How does face mask mandate effect online learning? From a regression discontinuity perspective

Yi Han (yh726)

May 2, 2022

1 Introduction

The COVID-19 Pandemic has disrupted learning for more than 56 million students in the United States. In the Spring of 2020, most states and local governments issued face mask mandate to stop the spread of the virus. In response, schools and teachers have attempted to reach students remotely through distance learning tools and digital platforms. Meanwhile, the concern is that education inequity will exacerbated by the COVID-19 pandemic.

The paper exploits a discontinuity in the participation and engagement in online learning that resulted from the issue of face mask mandate in public places. Our main finding is that there appears to be negative treatment effect on the face mask mandate on cyberspace learning which indicates people's reluctance and unfamiliarity to wear face mask when the mandate is issued. To find out the scope of the policy's effect to the education inequity, we also compare the face masks' effect among school districts varying in percentage of minorities and per pupil total expenditure from the local and federal government. The result shows that per pupil total expenditure plays a dominant role in children's continual preference in attending school wearing face mask and family's support in electronic appliances for online learning makes a difference when the governments' expenditures get close.

The paper is organized as follows. In section2, we will discuss our data and the research design. In section3, we will provide more detail about the classical RD framework and our methods in dealing with regression discontinuity. In section4, we present the results of running regression with discontinuity in assessing the effects of face mask mandate on online learning and compare the results among districts. In section5, we discuss the results and propose possible future improvements.

2 Preparation

2.1 Data

Dataset about online learning is derived from the Kaggle competition launched by education technology company LearnPlatform (https://www.kaggle.com/c/learnplatform-covid19-impact-on-digital-learning/data). The dataset provides a set of daily edtech engagement data from over 200 school districts in 2020, and includes three basic categories of data.

- Engagement Data: Based on LearnPlatform's Student Chrome Extension. The extension collects page load events of over 10K education technology products in our product library, including websites, apps, web apps, software programs, extensions, ebooks, hardwares, and services used in educational institutions. The engagement data have been aggregated at school district level, and each file represents data from one school district. The product file includes information about the characteristics of the top 372 products with most users in 2020.

- Districts Information: Information about the characteristics of school districts, including data from National Center for Education Statistics , The Federal Communications Commission, and Edunomics Lab.

- Product Information: Information about the characteristics of the top 372 products with most users in 2020.

COVID-19 related data is obtained from COVID-19 US State Policy Database on OPENICPSR (https://www.openicpsr.org/openicpsr/project/119446/version/V143/view). It basically provides the date when face mask mandate is issued in all states in the United States during the breakout of Covid-19.

Table 1: Original Variable Name and Description

Category	Name	Description
Engagement	time	date in "YYYY-MM-DD"
	lp_id	The unique identifier of the product
	pct_access	Percentage of students in the district have at least one page-load event of a given product and on a given day
	engagement_index	Total page-load events per one thousand students of a given product and on a given day
District	district_id	The unique identifier of the school district
	state	The state where the district resides in
	locale	NCES locale classification that categorizes U.S. territory into four types of areas: City, Suburban, Town, and Rural.
	pct_black/hispanic	Percentage of students in the districts identified as Black or Hispanic based on 2018-19 NCES data
	pct_free/reduced	Percentage of students in the districts eligible for free or reduced-price lunch based on 2018-19 NCES data
	pp_total_raw	Per-pupil total expenditure (sum of local and federal expenditure) from Edunomics Lab's National Education Resource Database on Schools (NERDS) project.The expenditure data are school-by-school, and we use the median value to represent the expenditure of a given school district.
Product	LP ID	The unique identifier of the product
	Sector(s)	Sector of education where the product is used
	Primary Essential Function	The basic function of the product. There are two layers of labels here. Products are first labeled as one of these three categories: LC = Learning & Curriculum, CM = Classroom Management, and SDO = School & District Operations. Each of these categories have multiple sub-categories with which the products were labeled
Covid-19	FM_ALL	The date a state mandated face mask use in public spaces by all individuals statewide.The order must apply state wide.

To adapt the original variables to the regression framework, we first transform the categorical variables **Sector(s)** and **Primary Essential Function** into several binary variables **Prek.12** / **Higher.Ed** / **Corporate** and **LC** / **CM** / **SDO** corresponding to each category in the variables.

Then more importantly, as for the time and treatment indicator: we generate the **diff_time** by the difference in **time** in the engagement data and **FM_ALL**. Further, we generate the treatment indicator **t**. Before the policy takes effect (**diff_time¡**0), we generated the indicator as 0 (t = 0) while after the policy takes effect (**diff_time¿**0), we generated the indicator as 1 (t = 1).

2.2 Districts Clustering

More than 300 districts varying in economic development and population structure are included. This leads to large variance in the data. Linear regression framework requires each covariate have the same marginal effect on the response variable for all samples in the dataset. However, the marginal effects seem to be influenced by some characteristics of the district. For example, the increase in percentage of access will be more significant in those wealthier districts where consequently the percent of children who are eligible to receive free or reduced price of meals are low. Therefore, we cluster all the samples into six categories by K-Means. The clustering is based on percentage of children that are black or hispanic, percent of children who are eligible to receive free or reduced price of meals and average per-pupil total expenditure by school.

The clusters are mainly based on the divergence in per pupil total expenditure and in the interval 10000-12000 which many districts fall into K-Means further cluster them into two categories based on the difference in the population structure of children. Figure 1 shows the distribution among categories. Table 2.2 shows districts falling into each category and their region types.

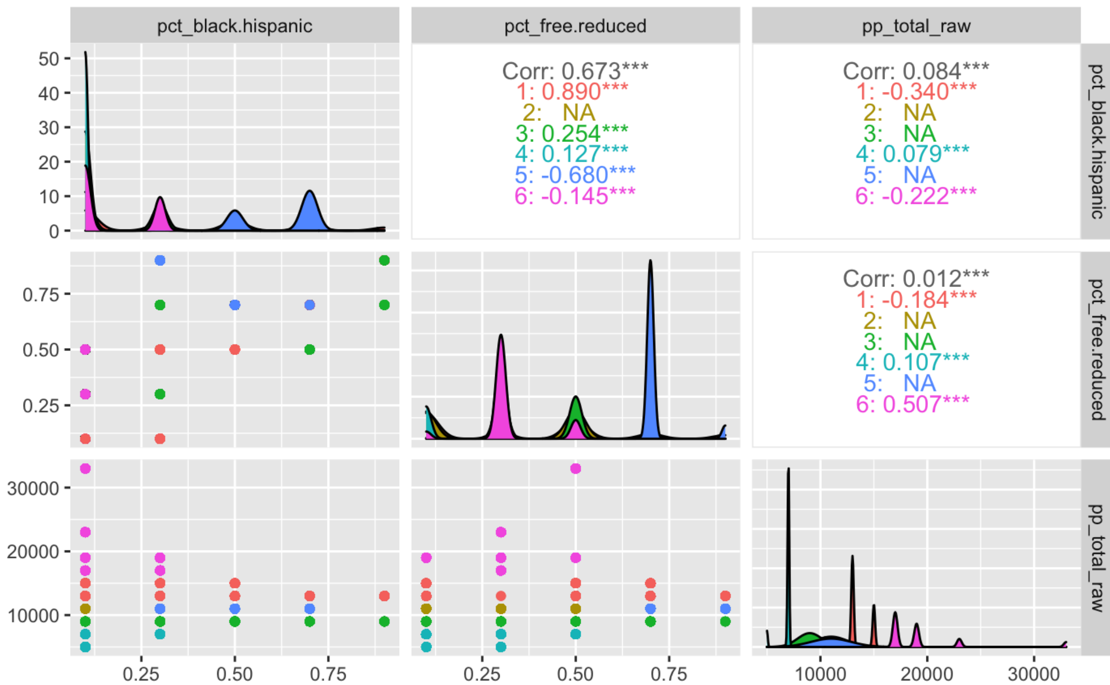

Figure 1: Clusters of Districts

Table 2: Districts in Each Category

	Category1			Category2			Category3			Category4			Category5			Category6	
4929	Virginia	Rural	3266	Utah	Town	4348	Missouri	Suburb	4921	Utah	Suburb	8784	Illinois	Suburb	8433	New Jersey	Suburb
8815	Illinois	Suburb	5987	Wisconsin	Suburb	7177	North Carolina	Suburb	3710	Utah	Suburb	5042	Illinois	Town	7798	New York	Rural
7457	Washington	City	1044	Missouri	Suburb	6584	North Carolina	Rural	9812	Utah	Suburb	7767	North Carolina	City	4629	Illinois	Suburb
1877	Illinois	Rural	2779	Illinois	Rural	7614	Utah	City	4936	Utah	Suburb	1791	Virginia	City	5890	Illinois	Suburb
9043	Illinois	Suburb	4051	Missouri	Suburb	4808	Indiana	City	2165	Utah	Suburb	6144	Michigan	Suburb	8520	New York	Rural
2167	Washington	Suburb	6577	Illinois	Suburb	1742	Missouri	Suburb	5600	Utah	Suburb				4775	New York	Suburb
3550	Minnesota	Suburb	5231	Utah	Town	3222	Indiana	City	6919	Florida	Suburb				1470	New York	Rural
9553	Illinois	Suburb	8902	Wisconsin	Suburb	1772	Utah	City	7387	Utah	City				1052	Illinois	Suburb
8937	Illinois	Rural	8328	Utah	Town	8884	Texas	City	9230	Utah	Rural				9515	New York	Rural
2130	Washington	Suburb	1549	Virginia	Suburb	3692	Utah	Suburb	6762	Utah	Suburb						
2393	Illinois	Suburb	7980	Virginia	Rural	4683	Texas	Suburb	7660	Utah	Suburb						
3248	Illinois	Suburb	7752	Wisconsin	Suburb	2956	Missouri	City	7308	Utah	Suburb						
2285	Illinois	Suburb				1536	Indiana	Rural	1204	Utah	Suburb						
5510	Washington	City				3772	Utah	Suburb	4668	Utah	Suburb						
2601	Illinois	Suburb				9778	Utah	Town	3864	Utah	Suburb						
9899	Illinois	Suburb				1324	Indiana	Rural									
1712	Illinois	Suburb				3160	Utah	Suburb									
2567	Washington	Suburb				5627	Indiana	Rural									
8256	New Jersey	Suburb				2870	Indiana	Suburb									
3732	Michigan	Suburb				3558	North Carolina	Suburb									
1705	Washington	City				4744	Utah	Suburb									
						7541	Utah	Suburb									
						1270	Utah	Rural									
						4183	Utah	Town									
						3228	Indiana	Town									
						5422	Missouri	Town									
						4373	Utah	City									
						2441	Utah	Town									

3 Methods

3.1 Sharp RD Design

In the standard RD design setting[CCFT19] , we have n i.i.d. random samples $\{Y_i, X_i, Z_i, t_i\}_{i=1}^n$, where Y_i is the outcome variable for the i th sample, $t_i \in \{0, 1\}$ is the binary treatment indicator, $X_i \in R$ is the running variable and $Z_i \in R^p$ is the covariate. In the sharp RD design, the treatment T_i is perfectly assigned through the running variable X_i relative to a known cutoff c. In the online learning dataset, we have

$$t_i = \mathbf{1}\left(X_i > 0\right)$$

We can write the observed outcome variable Y_i as

$$Y_i = Y_i(0) \cdot (1 - t_i) + Y_i(1) \cdot t_i$$

where $Y_i(0)$ and $Y_i(1)$ represent the potential outcomes without or with treatment. Under the assumption of all the covariates are continuous at the cutoff, the average treatment effect is defined as

$$\tau = E\left(Y_i(1) - Y_i(0) \mid X_i = 0\right)$$

3.2 Constant Treatment Effect Model

The average treatment effect of the sharp RD design needs the continuity assumption of all the covariates at the cutoff. However, the covariates in the model is not continuous at the cutoff. For example, the increase in percentage of access will be more significant in those wealthier districts where consequently the percent of children who are eligible to receive free or reduced price of meals are low. Therefore, we refer the the constant treatment effect model proposed in [PN19] to incorporate the influence of the discontinuous covariates to the treatment effect at the cutoff.

$$E\left(Y_i(t) \mid X_i, Z_i(t)\right) = \alpha + \tau \mathbf{1}(t = 1) + g\left(X_i\right) + Z_i(t)\gamma$$

for $t \in \{0, 1\}$, where $g(\cdot)$ is an arbitrary continuous function. The term interpreting the indirect effect of the The policy to the covariates is added the treatment effect at the cutoff:

$$\tau_{Indirect} = E\left(Z_i(1) \mid X_i = c\right) - E\left(Z_i(0) \mid X_i = c\right)\gamma$$

Two terms make up the estimand τ_{SRD}, both the direct and indirect terms.

$$\tau = \tau_{direct} + E\left(Z_i(1) \mid X_i = c\right) - E\left(Z_i(0) \mid X_i = c\right)\gamma$$

3.3 Model

Considering the interaction and polynomial terms of the covariates and the arbitrary continuous function g in the polynomial class, 5 models are constructed based on the linear regression. We start from a basic linear regression model, which is marked as Model 1, incorporating all the covariates in their original forms. Model 2 and 3 change the continuous function g to polynomials of degrees of freedom 2 and 3 respectively. From Figure 2, we can see that when degree of freedom(p) is 3, the majority of samples fall on or close to the regression line. Therefore, we stop to increase the degree of g.

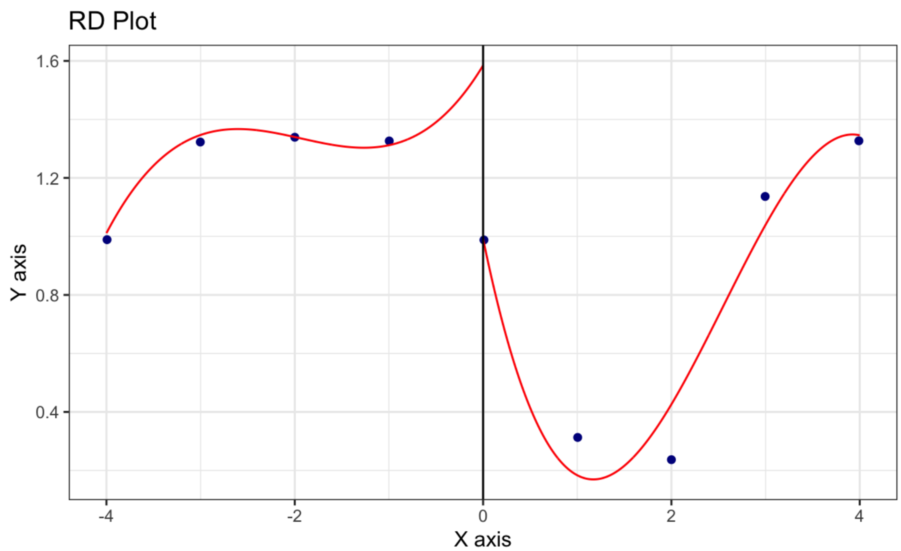

Figure 2: Regression with Discontinuity Plot with p=3

In Model 4 and 5, we keep the function g as a polynomial of degree of freedom 3 and change the form of the covariates to improve the performance of the regression model in terms of adjusted R^2. For Model 4, we replace the original **pct_free.reduced** with a polynomial with df=2 to see whether the **pct_free.reduced** has a quadratic influence on the response variable. In Model 5, we add an interaction between **pct_black.hispanic** and **pct_free.reduced** to validate our guess that the percentage of children eligible to receive reduced-priced/free lunch affects the marginal effect of the percentage of black and hispanic children on the engagement in online learning.

Apart from the basic linear regression framework, we apply more complex regression models, spline regression and the general addictive model.Table 3 shows all the regression models compared in the paper.

We perform seven regression models proposed in Table 3 and conducted model selection based on adjusted R^2.

Table 4 shows adjusted R^2 of the models applied to samples in each categories respectively. We can see that Model 6, which is the spline regression model, outperforms

Table 3: Regression Models

No.	Model Type	g(·)	Covariates
1	Linear	/	pct_black.hispanic , pct_free.reduced , pp_total_raw , PreK.12 , Higher.Ed , Corporate , Suburb , Rural , LC , CM , SDO
2	Linear	poly(2)	pct_black.hispanic , pct_free.reduced , pp_total_raw , PreK.12 , Higher.Ed , Corporate , Suburb , Rural , LC , CM , SDO
3	Linear	poly(3)	pct_black.hispanic , pct_free.reduced , pp_total_raw , PreK.12 , Higher.Ed , Corporate , Suburb , Rural , LC , CM , SDO
4	Linear	poly(3)	pct_black.hispanic , poly(pct free.reduced,2), pp_total_raw , PreK.12 , Higher.Ed , Corporate , Suburb , Rural , LC , CM , SDO
5	Linear	poly(3)	pct_black.hispanic , poly(pct free.reduced,2),pct_black.hispanic*pct free.reduced , pp_total_raw , PreK.12 , Higher.Ed , Corporate , Suburb , Rural , LC , CM , SDO
6	Spline	/	pct_black.hispanic , pct_free.reduced , pp_total_raw , PreK.12 , Higher.Ed , Corporate , Suburb , Rural , LC , CM , SDO
7	GAM	/	pct_black.hispanic , pct_free.reduced , pp_total_raw , PreK.12 , Higher.Ed , Corporate , Suburb , Rural , LC , CM , SDO

other regression model. It shows that when we use the percentage of access as a response variable, the original data fits better with a spline linear model than a polynomial model.

Table 5 shows rather than any complex linear regression model, the model adding an interaction between the percentage of children eligible to receive reduced-priced/free lunch and the percentage of black and hispanic achieves the highest adjusted R^2. It shows that when we use the engagement index as a response variable, the percentage of children eligible to receive reduced-priced/free lunch does affect the marginal effect of the percentage of black and hispanic children.

Table 4: Adjusted R^2 of Models for Percentage of Access

	Category1	Category2	Category3	Category4	Category5	Category6
1	0.04832	0.03887	0.02681	0.03613	0.04929	0.08991
2	0.04987	0.03995	0.03014	0.04257	0.05138	0.09267
3	0.05066	0.04091	0.03013	0.04254	0.05226	0.09281
4	0.05199	0.04141	0.03431	0.05312	0.05226	0.09281
5	0.05206	0.04141	0.04097	0.05317	0.05248	0.09271
6	0.0527	0.0415	0.0424	0.0558	0.052	0.0926
7	0.05265	0.0413	0.04239	0.0558	0.05164	0.09351

4 Results

In the following part of the paper discussing the results, we use Model 6 and Model 4 respectively in the percentage of access regression and engagement index regression. Table 6 and Table 7 show the regression coefficients.

Table 5: Adjusted R^2 of Models for Engagement Index

	Category1	Category2	Category3	Category4	Category5	Category6
1	0.01375	0.01384	0.01074	0.02546	0.01361	0.0268
2	0.01401	0.01367	0.01125	0.02676	0.01322	0.02751
3	0.01425	0.01383	0.01126	0.02677	0.01329	0.02748
4	0.01425	0.01383	0.01126	0.02567	0.01329	0.02748
5	0.01507	0.01413	0.01387	0.03004	0.01361	0.0277
6	0.0145	0.0142	0.0136	0.0289	0.0136	0.0273
7	0.0145	0.01382	0.1346	0.02947	0.01237	0.02715

To further discuss the effect of the per pupil total expenditure and the children population structure to the treatment effect, we transform the value of t into rank among categories and compare it with the rank of clustering criteria. The rank is shown in Table 8.

Negative treatment effect Our results show that the treatment effect of the face mask mandate to online learning participation, both the percentage of access and engagement index, is negative.

Online learning products considered in the dataset include three primary essential functions including curriculum, classroom management and school operations.

From children's perspective, after the face mask mandate take effect, every school-age children are asked to wear masks when they are attending schools but struggle to get familiar to the face mask. At the same time, some parents may prone to keep their children at home for fear of their children being infected by COVID-19 as the face mask mandate indicates a growing trend of the infection. These two reasons contribute to a sudden drop in children's attendance to schools. From schools' perspective, the face mask mandate affects schools' abilities to hold various activities as much as before owing to the restriction caused by face mask and social distance required during the pandemic. This results in a decrease in the percentage of access and engagement index for online learning products in classroom management and school operations sections.

For the curriculum section, although online courses are a good substitute for normal classes under the COVID-19 pandemic and those students who do not attend school will shift to online material, online courses need adequate setup for live lessons and preparation for prerecording lessons before all the students can have access to it. The transition period may also lead to a drop in online learning. To deal with the low attendance of children to school, schools need time to adjust their online learning resources from a supplementary role to a main role of teaching. Therefore, the use of online curriculum is not increasing immediately after the face mask mandate takes effect.

Table 6: Regression Coefficients for Percentage of Access

	Category1	Category2	Category3	Category4	Category5	Category6
(Intercept)	-1.69	0.25	-2.87	-0.78	-8.17	-4.73
pct_black.hispanic	-0.48	NA	6.96	-0.41	2.99	4.18
pct_free.reduced	-3.73	-2.83	5.24	0.74	6.07	1.75
pp_total_raw	0	NA	NA	0	NA	0
PreK.12	4.07	2.97	1.29	2.49	3.18	7.78
Higher.Ed	-0.35	-0.18	-0.27	-0.23	-0.38	-0.21
Corporate	1.37	1.09	0.56	0.69	0.83	2.05
Suburb	0.62	-0.84	0.78	0.28	0.88	-1.11
Rural	0.33	-0.55	-0.3	0.28	NA	NA
LC	-1.62	-1.37	-0.52	-1.07	-1.55	-3.55
CM	-0.3	-0.52	-0.14	-0.3	-0.64	-1.23
SDO	1.31	1.16	0.27	0.63	1.07	2.17
bs(diff_time, df = 8)1	1.74	1.42	0.98	1.75	-0.47	2.12
t	-0.93	-0.83	-0.6	-0.89	-0.01	-1.54

Per Pupil Total Expenditure Generally, the treatment effect is align with per pupil total expenditure. The rank of t accords with the rank of per pupil total expenditure especially when the per pupil total expenditure is largely above average, like more than 14000. High per-pupil total expenditure in one district indicates the local government's attention to primary education and potentially high education quality in the district. That means schools are receiving support strong enough to maintain the normal operating status of schools of all levels when the face mask mandate is issued. The support from the school can include free high-quality face mask provided for all students, regularly sanitizing school facilities and close monitor on both teachers and staffs health condition concerning COVID-19 related symptoms.

Children Population Structure Factors concerning children population structure in the paper are percentage of children who are black or hispanic and percentage of children who eligible for free or reduced-price lunch. The change of rank among Category2/4/5 shows some pattern that children population structure also affects online learning engagement.

The percentage of children who eligible for free or reduced-price lunch indicates that the family of these children may have a tight budget in supporting the equipment needed for online learning. Districts in Category 25 have similar per-pupil total expenditure but districts in Category5 have more minorities and more children who eligible for free or reduced-price lunch. To some degree, children in these districts are more likely

Table 7: Regression Coefficients for the Engagement Index

	Category1	Category2	Category3	Category4	Category5	Category6
(Intercept)	-194.17	75.55	-1000.02	-339.99	-2748.52	-1623.64
pct_black.hispanic	-348.88	NA	2581.53	-275.24	1023.89	1348.56
pct_free.reduced	-930.08	-1042.32	1417.96	304.45	1875.33	-23.57
pp_total_raw	-0.03	NA	NA	-0.03	NA	0.02
PreK.12	1105.58	852.77	398.34	1049.45	936.48	2182.9
Higher.Ed	-32.56	-7.99	-34	-29.33	-27.04	-6.17
Corporate	675.34	685.38	330.33	367.4	489.21	991.62
Suburb	215.06	-442.64	251.09	-37.26	247.23	-340.19
Rural	12.18	-218.2	-15.62	-79.88	NA	NA
LC	-304.94	-216.45	-61.98	-401.42	-343.25	-730.2
CM	-403.15	-450.23	-203.3	-451.33	-439.59	-946.44
SDO	142.83	41.61	-19.65	251.19	148.57	432.54
poly(diff_time, 3)1	6436.3	-665.3	5096.46	9360.08	-2759.99	10654.4
poly(diff_time, 3)2	7151.76	1904.3	3241.55	5907.65	2416.4	10205.69
poly(diff_time, 3)3	10558.95	6569.87	996.44	899	3301.77	3807.29
t	-162.14	-158.44	-49.23	-68.65	-24.41	-380.1
Interaction	927.75	NA	-5206.14	NA	NA	NA

Table 8: Rank

Name	Rank by category
pp_total_raw	6>1>2=5>3>4
treatment effect in pct_access regression	6>1>4>2>3>5
treatment effect in engagement_index regression	6>1>2>4>3>5

to have limited access to online learning owing to an equipment problem.Districts in category4 have the lowest per-pupil total expenditure but the percentage of minorities and children who eligible for free or reduced-price lunch of those districts are over average.

5 Discussion

In this paper, we mainly use regression with discontinuity design to obtain the treatment effect of face mask mandate on children's participation in online learning. The strength is that we take the discontinuity in the covariates into account. Moreover, we compare the effect among districts with different school expenditure and children population structure to evaluate the influence of COVID-19 pandemic on education inequity. Based on the data analysis above, we could conclude that the face mask mandate hold back children's participation once it is issued.

A potential threat to our research design comes from the possibility of discontinuities at the cutoff in other forms of COVID-19 related factors, which we do not include in our data. Besides, districts related variables in our data do not change along time but during the pandemic more people are allowed to work from home and thus consider to move to other cities. This leads to a change in the children structure. We can improve the performance of our model by applying time series of these variables.

References

[CCFT19] Sebastian Calonico, Matias D. Cattaneo, Max H. Farrell, and Rocío Titiunik. Regression Discontinuity Designs Using Covariates. *The Review of Economics and Statistics*, 101(3):442–451, July 2019.

[PN19] Sida Peng and Yang Ning. Regression Discontinuity Design under Self-selection. *arXiv:1911.09248 [econ, stat]*, November 2019. arXiv: 1911.09248.