

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Gust, Sarah

Working Paper (Not) going to school in times of climate change: Natural disasters and student achievement

ifo Working Paper, No. 413

Provided in Cooperation with: Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Gust, Sarah (2024) : (Not) going to school in times of climate change: Natural disasters and student achievement, ifo Working Paper, No. 413, ifo Institute - Leibniz Institute for Economic Research at the University of Munich, Munich

This Version is available at: https://hdl.handle.net/10419/305818

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



WWW.ECONSTOR.EU



(Not) Going to School in Times of Climate Change: Natural Disasters and Student Achievement

Sarah Gust



Imprint:

ifo Working Papers Publisher and distributor: ifo Institute – Leibniz Institute for Economic Research at the University of Munich Poschingerstr. 5, 81679 Munich, Germany Telephone + 49(0)89 9224 0, Telefax +49(0)89 985369, email ifo@ifo.de www.ifo.de

An electronic version of the paper may be downloaded from the ifo website www.ifo.de

(Not) Going to School in Times of Climate Change: Natural Disasters and Student Achievement *

Sarah Gust gust@ifo.de

ifo Institute at the University of Munich

October 23, 2024

Abstract

Climate change poses significant risks due to rising frequency and severity of natural disasters. These disasters can disrupt education through school closures, infrastructure damage, financial stress, and health issues. This study assesses their impact on student achievement by combining US county-level data on FEMA disaster declarations with the Stanford Education Data Archive. Exploiting variation in disasters across counties and years in a staggered two-way fixed effects design, event study estimates reveal a persistent negative effect of natural disasters on student achievement for up to five years post disaster. Counties with above-average per-pupil expenditure show quicker recovery. The study sheds light on effect heterogeneity by gender, socio-economic background, and disaster characteristics.

Keywords: Natural Disasters, Student Achievement, Education Economics, Disaster Resilience, Human Capital

JEL classification: Q54, I21, I25, O44

^{*}I gratefully acknowledge comments from Ludger Woessmann, Raphael Brade, Vera Freundl, Ingrid Haegele, Todd Jones, Lindsey Macmillan, Nico Pestel, Michel Serafinelli, Simon ter Meulen, Maria Waldinger, Katharina Werner, the seminar participants of the European Economic Association, the German Economic Association, the ifo Center for the Economics of Education, and the Canazei Winter School. I thank Fariba Dorpoush for excellent research assistance. Special thanks are also due to the SEDA team for their outstanding service to the research community and for providing additional information on school expenditure.

1 INTRODUCTION

The rising risk of natural disasters and extreme weather due to climate change poses significant challenges globally. In 2021, the United States experienced a record-breaking streak, with seven consecutive years of over ten billion-dollar disasters (Smith, 2022). Population growth and development patterns increase the damage costs of these events (Smith, 2022). Extending beyond immediate physical damage to broader socio-economic dimensions is crucial to comprehend the multifaceted impacts of natural disasters and long lasting effects of climate change. One critical aspect is the disruption caused to educational systems. Nearly half of the school closures in the US, before the Covid-19 pandemic, were attributed to natural disasters (Jahan et al., 2022). While school closures are perhaps the most visible manifestation of how natural disasters affect learning, it represents only one facet of a complex phenomenon. Natural disasters can affect child health (Currie and Rossin-Slater, 2013), school quality, and infrastructure damage, as well as induce family income instability (Deuchert and Felfe, 2015) and housing displacement with unequal effects for different socio-economic groups (Nguyen and Minh Pham, 2018).

This paper investigates the effects of natural disasters on student achievement and unravels the underlying mechanisms. Focusing on student achievement allows for a nuanced examination of the cognitive skill component of human capital, capturing variations in any skill inputs including ability, child health, family support, school resources, and institutional characteristics (Woessmann, 2016). I combine county level student achievement, achievement gaps, demographic compositions, and school financial information from the Stanford Education Data Archive (SEDA) (Reardon et al., 2019) with disaster declarations from the Federal Emergency Management Agency (FEMA).

The variation of natural disasters across US counties and school years serves as the foundation for my analysis in a two-way fixed effects framework. First, I study how the effect of natural disasters dynamically evolves over the school years following the event, using the event study approach by Sun and Abraham (2021). Sun and Abraham (2021) provide a method particularly for the event study setting, making it an especially appro-

priate baseline model for this study. Second, I estimate static two-way fixed effects models with several treatment specifications to study different intensity channels such as the type of natural disaster, the frequency by which natural disasters strike different counties, and disaster size along various dimensions such as damage, fatalities, and duration.

I find persistent negative effects of natural disasters on student achievement for up to five years after a natural disaster hits a county. Natural disasters adversely affect even those cohorts who were not yet enrolled at the time of the disaster, an effect that cannot be explained by school closures. Possible explanations include a decline in children's health, which hinders cognitive and social-emotional development, changes in the student composition of affected areas, and reduced school quality due to infrastructure damage or the loss of skilled teachers. I show suggestive evidence that financial distress in families may be a key mechanism. Specifically, natural disasters increase the share of students receiving free or reduced lunch, which seems to be unrelated to migration patterns. I present new evidence related to the resilience of schools to natural disasters and show that counties with above-average per-pupil expenditure recover more quickly from natural disasters.

Exploring different treatment intensities shows that not only very large disasters drive the results. Although point estimates are consistently higher for large disasters, belowaverage disasters have a significant negative effect on student achievement. In terms of the type of disaster, volcanic activity and landslides have the most devastating effect. Those disasters occur very rarely and might be harder to predict. Also hurricanes, fires, and earthquakes lead to significantly lower student achievement. Generally, predictability does not alleviate all damage. My results show that more frequent disasters are more harmful, indicating that a higher frequency leaves the county little time to recover and does not allow them to adapt sufficiently to disasters.

I contribute to the literature on the adverse effects of natural disasters during early life which shows that in-utero and post-birth exposure to natural disasters (Currie and Rossin-Slater, 2013) or pollution (Klauber et al., 2024) negatively affect various health outcomes, cognitive skills, and income (Karbownik and Wray, 2019). This strand of literature attributes the persistent negative results to a reduction in the health stock of children. While this might explain parts of the effects, I demonstrate that school inputs play a crucial role. The existing literature typically focuses on one specific major disaster event such as hurricanes (Özek, 2023; Sacerdote, 2012) or earthquakes (Di Pietro, 2018), which are often extreme outliers as the majority of disasters are not as severe. My paper is more closely related to Opper et al. (2023) who focus on the static effects of differently sized disasters on human capital and migration.¹ By leveraging recent advancements in the difference-in-difference methodology, I derive robust estimates regarding how these effects evolve over time.

The remainder of the paper is organized as follows: Section 2 illustrates the possible mechanisms and presents the existing empirical evidence. Section 3 describes the different data sources in more detail. Section 4 contains the two main empirical strategies, including the event study design and the static two-way fixed effects model. The results are presented in Section 5, and Section 6 concludes.

2 CONCEPTUAL FRAMEWORK AND EXISTING EVIDENCE

Natural disasters can affect educational outcomes through several mechanisms. ² First, natural disasters affect the shadow price of quality education, reflecting its accessibility (Nguyen and Minh Pham, 2018). This includes infrastructure damage, temporary or permanent school closures, as well as effects on teaching staff.

¹Opper et al. (2023), who developed their paper in parallel with this one, estimate the impact of disasters on net migration, average test scores, high school graduation rates, and post-secondary enrollment rates using FEMA disaster declarations and the SEDA database. Their study differs from my paper in three main ways: First, they focus on the effects of varying disaster sizes based on property damage, while this study also considers disaster frequency and type. Second, they present only static effects on first-difference outcomes, whereas this study provides robust event study results. Third, this paper offers additional insights into the impacts on student composition, adult mental health, and variations by gender, socio-economic background, and per-pupil spending.

 $^{^{2}}$ Nguyen and Minh Pham (2018) present a simple model to illustrate the different mechanisms through which natural disasters affect children's development, which guides the structure of this section.



Figure 1 Days of school closures in the US between 2011 and 2019

Note: The figure shows unplanned school closure days in the US by general causes on panel a) and by type of natural disaster in Panel b). Jahan et al. (2022) conducted daily systematic online searches to collect data on publicly announced unplanned school closures lasting at least one school days in the United States from August 1, 2011, through June 30, 2019

Natural disasters and extreme weather events account for the most frequent cause of prolonged unplanned school closures in the US prior to the Covid-19 pandemic, as depicted in Figure 1 panel a). Panel b) splits those natural disaster school closures by disaster type. Hurricanes, ice and snowfall cause most of the school closures. There is a wide range of literature showing that instruction time correlates with student performance (Aucejo and Romano, 2016; Jaume and Willén, 2019; Lavy, 2015; Pischke, 2007; Wedel, 2021).³ In Maryland, students experiencing reduced instruction time due to unscheduled closures during snowfall performed worse on high-stakes exams (Marcotte, 2007). Drawing from Massachusetts data, Goodman (2014) provides evidence that coordination problems in the classroom, as outlined by Lazear (2001), play an important role in this context. While heavy snowfall leads to coordinated school closures and no effect on achievement, moderate snowfall induces student absence and reduces math achievement by 0.05 standard deviations. The Covid-19 pandemic has also provided further insights into the

³See Blanden et al. (2023) for a detailed review of the evidence.

detrimental impact of school closures. Both cognitive and socio-emotional development were significantly impeded, with students from disadvantaged homes experiencing more severe setbacks (Engzell et al., 2021; Werner and Woessmann, 2023). However, natural disasters do not necessarily increase the cost of quality education. Sacerdote (2012) shows that students forced to switch school after the hurricanes Katrina and Rita experienced a sharp decline in test scores in the first year after the hurricanes. Yet, the long-run effects are mixed, with Sacerdote (2012) identifying an improvement in test scores among low-performing students placed in higher-quality educational settings.

Secondly, natural disasters can increase the costs of good health or impede the access to a healthy environment (Nguyen and Minh Pham, 2018). The evidence on wild fires is closely related to the broader literature on pollution exposure (Wu, 2022). Currie et al. (2009) show that a rise in pollution increases absence, possibly due to health issues, decreases cognitive skills (Lavy, 2015), but also hinders long-run knowledge acquisition due to impaired brain development (Block and Calderón-Garcidueñas, 2009). Currie and Rossin-Slater (2013) find negative effects of hurricane exposure during pregnancy on children. They demonstrate that the evidence regarding additional impacts of hurricanes during pregnancy on outcomes such as birth weight – subsequently influencing adult height, IQ, earnings, and education (Black et al., 2007) – is more mixed and depending on the specification. Klauber et al. (2024) observe no effect of cleaner air around birth on birth weights but show that children require less medication for at least five years. Fuller (2014) shows that hurricane exposure during pregnancy in North Carolina translates into lower standardized test scores in maths and reading by the third grade, while children exposed to floodings or tornadoes also exhibit somewhat diminished math performance. Karbownik and Wray (2019) study the long-run effects of hurricane exposure in utero and as an infant by using World War I draft records linked to census data and find that white males had 5% lower income. The literature on in-utero exposure links these effects to stress during pregnancy. To test this channel, I incorporate data on adult mental health.

Thirdly, natural disasters can affect children's education through household income, wages, and increased costs of other commodities (Arouri et al., 2015; Boustan et al., 2020; Masozera et al., 2007; Pleninger, 2022). Deuchert and Felfe (2015) demonstrate that damages to real estate redirect investment toward house reconstruction, potentially diverting resources away from children's health and education. I contribute to this channel by estimating the effects of natural disasters on the share of students receiving free or reduced lunch, which serves as an indicator of increased financial distress in affected households.

Certain characteristics influence the vulnerability to natural disasters (Cutter et al., 2008), including socio-economic status, race, ethnicity, and gender, as wealth and social safety nets can facilitate recovery. Boys often react more strongly to disruptive (family) events, showing fluctuations in test scores and increased disciplinary issues (Autor et al., 2016, 2019; Bertrand and Pan, 2013). My research enhances the literature by analyzing achievement patterns separately for boys and girls and assessing how natural disasters impact the gender achievement gap and the socio-economic gap.

Vulnerability is closely linked to disaster resilience, which Cutter et al. (2008) classify into social, economic, institutional, infrastructure, and community capital. Insurances play a crucial role in mitigation: Pleninger (2022) demonstrates that unemployment insurance effectively reduces the impact of natural disasters on income. Moreover, insurance and revenue diversification enhance the resilience of municipal bonds against price drops following natural disasters (Auh et al., 2022). However, municipalities with above-average racial minority compositions experience greater expenditure losses and a heightened debt default risk compared to average municipalities in the decade following a hurricane (Jerch et al., 2023). In turn, this constrains public expenditures: Deryugina (2017) illustrates that hurricanes lead to increased unemployment and disability insurance claims but a decline in educational assistance transfers. I contribute to this literature by studying whether counties with higher pre-existing per-pupil expenditures display enhanced resilience to natural disasters.

The literature underscores that the mechanisms might vary depending on the type of natural disaster. In this paper, I study the effect of different types of natural disasters. While numerous studies concentrate exclusively on individual types of disasters, such as hurricanes (Sacerdote, 2012) or earthquakes (Di Pietro, 2018), which are often extreme outliers and relatively infrequent occurrences, the majority of disasters are not as severe. I analyze heterogeneity across various dimensions of disaster size, including factors such as fatalities, costs, and duration.

3 DATA

3.1 Natural Disasters

The OpenFEMA Dataset by the Federal Emergency Management Agency (FEMA) of the Department of Homeland Security contains all major disaster declarations since 1964. A disaster declaration made by the President of the United States and only in strongly affected areas that struggle to deal with the consequences, which rules out any inconsequential natural disasters. Every disaster declaration includes the date the disaster was declared, the area, the type of incident, and which assistance program was declared. One disaster can cause multiple disaster events across different counties. The years are adjusted to align with school years, so the year variable matches the education data.

For large disasters, there is information on the number of deaths, the number of injured, the number of people that got homeless, reconstruction costs, insured damage and total damage in the international emergency events database (EM-DAT) that I merge via county and start date. Generally, EM-DAT considers only disasters that caused more than 10 deaths, that affected more than 100 people, or that called for international assistance or an emergency declaration. This information refers to overall fatalities and damage for a disaster and is not specified on a county level.

Panel a) Figure 2 shows the number of disaster declarations in FEMA by year. Each disaster event is only counted once, even when it affected multiple counties. In 2011, the year with the most disaster declarations, there were 177 events, while in 2015, there were the fewest declarations (79).⁴ Since then, the number has been increasing every year.

⁴Smith (2022) shows that the frequency and costs of severe natural disasters have been on the rise over the past four decades since the 1980s, with variations observed between individual years.





Note: Panel a) shows total events declared as disaster in FEMA. Every disaster is counted only once even when declared in multiple locations. Panel b) shows number of disasters by disaster type in FEMA between 2009 and 2018. Panel c) shows average disaster fatalities if information on fatalities exists. Panel d) shows average value of the damage in thousands of US dollars at the moment of the event adjusted for inflation using CPI. Note that EM-DAT includes only disasters with at least 10 fatalities, 100 affected or if a country called for international assistance or an emergency declaration.

Panel b) shows disasters by type: 502 fires, 414 storms, 4 earthquakes, and 2 volcanic activities. Panel c) displays the average fatalities for those disasters for which we can merge EM-DAT data. Hurricanes have the highest average fatalities with 45 deaths, followed by landslides with 43 fatalities even though those were only 4 events. Fires are the most frequent type, but they have no fatalities. Panel d) depicts the average disaster damage in thousands of dollars as recorded in EM-DAT. In EM-DAT, adjusted damage refers to the financial losses caused by a disaster, which are normalized to account

for factors such as inflation using CPI. Hurricanes emerge as the most economically devastating disasters, with storms following closely behind, causing approximately half the financial losses incurred by hurricanes. In contrast, landslides and fires are characterized by minimal direct financial impact.

3.2 Outcome Data

Student achievement: The Stanford Education Data Archive (SEDA) offers a unique source for school district- and county-level student achievement measured by standardized test scores for the 2008/09 to 2017/18 school years (Fahle et al., 2021). Moreover, it contains achievement gaps by gender and socio-economic status, demographic, and socio-economic data. Achievement in SEDA is based on the EDFacts data system that contains test data for all students in grades three to eight in maths and reading each year. EDFacts does not contain individual student-level data but the number of students in each school, subgroup, subject, grade, and year scoring at each performance level. Unfortunately, every state can design their own test and benchmark for performance levels, such that the data are not comparable across states. SEDA transforms these state specific benchmarks and places them onto a common scale across states and years using the National Assessment of Educational Progress (NAEP) and calculates achievement for different subgroups and geographical units. They exclude cases with low participation and insufficient data. Finally, the score is standardized by subtracting the average of the four national cohorts that were in fourth grade in 2009, 2011, 2013, and 2015 and dividing by the national grade-subject-specific standard deviation of this reference cohort. Consequently, a county mean of 0.5 indicates that the average student scored approximately 0.5 standard deviations higher than the average national reference cohort in that same grade. One standard deviation on this scale is approximately three grade levels. Estimates on this scale are comparable across the US and over time by relying on the stability of the NAEP scale over time.⁵

⁵See Reardon et al. (2019) for a more detailed description on the SEDA data construction.

Table 1 provides summary statistics extracted from the SEDA dataset for several key variables. The final analysis will focus on the youngest children in grade three. However, I will also show results for grades four and five.⁶ The math grades for grades three, four, and five have mean scores ranging from -0.016 to -0.064. The socio-economic gap in math for grade three has a mean of -0.229. Conversely, the male-female gap appears relatively low, with a mean of -0.023, suggesting minimal disparity between genders in math and grade three. On average, there are 13% Black students and 70% White students. 60% of the students get a free or reduced lunch. School size is captured through log enrollment, which has a mean value of 5.969. Furthermore, SEDA provides financial aspects of education, revealing mean values of \$11,524 for total per-pupil expenditure and \$11,572 for total per-pupil revenue. I will show results by high and low per-pupil expenditure, where high per-pupil expenditure exceeds \$11,000, which is close to the median value. Finally, Table 1 displays the population estimate for 2008 that I use to weight the regressions.

Table 1Summary statistics SEDA data

Statistic	Ν	Mean	Median	St. Dev.	Min	Max
Math grade 3	$15,\!930$	-0.016	-0.005	0.298	-1.513	1.120
Math grade 4	15,930	-0.047	-0.030	0.301	-1.621	1.052
Math grade 5	15,930	-0.064	-0.045	0.299	-1.589	1.075
Socio-economic gap	$15,\!663$	-0.229	-0.222	0.265	-1.539	0.778
Male-female gap	$15,\!624$	-0.023	-0.012	0.285	-1.496	1.053
Percent Black students	$15,\!930$	0.134	0.032	0.206	0.000	1.000
Percent White students	$15,\!930$	0.694	0.763	0.264	0.000	1.000
Percent free/reduced lunch	$15,\!930$	0.596	0.594	0.170	0.087	1.000
Log enrollment	15,930	5.969	5.814	1.230	2.890	11.158
Total per-pupil expenditure	15,860	11,524.460	10,936.790	2,519.827	4,920.042	30,126.450
Total per-pupil revenue	15,862	11,571.940	11,081.720	2,355.729	5,305.988	28,087.980
Population estimate 2008	15,930	82,383.660	27,729	$233,\!250.100$	1,469	$5,\!256,\!705$
Inflow	21,942	0.050	0.046	0.022	0.007	0.703
Outflow	21,942	0.049	0.046	0.019	0.012	0.419
Mental health	$23,\!013$	3.614	3.600	0.872	0.630	8.300

Note: Summary statistics for the county level data from the Stanford Educational Archive 4.1 for 2009 to 2018, county level in- and out-migration from the Statistics of Income Division migration data for 2010 to 2018, and average number of poor mental health days from CHR&R for 2010 to 2018.

SEDA is a repeated aggregate cross-section. However, Reardon et al. (2019) show that it highly correlates with longitudinal data on school and district level. The correlation

⁶Grades six to eight have substantially more missing values in SEDA.

is higher at the district level (r = 0.87) than at the school level (r = 0.80), which they attribute to higher mobility between schools than between districts. On the county level, the mobility-induced measurement error in learning rates will be even lower.

In- and out-migration: A natural disaster might increase mobility between counties. To study this impact, I add county level in- and out-migration from the Statistics of Income Division migration data. The Statistics of Income Division uses the number of personal exemptions claimed to approximate the number of individuals. Total values for migration to and from other US counties and abroad are available for 2010 to 2018. I divide the total inflow and total outflow by the population estimate from 2008. Table 1 contains the summary statistics. One limitation of the data is that those who are not required to file United States Federal income tax returns are not included. Thus, elderly and people of low socio-economic status are underrepresented. Additionally, the data contains estimates for the aggregate adjusted gross income of the in- and outflows.

Adult mental health: The county-level mental health data for 2010 to 2018 come from the County Health Rankings & Roadmaps (CHR&R) (Remington et al., 2015), a program of the University of Wisconsin Population Health Institute. The main outcome variable is the average number of poor mental health days among adults. Table 1 contains the summary statistics.

4 IDENTIFICATION STRATEGY

4.1 Difference-in-Differences Event Study Design

To explore the evolution of the impact of a natural disaster over time and evaluate the validity of the parallel trend assumption, I adopt a difference-in-differences event study design. Given the staggered occurrence of natural disasters across different time periods, various counties may exhibit learning curves or encounter changes in external conditions, resulting in heterogeneous treatment effects. The presence of such heterogeneous treatment effects complicates the identification of a clean control group and conventional DiD models are prone to generating biased estimates (Borusyak et al., 2024; Callaway and Sant'Anna, 2021; De Chaisemartin and d'Haultfoeuille, 2020; Gardner, 2022; Goodman-Bacon, 2021). Sun and Abraham (2021) illustrate this contamination concerning the coefficients related to lead and lag indicators within a dynamic specification of the two-way fixed effects model. They propose a re-weighting procedure to address this bias. Their target parameter is the cohort average treatment effect on the treated (CATT) for a treatment cohort e and a relative time period l:

$$CATT_{e,l} = E[Y_{i,e+l} - Y_{i,e+l}^{\infty}|E_i = e]$$

$$\tag{1}$$

In my setting, a treatment cohort e are counties that are treated in the same year. Panel a) of Figure A1 shows the first treatment period of each county. l are the periods to the initial natural disaster of county i within the observed time frame. $Y_{i,e+l}$ is the outcome in response to the treatment. $Y_{i,e+l}^{\infty}$ is the potential outcome of county i in a world where it is untreated. Each $CATT_{e,l}$ is then the average treatment effect l periods from initial treatment for all counties first treated at time e. The main identifying assumptions are parallel trends in the baseline outcomes and no treatment effect on pre-treatment periods, and potential treatment effect heterogeneity.

I estimate all $CATT_{e,l}$ using a linear two-way fixed effects event study model that interacts relative period indicators with cohort indicators, excluding the last treated as control C and the year before treatment l = -1:

$$Y_{i,t} = \alpha_i + \lambda_t + \sum_{e \notin C} \sum_{l \neq -1} \delta_{e,l} (1\{E_i = e\} \cdot D_{i,t}^l) + \epsilon_{i,t},$$

$$(2)$$

The county fixed effects α_i remove any geographic differences between the counties that are time-invariant, such as risk of experiencing a natural disaster, elevation, proximity to the coast or differences in time-constant spending. The year fixed effects λ_t capture any time-invariant factors across different years. $1\{E_i = e\}$ is an indicator for a county *i* belonging to treatment cohort *e*, and $D_{i,t}^l$ is an indicator for county *i* being *l* years away from treatment. Alternatively, one can also use never-treated counties as control, which I implement as a robustness check.

 $\hat{\delta}_{e,l}$ is the difference-in-differences estimator for $CATT_{e,l}$ that needs re-weighting. Sun and Abraham (2021) estimate the weights $Pr(E_i = e | E_i \in [-l, T - l])$ with total periods T by sample shares of each cohort in the relevant period $l \in g$ from its corresponding set g.

To derive the interaction-weighted estimator, Sun and Abraham (2021) take the weighted average over all estimates for $CATT_{e,l}$ multiplied by the sample share of each cohort in the period Pr(E = e):

$$\hat{v}_g = \frac{1}{|g|} \sum_{l \in g} \sum_e \hat{\delta}_{e,l} \widehat{Pr}(E_i = e | E_i \in [-l, T-l])$$
(3)

Under parallel trends, limited anticipation, and potential treatment effect heterogeneity, \hat{v}_g is consistent. Unless specified otherwise, standard errors are clustered at the county level.

4.2 Static Two-Way Fixed Effects

Some counties experience multiple disasters, and disasters vary in intensity. To illustrate these channels, I estimate a static two-way fixed effects model where several post treatment indicators are collected in a set. The baseline equation is the following:

$$Y_{i,t} = \alpha_i + \lambda_t + \beta D_{i,t}^{1-5} + \epsilon_{i,t}, \tag{4}$$

where α_i and λ_t are county and year fixed effects, respectively. First, I binarize the treatment following Callaway et al. (2024) such that $D_{i,t}^{1-5} = 1\{D_{i,t} > 0\}$ is a dummy of whether a county experienced at least one disaster in the past five years in a county and year, similar to Deryugina (2017) and Jerch et al. (2023).⁷ The control group are counties that did not experience a natural disaster in the past five years. As a robustness check,

⁷The dummy includes the past five years to capture longer term effects of the treatment.

I include state times year fixed effects to compare counties within states and account for time-varying factors specific to each state. Standard errors are clustered at the county level and all estimations are weighted by the county population in 2008.

The first intensification channel is the severity of a natural disaster. I estimate the equation:

$$Y_{i,t} = \alpha_i + \lambda_t + \delta_1 Major_{i,t}^{1-5} + \delta_2 Minor_{i,t}^{1-5} + \epsilon_{i,t},$$
(5)

 $Major_{i,t}^{1-5}$ is an indicator equal to 1 if a county experienced at least one major disaster in the past 5 years. $Minor_{i,t}^{1-5}$ is an indicator equal to 1 if the county experienced no major disaster but at least one minor disaster in the past five years, similar to Jerch et al. (2023). All other specifications are identical to Equation 4. $\hat{\delta}_1$ and $\hat{\delta}_2$ provide estimates of the impact of any type of disaster that falls into either of these two categories. I use three alternative methods to measure the severity of a disaster. In the first specification, I follow Boustan et al. (2020) and define major natural disasters as such that caused more than 25 deaths, which corresponds roughly to the median value of fatalities. Boustan et al. (2020) argue that the actual number of fatalities might be determined by economic development, which is why they avoid using the actual number of fatalities and prefer to use this simple threshold. Alternatively, one can distinguish between major and minor disasters based on costs. I define major natural disasters as such that cause more than one billion dollar (adjusted) total damage. Lastly, I split disasters on whether the event lasted more than 50 business days or less.

The second intensification channel is the number of natural disasters. More frequent disasters could be more detrimental if counties have no time to recover (Pleninger, 2022). Counties might also enhance their level of protection through increased exposure to disaster events.

Let the number of natural disasters be the dose d that a county experienced in the past five years. To study if more disasters cause more harm or whether counties are adapting, I estimate the following model:

$$Y_{i,t} = \alpha_i + \lambda_t + \sum_{j=1}^J \mathbb{1}\{D_{i,t}^{1-5} = d_j\}\gamma_j + \epsilon_{i,t},$$
(6)

where $\{D_{i,t}^{1-5} = d_j\}$ is a series of dummy variables equal to 1 if county *i* in year *t* experienced a dose of d_j natural disasters in the past five years, with untreated units as the omitted category. The spatial distribution of the number of natural disasters is shown in panel b) of Figure A1. This treatment specification follows the multi-valued discrete setting in Callaway et al. (2024) and the OLS coefficients $\hat{\gamma} = (\hat{\gamma}_1, ..., \hat{\gamma}_J)$ are estimators of the average level treatment effect over treatment dosages. However, comparison between different dosages requires stronger assumptions than standard parallel trends. Under strong parallel trends, the path of outcomes for lower-dose units must reflect the path of higher dose units had they received the lower-dose. In absence of this condition, the comparison across dose groups can still be interpreted as a causal response. However, it is contaminated by the selection bias (Callaway et al., 2024).

5 RESULTS

5.1 Event Study Results

To assess how student achievement evolves over time, I estimate the event study design from Equation 2, following Sun and Abraham (2021).

Figure 3 shows that there is a negative effect of natural disasters on student achievement from the year a county is hit by a natural disaster (year = 0) until up to five years later. Comparing different student cohorts over time, math achievement in grade three is about 0.025 of a standard deviation lower one year after the natural disaster. Since one standard deviation is approximately three grade levels, students fall around 0.075 grade levels behind due to the natural disaster. The effect is very persistent, such that students in grade three in counties that experienced a natural disaster five years earlier are still 0.03 standard deviations behind. The results look similar for grades four and five,

Figure 3 Event study: Math achievement



(a) Math achievement in grade three

(b) Math achievement in grade four

Note: The figure shows the main results from the event study analysis following Sun and Abraham (2021) for grade three, four, and five. The x-axis represents years relative to the first natural disaster between 2009 and 2018. The y-axis represents the estimate with 95% confidence intervals. The figure shows a decrease in math achievement after the natural disaster.

although the effects are somewhat smaller for grade four and again smaller for grade five. For grade five, only the effect one year after the disaster is statistically significant.

The analysis primarily focuses on grade three. Grades six to eight are not as extensively covered in the SEDA data. Despite these limitations, the results for grades six to eight are presented in Figure A2, which also indicate negative effects in the higher grades. However, these effects are not always significantly different from zero. Grade eight shows a positive point estimate in the year of the disaster, but it is close to zero and not statistically significant, making it difficult to draw any firm conclusions.

The coefficients in the years before the natural disaster in Figure 3 are all not significantly different from zero, reassuring that the parallel trends assumption holds. One issue discussed in the recent difference-in-differences literature is that pre-trend tests can be under-powered, such that one cannot reject the absence of pre-trends, nor can one reject the potential existence of pre-trends that would cause significant bias (Bilinski and Hatfield, 2018; Freyaldenhoven et al., 2019; Kahn-Lang and Lang, 2020; Roth, 2022).⁸ Roth (2022) finds that linear violations of parallel trends, which pre-trend tests detect only 50 percent of the time, can cause biases equal to or greater than the estimated treatment effect. Table A2 suggests that the pre-trend test from the baseline event study would detect a small linear trend of magnitude 0.006 with 50 percent power and a linear trend of 0.010 (0.009 for grade four and five) with 80 percent power. However, the low likelihood ratio of the observed coefficients under the linear trend of 0.010 relative to parallel trends favors parallel trends. For the small linear trend of 0.006 under 50 percent power, the likelihood ratio is still low for grade five but closer to one for grade three and four. I illustrate this linear trend as the red solid line in Figure A3. It is likely that a linear violation did not cause the estimated effect if it cannot fully explain the pattern in the event study. For grade three and five, the linear trend falls outside the confidence intervals. The linear trend would cause a bias of at most 0.013 for grade three in the year after the natural disaster, which is considerably smaller than the estimate. For grade four, the results are less robust and the bias would be larger or similar to the estimated treatment effect for the later post-treatment years if such a linear trend existed. I will conduct further robustness checks following Borusyak et al. (2024) and assess the sensitivity of the results in a number of ways.

⁸Pre-trend tests reverse the roles of a type I and a type II error. A 95 percent confidence interval sets the probability of finding a violation when parallel trends actually hold to 5 percent, but the probability of failing to detect a pre-trend can be much higher (Bilinski and Hatfield, 2018).

Another issue highlighted by Roth (2022) is that samples failing to detect a linear trend in the population means can suffer from selection bias. This bias often increases the bias arising from violations of parallel trends. To illustrate this, I assume that the true population means follow the red solid line in Figure A3. The dotted blue line shows the expected coefficients on average, conditional on not finding a significant pre-trend. However, in this analysis, the blue dotted line is very close to the red line, which suggests that the selection bias does not exacerbate the bias in my baseline event study.

Additionally, I estimate event study results following the imputation method by Borusyak et al. (2024). The imputation method proposed by Borusyak et al. (2024) separates pretrend testing from the estimation of treatment effects, removing the correlation between treatment effects and pre-trend estimators.⁹ This approach avoids the bias introduced by pre-testing, as highlighted by Roth (2022). Figure A4 shows that the pre-trend coefficients, with the exception in grade three and five, are statistically insignificant. The F-test cannot reject the hypothesis that coefficients are jointly equal to zero with a p-value of 0.184 for grade three and 0.706 for grade four, lending further support to the parallel trends assumption. Only for grade five, the F-test weakly rejects the hypothesis that coefficients are jointly equal to zero with a p-value of 0.025. The imputation method generally confirms the negative effects on student achievement from grades three to five. However, the method produces larger standard errors. Interestingly, there is a positive point estimate for grade three, four years after the disaster event. As with Sun and Abraham (2021), the period-four estimate is not significantly different from zero.

In the standard setting of Sun and Abraham (2021) in Figure 3, the last treated cohort serves as the control group. As a sensitivity analysis, one can augment the control group by using counties that were never treated as a control. Figure A5 illustrates that this specification yields virtually identical results.

The main specification with Sun and Abraham (2021) focuses on a single grade level, comparing different cohorts over time. Alternatively, we can pool all grade levels and

 $^{^9\}mathrm{Borusyak}$ et al. (2024) run the pre-trend test with a dynamic TWFE specification on the set of untreated observations only.

add cohort fixed effects to account for variations across cohorts. Panel a) of Figure A6 shows the pooled math score of all grade levels. Panel b) adds cohort fixed effects to the pooled math scores. Both plots appear nearly identical. The results in Figure A6 confirm a significant negative effect in the year of the natural disaster and the year after. In the following years, the point estimates show negative values, gradually approaching zero each year and becoming statistically insignificant. By the fifth year after the natural disaster, the point estimate slightly exceeds zero, but without significant deviation from it.

In summary, student achievement experiences a significant negative decline when a disaster strikes a county, particularly evident in the year of the natural disaster and one year after. These adverse effects persist for up to five years following the natural disaster. Notably, the impact is most pronounced and enduring among grade three students compared to those in grades four and five. This implies that disasters detrimentally affect cohorts, even those not yet enrolled in school at the time of the event. For example, a student showing lower math performance in third grade five years after the disaster would have been of preschool age at the time of the disaster, still three years away from starting elementary school. The negative effects may stem from a potential decline in preschool quality and supply. Part of this impact could also be linked to a reduction in children's health stock, consistent with findings by Fuller (2014), who reports negative impacts on grade three achievement for children whose parents were exposed to hurricanes during pregnancy. Other possible mechanisms include financial difficulties, a decline in infrastructure and school quality, especially if high-quality teachers relocate from the affected areas or if teachers and parents experience significant mental distress. Such distress can reduce parental support and affect the quality of teaching, ultimately impacting student achievement.

Achievement gaps: The persistent negative effects in grade three could affect vulnerable children more strongly than others and thus widen achievement gaps. The existing literature suggests that boys and girls respond differently to family and school





(a) Achievement of girls

(b) Achievement of boys

Note: The figure shows the main results from the event study analysis following Sun and Abraham (2021) for the male-female gap in grade three math achievement. The x-axis represents years relative to the first natural disaster between 2009 and 2018. The y-axis represents the estimate with 95% confidence intervals.

environments. (Autor et al., 2016, 2019; Bertrand and Pan, 2013), with boys often demonstrating more adverse outcomes in test scores and disciplinary issues in response to disruptive (family) events. Descriptively, the male-female gap in mathematics within the SEDA data is negligible. Figure 4 shows that in grade three, boys (panel b) face more negative effects than girls (panel a). However, panel c) indicates that the effect on the male-female gap is not significant, except for one year after the natural disaster. There is also no significant effect on the gender gap in grade four, as shown in Figure A7. However, in grade 5, there is a significant negative effect on the male-female gap, suggesting that boys may experience more pronounced setbacks in learning outcomes as a result of such events.



Figure 5 Event study: Socio-economic gap

Note: The figure shows the main results from the event study analysis following Sun and Abraham (2021) for the socio-economic gap in grade three math achievement. The x-axis represents years relative to the first natural disaster between 2009 and 2018. The y-axis represents the estimate with 95% confidence intervals.

Panel a) of Figure 5 displays the negative effect on economically disadvantaged children in grade three.¹⁰ However, panel b) shows no clear evidence of a widening socio-economic achievement gap post-disaster in grade three, except for the period five years after the natural disaster. For grade four and five, Figure A8 suggests a widening of the socioeconomic gap, indicating that students from economically disadvantaged backgrounds are more adversely affected by the disaster.

¹⁰SEDA does not provide a separate achievement score for children who are not economically disadvantaged.

Figure 6 Heterogeneity by per-pupil spending



Note: The figure shows the main results from the event study analysis following Sun and Abraham (2021) for math achievement by high and low per-pupil spending. The x-axis represents years relative to the first natural disaster between 2009 and 2018. The y-axis represents the estimate with 95% confidence intervals.

Per-pupil spending: In terms of effect heterogeneities, an alternative approach involves exploring different county attributes to investigate whether certain counties demonstrate greater resilience to natural disasters. I split the sample into ex-ante low and high per-pupil spending counties. Counties with per-pupil spending above USD 11,000 in 2009, which was approximately the average, are classified as high-spending counties. All other counties are classified as low-spending counties. Figure 6 shows that high per-pupil spending counties only experience significant negative effects in the year following the natural disaster, then appear to recover more quickly. In contrast, low per-pupil spending counties exhibit negative results until three years after the natural disaster. However, Table A1 indicates that counties with lower per-pupil spending also exhibit lower ex-ante student achievement and a higher proportion of students eligible for free or reduced-price

lunch. Consequently, the observed effect cannot be solely attributed to differences in per-pupil spending. Instead, it underscores the broader disparities in disaster resilience that are closely tied to financial resources, as highlighted by Cutter et al. (2008). The link between financial investment and resilience to natural disasters appears to be a consistent theme across different domains. Auh et al. (2022) find that municipal bonds backed by diversified revenue sources are generally resilient to natural disasters, except for those issued by municipalities in weak financial condition. This exception is attributed to the challenges faced by financially burdened municipalities in diversifying away the shock caused by severe natural disasters.

5.2 Heterogeneity by Disaster Characteristics

The previous section demonstrates that the effects on student achievement are consistently negative over time. Notably, the impact across the post-disaster periods remains similar, justifying their aggregation into a static effect. The specifications from Chapter 4.2 allow for more flexibility to study different channels, such as disaster severity, frequency, and type. Beginning with the baseline estimate for the static TWFE model in Equation 4, the initial column of Table 2 displays the outcomes without weighting. On average, third-grade students in counties experiencing at least one natural disaster within the past five years demonstrate a performance decline of 0.021 standard deviations. In the subsequent column, the preferred specification integrates population weights to address variations in population sizes across US counties. According to this specification, student achievement declines by 0.028 standard deviations. Model 3 incorporates state times school year fixed effects to enable comparisons among counties within states. The stateschool year fixed effects control for unobserved time-varying state effects such as state fiscal shocks. However, this specification may inadvertently absorb some of the treatment effect (Wolfers, 2006). As anticipated, the coefficient is half the size, yet the overall interpretation of the results remains consistent. The static estimates closely align with the average across the five post-treatment periods when employing the event study design in Figure 3.

Table 2Static effect on student achievement

	Math score grade 3				
	Model 1	Model 2	Model 3		
	(1)	(2)	(3)		
Disaster	-0.0206***	-0.0284***	-0.0141**		
	(0.0054)	(0.0080)	(0.0067)		
Population-weighted		\checkmark	\checkmark		
Observations	15,930	$15,\!930$	$15,\!930$		
\mathbb{R}^2	0.76518	0.87385	0.90586		
Within \mathbb{R}^2	0.00153	0.00693	0.00151		
County fixed effects	\checkmark	\checkmark	\checkmark		
School year fixed effects State-school year fixed effects	\checkmark	\checkmark	\checkmark		

Note: The table presents the results from estimating the TWFE model in equation 4 for math achievement in grade three. Model 1 contains only county and school-year fixed effects and no weights. Model 2 contains county and school-year fixed effects and is weighted by the ex-ante county population size. Model 3 contains county and state times school year fixed effects and population weights. Clustered (county) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.11

Severity of disasters: The negative coefficient observed in Table 2 may be driven by severe natural disasters, while smaller-scale disasters may not have a significant impact. To investigate this possibility, I estimate Equation 5 and differentiate between minor and major natural disasters. In the first column of Table 3, I define major disasters as those resulting in more than 25 deaths, following the definition by Boustan et al. (2020). All regressions are population-weighted. Counties in the aftermath of a major natural disaster exhibit a performance decline of 0.038 standard deviations compared to counties unaffected by natural disasters. However, the coefficient on minor natural disasters remains close to the overall effect with a significant of coefficient of -0.027, indicating that also below-average natural disaster cause severe harm.

In column two, I dissect disasters by their costs, employing a threshold of 1 billion dollars in damages as the cutoff point. The point estimate for major disasters is -0.042, but it is not only the 1 billion dollar disasters that impede learning. Minor disasters in this specification decrease student achievement by 0.025 of a standard deviation. Lastly, I

Table 3Heterogeneity by severity of disasters

	Math score grade 3			
	By fatalities	By damage	By duration	
	(1)	(2)	(3)	
Minor disaster	-0.0274^{***}	-0.0251***	-0.0247***	
	(0.0079)	(0.0080)	(0.0077)	
Major disaster	-0.0377***	-0.0423***	-0.0680***	
	(0.0137)	(0.0115)	(0.0131)	
Observations	15,930	15,930	15,930	
\mathbb{R}^2	0.87391	0.87422	0.87516	
Within \mathbb{R}^2	0.00738	0.00980	0.01723	
County fixed effects	\checkmark	\checkmark	\checkmark	
School year fixed effects	\checkmark	\checkmark	\checkmark	

Note: The table presents the results from estimating the TWFE model in equation 5 with county and school year fixed effects for math achievement in grade three. Major disasters are those that caused more than 25 deaths (column 1), more than 1 billion dollar (adjusted) total damage (column 2), or that lasted more than 50 business days (column 3). All models contain population weights. Clustered (county) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.11

redefine major disasters based on the duration of the disaster declaration, with durations exceeding 50 days classified as major disasters. Counties experiencing a disaster that lasted more than 50 days have a -0.068 lower student achievement, but again, also shorter disasters cause significant harm in student achievement with a significant coefficient of -0.025. These results align with Opper et al. (2023), who also find the largest effects from very large disasters, defined as those causing over \$500 per capita property damage. They also show that even disasters exceeding \$100 per capita property damage lead to significant negative impacts on student achievement.¹¹

Multiple disasters: In some counties, multiple natural disasters occurred within the observed time frame. The occurrence of several disasters could intensify their adverse effects, leaving counties with limited time to recover. However, another perspective suggests that counties may develop greater resilience and adaptability to natural disasters over

¹¹Unlike Equation 5, Opper et al. (2023) only consider disasters that occurred within the same year and estimate their effect on the first difference of student achievement.

time, potentially reducing the harm caused by frequent occurrences. I estimate Equation 6, which includes a dummy variable for each count of disasters. The counterfactual are counties that did not experience any disaster in the past five years.

Figure 7 Heterogeneity by number of disasters



Note: The figure shows the results from estimating equation 6 that contains county and school-year fixed effects, population weights and a dummy for each disaster count over the past five years.

Figure 7 demonstrates a clear trend: as the number of natural disasters increases, students in affected counties experience progressively worse academic performance. This suggests that frequent occurrences of natural disasters may leave counties with insufficient time for recovery, contributing to the persistent decline in academic performance.

Disaster type: Certain types of natural disasters may have a more pronounced impact on student achievement than others. As depicted in Figure 8, volcanic activity and landslides cause the greatest harm. However, those are relatively rare events and are thus unlikely to drive the results. Earthquakes, though rare, can cause substantial infrastructure damage and incur high costs. These events are associated with a 0.05 standard deviation decrease in student achievement. More frequent events, such as hurricanes and fires, reduce student achievement by 0.05 standard deviations, while the effects of storms and floods are close to zero and not significant.

Figure 8 Heterogeneity by disaster type



Note: The figure shows the results from estimating equation 6 that contains county and school-year fixed effects, population weights and a dummy for whether a disaster type occurred in the past five years.

5.3 Mechanisms

The results show that disasters detrimentally affect cohorts, even those not yet enrolled in school at the time of the event. Thus, missed days in school cannot be the only factor at play. There are many potential mechanisms that may play a role, including long-lasting infrastructure damage, relocation of higher-performing students or high-ability teachers, financial difficulties of families, and a decline in the health stock of children and caregivers. Some of these mechanisms can be tested directly with the available data. In this section, I discuss them in more detail, focusing on migration patterns, changes in the student composition, and mental health of the adult population.

Migration: The loss of human capital following a natural disaster poses a significant detriment to a region or county, whether it be through the decline in cognitive abilities of the population or through a brain drain. One concern is that some families might have moved away after the disaster event. If families move within counties, this is still reflected

in the student achievement county score. However, the migration of selected families from or into the county might partially account for the decline in student achievement. Overall, around 5% of the population of the county baseline population leaves the county each year and 5% of the county baseline population moves in from another county.

(b) Migration outflow

Figure 9 Event study: Migration



(a) Migration inflow

Note: The figure shows the results from the event study analysis following Sun and Abraham (2021) for migration inflow and migration outflow. The x-axis represents years relative to the first natural disaster between 2010 to 2018. The y-axis represents the estimate with 95 confidence intervals.

Figure 9 shows the event studies for in- and out-migration of the county. Interestingly, there is an increase in the inflow and the outflow. Both variables are highly correlated. However, this increase is very small, with the largest point estimate being 0.003 percentage points. Although the effects on migration are small, the second question is if these migration patterns are selective, because certain groups have better financial resources and job opportunities to choose their place of residence more freely. If migration outflow was positively selected while migration inflow was negatively selected, it could lead to an overestimation of the effect on student achievement. However, Figure A9 demonstrates that the aggregate adjusted gross income (AGI) of the out-migrating population remains unchanged after a disaster event. This suggests that out-migrating families are not more positively selected after a disaster. On the contrary, the increase in aggregate

AGI indicates that the incoming households may be slightly more positively selected compared to those before a disaster. Since high-SES students tend to perform better academically than their low-SES peers, this implies that if there is any impact, it may be an underestimation of the effect of natural disasters on student achievement.

Figure 10 Event study: Enrollment and student composition



Note: The figure shows the results from the event study analysis following Sun and Abraham (2021) for enrollment, share of White students, the share of Black students, and share of students receiving free or reduced lunch in grade three. The x-axis represents years relative to the first natural disaster between 2010 to 2018. The y-axis represents the estimate with 95% confidence intervals.

(a) Enrollment

(b) Black population

Student composition: Next, I test if natural disasters shift the student composition. As depicted in Figure 10, overall enrollment, the share of Black students and the share of White students remain unchanged. This underscores that selective migration or a selective relocation of students is unlikely. There is a significant rise in the number of students receiving free or reduced lunches two to four years after the disaster event. Given the results on aggregate adjusted gross income for both in-migration and out-migration, this shift is unlikely to be driven by migration. Instead, it appears to be more closely related to a financial deterioration of the population following a natural disaster. Although not entirely clear, this suggests that financial distress within families could be a key channel through which natural disasters negatively impact student achievement in the long term.

Figure 11 Event study: Adult mental health



Note: The figure shows the results from the event study analysis following Sun and Abraham (2021) for mental health. The x-axis represents years relative to the first natural disaster between 2009 and 2018. The y-axis represents the estimate with 95% confidence intervals.

Adult mental health: Long-run effects of early life exposure to natural disasters may be linked to a reduction in the mental and physical health stock (Karbownik and Wray, 2019). Unfortunately, I cannot directly test the effect on children's mental health due to data limitations. However, mental health of adults could serve as a critical channel. Adults experiencing distress, anxiety, or depression following a natural disaster may struggle to provide the necessary support and stability for children, at home, at school, and within the broader community, which could have long-term negative consequences for children's cognitive and socio-emotional development. Moreover, adults grappling with their own mental health may find it challenging to model positive coping strategies or maintain consistent communication with schools, hindering the implementation of effective interventions to support student well-being and academic success. However, Figure 11 shows no effect on the average number of mentally unhealthy days.

6 DISCUSSION AND CONCLUSION

This research unveils a persistent negative impact on student achievement following natural disasters, with students experiencing setbacks for up to five years post-event. These dynamics signal a depletion in the human capital reservoir of these regions, resulting in long-term economic damage (Gust et al., 2024). Leveraging data from the SEDA database spanning 2009 to 2018 and FEMA disaster declarations, I apply state-of-the-art difference-in-differences techniques by Sun and Abraham (2021) to produce event study estimates that account for heterogeneous treatment effects in this staggered framework. Boys and low-SES children experience somewhat stronger effects. Alternative differencein-differences methods by Borusyak et al. (2024), along with an alternative control group specification, confirm the detrimental effects on students. These findings are particularly concerning given the increasing frequency and severity of natural disasters. Except for the Covid-19 pandemic, natural disasters are the most frequent reason for prolonged unplanned school closures in the US.

However, school closures represent just one facet through which natural disasters impact human capital. The observed negative effects on cohorts not yet enrolled during disasters suggest a depletion of health resources among younger children, damage of school buildings and facilities, lower financial resources of families and communities, and potential shifts in student and teacher composition of affected areas as contributing factors. While not all of these channels are directly testable, the increase in the share of students receiving free or reduced lunch suggests financial distress in families may be a key mechanism.

Significant divergences emerge when considering pre-disaster investment levels, with counties with higher per-pupil expenditure demonstrating swifter recovery compared to their lower-spending counterparts. However, countries with lower pre-disaster investment levels also have an ex-ante higher share of children receiving reduced or free lunch. Nevertheless, this finding holds important policy implications, indicating that augmenting perpupil spending can enhance community resilience against human capital erosion caused by natural disasters. Future inquiries should delve into the specific types of investments pivotal in shielding against disaster-induced damage. Unfortunately, empirical evidence by Deryugina (2017) suggests that governments typically curtail education spending in the aftermath of natural disasters, which could worsen the situation in future disasters.

The results in this paper show that having multiple disasters in a row can cause more damage. Furthermore, the severity and nature of disasters play pivotal roles, with major events such as hurricanes exerting the most pronounced adverse impact on student performance. These findings underscore the imperative for proactive disaster preparedness and response measures, alongside targeted interventions.

REFERENCES

- Arouri, Mohamed, Cuong Nguyen and Adel Ben Youssef (2015), 'Natural disasters, household welfare, and resilience: Evidence from rural Vietnam', World development 70, 59–77.
- 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.
- Auh, Jun Kyung, Jaewon Choi, Tatyana Deryugina and Tim Park (2022), Natural disasters and municipal bonds, Technical report, National Bureau of Economic Research.
- Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth and Melanie Wasserman (2016), 'School quality and the gender gap in educational achievement', American Economic Review 106(5), 289–295.
- Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth and Melanie Wasserman (2019), 'Family disadvantage and the gender gap in behavioral and educational outcomes', American Economic Journal: Applied Economics 11(3), 338–381.
- Bertrand, Marianne and Jessica Pan (2013), 'The trouble with boys: Social influences and the gender gap in disruptive behavior', *American economic journal: applied economics* 5(1), 32–64.
- Bilinski, Alyssa and Laura A Hatfield (2018), 'Seeking evidence of absence: Reconsidering tests of model assumptions', arXiv preprint arXiv:1805.03273.
- Black, Sandra E., Paul J. Devereux and Kjell G. Salvanes (2007), 'From the cradle to the labor market? The effect of birth weight on adult outcomes', *The Quarterly Journal of Economics* 122(1), 409–439.
- Blanden, Jo, Matthias Doepke and Jan Stuhler (2023), Educational inequality, *in* 'Handbook of the Economics of Education', Vol. 6, Elsevier, pp. 405–497.
- Block, Michelle L. and Lilian Calderón-Garcidueñas (2009), 'Air pollution: Mechanisms of neuroinflammation and CNS disease', *Trends in neurosciences* **32**(9), 506–516.
- Borusyak, Kirill, Xavier Jaravel and Jann Spiess (2024), 'Revisiting event-study designs: robust and efficient estimation', *Review of Economic Studies* p. rdae007.
- Boustan, Leah Platt, Matthew E. Kahn, Paul W. Rhode and Maria Lucia Yanguas (2020), 'The effect of natural disasters on economic activity in US counties: A century of data', *Journal of Urban Economics* 118, 103257.
- Callaway, Brantly, Andrew Goodman-Bacon and Pedro HC Sant'Anna (2024), Differencein-differences with a continuous treatment, Technical report, National Bureau of Economic Research.
- Callaway, Brantly and Pedro HC Sant'Anna (2021), 'Difference-in-differences with multiple time periods', *Journal of econometrics* **225**(2), 200–230.

- Currie, Janet, Eric A. Hanushek, E. Megan Kahn, Matthew Neidell and Steven G. Rivkin (2009), 'Does pollution increase school absences?', *The Review of Economics and Statistics* 91(4), 682–694.
- Currie, Janet and Maya Rossin-Slater (2013), 'Weathering the storm: Hurricanes and birth outcomes', *Journal of health economics* **32**(3), 487–503.
- Cutter, Susan L., Lindsey Barnes, Melissa Berry, Christopher Burton, Elijah Evans, Eric Tate and Jennifer Webb (2008), 'A place-based model for understanding community resilience to natural disasters', *Global environmental change* 18(4), 598–606.
- De Chaisemartin, Clément and Xavier d'Haultfoeuille (2020), 'Two-way fixed effects estimators with heterogeneous treatment effects', American Economic Review **110**(9), 2964–96.
- Deryugina, Tatyana (2017), 'The fiscal cost of hurricanes: Disaster aid versus social insurance', American Economic Journal: Economic Policy 9(3), 168–198.
- Deuchert, Eva and Christina Felfe (2015), 'The tempest: Short-and long-term consequences of a natural disaster for children's development', *European Economic Review* **80**, 280–294.
- Di Pietro, Giorgio (2018), 'The academic impact of natural disasters: Evidence from L'Aquila earthquake', *Education Economics* **26**(1), 62–77.
- Engzell, Per, Arun Frey and Mark D Verhagen (2021), 'Learning loss due to school closures during the COVID-19 pandemic', *Proceedings of the National Academy of Sciences* 118(17), e2022376118.
- Fahle, Erin M, Belen Chavez, Demetra Kalogrides, Benjamin R Shear, Sean F Reardon and Andrew D Ho (2021), 'Stanford education data archive (version 4.1)', *Retrieved* from http://purl.stanford.edu/db586ns4974. 1.
- Freyaldenhoven, Simon, Christian Hansen and Jesse M Shapiro (2019), 'Pre-event trends in the panel event-study design', *American Economic Review* **109**(9), 3307–3338.
- Fuller, Sarah C. (2014), 'The effect of prenatal natural disaster exposure on school outcomes', *Demography* 51(4), 1501–1525.
- Gardner, John (2022), 'Two-stage differences in differences', Working Paper.
- Goodman-Bacon, Andrew (2021), 'Difference-in-differences with variation in treatment timing', *Journal of Econometrics* **225**(2), 254–277.
- Goodman, Joshua (2014), Flaking out: Student absences and snow days as disruptions of instructional time, Technical report, National Bureau of Economic Research.
- Gust, Sarah, Eric A. Hanushek and Ludger Woessmann (2024), 'Global universal basic skills: Current deficits and implications for world development', *Journal of Development Economics* 166, 103205.

- Jahan, Ferdous A, Nicole Zviedrite, Hongjiang Gao, Faruque Ahmed and Amra Uzicanin (2022), 'Causes, characteristics, and patterns of prolonged unplanned school closures prior to the COVID-19 pandemic—United States, 2011–2019', *PLoS One* 17(7), e0272088.
- Jaume, David and Alexander Willén (2019), 'The Long-Run Effects of Teacher Strikes: Evidence from Argentina', *Journal of Labor Economics* **37**(4), 1097–1139.
- Jerch, Rhiannon, Matthew E. Kahn and Gary C. Lin (2023), 'Local public finance dynamics and hurricane shocks', *Journal of Urban Economics* **134**, 103516.
- Kahn-Lang, Ariella and Kevin Lang (2020), 'The promise and pitfalls of differences-indifferences: Reflections on 16 and pregnant and other applications', *Journal of Business* & Economic Statistics 38(3), 613–620.
- Karbownik, Krzysztof and Anthony Wray (2019), 'Long-Run Consequences of Exposure to Natural Disasters', Journal of Labor Economics 37(3), 949–1007.
- Klauber, Hannah, Felix Holub, Nicolas Koch, Nico Pestel, Nolan Ritter and Alexander Rohlf (2024), 'Killing prescriptions softly: Low emission zones and child health from birth to school', American Economic Journal: Economic Policy 16(2), 220–248.
- Lavy, Victor (2015), 'Do differences in schools' instruction time explain international achievement gaps? Evidence from developed and developing countries', *The Economic Journal* 125(588), F397–F424.
- Lazear, Edward P. (2001), 'Educational production', The Quarterly Journal of Economics 116(3), 777–803.
- Marcotte, Dave E. (2007), 'Schooling and test scores: A mother-natural experiment', Economics of Education Review **26**(5), 629–640.
- Masozera, Michel, Melissa Bailey and Charles Kerchner (2007), 'Distribution of impacts of natural disasters across income groups: A case study of New Orleans', *Ecological economics* **63**(2-3), 299–306.
- Nguyen, Cuong Viet and Nguyet Minh Pham (2018), 'The impact of natural disasters on children's education: Comparative evidence from ethiopia, india, peru, and vietnam', *Review of Development Economics* **22**(4), 1561–1589.
- Opper, Isaac M., R. Jisung Park and Lucas Husted (2023), 'The effect of natural disasters on human capital in the United States', *Nature human behaviour* pp. 1–12.
- Özek, Umut (2023), 'Examining the educational spillover effects of severe natural disasters: The case of hurricane maria', *Journal of Human Resources* 58(2), 421–451.
- 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.
- Pleninger, Regina (2022), 'Impact of natural disasters on the income distribution', World Development 157, 105936.

- Reardon, Sean F., John P. Papay, Tara Kilbride, Katherine O. Strunk, Joshua Cowen, Lily An and Kate Donohue (2019), 'Can Repeated Aggregate Cross-Sectional Data Be Used to Measure Average Student Learning Rates? A Validation Study of Learning Rate Measures in the Stanford Education Data Archive. CEPA Working Paper No. 19-08.', Stanford Center for Education Policy Analysis.
- Remington, Patrick L, Bridget B Catlin and Keith P Gennuso (2015), 'The County Health Rankings: Rationale and methods', *Population Health Metrics* **13**(1), 11.
- Roth, Jonathan (2022), 'Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends', American Economic Review: Insights 4(3), 305–322.
- 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), 109–35.
- Smith, Adam B. (2022), '2021 U.S. billion-dollar weather and climate disasters in historical context', *Technical Report*.
- Sun, Liyang and Sarah Abraham (2021), 'Estimating dynamic treatment effects in event studies with heterogeneous treatment effects', *Journal of Econometrics* **225**(2), 175–199.
- Wedel, Katharina (2021), 'Instruction time and student achievement: The moderating role of teacher qualifications', *Economics of Education Review* **85**, 102183.
- Werner, Katharina and Ludger Woessmann (2023), 'The legacy of covid-19 in education', Economic Policy **38**(115), 609–668.
- Woessmann, Ludger (2016), 'The economic case for education', *Education Economics* **24**(1), 3–32.
- Wolfers, Justin (2006), 'Did unilateral divorce laws raise divorce rates? A reconciliation and new results', *American Economic Review* **96**(5), 1802–1820.
- Wu, Ge (2022), Do Wildfilres Harm Student Learning?, Technical report, University of Tennessee, Department of Economics.

A APPENDIX A: ADDITIONAL TABLES AND FIGURES

Figure A1

Map: First year of disaster and number of disasters

(a) First disaster event



(b) Average number of disasters in past five years



Note: The figure shows the first year of a FEMA disaster declaration between 2009 and 2018 for the Sun and Abraham (2021) estimations and the average number of FEMA disaster declarations in the past five years.

Figure A2 Event study: Math achievement in higher grades



Note: The figure shows the results from the event study analysis following Sun and Abraham (2021) for grade six, seven, and eight. The x-axis represents years relative to the first natural disaster between 2009 and 2018. The y-axis represents the estimate with 95 confidence intervals. The figure shows a decrease in math achievement after the natural disaster.





Note: The plots show the pre-trend diagnostics by Roth (2022) for the Sun and Abraham (2021) estimates. The red line is the a linear violation of parallel trends that a pre-trend test would detect with 50 percent power. The dashed blue the expected coefficients conditional on not finding a significant pre-trend if the true population means were the hypothesized red line.

Figure A4 Event Study: Math achievement with imputation method



Note: Event study analysis following the imputation method by Borusyak et al. (2024) for grade three, four, and five. The x-axis represents years relative to the first natural disaster. The y-axis represents the estimate with 95% confidence intervals.





Note: The figure shows the results from the event study analysis following Sun and Abraham (2021) for grade three, four, and five with never-treated counties as control group. The x-axis represents years relative to the first natural disaster between 2009 and 2018. The y-axis represents the estimate with 95% confidence intervals. The figure shows a decrease in math achievement after the natural disaster.

Figure A6 Event study: Math achievement with pooled grades and cohort fixed effects



(a) Math achievement pooled

Figure A7 Event study: Male-female achievement gap in grade four and five



Note: The figure shows the main results from the event study analysis following Sun and Abraham (2021) for the male-female gap in grade four and five math achievement. The x-axis represents years relative to the first natural disaster between 2009 and 2018. The y-axis represents the estimate with 95% confidence intervals.

⁽b) Math achievement with cohort fixed effects

Note: The figure shows the results from the event study analysis following Sun and Abraham (2021) for grade three to eight a) pooled and b) including cohort fixed effects. The x-axis represents years relative to the first natural disaster between 2009 and 2018. The y-axis represents the estimate with 95% confidence intervals. The figure shows a decrease in math achievement after the natural disaster.

Figure A8

Event study: Socio-economic achievement gap in grade four and five



Note: The figure shows the main results from the event study analysis following Sun and Abraham (2021) for the socio-economic gap in grade four and five math achievement. The x-axis represents years relative to the first natural disaster between 2009 and 2018. The y-axis represents the estimate with 95% confidence intervals.

Figure A9 Event study: Aggregate adjusted gross income



Note: The figure shows the main results from the event study analysis following Sun and Abraham (2021) for the aggregate adjusted gross income. The x-axis represents years relative to the first natural disaster between 2009 and 2018. The y-axis represents the estimate with 95% confidence intervals.

(a) Socio-economic gap grade four (b) Socio-economic gap grade five

	High (N=659)		Low (N=928)			
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	Std. Error
Math grade 3	-0.024	0.308	-0.084	0.272	-0.060	0.015
Percent Black students	0.133	0.211	0.144	0.209	0.010	0.011
Percent White students	0.681	0.273	0.733	0.251	0.052	0.013
$Percent \ free/reduced \ lunch$	0.511	0.177	0.564	0.151	0.053	0.008
Log enrollment	6.069	1.427	5.934	1.037	-0.135	0.065

Table A1Balancing table: High and low per-pupil spending

Note: The table shows the mean and the standard deviation by high and low per-pupil spending, the difference in means, and the standard error.

Table A2Pre-trend test results

Metric	Grade 3	Grade 4	Grade 5
Power	50	50	50
Hypothesized trend	0.006	0.006	0.006
Bayes factor	0.585	0.586	0.235
Likelihood ratio	1.038	1.105	0.526
Power	80	80	80
Hypothesized trend	0.010	0.009	0.009
Bayes factor	0.234	0.234	0.235
Likelihood ratio	0.336	0.372	0.526

Note: Summary statistics of power, slope and bias calculation (Roth, 2022). The table shows the probability that we would find a significant pre-trend (set to 50% or 80%), the slope of the differential trend that we would be able to detect with that power, the ratio of the probability of "passing" the pre-test under the hypothesized trend relative to under parallel trends (a smaller Bayes factor favors parallel trends over the hypothesized trend when the pre-trend is insignificant), the likelihood ratio of the observed coefficients under the hypothesized trend relative to under parallel trends.