Do Immigration Raids Deter Head Start Enrollment?

By Robert Santillano, Stephanie Potochnick, and Jade Jenkins*

We investigate the local deterrence effect of immigration raids on Hispanic Head Start enrollment. Using a nationwide panel of raids from 2006 to 2008, a time of intensified community raids in the U.S., we find robust evidence that raids decreased Hispanic Head Start enrollment by around 10%. We further disentangle this effect by showing it is driven by a deterrence effect instead of a mobility effect. In other words, of families who are influence, most are keeping their children at home.
JEL: I24, J15, J18
Keywords: Early Childcare Education, Immigration, Raids

I. Introduction

Immigration raids can deter families from utilizing public services. Behaviorally, local raids by federal immigration officials increase the perceived probability of detection and deportation for mixed-status families—those with at least one undocumented family member. Due to this perceived risk, mixed-status families may attempt strategies to lower the probability of detection, such as by moving or disengaging with services. Because the vast majority of mixed-status families in the U.S. are Latino (Passel, 2011),

* Santillano: Mathematica, rsantillano@gmail.com; Potochnick, University of North Carolina Charlotte, spotochn@uncc.edu; Jenkins: University of California at Irvine, jvjenkin@uci.edu. All errors are our own.

and one quarter of Latino children have a parent who is an unauthorized immigrant (Clarke, Turner and Guzman, 2017), Hispanic children may be particularly susceptible to these effects.

This study investigates the local deterrence effect of immigration raids on Hispanic Head Start enrollment. Head Start is the largest federal early child-care education program in the United States and provides education, health, and other services to low income families (Zigler and Styfco, 2004). Children in mixed-status families are particularly vulnerable if this deterrence effect exists since they face multiple disadvantages (Karoly and Gonzalez, 2011), and benefits from attending preschool, particularly for English learners, have been found (Gormley Jr, 2008; Magnuson, Lahaie and Waldfogel, 2006). In fact, of the subgroups studied by the experimental nationwide study of Head Start, Hispanic English language learners benefited the most (Puma et al., 2010). At the same time, qualitative evidence suggests that immigration enforcement around Head Start centers hinders efforts to engage Hispanic families (Murguía, 2008).

Using a dataset of nationwide raids from 2006 through 2008, we extend the evidence in four ways. First, although many immigration enforcement studies have been conducted, we provide the first large-scale evidence on the causal impact of local immigration raids.[1] Second, we find robust evidence that immigration raids decrease Hispanic enrollment in Head Start by over 10 percent. Third, to our knowledge, we are the first to propose an empirical strategy that disentangles a mobility effect from a deterrence effect in response to local immigration enforcement. Finally, we find evidence that

[1]The closest study we know of in terms of scale and type of enforcement is Watson (2014) who provides evidence of decreased Medicaid enrollment for children of non-citizens when "deportable" apprehensions in a region rise.

the decrease in Hispanic enrollment is driven by a deterrence effect. In other words, of effected families, more are staying in their communities but not enrolling in Head Start.

II. Data

The data for this study were compiled from a range of sources. The most novel of these is a comprehensive panel of county locations and dates of federal immigration enforcement raids on workplaces, homes, and communities that were conducted between 2006 and 2008.[2] These were obtained by cross-checking raid listings from three immigrants-rights organizations—Centro Latino, Detention Watch Network, and Catholic Legal Immigration Network. These organizations tracked immigration raids through a variety of sources, including Freedom of Information Act (FOIA) requests, news tracking, and immigrant networks. The strength of these data are that we have high confidence that, nationwide, we know when a raid occurred and where it occurred. The primary limitation is that we do not know all the details of the raids. Sometimes information on the size, location, and number detained were collected, but these data appear incomplete. Because of that, we make no attempts in this study to classify raids for nuanced analyses.

The remaining data come from administrative sources and decennial censuses. For Head Start enrollment, we compiled data covering the 2003-04 through 2008-09 academic years from Head Start's Program Information Report. We also included enrollment data over the same time period for

[2] Immigrations and Customs Enforcement (ICE), the federal agency responsible for conducting raids, adopted an intensifed interior enforcement strategy in 2006. Based on the National Fugitive Operations Program of 2002, these raids were intended to be highly public and conducted in communities. In 2008, the Obama administration adjusted the nature of interior enforcement to targeted approaches for those suspected of criminal activity.

students in first grade from the National Center for Education Statistics' Common Core of Data. As discussed in the Methodology section, first grade enrollment is used to disentangle mobility from deterrence. Finally, we obtained county-level demographic characteristics from the 1990 and 2000 decennial censuses. These were obtained for descriptive purposes, as well as to help improve the comparability of counties given important demographic shifts in the location of Hispanics and "new destination" areas over our study time period (Lichter and Johnson, 2009).

Summary statistics for the universe of raided and never-raided Head Start counties are given in columns (1a) and (1b) of Table 1, respectively. Overall, there are 207 raided and 699 never-raided counties. Panel A of the table presents summary statistics on 3-year average enrollment across the 2003-04 to the 2005-06 academic years—before we observe any raid. Although raids could have occurred during this time period, they would have occurred before ICE intensified their interior enforcement efforts, and so we refer to it as the "pre-raid" time period (see footnote 2). Panel B of the table presents demographic characteristics covering overall population size in 2000, the Hispanic share of the population in 1990, and the level percentage change of Hispanics in the population from 1990 to 2000. The most important thing to notice from these first two columns is the large size and compositional differences between raided and non-raided counties.

III. Methodology

There are two primary challenges to identifying the deterrence effect of raids on Head Start enrollment. The first is the non-random location of raids, which could relate to other demographic trends that influence Head

Start enrollment. This challenge is confirmed by comparing the first two columns of Table 1. The second challenge is separating the deterrence effect of not enrolling in a public program and a mobility effect of moving to a location that is perceived to be less hostile. When focusing on locally enforced immigration policies, such as local partnerships between local law enforcement and federal officers (e.g., 287(g) and Secure Communities), both overall mobility of Hispanics and mobility of Hispanic school-aged children have been documented (Dee and Murphy, 2019; Watson, 2013). However, since Head Start is voluntary, any change in enrollment will represent a combination of both deterrence and mobility effects.

We address the identification challenges through a combination of flexible local-area matching, a comprehensive set of robustness checks, and a triple-difference design. For clarity, we discuss the triple-difference first. Given the policy is implemented on a subset of counties at different points in time, we first consider a difference-in-difference (DD) approach to estimate the impact of raids on Head Start enrollment. For the moment, assume the parallel trends assumption holds. In that case, the impact could be estimated with the following basic specification:

$$\ln(y_{ct}) = \alpha_c + \beta \times \text{PostRaid}_{ct} + \pi_t + \varepsilon_{ct}, \tag{1}$$

where y is enrollment in county c at time t, α and π are fixed effects, ε is an error term, and β represents the impact of interest, which is interpreted as the percentage change on enrollment. Since Head Start is voluntary, β is comprised of both a mobility effect and a deterrence effect. However, if we were to apply the same specification in (1) to first-grade Hispanic enrollment, where attendance is compulsory, the impact of raids would just represent the mobility effect for Hispanic students in the community. This is similar to how Dee and Murphy (2019) considered mobility when studying Hispanic students in grades K-12. We focus on first grade because it avoids any issues with potential "red-shirting" from kindergarten, yet it is for students who are closest in age to Head Start. To proceed, we fully interact (1) by each grade, stack models, and simultaneously estimate two DD impacts: β^{HS} and β^{G01} for Head Start and first grade, respectively. Finally, using a post-estimation approach, we take the difference between the two parameters to estimate a triple difference (DDD) in order to isolate the deterrence effect:

$$\text{Deterrence:} \qquad \beta^{\mathrm{DDD}} = \beta^{\mathrm{HS}} - \beta^{\mathrm{G01}}. \tag{2}$$

Due to well known auto-correlation issues with these models, we cluster all standard errors at the county level (Bertrand, Duflo and Mullainathan, 2004).

We now turn to a matching strategy to improve the credibility of comparisons. If the parallel trends assumption were actually satisfied, the above DD and DDD estimates would produce un-biased impact estimates. However, the already presented evidence makes this hard to believe. A common approach to check this assumption is to turn equation (1) into an event-study specification with leads and lags and test that the leads are not statistically significant (Angrist and Pischke, 2008). Because we have a relatively short pre-raid panel, and we also worry about the statistical power of such a test, we instead address the observed differences directly.

Taking inspiration from the synthetic control approach (Abadie, Diamond and Hainmueller, 2010), we created a full set of ordered matches for each raided county with a focus on balancing pre-raid enrollment patterns. Specifically, we matched raided counties to the donor pool of never-raided counties such that their pre-raid enrollment patterns were similar. To do this, we first removed all counties that did not have a balanced panel—meaning we only retained those with positive Hispanic enrollment for both Head Start and first grade across the six years.[3] This resulted in 174 raided counties and 418 potential donor counties. Next, focusing on one raided county at a time, we created three distance measures to each of the never-raided counties: (1) root mean squared enrollment differences across pre-raid years for Head Start; (2) root mean squared enrollment differences across pre-raid years for first grade; and (3) the Mahalanobis distance using the county population in 2000, the percent Hispanic in 1990, and the percentage point growth of Hispanics from 1990 to 2000 (Rosenbaum and Rubin, 1985). We created distances in this way because it emphasizes the importance of the full pre-raid enrollment trends for both Head Start and first grade, which is important given they represent outcomes of interest. Finally, we joined the three measures into one by again calculating the Mahalanobis distance between each treated county and all 418 never-raided counties.

Particular strengths of this approach are that it is flexible, non-parametric, and sets up a series of alternative samples that could be used as robustness

[3]It is important to note that imbalanced panels were mostly caused by missing data on all students, not just Hispanics. This could happen if a county were starting or ending a Head Start program and it is important to exclude them because these changes at the extensive margin are not likely to be influenced by raids. For raided counties, in cases where only Hispanics had zero enrollment for any year, we confirmed that all had experienced zero enrollment before the first documented raid, which implies that raids did not cause these counties to be imbalanced and, thus, excluded.

checks. The flexibility largely comes from the fact that we are able to match all pre-raid years for each of the raided counties even though raids occurred in different years. After the matching is done, we can estimate model (1) simply and with no additional controls given we worked to make the samples more similar. Finally, we can use the distances to create a range of intuitive samples to estimate impacts. We implement this in two ways. First, we exclude hard-to-match raided counties from the analysis sample based on those that have the largest distance from their closest match. Intuitively, the more hard-to-match raided counties we exclude, the more internal validity the impact estimates will have. The trade-off, however, is the external validity of the impacts to all raided counties. Second, we can include anywhere from 1 (e.g. nearest neighbor) to all 418 never-raided counties for each of the raided counties in the analysis. We match with replacement and give a donor a weight equal to the number of times it was matched with a raided county. We then normalize never-raided county weights to average one and we estimate weighted statistics for all matched samples. This results in estimates that can be thought of as treatment-on-the-treated impacts (Imbens, 2015).

We assess the credibility of our identification strategy in three ways. First, we compare pre-raid characteristics across raided and never-raided counties after matching. Second, we assess the robustness of the impacts across a wide range of matched samples after excluding up to 20 hard-to-match raided counties and including up to 40 never-raided matches for each raided county. Finally, we recreate the matched samples after excluding counties with an active 287(g) agreement over this time period. These agreements allow cooperation between federal immigration officers and local law enforcement, and Dee and Murphy (2019) found that they lead to a decrease in Hispanic K-12 student enrollment. For our sample of Head Start counties, this results in the exclusion of 30 raided counties and 16 never-raided counties from the donor pool.

Before turning to results, we need to more fully describe how post-raid treatment is defined. We define "post-raid" periods as all academic years following a raid occurring before October. For example, any raid that occurred between October 1, 2016 and September 30, 2017 meant that the 2017-18 academic year and all that followed were post-raid periods. This is because Head Start programs generally have a goal of being full by October, and that is the month that first grade enrollment is captured. There are a few limitations with this approach. First, Head Start enrollment is not based on a single point in time, but rather includes students who were touched by the program across the year. This means that a raid after October 1 could still influence enrollment. This may add a positive bias to the estimated impact if it increases enrollment "churn" for Hispanics in the first year of the raid, but only if this increase does not persist in following years. The second limitation is that we are unable to estimate dynamic effects. However, understanding dynamics is complicated by the fact that each county experiences a unique raid pattern. In separate work, we study these raids patterns, but for the purpose of this study, we do not believe we have to estimate dynamic impacts to convincingly answer the research question.

IV. Matching Results

Matching improves the comparability of raided and never-raided counties considerably. Table 1 presents two sets of columns reflecting two different matched samples. The first set of matches in columns (2a) and (2b) reflect all of the raided counties with a balanced panel and their best match. There is no longer a statistically significant difference across enrollment or overall county population. However, the magnitude of the differences are still meaningfully large, so the lack of statistical significance could reflect smaller sample sizes. Notice, the number of unique never-raided counties in (2b) is 91, which implies that many donors are repeat matches. Also notice that even though there is considerable improvement across the two groups, there are persistent differences in the percent of Hispanics in 1990 and the percentage point growth in Hispanics from 1990 to 2000.

Selecting a matched sample that excludes hard-to-match raided counties and allows more matches improves comparability even more. In columns (3a) and (3b), we exclude the 5 hardest-to-match counties and allowed up to three matches, which is the minimum number needed to get at least the same number of unique donor counties as raided counties.[4] Although allowing more matches improves statistical precision, it also decreases balance relative to the best match. Again, the comparability of enrollment and population is greatly improved, but the historic demographic differences for Hispanics persists. However, we want to emphasize that our identification assumption is parallel trends, not selection on observables. Therefore, even though we could not achieve balance on all characteristics, overall, we take this as

[4]The list of the five hardest-to-match counties are: (1) Los Angeles, CA; (2) Cook County, IL, including Chicago; (3) Harris County, TX, including Houston; (4) Maricopa County, AZ, including Phoenix; and (5) Orange County, CA.

evidence that the matching strategy worked.

V. Impact Results

Impact estimates for all samples included in Table 1, as well as those after excluding 287(g), are presented in Table 2. Across all samples, the negative DD impact on Hispanic Head Start enrollment, reflecting deterrence and mobility, is convincingly robust. It ranges from -8.6 percent to -12.7 percent, and is statistically significant at the 5 percent level across all but one sample. At around -3 percent, the negative DD impacts on Hispanic grade 1 enrollment, reflecting mobility, are robust in magnitude, but less so in statistical significance. Finally, when taking the difference across the two grades for the DDD impacts, the negative impact, reflecting deterrence, is consistent in magnitude although it is not a statistically strong finding. Combined, we take this as suggestive evidence that the deterrence effect is driving the decrease in Head Start enrollment for Hispanic families. Finally, Figure 1 presents a series of impact estimates for Hispanic Head Start enrollment for a range of matched samples—further demonstrating that the impact estimates are not influenced by any specific outliers.

VI. Conclusion

Similar to studies of other immigration enforcement efforts (e.g., 287g and Secure Communities), we find that federal immigration raids strongly reduce public service use (Amuedo-Dorantes and Lopez, 2017; Rhodes et al., 2015; Watson, 2014). We find that post-raid Head Start enrollment for Hispanic children decreased by over 10 percent. Furthermore, we find suggestive evidence that this decrease in enrollment is not solely due to the out-migration effects found in other studies (Dee and Murphy, 2019). Specifically, around

70 percent of the decrease in enrollment is explained by a deterrence effect where Hispanics in these communities are not enrolling in Head Start programs. This deterrence effect affirms the expressed fears of Head Start administrators (Murguía, 2008) and aligns with qualitative evidence on immigration raids (Capps et al., 2007; Chaudry et al., 2010). Importantly, unlike other immigration enforcement efforts, raids are not place based, so the relative importance of deterrence versus mobility is an important finding for understanding the mechanisms of these effects.

Finally, the decrease in enrollment raises further concerns that immigration enforcement is hindering the formative child-development years of young children of immigrants, the vast majority of whom are US citizens (Passel, 2011). Prior research indicates that as a result of immigration enforcement, Hispanic mothers access fewer pre-natal care services (Rhodes et al., 2015). Our study demonstrates that Hispanic families are less likely to access Head Start, which provides comprehensive health and educational services for preschool aged children and families (Zigler and Styfco, 2004). In combination, these results provide insights into federal efforts focused on revitalizing mass deportation: such efforts are likely to have harmful effects on the well-being of vulnerable, young Hispanic children.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. “Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program.” *Journal of the American statistical Association*, 105(490): 493–505.

Amuedo-Dorantes, Catalina, and Mary J Lopez. 2017. “The hidden educational costs of intensified immigration enforcement." *Southern Economic Journal*, 84(1): 120–154.

Angrist, Joshua D, and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion.* Princeton university press.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How much should we trust differences-in-differences estimates?" *The Quarterly journal of economics*, 119(1): 249–275.

Capps, Randolph, Rosa Maria Castañeda, Ajay Chaudry, and Robert Santos. 2007. "Paying the price: The impact of immigration raids on America's children."

Chaudry, Ajay, Randy Capps, Juan Manuel Pedroza, Rosa Maria Castañeda, Robert Santos, and Molly M Scott. 2010. "Facing Our Future: Children in the Aftermath of Immigration Enforcement." *Urban Institute*.

Clarke, Wyatt, Kimberly Turner, and Lina Guzman. 2017. "One quarter of Hispanic children in the United States have an unauthorized immigrant parent." *National Research Center on Hispanic Children & Families*, , (2017-28).

Dee, Thomas, and Mark Murphy. 2019. "Vanished classmates: The effects of local immigration enforcement on student enrollment." *American Educational Research Journal*.

Gormley Jr, William T. 2008. "The effects of Oklahoma's pre-k program on Hispanic children." *Social Science Quarterly*, 89(4): 916–936.

Imbens, Guido W. 2015. "Matching methods in practice: Three examples." *Journal of Human Resources*, 50(2): 373–419.

Karoly, Lynn A, and Gabriella C Gonzalez. 2011. "Early care and education for children in immigrant families." *The Future of Children*, 71–101.

Magnuson, Katherine, Claudia Lahaie, and Jane Waldfogel. 2006. "Preschool and school readiness of children of immigrants." *Social science quarterly*, 87(5): 1241–1262.

Murguía, Janet. 2008. "The implications of immigration enforcement on America's children."

Passel, Jeffrey S. 2011. "Demography of immigrant youth: Past, present, and future." *The Future of Children*, 19–41.

Puma, Michael, Stephen Bell, Ronna Cook, Camilla Heid, Gary Shapiro, Pam Broene, Frank Jenkins, Philip Fletcher, Liz Quinn, Janet Friedman, et al. 2010. "Head Start Impact Study. Final Report." *Administration for Children & Families.*

Rhodes, Scott D, Lilli Mann, Florence M Simán, Eunyoung Song, Jorge Alonzo, Mario Downs, Emma Lawlor, Omar Martinez, Christina J Sun, Mary Claire O'Brien, et al. 2015. "The impact of local immigration enforcement policies on the health of immigrant Hispanics/Latinos in the United States." *American Journal of Public Health*, 105(2): 329–337.

Rosenbaum, Paul R, and Donald B Rubin. 1985. "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score." *The American Statistician*, 39(1): 33–38.

Watson, Tara. 2013. "Enforcement and immigrant location choice." National Bureau of Economic Research.

Watson, Tara. 2014. "Inside the refrigerator: immigration enforcement and chilling effects in Medicaid participation." *American Economic Journal: Economic Policy*, 6(3): 313–38.

Zigler, Edward, and Sally J Styfco. 2004. *The Head Start Debates.* ERIC.

Table 1—Summary statistics for samples of raided and never-raided Head Start counties

	Full sample: Unbalanced panel		Matched sample: Exclude 0, Match 1		Matched sample: Exclude 5, Match 3	
	Raided	Not Raided	Raided	Not Raided	Raided	Not Raided
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)
Panel A: Average pre-raid 3-year enrollment						
Head Start: Hispanics	806	106†††	920	581	644	555
	(2,264)	(415)	(2,449)	(1,458)	(1,061)	(1,388)
Grade 1: Hispanics	2,593	212†††	2,976	1,781	2,006	1,450
	(7,314)	(841)	(3,595)	(3,570)	(3,492)	(3,268)
Panel B: County demographic characteristics from decennial censuses						
2000 Population (1,000s)	632	115†††	680	559	557	505
	(922)	(162)	(986)	(483)	(533)	(413)
Hispanics in 1990 (%)	9.4	5.3†††	10	6.4††	9.6	6.7††
	(13.6)	(12.8)	(13.9)	(11.3)	(13.9)	(13.1)
Hisp. 90 to 00 level Δ (%)	3.7	1.7†††	3.9	3.2†	3.8	2.8†††
	(3.1)	(2.5)	(3.2)	(2.9)	(3.1)	(2.6)
N: Counties	207	699	174	91	169	178

Note: For data source descriptions, see Section 2. For matched-sample descriptions, see Section 3. Standard errors are in parentheses, and a means test, which is weighted for the matched samples, is performed for each characteristic across paired columns.
† $p < 0.1$, †† $p < 0.05$, ††† $p < 0.01$.

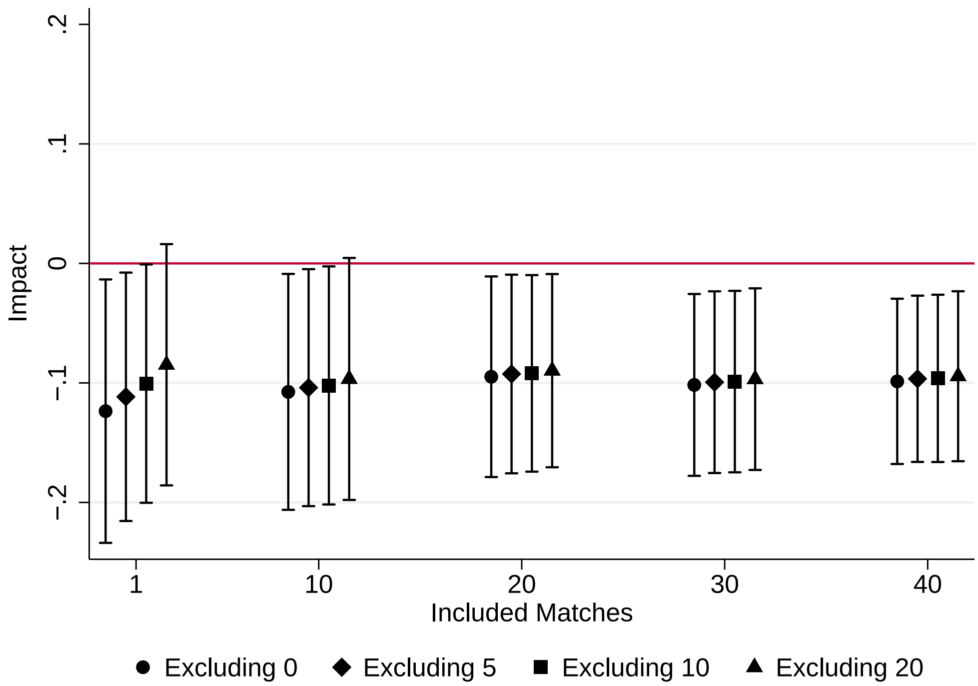
Table 2—Impacts of raids on ln(enrollment) from DD and DDD estimates for select samples

	Head Start	Grade 1	
	DD	DD	DDD
	mobility + deterrence	mobility	deterrence
Panel A: Unbalanced panel			
All Head Start counties	-0.111***	-0.033**	-0.078***
	(0.029)	(0.015)	(0.030)
Panel B: Balanced panel matches			
Exclude 0, Match 1	-0.124**	-0.037	-0.086
	(0.056)	(0.026)	(0.065)
Exclude 5, Match 3	-0.127***	-0.033*	-0.094**
	(0.046)	(0.019)	(0.044)
Panel C: Balanced panel matches excluding 287(g) counties			
Exclude 0, Match 1	-0.086*	-0.044**	-0.042
	(0.051)	(0.022)	(0.053)
Exclude 5, Match 3	-0.122**	-0.037*	-0.086*
	(0.049)	(0.019)	(0.045)

Note: The panel covers six academic years from 2003-04 to 2008-09. Difference-in-difference (DD) and Triple-difference (DDD) results from Panels A and B are estimated from the same samples presented in Table 1. See Section 3 for a description of excluded 287(g) counties. Heteroskedastic-robust standard errors clustered at the county level are in parentheses.

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

Figure 1. Impacts of raids on Hispanic Head Start enrollment (%), by matched sample



Note: Impact estimates from excluding hard-to-match raided counties and allowing for multiple matches. Bars represent 95% confidence intervals.