

Which Students are Most Impacted by Extended School Closures?

Melinda Morrill*

John Westall^{†‡}

This Version: November 6, 2021

Abstract

In September 2018, Hurricane Florence caused widespread and extended school closures throughout North Carolina. We leverage this geographic variation to assess whether the impacts on learning depends on students' baseline characteristics. There is robust evidence of small declines on reading and mathematics end-of-grade tests. Importantly, those performing at the lowest levels in the prior year did not experience disproportionate losses. Our estimates suggest that most students felt the effects irrespective of baseline human capital or demographic group.

*Corresponding Author: Melinda Morrill, Professor, Department of Economics, North Carolina State University and NBER. Contact: melinda_morrill@ncsu.edu

[†]John Westall, Postdoctoral Research Fellow, Education Policy Innovation Collaborative, Michigan State University. Contact: westall2@msu.edu

[‡]Acknowledgements: The authors would like to thank seminar participants at the AEFPP Annual Conference 2021 and the SIEPR 2020 Working Group, as well as Sarah Crittenden Fuller and Thayer Morrill for helpful feedback and comments. The authors thank Eleanor Warren for excellent research assistance.

1 Introduction

In September 2018, flooding from Hurricane Florence caused widespread and extended school closures in North Carolina. With climate change-related increases in average sea and air temperatures, scientists predict that extreme weather events, such as hurricanes, and weather-related natural disasters, such as forest fires, will become more frequent and more severe (Abatzoglou and Williams, 2016; Walsh et al., 2016).¹ These events often result in schools having to close for an extended period.² Thus, it is important to understand which students are likely to be harmed the most by school closures. We provide new insights on the heterogeneous impacts of extended school closures by studying the aftermath of Hurricane Florence.

School closures may disproportionately impact students who were previously struggling and performing at the lowest levels because it is more difficult for these students to catch up on missed material. It may also be the case that the most disadvantaged students are less resilient to school closures.³ In this paper, we estimate learning losses by prior achievement levels and by measures of economic disadvantage, racial/ethnic group, gender, and limited English proficiency status. Further, middle school students may be more capable of independent learning and could, theoretically, make up missed material outside of classroom hours. Therefore, we also explore whether elementary school students experienced greater impacts than middle school students.

Using geographic variation in school closures in North Carolina following Hurricane Florence, we estimate a regression model with school fixed effects to compare end-of-grade test performance of elementary (4th and 5th grade) and middle school (7th and 8th grade) students using data spanning AY2014-15 through AY2018-19.⁴ We find a small but statistically significant and robust drop, on average, in both reading and mathematics end-of-grade examinations. Importantly, the effect is concentrated near the median of mathematics performance. On the other hand, all but those with the highest performance on prior-year reading exams saw significant reading test score losses. We fail to find evidence that economically disadvantaged or racial minority students experienced larger learning losses. Indeed, nearly every subgroup studied experienced some learning loss due to the school closures.

¹For example, Abatzoglou and Williams (2016) estimate that human-caused climate change nearly doubled the number of forest fires in the western United States between 1984 and 2015.

²For example, in California between 2002 and 2019, wildfires resulted in over 21,000 cumulative missed school days across from 6,500 schools (Lardieri, 2019). In the last week of October 2019 alone, wildfires and planned power outages in California caused the closure of 1,510 schools, which served over 500,000 students (Lambert, 2019).

³While it is also true that students previously scoring highest might struggle to perform at those same levels following school closures, this type of mechanism is unlikely to manifest itself in elementary and middle school.

⁴As described in more detail below, Mathematics EOG scores are not available for all 8th graders.

An important caveat to these findings is that they measure the short run test score declines net of any actions taken by schools, parents, students, teachers, etc. It is outside the scope of this study to explore what strategies can be most effective at mitigating the harm caused by school closures. Moreover, we cannot provide evidence on whether the impacts persist or what remedial actions may be effective. Further, the estimated impact of school closures will capture both the learning losses from missed instruction and any disruption that occurs in the student’s life outside of school. We explore this phenomenon and find that both effects are likely present in our setting. Together, the findings in this study suggest that the impacts of extended school closures are broadly experienced by students in middle and elementary school across the distribution of prior performance levels.

2 Background and Methods

2.1 Prior Literature

Natural disasters lead to disruptions in students’ home lives that are idiosyncratic and extreme, with some students suffering material hardship outside of any impacts from school closures. For example, Hurricane Katrina caused major evacuations and long-term damage in New Orleans in 2005. Sacerdote (2012) tracks the long-term academic performance of evacuees, while Imberman, Kugler, and Sacerdote (2012) considers the impacts of those students on peers in receiving schools. One distinct aspect of Hurricane Katrina is that many students were displaced and forced to change school districts, but students’ outcomes improved in the long term because many moved to higher-quality schools. In related work, Davis, Cannon, and Fuller (2021) interview school districts following Hurricane Harvey (2017) in Texas and Hurricane Matthew (2016) in North Carolina. Their qualitative findings suggest that hurricanes constrain instructional time and increase the social-emotional needs of students.⁵

Our research builds on prior work using weather-related school closings to identify the causal effect of instructional time on student test performance. For example, Marcotte and Hemelt (2008) use snowfall as an instrument to estimate the effect of unexpected school closures on academic performance. Goodman (2014) uses a similar strategy to show the impact of “coordinated” school closures. They contrast school closures with the impact of individual student absences, where the former has little impact, but the latter is harmful to test performance.

⁵In a conference presentation, Fuller and Davis (2020) use North Carolina school data to study the effects of Hurricanes Matthew and Florence on student achievement and behavior. Their preliminary findings are mixed but provide suggestive evidence of test score losses in mathematics.

Our study also contributes to a large literature examining the impact of instructional time on student outcomes through expansion of the school day, the school year, or international comparisons of instructional time (see Huebener, Kuger, and Marcus (2017) and Rivkin and Schiman (2015) for excellent reviews). Huebener, Kuger, and Marcus (2017) exploit a reform in Germany that expanded instructional hours, and find small treatment effects concentrated among the highest performing students. Similarly, using international comparisons, Rivkin and Schiman (2015) find that increased instructional time positively impacted the performance of students in high-quality classroom environments. Related research explores how instructional time impacts student performance by exploiting random variation in testing dates and within-classroom absences (e.g., Aucejo and Romano, 2016; Fitzpatrick, Grissmer, and Hastedt, 2011; Gershenson, Jacknowitz, and Brannegan, 2016; Hansen, 2011; Sims, 2008). Aucejo and Romano (2016) find a small positive effect of instructional days before the testing date, but this impact is considerably smaller than that due to days missed because of student absences. Fitzpatrick, Grissmer, and Hastedt (2011) also find significant learning gains from instructional days with similar impacts across the distribution of student background characteristics. Sims (2008) explores how schools strategically delay testing dates to increase instructional days and finds small positive effects for mathematics but not reading. Most similar to our study, Hansen (2011) finds that test score impacts following weather-related closures are consistent across performance and grade levels. Other research has exploited variation in attendance resulting from teacher’s professional development days (Jacob and Lefgren, 2003). Jacob and Lefgren (2003) find attending school reduces juvenile property crimes but increases juvenile violent crimes. Another strand of literature examines summer learning losses (e.g., Alexander, Entwisle, and Olson, 2001; Fryer and Levitt, 2004). This research suggests that, while students’ test scores grow over the summer, the test-score growth rate is about 0.1 standard deviations per month slower relative to during the school year (Alexander, Entwisle, and Olson, 2001).

While these studies provide policy-relevant information on the importance of instructional time or short-term disruptions, the implications are less clear for understanding the impacts of extended school closures. For example, Pischke (2007) shows that a school year shortened by two-thirds lead to increased grade retention and fewer students attending higher tracks, but no later adverse effects on earnings and employment. Similarly, Meyers and Thomasson (2021) find the polio epidemic of 1916 substantially reduced educational attainment. Interestingly, Meyers and Thomasson (2021) do not find a significant impact of the 1918 flu pandemic, which had a smaller impact on children. This latter finding is consistent with Ager et al. (2020), which finds precise null effects of the 1918 flu pandemic

on attainment and adult outcomes. Our research contributes to this small literature by considering the short-run effects of school closures that are contiguous and sustained, but we cannot consider any longer-term impacts on educational attainment or labor market outcomes within this study design.

2.2 Setting and Closure Data

In the wake of Hurricane Florence in September 2018, the Federal Emergency Management Agency (FEMA) declared 34 (out of 100) counties in North Carolina disaster areas.⁶ We use published statistics from the North Carolina Department of Public Instruction (NCDPI) on the number of days schools were closed following Florence.⁷ Our analysis sample consists of 1,503 schools for which we have Florence closure information.⁸

Figure 1 illustrates the distribution of school closure duration due to Florence, net make-up days. Approximately 24 percent of schools in our sample did not close due to Hurricane Florence, while around two percent closed for fewer than 1 days. The longest school closures, of 26.5 days, affected 0.13 percent of schools. Although the vast majority of school districts closed all their schools for the same period of time, approximately 10 percent of schools are in districts where some schools closed longer than others. Schools closed longer than others in their district tend to have been more severely affected by the storm. For instance, all but four schools closed more than 14 days were closed longer than their districts' modal closure length. Our analysis includes school fixed effects to account for any time invariant school-level characteristics that may be correlated with both the propensity to close for longer and student test score gains, such as older buildings or a teacher workforce drawn from larger geographic distances.

2.3 Days of Instruction vs. Florence Closure Days

North Carolina's school calendar policies are formalized in 2012 Senate Bill 187 and apply as of the 2013-2014 school year.⁹ Testing dates are typically set before the academic year commences, and school calendars typically are constructed assuming a small number of missed school days due to inclement

⁶See: <https://www.fema.gov/disaster/4393>, [accessed July 2020].

⁷Data are provided at: https://files.nc.gov/dpi/spg-report2019_final.xlsx, [accessed June 2020]. These data report the number of school days missed net of makeup days. As described in the Appendix, we gathered supplemental data from newspaper articles and other public data sources on the total number of days school were closed irrespective of makeup days.

⁸The student-level data are described in Section 2.4 and the Data Appendix, where we explain how we construct the sample of 1,503 schools used in the analysis.

⁹See: <https://www.dpi.nc.gov/districts-schools/district-operations/financial-and-business-services/school-calendar-legislation>.

weather. North Carolina requires that EOG examinations are held during the final 10 instructional days of the school year. Schools must hold at least 185 days of instruction with at least 5.5 instructional hours per day or at least 1,025 instructional hours in the school year. Schools must start no earlier than the Monday closest to August 26 and end no later than the Friday closest to June 11, unless a weather-related calendar waiver is approved. Following Hurricane Florence, North Carolina allowed districts in counties under federal disaster declarations to waive 20 days of school in the 2019 academic year. The special rules did not dictate anything about the testing window for EOG examinations. Therefore, a priori, days of closure may not translate directly into lost days of instruction either before the EOG (our outcome measure) or for the academic year.

Our data include student-level “membership days,” which corresponds to the total days of attendance in school as measured on the date of the EOG examinations.¹⁰ For this exercise, we define a school’s days of instruction as the median of the student-level membership days in each grade by year.¹¹

Figure 2 plots the days a school was closed for Hurricane Florence by the median days of instruction in 2018 using a flexible polynomial fit. The relationship is u-shaped and suggests that, if anything, there is a slight positive relationship whereby schools with more days of instruction in the 2017-2018 academic year actually missed more school days following Hurricane Florence in September 2018. Figure 3 illustrates how days of school missed due to Hurricane Florence maps into total days of instruction prior to EOG examinations. Most notably, there are 6 schools with days of instruction below 150, none of which were severely impacted by Hurricane Florence. Because these schools are small in size, removing them does not alter the basic relationship. Days of instruction is negatively related to Florence closure days, but the relationship is slightly non-linear. We explore this further below at the student level.¹²

2.4 Student Data

The student-level data are derived from restricted-access administrative data records from the North Carolina Education Research Data Center (NCERDC). The number of local education authorities (LEAs), schools, and students at each data creation step are illustrated in Appendix Table A1. Roughly speaking, in North Carolina there are about 112K students per grade per year. We assign students to

¹⁰Prior to 2018, these data are reported alongside test score information. For 2018 and 2019, the data are only available in attendance files.

¹¹The student-level data construction is described below in Section 2.4 and the Data Appendix, where we explain how we construct the sample of 1,503 schools.

¹²In results not shown, a univariate regression at the school by grade level of days of instruction in 2019 on days missed due to Florence yields a coefficient of only -0.83, indicating one day closed resulted in only 0.83 fewer days of instruction.

the school attended at time (t-1) to abstract away from any school switching that may have happened in 2019. So, for elementary school we study 4th and 5th grade students, and the middle school sample includes only 7th and 8th graders. Thus, starting with a dataset of 3rd, 4th, 6th, and 7th graders who have a valid demographics and EOG test scores in the years 2016-2018, we have 1,394,825 students in 2,081 schools in 284 LEA's. We restrict our attention to "regular" (i.e., not alternative or special education) schools, eliminating 0.22 percent of students, and remove charter schools, dropping about 7 percent of students. As shown in the second row of Appendix Table A1, this leaves 115 school districts, 1,844 schools, and 1,294,753 students. We further restrict to schools with traditional calendars (i.e., not year-round), remove three school districts that were severely impacted by Hurricane Matthew in 2018, and remove schools with grade span configurations that are incompatible with our identification strategy of assigning students to their t-1 school.¹³ Finally, we only keep students at schools where the median days of membership was between 150-200 days to remove extreme outliers. These restrictions yield 107 school districts, 1,503 schools, and 1,073,442 students. For analysis on time t test score outcomes, roughly 5 percent of students do not have a valid, on-time test score at time t . This yields our main analysis sample of 107 school districts, 1,503 schools, and 1,020,384 students.

The main test scores of interest are end-of-grade (EOG) examinations in mathematics and reading. For elementary school students, we estimate impacts for 4th and 5th graders, controlling for their prior-year performance and linking to their prior-year school. We focus on 7th and 8th grade for middle school students, where we again know the school attended and performance in the prior year. Starting in the 2017-18 school year, students enrolled in algebra in 8th grade were given a different examination (called "MATH1") than the standard 8th-grade mathematics EOG. The students taking algebra are positively selected by prior year performance. Therefore, we only estimate mathematics EOG performance in middle school for 7th graders. The EOG examination scores are normalized to have a mean of zero and a standard deviation of one within each grade in each academic year. We perform the normalization before any sample restrictions.

Table 1 presents summary statistics for the main sample. Approximately half of the students are male, 25 percent are Black, 19 percent are Hispanic, and 9 percent are another, non-White race. About 12 percent of students had a disability, meaning they have a physical or mental impairment that substantially limits one or more major life activities.¹⁴ Approximately 50 percent were economically

¹³The data include schools with the following grade spans: PK-5, PK-6, PK-8, PK-9, PK-12, K-5, K-6 (if not 6th grader), K-8, K12, 1-5, 2-5, 3-5, 6-8, 6-9, 6-10, 6-12, and 6-13. The most common configurations are PK-5 (31%), K-5 (14%), and 6-8 (47%).

¹⁴For a formal definition, see: <https://www.dpi.nc.gov/students-families/parents-corner/students-disabilities>, [accessed

disadvantaged, defined as being eligible for free- and reduced-price lunch. About 10 percent had limited English proficiency, indicated by participating in North Carolina’s English Language Learner program.¹⁵ The prior year mathematics and reading scores are normalized to be mean zero and standard deviation one prior to implementing the sample restrictions, so the means here are slightly above zero.

Panel B of Table 1 presents statistics on the average days of instruction, measured as membership days at the time of the EOG examinations. On average, the mean and standard deviation of instructional days is similar for all three years in our sample. The average number of days missed due to Hurricane Florence is 3. For those students with positive days missed, the average days missed is 4.

2.5 Empirical Methods

Our empirical strategy relies on the exogenous timing and geographic location of school closures following Hurricane Florence in 2018 for identification. Our econometric model includes both school and academic year fixed effects, allowing us to control for time invariant characteristics of schools and any idiosyncratic differences by year that are common to all schools. The model includes controls for prior year test performance, so the outcomes of interest can be interpreted as annual test score gains. All regression equations also include grade fixed effects and controls for student gender, race/ethnicity, disability status, limited English proficiency (LEP), and economically disadvantaged status (EDS). The key identifying assumption in our model is that the extent of closure days due to Hurricane Florence is not correlated with time-varying differences across schools that are also correlated with test scores. We provide evidence supporting this assumption below, although reading test scores exhibit some differential prior trends. A second identifying assumption is that no other contemporaneous shocks affect students differentially according to days missed. In this case, note that the Florence closure days variable will capture all ways in which the Hurricane affected students, including disruption in students’ communities and homes. Thus, the estimated coefficient on days closed should be interpreted as reflecting all impacts of the Hurricane on student learning.

The regression equation takes the following form:

$$(1) \quad Y_{ist} = \alpha + \lambda Y_{is,t-1} + \beta \text{Florence Closure Days}_{st} + X_{is,t-1} \Gamma + \tau_t + \psi_s + \epsilon_{ist}$$

October 2020].

¹⁵See: <https://www.dpi.nc.gov/districts-schools/testing-and-school-accountability/testing-policy-and-operations/testing-students-identified-english-learners>, [accessed October 2020].

Here, $Y_{i,s,t}$ is the standardized test score of student, i , in school, s , in year t . The coefficient of interest, β , measures how much more students' standardized test scores changed, on average, as a function of the Florence closure days, holding constant time invariant characteristics of schools and any idiosyncratic year effects that are common across students. The model includes controls for performance on mathematics and reading EOG examinations in the prior year, $Y_{i,s,t-1}$, so our coefficient estimates are appropriately interpreted as changes from students' baseline performance level.

3 Results

3.1 Average Test Score Losses

As discussed above, when schools close following a natural disaster, we anticipate that student learning will fall due to both a loss of instructional days and any disruption effects. To begin, Table 2 estimates the relationship between the number of days a school was closed following Hurricane Florence and students' mathematics and reading EOG test scores. We estimate versions of equation (1), where the key coefficient of interest, β , measures the effect of an additional closure day on student performance. In column (1), the model includes grade and year fixed effects, as well as controls for student gender, race/ethnicity, disability status, LEP status, and economically disadvantaged status. Panel A presents estimates for mathematics and Panel B for reading. We find that one additional closure day due to Florence is associated with a significant decrease in mathematics scores of 0.0029 standard deviations, while we detect no significant impact on reading scores.

Column (2) adds school fixed effects, and our estimates suggest that much of the school closure effect on mathematics is mediated by time invariant school characteristics; we estimate an additional school closure day reduces mathematics test scores by 0.0019 standard deviations. On the other hand, for reading, the estimated impact is larger in magnitude when we include school fixed effects. The estimates in Table 2, Column (2), Panel B indicate that an additional missed school day leads to 0.0023 standard deviation lower reading test score. These estimates are consistent with Aucejo and Romano (2016), which finds that one additional calendar day increases mathematics by 0.0017 standard deviations and reading by 0.0008 standard deviations. Similarly, Sims (2008) finds an additional instruction day leads to 0.0052 standard deviations higher mathematics scores and a statistically insignificant 0.001 standard deviation higher reading scores.

Figure 3 illustrated that school closures following Hurricane Florence did not translate one-to-one

to reduced instructional days before examinations. The estimated impact of net school closure days will necessarily combine both the human capital losses from lost instructional time (i.e., the “opportunity cost”) and whatever disruption effects students experience. In Column (3), Table 2, we add school-level days of instruction as an additional control variable. Note that days of instruction are highly correlated, but not perfectly collinear, with school closure days. Including days of instruction as a control enables us disentangle the impacts of missed instructional time and other disruptions due to the storm. Panel A shows that the average impact of an additional instruction day is a 0.0009 standard deviation increase in mathematics, while the effect for Florence closure days conditional on instruction time is diminished and no longer statistically significant. The results are similar when the instructional days in the prior year are included. For reading, in Panel B, the average impact of instructional days is similar to that of mathematics, but the main effect of Florence closure days is only slightly smaller and is still statistically significant indicating a 0.0018 standard deviation decrease in reading scores. These results suggest that the impact on students from Hurricane Florence is not entirely explained by missed days of instruction, but, rather, combines the impacts missed instructional time and other hurricane-related disruptions.

In a further attempt to disentangle the school closure and disruption effects of Hurricane Florence on students’ academic performance, we use data on valid registrations for FEMA’s Individuals and Households Program (IHP) following Hurricane Florence (disaster declaration number 4393) to proxy for damage and disruption in students’ home environment.¹⁶ The Individuals and Households Program provides financial and direct services to eligible individuals and households who experienced uninsured or under-insured necessary expenses. The data include total valid registrations and the IHP amount awarded in dollars collapsed to the county by city by zip code level.¹⁷ To link these measures of damage to North Carolina students, we utilized a Census Block Group to ZCTA crosswalk that matches Census Block Group centroids to the ZCTA in which they are located.¹⁸ For this exercise, we restrict the sample to students living a block group that is linkable to a ZCTA with a population size larger than 90.

¹⁶FEMA provides these data at <https://www.fema.gov/openfema-data-page/registration-intake-and-individuals-household-program-ri-ihp-v2>, [accessed January 2021]. The data provided by FEMA are raw and unedited. To get more accurate estimates at the Zip Code Tabulation Area (ZCTA) level, we edited obvious typographical errors and then matched the data by zip code to ZCTA. The data were then collapsed to be unique by ZCTA and merged to 2019 ACS 5-Year ZCTA-level population counts. Using the ZCTA-level population data, FEMA IHP registrations and dollar amounts were normalized per 100 people.

¹⁷More information on IHP can be found at <https://www.fema.gov/assistance/individual/program>, [accessed February 1st, 2021].

¹⁸This crosswalk was constructed using 2018 North Carolina Census Block Group and 2010 United States ZCTA TIGER/Line Shapefile data. The unique block group identifier GEOID was then matched to the North Carolina Education Research Data Center (NCERDC) block group identifier BLOCK2010 to link the normalized FEMA data to our sample.

For comparison, we begin by replicating the main estimates on the sub-sample of students with valid information on hurricane damage in their ZCTA in Table 2, Column (5). The effects are slightly smaller, but are qualitatively similar and still statistically significant. Column (6) includes both the number of valid registrations per 100 people and the total IHP dollars per 100 people. Again, the impacts differ between mathematics and reading. For mathematics, the number of FEMA claims entirely explains the school closure effect. For reading, the point estimate is slightly smaller and no longer statistically significant, but the hurricane damage variables do not significantly predict reading test scores.

Prior work has found that individual student absences have a substantially larger impact on test scores than days of instruction (Aucejo and Romano, 2016). The main mechanism that has been proposed is that it is more difficult for individual students to catch up on missed material, rather than a whole class having missed instructional time. Our results suggest a potential additional mechanism: disruptions in a student’s life, such as illness, a family emergency, or other situation that leads the student to be absent, may diminish a student’s ability to acquire human capital above and beyond the impact of instructional time missed. This may be an interesting direction for future work.

The remainder of the results focus on the preferred specification in Table 2, Column (2), which includes school fixed effects to account for any time invariant school-level characteristics that might be correlated with the extent of school days missed due to Hurricane Florence. These results should be interpreted as the net effect of hurricane exposure, as they incorporate the impact both of lost instructional time and potential disruptions in students’ home environments.

3.2 Specification Checks, Attrition, and Robustness

The linear specification in Table 2 provides a tractable framework for understanding the impact of school closure days due to Hurricane Florence. However, the impact of school closures may be non-linear.¹⁹ Table 3 presents estimates from a flexible parameterization of the model where we include separate indicator variables for each day of closure due to Florence in lieu of a linear parameterization. All regression estimates still include school fixed effects to account for any time invariant characteristics of schools that might be correlated with their school days missed. The models also include grade and year fixed effects, as well as indicators for students’ gender, race/ethnicity, disability status, LEP, and economically disadvantaged status.

¹⁹In results not shown, higher order terms in Florence Closure Days do not suggest any non-linearity.

The first column of Table 3 illustrates how closure days translated into days of instruction at the student level. This parallels Figure 3, which modeled how the school-level median days of instruction was related to Florence Closure Days. The regression estimates presented in Table 3, Column (1) are at the student level, but tell a similar story. Students whose school closed for 0.6 or 1 day did not see a statistically significant reduction in days of instruction. However, for more extended school closures the impact on instructional days is approximately linear. Table 3, Columns (2) and (3) present the EOG test scores for mathematics and reading, respectively. The estimated coefficients on test scores indicate a significant and large *benefit* to closing for a small number of days due to Hurricane Florence. Further, while most of the estimated coefficients on days missed are negative, there are some notable exceptions.

These estimates are net of any remediation engaged in by parents, schools, teachers, or students. Furthermore, they are a function of both lost instructional time and any disruptions at home. Moreover, these analyses only examine one event, thus we anticipate some idiosyncrasies across schools. For example, perhaps the decision to close in September led some schools to be more reluctant to close following inclement weather later in the year. If a missed school day in September is less impactful on test scores than, say, a missed school day in February, then the students exposed to a small number of closure days following Hurricane Florence might benefit relative to those in schools that did not close because of Florence. Future work might explore heterogeneity not just in the extent of school closure but also the timing during the school year. The estimates in Table 3 are consistent with the preferred specification of equation (1), on average one additional Florence Closure Day leads to a small, but statistically significant, drop in mathematics and reading test scores. Before moving to the analysis of heterogeneity, we explore several other specification checks.

A potential concern with studying the impacts of a natural disaster early in the school year is student attrition prior to testing in the Spring semester. Some students might be displaced or simply not return to school, or students might switch schools due to closures. The main analysis sample consists of students who have valid test scores in the prior year, and school closure days are assigned based on the school attended in year $t - 1$. The regression estimates so far constrain the sample to include only students with valid time t test scores in the appropriate grade. To test for differential test taking rates, we model whether students with valid prior-year test scores have an on-time, correct-grade test score in the current year. Table 4, Column (1) tests whether exposure to an additional closure day changed the probability that students' have a valid time t test score. We do not detect a statistically

significant relationship between closure days and valid year t test scores. However, the point estimate is negative suggesting that exposure to an additional closure day may have reduced the probability of test taking. If attriting students were also lower performing, then this type of attrition would cause the estimated impact on test scores to be biased towards zero. Next, Column (2) of Table 4 shows that, among students with valid test scores, school switching is more likely following Hurricane Florence closures. Missing 10 days is associated with a 1.9 percentage point higher probability of switching schools relative to a baseline school switching rate of 11 percent. This suggests that school switching, and the associated disruption, may be one potential mechanism whereby Hurricane Florence negative impacted students.

Equation (1) includes school and year fixed effects that control for all time invariant characteristics of schools that might also impact test scores and all shocks common across schools by year. The key identifying assumptions are common trends (i.e., that the exposed students' test scores would have trended similarly to non-exposed students' test scores absent the hurricane) and no contemporaneous shocks that differentially affected exposed students. As discussed in Section 3.1 and Table 2 above, for the latter assumption, it is important to note that the hurricane may have affected all aspects of students lives beyond just a closed school. In that sense, our estimates should reflect the net impact of experiencing a severe hurricane, as proxied by the number of days the school was closed. Columns (3)-(8) of Table 4 provide an empirical test of common trends. In Columns (3)-(5), we focus on the main analysis sample from 2017-2019. We add to the model a placebo, or "false," measure of days missed in 2017, meaning we interact closure days due to Hurricane Florence in 2019 with an indicator for 2017. We use the year prior to Hurricane Florence, 2018, as the omitted, reference year. In Column (3), the dependent variable is the student-level days of instruction. The "true" closure days lead to -0.87 fewer days of instruction. The estimated coefficient on "false" closure days in 2017 is small and positive, suggesting that, if anything, the schools with more closures in 2019 had 0.07 more days of instruction in 2017 relative to the 2018 school year. Columns (4) and (5) present estimates of the "true" closure days in 2019 and the "false" measure of closure days applied to 2017. The estimates for mathematics do not suggest differential prior trends. However, for reading scores, the estimated coefficient "false" closure days in 2017 is positive and statistically significant, suggesting differential trends in reading scores; reading scores in the exposed schools may have been trending downward in the years before Florence relative schools that did not close.

To further explore robustness, we expand the sample back include 2015 to 2019. Table 4, Columns

(6)-(8) reports estimates for the longer test of common trends. The omitted year is still 2018. In Column (6), we see that days of instruction in year (t) is again about 0.07 days higher when the 2019 school closure days are falsely applied to 2017. But, we do not see any other statistically significant differences for the other placebo years. In Column (7), the estimates still show no statistically significant differential in trends for mathematics scores. However, evidence of differential trends in reading scores persists. In 2017, “false” days missed is still positive and statistically significantly related to reading scores. Furthermore, “false” days missed in 2015 are statistically significant, negative, and of the same magnitude as the estimated impact of actual days missed. Thus, we conclude that reading test scores may have been trending differentially by school closure duration before Hurricane Florence.

3.3 Heterogeneity by Prior Achievement

This study’s central question is, which students are most impacted by school closures due to a natural disaster? To address this question we first examine heterogeneity by student’s prior-year test performance. In Table 5, we interact the main closure days effect with indicators for the quintile of performance on the EOG test in the preceding year. The reference category is the third quintile (i.e., the median quintile). Additionally, instead of single, linear control variables for prior mathematics and reading test scores, the regressions include quintile-specific polynomials in prior year reading and mathematics test scores.

Table 5, Panel A, shows that the effect of an additional school closure day for the students in the third quintile of prior year mathematics performance is about twice as large as the estimated average effect in Table 2, Column (2). Interestingly, we do not detect negative impacts on mathematics scores in either the first or fifth quintile of prior-year performance. This is in direct contrast to our hypotheses that students at the bottom (or top) of the test score distribution may be particularly affected. It should be noted that these bins are coarse, so this exercise cannot rule out that a small subgroup within each category is significantly impacted. However, the results do suggest that the negative impacts of school closures are felt throughout the distribution and seem to be more concentrated near the median of prior performance.

On the other hand, in Panel B of Table 5, we see that the impact of reading is most strongly affected for the bottom quintile of students in terms of prior reading test score performance. Still, there is a large and statistically significant impact for students in all but the top quintile of prior performance. These results fail to support the hypothesis that impacts are concentrated on students who were the

lowest performing in the prior year.

3.4 Heterogeneity by Demographic Characteristics

Next, we consider other student subgroups to determine the characteristics of students most impacted by the school closures in Table 6. Column (1) explores whether effects are larger for elementary or middle school students. Middle school students may have greater ability to make-up missed material independently, but may not have the same level of parent involvement or supervision. Here the reference category is elementary school students (grades 4 and 5). The difference in the impact on performance is not statistically significant for mathematics or reading, but the mathematics estimates suggest that middle school students might experience a larger learning loss following the school closures relative to elementary school students. These results fail to support targeting elementary or middle school students for remediation.

Table 6 presents results by students' gender. Heterogeneity by gender may reflect behavioral differences or differences in learning styles. For example, boys are four times as likely to have an ADHD diagnosis than girls. However, it is not clear how this might translate into differential impacts of school closures. For both mathematics and reading, boys experience a smaller, but still statistically significant, impact of school closures. In results not shown, the sum of the coefficients for males is not statistically significant for mathematics but suggests a -0.0020 (SE 0.0007) standard deviation decrease in reading test scores.

Our data include only a coarse measure of economic disadvantage, as described in Section 2.4. Roughly half of the students in our sample are classified as being from economically disadvantaged families using this measure. Along this dimension, we fail to find statistically significant differences in the impact of school closures on either mathematics or reading tests. The next column of Table 6 considers groupings by limited English proficiency (LEP) status, as defined in Section 2.4. LEP students see significantly *smaller* mathematics and reading test score declines (and no effect overall). Recall that these estimates are net of any remediation that occurred prior to the end of grade testing dates.²⁰

When looking across different racial/ethnic groups, we find that White students felt the smallest impacts in mathematics. Hispanic students' mathematics test scores fell by over twice as much. Simi-

²⁰In results not shown, we find evidence that on-time test taking falls significantly more for students classified as LEP (in the prior year) due to Florence closure days. If the lowest performing students are granted exemptions at a higher rate due to the school closures, then the results will be biased upwards. This might explain why we find differences in this direction by prior year's LEP status.

larly, the group “other non-White students,” which is a mixture of students from various different racial and ethnic backgrounds, experience a decline in mathematics test scores that is over twice as large as that for White students. However, it is difficult to interpret the estimates for such a heterogeneous group. Reading test score impacts do not vary significantly by racial/ethnic group. These results again point to impacts across a broad range of students.

4 Discussion and Conclusion

It is not clear, a priori, which students might be most able to recover from a school closure due to a natural disaster. Students who previously were struggling may lose additional ground. Some families may be more directly impacted by the flooding following a hurricane, while others may have the resources to provide supplemental instruction to their children. Understanding which students might be least able to recover following an event that causes school closures is key to allocating resources effectively. However, it is imperative to note that the test score losses we measure here account for all efforts that teachers, students, parents, and schools made following Hurricane Florence to mitigate the impacts on student learning. We cannot say what the test score losses might have been absent these efforts. Rather, we interpret these estimates as test score losses net of all efforts to recover following the Hurricane.

The evidence presented here suggests that many students feel the impacts of extended school closures. Moreover, we fail to find strong evidence that particular demographic groups or students with higher or lower levels of past performance experience disproportionate learning losses. The magnitudes of these estimates are small. For mathematics, students from the second through fourth quintiles in prior achievement experienced statistically significant learning losses, with no significant impacts for the lowest or highest performers. On the other hand, reading scores dropped for all except the group of students performing at the highest levels in the prior year. These results fail to support targeting amelioration efforts at the lowest performers and suggest broad impacts of school closures. When looking at demographic groups, we fail to find evidence that effects are concentrated by racial/ethnic group or by economic disadvantage.

Extreme weather events, such as wildfires or hurricanes, are becoming more frequent and more intense and will likely lead to an increasing number of extended school closures similar to those studied here. The results of this study should inform how policymakers view the aftermath of extended school closures as they plan for, and grapple with, how to mitigate the consequences for students. This study

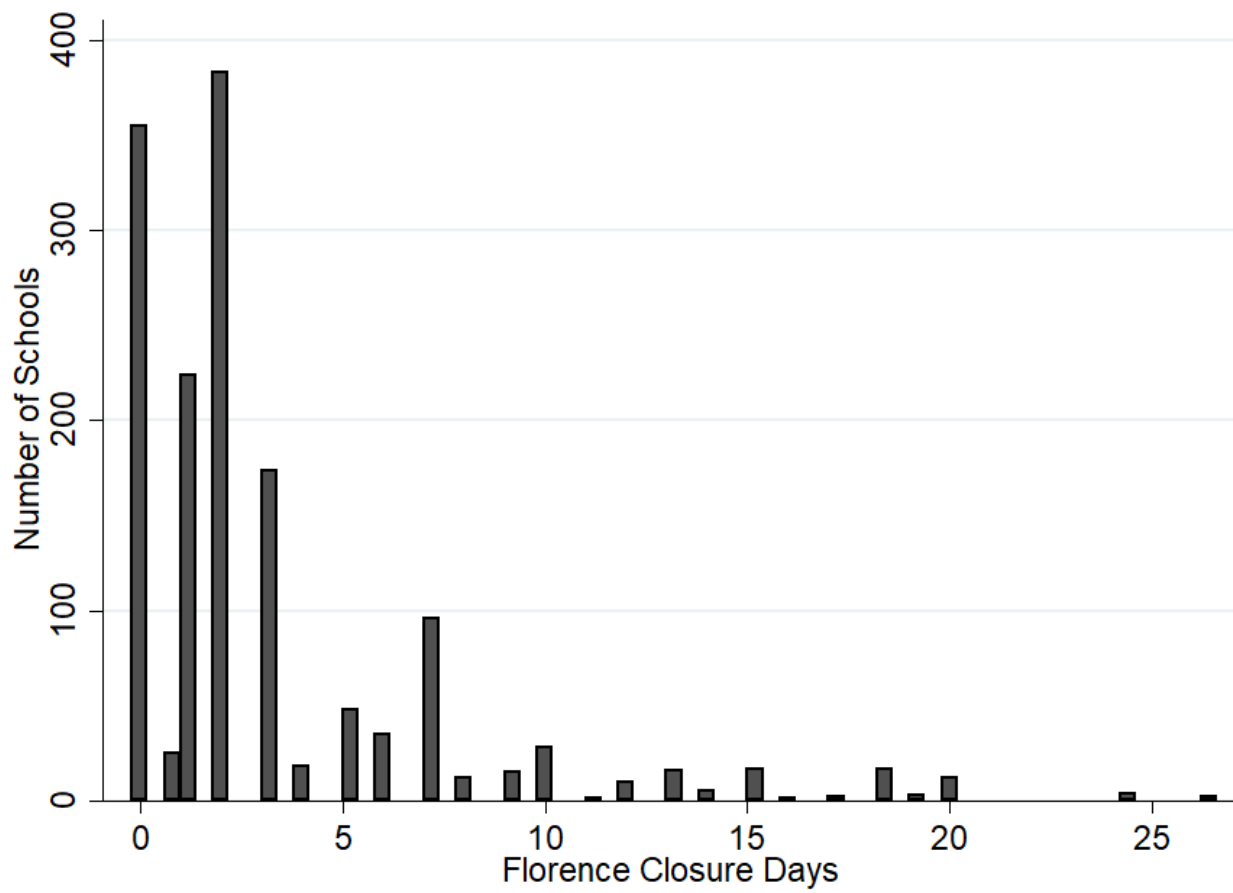
does not address the long-term impacts of school closings, nor the efficacy of distance learning and other strategies to mitigate learning loss. Indeed, the time period under study is before COVID-19 pandemic-related closures made distance learning more accessible for some students. Furthermore, future work should consider whether school closures due to extreme weather events have compounding effects on students, or whether schools, parents, and students can develop strategies that ameliorate the learning losses felt following school closures.

References

- Abatzoglou, John T. and A. Park Williams. 2016. "Impact of Anthropogenic Climate Change on Wildfire across Western US Forests." *Proceedings of the National Academy of Sciences* 113 (42):11770–11775.
- Ager, Philipp, Katherine Eriksson, Ezra Karger, Peter Nencka, and Melissa Thomasson. 2020. "School Closures During the 1918 Flu Pandemic." Tech. Rep. w28246, National Bureau of Economic Research, Cambridge, MA.
- Alexander, Karl L., Doris R. Entwisle, and Linda S. Olson. 2001. "Schools, Achievement, and Inequality: A Seasonal Perspective." *Educational Evaluation and Policy Analysis* 23 (2):171–191.
- Aucejo, Esteban M. and Teresa Foy Romano. 2016. "Assessing the Effect of School Days and Absences on Test Score Performance." *Economics of Education Review* 55:70–87.
- Davis, Cassandra R., Sarah R. Cannon, and Sarah C. Fuller. 2021. "The Storm after the Storm: The Long-Term Lingering Impacts of Hurricanes on Schools." *Disaster Prevention and Management: An International Journal* 30 (3):264–278.
- Fitzpatrick, Maria D., David Grissmer, and Sarah Hastedt. 2011. "What a Difference a Day Makes: Estimating Daily Learning Gains during Kindergarten and First Grade Using a Natural Experiment." *Economics of Education Review* 30 (2):269–279.
- Fryer, Roland G. and Steven D. Levitt. 2004. "Understanding the Black-White Test Score Gap in the First Two Years of School." *The Review of Economics and Statistics* 86 (2):447–464.
- Fuller, Sarah C. and Cassandra R. Davis. 2020. "Investigating the Impact of Hurricane Exposure on Student Achievement and Behavior." In *AEFP Panel Presentation, March 2020*.
- Gershenson, Seth, Alison Jackowitz, and Andrew Brannegan. 2016. "Are Student Absences Worth the Worry in U.S. Primary Schools?" *Education Finance and Policy* 12 (2):137–165.
- Goodman, Joshua. 2014. "Flaking Out: Student Absences and Snow Days as Disruptions of Instructional Time." Working Paper 20221, National Bureau of Economic Research.
- Hansen, Benjamin. 2011. "School Year Length and Student Performance: Quasi-Experimental Evidence." SSRN Scholarly Paper ID 2269846, Social Science Research Network, Rochester, NY.

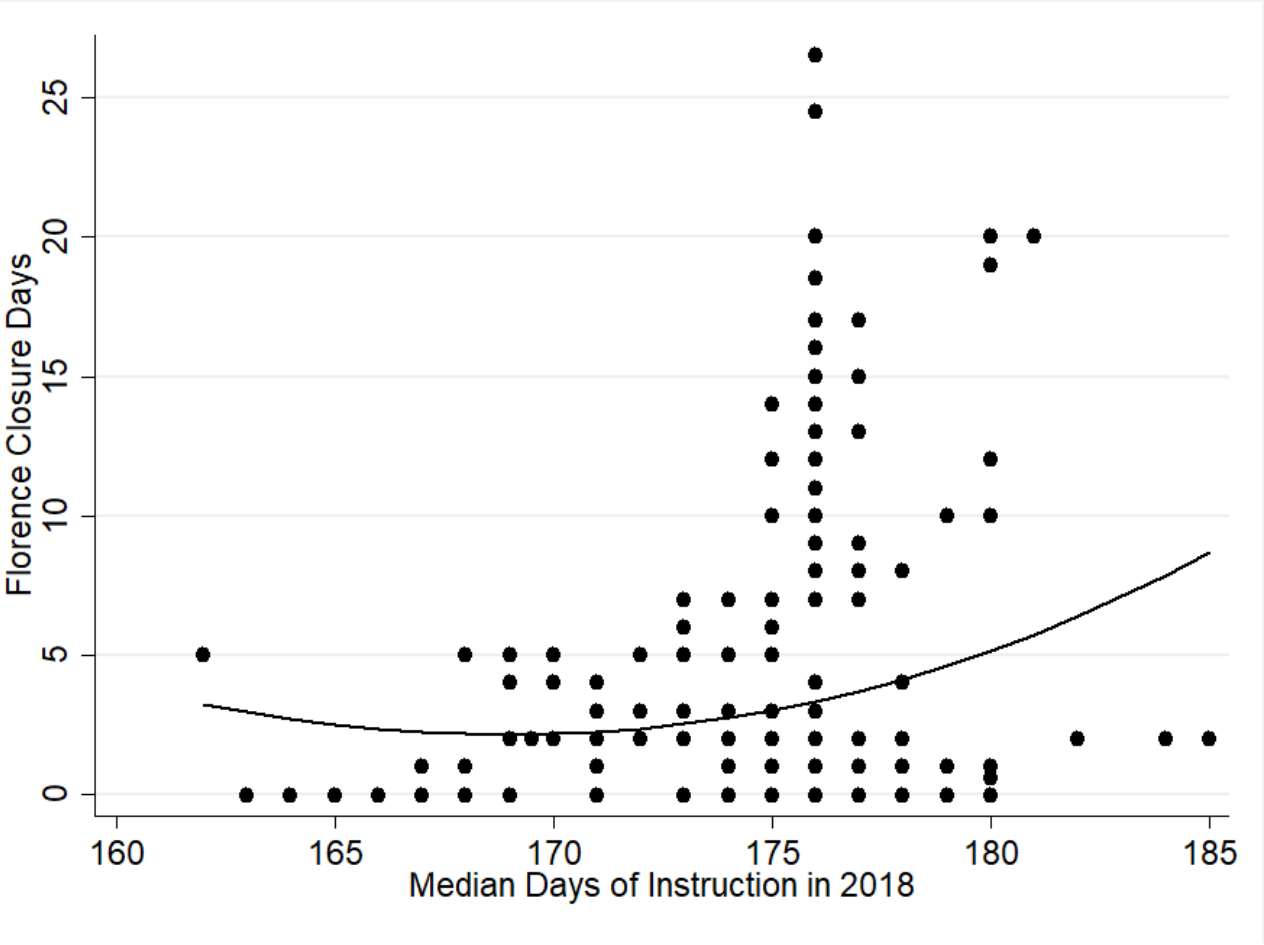
- Huebener, Mathias, Susanne Kuger, and Jan Marcus. 2017. “Increased Instruction Hours and the Widening Gap in Student Performance.” *Labour Economics* 47:15–34.
- Imberman, Scott A., Adriana D. Kugler, and Bruce I. Sacerdote. 2012. “Katrina’s Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees.” *The American Economic Review* 102 (5):2048–2082.
- Jacob, Brian A. and Lars Lefgren. 2003. “Are Idle Hands the Devil’s Workshop? Incapacitation, Concentration, and Juvenile Crime.” *American Economic Review* 93 (5):1560–1577.
- Lambert, Diana. 2019. “California Schools Closed for Unprecedented Number of Days Due to Fire, Power Outages.” *EdSource* .
- Lardieri, Alexa. 2019. “Another Victim of the California Wildfires: Education.” *US News & World Report* .
- Marcotte, Dave E. and Steven W. Hemelt. 2008. “Unscheduled School Closings and Student Performance.” *Education Finance and Policy* 3 (3):316–338.
- Meyers, Keith and Melissa A. Thomasson. 2021. “Can Pandemics Affect Educational Attainment? Evidence from the Polio Epidemic of 1916.” *Cliometrica* 15 (2):231–265.
- Pischke, Jörn-Steffen. 2007. “The Impact of Length of the School Year on Student Performance and Earnings: Evidence from the German Short School Years.” *The Economic Journal* 117 (523):1216–1242.
- Rivkin, Steven G. and Jeffrey C. Schiman. 2015. “Instruction Time, Classroom Quality, and Academic Achievement.” *The Economic Journal* 125 (588):F425–F448.
- Sacerdote, Bruce. 2012. “When the Saints Go Marching Out: Long-Term Outcomes for Student Evacuees from Hurricanes Katrina and Rita.” *American Economic Journal: Applied Economics* 4 (1):109–135.
- Sims, David P. 2008. “Strategic Responses to School Accountability Measures: It’s All in the Timing.” *Economics of Education Review* 27 (1):58–68.
- Walsh, Kevin J. E., John L. McBride, Philip J. Klotzbach, Sethurathinam Balachandran, Suzana J. Camargo, Greg Holland, Thomas R. Knutson, James P. Kossin, Tsz-cheung Lee, Adam Sobel, and Masato Sugi. 2016. “Tropical Cyclones and Climate Change.” *WIREs Climate Change* 7 (1):65–89.

Figure 1: Florence Closure Days in 2019, North Carolina Schools



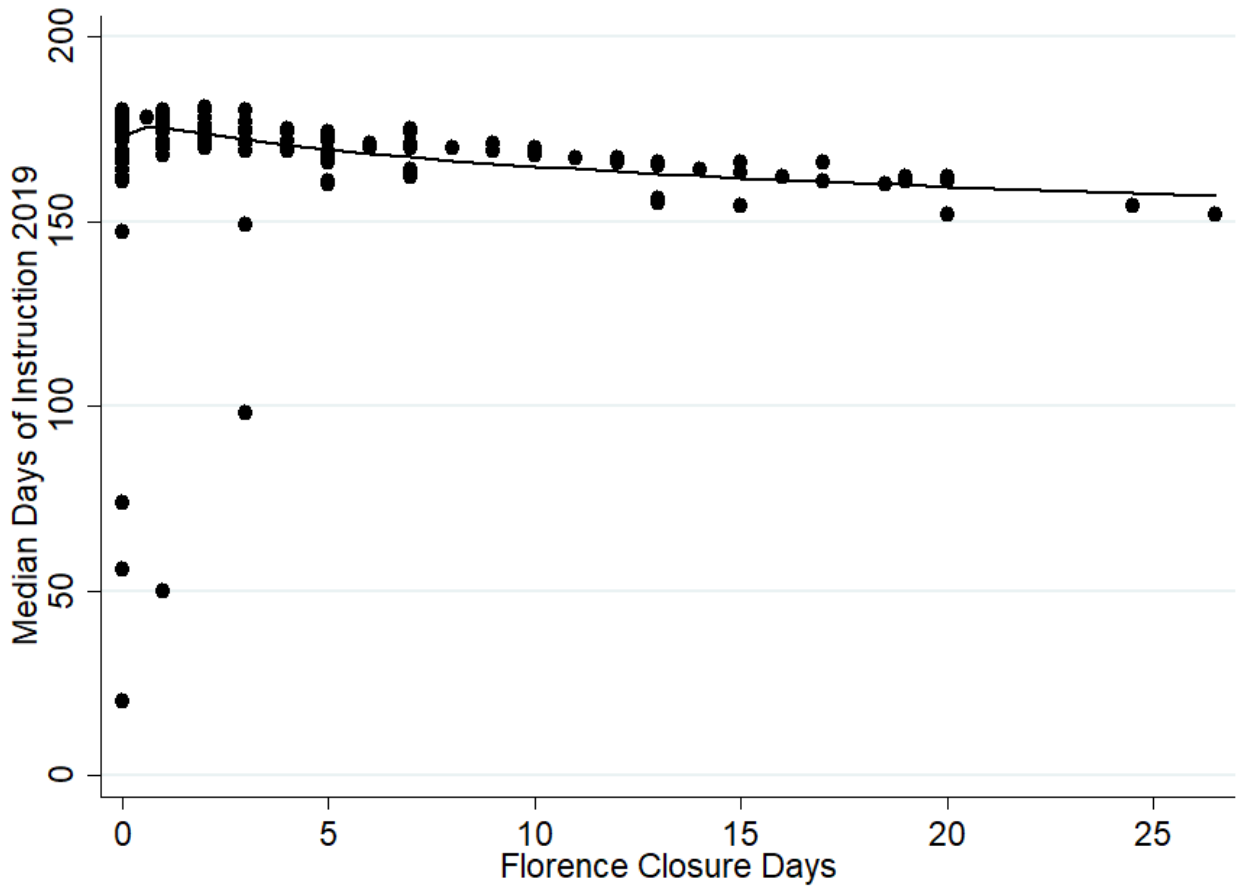
Notes: N = 1,503 schools.

Figure 2: Florence Closure Days in 2019 by 2018 Days of Instruction, North Carolina Schools



Notes: N = 3,186 grade by school observations. Median days of instruction is defined as the median of membership days at the time of the EOG examinations.

Figure 3: 2019 Days of Instruction by Florence Closure Days, North Carolina Schools



Notes: N = 3,186 grade by school observations. Median days of instruction is defined as the median of membership days at the time of the EOG examinations.

Table 1: Sample Summary Statistics

Panel A: Student Characteristics, Time $t - 1$			
	Mean	Std. Dev.	Observations
Male	0.51	0.50	1020384
Black	0.25	0.43	1020384
Hispanic	0.18	0.39	1020384
Other (Non-White)	0.08	0.28	1020384
Disability	0.12	0.32	1020384
Economically Disadvantaged	0.50	0.50	1020384
Limited English Proficiency	0.10	0.31	1020384
Grade 3 (t-1)	0.26	0.44	1020384
Grade 4 (t-1)	0.26	0.44	1020384
Grade 6 (t-1)	0.24	0.43	1020384
Grade 7 (t-1)	0.25	0.43	1020384
Math EOG t-1	0.02	1.00	1020384
Reading EOG t-1	0.01	0.99	1020384

Panel B: Student Instruction Days by Year t			
	Mean	Std. Dev.	Observations
Days of Instruction 2017	165	19.5	335873
Days of Instruction 2018	172	18.1	336680
Days of Instruction 2019	169	18.3	343853
Florence Days Missed 2019	3	4.1	345649
Florence Days Missed 2019, if positive	4	4.2	275496

Notes: Data are derived from the North Carolina Education Research Data Center. Panel A presents background characteristics measured at time $t - 1$, while Panel B presents the sample means for days of instruction in 2017-2019 and days of school missed due to Hurricane Florence in 2019. See Section 2.4 and the Data Appendix for a complete list of sample restrictions.

Table 2: School Closures and Test Score Losses

	Linear (1)	+ School FE (2)	+ Days Instruction (t) (3)	+ Days Instruction (t-1) (4)	FEMA Sample (5)	FEMA Controls (6)
Panel A: Mathematics EOG						
Florence Days Missed	-0.0029*** (0.0009)	-0.0019** (0.0008)	-0.0011 (0.0008)	-0.0011 (0.0008)	-0.0015* (0.0009)	0.0016 (0.0014)
Membership Days (t)			0.0009*** (0.0000)	0.0009*** (0.0000)		
Membership Days (t-1)				-0.0001** (0.0000)		
FEMA Regs per 100						-0.0057*** (0.0022)
IHP Amount per 100 pop						0.0079 (0.0081)
ZCTA Population					0.0000 (0.0000)	0.0000 (0.0000)
Observations	768841	768841	765705	765259	534795	534795
Mean of Dep Var	0.0300	0.0300	0.0332	0.0339	0.0509	0.0509
Panel B: Reading EOG						
Florence Days Missed	-0.0007 (0.0005)	-0.0023*** (0.0006)	-0.0018*** (0.0006)	-0.0018*** (0.0006)	-0.0017*** (0.0007)	-0.0013 (0.0010)
Membership Days (t)			0.0006*** (0.0000)	0.0006*** (0.0000)		
Membership Days (t-1)				0.0001** (0.0000)		
FEMA Regs per 100						-0.0012 (0.0015)
IHP Amount per 100 pop						0.0040 (0.0052)
ZCTA Population					-0.0000 (0.0000)	-0.0000 (0.0000)
Observations	1020384	1020384	1016406	1015836	712108	712108
Mean of Dep Var	0.0146	0.0146	0.0178	0.0185	0.0264	0.0264

Notes: The sample includes fourth, fifth, seventh, and eight (reading only) grade students from AY2016-17 through AY2018-19. The regression specifications include school, grade, and year fixed effects, as well as indicators for student gender, race/ethnicity, disability status, LEP, and economically disadvantaged status. The dependent variable in Panel A is Mathematics EOG and in Panel B is Reading EOG. In columns (5) and (6), the sample is further restricted to students living in a Census Block Group that can be linked to a ZCTA with a population size of at least 90. Hurricane damage controls are derived from publicly available FEMA data on the Individuals and Households Program (IHP), described in detail in the text. Standard errors are clustered at the time $t - 1$ school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Spline in Florence Days Missed

	Instruction Days (1)	Mathematics (2)	Reading (3)
0.6 Days	-0.2106 (0.3785)	0.0873*** (0.0213)	0.0875*** (0.0193)
1 Day	-0.5100 (0.5022)	0.0425*** (0.0113)	0.0206*** (0.0078)
2 Days	-1.3699*** (0.3295)	0.0345*** (0.0098)	0.0369*** (0.0067)
3 Days	-0.2256 (0.3358)	0.0092 (0.0119)	0.0070 (0.0077)
4 Days	-1.4375*** (0.5219)	0.0828*** (0.0285)	0.0096 (0.0139)
5 Days	-2.6404*** (0.7056)	0.0861*** (0.0196)	0.0169* (0.0101)
6 Days	-1.8832*** (0.3697)	0.0241 (0.0168)	-0.0110 (0.0103)
7 Days	-7.5930*** (0.5891)	-0.0616*** (0.0149)	-0.0186** (0.0084)
8 Days	-4.1379*** (0.5392)	0.0623* (0.0369)	0.0076 (0.0186)
9 Days	-3.7363*** (0.6045)	-0.0993*** (0.0285)	-0.0324** (0.0145)
10 Days	-9.5518*** (0.5166)	-0.0038 (0.0203)	-0.0224 (0.0157)
11 Days	-4.5164*** (0.2487)	-0.1230*** (0.0073)	-0.0219*** (0.0047)
12 Days	-7.8687*** (0.6170)	-0.1040** (0.0430)	-0.0523* (0.0300)
13 Days	-17.2320*** (1.0561)	0.0782*** (0.0298)	-0.0244 (0.0234)
14 Days	-10.3634*** (0.5866)	-0.0681** (0.0311)	0.0465** (0.0228)
15 Days	-10.7846*** (0.9928)	0.0072 (0.0376)	0.0318 (0.0271)
16 Days	-13.3345*** (0.2490)	-0.1908*** (0.0073)	-0.0653*** (0.0047)
17 Days	-13.1898*** (2.1417)	0.0691 (0.1148)	-0.0052 (0.0054)
18.5 Days	-16.1952*** (0.5294)	-0.0068 (0.0322)	-0.0054 (0.0239)
19 Days	-17.7741*** (0.3786)	0.0670 (0.0814)	0.0470*** (0.0101)
20 Days	-18.4562*** (0.5909)	0.0280 (0.0363)	-0.0021 (0.0304)
24.5 Days	-22.8763*** (0.3574)	0.0054 (0.0405)	-0.0939** (0.0471)
26.5 Days	-24.5794*** (0.2493)	0.0142 (0.0696)	-0.0043 (0.0067)
Observations	1016406	768841	1020384
Mean of Dep Var	168.6463	0.0300	0.0146

Notes: The dependent variables in each column are: (1) Days of Instruction, (2) EOG Mathematics, and (3) EOG Reading. The “days” variables are dummy variables for days missed due to Hurricane Florence in 2019. The sample includes fourth, fifth, seventh, and eighth (reading only) grade students from AY2016-17 through AY2018-19. The regression specifications include school, grade, and year fixed effects, as well as indicators for student gender, race/ethnicity, disability status, LEP, and economically disadvantaged status. Standard errors are clustered at the time $t - 1$ school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Specification Checks

	Valid t Test (1)	Switch Schools (2)	Days Instruction (3)	EOG Mathematics (4)	EOG Reading (5)	Days Instruction (6)	EOG Mathematics (7)	EOG Reading (8)
Florence Days Missed x 2019	-0.0002 (0.0001)	0.0019* (0.0011)	-0.8728*** (0.0271)	-0.0015* (0.0009)	-0.0017** (0.0007)	-0.8744*** (0.0269)	-0.0016* (0.0009)	-0.0017** (0.0007)
(False) Days x 2015						0.0459 (0.0303)	0.0007 (0.0012)	-0.0018** (0.0008)
(False) Days x 2016						0.0352 (0.0272)	0.0009 (0.0011)	0.0008 (0.0006)
(False) Days x 2017			0.0686*** (0.0247)	0.0008 (0.0009)	0.0013* (0.0007)	0.0695*** (0.0249)	0.0008 (0.0009)	0.0012* (0.0007)
Observations	1073442	1020384	1022451	771515	1023053	1690101	1265110	1689685
Mean of Dep Var	0.9506	0.1108	168.5512	0.0274	0.0129	167.9189	0.0331	0.0224

Notes: The sample in Column (1) is all students who have a valid time $t - 1$ test score, and the dependent variable is having a valid year t reading test score. The dependent variable in Column (2) is having a time t test score at a different North Carolina school than that attended at time $t - 1$. The dependent variables in Columns (3) and (6) are days of instruction, Columns (4) and (7) are mathematics EOG, and Columns (5) and (8) are reading EOG. Columns (1)-(5) include AY2016-17 through AY2018-19. Columns (6)-(9) add AY2014-15 and AY2015-16. "False" treatments assign the days closed due to Hurricane Florence to the year as indicated. All specifications include year, grade, and time $t - 1$ school fixed effects, as well as controls for the students' gender, race/ethnicity, disability status, eligibility for free or reduced price lunch, and having limited English proficiency. Standard errors are clustered at the time $t - 1$ school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Heterogeneity by Prior Achievement

(1)	
Panel A: Mathematics EOG, N = 768,841	
Florence Days Missed (Ref: 3rd Quintile)	-0.0040*** (0.0010)
x 1st Quintile	0.0040*** (0.0009)
x 2nd Quintile	0.0002 (0.0007)
x 4th Quintile	0.0010 (0.0006)
x 5th Quintile	0.0069*** (0.0009)
Panel B: Reading EOG, N = 1,020,384	
Florence Days Missed (Ref: 3rd Quintile)	-0.0017** (0.0007)
x 1st Quintile	-0.0033*** (0.0007)
x 2nd Quintile	-0.0010* (0.0006)
x 4th Quintile	-0.0003 (0.0006)
x 5th Quintile	0.0014** (0.0006)

Notes: All specifications include covariates as described in Table 2, including school fixed effects, except that these models include quintile-specific, linear prior test score controls. Standard errors are clustered at the time $t - 1$ school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Heterogeneity by Grade and Demographic Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Mathematics EOG, N = 768,841						
Florence Days Missed	-0.0014 (0.0010)	-0.0027*** (0.0009)	-0.0019** (0.0009)	-0.0021** (0.0008)	-0.0015* (0.0009)	-0.0018 (0.0011)
x Middle School	-0.0013 (0.0014)					-0.0013 (0.0014)
x Male		0.0015*** (0.0004)				0.0015*** (0.0004)
x EDS			-0.0000 (0.0005)			0.0000 (0.0005)
x LEP				0.0032*** (0.0010)		0.0046*** (0.0010)
x Black					-0.0006 (0.0008)	-0.0006 (0.0008)
x Hispanic					-0.0004 (0.0008)	-0.0021*** (0.0008)
x Other NW					-0.0021** (0.0009)	-0.0025*** (0.0009)
Panel B: Reading EOG, N = 1,020,384						
Florence Days Missed	-0.0021*** (0.0007)	-0.0027*** (0.0006)	-0.0025*** (0.0007)	-0.0025*** (0.0006)	-0.0026*** (0.0006)	-0.0028*** (0.0008)
x Middle School	-0.0005 (0.0009)					-0.0005 (0.0009)
x Male		0.0007* (0.0004)				0.0007* (0.0004)
x EDS			0.0004 (0.0004)			0.0001 (0.0005)
x LEP				0.0027*** (0.0008)		0.0019* (0.0010)
x Black					0.0004 (0.0006)	0.0003 (0.0006)
x Hispanic					0.0020*** (0.0006)	0.0012 (0.0008)
x Other NW					-0.0005 (0.0007)	-0.0007 (0.0007)

Notes: EDS refers to Economically Disadvantaged Status, LEP is Limited English Proficiency, and Other NW is all students who do not fit into the categories non-Hispanic White, non-Hispanic Black, or Hispanic. All specifications include covariates as described in Table 2, including school fixed effects. Standard errors are clustered at the time $t - 1$ school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A Online Appendix

The data are derived from restricted-access administrative data records from the North Carolina Education Research Data Center (NCERDC). The analysis dataset includes students from AY2015-16 through AY2018-19, allowing the construction of information for the prior year ($t-1$). The Master Build files contain students' test scores on end-of-grade (EOG) examinations and data on accommodations, exemptions, and demographics and socioeconomic status, including race/ethnicity, gender, economic disadvantage, English learner status, academically or intellectually gifted status, and disability status.²¹ This is supplemented with data on attendance. For the small number of duplicate observations, we retain the highest test score and the most number of membership days (if available). North Carolina requires students in grades 3 through 8 to take standardized end-of-grade tests (EOG) in the spring of each academic year. Students are tested in reading each year, in mathematics in grades 3 through 7, and in science in grades 5 and 8. Starting in the 2017-18 school year, 8th grade students enrolled in Math 1 were exempted from taking the 8th grade EOG and instead complete the Math 1 EOC examination. Therefore, our analysis excludes 8th grade mathematics. The sample is restricted to traditional, non-charter schools, with non-missing values for the days closed due to Hurricane Florence. For the analysis, student school characteristics and individual characteristics are all measured as of time ($t-1$), with only test scores and time t instructional days measured at time t . Therefore, the sample only includes schools where a student in 4th, 5th, 7th, or 8th grade would have been in the same school in year $t-1$. For the extended dataset used in Table 4, days of instruction are not available in the data, so the sample is not restricted based on having between 150-200 days of instruction at time ($t-1$).

²¹NCDPI demographic and economically disadvantaged student information can be found here <https://files.nc.gov/dpi/documents/fbs/resources/data/factsfigures/2015-16figures.pdf> and here <https://childnutrition.ncpublicschools.gov/information-resources/eligibility/data-reports/sy15-16.xlsx>

Table A1: Dataset Construction

	LEA	Schools	Students
Years (t-1) 2016-2018, Grade (t-1) 3rd, 4th, 6th, and 7th, with valid t-1 demographics and test scores	284	2,081	1,394,825
<u>School-level restrictions</u>			
Restrict to regular and non-charter schools	115	1,844	1,294,753
Non-missing values for days missed (NCDPI data)	115	1,807	1,282,256
Restrict to traditional calendar	115	1,684	1,192,995
Remove 3 districts impacted severely by Hurricane Matthew	112	1,633	1,158,494
Remove incompatible grade span configurations	107	1,503	1,074,916
Keep only schools with (t-1) median membership days between 150-200	107	1,503	1,073,442
<u>Time (t) student-level restrictions</u>			
Restrict to valid time (t) reading test scores	107	1,503	1,020,384
Restrict to valid time (t) mathematics test scores	107	1,499	768,841