The Effects of School Consolidation on Students and Teachers: Evidence from an Underperforming System

Gustavo J. Bobonis

Reza Moradi

Orlando J. Sotomayor

Jeancarlo Vélez-Lebrón

Jessica Wagner*

November 5, 2026

Abstract

In recent years, a number of states and school districts in the US have engaged in large scale school consolidation reforms driven by infrastructure underutilization and the objective of reallocating scarce resources across schools. We investigate the determinants of this type of school consolidations, their effects on student enrollment and achievement, as well as the consequences for teachers, using data from over 400 consolidations of public elementary and middle schools across Puerto Rico during the period 2010-18. We document that school closures were orthogonal to students' academic performance at baseline; the strongest predictors of school closure are low aggregate enrollment levels and small average class sizes. Using a matched event study design that controls for students' enrollment histories, We find that school closures led to a 2 percentage points decline in enrollment and no effects on students' achievement in mathematics or language standardized tests, on average. Moreover, students displaced from under-(over-) performing schools experience large achievement gains (losses) relative to their non-displaced peers. These positive effects coincide with accessing teachers with higher (lower) baseline value added, and are much larger in magnitude compared to estimates from higher-performance school districts. We find limited evidence of spillovers from student and teacher displacement on receiving schools. Altogether, the results indicate that cost-saving consolidations could be made while limiting harm to achievement when schools are closed in the vicinity of adequate alternatives.

*Bobonis: University of Toronto, gustavo.bobonis@utoronto.ca; Moradi: University of Toronto, reza.moradi@mail.utoronto.ca; Sotomayor: University of Puerto Rico, Mayagüez, orlando.sotomayor@upr.edu; Vélez-Lebrón: University of Toronto, j.velezlebron@mail.utoronto.ca; Wagner: University of Toronto, j.wagner@mail.utoronto.ca. We are grateful to Jishnu Das, Robert McMillan, Michael Stepner, seminar and conference participants at the LACEA 2022 Annual Meetings for helpful comments; Alex Hogarty, Tyler Skura, and especially Cameron Panter for outstanding research assistance, and Dr. Damarys Varela-Vélez, Dr. Beverly Morro, Lydiana I. López Díaz, Emily Goldman, and Nicolás Riveros-Medelius for general support throughout. We acknowledge financial support from the William T. Grant Foundation, Spencer Foundation, and SSHRC Canada Research Chairs Program. All errors are our own.

1 Introduction

As patterns in where young people live and attend school fluctuate, whether they be gradual demographic shifts or sudden, persistent shocks – such as those affecting districts in the wake of the COVID-19 pandemic (Hubler, 2022) – education systems must adapt to changing demand. Jurisdictions facing declining enrollments and the under-utilization of infrastructure are confronted with the particularly difficult decisions of closing underutilized schools and reassigning displaced students and teachers across remaining schools. Though these consolidations may cause considerable disruption for both displaced students and those in schools that absorb them, they could also present opportunities to better assign students to effective teachers and schools.[1] Moreover, the distributional effects may vary significantly depending on the students who are impacted, the scale of consolidations, and the context and overall performance of the school system in which these take place.[2] Understanding the consequences of consolidating schools under varying conditions is critical for policy makers looking to minimize these disruptive costs while optimally reallocating students and teachers.

A comprehensive understanding of the consequences of closing schools requires analysis of a considerable number of heterogeneous closure events which are unrelated to students' prior achievement. The present study takes advantage of a unique setting that provides us with the opportunity to answer these important policy questions. During the period 2010-18, the Puerto Rico Department of Education (PRDE) consolidated over four hundred public schools due to chronically low and declining enrollment.[3] Further, closures occurred across both urban and rural areas, affected a variety of students who were displaced to schools of varying quality, distance to the closed schools, and peers as well as teachers' composition. Using linked longitudinal administrative data on both students and teachers for this period and leveraging this rich source of heterogeneity allows us to provide more granular evidence as to what drives the average and distributional effects of closures and offer more context for informing future policy decisions.

We first conduct a descriptive analysis to establish a set of stylized facts regarding the determi-

[1] Children tend to suffer academically from disruptions brought by school switching (Hanushek et al., 2004), or large intakes of new students into their schools (Imberman et al., 2012; Özek, 2021). Yet, there is also ample evidence that access to higher achieving peers (Sacerdote, 2011) and more effective schools (Angrist et al., 2017) and teachers (Chetty et al., 2014b) are highly beneficial.

[2] While the systems studied in prior work on school consolidations exhibit little variation in context or overall performance, a related literature on private school vouchers documents widely varying effects across different policy contexts and school systems (Epple et al., 2017).

[3] Puerto Rico experienced a drop in the population of primary school-aged youth (5-14 years old) from 504,912 to 343,697 or 32 percent, over this time period (authors' calculation using the Puerto Rico Community Survey).

nants of school closures in the PRDE context. We find that both low aggregate enrollment levels and small average class sizes were the main drivers of school consolidation in this context, consistent with PRDE policy. The relative influence of these variables shifted over time, with average class sizes being the strongest determinant of consolidation from 2010 to 2016, and low total enrollment gaining predictive power from 2013 to 2016 and becoming the key factor in 2017 and 2018 closures. Importantly, we document that school closures are orthogonal to students' academic performance at baseline, supporting our efforts to estimate causal effects of displacement for both students and teachers.[4]

To estimate the effects of consolidation on displaced students, we leverage the staggered timing of consolidation events across the territory and estimate a series of event study models of the effects of closures on students' enrollment and academic achievement. Specifically, we compare the trajectories of students displaced due to a school closure relative to those of similar non-displaced students with analogous enrollment histories in schools receiving no displaced students within the same region.[5] Further, we address the potential for spillover effects on those exposed to displaced peers by estimating spillovers directly in an analogous analysis that compares students in absorbing schools relative to similar students in non-absorbing ones.

We find that being forced to relocate due to a school closure leads to a considerable 1-4 percentage points decrease in students' enrollment in the public school system up to 4 years after displacement. Similarly, we find no average causal effect of displacement on achievement in Math, Spanish, or English scores in overall terms. We also document no average detrimental spillover effects of exposure to displaced peers or teachers among students in receiving schools. This evidence is generally consistent with the existence of some disruption effects of school consolidation efforts on students in the public school system.

We further examine the distributional effects of the consolidation policy, using the rich variation in closure conditions to conduct a comprehensive investigation of heterogeneity underlying these average effects. First, we find that students displaced from schools near higher-achievement alternatives experience large achievement gains of up to 0.38 of a standard deviation in the four years

[4] In general, quantifying the effects of consolidation on student achievement presents an empirical challenge as schools are often selected for closure on the basis of low performance, confounding naïve comparisons of the displaced to the undisturbed (see, e.g., Sunderman and Payne (2009), Engberg et al. (2012), and Steinberg and MacDonald (2019).)

[5] With repeated student observations and many closures occurring each year, we are able to control for characteristics of the displaced and control students that are fixed over time, as well as one-time shocks that are felt across the island, using a two-way fixed effects estimator. We also implement newly developed techniques from the emergent literature on staggered two-way fixed effects models to correctly estimate dynamic treatment effects in the presence of heterogeneity across closure periods.

following consolidation. In contrast, those whose neighboring schools had lower average scores at baseline saw students do worse, with losses of up to 0.21 of a standard deviation in test scores. In terms of mechanisms, we show that this heterogeneity is systematically related to students having access to schools with more or less effective teachers as measured by their value-added at baseline, and to a lesser degree by access to schools with smaller or larger class sizes. In summary, we find that access to higher achieving peers and more effective schools and teachers are beneficial for displaced students.[6]

We also consider whether displacements that are more disruptive to the peer and community network can lead to detrimental effects on students' achievement.[7] To do so, we examine heterogeneity in impacts for students with a greater concentration of peers displaced to the same schools (a measure of peer network continuity) and for those with shorter distances to the schools of origin. Consistent with this channel, we estimate stronger academic achievement effects in cases where students are displaced shorter distances, those with a greater share of their original peers in the same receiving school, and those with more local schooling options. Finally, we document effects of the consolidations on displaced teachers: although they are less likely to be assigned to another school in the year following the consolidation, they are approximately 3 percentage points more likely to remain in the public school system in subsequent years. This also points to modest short-term disruptive effects for teachers exposed to such decisions.

We contribute to the literature on the effects of school closures on student and teacher outcomes in several ways. Most of this work examines consolidations in the United States and Northern Europe, where average incomes are high. This paper is the first to comprehensively estimate student achievement effects in a setting where students come from relatively poor households, and aggregate performance across the system is low relative to average incomes.[8] We expand on prior work that finds closing remote schools in rural China lowered female school enrollment (Hannum et al., 2021). In Puerto Rico primary schools, dropout is unusual across genders, likely due to compulsory schooling laws and high demand for primary education, though system exit to private

[6] In contrast, we do not find strong evidence of heterogeneity in displacement effects by student's gender, special education status, or achievement at baseline.

[7] This analysis contributes to the growing literature documenting the consequences of school closures and consolidations on community networks (Brummet, 2014; Steinberg & MacDonald, 2019).

[8] According to data from the U.S. Bureau of Economic Analysis, Puerto Rico's GDP per capita was just under half that of the United States in 2021, and is significantly lower than that of the poorest state. Yet, the territory is considered high income by World Bank standards, with average incomes on par with countries like Portugal and Greece (The World Bank, 2022). Despite this, students in Puerto Rico score relatively low on the Program for International Student Achievement (PISA) tests. In 2012, 15-year-old Puerto Rico students sampled from both the public and private system scored 382 on the PISA math component, compared to 481 in the United States, and more in line with middle-income Latin American countries such as Colombia, Peru, and Brazil (Chan et al., 2014)

schools and emigration are common. In addition, we might expect less negative effects given closures in Puerto Rico did not reassign students across such far distances. Indeed, we observe modest declines in enrollment following consolidation, and produce novel estimates of achievement effects in a developing country context.

We also complement findings on the achievement effects of consolidation in high-income countries. This research typically finds average effects on achievement that are small, negative, and often statistically insignificant, both in cases where the primary motivation for closing schools was under-utilization, as in studies on Michigan (Brummet, 2014) and Denmark (Beuchert et al., 2018), or a combination of under-utilization and low performance, as in Sweden (Taghizadeh, 2020a, 2020b), Chicago (De la Torre & Gwynne, 2009), Philadelphia (Steinberg & MacDonald, 2019), and Pittsburgh (Engberg et al., 2012).[9] Studies based in the U.S. show that students tend to improve when they are displaced to higher performing schools, and non-displaced students experience negative spillovers from having displaced peers (Brummet, 2014; Engberg et al., 2012; Steinberg & MacDonald, 2019), though null effects are documented in Sweden (Taghizadeh, 2020a). While Puerto Rico is a U.S. commonwealth, average incomes are substantially lower than in these locales, the private school system enrolls a larger share of total students, and students are underperforming even when compared to many high middle-income Latin American countries. Prior research has found the effects of education policies and interventions to be highly context specific. For example, studies have shown positive effects of attending high performing schools in lower-income Romania (Pop-Eleches & Urquiola, 2013), but not in New York or Boston (Abdulkadiroglu et al., 2011).[10] Relative to the literature on high-income countries, we find largely consistent consolidation effects, though the mediating role of school quality is much larger in magnitude. These results further our knowledge of the consequences of consolidation disruptions and changing school quality in an education system whose performance resembles that of other middle-income countries.

A closely related study by Caraballo-Cueto (2023) examines the large wave of school closures in Puerto Rico in 2017–2018 and documents negative effects on displaced students' academic achievement. We build on this evidence by examining a broader set of school consolidations over a longer period, using a more comprehensive longitudinal research design, and investigating how differences in school and teacher quality shape the consequences of displacement. While the previous study

[9] Studies of purely performance-based closures have documented positive achievement effects under these circumstances, as most students are displaced to higher performing schools (Bifulco & Schwegman, 2020; Carlson & Lavertu, 2016).

[10] A related body of work finds highly setting-dependent results of private school vouchers (see Epple et al. (2017) for a review of this literature).

finds negative average achievement effects, our estimates suggest that these effects are more heterogeneous, with the consequences of displacement depending importantly on the quality of the educational alternatives available to students.[11]

Lastly, we contribute a new understanding of the mechanisms underlying consolidation effects by investigating the role of a rich set of learning inputs, as well as placing added emphasis on enrollment and sorting factors, which are often overlooked. This study is the first to analyze whether consolidations provide students access to higher value-added teachers, though a recent study examines other proxy measures of teacher quality such as experience (Taghizadeh, 2020b). We also provide a novel analysis of class size effects, which are especially relevant in the case when schools are closed due to low enrollment. We follow the prior literature in estimating heterogeneous treatment effects by displaced student concentration, and build on this using additional related measures of the median distance between closed and receiving schools and the number of school options in a five kilometer radius. Finally, given the potential for system exit that may be endogenous, we make careful consideration of enrollment effects of treatment. Sorting and attrition effects are given little attention in the literature, with the exception of Beuchert et al. (2018), and are especially salient for enrollment-based closures. We confirm several findings of the earlier literature while obtaining new insights into the diverging pathways that lead students to benefit or lose out as a result of a school consolidation. These insights provide a useful guide for optimizing school consolidations in the future. It remains to be determined how these academic trade-offs weigh against the cost-savings generated from downsizing school infrastructure.

The remainder of the paper is organized as follows. Section 2 provides background information on the PRDE and the political context of consolidations; Section 3 describes the data used for analysis, including a descriptive summary of the data and closures; Section 4 outlines our empirical method; Section 5 details results, and Section 6 concludes.

[11] Caraballo-Cueto (2023) examines the 2017 and 2018 school closures using student-level standardized test data from 2017–2019 and difference-in-differences and propensity-score matching estimators. Our analysis extends this work in several important dimensions. First, we study the broader set of school consolidations over 2012–2018, rather than focusing exclusively on the two most recent waves of closures. Second, we exploit detailed longitudinal administrative records linking students to schools and teachers over time, and implement a matched event-study design that compares displaced students with non-displaced students with similar prior enrollment histories, carefully selecting comparison students to limit potential contamination from displacement spillovers. Third, we use more detailed, subject-specific standardized test scores rather than categorical proficiency measures, allowing us to examine achievement trajectories and treatment-effect heterogeneity more precisely. Finally, we investigate the mechanisms underlying these effects by measuring the quality of both closed and alternative schools, including differences in baseline academic performance and teacher value added. These extensions allow us to distinguish the disruptive effects of displacement from the potential benefits or costs of accessing schools and teachers of different quality, and to characterize the substantial heterogeneity underlying average consolidation effects.

2 Background and Context

2.1 Puerto Rico Department of Education

The PRDE operates a unified school district serving the entire territory. Serving a population where approximately 57% of children live in poverty, the PRDE has historically lagged behind all other state-level school systems in the US in academic performance (Ladd & Rivera-Batiz, 2006). National Assessment of Education Progress (NAEP) test scores support this assertion: 85% of fourth graders in PR do not demonstrate basic proficiency in mathematics, as compared to 20% in the nation or 29% in the worst performing state (National Center for Education Statistics, 2019). The academic achievement of public school students in the territory also lags that of their private school counterparts. According to the Programme for International Student Assessment (PISA) 2012 mathematics literacy test, the scores of 15-year old students on the PISA scale for students attending public schools are significantly lower on average than those attending private schools, and are comparable to that of students in severely underperforming middle-income countries such as Colombia, Peru, Brazil, Indonesia, Jordan, and Tunisia (Chan et al., 2014). As a reflection of citizens' perception of the quality of state-provided education, about a quarter of students enroll in private schools, a significantly higher rate than in the mainland US (Ladd & Rivera-Batiz, 2006).

2.2 School Consolidations in the PRDE

Once one of the largest school districts in the US, enrollment in the PRDE began falling in the 1980s as a result of increased enrollment in private schools, a decline in fertility, and migration outflows that further reduced the school-age population (Ladd & Rivera-Batiz, 2006; Hinojosa et al., 2019; Abel & Deitz, 2014). Despite the long-term trend, substantial school closures did not take place until the 2010s. During the 2008-12 political cycle governed by the right-of-center New Progressive Party (PNP), 38 out of a total of 1509 schools operating at the time were closed in AY 2011 in response to the enrollment decline, with an additional 25 schools closed in the following two years. During the same time period, five new schools opened and 95 other were remodeled.[12]

The first year of the following political term, governed by the left-of-center Popular Democratic Party (PPD), began with just 9 closures, but a system restructuring initiative led to 79, 60, and 43 closures in the following three years, respectively (Comisión de Derechos Civiles, 2018). Depart-

[12] Comisión de Derechos Civiles de Puerto Rico (Comisión de Derechos Civiles, 2018), https://de.pr.gov/uncategorized/escuelas-para-el-siglo-21-2/.

ment policy established in the early 2000s and marginally revised over time, stated that closure evaluation committees had to take into account a number of criteria that included actual and projected enrollment, number of employees and operational costs, health and safety indicators, state of the infrastructure, academic indicators, absence of schools nearby, and schools offering unique academic programs. Underscoring the importance of enrollment in making closure decisions, regulations mandated an inventory of schools with enrollment under 100 students, regardless of whether these schools were recommended for closure or not (see Departamento de Educación (2006, 2007, 2009, 2013, 2014, 2015)).[13] An additional factor weighting on closure decisions was PRDE policy establishing minimum class size limits.[14]

This system restructuring followed a Boston Consulting Group study that established school occupancy rates averaging 71% in 2014 and falling to 55% in 2021 if enrollment trends continued. To reach a goal of 90% occupancy, it estimated that 300 schools would have to be reconfigured by AY 2016 and a total of 580 would have to close by 2021. The study suggested moving displaced students to schools (i) with equal or better performance in Mathematics and Spanish standardized tests, (ii) with better infrastructure, and (iii) within 3 to 4 miles of sending schools. A re-organization conducted in this manner would not only have the potential for improving academic achievement, but would also free up resources in the US state or territory devoting the largest share of its budget to administrative costs (Boston Consulting Group, 2014).

In May of 2017, the Secretary of Education appointed by the newly-elected PNP governor announced its intention of closing 179 schools in the 2017-18 academic year, with additional closures to be scheduled for 2018-19. The Secretary stated that over 40% of schools had an occupancy rate of under 60% and the closure rounds aimed to increase the system's rate to 85%. Academic achievement did not play a role in decision-making (Comisión de Derechos Civiles, 2018). In the end, the damage and destruction caused by the passing of Hurricane Maria in September of 2017 resulted in the closure of 259 schools in 2017-18 and an announcement of additional round of 283 closures scheduled for 2018-19, leaving 828 schools open during this academic year (Comisión de Derechos Civiles, 2018). These closures were met with substantial social, political, and academic resistance. Studies challenged the proposition that the closures resulted in cost savings or improved academic achievement, and reports lamented the increased incidence of derelict schools (Cueto,

[13] In 2014 a "system restructuring" criterion was given prominence in closure regulations. In 2015 the threshold for inclusion in the inventory of schools was changed to any school undergoing substantial and continued decline in enrollment.

[14] See Departamento de Educación (2007).

2020; Lafarga Previdi & Vélez Vega, 2020). The Law of Educational Reform of 2018 made it more difficult for closures to take place, mandating public hearings, and publicly-available studies outlining the reasons and ascertaining the potential impacts of school closures. Since then, few schools have closed as a result of falling enrollment. However, the number of students served by the school district continues to fall steadily.

Figure 1 depicts the total number of primary schools closed each year and the number remaining in operation, where closures are registered in the academic year when school enrollment falls to zero.[15] The figure illustrates that closures were common throughout the decade, but were largest in AYs 2017/18 and 2018/19. The last wave of consolidations removed a quarter of schools from operation. From here forth we will refer to the academic year using the second calendar year as our key outcome of interest, standardized test scores, are measured in Spring. Table 1 provides a summary of school-level measures of interest across closed and operating schools in each political term: 2010 to 2012, 2013 to 2016, and 2017 to 2018. Closed schools consistently had lower enrollment and class sizes over time, with slightly higher shares of high-poverty students and those with learning impediments. They also tended to enroll higher performing students, on average, and be disproportionately rural. Closed schools also appear somewhat more likely to occur in municipalities for which the territory's opposition party is in power locally, potentially implying favoritism in the decision to close a school.

3 Data and Descriptive Statistics

3.1 Data

We employ a large panel of administrative data from the universe of students, teachers, and schools in the Puerto Rico public education system from academic years 2010 to 2018. In total, 398 primary schools (serving kindergarten to grade 8) and 12 secondary schools were closed in this period.[16] Given the few high school consolidations, we limit our focus to primary schools. Focus is also placed on consolidations between 2010 and 2018, as we consider the post-Maria (2019) closure round to be a special event whose analysis is further hindered by the covid-19-related absence of standardized test data in 2020 and 2021. The resulting sample includes about 2.7 million student-year records

[15] A closure in (say) AY 2011/12 indicates that the school's final year of operation was AY 2010/11. The true date of the closure would be some time in Summer 2011.

[16] We observe two cases in which a school is relocated to a nearby facility (within 0.5 kilometer) with nearly identical student compositions. We consider these to be relocations, rather than closures.

describing 650,000 students, of which 46,068 are affected by a closure. The data link students to their teachers and schools each year and include detailed student attributes from enrollment records, such as gender, age, grade, poverty status, and special education status, and full academic histories from transcripts and standardized tests.

We rely on standardized test performance as our key measure of academic achievement. The PRDE administered the *Pruebas Puertorriqueñas de Aprovechamiento Académico* (PPAA) exams in April of each year to students in grades 3 through 8 in Spanish, Mathematics, and English. As in many state education settings, these tests were developed to be compliant with the common standards required by the No Child Left Behind Act. We normalize test score outcomes within grade and year to have a mean of zero and standard deviation of one.

We use linked student-teacher data to estimate teacher value added (TVA) for those who taught tested grade-subjects in the base period of 2010 to 2013.[17] We follow the methodology developed by Chetty et al. (2014a) for estimating teacher value added, implementing a fixed effects approach that residualizes test scores on a control vector including: a cubic in lagged own- and cross-subject scores, by grade; student-level characteristics including gender, age, and indicators for poverty status and special education status; school-grade means of student-level characteristics; and grade and year fixed effects. To accommodate our relatively short estimation window, we do not allow teacher quality to drift over time and therefore use a constant factor to shrink value added estimates toward the sample mean. In this regard, our estimation strategy closely parallels that developed by Kane and Staiger (2008).

3.2 Descriptive Summary of School Closures

We begin our analysis by documenting the timing and nature of school closures over the sample period. To establish the nature of school closures and how these varied across PRDE administrations, we examine correlates by political term. We test the relative importance of a number of school attributes in predicting closure using a simple regression framework. Figure 2 presents coefficient estimates and 95% confidence intervals from a linear probability model regressing relevant school-level factors on an indicator for term-specific closure.[18] Despite average differences in many measures depicted in Table 1, the regression estimates suggest that enrollment factors were the

[17] We restrict this estimation for a baseline period as using data from later years would incorporate treatment effects of consolidations into teacher effects. Any analysis using these measures will exclude the baseline period from which TVA is estimated.

[18] Schools are included in the control group if they did not close in a given term, but may have closed in a future term. For closed schools, measures are taken from the last year before closure.

driving force behind school closures across political administrations. Low average class sizes (below 15 students) were associated with a 5 percentage point increase in closure probability between 2010 and 2012, and a nearly 20 percentage point increase between 2013 and 2016. Low enrollment is strongly predictive of the most recent and largest wave of closures, with schools in the bottom quintile of total enrollment having a 25 percentage point higher closure probability than schools in the second quintile.

After conditioning on enrollments, school closures do not appear to be selected based on the average performance of students, a one-year change in math performance, student personal attributes, or local political environments. Although smaller schools tended to be in rural areas and enroll higher performers, we uncover no evidence that these factors motivated closure decisions independent of school size. These selection patterns are of key importance as we consider estimating the causal effect of being displaced by a closure on student achievement.

Summary statistics for the full student sample are provided in column (1) of Table 1. The sample is evenly split on gender and rural school attendance, and nearly a quarter of students are classified as requiring a special education accommodation at some point in the sample period. The average age of students is just under 12 years old and the typical class held 23 students. Test score outcomes are standardized to have mean zero and standard deviation one, and average baseline TVA is also close to zero by construction. Columns (2) to (4) present summary statistics for the empirical samples described in the following section, and will be discussed in turn, as will outcomes of interest and measures of closure conditions.

4 Empirical Methodology

4.1 Average Direct Effects of Displacement

We estimate the causal effect on enrollment and student achievement of being displaced from a closed school using an event study model. Students are defined as having been exposed to a school closure if they attended a school in its final operating year and were not enrolled in its final grade, in which case they would be expected to relocate regardless.[19] Sample characteristics of these students are summarized in column (2) of Table 1. Displaced students are somewhat more likely than the overall sample to be special education status or attend rural schools, though they perform

[19] In rare cases (approximately 3% of students affected by closures) where students are displaced by more than one closure in the sample period, we define their exposure only by way of the initial closure.

similarly on standardized tests. Their teachers have slightly higher value added at baseline, and they experience smaller class sizes of under 21 students per class.

The panel nature of the data allows us to follow individual students as they leave closed schools and enroll in receiving schools, and compare their enrollment and performance on standardized tests before and after these events. Thus, as in a difference-in-difference approach, we compare treated students to control students pre- and post-treatment, using binary measures of time since treatment.Let y_{it} denote the standardized test score of student i in year t, and C_i the first year in which student i is forced to attend a new school due to their existing school closing. A series of treatment measures are defined by the year relative to the displacement event, $D_{it}^k = \mathbb{1}\{k = t - C_i\}$. The event study regression model is as follows:

$$y_{it} = \sum_{k=-4}^{-2} \tau_k D_{it}^k + \sum_{k=0}^{3} \tau_k D_{it}^k + \alpha_i + \lambda_t + \delta_r + \epsilon_{it} \tag{1}$$

where α_i are student fixed effects capturing differences across treated and untreated students and any other fixed student factors, λ_t are year fixed effects, and δ_r are seven regional fixed effects. The period relative to treatment is denoted k, and we omit $k = -1$ as the base period. The parameters of interest, τ_k, capture the displaced students' change in test scores from the year before displacement to k years after displacement, relative to non-displaced students in the same years. The control group comprises students who were not affected by a school closure at any time in the sample period, as we exclude observations for treated students outside the four-year window pre- and post-displacement.

Under certain conditions, this two-way fixed effects specification extends the causal difference-in-difference model to allow for staggered treatment timing and dynamic effects of the treatment. The key assumption for identification of the treatment effects, τ_k, is that in the absence of a closure disruption, displaced students would have experienced parallel achievement trajectories to those who were not displaced. This is commonly referred to as the parallel trends assumption. The nature of Puerto Rico's school closures present a potential challenge to this assumption. Recall that schools were closed to address very low, and in some cases declining enrollment. As a result, students who remain in schools up until they close may be more attached to their local public school for reasons that correlate with the path of academic performance.

In fact, in any given year students never displaced by a closure are observed in the sample 81.6% of the time that we would expect them to be had they enrolled in the public system for all

primary grades.[20] On the other hand, displaced students are, by definition, always in-sample the year before a closure, as displacement is based on enrollment in a closing school the last year it is open. The year prior 93.8% of these yet-to-be treated students are still in the sample, and this declines to 90.8% another year earlier – still much higher than the untreated average. Thus, we see some evidence that students affected by closures exhibit diverging enrollment histories in the pre-treatment period relative to the control group. De-meaning on grade, year, and region reduces these differences only somewhat.

To address this potential source of bias, we employ a matching strategy that balances enrollment histories across treated and control units. Specifically, we use the coarsened exact matching approach developed by Iacus et al. (2008), enforcing an exact match on region and a 3-year enrollment history and allowing for coarser matches on grade.[21] We exclude students who may have been indirectly affected by closures from the pool of potential matches, dropping those for whom more than 5% of their peers were displaced the past 4 years.[22] Students are matched with replacement, where each treatment cohort is matched to never-displaced students whose enrollment history, region of enrollment, and grade resembles those of the displaced students in the year before they are displaced. The resulting match weights balance the treated and control groups along these dimensions within each match cohort. Our preferred specification uses the matched data and amends equation (1) to include match weights and match-cohort-by-year fixed effects, and drops region fixed effects which become redundant, as follows:

$$y_{it} = \sum_{k=-4}^{-2} \tau_k D_{it}^k + \sum_{k=0}^{3} \tau_k D_{it}^k + \alpha_i + \lambda_{tm} + \epsilon_{it} \tag{2}$$

where match weights w_{im} for student i in match cohort m are applied, and λ_{mt} are match cohort-by-year fixed effects.

It remains possible that trends in school and class size which motivate closure have confounding effects on academic achievement. An advantage of the event study framework is that we may visually inspect the time path of relative outcomes across the treatment and control groups in the pre-treatment period, and conduct statistical tests for jointly nonzero pre-treatment coefficients.

[20] We calculate this share by expanding each student's panel backwards to kindergarten (or 2010, if it comes first) and forwards to grade 8 (or 2018, if it comes first), and dividing the total *observed* student-years by the total number of student-years in the expanded panel.

[21] Coarsened exact matching is implemented computationally using the Stata command `cem` from Blackwell et al. (2009).

[22] An analysis of spillover effects on non-displaced peers at schools which receive displaced students is carried out in Section 5.3.

We report the p-value from these joint tests in all results figures, and explore the role of evolving enrollments and class sizes as a potential mechanism in Section 5.4.4.

Two additional identification assumptions are required in order to interpret τ_k as the average treatment effects on the treated. First, the no anticipation assumption asserts that no causal effects of school closure be realized prior to the actual closure of the school. This would be a concern if students or their parents, teachers, and principals are aware of the impending closure in advance and change their behavior as a result. In Puerto Rico, school consolidations were typically announced suddenly by the PRDE, with little advance knowledge by local education administrators and parents. As with parallel trends, inspection of pre-treatment coefficient estimates provide empirical evidence for whether this assumption is satisfied.

Lastly, recent work examining the properties of difference-in-difference and event study estimators that employ two-way fixed effects models have shown that estimates may be biased in the presence of heterogeneous effects across cohorts treated at different times (see Roth et al. (2022) for a recent review). Thus, a third identification assumption is that average treatment effects on the treated do not differ across students displaced at different points in time. Sun and Abraham (2021) evaluate the event study model in particular and develop an estimation method to correct bias in the presence of these heterogeneous treatment effects. Given the varying contexts of closures over time, it is plausible that this assumption is violated in our context. Thus, we use Abraham and Sun's interaction-weighted estimator to correct for this potential bias.[23]

4.2 Spillover Effects on Non-Displaced Students

We also estimate the effects of school consolidations on those students who were not displaced, but experienced an influx of new peers as a result local closures. Adapting the approach from Brummet (2014), we define f^P_{sgtk} as the share of a non-displaced student's peers in school s, grade g, and year t who were displaced by a closure k years ago. Exposure to displaced teachers is defined through an analogous measure, f^T_{sgtk}. We estimate the following regression model for non-displaced students alone:

$$y_{it} = \sum_{k=0}^{3} [\pi^P_k f^P_{sgtk} + \pi^T_k f^T_{sgtk}] + \alpha_i + \lambda_t + \delta_r + \nu_{it} \tag{3}$$

[23] Sun and Abraham (2021) show that the average treatment effect for a given relative period k can be decomposed into a linear combination of cohort-specific average treatment effects for the period of interest and cohort-specific average treatment effects from other relative treatment periods. Under homogeneous treatment effects, these contamination effects cancel out when taking a weighted sum across cohorts. Their interaction-weighted method estimates cohort-specific average treatment effects before aggregating them to an overall average treatment effect, eliminating the source of contamination.

where π_k^P and π_k^T are the parameters of interest measuring spillover effects from closures. Table 1 summarizes the characteristics of the spillovers sample sample in column (4). These students reflect the overall sample in most respects, as they make up the vast majority of students. Though the average share of peers recently displaced is low overall, conditioning on any displaced peers generates more variation. The median student in schools receiving any just-displaced students or teachers had 3% newcomer peers, or about 2 students, and 7% newcomer teachers. Values in the right tail of the distribution approach 20% and 25% of peers and teachers, respectively.

4.3 Heterogeneous Treatment Effects by Consolidation Characteristics

Whether students realize benefits to or costs from school consolidations is likely to depend on the nature of both their closed schools, and the schools to which they are reassigned. To better understand the experiences of displaced students, we define several measures that aim to capture changes in peer, teacher, and school inputs resulting from closures. First, we follow Brummet (2014) in measuring changes in school quality by calculating the difference between average test scores of the closed school and the average test score of all schools within 5 kilometers of the closed school, in the base years of 2010 to 2012. These schools enroll over 85% of displaced students, and the resulting pre-determined measure is free of potential bias from student selection of their actual school. We define $\Delta^S_{s(i,C_i)}$ as the relative performance of school s in 2010 to 2012, attended by student i in the year before closure C_i, or Δ_i^S for short. Similarly, we measure relative teacher performance by differencing estimates of teacher value added across the same set of schools in the base years of 2010 to 2012, and denote this Δ_i^T.[24] We calculate the median distance between students' closed and new schools to capture the degree of geographic displacement. A fourth and related measure estimates the continuity of displaced peer networks by calculating the share of students who are displaced to the modal receiving school.

Panel C of Table 1 summarizes variables that characterize the consolidation conditions experienced by displaced (column (2)) and non-displaced students (column (3)). It depicts the typical treated student as displaced from a school that had similar average test scores and teacher value added to neighboring schools at baseline, and sent students to receiving schools 2.34km away with 47% of their peers. There is substantial variation in these measures, allowing for a rich examination to uncover the ideal conditions for school consolidation.

[24] We define analogous variables for students in receiving schools by taking the average of each of these measures among their displaced peers and scaling this by the share of displaced peers.

We employ our measures of consolidation conditions to investigate their mediating effects on achievement outcomes. We use both measures that are predetermined, and thus not subject to selection effects, and those that can only be determined post-treatment, all of which are summarized in Table 1 and described above. For each measure, we group displaced students roughly into terciles of the mediating variable. For example, using the relative achievement at baseline measure, we categorize closed schools into three groups: underperforming, average performing, and overperforming, as follows:

$$
\begin{aligned}
low\Delta_i^S &= \mathbb{1}\{\Delta^S_{s(i,C_i)} \in (-\infty, -0.25)\} \\
med\Delta_i^S &= \mathbb{1}\{\Delta^S_{s(i,C_i)} \in [-0.25, 0.25]\} \\
high\Delta_i^S &= \mathbb{1}\{\Delta^S_{s(i,C_i)} \in (0.25, \infty)\}
\end{aligned}
$$

We then estimate the following regression, interacting dummy variables for each group with our event study parameters:

$$
y_{it} = \sum_k \left[\tau_k^{\ell} D_{it}^k low\Delta_i^S + \tau_k^m D_{it}^k med\Delta_i^S + \tau_k^h D_{it}^k high\Delta_i^S \right] + \alpha_i + \lambda_{tm} + \epsilon_{it} \tag{4}
$$

where λ_{tm} are year by match-year fixed effects. For relative school and teacher performance, the regression is estimated using years since 2012 to avoid confounding the outcome, y_{it}, with the source of heterogeneity, which is calculating using y_{it} for these variables.

5 Results

5.1 Average Direct Effects on Enrollment

Figure 3 plots event study estimates where the outcome is being observed in the sample in a given year. For each student, we create a full panel from kindergarten (or 2010 if it comes first) to grade 8 (or 2018 if it comes first) and record whether they were actually observed in a given year in the PRDE system. The figures illustrates the parallel trends in pre-treatment enrollment ensured by the matching approach, alleviating concerns about selection of the treated students on system attachment.[25] As expected, school consolidations causes some system exit. As the panel does not

[25] To rule out selection on other time-varying observables that may suggest a parallel trends violation, we estimate this event study on additional measures. No such concern arises for measures of student poverty, special education status, enrollment in a rural school, or grade point average.

extend beyond eighth grade or 2018, we investigate whether this possible fade-out is due to a higher rate of exit for cohorts treated close to the end of the panel or in later grades. We find no evidence in support of this explanation, and instead observe displaced students leaving from and returning to the public system, possibly enrolling temporarily in the private system.[26]

5.2 Average Direct Effects on Achievement

We present results for standardized math test scores, expressed in standard deviations, in the main section of the paper, with results for Spanish and English, which are highly similar, left to and appendix for brevity (see Appendix Figure A1). Figure 4 displays point estimates and 95% confidence intervals of the effects of being displaced by a school closure on math scores, using our matched event study model and estimated using the interaction-weighted estimator. The figure illustrates a lack of diverging trends across the treated and matched control students in the pre-treatment period, and no anticipatory effect prior to the closure. Indeed, a test of pre-treatment coefficients fails to reject that they are jointly zero, with a p-value of 0.71. The post-closure estimates are small and statistically insignificant, indicating null average treatment effect on the displaced students. This result aligns with those in Brummet (2014), which similarly uncover small or insignificant average achievement effects from consolidating schools with declining enrollments, as well as with studies on the closure of low-achieving schools (see e.g., Engberg et al. (2012)).

5.3 Spillover Effects on Non-Displaced

We now turn to the students who weren't themselves displaced, but experienced changes to their school environment as a result of local closures. While the median student had no peers or teachers who experienced a closure in the past three years, those at schools receiving displaced students varied widely in their exposure to these events. We investigate the effects of spillovers from consolidations on both attrition and achievement using equation (3). Figure 5a presents estimates of the effect of receiving displaced peers on the likelihood that students exit the sample, and on the likelihood that they change school within the sample in Figure 5b.[27] We find that exposure to displaced students causes a small but statistically significant increases in the rate of sample exit, with a 10 percentage point increase in displaced peers being associated with a 0.2 to 0.6 percentage

[26] We do not believe this to be driven by administrative error during relocation. The raw data allow us to observe every instance in which a student's record was updated, and we assign the last known place of enrollment in the school year as their attended school.

[27] For estimating the effects on sample exit, we replace student fixed effects with school fixed effects, as students are likely to only leave the sample once whereas the rate of attrition likely varies across schools.

point increase in the likelihood that a student leaves the sample. Contrary to this, we find that an increase in displaced peers is also associated with a lower rate of school switching within the system, by about 1 to 2 percentage points. Though this could be due to fewer local options as a result of consolidations, it may also signal a preference for the school that has been chosen to remain open. Whether students benefit academically could shed light on these decisions.

Estimates of the spillover effects on achievement are presented in Figure 6. The figure plots coefficients and confidence intervals corresponding to the change in test scores, expressed in standard deviations, that result from a 10 percentage point increase in the share of peers who were displaced this year and one, two, or three years ago. It shows small negative effects of displaced students on their non-displaced peers, which are statistically insignificant with the exception of one year post-closure. Analogous estimates of spillovers from displaced teachers, which are estimated in the same regression, are presented in **??** and are small and statistically insignificant. Thus, we find closures have no statistically discernible average effects on achievement, either directly or indirectly.

5.4 Heterogeneity and Mechanisms

Despite uncovering no substantial average effect on achievement from closing schools, we cannot be certain that closures didn't create winners and losers, so to speak. Consolidations result in a wide range of changes in education inputs for displaced students, as shown in Table 1. Displaced students may end up in schools with higher or lower achieving peers, more or less effective teachers, shifting class sizes, and varying distances travelled between schools. Moreover, concentrating displaced students into the same receiving schools with one another could provide beneficial continuity in peer networks, but may also disrupt opportunities for new and advantageous peer connections. We examine the role of each of these factors in mediating the average treatment effects below.

5.4.1 Student Attributes

We investigate whether average effects vary across students types by interacting the event study model, including match cohort-by-year fixed effects, λ_{mt}, with indicators of student attributes. These attributes include gender, special education status, above-median baseline achievement (as measured by grade 3 test scores), and whether the student resided in urban or rural areas.[28] These results, provided in Appendix Figure A4, illustrate consistent null effects of closure across several

[28]The data include a measure of whether the school was located in an urban or rural area. For displaced students, we use the urban/rural status of the closed school, as this is of direct interest. For non-displaced students, we use the modal urban/rural status of schools they attended in the sample period.

student types, with the exception of those displaced from urban schools, who experience some achievement gains one to two years after closure, which dissipate in the third year. It is possible that this effect for urban-based students is driven by a larger selection of high quality school options in denser areas, aspects of the closure which we will explore further in the following sections.

5.4.2 Relative School and Teacher Quality

We employ our measures of the relative school performance and teacher value added of closed and nearby schools to investigate how changes in peer and teacher performance mediate consolidation effects. The average achievement of students in a school is a rough measure of educational output comprised of idiosyncratic student factors along with teacher and school inputs. In comparing the baseline achievement of closed schools to those in the near vicinity, we capture the degree to which the typical displaced student experienced a change in these combined factors. We begin by analyzing this aggregate measure of relative school quality, before isolating the role of teacher inputs alone.

The estimated heterogeneous treatment effects are displayed in Figure 7. In the left panel, students who were displaced from underperforming schools experience large improvements in math test scores following consolidation, of about 0.40 standard deviation after three years. Those from schools with test scores more closely resembling their neighbours (middle panel) improved more modestly, by about 0.10 standard deviation after three years and 0.15 standard deviation after four years. Those in overperforming closed schools saw slight declines in achievement of around 0.10 standard deviation the first two years of displacement, which are less precisely estimated over time. These findings are again consistent with prior research, though the magnitude of achievement gains are unseen in the existing literature.[29] They suggest that baseline standardized test outcomes measured at the school level capture elements of the learning environment that are highly influential to incoming students.

We now turn to opening this black box of education quality by investigating a key input to learning: teachers. We begin by using our estimates of teacher value added (TVA) to define a predetermined measure of baseline relative teacher quality, Δ_i^T, summarized in Table 1. As above, this type of measure can inform which schools are more or less beneficial to close. We use the baseline relative teacher performance to categorize a closed school's teachers as underperforming,

[29] Brummet, for example, estimates math score improvements of 0.09 standard deviation after 2 years when closed schools underperformed neighbors by 0.5 standard deviation. Using their approach, we estimate a 0.20 standard deviation increase at this level, more than double the magnitude.

with baseline average TVA 0.06 lower than neighboring schools, overperforming, with average TVA 0.06 higher than neighboring schools, and average performing for intermediate values. Though Δ_i^T is calculated using the average teacher's contributions to improving test scores, it is highly correlated with Δ_i^S which is calculated from only raw, and not residualized scores, with a correlation coefficient of 0.61 in the base period. Thus, it is not surprising that the results, presented in Appendix Figure A5, strongly mirror those using that rougher measure of educational quality. The results in the left panel show that students displaced from schools which had lower performing teachers than their neighboring schools improved test scores by about 0.35 standard deviation as a result of consolidation. This result is robust to controlling for relative school quality, Δ_i^S, demonstrating that teacher effectiveness is both an important input to achievement and mediator in closure effects.

While this analysis tells us how one might optimally select a school for closure based on pre-determined characteristics of its teachers and neighboring schools, we have yet to uncover the actual reallocation of displaced students to new teachers. Examining teacher value added as an outcome in itself can illuminate whether consolidations improved access to effective teachers, and how this translates into academic achievement. We relate changes in this outcome to achievement effects by replicating **??** using teacher value added as the outcome. Teacher value added is estimated in the baseline years of 2010 to 2013, for those teaching the relevant subject in grades 4 through 8.[30] From 2014 to 2018, we define the teacher quality experienced by a student as that which was estimated at baseline by their currently assigned math teacher.[31]

Heterogeneous treatment effects on math TVA by baseline relative school quality are shown in Figure 8. Displacement from an under- or average-performing school resulted in students being assigned to modestly higher quality teachers in the first two years following displacement, followed by larger improvements for those from under-performing schools. Notably, a strong pattern of declining teacher quality emerges for those students who were displaced from overperforming schools, with an initial effect of -0.10 growing to -0.30 after three years. This is over a one standard deviation decline in TVA, or a drop of 35 percentiles in the teacher quality distribution. Recall that achievement also suffers for these students as a result of closure. It appears as though the high quality teachers from the overperforming closed schools did not follow their students to new schools, whether by choice or assignment, likely contributing to the approximately 0.10 standard

[30] The first grade for which we estimate value added is 4 since we require one year of lagged scores in the control vector.

[31] Where a student's assigned teacher does not have an estimate of TVA, either because they were not in the system in 2010 to 2013 or were not teaching the required grades, we use the average TVA in the school in the relevant subject.

deviation decline in math test scores for those students.

5.4.3 Distance between Schools and Peer Continuity

A frequently cited downside of consolidating underutilized schools is that it is highly disruptive for students, who are forced to adjust to new settings, peers, and teachers. These disturbances could be ameliorated somewhat if students are displaced to nearby schools or most of their peers are reassigned to a common school, creating some continuity in social networks. On the other hand, greater dispersion of students across receiving schools could reflect more schooling options, allowing students to select a more beneficial assignment. We examine each of these related mediating influences in turn. As described in Section 3, we measure distance of displacement using the median distance to new schools among students displaced by a given closure, and the continuity of peers by the share of students displaced by a given closure who go to the modal receiving school. These measures are designed to abstract from the individual student's experience, which may be driven by selection, by aggregating to the typical student experience arising from a given closure. For estimating the heterogeneous effects using either variable, we estimate the regression model allowing for linear effects of relative school quality, that is $\Delta_i^S \sum_k \tau_k^{\Delta} D_{it}^k$.

We group displaced distances into under a kilometer, from one to 2.75 kilometers, and further than 2.75 kilometers. The results in Appendix Figure A6 show that achievement gains are more likely to be realized when students are reallocated to schools in close geographic proximity, holding fixed the average relative quality of schools in a 5 kilometer radius. It remains possible that this result is driven by larger average effects for students who access higher quality schools, however distance of displacement is clearly independently important to lessening student disruptions.

Turning to peer continuity, we again discretize the concentration of displaced students into three groups: lower than 35 percent attending the modal receiving school, between 35 and 55 percent attending the modal receiving school, and over 55 percent attending the modal receiving school. Appendix Figure A7 presents the results, again controlling for the performance gap between closed and nearby schools at baseline. Consistent with prior literature, we uncover some beneficial effects associated with concentrating displaced students into the same receiving schools, though students also improve when they are more dispersed. We conjecture that this conflicting result may be driven by a correlation between student dispersion across many receiving schools and a wealth of school options from which students may choose. Indeed, controlling for the number of schools in a 5 kilometer radius drives the effects in the left panel to zero, while maintaining large estimated

improvements for students who concentrate in common receiving schools.

Altogether, these patterns are less distinct and display more variance relative to comparisons across baseline school and teacher performance, suggesting distance and peer continuity are of secondary importance to driving consolidation effects.

5.4.4 Class Size

Considering that school closures in Puerto Rico are motivated by declining enrollment and underutilized facilities, displaced students likely experience a downward trend in class size preceding closure followed by an increase as they join the fuller classes of their receiving schools. Some work examining the influence of class size as an education input has found that lower class sizes improve achievement (Lazear, 2001; Krueger & Whitmore, 2001), sometimes nonlinearly (Connolly & Haeck, 2022; Hojo, 2013; Urquiola, 2006), though there are studies that report null effects (Hoxby, 2000; Leuven et al., 2008). We analyze the role of changing class sizes in driving the above-documented heterogeneous treatment effects. In particular, we once again estimate the heterogeneous effects across schools with varying baseline achievement, this time for the outcome class size.

The resulting event study estimates are presented in Figure 9. The right and middle panels depict the expected trend: class sizes gradually shrink in the four years leading up to consolidation, at which point they jump back up to their earliest levels or slightly higher. The pattern is quite muted for students displaced from schools that underperformed relative to their neighbors, who saw a small and short-lived increase in class size. Thus, not only were these students fortunate to attend schools with more effective teachers, they were also spared from the shock of increasing class sizes typically experienced by the displaced. Yet, it is worth noting that despite the magnitudes of class size increases and teacher value added decreases, achievement losses suffered by students displaced from overperforming schools are smaller and less persistent when compared to those from underperforming closed schools, where these inputs responded to a lesser degree in the opposite direction. This may suggest that peers are really the driving force behind test score impacts of closure, as students appear insulated from substantial declines in teacher quality or increases in class size. Whether displaced students had a similar effect on their counterparts in receiving schools is the subject of the remaining analysis.

5.4.5 Heterogeneity in Spillover Effects

As with displaced students, we investigate the potential for heterogeneous treatment effects in both student and closure attributes. We estimate no statistically significant spillover effects on students across gender, special education status, baseline achievement, or urban versus rural residency. We define closure conditions for non-displaced students by taking the average of closure characteristics among their displaced peers, and scaling this by the share of displaced peers. We do not uncover any heterogeneous treatment effects by relative school quality, relative teacher quality, distance between closed and new schools, or peer continuity.[32]

Lastly, we estimate spillover effects on two education inputs of interest: teacher quality and class size. As schools receive an influx of students from closed schools, they may also hire new teachers – potentially from the closed schools themselves – to accommodate the increased enrollment, or simply increase class sizes. Depending on the resulting configuration of new teachers and students, class sizes could also decrease. We estimate the effect of a 10 percent increase in displaced peers on math teacher value added and class sizes in Figure A8a and Figure A8b, respectively. Once again, we estimate no persistent statistically significant effects on either outcome, indicating that consolidations did not alter the learning environment enough to meaningfully impact those who were not displaced, though a very small portion of them are induced to leave the system.

5.5 Teacher Relocation

While we have shown that students displaced from overperforming schools experienced diminished access to high-quality teachers, and vice-versa for underperforming closures, we have yet to explore how teachers are reallocated to produce these outcomes. A particular concern when consolidating a school is that teachers may be unable to secure a new position, or may exit the public system to private schools or retirement. This would impair student learning if the system loses high performing teachers, or exits cause disruptions in community networks for children. We explore teacher exit through a simple regression framework that compares teachers working in schools just before they closed to those who never experienced a closure, controlling for year, age, age squared, and region. We use the teacher panel for this analysis, and expand the panel one year past the exit year to flag the year in which they no longer appear in the sample.

Figure 10 displays coefficient and standard error estimates for the effect of displacement on

[32] Results available upon request.

permanent exit from the PRDE (within the time frame of 2010 to 2018) and termporary exit. We find that school closure was associated with an initial 5 percentage point increase in PRDE exit for teachers, though teachers who did not exit in the first year post-closure were 3 percentage points more likely to stay on than those in schools that remained open. A smaller increase in temporary exit of 2 percentage points is estimated in the right panel, suggesting some teachers were simply not reassigned right away, or chose to return for other reasons. We examine whether high value-added teachers are more prone to remain in the system in Appendix Figure A9. While we see not systematic relationship with regards to exit and measured teacher quality, we see that the increased exit is driven entirely by teachers for whom we are unable to measure value added. These are those who did not teacher the core subjects in grades 4 through 8, and are likely more junior. Work is ongoing to further investigate the role of teacher reallocation.

6 Conclusion

This paper has investigated the effects of school consolidation on public school enrollment and student achievement. Examining a large number of enrollment-based consolidations in Puerto Rico's public school system, we leverage the staggered timing of closures and a panel of student data to recover causal estimates from an event study model. Unlike in previous studies, we show that schools are not closed on the basis of low achievement, and test score outcomes across the treated and control students follow parallel trends leading up to closure. We conduct a careful study of enrollment effects, both to understand treatment effects on system exit as well as to be conscious of differential attrition, which motivates us to employ a matching procedure that balances student enrollment histories.

We find that school consolidations lead to improved academic scores when the schools that close underperform (at baseline) the neighboring schools that accept most displaced students. The reverse is true of overperforming schools that close, and in Puerto Rico the distribution of consolidation conditions yields null effects on average. These findings support similar results taken from closures in higher-income and higher-performing settings, though the magnitude of effects are around twice as large than in the most comparable study (Brummet, 2014). Contrary to those papers, we uncover no evidence that displaced students have negative spillover effects on non-displaced peers, though they do drive a small number to exit the public system. This holds regardless of the achievement levels of displaced peers, suggesting that while displaced students seem to be affected by their new

peers, the reverse is not true. We shed further light on this by estimating how other school inputs – teacher value added and class size – adjust in the wake of consolidations, showing that changes in these factors likely play a role in contributing to the academic results, independent of peer quality.

This research advances a new understanding of the consequences of school consolidations by applying rigorous empirical techniques and analyzing a rich array of closures conditions and outcomes in a novel setting. The findings have important implications for education policy. As new consolidations are inevitably contemplated in the future, careful consideration should be given to what alternatives are nearby when selecting a school for closure. We have shown that this is even more true in low performing settings, where the relative benefits of optimally reallocating students following a consolidation are large. In light of our findings of coinciding gains in test scores and teacher quality, further work is needed to better understand whether teacher reallocations can be made more strategically, perhaps to counteract suboptimal school alternatives. Without data on the state of physical infrastructure and the cost-savings realized from ceasing enrollment at a school, counterfactual policy and cost-benefit analyses are beyond the scope of this paper. Finally, we exclude an analysis of school closures in PRDE in 2018 which dramatically shrunk the size of the school system. This rare occurrence warrants independent study to understand whether large-scale consolidations of this magnitude have general equilibrium effects that differ from the partial estimated in this article.

References

Abdulkadiroglu, A., Angrist, J., & Pathak, P. (2011). The elite illusion: Achievement effects at boston and new york exam schools. *Econometrica*, *82*. https://doi.org/10.3982/ECTA10266

Abel, J. R., & Deitz, R. (2014). The causes and consequences of puerto rico's declining population. *Current Issues in Economics and Finance*, *20*(4).

Angrist, J. D., Hull, P. D., Pathak, P. A., & Walters, C. R. (2017). Leveraging Lotteries for School Value-Added: Testing and Estimation. *The Quarterly Journal of Economics*, *132*(2), 871–919.

Beuchert, L., Humlum, M. K., Nielsen, H. S., & Smith, N. (2018). The short-term effects of school consolidation on student achievement: Evidence of disruption? *Economics of Education Review*, *65*, 31–47.

Bifulco, R., & Schwegman, D. J. (2020). Who benefits from accountability-driven school closure? evidence from new york city. *Journal of Policy Analysis and Management*, *39*(1), 96–130.

Blackwell, M., Iacus, S., King, G., & Porro, G. (2009). Cem: Coarsened exact matching in stata. *The Stata Journal*, *9*(4), 524–546.

Boston Consulting Group. (2014). Hallazgos y recomendaciones al departamento de educación de puerto rico.

Brummet, Q. (2014). The effect of school closings on student achievement. *Journal of Public Economics*, *119*, 108–124.

Caraballo-Cueto, J. (2023). Unintended consequences of the largest (austerity-driven) school consolidation on academic achievement. *CENTRO Journal*, *35*(3), 75–92.

Carlson, D., & Lavertu, S. (2016). Charter school closure and student achievement: Evidence from ohio. *Journal of Urban Economics*, *95*, 31–48.

Chan, J. Y., Lemanski, N., Perkins, R., Kastberg, D., Roey, S., & Ruess, R. (2014). *Performance of puerto rico 15-year-old students in mathematics, science, and reading literacy in an international context* (tech. rep.). Organization of Economic Development.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014a). Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American Economic Review*, *104*(9), 2593–2632.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014b). Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. *American Economic Review*, *104*(9), 2633–79. https://doi.org/10.1257/aer.104.9.2633

Comisión de Derechos Civiles de Puerto Rico. (2018). Informe sobre el cierre de escuelas, departamento de educación de puerto rico, complaint number 2018-4-17371.

Connolly, M., & Haeck, C. (2022). Nonlinear class size effects on cognitive and noncognitive development of young children. *Journal of Labor Economics*, *40*(S1), S341–S382.

Cueto, J. C. (2020). *Aprovechamiento académico y el cierre de escuelas en puerto rico* (Study No. 5). University of Puerto Rico.

De la Torre, M., & Gwynne, J. (2009). *When schools close: Effects on displaced students in chicago public schools. research report.* ERIC.

Departamento de Educación. (2006). Carta circular no. 22-2006-2007: Normas y procedimientos para el cierre, consolidación y/o reorganización de escuelas.

Departamento de Educación. (2007). Carta circular no. 12-2007-2008: Política pública sobre la organización de las escuelas elementales y secundarias del departamento de educación.

Departamento de Educación. (2009). Carta circular no. 4-2009-2010: Normas y procedimientos para el cierre, consolidación y/o reorganización de escuelas.

Departamento de Educación. (2013). Carta circular no. 2-2013-2014: Normas y procedimientos para el cierre, consolidación y/o reorganización de escuelas.

Departamento de Educación. (2014). Carta circular no. 25-2014-2015: Normas y procedimientos para el cierre, consolidación y/o reorganización de escuelas.

Departamento de Educación. (2015). Carta circular no. 21-2015-2016: Normas y procedimientos para el cierre, consolidación y/o reorganización de escuelas.

Engberg, J., Gill, B., Zamarro, G., & Zimmer, R. (2012). Closing schools in a shrinking district: Do student outcomes depend on which schools are closed? *Journal of Urban Economics*, *71*(2), 189–203.

Epple, D., Romano, R. E., & Urquiola, M. (2017). School vouchers: A survey of the economics literature. *Journal of Economic Literature*, *55*(2), 441–92. https://doi.org/10.1257/jel.20150679

Hannum, E., Liu, X., & Wang, F. (2021). Estimating the effects of educational system consolidation: The case of china's rural school closure initiative. *Economic Development and Cultural Change*, *70*(1), 485–528.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Disruption versus Tiebout improvement: the costs and benefits of switching schools. *Journal of Public Economics*, *88*(9-10), 1721–1746. https://ideas.repec.org/a/eee/pubeco/v88y2004i9-10p1721-1746.html

Hinojosa, J., Meléndez, E., & Severino Pietri, K. (2019). Population decline and school closure in puerto rico. *New York: Center for Puerto Rican Studies*.

Hojo, M. (2013). Class-size effects in japanese schools: A spline regression approach. *Economics Letters*, *120*(3), 583–587.

Hoxby, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. *The Quarterly Journal of Economics*, *115*(4), 1239–1285.

Hubler, S. (2022). With plunging enrollment, a 'seismic hit' to public schools. *The New York Times*. https://www.nytimes.com/2022/05/17/us/public-schools-falling-enrollment.html

Iacus, S. M., King, G., & Porro, G. (2008). Matching for causal inference without balance checking. *Available at SSRN 1152391*.

Imberman, S. A., Kugler, A. D., & Sacerdote, B. I. (2012). Katrina's children: Evidence on the structure of peer effects from hurricane evacuees. *American Economic Review*, *102*(5), 2048–82. https://doi.org/10.1257/aer.102.5.2048

Kane, T. J., & Staiger, D. O. (2008). *Estimating teacher impacts on student achievement: An experimental evaluation* (tech. rep.). National Bureau of Economic Research.

Krueger, A. B., & Whitmore, D. M. (2001). The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from project star. *The Economic Journal*, *111*(468), 1–28.

Ladd, H. F., & Rivera-Batiz, F. (2006). Education and economic development. *The Economy of Puerto Rico: Restoring Growth*, 189–238.

Lafarga Previdi, I., & Vélez Vega, C. M. (2020). Health disparities research framework adaptation to reflect puerto rico's socio-cultural context. *International Journal of Environmental Research and Public Health*, *17*(22), 8544.

Lazear, E. P. (2001). Educational production. *The Quarterly Journal of Economics*, *116*(3), 777–803.

Leuven, E., Oosterbeek, H., & Rønning, M. (2008). Quasi-experimental estimates of the effect of class size on achievement in norway. *The Scandinavian Journal of Economics*, *110*(4), 663–693.

Özek, U. (2021). Examining the educational spillover effects of severe natural disasters: The case of hurricane maria. *Journal of Human Resources*. https://doi.org/10.3368/jhr.58.4.0520-10893R2

Pop-Eleches, C., & Urquiola, M. (2013). Going to a better school: Effects and behavioral responses. *American Economic Review*, *103*(4), 1289–1324. https://doi.org/10.1257/aer.103.4.1289

Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2022). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. https://doi.org/10.48550/ARXIV.2201.01194

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? In E. Hanushek, S. Machin, & L. Woessmann (Eds.). Elsevier.

Steinberg, M. P., & MacDonald, J. M. (2019). The effects of closing urban schools on students' academic and behavioral outcomes: Evidence from philadelphia. *Economics of Education Review*, *69*, 25–60.

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199.

Sunderman, G. L., & Payne, A. (2009). Does closing schools cause educational harm? a review of the research. information brief. *Mid-Atlantic Equity Center*.

Taghizadeh, J. L. (2020a). Are students in receiving schools hurt by the closing of low-performing schools? effects of school closures on receiving schools in sweden 2000–2016. *Economics of Education Review*, *78*, 102020.

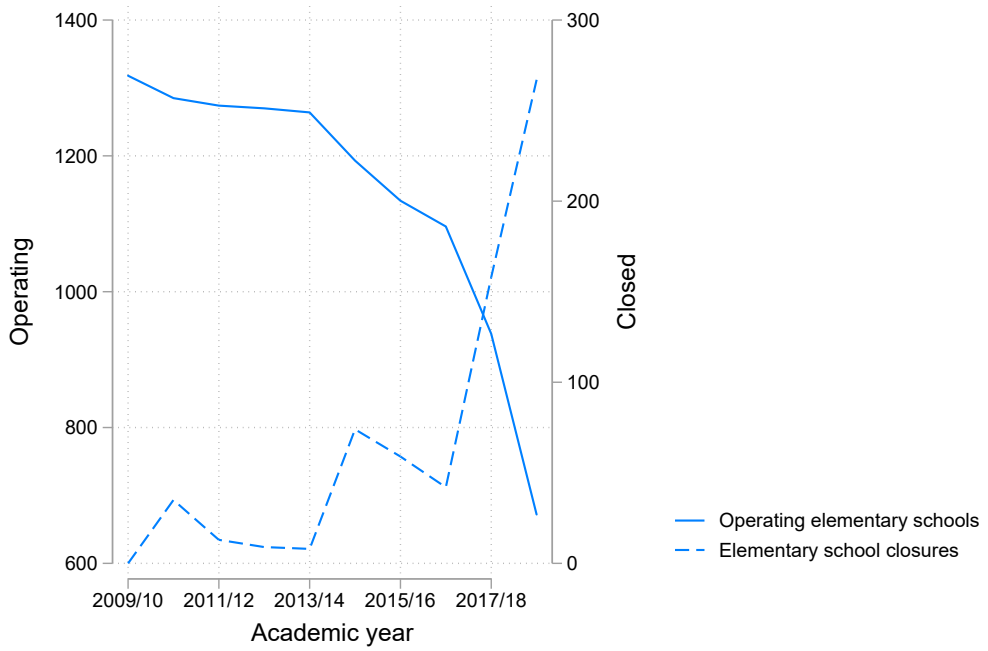
Taghizadeh, J. L. (2020b). Effects of school closures on displaced students and future cohorts. *Labour Economics*, *67*, 101910.

The World Bank. (2022). *Gdp per capita, ppp (current international $)* (tech. rep.). The World Bank. Washington, DC. Retrieved June 15, 2022, from https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD

Urquiola, M. (2006). Identifying class size effects in developing countries: Evidence from rural bolivia. *Review of Economics and Statistics*, *88*(1), 171–177.

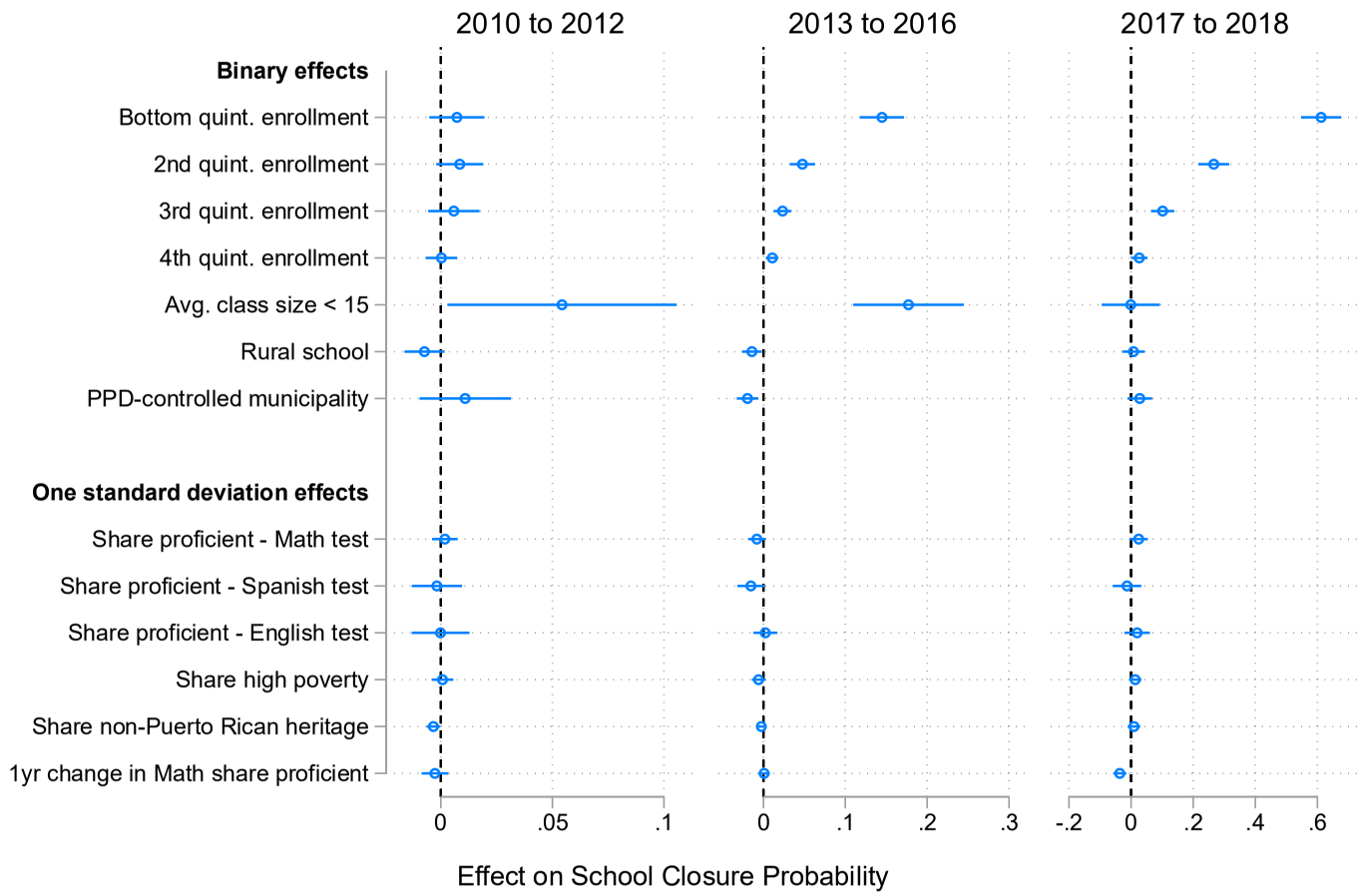
Figures

Figure 1: Number of PRDE Primary Schools Closed and Operating



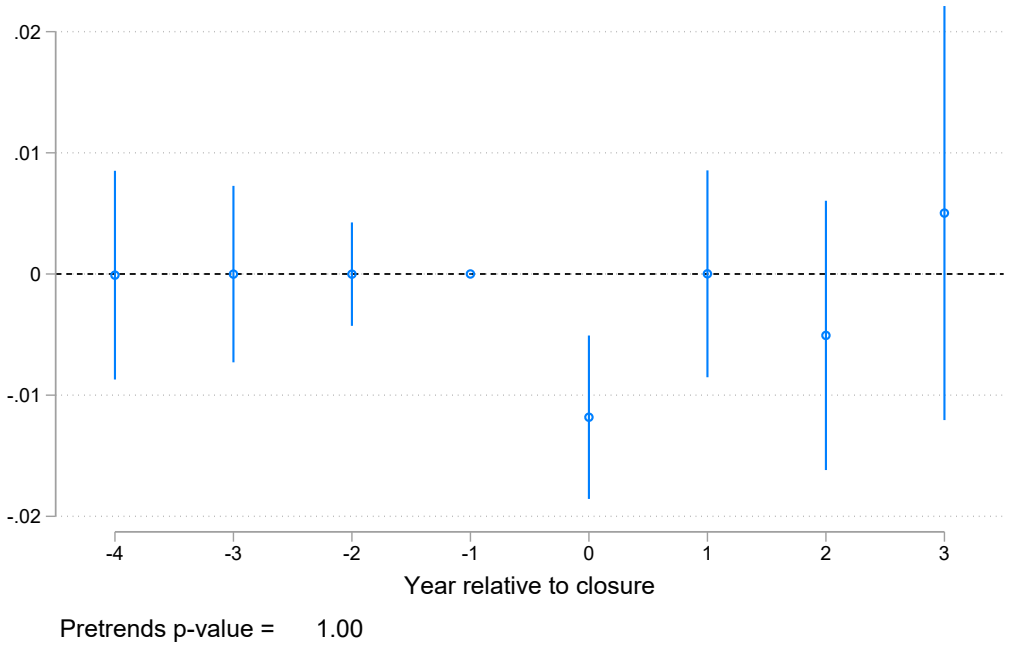
Notes: Closures are counted in the first academic year for which a school stopped enrolling students.

Figure 2: Predictors of School Closures, by Term



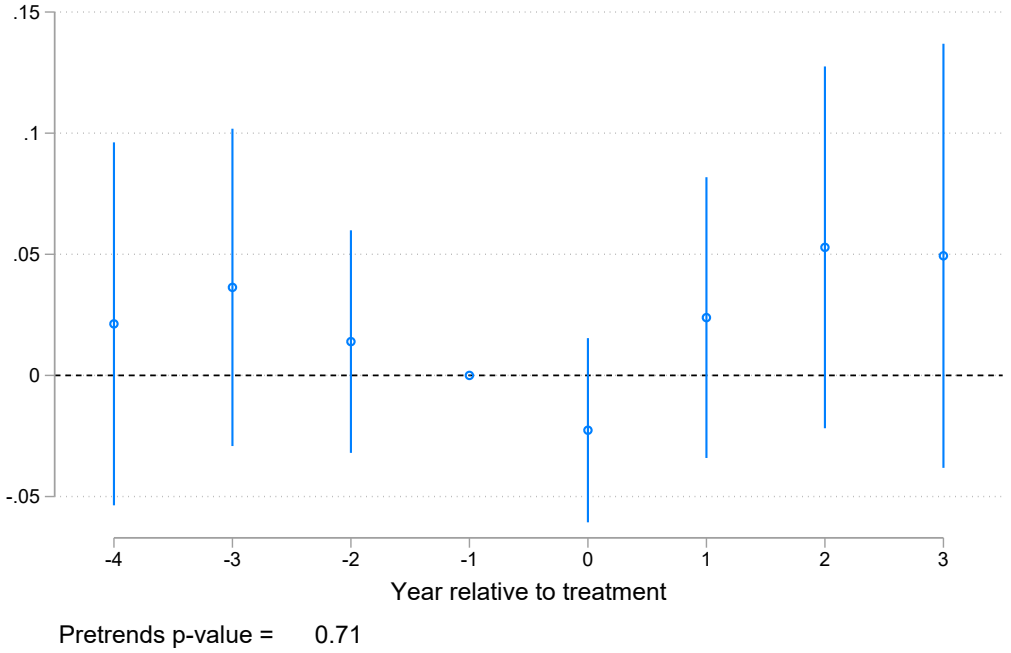
Notes: Coefficients and 95% confidence intervals are derived from a linear probability model regressing an indicator variable for whether a school closed (in a given electoral term) on characteristics of the school, as well as year and region fixed effects. The sample includes one observation per school-year, with closed schools only observed pre-closure. All regressors are interacted with indicator variables for the political term, with the exception of year fixed effects. Coefficients on continuous variables are scaled to be in units of standard deviations.

Figure 3: Effects of Displacement on Student Enrollment in the Public System



Notes: Coefficients and 95% confidence intervals are estimated using equation (2), where the outcome captures continued enrollment. For each student, we create a full panel from kindergarten (or 2010 if it comes first) to grade 8 (or 2018 if it comes first) and record whether they were actually observed in a given year in the PRDE system. Changes are measured relative to the likelihood of being observed one year prior to closure ($t = -1$). The treated cohort (those ever exposed to a school closure) are matched to a control cohort (students never exposed to a closure and not attending a receiving school) following a coarsened exact matching procedure on 3-year enrollment history, grade, and region. The sample includes all students in grades K–8. Standard errors are clustered at the municipal-district level.

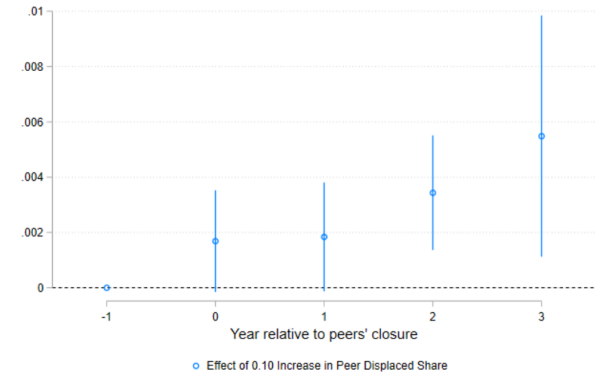
Figure 4: Average Effects of School Closures on Students' Math Test Scores



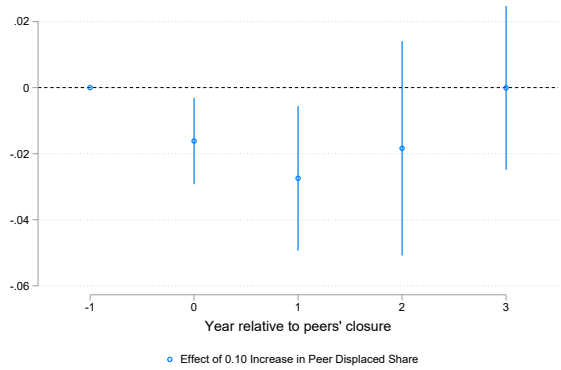
Notes: Coefficients and 95% confidence intervals are estimated using equation (1), with math achievement measured in student-level standard deviation units. The treated cohort (those ever exposed to a school closure) are matched to a control cohort (students never exposed to a closure and not attending a receiving school) following a coarsened exact matching procedure on 3-year enrollment history, grade, and region. Treatment is defined at $t = -1$, the final year in which a closed school is observed in operation. Event study estimates are adjusted using Sun and Abraham (2021)'s interaction-weighted estimator, and standard errors are clustered at the municipal-district level. The sample includes all tested students in grades 3 to 8.

Figure 5: Spillover Effects on Attrition and Within-Sample Mobility

(a) Effect on Sample Exit

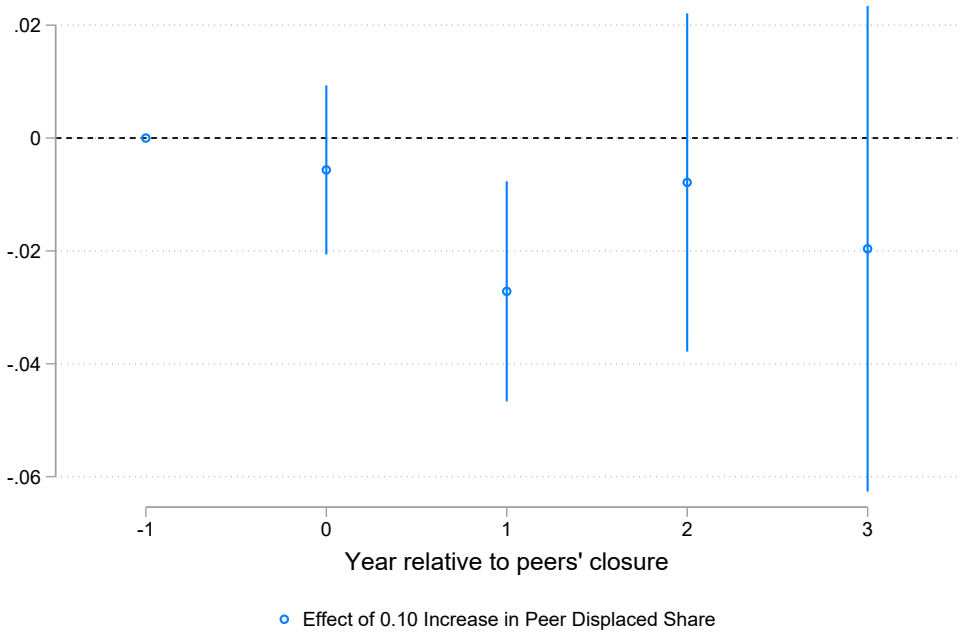


(b) Effect on Exit to Another School



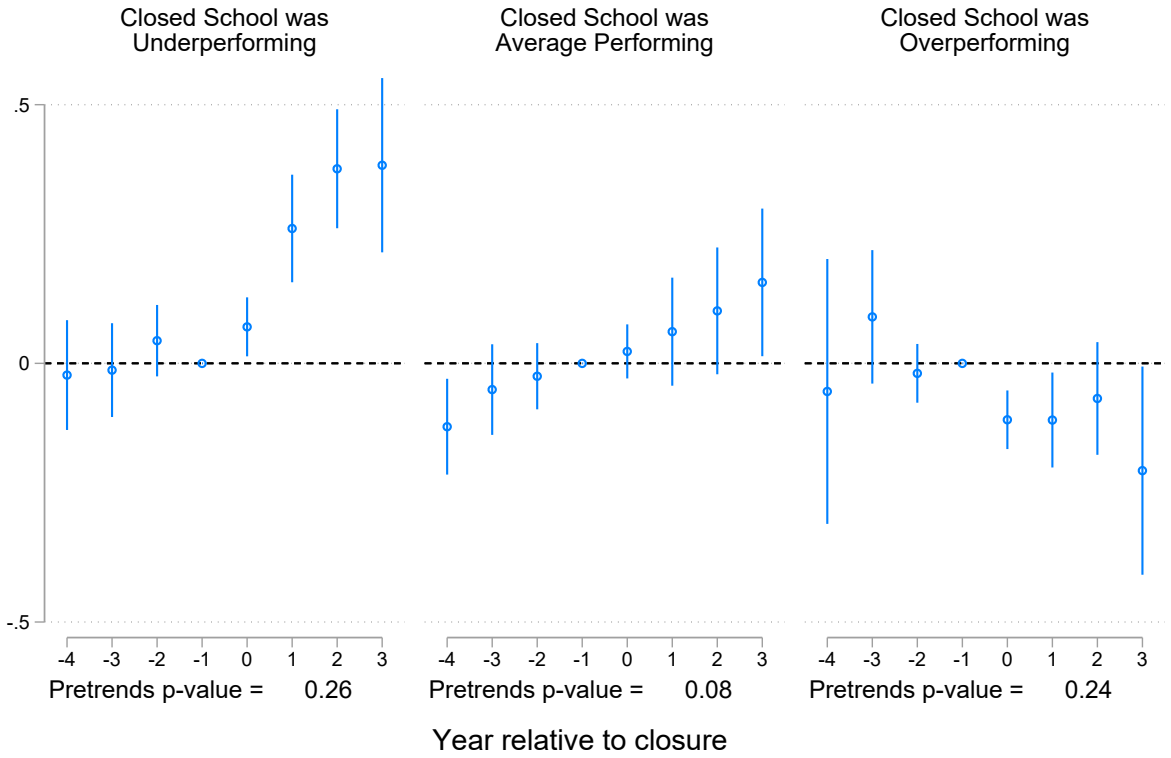
Notes: Coefficients and 95% confidence intervals are estimated using equation (3), where sample exit is a binary variable indicating that a student currently enrolled does not appear in the system the following year, and exit to another school in a binary variable indicating that students changed schools within the system before reaching a terminal grade. Standard errors are clustered at the municipal-district level and the sample includes all tested students in grades 3 to 8 who were not displaced in the sample period.

Figure 6: Spillover Effects on Math Test Scores



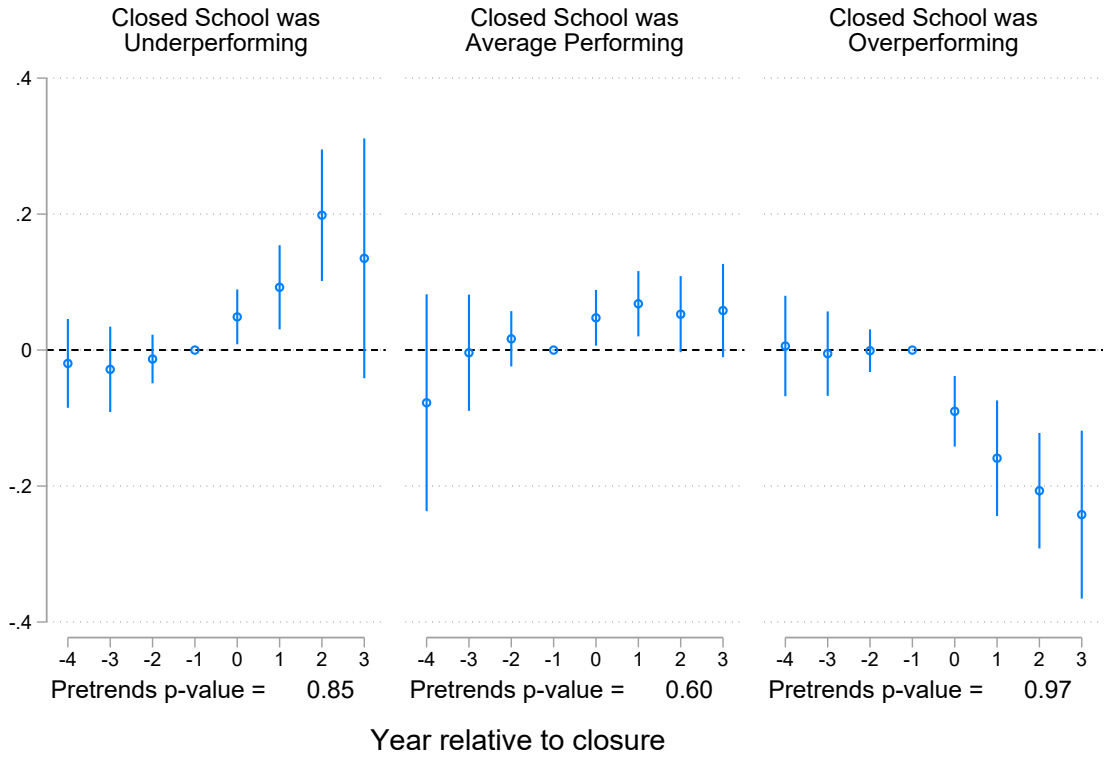
Notes. Coefficients and 95% confidence intervals are estimated using equation (3), with math achievement measured in student-level standard deviation units. Standard errors are clustered at the municipal-district level and the sample includes all tested students in grades 3 to 8 who were not displaced in the sample period.

Figure 7: Effects of Displacement on Math Test Scores, by Baseline Relative Performance of Closed School



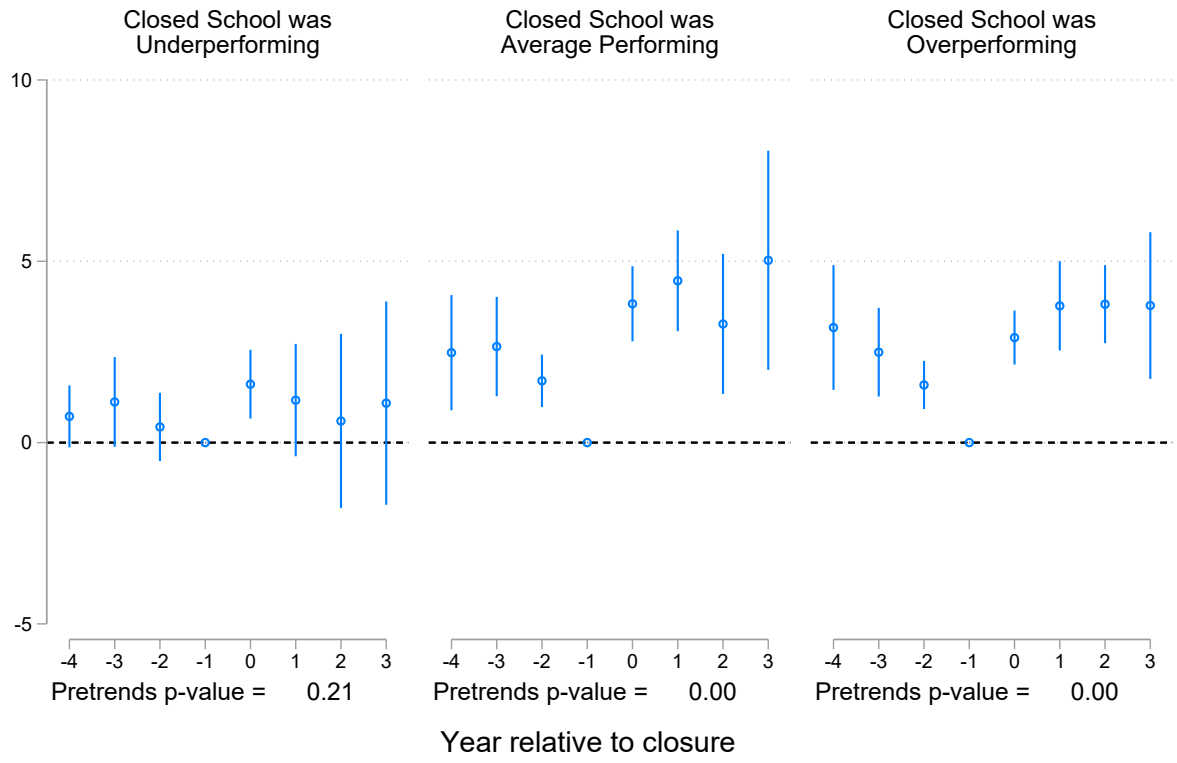
Notes: Coefficients and 95% confidence intervals are estimated using equation (4), with math achievement measured in student-level standard deviation units. Relative school performance at baseline is the difference between average test scores in the closed school in 2010 to 2012 and schools within a 5km radius. The treated cohort (those ever exposed to a school closure) are matched to a control cohort (students never exposed to a closure and not attending a receiving school) following a coarsened exact matching procedure on 3-year enrollment history, grade, and region. Treatment is defined at $t = -1$, the final year in which a closed school is observed in operation. Event study estimates are adjusted using Sun and Abraham (2021)'s interaction-weighted estimator, and standard errors are clustered at the municipal-district level. The sample includes all tested students in grades 3 to 8 and excludes the baseline years 2010 to 2012.

Figure 8: Effects of Displacement on Math Teacher Value Added, by Baseline Relative Performance of Closed School



Notes: Coefficients and 95% confidence intervals are estimated using equation (4), with the outcome teacher value added measured following the methodology developed in Chetty et al. (2014a), using 4 years of test scores from 2010 to 2013. Relative school performance at baseline is the difference between average test scores in the closed school in 2010 to 2012 and schools within a 5km radius. The treated cohort (those ever exposed to a school closure) are matched to a control cohort (students never exposed to a closure and not attending a receiving school) following a coarsened exact matching procedure on 3-year enrollment history, grade, and region. Treatment is defined at $t = -1$, the final year in which a closed school is observed in operation. Event study estimates are adjusted using Sun and Abraham (2021)'s interaction-weighted estimator, and standard errors are clustered at the municipal-district level. The sample includes all tested students in grades 3 to 8 and excludes the baseline years 2010 to 2012.

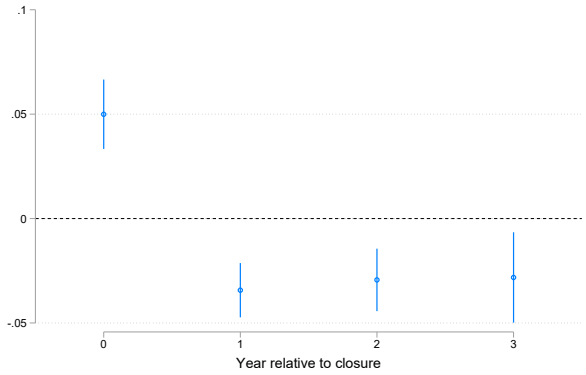
Figure 9: Effects of Displacement on Class Size, Baseline Relative Performance of Closed School



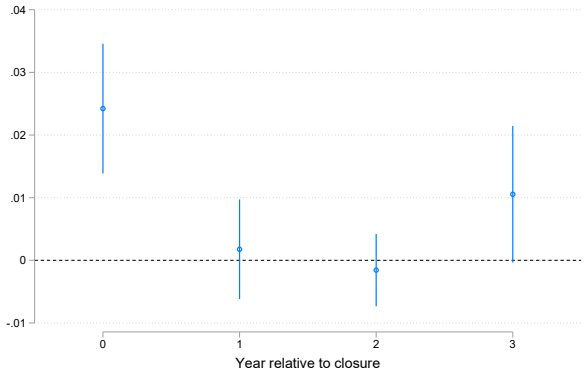
Notes: Coefficients and 95% confidence intervals are estimated using equation (4), with the outcome class size defined as the average class size in the student's school and grade in a given year. Relative school performance at baseline is the difference between average test scores in the closed school in 2010 to 2012 and schools within a 5km radius. The treated cohort (those ever exposed to a school closure) are matched to a control cohort (students never exposed to a closure and not attending a receiving school) following a coarsened exact matching procedure on 3-year enrollment history, grade, and region. Treatment is defined at $t = -1$, the final year in which a closed school is observed in operation. Event study estimates are adjusted using Sun and Abraham (2021)'s interaction-weighted estimator, and standard errors are clustered at the municipal-district level. The sample includes all tested students in grades 3 to 8 and excludes the baseline years 2010 to 2012.

Figure 10: Effects of Displacement on Teacher Exit from Public System

(a) Permanent Sample Exit



(b) Temporary Sample Exit



Notes: Coefficients and 95% confidence intervals are estimated using a teacher panel and regressing an indicator variable for whether the teacher disappeared from the sample in the current year on the year since experiencing a closure. Treatment is defined at $t = -1$, the final year in which a closed school is observed in operation. Since teachers could not, by construction, exit the sample prior to experiencing the closure we include only years of observations after the closure. The regression controls for year, a quadratic in teacher age, and region. We estimate baseline teacher value added following the methodology developed in Chetty et al. (2014a), using 4 years of test scores from 2010 to 2013. Standard errors are clustered at the municipal-district level.

Tables

Table 1: Summary Statistics of Schools, by Political Term

	AY 09/10 to 11/12		AY 12/13 to 15/16		AY 16/17 to 17/18	
	Schools one year before closure	Schools w/ no closure	Schools one year before closure	Schools w/ no closure	Schools one year before closure	Schools w/ no closure
School Enrollments						
End of year enrollment	137.74	386.41	143.18	343.95	155.60	315.59
	(114.48)	(156.69)	(69.90)	(140.75)	(77.63)	(131.45)
Average class size (core courses)	16.38	22.64	18.56	22.62	19.68	22.25
	(5.90)	(3.33)	(5.12)	(3.38)	(4.29)	(3.82)
Student Body Characteristics						
Share w/ ethnicity other than Puerto Rican	0.01	0.02	0.02	0.02	0.03	0.02
	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
Share of students w/ poverty status on PPAA exam	0.85	0.80	0.85	0.81	0.84	0.80
	(0.11)	(0.10)	(0.13)	(0.11)	(0.10)	(0.12)
Share of special education students	0.22	0.20	0.31	0.29	0.34	0.33
	(0.12)	(0.11)	(0.10)	(0.09)	(0.09)	(0.10)
Share of students left the PRDE system	0.02	0.02	0.02	0.02	0.01	0.01
	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
Academic Performance						
Average GPA	2.80	2.73	2.96	2.87	3.05	2.96
	(0.51)	(0.39)	(0.40)	(0.35)	(0.38)	(0.35)
Gr 3-8 Math share proficient	0.40	0.30	0.41	0.34	0.56	0.41
	(0.27)	(0.23)	(0.27)	(0.25)	(0.28)	(0.27)
Gr 3-8 Spanish share proficient	0.44	0.45	0.46	0.48	0.57	0.50
	(0.24)	(0.19)	(0.23)	(0.20)	(0.24)	(0.20)
Gr 3-8 English share proficient	0.42	0.40	0.42	0.43	0.49	0.43
	(0.24)	(0.21)	(0.27)	(0.23)	(0.27)	(0.22)
Political Environment						
PNP in power at municipal level	0.86	0.94	0.61	0.52	0.64	0.69
PPD in power at municipal level	0.14	0.06	0.39	0.48	0.36	0.31
Opposition party in power at municipal level	0.14	0.06	0.61	0.52	0.36	0.31
Observations	57	1,958	183	2,642	425	1,328

Notes: The table reports means and standard deviations of relevant school-level measures, taken the year before closure for closed schools or pooling all school-years for open schools. Consolidation events are grouped by the political term in which the schools were selected for closure. Schools with no closure in a term include those that may close in a later term.

Table 2: Summary Statistics of Students, by Key Samples

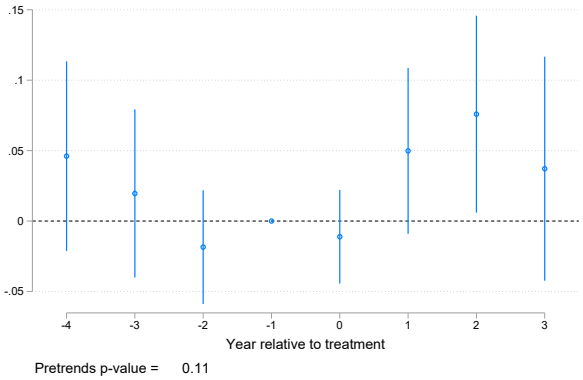
	Overall (1)	Treated (2)	Control (3)	Spillovers (4)
Panel A: Student Characteristics				
Female Student	0.49	0.48	0.49	0.49
Special Education Student	0.23	0.33	0.23	0.22
Rural Student	0.50	0.56	0.50	0.49
Student Age	11.93	11.28	13.06	12.09
	(3.52)	(2.41)	(2.77)	(3.54)
Panel B: Outcomes				
Student Remains Enrolled in PRDE Next Year	0.92	0.87	0.95	0.92
	(0.27)	(0.34)	(0.22)	(0.27)
Normalized Math Test Score	0.00	0.02	0.01	0.00
	(1.00)	(0.63)	(0.96)	(1.00)
Normalized Spanish Test Score	0.00	-0.01	0.01	0.00
	(1.00)	(0.79)	(0.97)	(1.00)
Normalized English Test Score	0.00	-0.00	0.00	0.00
	(1.00)	(0.85)	(0.98)	(1.00)
Teacher's Math Value Added at Baseline	0.02	0.06	0.01	0.01
	(0.30)	(0.30)	(0.30)	(0.30)
Teacher's Spanish Value Added at Baseline	0.01	0.04	0.01	0.01
	(0.24)	(0.25)	(0.24)	(0.24)
Teacher's English Value Added at Baseline	0.01	0.04	0.01	0.01
	(0.27)	(0.28)	(0.27)	(0.27)
Class Size in Core Courses	23.09	20.83	23.30	23.17
	(4.68)	(5.76)	(4.61)	(4.63)
Panel C: Closure Conditions				
Share of peers displaced this year	0.01		0.01	0.01
	(0.04)		(0.03)	(0.04)
Share of teachers displaced this year	0.01		0.01	0.01
	(0.05)		(0.04)	(0.05)
Baseline Relative Performance of Closed School	0.16	-0.01		
	(0.82)	(0.82)		
Baseline Relative TVA of Closed School	0.02	0.02		
	(0.20)	(0.19)		
Share of Displaced Attending Modal Receiving School	0.46	0.47		
	(0.17)	(0.16)		
Distance Between Closed and Receiving Schools (km)	2.26	2.34		
	(2.19)	(2.52)		
Student-Year Observations	2,646,040	161,375	1,741,917	2,421,137
Student Observations	641,227	24,909	337,906	595,242

Notes: The table reports means and standard deviations of relevant student-level measures. The first column summarizes the overall sample. The second column reports values from the year before closure for treated students in the matched sample (i.e. those in tested grades). The third column summarizes the control population that neither experienced a closure nor were exposed to many displaced peers, and who are matched to treated students. The final column includes the panel of all students who were not displaced by a school closure and thus comprise the spillovers sample.

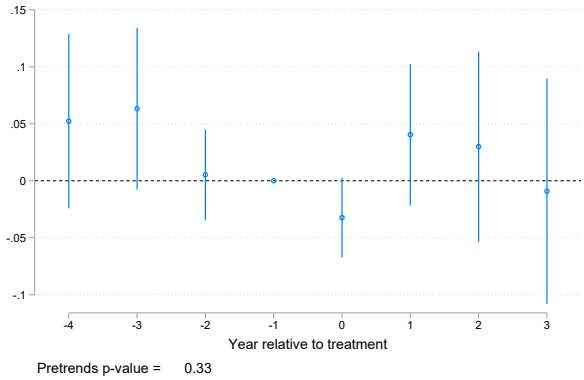
A Appendix Figures

Figure A1: Effects of Displacement on Language Test Scores

(a) Spanish

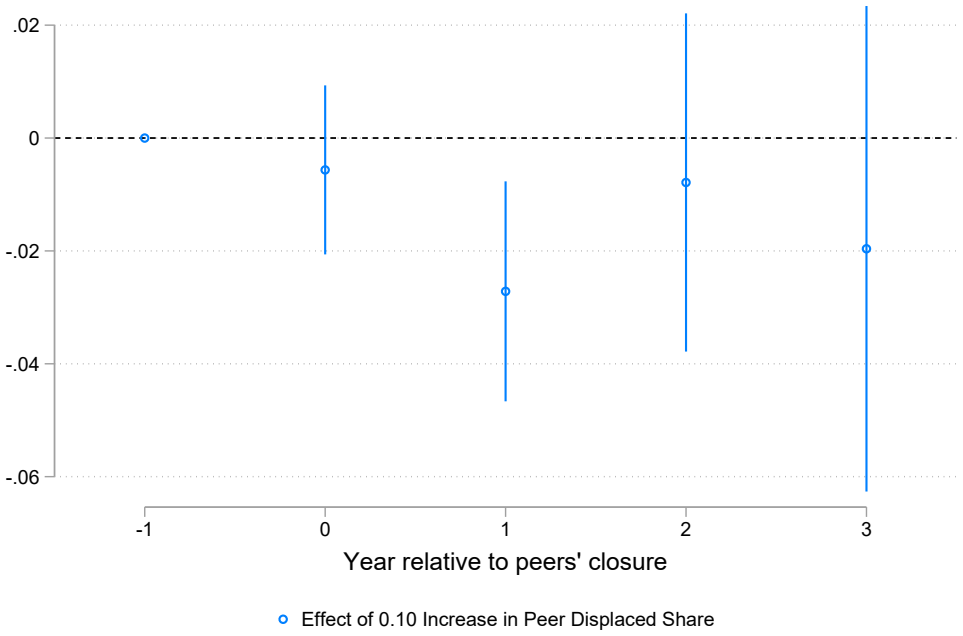


(b) English



Notes: Coefficients and 95% confidence intervals are estimated using equation (1), with achievement measured in student-level standard deviation units. The treated cohort (those ever exposed to a school closure) are matched to a control cohort (students never exposed to a closure and not attending a receiving school) following a coarsened exact matching procedure on 3-year enrollment history, grade, and region. Treatment is defined at $t = -1$, the final year in which a closed school is observed in operation. Event study estimates are adjusted using Sun and Abraham (2021)'s interaction-weighted estimator, and standard errors are clustered at the municipal-district level. The sample includes all tested students in grades 3 to 8.

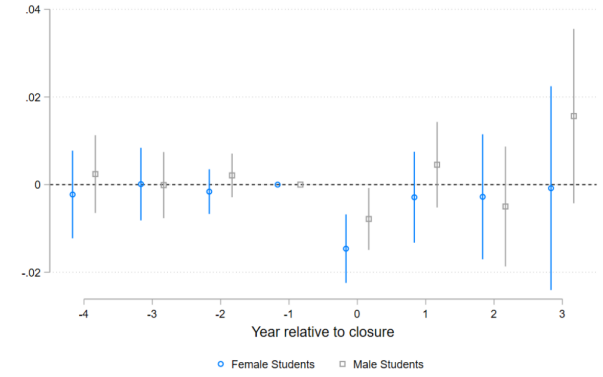
Figure A2: Spillover Effects of Displaced Teachers on Math Test Scores



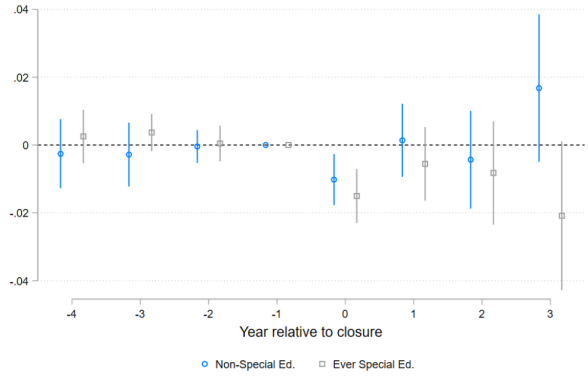
Notes. Coefficients and 95% confidence intervals are estimated using equation (3), with math achievement measured in student-level standard deviation units. Standard errors are clustered at the municipal-district level and the sample includes all tested students in grades 3 to 8 who were not displaced in the sample period.

Figure A3: Effects of Displacement on Student Enrollment in the Public System, by Student Attributes

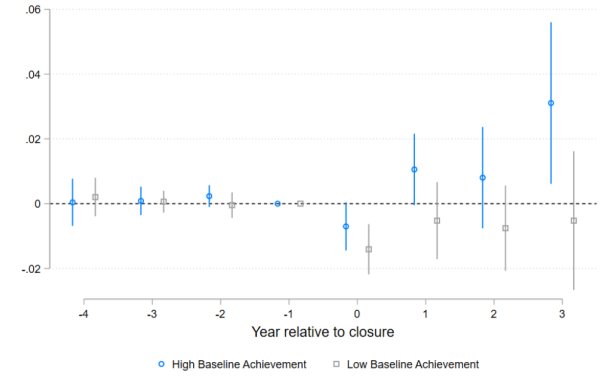
(a) Gender



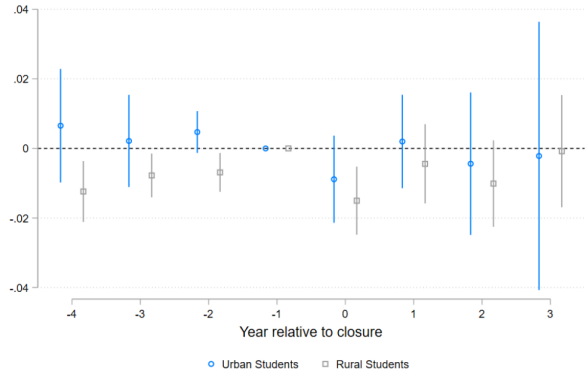
(b) Special Education Status



(c) Baseline (Grade 3) Achievement



(d) Urban or Rural

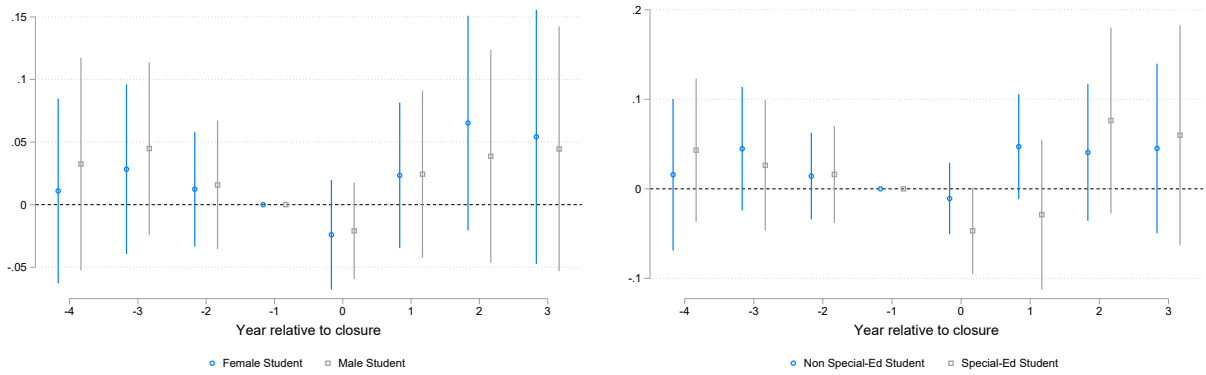


Notes: Coefficients and 95% confidence intervals are estimated using equation (1) interacted with binary variables for the student attribute. Math achievement is measured in student-level standard deviation units. The treated cohort (those ever exposed to a school closure) are matched to a control cohort (students never exposed to a closure and not attending a receiving school) following a coarsened exact matching procedure on 3-year enrollment history, grade, and region. Treatment is defined at $t = -1$, the final year in which a closed school is observed in operation. Event study estimates are adjusted using Sun and Abraham (2021)'s interaction-weighted estimator, and standard errors are clustered at the municipal-district level. The sample includes all tested students in grades 3 to 8.

Figure A4: Effects of Displacement on Math Test Scores, by Student Attributes

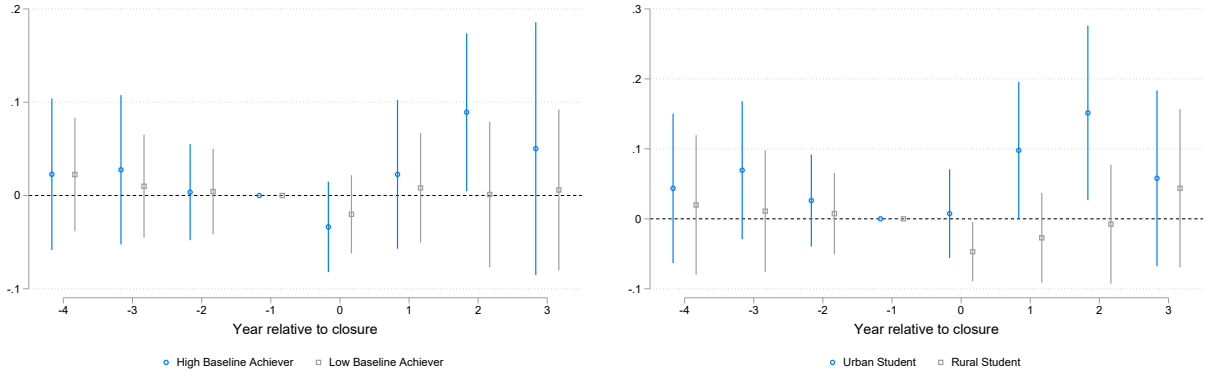
(a) Gender

(b) Special Education Status



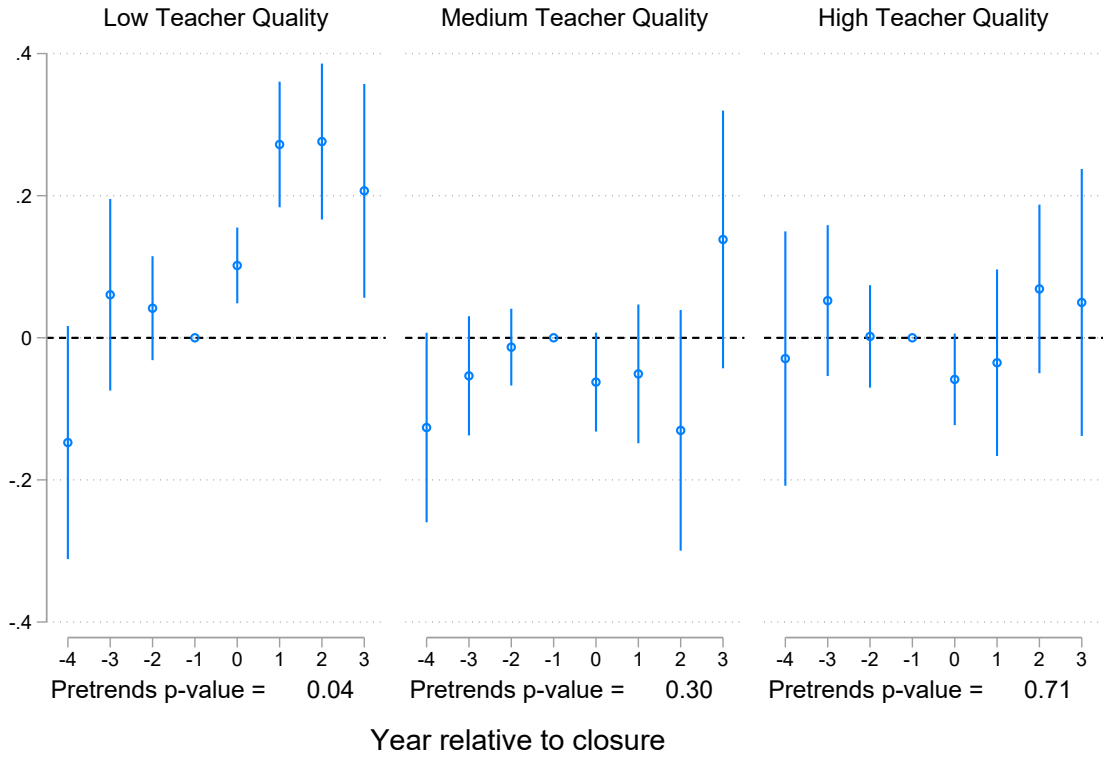
(c) Baseline (Grade 3) Achievement

(d) Urban or Rural



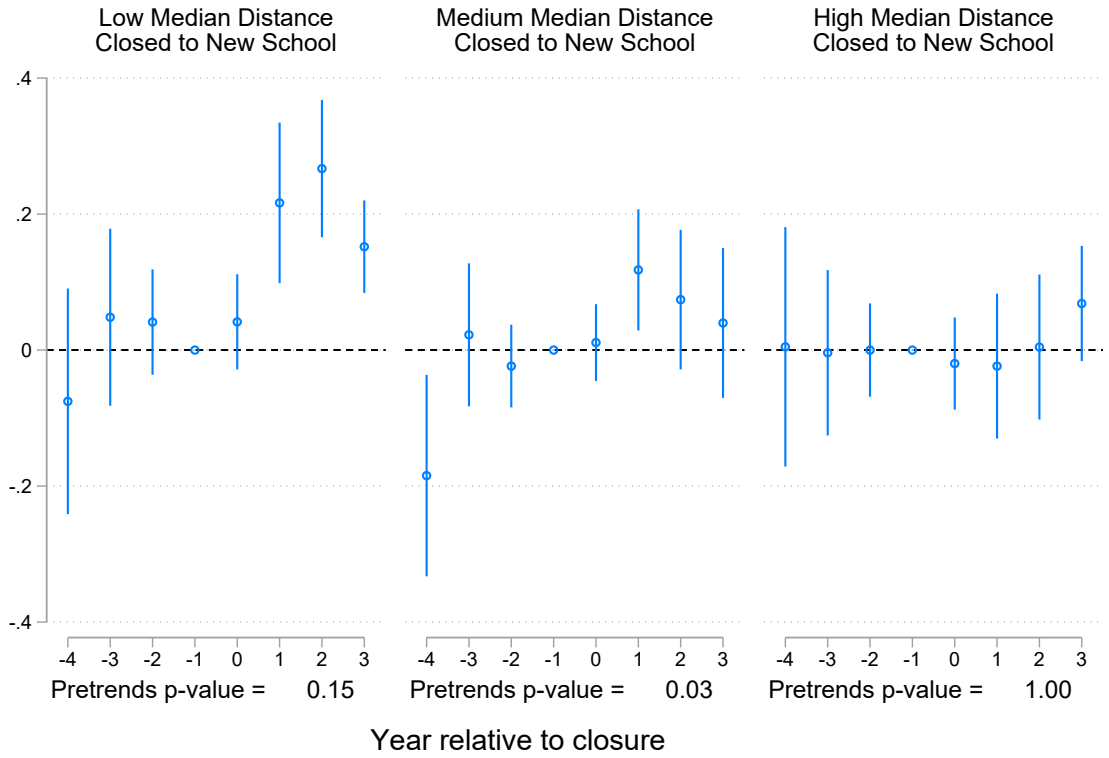
Notes: Coefficients and 95% confidence intervals are estimated using equation (1) interacted with binary variables for the student attribute. Math achievement is measured in student-level standard deviation units. The treated cohort (those ever exposed to a school closure) are matched to a control cohort (students never exposed to a closure and not attending a receiving school) following a coarsened exact matching procedure on 3-year enrollment history, grade, and region. Treatment is defined at $t = -1$, the final year in which a closed school is observed in operation. Event study estimates are adjusted using Sun and Abraham (2021)'s interaction-weighted estimator, and standard errors are clustered at the municipal-district level. The sample includes all tested students in grades 3 to 8.

Figure A5: Effects of Displacement on Math Test Scores, by Baseline Relative Teacher Value Added of Closed School



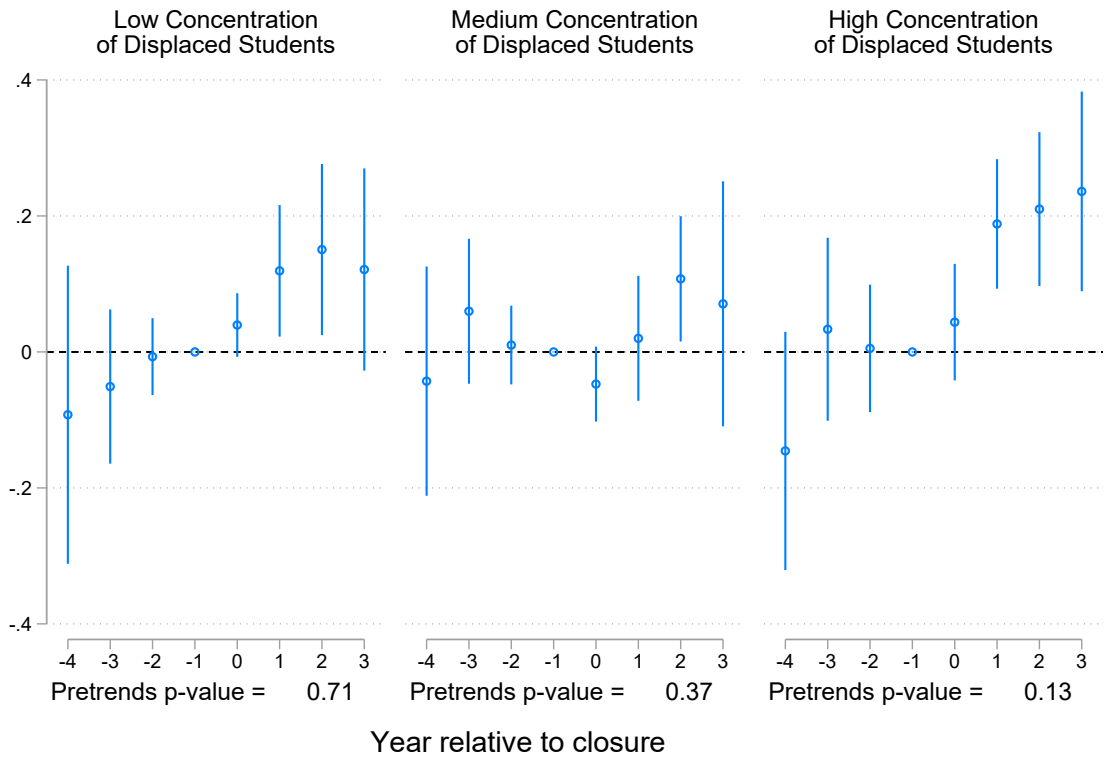
Notes: Coefficients and 95% confidence intervals are estimated using an a analog of equation (4), with math achievement measured in student-level standard deviation units. We estimate baseline teacher value added following the methodology developed in Chetty et al. (2014a), using 4 years of test scores from 2010 to 2013. We take the difference in this average baseline value-added between the closed school and schools in a 5km radius. The treated cohort (those ever exposed to a school closure) are matched to a control cohort (students never exposed to a closure and not attending a receiving school) following a coarsened exact matching procedure on 3-year enrollment history, grade, and region. Treatment is defined at $t = -1$, the final year in which a closed school is observed in operation. Event study estimates are adjusted using Sun and Abraham (2021)'s interaction-weighted estimator, and standard errors are clustered at the municipal-district level. The sample includes all tested students in grades 3 to 8 and excludes the baseline years 2010 to 2012.

Figure A6: Effects of Displacement on Math Test Scores, by Median Distance Between Closed and New School



Notes: Coefficients and 95% confidence intervals are estimated using an analog of equation (4) using discretized median distance and including linear controls are included for baseline relative school quality. Math achievement is measured in student-level standard deviation units. Low median distance schools are those whose median student attended a new school less than 1km away following consolidation, while high median distance includes those schools whose median student travelled at least 2.75km. The treated cohort (those ever exposed to a school closure) are matched to a control cohort (students never exposed to a closure and not attending a receiving school) following a coarsened exact matching procedure on 3-year enrollment history, grade, and region. Treatment is defined at $t = -1$, the final year in which a closed school is observed in operation. Event study estimates are adjusted using Sun and Abraham (2021)'s interaction-weighted estimator, and standard errors are clustered at the municipal-district level. The sample includes all tested students in grades 3 to 8 and excludes the baseline years 2010 to 2012.

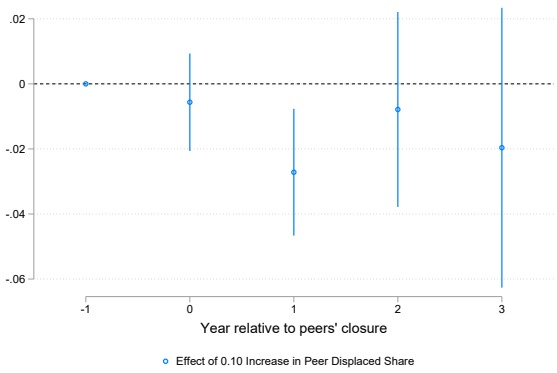
Figure A7: Effects of Displacement on Math Test Scores, by Concentration of Displaced Students



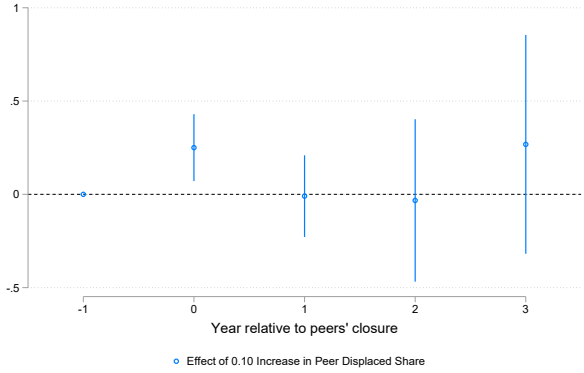
Notes. Coefficients and 95% confidence intervals are estimated using an analog of equation (4) using discretized displacement concentration and including linear controls are included for baseline relative school quality. Math achievement is measured in student-level standard deviation units. The concentration of displaced students is defined as the share of students from the closed school who attend the modal receiving school. The treated cohort (those ever exposed to a school closure) are matched to a control cohort (students never exposed to a closure and not attending a receiving school) following a coarsened exact matching procedure on 3-year enrollment history, grade, and region. Treatment is defined at $t = -1$, the final year in which a closed school is observed in operation. Event study estimates are adjusted using Sun and Abraham (2021)'s interaction-weighted estimator, and standard errors are clustered at the municipal-district level. The sample includes all tested students in grades 3 to 8 and excludes the baseline years 2010 to 2012.

Figure A8: Spillover Effects on Learning Inputs

(a) Effect on Teacher Value Added

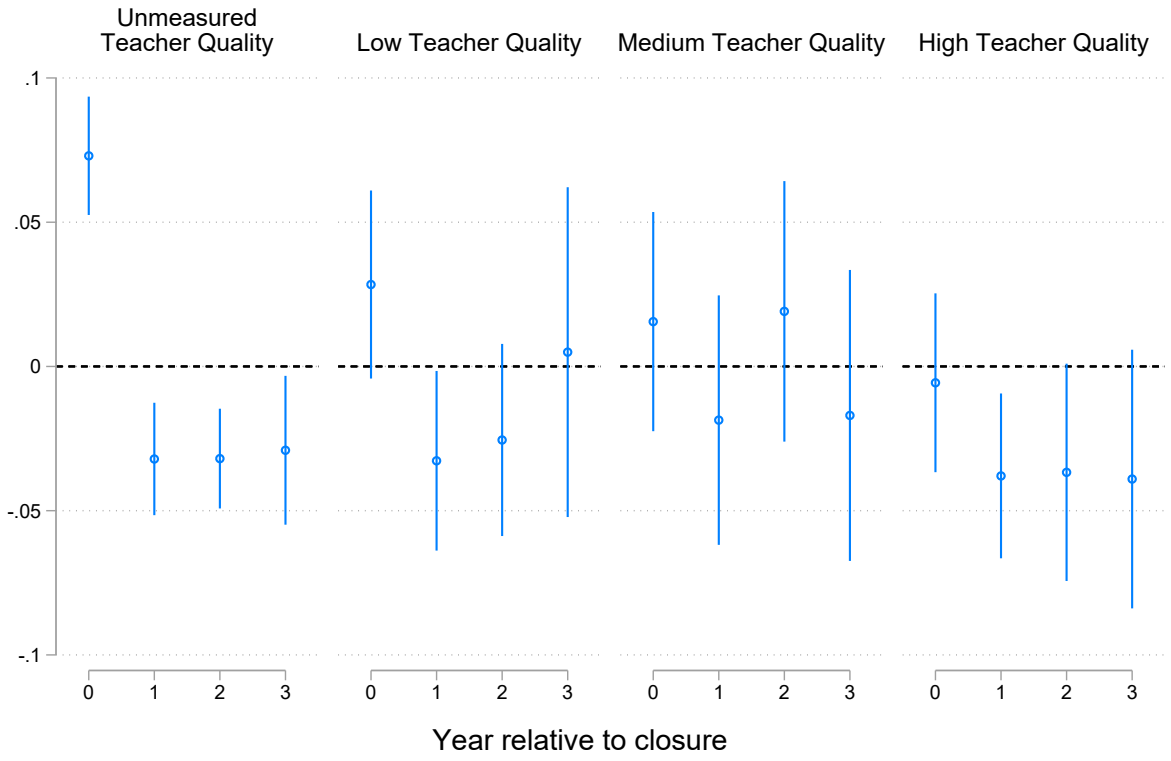


(b) Effect on Class Size



Notes. Coefficients and 95% confidence intervals are estimated using equation (3). In panel (a), the outcome teacher value added is measured following the methodology developed in Chetty et al. (2014a), using 4 years of test scores from 2010 to 2013. In panel (b), the outcome class size is defined as the average class size in the student's school and grade in a given year. Standard errors are clustered at the municipal-district level and the sample includes all tested students in grades 3 to 8 who were not displaced in the sample period.

Figure A9: Effects of Displacement on Teacher Exit from Public System, by Baseline Math TVA



Notes: Coefficients and 95% confidence intervals are estimated using a teacher panel and regressing an indicator variable for whether the teacher disappeared from the sample in the current year on the year since experiencing a closure. Treatment is defined at $t = -1$, the final year in which a closed school is observed in operation. Since teachers could not, by construction, exit the sample prior to experiencing the closure we include only years of observations after the closure. The regression controls for year, a quadratic in teacher age, and region. We estimate baseline teacher value added following the methodology developed in Chetty et al. (2014a), using 4 years of test scores from 2010 to 2013, and divide teachers roughly into terciles with respect to this measure. Standard errors are clustered at the municipal-district level and the sample includes all teachers who taught core subjects in grades 4 to 8 at some point from 2010 to 2013.