



COVID'S IMPACT ON STUDENTS' LEARNING OUTCOMES

Interdisciplinary Focus on Gendered Outcomes For Schools in New York

TEAM 9

Kashaf Ali
Beibei Du
Eric Rios Soderman

EXECUTIVE SUMMARY



It is no secret that the Covid-19 pandemic affected every sector of human civilization humanly imaginable, and we decided to focus specifically on Covid-19's long-term effects on the educational outcomes across genders in the New York school system. Pandemic regulations and online learning modalities stressed the educational system and affected student well-being. As students were returning to school and were studying online as well, they faced an amalgamation of issues such as social isolation, increased anxiety, and strong familial and financial instability. According to a study conducted by McKinsey, students were falling behind by an average of five months in mathematics [1]. When considering this information, we devoted our attention to quantifying this impact under a gendered lens, given that we also know how females are impacted in a more nuanced way during public health emergencies.

In our search for answers, we looked at four academic years: 2018 and 2019, categorized as pre-pandemic, and 2021 and 2022, as post-pandemic. We centered our attention on the mean scores per school for the subjects of English, Math, and Science. Here we are running a fixed effects model on a school level to account for school-level variations as well as the time variations of the given years. We found that, prior to the pandemic, females scored about 3 mean score points higher than their male counterparts, substantiating the existence of a performance gap for the genders. When examining the interplay of how the pandemic's effects varied academic performance per gender, the presence of the pandemic led to a decrease of 1.36 mean score points larger for girls when compared to boys' decrease in scores. This signifies, that although females perform better than their male counterparts in terms of academic performance, the pre-existing gap between them was narrowed due to how females reacted to the pandemic's harrowing effects, which is a finding we wanted to verify.

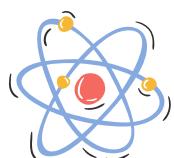


EXECUTIVE SUMMARY

As for our general results, a decreasing slope was observed for the mean scores of most of the English assessments, before and after the pandemic, but, despite this trend, the post-pandemic scores were higher, with females scoring more than males. Although, decreasing scores were detected for the assessments of lower grades. For Math assessments, scores generally improved post-pandemic compared to their pre-pandemic levels. However, most females scored less than their male counterparts. Lastly, for Science, both genders exhibited increasing scores after the pandemic and decreasing ones pre-pandemic. Unfortunately, regardless of trends, the scores themselves, despite the increasing trend, are lower post-pandemic. Additionally, females performed equally or slightly less than males in this subject matter.

The main takeaway is that this study can contribute to the exploration of how the pandemic has affected the students who are to become tomorrow's future. This future includes boys and girls, and, although both were impacted, this study helped us notice the nuanced effects on the academic performance of female students impacted by this pandemic. We hope that more studies like this can quantitatively shed light on some of the ways the pandemic has affected students.

THE PANDEMIC GENERATED A DECLINE OF 1.36 POINTS LARGER FOR FEMALE SCORES IN COMPARISON TO THEIR MALE COUNTERPARTS' DECLINE IN SCORES



PROBLEM STATEMENT

While the Covid-19 pandemic affected lives and livelihoods all across the globe, it also took a dramatic toll on the education system in the United States. Several post-peak pandemic retrospections shed light on how the pandemic not only disrupted school life, but also adversely impacted student learning and well-being. Although rising vaccination levels and adaptation to remote learning has enabled students to eventually return back to school, it aided in creating or worsening pre-existing, significant learning gaps that the education system continues to cope with now.



It is also important to recognise that the pandemic widened pre-existing opportunity and achievement gaps in the United States, and historically disadvantaged students suffered the brunt of it. UNESCO's analysis on the pandemic's effect on education also argues that the most vulnerable learners were hit the hardest by the lack of schooling [2].

THE PANDEMIC NEGATIVELY IMPACTED K-12 STUDENT LEARNING WITH STUDENTS FALLING BEHIND BY AN AVERAGE OF FIVE MONTHS IN MATHEMATICS AND FOUR MONTHS IN READING (MCKINSEY)

In addition to this, students who returned to school also suffered from mental health concerns due to increased anxiety, stress, depression, social isolation, family financial instability, and uncertainty about the future. It is also important to note that many researchers have also predicted the ripple effects of the pandemic on students' college and job market performance. McKinsey's analysis of the lingering effects of the pandemic suggests that today's students are likely to earn \$49,000 to \$61,000 less over their lifetime owing to the effect of the pandemic on their schooling than their counterparts today unless steps are taken to address unfinished learning [1].

In addition to this, it is also critical to evaluate the impact of Covid-19 from a gendered lens, where learning outcomes for females may be impacted differently from their male counterparts. While previous research sheds light on how girls and women tend to experience a higher risk of dropping out and negative effects on their learning outcomes during public health emergencies, it is important to note that these affected women also become more vulnerable to COVID-19-related economic effects because of existing gender inequalities in the long-run [3]. Moreover, it is also critical to note that school girls belonging to marginalized racial groups or possessing different English language skills may be affected differently than their white counterparts, and hence, it is critical to understand how the pandemic affected their learning outcomes through an interdisciplinary approach.

Therefore, our project will focus on understanding the impact of the Covid-19 pandemic on student's learning outcomes, more specifically through student's mean test scores, and how these learning outcomes varied for female students, i.e. those who are in elementary or middle school and belong to different demographic groups (race, English language speakers etc.) More specifically, our focus is on schools in New York. As a lot of research in this space is focused on students' learning outcomes in developing or underdeveloped countries, we aim to understand if gendered differences in learning outcomes were also exacerbated during the pandemic in a more developed city. Since New York City was one of the epicenters of the outbreak in the United States, we decided to limit our study to schools in this city.



RESEARCH QUESTION

MAIN CAUSAL QUESTION

WHAT IS THE CAUSAL
EFFECT OF THE COVID-
19 PANDEMIC ON
STUDENT'S TEST
SCORES BY GENDER IN
NEW YORK CITY?



SECONDARY RESEARCH QUESTION:

HOW DOES THIS
GENDERED EFFECT
VARY FOR STUDENTS
FROM DIFFERENT
DEMOGRAPHIC GROUPS
(RACES, ENGLISH
LANGUAGE LEARNERS)?



RESEARCH DESIGN



NEW YORK STATE EDUCATION DATASET

We used the Report Card dataset from the official website of the New York State Education Department for four academic years, namely, 2017-2018, 2018-2019, 2020-2021, and 2021-2022 to understand our problem [4]. This school-level dataset provides a wide range of assessment results for elementary and intermediate-level for ELA, Math, and Science assessments.

THE COMBINATION OF DIFFERENCE-IN-DIFFERENCES AND FIXED EFFECTS IS A POWERFUL METHOD FOR ESTIMATING CAUSAL EFFECTS IN PANEL DATA, AS IT ALLOWS US TO CONTROL FOR BOTH OBSERVED AND UNOBSERVED SOURCES OF VARIATION ACROSS SCHOOLS, AND OBTAIN MORE ACCURATE ESTIMATES OF THE IMPACT OF THE PANDEMIC ON ACADEMIC ACHIEVEMENT.

Our study used a panel dataset, containing information regarding assessment scores from the same schools before and after the COVID-19 pandemic outbreak. We observed pre-pandemic and post-pandemic mean test scores by gender and subject, and other demographic variables such as race, and whether English is their first language for different schools in New York. The design of our causal inference study used a combination of difference-in-difference (Diff-in-Diff) and fixed effects (FE) to estimate the impact of the COVID-19 pandemic on academic achievement while controlling for unobserved variation across schools.

Difference-in-Difference is a statistical technique that measures the effect of an event or treatment by comparing changes in a group that was exposed to the event or treatment to changes in a similar group that was not exposed. We used Diff-in-Diff to compare the change in mean test scores for females and males at the school level across different subject assessments before and after the pandemic. This approach will provide an estimate of the causal effect of the pandemic on female test scores. However, this approach may not account for unobserved variation between different schools, such as differences in school quality, student demographics, or teacher quality.

To account for this variation, we added **fixed effects** at the school level to the Diff-in-Diff model. The fixed effects will capture the differences in the unobserved school-level characteristics that are constant over time, such as school culture or resources. By controlling for these factors, we can better isolate the impact of the pandemic on academic achievement. In a fixed effects model, the analysis focuses on changes within each school over time, rather than differences between schools. This means that we are looking at the impact of the pandemic on changes in test scores within each school, rather than comparing average test scores between different schools. Additionally, the use of fixed effects helps to reduce bias from unobserved time-invariant factors that vary across schools.

FINDINGS

We compared the mean scores for females and males across different subject assessments both before and after the pandemic to analyze how mean test scores changed before and after the pandemic for both gender groups. It is evident that the mean test scores changed after the COVID-19 pandemic for each subject collectively (Figure 1, 2 and 3) and at the assessment-level (**Appendix I**). In our analysis, the gender variable was coded such that 1 represents females and 0 represents males.

ENGLISH (ELA)

English (ELA) mean scores for all assessments showed a decreasing trend both before and after the COVID-19 pandemic outbreak. Although the downward slope of the trend became steeper after the pandemic outbreak, post-pandemic mean scores for all English assessments were higher than pre-pandemic ones, with females scoring more than males. Individual mean scores for assessments ELA3, ELA4, ELA5, and ELA8 witnessed decreasing trends in the post-pandemic period for both genders and mean scores for ELA3, and ELA4 decreased even below their pre-pandemic levels for both genders. In contrast, ELA6 and ELA7 observed an increasing trend compared to a decreasing trend pre-pandemic. These trends indicate that for some grades, although scores were higher initially, they might continue to decline, indicating worse long term effects of the pandemic.

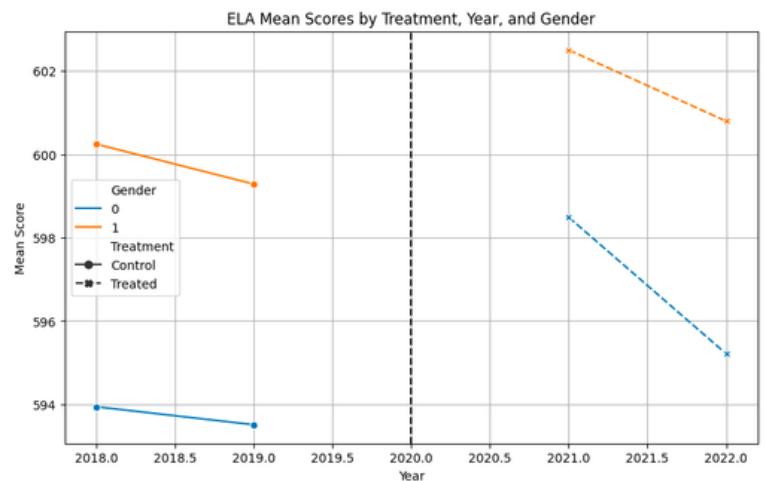


Figure 1: Mean English language scores decrease before and after Covid-19 for both genders.

MATH

Math mean scores in Figure 2 showed an increasing trend before and after the pandemic. Nevertheless, for the male group, the increasing trend was not as sharp as the female group's one after the pandemic. Furthermore, individual assessment mean scores were also lower post-pandemic, indicating that all grade levels and genders were impacted negatively due to the pandemic. However, all assessment mean scores witnessed an upward trend post-pandemic, regardless of what their trend was pre-pandemic, except Math5 for both genders and Math8 for males; this might indicate that scores for most grade levels are improving post-pandemic, but they are still lower than what they used to be before the pandemic. Although there was less variation in female and male scores pre-pandemic, during post-pandemic, most females scored less than their male counterparts.

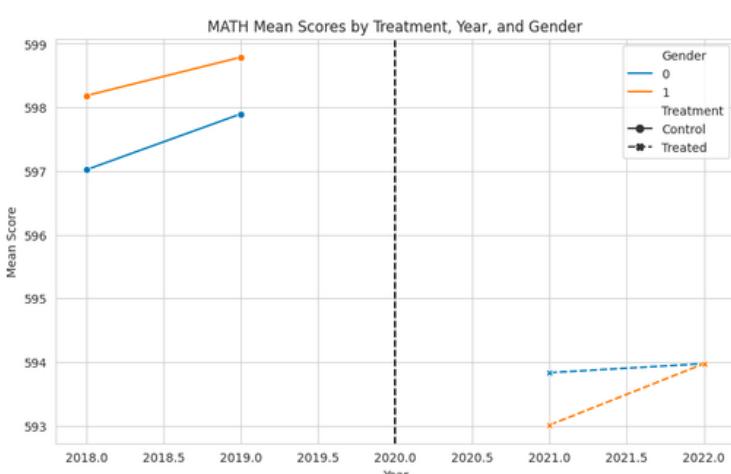


Figure 2: Math assessment scores have generally shown an increasing trend for both genders in both pre-and post-COVID eras.

FINDINGS

SCIENCE

For Science in Figure 3, both groups exhibited a decreasing trend before the pandemic and an increasing trend after the pandemic. Moreover, individual assessment mean scores were also lower post-pandemic, even though they depicted an upward trend for both genders; this probably indicates that mean scores are improving but they have not reached the pre-pandemic levels. It is important to note that we only had data available for mean scores for grade 4 and 8 for science. It was interesting to note that, although females performed better than their male counterparts, their performance reduced or became equal to their male counterparts post-pandemic.

Controlling for schools as entity effects and subject assessments in our fixed effects regression in addition to time effects, our model reveals some key insights (Model Results shown in **Appendix II**). Without COVID-19, females scored about 3 mean score points higher than their male counterparts. When accounting for the interaction effect between the COVID-19 pandemic and gender, it allows us to compare the change in mean scores for both genders. The presence of COVID-19 produced a decrease of 1.36 mean score points larger for girls than the decrease in mean scores for boys, indicating that although females continued to perform better than their male counterparts in terms of mean scores by 3 mean score points, the gap between the mean scores of both genders decreased as females were affected by the pandemic more. In addition, we considered if the changes in gender composition of the schools after the pandemic's outbreak skewed these results, but our tests proved it was statistically insignificant (Compositional Tables and T-Test in **Appendix III**).

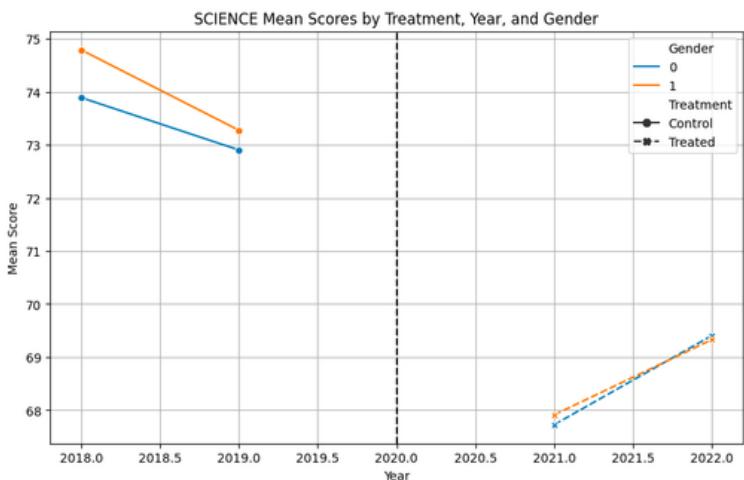


Figure 3: The mean scores for science showed a decreasing trend for both genders before the pandemic, but an increasing trend after the pandemic.

FEMALES CONTINUED TO PERFORM BETTER THAN MALES IN TERMS OF MEAN SCORES BY 3 MEAN SCORE POINTS, BUT THE GAP BETWEEN THE MEAN SCORES OF BOTH GENDERS DECREASED AS FEMALES WERE Affected BY THE PANDEMIC MORE.

FINDINGS

Moreover, we consider the presence of racial and English language proficiency gaps. We define these gaps in the following manner. A racial gap indicates a school where the white mean scores surpassed the non-white ones, and an English language learner (ELL) gap refers to the gap between the mean scores of English language learners and non-English Language Learners. Results show that, if a given school exhibits a racial gap, it would expect to see a rise in its mean scores by 0.63, which comes across as quite unintuitive. Likewise, the presence of an ELL gap means that when a given school demonstrates said gap, its overall mean scores will increase by 0.34 points. However, both of these gaps, in this case, are statistically insignificant (as shown in Figure 4 below in gray), likely because the racial composition and the English language learner composition in a school do not change much over time when we account for school-level entities and time effects.

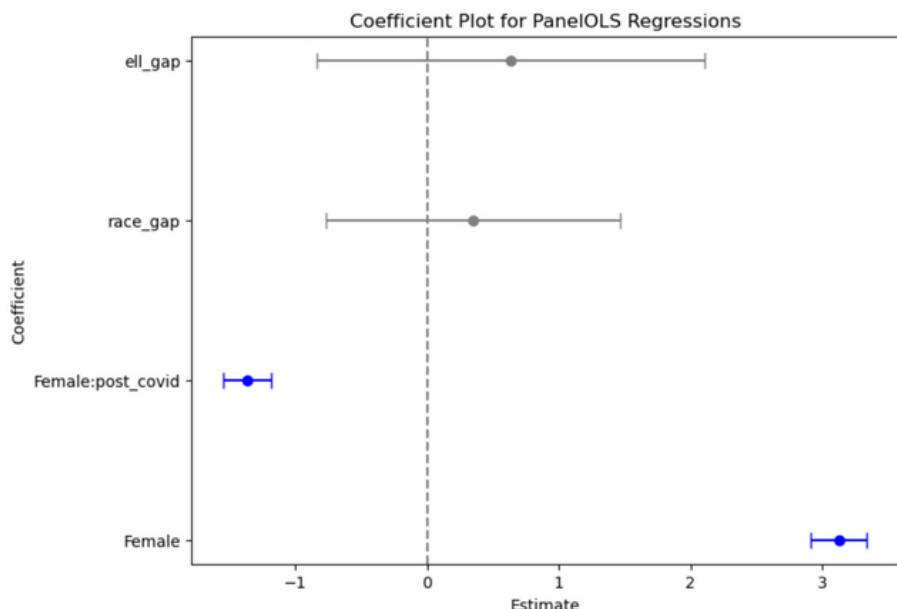


Figure 4: The coefficient plot shows the fixed effects model regression results, with statistically significant coefficients in blue. Results show that females continued to perform better than their male counterparts in terms of mean scores by 3 mean score points.

LIMITATIONS

MISSING DATA ISSUES

Since our dataset had a lot of missing values and the survey questions varied across some years, we had to limit our analysis to schools that completed the survey in all four years to ensure our results were for the same schools. To address this, we used mean imputation based on the same school, year, and subject. Although we acknowledge that this may introduce bias in our causal inference estimate, mean imputation is a common method to deal with missing values in data analysis. Furthermore, the date of the mean score was not available for 2020 due to the COVID-19 pandemic in that year. However, we used that year as our natural pre and post-pandemic comparison and considered two years before and after that for our comparison. The section pertaining to detailed data imputation is provided in **Appendix IV**. While we undertook the task of imputing the missing data, we acknowledge the possibility of bias in our causal inference study. However, based on our analysis and understanding of the data, we anticipate that any potential bias resulting from the imputation process will not significantly impact the results of our study.

EXTERNAL VALIDITY

Our study focuses solely on elementary and middle school districts in New York, which might limit the generalizability of our findings. It is possible that the results may not be applicable to other states or countries with different education systems across the globe. Additionally, the study only examines data from schools in more developed areas since New York is more developed compared to other places in the United States, and other places in the world, and hence, may not be representative of schools in less developed regions. Therefore, caution should be exercised when interpreting our findings and making conclusions that extend beyond the context of our study. To enhance our study on Covid-19's impact on learning in the New York school system, we suggest adding more years to analysis, subject to data availability. Since our focus on the gendered dimension of the pandemic's effects highlights the potential disproportionate impact on females, further analysis of additional years could provide a more comprehensive understanding of the pandemic's impact on academic performance and identify trends to inform policies to support disadvantaged groups.



CONCLUSION



The Covid-19 pandemic affected every facet of human life. Within those facets, we chose to focus on the test assessment scores of students from the New York school system (elementary and middle school) because there is a well-documented understanding of the severe disruptions that global education suffered. The online modalities and partial assistance to school brought a mix of negative effects on students' well-being, such as social isolation and heightened anxiety.

When examining the gendered dimension to this view, we also develop a somewhat panoramic appreciation of the struggles students traverse. It bears to mention that both genders suffered this impact, without room for doubt. Also, during public health emergencies, girls and women tend to bear a higher brunt of the blow that may lead to them dropping from school faster or, as seen in our study, performing academically worse.

For these reasons, we decided to focus on the test assessment scores for the genders, to verify if the pandemic was impacting them differently. Fairly mixed results were observed. In some cases like Math and English assessments, higher post-pandemic scores were observed. In some cases, female students would consistently outperform their male counterparts, but, in others, like Math assessments, female students would score less than males. For other assessments, the post-pandemic scores were consistently lower, like Science, further demonstrating the extent of the varied effects brought forth by the pandemic.

- Female students are at **higher risk of suffering from the negative academic effects of COVID-19 pandemic**
- **The Gender gap between the mean scores of both genders decreased as females were affected by the pandemic more.**

CONCLUSION



Lastly, when extracting the insights from our fixed effects model, we discovered that, although a performance gap of 3 mean score points existed in favor of female students, the interaction effect of being female students and the pandemic caused a decrease of 1.36 mean score points in comparison to males. This evidenced how female students were impacted a little differently than male ones during a pandemic that spared no one its effects.

In this same vein, we worry if there are more areas where young boys and girls were left suffering some of the nuanced effects of the pandemic, where each of these effects could be its own research topic. As for our research focus, we cannot conclusively quantify the extent to which female students suffered in comparison to male students, or vice-versa, given our limitations. However, we hope to get an idea for further research that can aid in the protection of disadvantaged groups in terms of gender, so that the effects of the post-pandemic world can be curbed as much as possible when it comes to the education of students in general and, more specifically, the students in New York schools as well.

- **Gender disparity**
found: Girls scored lower on assessments than boys during pandemic. Further research needed to understand extent of impact.

REFERENCES

- [1] Dorn, E., Hancock, B., & Sarakatsannis, J. (2021). COVID-19 and education: The lingering effects of unfinished learning. McKinsey.
- [2] World Economic Forum. (2022, November 11). Covid-19 education impact: What will be the legacy? Retrieved from <https://www.weforum.org/agenda/2022/11/covid19-education-impact-legacy/>
- [3] McKinsey & Company. (2020, July 15). COVID-19 and gender equality: Counteracting the regressive effects. <https://www.mckinsey.com/featured-insights/future-of-work/covid-19-and-gender-equality-counteracting-the-regressive-effects>
- [4] New York State Education Department. (n.d.). Downloads. Retrieved from <https://data.nysed.gov/downloads.php>

APPENDIX I: TABLES OF GENDER MEAN SCORE IN THREE ASSESSMENTS

In our analysis, the gender variable was coded such that 1 represents females and 0 represents males, and the number of the Assessment corresponds to the school grade.

Table 1: ELA Scores by Year and Test

Year	ELA3		ELA4		ELA5	
	Test	0	1	0	1	0
2018	595.385	600.163	594.345	599.688	595.093	599.883
2019	595.021	598.800	594.219	598.859	593.447	599.775
2021	598.590	600.274	600.717	603.494	599.002	603.005
2022	593.833	597.883	591.578	596.220	596.064	601.277

Year	ELA6		ELA7		ELA8	
	Test	0	1	0	1	0
2018	593.979	600.574	592.890	600.418	592.103	600.732
2019	592.575	598.161	593.107	599.972	592.569	600.184
2021	597.728	603.417	596.472	601.287	598.156	603.489
2022	598.102	603.828	598.261	604.641	593.138	601.171

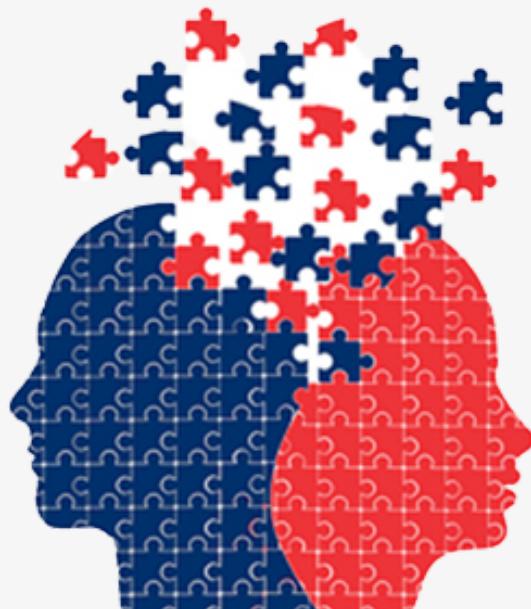


Table 2: MATH ASSESSMENT NAME Scores by YEAR and Gender

Year	MATH3		MATH4		
	Gender	0	1	0	1
2018		598.427	598.946	597.448	598.051
2019		598.739	598.224	599.011	598.202
2021		593.727	591.899	593.836	592.857
2022		594.295	593.599	594.274	593.218

Year	MATH5		MATH6		
	Gender	0	1	0	1
2018		596.943	597.007	598.257	598.119
2019		597.468	598.242	598.694	598.904
2021		593.636	592.641	594.807	592.523
2022		593.321	591.931	595.562	594.968

Year	MATH7		MATH8		
	Gender	0	1	0	1
2018		595.422	598.221	595.725	598.796
2019		597.642	599.450	595.841	599.707
2021		593.374	594.033	593.688	594.247
2022		594.179	595.505	592.336	594.608

Table 3: Science Scores by Year and Test

Year	Science4		Science8	
	Test	0	1	0
2018	81.062857	81.594688	66.724706	68.034772
2019	79.869318	80.898501	66.036655	66.364035
2021	75.178784	74.756786	59.967742	60.090571
2022	75.733087	75.673861	63.268176	63.172093

APPENDIX II: RESULT OF THE FIXED EFFECTS MODEL

	Parameter Estimates					
	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI
Intercept	595.96	0.6336	940.62	0.0000	594.71	597.20
ASSESSMENT_NAME[T.ELA4]	0.0005	0.1374	0.0040	0.9968	-0.2688	0.2699
ASSESSMENT_NAME[T.ELA5]	0.9704	0.1390	6.9806	0.0000	0.6979	1.2428
ASSESSMENT_NAME[T.ELA6]	0.9560	0.1994	4.7947	0.0000	0.5652	1.3468
ASSESSMENT_NAME[T.ELA7]	0.8527	0.2431	3.5072	0.0005	0.3762	1.3292
ASSESSMENT_NAME[T.ELA8]	0.2808	0.1999	1.4051	0.1600	-0.1109	0.6725
ASSESSMENT_NAME[T.MATH3]	-1.5046	0.2515	-5.9823	0.0000	-1.9976	-1.0117
ASSESSMENT_NAME[T.MATH4]	-1.7764	0.3338	-5.3222	0.0000	-2.4305	-1.1222
ASSESSMENT_NAME[T.MATH5]	-2.3445	0.2854	-8.2161	0.0000	-2.9038	-1.7852
ASSESSMENT_NAME[T.MATH6]	-0.8829	0.3516	-2.5107	0.0120	-1.5721	-0.1937
ASSESSMENT_NAME[T.MATH7]	-1.5708	0.3502	-4.4858	0.0000	-2.2571	-0.8844
ASSESSMENT_NAME[T.MATH8]	-1.8771	0.5208	-3.6041	0.0003	-2.8978	-0.8563
ASSESSMENT_NAME[T.Science4]	-519.47	0.6694	-775.98	0.0000	-520.78	-518.15
ASSESSMENT_NAME[T.Science8]	-533.36	1.1106	-480.24	0.0000	-535.53	-531.18
Female	3.1232	0.1085	28.791	0.0000	2.9106	3.3359
ell_gap	0.6353	0.7489	0.8482	0.3963	-0.8326	2.1031
race_gap	0.3470	0.5688	0.6101	0.5418	-0.7678	1.4618
Female:post_covid	-1.3608	0.0925	-14.704	0.0000	-1.5422	-1.1794

Figure 5. PanelOLS result from our fixed effect model

Figure 5: The table above is the PanelOLS results for the fixed effect model. The p-values and the coefficient are denoted in the parameter and P-value columns.

APPENDIX III: SUMMARY STATISTICS OF DEMOGRAPHIC COMPOSITION

Summary Statistics of Gender Composition of New York Schools

Treatment Arm	All Students	Male Share	Female_Share
Control	471.70	0.509	0.4899
Treatment	436.46	0.510	0.4892

T-Test for the Differences in Gender Compositions Percentages Before and After the Pandemic Across All Schools

Gender	P-value	T-Statistic
Male	-1.527	0.127
Female	0.906	0.365

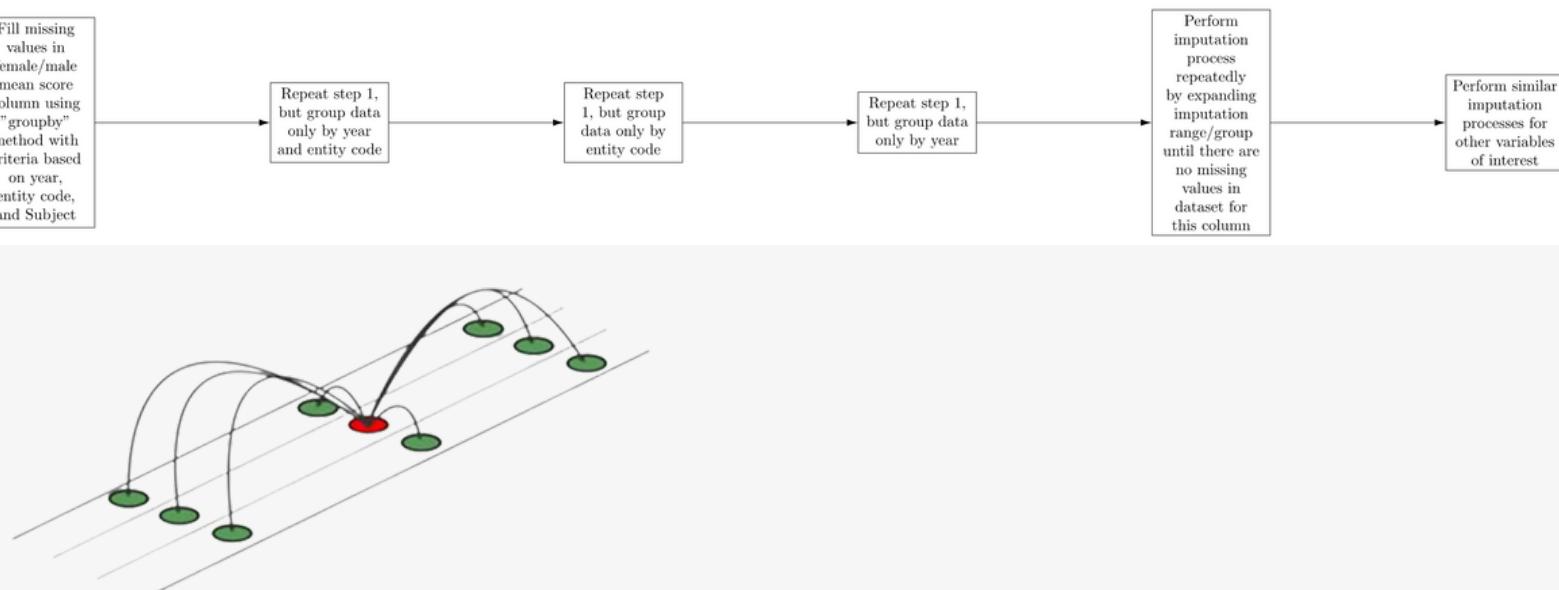
APPENDIX IV: IMPUTATION METHODS

In order to account for the missing values in both the female and male mean scores, we need to consider filling in the missing values in our dataset strategically to perform the causal inference analysis. The missing values are replaced by the mean of the corresponding data group based on a specific criteria using the "group by" method. The criteria for grouping data is determined based on the year, school id, and subject.

The imputation process is carried out in four steps:

1. Firstly, missing values in the mean score column are filled using the mean values grouped by year, school id, and subject (three options: Math, ELA, and Science).
2. The second step repeats the process for those values that are still missing after step 1 and replaces missing values by using the mean scores grouped by year and school id only.
3. The third step is applied to values that remain missing after step 2 and groups data only by school id.
4. The fourth step groups data only by year. It will eventually impute all missing values in the dataset.

The imputation process is performed repeatedly by expanding the imputation range/group until there are no missing values in the dataset for this column. Similar imputation processes are performed for other variables of interest, such as mean scores for different races (white vs. no white) and English language learners.



APPENDIX V: SUMMARY STATISTICS OF THREE ASSESSMENTS MEAN SCORE BY GENDER

Table 4: Summary statistics for Math Mean Scores by Gender

	Male_MEAN_SCORE	Female_MEAN_SCORE
count	41419.000000	41419.000000
mean	595.734470	595.999879
std	6.345676	6.498444
min	578.000000	578.000000
max	611.000000	612.000000
difference between 25% and 75%	6.0	6.0

Table 5: Summary statistics for Math Mean Scores by Gender

	Male_MEAN_SCORE	Female_MEAN_SCORE
count	16069.000000	16069.000000
mean	71.471053	71.667334
std	8.778889	8.798180
min	39.000000	40.000000
max	88.000000	88.000000
difference between 25% and 75%	12.000000	12.000000

Table 6: Summary statistics for Mean Scores by Gender

	Male_MEAN_SCORE	Female_MEAN_SCORE
count	42512.000000	42512.000000
mean	595.324308	600.713257
std	6.018730	5.654198
min	580.000000	584.000000
max	616.000000	619.000000
difference between 25% and 75%	8.0	7.0