–2018, in Southeastern School Districts



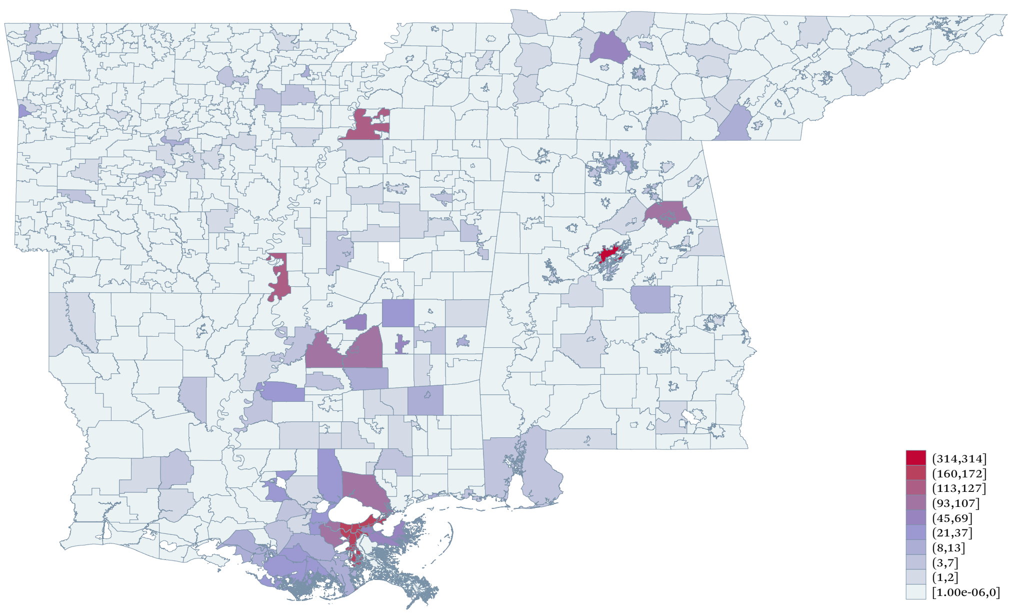
Source: Merged ICE x SEDA datafile.

Figure 4. All CAP Arrests, 2008–2018, in Southeastern School Districts



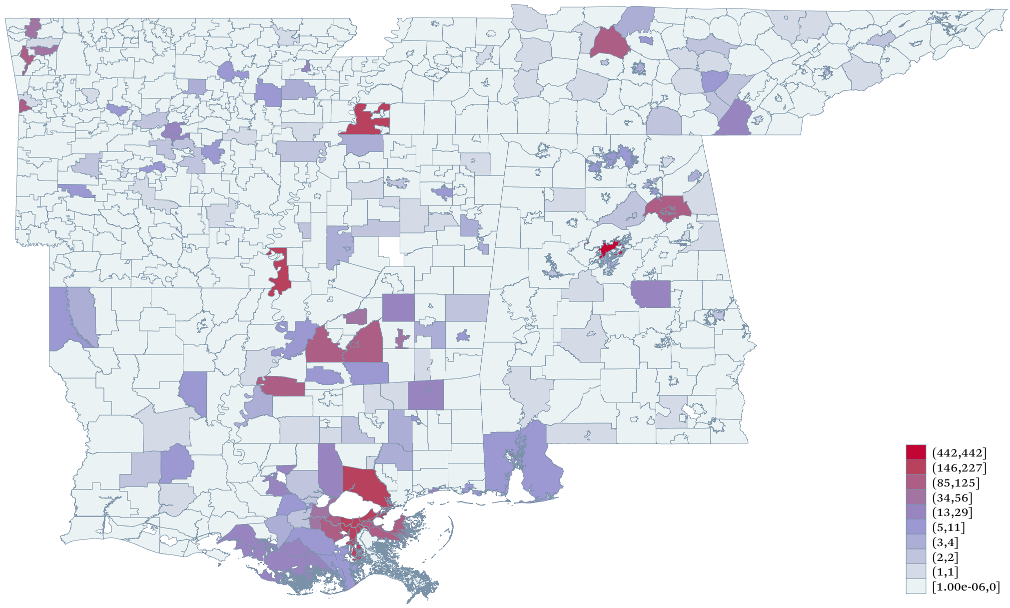
Source: Merged ICE x SEDA datafile.

Figure 5. All Located ICE Arrests, 2008–2018, in Southeastern School Districts



Source: Merged ICE x SEDA datafile.

Figure 6. All Located and Non-Custodial ICE Arrests, 2008–2018, in Southeastern School Districts



Source: Merged ICE x SEDA datafile.

variable across geographies, which partially reflects the holding jails of ICE (Mississippi and Louisiana) as well as the inherent randomness of transferring individuals in and out of custody.

Education Data

To assess student test scores, we use the nationwide standardized test score database from SEDA (Fahle et al. 2021). A large literature has grown from these data in not only education (Bellows 2019; Chin et al. 2020; Reardon et al. 2019; Welsh and Swain 2020; Rossin-Slater et al. 2020) but also economics (Abott et al. 2020; Baum-Snow et al. 2019; Bergman et al. 2019; Park et al. 2021). We specifically use SEDA Version 4.1 with estimates spanning from the 2008–2009 to 2017–2018 academic years. For efficiency, we simplify 2008–2009, for example, to academic year 2009 for the remainder of the article. The SEDA data leverages test score data required by the No Child Left Behind Act, which was then reported to the National Center for Education Statistics. SEDA is a uniquely comprehensive dataset that aggregates approximately 430 million standardized test scores from the 2009 through 2018 academic years. The dataset places state-administered assessments onto a unified national scale, facilitating consistent comparisons of educational opportunity across states, counties, school districts, and individual schools (Fahle et al. 2021). Rather than being sampled, it is population-level, encompassing nearly all US public school students. This allows for a direct, census-like description of student achievement nationwide.

From SEDA, we extract standardized test scores for grades 3 through 8 across all US school districts. The data include subgroup estimates disaggregated by student race, ethnicity, and economic disadvantage status. Our analyses incorporate scores from both mathematics and reading language arts (RLA) assessments. All achievement metrics are standardized and reported in terms of standard deviations relative to the national distribution of student-level test scores.

This makes SEDA data particularly appealing given that it is normalized to a common scale, allowing test scores at each grade, subject, and ethnic group to be compared across county and state lines (Fahle et al. 2021). In addition to average achievement levels, we also incorporate district-level achievement gaps reported by SEDA, where available. These gaps quantify the difference in academic performance between specified student groups and serve as key indicators of educational inequality. Because our ICE data precede our SEDA data by one year, we limit our analysis to academic years 2009–2018, focusing on grades 3 through 8. Our reported estimates reflect the effect size; or explicitly, a one unit increase in average achievement is as if a school district's average Hispanic student scored one standard deviation higher than the national grade reference achievement score (Bennett 2020). For context, David E. Frisvold (2015) estimated a positive 0.09 SD increase in test scores following the adoption of universal school lunches, while Eunsik Chang and Maria Padilla-Romo (2023) estimated a 0.11 SD test score drop for students who were exposed to violent crimes near their schools.

ANALYTIC STRATEGY AND RESULTS

To assess the impact of ICE enforcement on schoolchildren, we leverage the depth of the SEDA database to estimate the difference in the achievement gap between White and Hispanic students, across districts that receive high levels of ICE enforcement compared to those that do not (table 1). This affords an extra layer of robustness as compared to a classic difference-in-differences methodology, which might empirically fit where students who receive treatment are differenced out from those who do not before and after the event. Characteristics endogenous to the student population will likely be captured in our within-district estimates, specifically anything affecting all students in the district. This is accounted for not only in the district-level fixed effect but also in the fact that we estimate the gap between White and Hispanic students in the same district. For example, enforcement is endogenous to the makeup of the local community. ICE enforcement events rarely occur in predominately non-Hispanic school districts. Because our outcome variable captures the differences between white and Hispanic students within a school district, it should control for unobserved heterogeneity between school districts. We choose to look at White students as a control cohort because White students typically are unaffected by ICE enforcement (Meadows 2021; Bellows 2019) making them a good control group. Explicitly, we model a difference-in-differences model per Brantly Callaway and Pedro H. C. Sant'Anna (2021), but also leverage a key feature of the SEDA dataset—the White-Hispanic test score gap. While enforcement is endogenous to the makeup of the local community, as ICE enforcement events rarely occur in predominately White school districts, our outcome variable captures the differences between the target population and others within a school district.

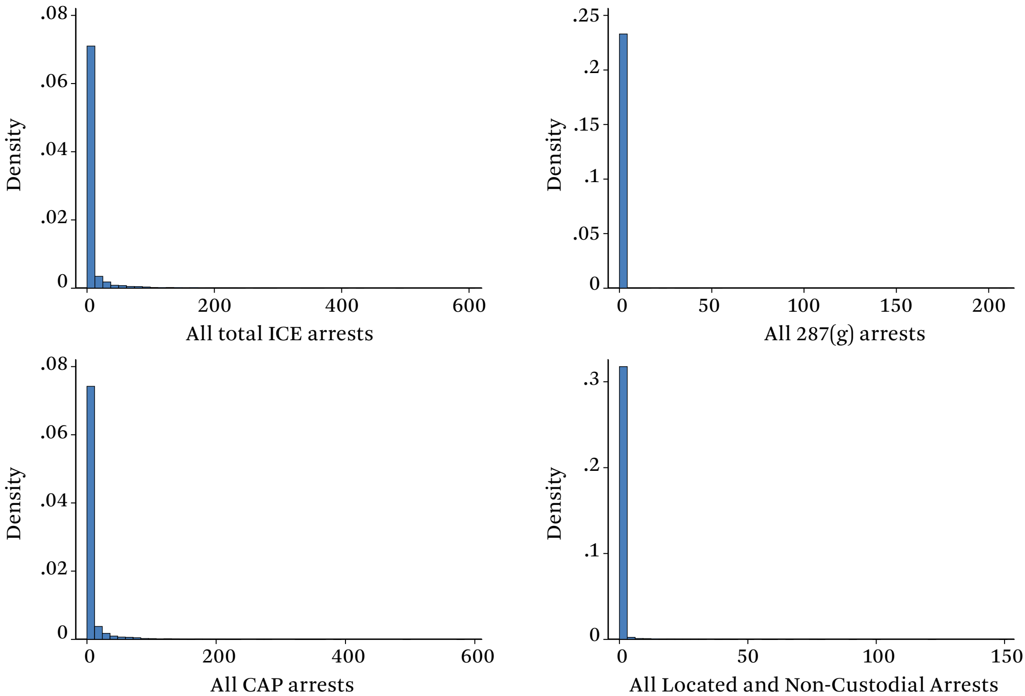
This outcome variable is advantageous, as it allows us to compare the difference in test scores of Hispanic students to White students within the same school district against the difference in test scores of Hispanic and White students in school districts that never receive ICE enforcement after an ICE enforcement event. This pseudo-triple-difference allows us to account for unobserved differences between school districts (including geographic sorting of families), unobserved differences between White and Hispanic students, and any temporal or other cross-boundary events that could confound our analysis.

Table 1. Summary Statistics—Achievement Scores

	Math Achievement Scores			RLA Achievement Scores		
	N	Mean	SD	*N*	Mean	SD
All students	40470	−.2493579	.3782498	41247	−.1717547	.342783
White	34747	−.0455919	.3399081	35864	.0395518	.3095359
Hispanic	7967	−.3169334	.3130288	7797	−.3163735	.2983609
Black	23200	−.6157642	.3082515	23862	−.522243	.2765474
Female	36281	−.1964859	.3675684	37355	−.0084557	.3460839

Source: Merged ICE x SEDA datafile.

Figure 7. Density of ICE Enforcement Events Within School Districts



Source: Merged ICE x SEDA datafile.

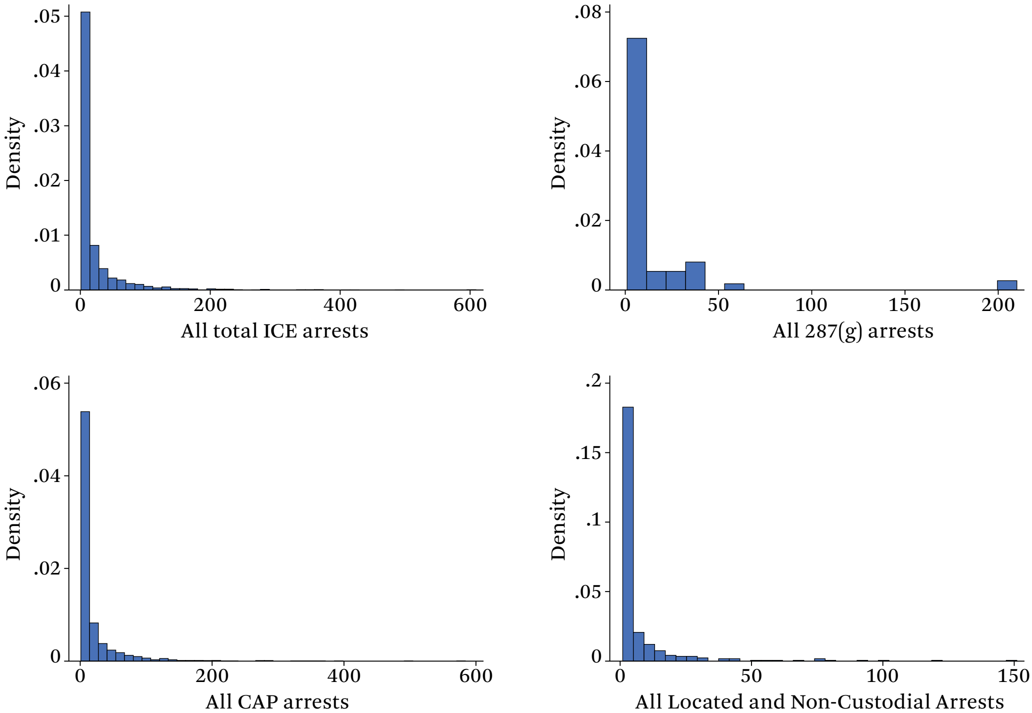
As previously mentioned, ICE enforcement covers most geographies in our dataset, but at minimal levels. For instance, 64 percent of school districts never have an ICE enforcement event within a school year. Figure 7 reveals this gap—the most common ICE enforcement event is none at all. A large portion of events are small scale, with between one and three arrests total across a decade, which might not factor prominently into students' annual test scores. Instead, high-profile events might affect students the most. To this end, we focus on right-tail events. Figures 7 and 8 show the density of immigration events at the school district level. Depending on the arrest type, the discrete number of arrest events in the right tail might change—meaning, a statistical cut point related to a percentage or standard deviation would lead to an unbalanced treatment scheme.[1] So for clarity and simplicity, we run multiple, separate regressions, selecting only one type of ICE enforcement at a time and coding the ten arrest events with the highest number of apprehensions as treatment.[2] School district-year observations within the largest ten arrest events are in our treatment group. All other school district-year observations are in the control group.

Given the staggered timing of the ICE events across our panel, combined with the need to control for unobserved heterogeneity across school districts, we estimate the two-way estimator proposed by Callaway and Sant'Anna (2021) to account for the well-established issue of "forbidden" comparisons and negative weighting. Specifically, we use the formulation from Callaway and Sant'Anna (2021) that estimates long gaps, using t–1 as the base period,

1. Alternatively, we could select cut points based on the shape of the enforcement histograms, but this would lead to differing quantities of treatment dates for each enforcement type, since a histogram cut point would leave us with 287(g) arrests receiving one treatment date and CAP arrests receiving eight.

2. A tabulated list of these events can be found in the appendix tables A.1–A.8.

Figure 8. Density of Non-Zero ICE Enforcement Events Within School Districts



Source: Merged ICE x SEDA datafile.

likening the estimand—the effect the model aims to measure—to a standard difference-in-differences setup. Explicitly, we estimate a difference-in-difference model using Callaway and Sant'Anna (2021) by

$$(1) \quad Y_{igt} = \eta_i + \eta_t + \beta \times ICE_i \times POST_t,$$

where Y is the respective test score gap between White and Hispanic students (reading or math) in i, the individual school district; g, the grade level; and t, the school year. Fixed effects η_i and η_t, help control for spatial and temporal unobserved heterogeneity. ICE is an indicator variable that represents those school districts that receive one of the top ten ICE enforcement events, and *POST* indicates the years that follow said event. Whereas there has been a statistical reckoning with difference-in-differences models, we address these concerns in two ways. We use the Callaway and Sant'Anna estimator to estimate a pseudo-triple difference by comparing the gap between White and Hispanic students within the same district to the gap in untreated districts. This approach helps explain some confounding, unobserved heterogeneity, as discussed by Andreas Olden and Jarle Møen (2022).

Tables 2 and 3 report the estimates of high ICE enforcement events, as denoted by the ICE variable, on the math and RLA achievement, respectively, differences between Hispanic and White students, relative to the difference of those students who did not have an ICE enforcement event in their school district. All aggregated ICE arrests increase the gap between White and Hispanic test scores by 0.12 and 0.067 of a standard deviation for math and RLA, respectively, although RLA is not precisely estimated. CAP and 287(g) arrests are both imprecisely estimated. Located Arrests widen the gap between White and Hispanic students by 0.105 and 0.10 of a standard deviation in both math and RLA, respectively. Table 4 reports the summary statistics by treatment level.

Given the possibility of multiple treatments (see appendix tables A.1–A.8), we assess the effect of a dosage style treatment by comparing

Table 2. Achievement Gap, Math, White-Hispanic

	All Total ICE Arrests	All CAP Local Arrests	All 287(g) Arrests	All Located Arrests
POST x ICE	0.118*	0.0760	−0.0126	0.105**
	(0.0510)	(0.0413)	(0.0451)	(0.0343)
Observations	26,816	26,673	26,973	27,011

Source: Merged ICE x SEDA datafile.
Note: Standard errors in parentheses.
* $p < .05$; ** $p < .01$; *** $p < .001$

Table 3. Achievement Gap, RLA, White-Hispanic

	All Total ICE Arrests	All CAP Local Arrests	All 287(g) Arrests	All Located Arrests
POST x ICE	0.0668	0.0514	0.0620	0.100*
	(0.0395)	(0.0317)	(0.0325)	(0.0393)
Observations	26,318	25,867	26,852	26,657

Source: Merged ICE x SEDA datafile.
Note: Standard errors in parentheses.
* $p < .05$; ** $p < .01$; *** $p < .00$

Table 4. Demographic and Summary Statistics by Treatment

	ICE Enforcement Area		
	No	Yes	All
RLA scores, White students	0.0390683	0.243052	0.0400834
RLA scores, Hispanic students	−0.3142646	−.4383601	−0.3164618
RLA White-Hispanic gap estimate	0.3413656	0.575192	0.3425821
Math scores, White students	−0.0458852	0.1209481	−0.0450476
Math scores, Hispanic students	−0.31501	−0.4178571	−0.3169729
Math White-Hispanic gap estimate	0.2741023	0.4739589	0.2752574
Poverty rate	0.190836	0.1904757	0.1908344
% Hispanic	0.0504732	0.095711	0.0506789
% economically disadvantaged	0.6601144	0.6294719	0.659972
Unemployment rate	0.0879265	0.0870962	0.0879227

Source: Merged ICE x SEDA datafile.

districts that received multiple large-scale ICE enforcement events in the panel, as compared to those that did not. Explicitly, we drop the districts that were treated once and compare the estimates from the multiply treated districts to the control (which still remains never treated) and use the same estimation strategy as before. In appendix tables A.5 and A.6, we see that the estimates are more muted and imprecisely estimated, meaning that the multiply treated districts are not driving our results.

DISCUSSION

In this article, we matched ICE enforcement data (GPS or landmark-location data) to the school district level—specifically, to achievement data—to understand the effect immigration enforcement has on schoolchildren. We find tightly estimated and strongly negative test score drops among the Located Arrests and Non-Custodial Arrests categories of ICE enforcement. While these precise estimates are informative, they are unsurprising. Given what researchers have established about traumatic events (Shany 2016; Chang and Padilla-Romo 2023), the fact that test scores drop when ICE executes a large-scale operation within the school district might be expected. However, contextualizing how traumatic this event might be is additionally important.

Here, we rely on Matthew A. Kraft's (2020) proposed guidelines and effect size benchmarks to better understand our results. These guidelines help set expectations about the effect sizes and place them into context, given the research design. To develop them, Kraft evaluated almost two thousand effect sizes from 747 randomized controlled trials (RCTs) and proposed new benchmarks for education research: Less than 0.05 is small, between 0.05 and 0.20 is medium, and greater than 0.20 is large. While we do not study a positive education intervention, these benchmarks are appropriate since we use standardized achievement scores as our outcome. We can situate the results presented here, while negative, against these positive benchmarks.

To put our results in context with other similar studies, it is important to analyze the results in light of the features of our specific research design. The magnitude of the effect sizes we present is affected by what outcome we used and when it was measured.

One could describe the annual, district-level achievement scores we used as sticky—or difficult to change—especially in large districts, where treated students might make up a small minority. Given this, the medium-sized effect we estimate on the gap between White and Hispanic students is quite considerable. One might think that documented Hispanic students would respond to immigration enforcement in a milder way than undocumented students. However, our outcome data only identified Hispanic students, which might have minimized the true effects. In addition, one would expect the effect size to be larger in our situation if we could measure the effects of a standardized test taken immediately in the wake of an immigration enforcement incident, as done in Eunsik Chang and Maria Padilla-Romo's (2023) study. We imagine the salience of any existing trauma could affect student performance much more than in our annual sample.[3]

In addition, the treatment and control samples selected play a role in the size and interpretation of the result. We used data aggregated at the district level, relying on large numbers of enforcement incidents to identify treatment. Some school districts are large enough to have multiple schools, some of which might not have been affected by an isolated immigration-related apprehension. If areas within a treated observation acted like a control observation, we would expect, again, a larger treatment effect. Being able to aggregate the data at the classroom level or gain access to individual-level data would have better isolated the treatment.

Given all this discussion of effect sizes and Kraft's (2020) updated benchmarks, we consider our results a medium-sized, lower-bound estimate of the potential educational achievement effects of a large-scale immigration enforcement incident. The realized effects could be much greater than our model was able to detect.

We propose two mechanisms that could be causing the increase in the White-Hispanic achievement gap, both of which have a direct effect on children's performance on standardized test scores. First, Located Arrests could be creating pockets of absenteeism among Hispanic students not identified in the aggregate data. Second, these events could be placing undue amounts of stress on children with undocumented family members. Even if students are not missing school, the atmosphere outside the classroom could make learning difficult, impeding success on the standardized tests, as detailed thoroughly in Frank D. Bean and col-

3. The SEDA data is undated within an academic year. Because states administer standardized tests at different points in the year, we cannot precisely align enforcement events with the timing of test administration.

leagues' (2015) work. Both absenteeism and stress could result from an actual Located Arrest, but they could also be the result of the threat of increased enforcement. While the threat of increased activity is nearly impossible to measure, it is likely present in our control districts. If control districts are substantially affected by this threat in the same way as treated districts, the result we find might be moderated.

It is important to note that this article documents learning losses rather than improvements. Therefore, these contextualizations are likely to be conservative in nature. For instance, David E. Frisvold (2015) estimates a positive 0.09 standard deviation increase in test scores following the adoption of universal school lunches. Our estimate of Located Arrest events on Hispanic student test scores is 65 percent of that estimate. Additionally, Eunsik Chang and Maria Padilla-Romo (2023) estimated a 0.11 standard deviation test score drop for students who were exposed to violent crimes near their schools. Thus, a rough comparison of effect sizes shows an ICE enforcement event is half as detrimental as having a violent crime near the school. So, while our estimates are expected on the policy scale, they are nonetheless consequential by comparison. Research by Benjamin Meadows (2021) shows that undocumented and documented students share the burden of immigration enforcement legislation, suggesting that the human capital losses in this study are likely distributed between citizens and noncitizens alike.

APPENDIX. ICE ENFORCEMENT LEVELS, DISTRICT ACHIEVEMENT, AND EVENT STUDY RESULTS

Table A.1. List of Districts with the Highest Enforcement

	2010	2011	2012	2013	2014–2018	Total
Birmingham City					×	×
Davidson County				×		×
Natchez-Adams School District					×	×
Rankin County School District		×				×
Shelby County	×	×	×	×	×	× × × × ×

Source: Merged ICE x SEDA datafile.

Table A.2. List of Districts with the Highest Enforcement: CAP Local

	2010	2011	2012	2013	2014	2015	Total
Davidson County					×	×	× ×
Gulfport School District	×						×
Knox County		×					×
Murfreesboro			×				×
Rankin County School District		×					×
Shelby County		×	×	×	×		× × × × × × ×

Source: Merged ICE x SEDA datafile.

Table A.3. List of Districts with the Highest Level of Section 287(g) Enforcement

	2013	2014	2018	Total
Bentonville School District		×	×	××
Davidson County	×			× ×
East Baton Rouge Parish		× ×		× ×
Etowah County	×	×		× ×
Fayetteville School District		×	×	× ×
Gadsen City		×	×	× ×
Knox County		×		×

Source: Merged ICE x SEDA datafile.

Table A.4. List of Districts with the Highest Level of Located Enforcement

	2009	2010	2011	2012	2014	2016	2017	2018	Total
Birmingham City					×		×	×	× × ×
Davidson County		×							×
Etowah County		×							×
Jefferson Parish	× ×		× ×	× ×					× × × × ×
Orleans Parish					×				×
Rankin County School District				×					×
Shelby County		×						×	× ×

Source: Merged ICE x SEDA datafile.

Table A.5. Achievement Scores for Districts Receiving Multiple Enforcement Events (Math)

	All Total ICE Arrests	All CAP Local Arrests	All 287(g) Arrests	All Located Arrests
POST × ICE	−0.0221	−0.00752	−0.0134	−0.0176
	(0.0277)	(0.0294)	(0.0509)	(0.0394)
Observations	21,247	21,478	26,190	20,669

Source: Merged ICE x SEDA datafile.
Note: Standard errors in parentheses.
* $p < .05$; ** $p < .01$; *** $p < .001$

Table A.6. Achievement Scores for Districts Receiving Multiple Enforcement Events (RLA)

	All Total ICE Arrests	All CAP Local Arrest	All 287(g) Arrests	All Located Arrests
POST × ICE	−0.0197	−0.0447	0.0780*	−0.00325
	(0.0180)	(0.0288)	(0.0385)	(0.0327)
Observations	20,770	21,035	25,867	20,291

Source: Merged ICE x SEDA datafile.
Note: Standard errors in parentheses.
* $p < .05$; ** $p < .01$; *** $p < .001$

Table A.7. White-Hispanic Achievement Gap in Math, Any ICE Enforcement

	All Total ICE Arrests	All CAP Local Arrests	All 287(g) Arrests	All Located Arrests
POST × ICE	0.0185	0.0197	0.0608	0.0426
	(0.0210)	(0.0208)	(0.0317)	(0.0241)
Observations	18,568	18,937	26,399	23,910

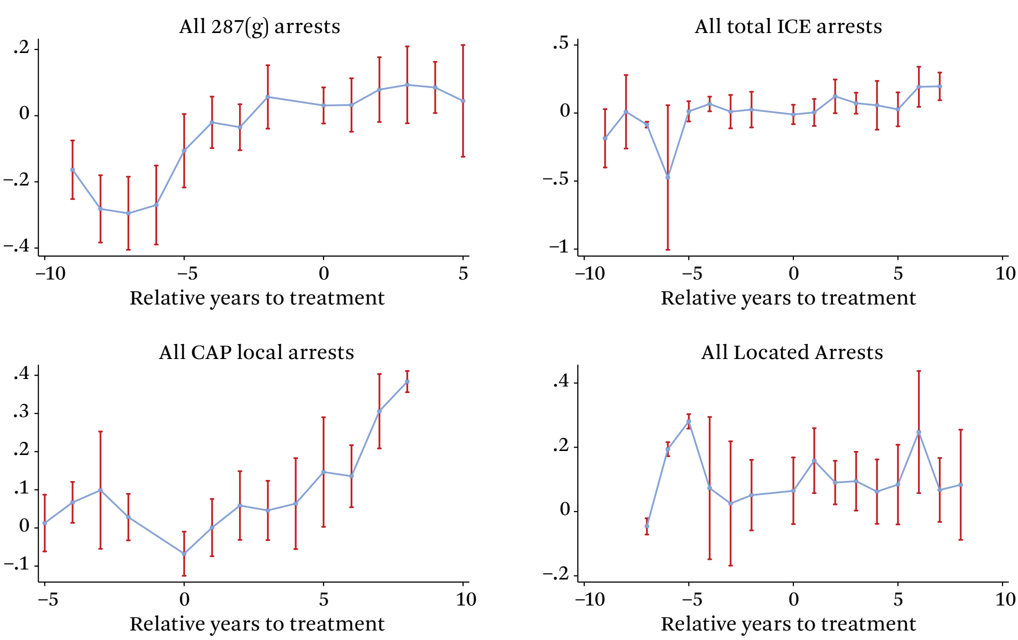
Source: Merged ICE x SEDA datafile.
Note: Standard errors in parentheses.
* $p < .05$; ** $p < .01$; *** $p < .001$

Table A.8. White-Hispanic Achievement Gap in RLA, Any ICE Enforcement

	All Total ICE Arrests	All CAP Local Arrests	All 287(g) Arrests	All Located Arrests
POST × ICE	0.00678	0.00666	0.0574	0.0317
	(0.0203)	(0.0202)	(0.0331)	(0.0251)
Observations	17,500	17,703	25,987	23,239

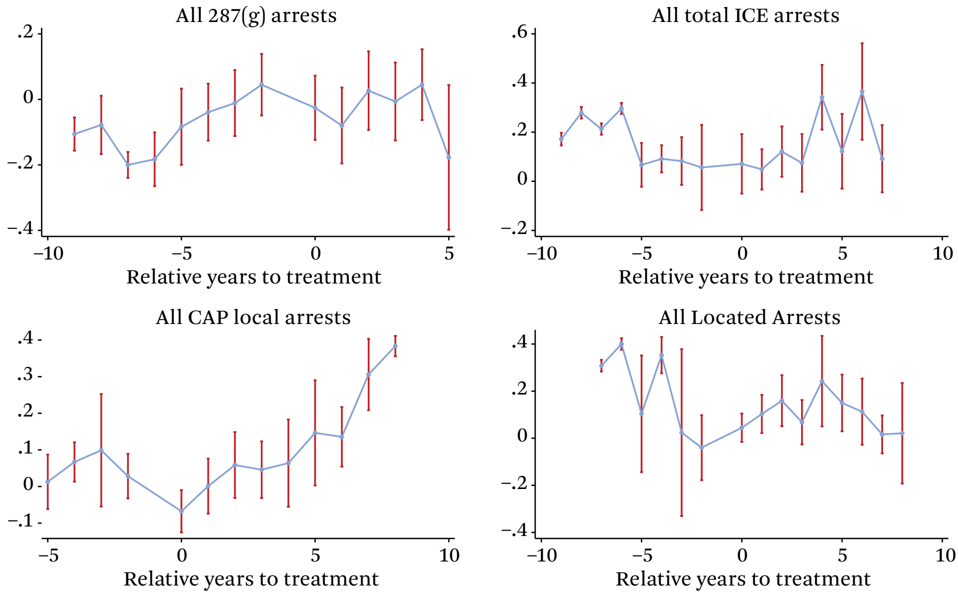
Source: Merged ICE x SEDA datafile.
Note: Standard errors in parentheses.
* $p < .05$; ** $p < .01$; *** $p < .001$

Figure A.1. Event Study for Each Type of ICE Enforcement, RLA



Source: Merged ICE x SEDA datafile.

Figure A.2. Event Study for Each Type of ICE Enforcement, Math



Source: Merged ICE x SEDA datafile.

REFERENCES

Abott, Carolyn, Vladimir Kogan, Stéphane Lavertu, and Zachary Peskowitz. 2020. "School District Operational Spending and Student Outcomes: Evidence from Tax Elections in Seven States." *Journal of Public Economics* 183: 104142.

Alsan, Marcella, and Crystal S. Yang. 2024. "Fear and the Safety Net: Evidence from Secure Communities." *Review of Economics and Statistics* 106(6): 1427–41.

Amuedo-Dorantes, Catalina, and Francisca M. Antman. 2022. "De Facto Immigration Enforcement, ICE Raid Awareness, and Worker Engagement." *Economic Inquiry* 60(1): 373–91.

Amuedo-Dorantes, Catalina, Esther Arenas-Arroyo, and Almudena Sevilla. 2018. "Immigration Enforcement and Economic Resources of Children with Likely Unauthorized Parents." *Journal of Public Economics* 158: 63–78.

Asad, Asad L., and Matthew Clair. 2018. "Racialized Legal Status as a Social Determinant of Health." *Social Science & Medicine* 199: 19–28.

Baum-Snow, Nathaniel, Daniel A. Hartley, and Kwan Ok Lee. 2019. "The Long-Run Effects of Neighborhood Change on Incumbent Families." *Working Papers* 2019-02. Federal Reserve Bank of Chicago.

Bean, Frank D., Susan K. Brown, and James D. Bachmeier. 2015. *Parents Without Papers: The Progress and Pitfalls of Mexican American Integration*. Russell Sage Foundation.

Bellows, Laura. 2019. "Immigration Enforcement and Student Achievement in the Wake of Secure Communities." *AERA Open* 5(4): 2332858419884891.

Bennett, Cora. 2020. "It's Dolly's World, We're Just Reading in It: The Effects of an Early Childhood Literacy Program." Poster presented at the Association for Public Policy Analysis and Management (APPAM) Fall Research Conference. Virtual (November 11–13, 2020).

Bergman, Peter, Raj Chetty, Stefanie DeLuca, Nathaniel Hendren, Lawrence F. Katz, and Christopher Palmer. 2019. "Creating Moves to Opportunity: Experimental Evidence on Barriers to Neighborhood Choice." *NBER Working Paper No.* 26164. National Bureau of Economic Research.

Callaway, Brantly, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics* 225(2): 200–30.

Chang, Eunsik, and Maria Padilla-Romo. 2023. "When Crime Comes to the Neighborhood: Short-Term Shocks to Student Cognition and Secondary Consequences." *Journal of Labor Economics* 41(4).

Chin, Mark J., David M. Quinn, Tasminda K. Dhaliwal, and Virginia S. Lovison. 2020. “Bias in the Air: A Nationwide Exploration of Teachers’ Implicit Racial Attitudes, Aggregate Bias, and Student Outcomes.” *Educational Researcher* 49(8): 566–78.

Corral, Daniel. 2021. “The Relationship Between Immigration Enforcement and Educational Attainment: The Role of Sanctuary Policies.” *AERA Open* 7: 23328584211037253.

Dee, Thomas, and Mark Murphy. 2018. “Vanished Classmates: The Effects of Local Immigration Enforcement on Student Enrollment.” *NBER Working Paper No.* 25080. National Bureau of Economic Research.

East, Chloe N., Annie L. Hines, Philip Luck, Hani Mansour, and Andrea Velásquez. 2023. “The Labor Market Effects of Immigration Enforcement.” *Journal of Labor Economics* 41(4): 957–96.

Fahle, Erin M., Belen Chavez, Demetra Kalogrides, Benjamin R. Shear, Sean F. Reardon, and Andrew D. Ho. 2021. *Stanford Education Data Archive Technical Documentation, Version 4.0, February 2021* [technical report]. Stanford Graduate School of Education.

Flores, René D., and Ariela Schachter. 2018. “Who Are the ‘Illegals’? The Social Construction of Illegality in the United States.” *American Sociological Review* 83(5): 839–68.

Frisvold, David E. 2015. “Nutrition and Cognitive Achievement: An Evaluation of the School Breakfast Program.” *Journal of Public Economics* 124: 91–104.

Gonzalez, Gabriela, and Caitlin Patler. 2021. “The Educational Consequences of Parental Immigration Detention.” *Sociological Perspectives* 64(2): 301–20.

Jacome, Elisa. 2022. “The Effect of Immigration Enforcement on Crime Reporting: Evidence from Dallas.” *Journal of Urban Economics* 128: 103395.

Kirksey, J. Jacob, and Carolyn Sattin-Bajaj. 2021. “Immigration Arrests and Educational Impacts: Linking ICE Arrests to Declines in Achievement, Attendance, and School Climate and Safety in California.” *AERA Open* 7: 23328584211039787.

Kirksey, J. Jacob, and Carolyn Sattin-Bajaj. 2025. “Future, Interrupted: Examining the Impact of a Large Worksite Enforcement Operation on Students’ Educational and Workforce Pathways.” *RSF: The Russell Sage Foundation Journal of the Social Sciences* 11(4): 123–41. https://doi.org/10.7758/RSF.2025.11.4.06.

Kraft, Matthew A. 2020. “Interpreting Effect Sizes of Education Interventions.” *Educational Researcher* 49(4): 241–53.

Laurito, Agustina, and Ashley N. Muchow. 2025. “Climate of Exclusion: Spillover Effects of Home-Country Natural Disasters on Immigrant Removals from the United States.” *RSF: The Russell Sage Foundation Journal of the Social Sciences* 11(4): 78–101. https://doi.org/10.7758/RSF.2025.11.4.04.

Legewie, Joscha, and Jeffrey Fagan. 2019. “Aggressive Policing and the Educational Performance of Minority Youth.” *American Sociological Review* 84(2): 220–47.

Legewie, Joscha, Amy Hsin, Niklas Harder, and Linna Martén. 2022. “Local Policing and the Educational Outcomes of Undocumented College Students.” *Sociological Science* 9: 406–29.

Meadows, Benjamin. 2021. “Undocumented and Under Threat of Deportation: Immigrant Students in the Classroom.” *Journal of Human Resources* 0621-11738R1.

Meadows, Benjamin, and Griffin Edwards. 2023. “Doubling the Cost by Splitting Families: The Effect of Immigration Laws Upon Hispanic Children’s Human Capital Formation.” Working paper. University of Alabama.

Menjívar, Cecilia. “The Racialization of ‘Illegality.’” *Daedalus* 150(2): 91–105.

Moinester, Margot. 2024. “The Control Boom: US Interior Immigration Enforcement, 1971–2010.” *American Journal of Sociology* 129(5): 1447–92.

Olden, Andreas, and Jarle Møen. 2022. “The Triple Difference Estimator.” *Econometrics Journal* 25(3) 531–53.

Park, R. Jisung, A. Patrick Behrer, and Joshua Goodman. 2021. “Learning Is Inhibited by Heat Exposure, Both Internationally and Within the United States.” *Nature Human Behaviour* 5(1): 19–27.

Patler, Caitlin, and Bradford Jones. 2025. “The US Deportation System: History, Impacts, and New Empirical Research.” *RSF: The Russell Sage Foundation Journal of the Social Sciences* 11(4): 1–24. https://doi.org/10.7758/RSF.2025.11.4.01.

Peacock, Ian G. 2025. “Imitate, Then Escalate: Social Influence and Its Consequences for the Subfederal Deportation System in the United

States." *RSF: The Russell Sage Foundation Journal of the Social Sciences* 11(4): 26–48. https://doi.org/10.7758/RSF.2025.11.4.02.

Reardon, Sean F., Erin M. Fahle, Demetra Kalogrides, Anne Podolsky, and Rosalía C Zárate. 2019. "Gender Achievement Gaps in U.S. School Districts." *American Educational Research Journal* 56(6): 2474–508.

Rossin-Slater, Maya, Molly Schnell, Hannes Schwandt, Sam Trejo, and Lindsey Uniat. 2020. "Local Exposure to School Shootings and Youth Antidepressant Use." *Proceedings of the National Academy of Sciences*, 117(38): 23484–89.

Shany, Adi. 2016. "Too Scared for School? The Effects of Terrorism on Israeli Student Achievement." Working paper. Hebrew University of Jerusalem.

Welsh, Richard O., and Walker A. Swain. 2020. "(Re) Defining Urban Education: A Conceptual Review and Empirical Exploration of the Definition of Urban Education." *Educational Researcher* 49(2): 90–100.