

COVID School Closures and Their Impact on Students*

Cristina Burca, Yan Mezhiborsky, Sakura Noskor

February 12, 2024

Abstract

The COVID-19 pandemic necessitated widespread institution closures, including schools, for a duration of 2-3 years. This study examines the impact of these closures on students and educational outcomes, specifically in the United States, initially focusing on global trends in school closures during the pandemic. Moreover, we analyze and compare affected standardized test scores from before, during and after the pandemic as well as the enrolment rates during the pandemic, and find that school closures have effected enrollment rates greatly and so has the test results of the standerdized test. This is an important research as this research directly shows what negatively impacts children's performance in school.

Table of contents

Introduction

2

*Code and some data from this paper are available at: [github repo](#) and the [doi](#)

Data	4
Data Source and Collection	4
Data Cleaning / Methodology	4
Measurment	7
Results	8
School Closures	8
Standardized Tests vs COVID	10
Enrollment rate in the US	13
Discussion	18
Categorization of Continents	18
Change in Test Scores of 2022	18
Change in enrollment rate	19
Kindergarten vs Total Enrollment	19
Total Enrollment in 2020 vs Total Enrollment in 2019	19
Total Enrollment for in-person learning vs virtual	20
Finishing Thoughts	20
Ethics and Bias	21
Weakness and Limitations	22
Translating code	22
Lack of Data description	23
Solution to Limitation	23
Furthur research	24
Appendix	24
Reference	26

Introduction

This paper will discuss several ways that the Covid-19 pandemic may have affected elementary school students and education as a whole. We

will be looking at data produced from all around the world, and majorly be examining data collected from the United States of America. The original paper we have derived our data from has not included data from the year 2022, and thus our paper extended the research for the year 2022 and analyzes aspects of the original paper, inspecting slightly different variables. Our aim is to consider the after affects of Covid-19, and analyze data of the consequences of the restrictions on students learning virtually for an extended period of time. Many claim that social interaction is essential for a child's well being, and we aim to support or refute this statement by examining similar correlations as the original paper, and comparing them with the findings from the year of 2022, when schools were once again open.

First, we will be investigating school closures in the five major continents and some deductions we can make from them, as well as comparing the number of school closures in between countries. Next, we will look at test scores for the USA in Mathematics as well as ELA(English Learning Arts), and how they changed from year to year, before, during, and after the pandemic to infer on the trend of test scores over the years. Lastly, the outcomes of this paper are that there is an overall increase in the percent of enrollment as more in person schools are opened in the educational districts of the United States. The concept of educational districts will be described later in the data section for the user's aid but in this context it is not the most important.

From our analysis, we were able to deduce the fact that students in the United States were required to abide by institutional lockdowns and virtual learning for a longer period compared to other countries, after stay-at-home restrictions were lifted. Average grades of ELA and Math test scores from U.S. students have decreased during 2021, where virtual learning was at its peak. Furthermore, tests before and after the pandemic were scored higher, with Math scores from 2022 spiking after they dropped in 2021. The decline in test scores may likely be due to the lack of in-person teaching, social interaction, and excessive screen exposure. Also looking at the enrollment rate we see that students much prefer the in-person learning

methods as enrollment rates increase with more in-person delivery modes in school. We also see this when the enrollment rate drops with more schools going virtual through out the districts.

Studies of this variety are important, as they explore and optimize study methods and environments most suitable for children’s learning and development. Studies help evaluate the extent to which the pandemic has affected students and their learning outcomes, as well as provide insight into the effectiveness of various teaching mediums on students of different ages. Moreover, by examining changes in the academic performance and well-being of students over years, the impact of social isolation on various aspects of development can be researched.

Data

Data Source and Collection

We use R Core Team (2022) to make this paper as well as the graphs and topic were taken inspiration from Jack and Oster (2023). Various helpful packages were used in order to clean, sort and graph this paper in a way such that the reader will not have difficulty undertsnading neither the topic nor the data sets of this paper. The packages are, Wickham (2016), Wickham et al. (2019), Wickham et al. (2023), Wickham, Hester, and Bryan (2023), Xie (2014), Firke (2023), Zhu (2021), Wickham, Vaughan, and Girlich (2024), Wickham and Miller (2021), Hyndman and O’Hara-Wild (2021), Arel-Bundock, Enevoldsen, and Yetman (2018).

Data Cleaning / Methodology

The first set of graphs introduced in this paper analyze the number of school closures by continent and by country due to the Covid-19 stay-at-home restrictions that took effect from the beginning of 2020 to the

beginning of 2022, varying by country. The variables considered from the dataset are Date, Year, Country, School.Closing, and Continent. To explain the variables, Date is written in the format YYYY-MM-DD, Year was created to assist with data manipulating, Country lists the Country name, School.Closing lists an integer value of 0 if countries's institutions were open on the listed date, and 1 if they were closed. The data is organized in sequential order from January 1st, 2020 to December 31st, 2022. It was cleaned by firstly selecting the variables of interest and creating the Year column derived from Date to allow for comparison by year. The package Arel-Bundock, Enevoldsen, and Yetman (2018) was used to create a new column that categorizes the continent the country is in. Data from the original paper was organized and presented under a different list of continents.

The third set of graphs shown in this paper is the enrollment rates in the different educational districts of the US. The original data provided was from US department of Education through NCES (National Institute for Education Statistics). The variables included were year, leaid, fips, grade, race, sex, and enrollment. Starting by defining the variables, leaid stands for local education agency id number. This allows the NCES to identify the districts with their unique id code. Next is fips, which stands for federal information processing standards. This number is a rate/level given by the US government to a specific district which essentially rates the technological security of each district (Editor (n.d.)). The grade essentially says the grade the students are in from 0-12 to 99 where 0 is kindergarten and 99 is representing the total enrollment for all grades. The data set was also made through another raw data set from COVID School Data Hub (n.d.) which contains different abbreviations and id numbers corresponding to each district, including leaid, and they also have the share of in-person, virtual and hybrid schools open in that district. The share of in-person school represents the amount of schools that were in_person in that district. Similar explanations for virtual and hybrid shares.

The process of cleaning this raw data was simple. First, we look at the NCES data. Since the paper only looks at enrollment rates for all grades

and kindergartners, a filter is applied for grade to be 0 or 99 with another filter of race to be 99 as 99 represents all. Then another filter is applied where, the data only contains fips that has a value less than 56 as there are only 56 levels in fips. Then the entire data is transformed into a wide format where the new following new columns appear: enrollment for males, enrollment for females, enrollment for all (Not the real names). There are other columns like enrollment9 that gets created but it is dropped as there are no values contained within it. Then the values of enrollment for all columns are filtered such that if they contain an invalid value or such as not available or 0, then the entire row is deleted. After that, the entire data is again reshaped such that we get rid of the grades columns and have new columns as such; enrollment for males in total, enrollment for females in total, enrollment for males in kindergarten, enrollment for females in kindergarten, enrollment for both gender in total, enrollment for both gender in kindergarten. So currently have the rows listed before, the year, and the leaid. Then the current data is merged with the other raw data provided by COVID School Data Hub(n.d.). They were merged based on the leaid and any access columns such as NCES id or state id was deleted as it was of no use to the paper. Then the computations begin. In order to find the rate of enrollment, the formula was to subtract the previous value from the current and divide by the previous value. So for example to find the enrollment rate for males that attend in-person, the `lag()` of enrollment for males in total was taken and subtracted to it self and stored in a variable called `changeTotal`. Then the `changeTotal` variable was divided again by the lag of “enrollment for males in total” column. This was repeated for all the columns that had the enrollment values for each subcategory. This value was then binned according to the share in-person column or the share virtual column based on the graph that was produced. However this was not enough as there were many points, to the point where the analysis of that data set was not worth anything and hence, per guidance to the original paper, a weighted mean was taken. An ordinary average was not taken as this some states would have higher in-person rates then others which would mean there are outliers in this case. So an ordinary average function would have lost some of the important

trends we see when we take a weighted average. Then using a simple algorithm, we found the line of best fit helping the reader see the trends better.

Measurement

The first set of graphs introduced in this paper analyze the number of school closures per year from 2020 to 2022. The data originally included a figure displaying the average number of days with School Closed, by Region, from January 2020 to December 2021, with the Regions of Sub-Saharan Africa, Europe and Central Asia, East Asia and Pacific, Middle East and North Africa, Latin America and Caribbean, South Asia, and North America. This paper instead analyzed the number of school closures per year by continent, including the year 2022. Moreover, since this paper focuses on data from the United States, the second graph examines data from the United States in comparison with other various countries.

The second set of this paper's graphs focused on test scores of students in the United States of America, specifically their score changes in ELA(English Language Arts) and Mathematics. The data this was compared to was the percentage of in person attendance as well as the geographic locale of the students.

As mentioned before the third claim that this paper looks at is the enrollment rates with respect to the share of in-person learning or virtual learning. For this category there are 2 measurements. The data originally come with the number of enrollments in each lead for each year however, the data is cleaned up so the reader may see the rates related to it. The x-axis show the share of in-person or virtual schools in the district. The share of in-person was only changed in terms of how this paper binned the data; the share value was from 0 to 100% and it was cut into 11 equal pieces. The measurement for the x-axis is the same as it is on the raw datasheet, which is percentage representing share of districts that were either in-person or virtual, values from 0 to 100%. The measurement for

the y axis is the enrollment rate. This has been changed from the raw data set as originally it recorded the number of students enrolled for each grade by lead for a certain year. As mentioned in the methodology, the original value was changed into a rate then a percentage to better show the reader what the actual changes were. If this step was not done then comparing the declines and the inclines of enrollment rate would have been hard due to the large numbers. Anyhow, the measurement for the y-axis is the enrollment rate in percentage and it doesn't have a limit.

Results

School Closures

As global stay-at-home orders were enforced due to Covid-19 safety mandates in the beginning of 2020, the number of elementary schools regulated to transition to virtual learning increased rapidly. In this section, we will explore general trends of the total school closures across the world from 2020 to 2022, categorized by continent. Furthermore, we will investigate school closures trends over the years of the United States in comparison with other countries.

In Figure 1, we graph the total number of schools closing due to Covid-19 safety mandates and alternated to remote learning by continent, for years 2020 to 2022. Note that the data for South America and North America are combined under the category 'Americas'. We notice some evident features; Asia has the overall largest number of school closures, this can attribute to a larger population compared to other continents, or a greater number of schools relative to the population. All continents experienced a correspondent decline in the number of school closures each year, apart from Africa. Although this may be compelling to investigate further, this is not relevant to this paper. Let us explore the overall decline in school closings between the years 2020 and 2021 in Table 1.

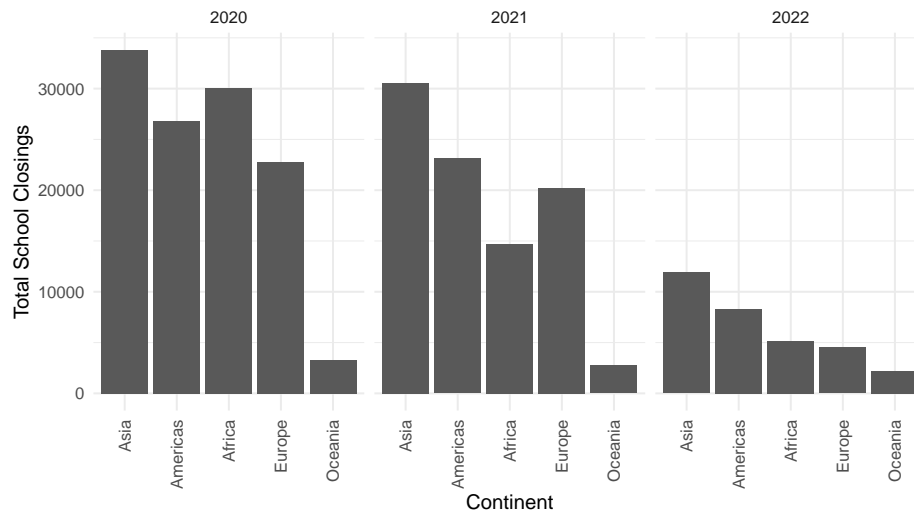


Figure 1: Number of school closures per Contient per year

Table 1: Change in school closures percentage by Continent from 2020 to 2021

Year	Continent	# of Closings	Percent
2021	Asia	30532	-9.71
2021	Americas	23119	-13.81
2021	Africa	14691	-51.13
2021	Europe	20187	-11.41
2021	Oceania	2742	-16.56

The data across the continents, with the exclusion of Africa, exhibited a similar decrease of about 13%. This could a result of the international consensus of implementation of laws in respect social and institutional closures. We now analyze the data pertaining to the United States in Figure 2.

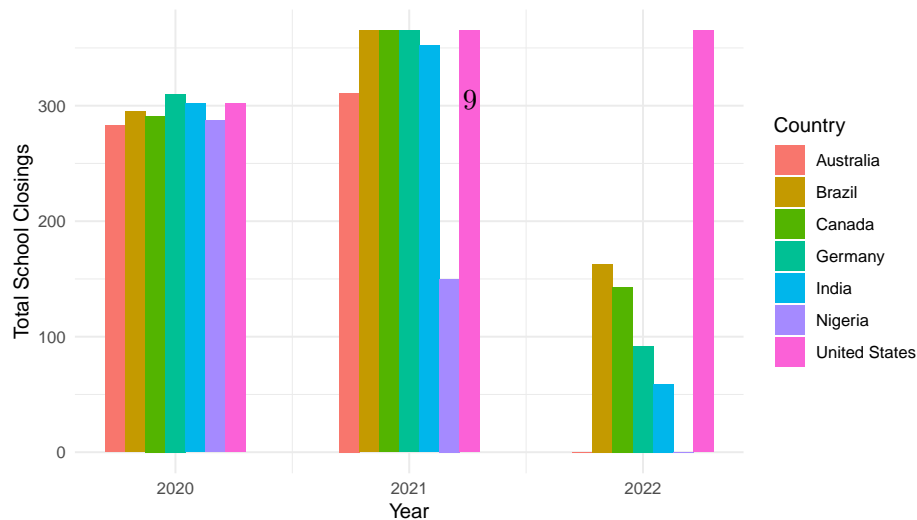


Figure 2: Number of school closures per country per year

I have chosen to show the number of school closures per year with a list of countries for comparison with the United States. Australia, Brazil, Germany, India, and Turkey were chosen due to their standing as some of the most populated cities representing each continent, and Canada for standard of comparison. Evidently, we observe that the United States experienced national school closures for the entirety of 2021, 2022, and majority of 2020. This is in comparison to the alternate countries, which collectively exhibited a significant decrease in school closures in 2022, when Covid-19 restrictions were officially ceased. Notably, students in the United States were compelled to learn in a remote environment for a much longer term than other countries

Standardized Tests vs COVID

Here we will look at the change in percentage of test score of students between the years of 2017 and 2022. Covid-19 had a major impact on student test scores, whether it be because of the quality of remote learning or the resources the students may have or not have had while learning remotely. the first of the two figures is a near identical replication of the the on in the original paper we are replicating while the second figure is gives us some additional information to analyze that wasn't examined as thoroughly in the original paper.

The first observation that can be made is the percent change in test score between Spring 2019 and Spring 2021 where the students who who attended class less than 25% of the time had an average of 15% less in their Math tests and approximately 9% in their ELA tests. Although it can be seen that the more the students attended in person the less their test scores would decrease, never the less the students still saw a decrease in their test scores between the two years even if they had gone to school in person. One thing the original paper failed to include although it was in their data set, was the change in test scores of Spring 2022. While the test

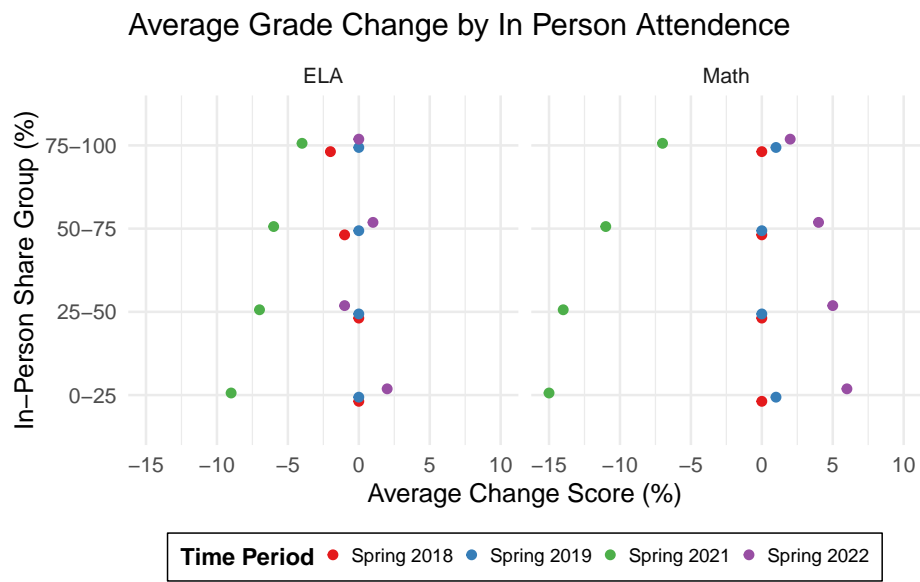


Figure 3: Average Grade Change by In Person Attendance

scores do increase in Math and ELA, ELA less so than Math, they do not increase as much as they fell in the previous year.

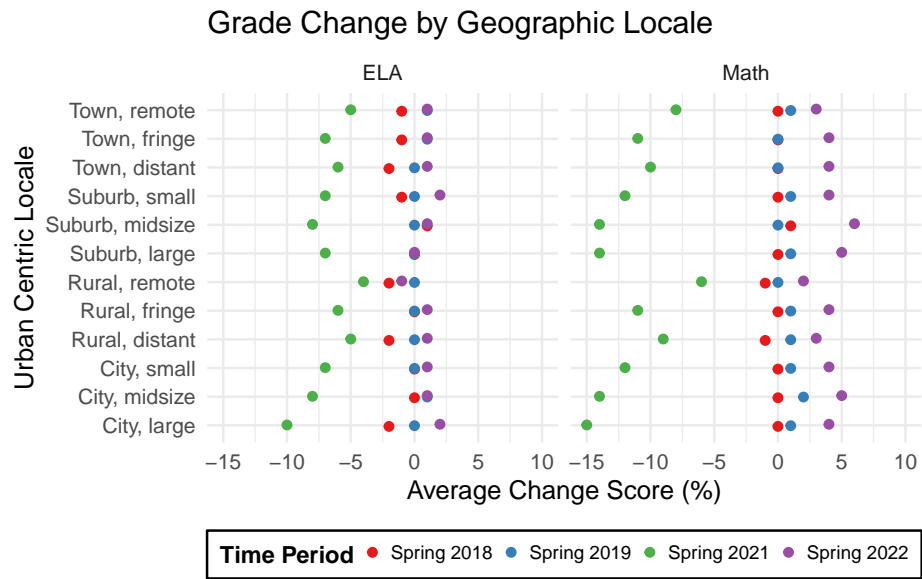


Figure 4: Average Grade Change by In Person Attendance

Here we will look at the Geographic Locales of students, and how their grade have changed based on what type of municipality they live in. We can observe that Cities and suburbs of all sizes both in Math and ELA Scores. This may be due to a variety of reasons among them the Covid-19 in each of the specific localities. As cities are densely populated they are more likely to have had stricter lock down rules during the outbreak of Covid-19. Furthermore, suburbs are usually near cities and would likely have similar lock down rules to the cities they are near.

Enrollment rate in the US

It is no secret that COVID-19 affected the school enrollment rates, when the students were forced to learn in a virtual environment. In this section the paper will uncover any trends in enrollment rates with relation to the number of in-person learning offered in each district. For all of the graphs below the y axis will be from -0.12 to 0.05 to ensure reader us able to comapre the trends right beside each other.

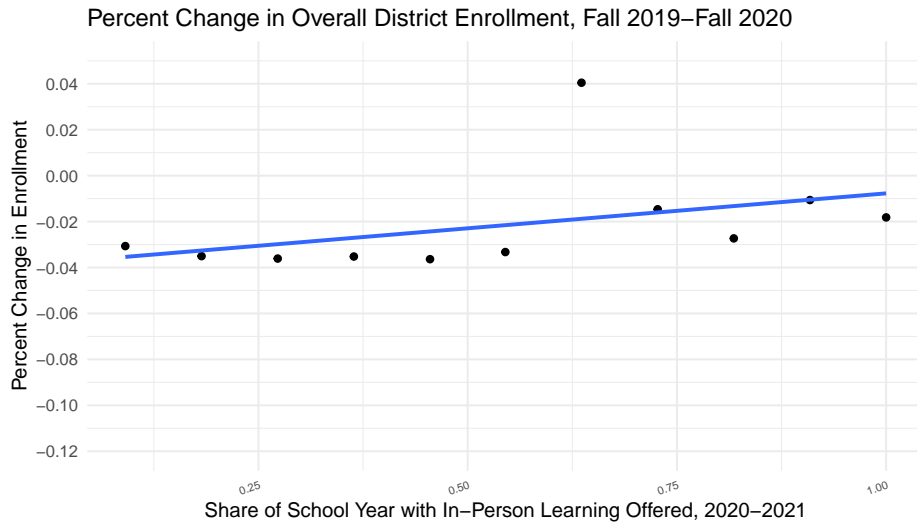


Figure 5: The weighted mean of the percent of overall district enrollment was calculated and graphed. The weights were decided on the bins which were cut in eleven sections based on share of in person school overall.

First, we take a look at the overall enrollment through out the districts Figure 5. We can see that most values are in the negatives with one value in the positives in the [0.545, 0.6360] bin. We however see an upward trend signifying that we have a a positive change in the percent of rate of

change. This means that the enrollment rate increases with more shares of the district going in person. Diving in deeper to look at the values (provided in the Appendix).The graph shows that for 0% to 9% of in-person schooling, we have a decrease of 3% in the enrollment rate. For 9% to 55% share or in person schooling, the rate of enrollment is around -35% meaning there is decrease of 35% and the rate goes to 4% increase in enrollment rate for 55% to 64% of in person schooling in the district. Then the value drops to 1% decrease in enrollment rate when the share of in person offered is 73% to 64% approximately. Then it jumps to a 2% decrease in enrollment then it levels out to 1% decrease in enrollment rate for the last 73% to 100% share in in-person learning. The overall trend is increasing but there are brief moments of decreasing enrollment rates.

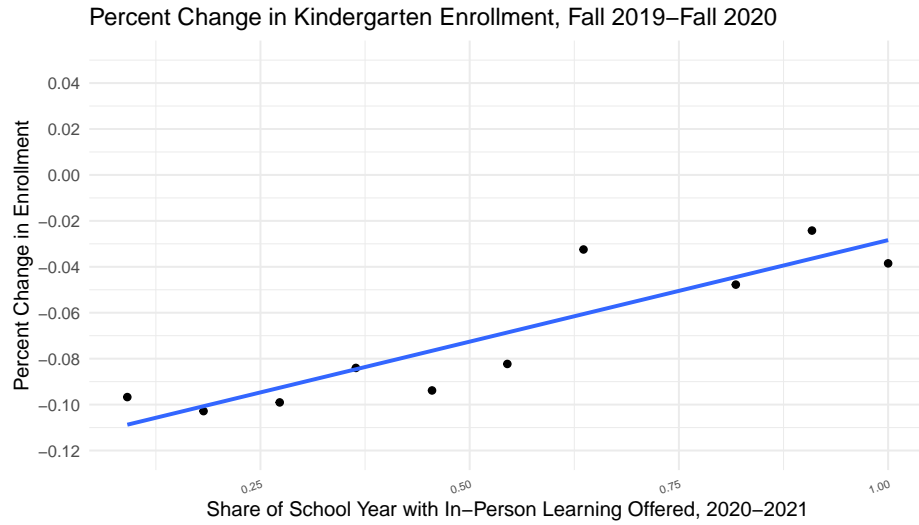


Figure 6: The weighted mean of the percent of kindergarden enrollment in all districts was calculated and graphed. The weights were decided on the bins which were cut in eleven sections based on share of in person school overall.

In comparison we look at the kindergarden enrollment in the same year

Figure 6. Looking at the values individually, the graph shows that for the initial shares 9%, 18% and 27%, we get an enrollment rate of kindergartners in the US is -9%, -10% and 10% respectively. They all show decreasing enrollment rate. Then for the 36% of in-person schooling we get a 8.4% enrollment rate and there is a decrease of 9.4% and 8.2% in enrollment rate when there is 45% and 55% share of in person school respectively, for kindergartens. Then the enrollment rates increase to -3.2% and 7% for the proceeding 64% and 73% of in person school share. Afterwards there is a decrease 4.8% in enrollment rates when 82% of schools are in person. Then there is another decrease to -2.4% in enrollment rate when there is 91% share of in person rates. And lastly for 91% to 100% of the in person shares the enrollment rates are -3.9%. This means that although the enrollment rates are negative for the most part there is an increasing trend. This means that the enrollment rates were increasing as more kindergartners were attending in person. Also compared to Figure 5, the slope of the line of best fit is steeper meaning that the rate of increase in enrollment rates is higher. More on the reason to be discussed in later sections.

The purpose of this paper is to compare and analyse the outcomes of the school closures during the pandemic to better understand the impact of it. To see the comparison, this paper also analyses the enrollment rate in 2019 Figure 7. The graph shows that the initial 9% share of in-person school has a decrease of 1% in enrollment. These values stay relatively the same - hovering around the -0.10% to -0.019% enrollment rate- until the 55% to 64% share of in-person school where the enrollment rate jumps to 0.02. Then in the 73th percentile, we see a decrease to -1% in enrollment rate but the proceeding values even out to approximately 0\$ (or more specifically 0.001% and 0.004%). Although the values are negative for the large portion of the graph, there is still an upward trend we saw in previous graphs where the more proportion of in-person schooling, the larger the enrollment rate.

Lastly, in order to properly make a conclusion about enrollment rates and in-person schooling the reader must also see the opposite correlation, hence

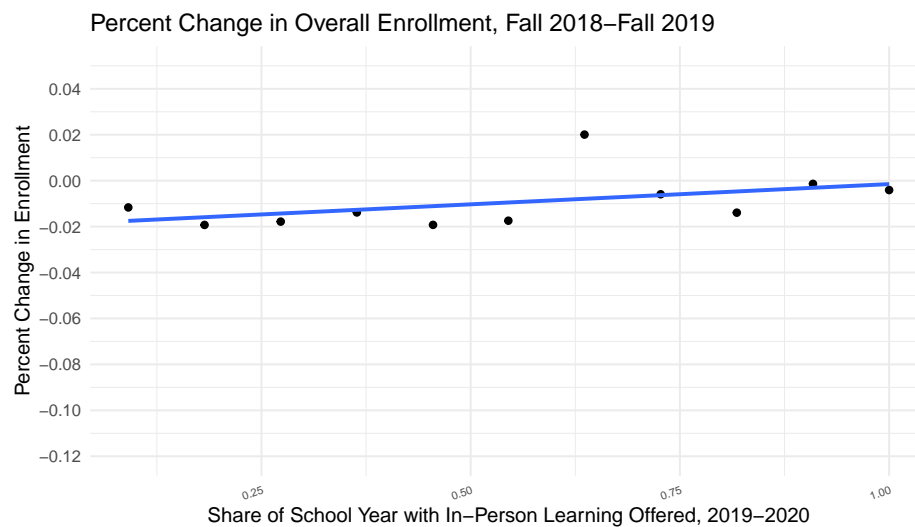


Figure 7: The weighted mean of the percent of total enrollment in 2019 accross all districts was calculated and graphed. The weights were decided on the bins which were cut in eleven sections based on share of in person school overall.

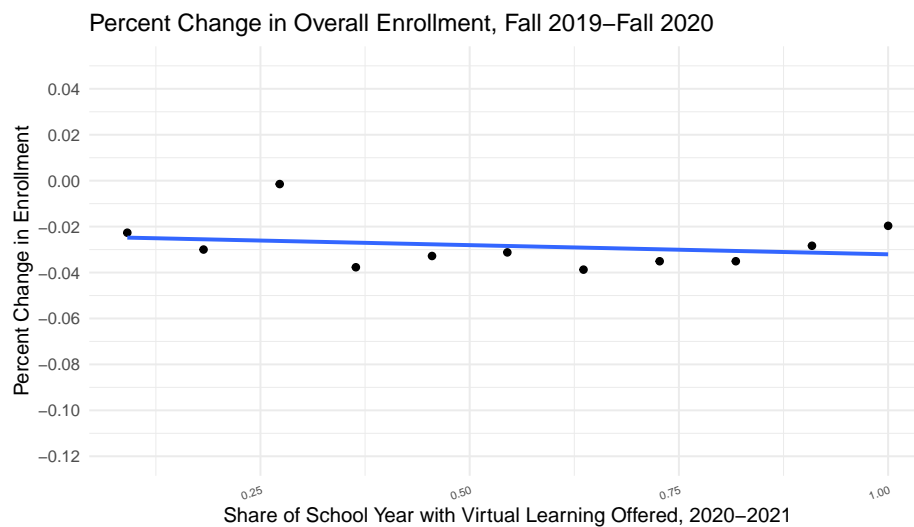


Figure 8: The weighted mean of the percent of total enrollment in 2020 accross all districts was calculated and graphed. The weights were decided on the bins which were cut in eleven sections based on share of virtual school overall.

the paper looks at the enrollment rates for share of virtual learning as well Figure 8. The graphs shows that for the first 9% of schools doing virtual learning in the districts have a enrollment rate of -2%, which increase to approximately -3% in the 18% share of virtual school. Then the value jumps up to the maximum value in the entire graph at 0.001% decrease in enrollment rate when 18% to 27% of schools are virtual. Then for the shares from 27% to 72% we see that the enrollment rate fluctuates between 3.1 to 3.7% decrease in enrollment rate where the change is so small the difference is not changing the graph at all. Finally we see a slight increase in enrollment rate with -2.8% and then proceeding with an even higher enrollment rate of -1.9% for 90% and 100% share of virtual school respectively. Looking at the graph overall, the line of best fit seems to have a negative slop compared to the line of best fit computed for the other graphs.

Discussion

Categorization of Continents

In the similar replication Figure 1 of the graph Figure 2 in the original paper (Jack and Oster 2023), the authors decided to analyze data spanning from January 2020 to December 2021, excluding any data recorded from the year 2022. While this could be due to various reasons, it is imperative to analyze the period following the conclusion of the Covid -19 pandemic to gain insight into the implications of school closures with no stay-at-home mandates.

Change in Test Scores of 2022

In the near identical Replication of the graph in the original paper(Figure 3), we have included the year 2022 in the change in test scores,

where the original paper failed to do so. This could be due to a variety of reasons, among them possibly the fact that the original authors may have thought it to be unimportant, or they may have been trying to make the change in test scores to seem more permanent, as if the virus had more long term effect. While the test scores did not increase as much as they fell, they increased never the less, although we do not have data for spring 2023, it is possible we could have made further conclusions with this additional information.

Change in enrollment rate

Kindergarten vs Total Enrollment

The third set of graphs that were shown in the results section was graphs that showed the change in enrollment with the shares of in-person school that started (Figure 5, Figure 6). As discussed in the section it seemed like the overall trend of the graph was that enrollment rates were higher the more in-person school there was. This was seen in both cases but more predominantly in the kindergarten graph (Figure 6). This result was interesting as this shows with younger ages the enrollment rates being higher means the children enjoy going to school in person. This is important as kindergarten is when most children start developing the fundamental social skills required to interact with other people.

Total Enrollment in 2020 vs Total Enrollment in 2019

The graph Figure 7 also shows the share of in-person schooling vs the enrollment rates and the result is that the enrollment rate is a lot higher then the rate for 2020 however, the over all trend stays the same. The rates are slightly higher meaning that in 2019 the enrollment rate was decreasing but not by the magnitude it was decreasing at in 2020. This shows that COVID-19 greatly affected the students, to the point where some students

may have dropped out. This comparison can be supported by the fact that during COVID-19 the mental health of young adults decreased by a large amounts. This result was cited in the original paper as well, where they state the NCES reported an “accelerated decrease in youth mental health levels” (Jack and Oster 2023). There were no other factors that would greatly influence the results shown in these 2 figures Figure 7 other then the limitation of social interactions during the pandemic.

Total Enrollment for in-person learning vs virtual

Last relation is the comparison between having an in person class and virtual class. If it wasn't clear from the previous three graphs (Figure 5, Figure 6 and Figure 7), in-person learning results in higher increase in enrollment rates, while in this section one notices the negative correlation with virtual environment with enrollment rates. In Figure 8 we see that the line of best fit is negative compared to the previous graphs. This means that as more schools went virtual the larger the decrease was in the enrollment rates. This result supports the papers hypothesis that students are more willing to learn in a environment where they are able to physically interact with other people. Part of this might be due to the fact that a lot of students, when in person feel more inclined to work harder as the task they are doing is not as easy to ignore compared to a virtual environment where they can simply log off.

Finishing Thoughts

In both of these comparisons, it can be seen that the major cause to the trends seen in the graphs shown above is social interactions or the lack there of during the pandemic. What is left to learn is what to do with such data. The reason why this is so interesting was that during this time there was no other options but to put the students on lock down to stop the spread of the disease but it came at the great cost of the future and the

mental health of the students. The interesting thing was the towards the end of the pandemic, educators seemed to notice the gradual decrease in enrollment and saw the correlation with social interactions. Hence many different programs were implemented so that students don't feel the gap in the interactions. Some of the measures taken were putting students on a zoom call with the camera's turned on or allowing guidance support from the school.

Ethics and Bias

Some of the ethical concerns of this research is the idea of consent. All of this data is publicly available so, it is more important that consent is gathered when performing these surveys. Other ethical concerns lie in the area of having the choice of in-person vs virtual schooling. Although this paper shows that in-person schooling has allowed a better enrollment rate and such, however that does not necessarily mean that students were doing better in-person. many studies during the pandemic found that course average had been either very high due to the fact that classes were taken online and students had the flexibility required to maintain a proper school-life balance but on the other hand other districts saw lower averages as the district school board would decide to test on harder material to fight against the real possibility of plagiarism during virtual school. Plagiarism is another big ethics topic to cover when talking about virtual school as attending a school online makes it very difficult for educators to monitor every students move and take action again every academic offence case that comes up. This leads to a difficult question for policy makers on whether they allow the virtual school format with stricter rules to combat such offence or to come up with a better more relaxed approach where students aren't restricted by the fear of accidentally committing plagiarism.

In terms of bias in this paper, there is not a lot of it. Most of the data would not be considered bias as the data was obtained from a reputable government website which by law requires them to publish truthful data.

Also the fact that the US is considered a developed country diminishes the chance missing major data (such as the enrollment rate or test scores) from educational districts, as they are all accounted for in organised manner. A possible bias could be the fact that the data only looks at the enrollment rates for grade 12 and lower where in the states this is almost mandatory. This means that change in enrollment is never due to the fact that students just stopped going to school. A possible way to combat this bias is to have the same data but for university/college students or for the workplace. This allows a more in-depth analysis about the mental health issues that might surround people during the pandemic, resulting in them either having higher productivity -which would be showcased by an increase in “enrollment” as there is more virtual options- or a decrease in productivity resulting in dropping out - which would be showcased in the graph as either an increase in “enrollment” as more in-person options open up or decrease in “enrollment” as more virtual option open up.

Weakness and Limitations

Translating code

One significant challenge encountered throughout this paper pertains to the translation of specific variables, functions, and lines of code from Stata to R. This posed difficulty due to differences in syntaxes and data structures. Much difficulty lied in recreating the continent category in Figure 1, as Figure 2 in the original paper is divided by criteria chosen by the author. In this paper, the data was categorized into continents for analysis using the package `countrycode` (Arel-Bundock, Enevoldsen, and Yetman 2018).

Some of the weaknesses seen in the enrollment data was that there were a lot of sections that had “NA” as a value meaning they didn’t have data for that section. These values were always in the sex, grade or race column meaning that the enrollment rates were accounted for but the small

data associated with it was not. This was particularly difficult during the cleaning process because when trying to extend the data by adding new columns the function `pivot_wider()` was used. Initially the plan was to use the `reshape()` function but when trying to reshape the data set while some of the columns had no value was difficult. This was due to the fact that it would make an entire row empty if this was the case. This problem was tackled using the `r` help book (R Core Team 2022) and was replaced with `pivot_wide` instead.

Lack of Data description

The original file of the paper contained many data sets, all contained specific clean information used in various parts of the paper. However, the data does not include a description of the variables used or the column names. For some variables, the values were arbitrary, such as integer values of 0,1,9, and 99 for gender, or values 0,1,2 and 3 for school closure. With description of the variables, further analysis could have been conducted by the introduction of other variables. There were many variables such as `leadid` and `fips`, that had to be [googled](#) and searched for in order to find out what they meant as well.

Solution to Limitation

To help with the translation and replication of code, detailed comments should be included to explain the functionality of code. Moreover, variables and columns with their values should be explained. Furthermore, the best way to see the how test scores would change would be to wait several years and analyze the data that would be collected after a few ages complete these tests. While in the short term the test scores of Spring 2023 would be beneficial information to have before making any more conclusions than we have.

Furthur research

As stated previously we do not have data for the test scores in 2023, which could give us additional information about how the pandemic affected students grades, as well as the younger students who were in elementary and middle schools when the pandemic began and how their educational foundations may have been affected. If analyzing data specifically from the United States, it would be compelling to explore variation in educational foundations of children in different states, as Covid-19 restrictions were decided by the state government in the United States, and thus social and institutional restrictions varied by state.

Appendix

Table 2: Enrollment rates throughout all districts in the U.S. 2020 catagorized by share of inperson learning

Share of inperson	average total enrollment rate
0.0909	-0.0306700
0.1820	-0.0350468
0.2730	-0.0360982
0.3640	-0.0352125
0.4550	-0.0363727
0.5450	-0.0332659
0.6360	0.0404678
0.7270	-0.0146317
0.8180	-0.0273175
0.9090	-0.0106241
1.0000	-0.0181826

Table 3: Enrollment rates for kindergartners throughout all districts in the U.S. 2020 catagorized by share of inperson learning

Share of inperson	average kindergarten enrollment rate
0.0909	-0.0967289
0.1820	-0.1028427
0.2730	-0.0990449
0.3640	-0.0840324
0.4550	-0.0938456
0.5450	-0.0822999
0.6360	-0.0324613
0.7270	0.0748270
0.8180	-0.0477577
0.9090	-0.0242239
1.0000	-0.0385117

Table 4: Enrollment rates throughout all districts in the U.S. 2019 catagorized by share of inperson learning

Share of inperson	average enrollment rate
0.0909	-0.0116537
0.1820	-0.0192675
0.2730	-0.0178225
0.3640	-0.0138305
0.4550	-0.0192475
0.5450	-0.0174195
0.6360	0.0200701
0.7270	-0.0059215
0.8180	-0.0139261
0.9090	-0.0013853
1.0000	-0.0040847

Table 5: Enrollment rates throughout all districts in the U.S. 2020 catagorized by share of virtual learning

Share of virtual	average enrollment rate
0.0909	-0.0226540
0.1820	-0.0299878
0.2730	-0.0014796
0.3640	-0.0377017
0.4550	-0.0327944
0.5450	-0.0312298
0.6360	-0.0387331
0.7270	-0.0350839
0.8180	-0.0350736
0.9090	-0.0283413
1.0000	-0.0196509

Reference

- n.d. *COVID 19*. Covid school data hub. <https://www.covidschooldatahub.com/data-resources>.
- Arel-Bundock, Vincent, Nils Enevoldsen, and CJ Yetman. 2018. “Countrycode: An r Package to Convert Country Names and Country Codes.” *Journal of Open Source Software* 3 (28): 848. <https://doi.org/10.21105/joss.00848>.
- Editor, CSRC Content. n.d. “Federal Information Processing Standard (FIPS) - Glossary: CSRC.” *CSRC Content Editor*. https://csrc.nist.gov/glossary/term/federal_information_processing_standard.
- Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. <https://CRAN.R-project.org/package=janitor>.
- Hyndman, Rob, and Mitchell O’Hara-Wild. 2021. *Tsibble: Tidy Temporal Data Frames and Tools*. <https://CRAN.R-project.org/package=tsibble>.

- Jack, Rebecca, and Emily Oster. 2023. “COVID-19, School Closures, and Outcomes.” *Journal of Economic Perspectives* 37 (4): 51–70. <https://doi.org/10.1257/jep.37.4.51>.
- R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2023. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.
- Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2023. *Readr: Read Rectangular Text Data*. <https://CRAN.R-project.org/package=readr>.
- Wickham, Hadley, and Evan Miller. 2021. *Haven: Import and Export ‘SPSS’, ‘Stata’ and ‘SAS’ Files*. <https://CRAN.R-project.org/package=haven>.
- Wickham, Hadley, Davis Vaughan, and Maximilian Girlich. 2024. *Tidyr: Tidy Messy Data*. <https://tidyr.tidyverse.org>.
- Xie, Yihui. 2014. *Knitr: A Comprehensive Tool for Reproducible Research in R*. Edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC. <http://www.crcpress.com/product/isbn/9781466561595>.
- Zhu, Hao. 2021. *kableExtra: Construct Complex Table with ‘Kable’ and Pipe Syntax*. <http://haozhu233.github.io/kableExtra/>.