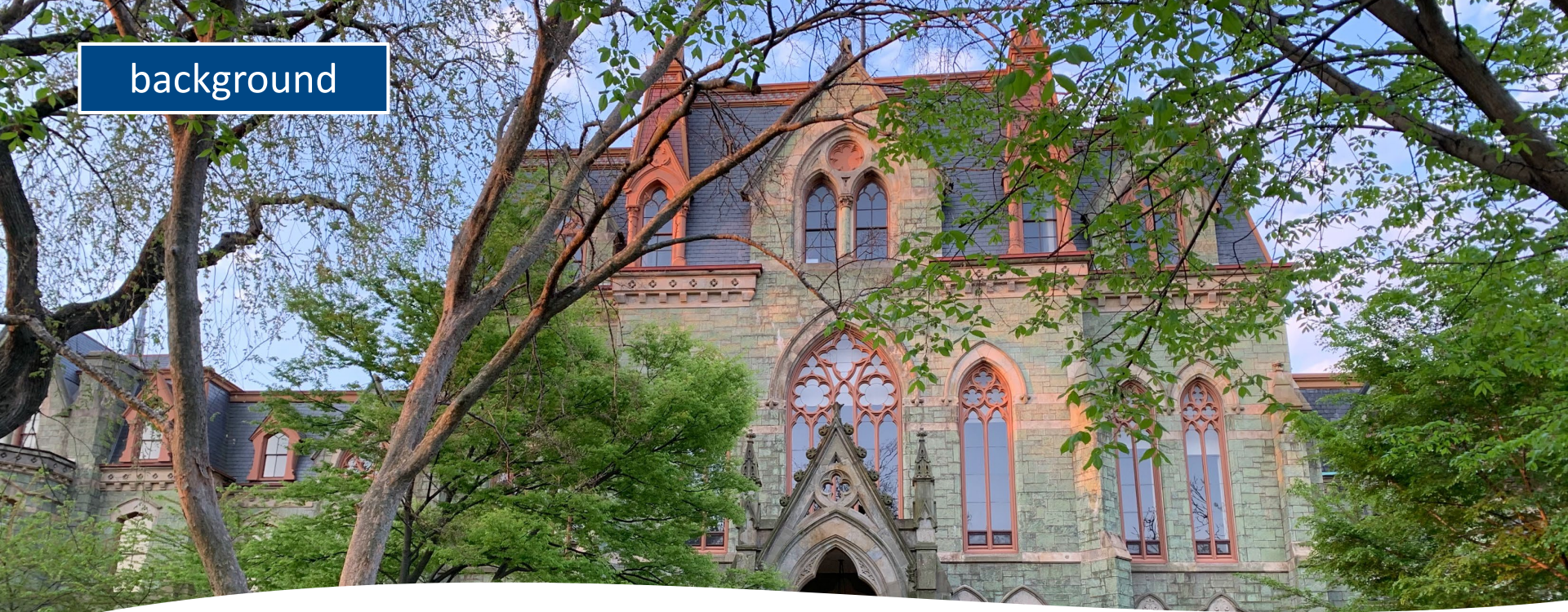# WHARTON
UNIVERSITY of PENNSYLVANIA

Impact of Universities' Online Covid-19 Dashboards on County Covid Cases

Alex Coble & Joe Moran

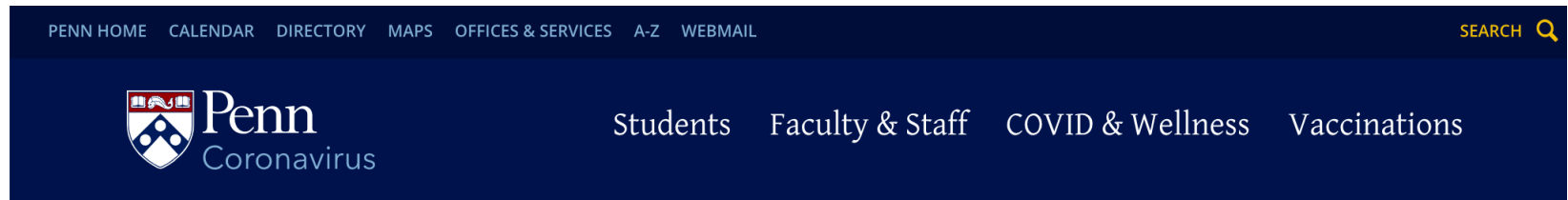STAT 571 | Data Science Live | April 30, 2021

Covid-19's Effect on Universities

- The onset of the pandemic forced many universities to close their doors to students in March 2020 across the nation and worldwide

- As the pandemic continued, a large variation in university policies for the Fall 2020 semester emerged. Some schools offered courses in person while others adopted a hybrid format or moved fully online.

University Covid Dashboards

- Many universities created webpages completely dedicated to tracking and monitoring the number of covid cases among their population of students, faculty, and staff ("dashboards")
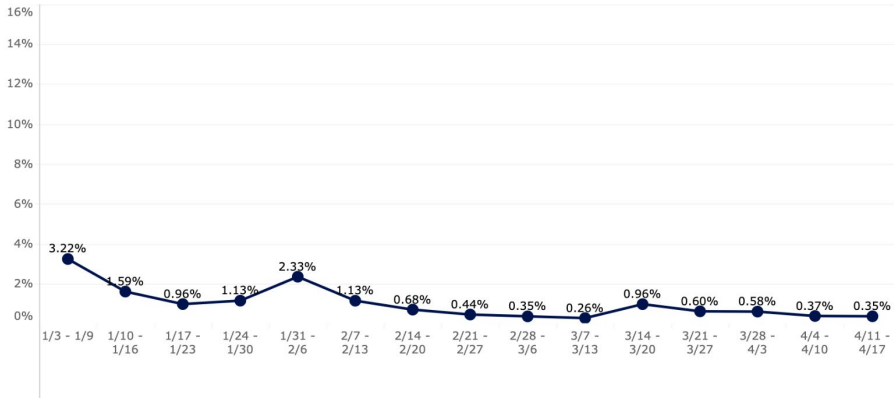
U Penn's dashboard: https://coronavirus.upenn.edu/content/dashboard

PENN HOME   CALENDAR   DIRECTORY   MAPS   OFFICES & SERVICES   A-Z   WEBMAIL   SEARCH 🔍

Penn Coronavirus   Students   Faculty & Staff   COVID & Wellness   Vaccinations

COVID-19 Dashboard

Announcements | President's Messages | **Dashboard** | FAQs | Contact | PennOpen Pass

The University collects and shares aggregate results of faculty, postdocs, students, and staff who are on campus regularly and participate in the Penn Cares program.
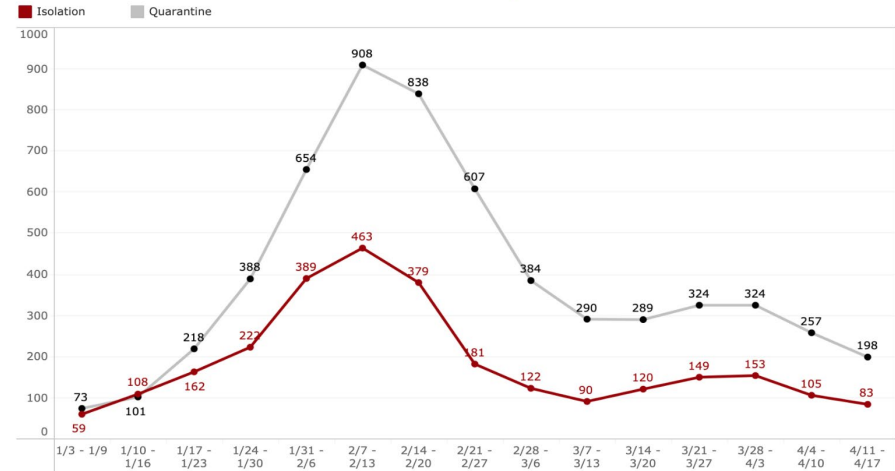
The Penn Cares program is informed by the University's profound respect for the impact its public health initiatives have on the health and safety of our friends, families, and neighbors.

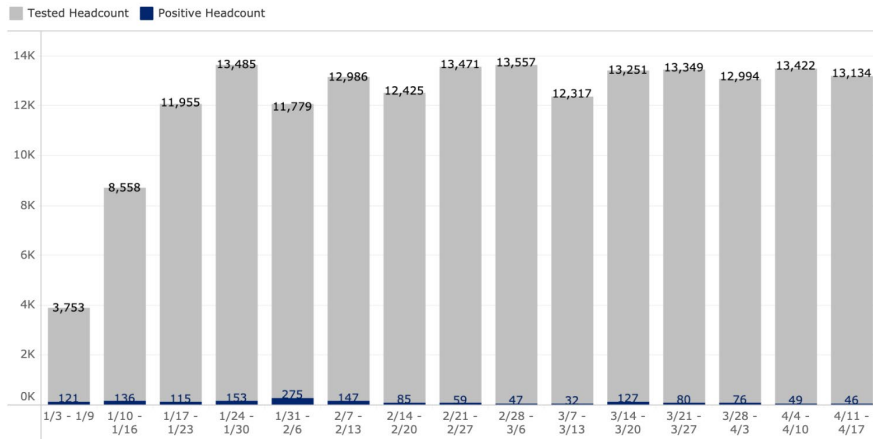Wharton

Positivity Rate
7-day average
Total

Values shown: 3.22%, 1.59%, 0.96%, 1.13%, 2.33%, 1.13%, 0.68%, 0.44%, 0.35%, 0.26%, 0.96%, 0.60%, 0.58%, 0.37%, 0.35%

Students in Isolation or Quarantine
Headcounts by Week

Legend: ■ Isolation   ■ Quarantine

Quarantine: 73, 101, 218, 388, 654, 908, 838, 607, 384, 290, 289, 324, 324, 257, 198
Isolation: 59, 108, 162, 222, 389, 463, 379, 181, 122, 90, 120, 149, 153, 105, 83

Tested and Positive Headcounts by Week
Total

Legend: ■ Tested Headcount   ■ Positive Headcount

Tested Headcount: 3,753, 8,558, 11,955, 13,485, 11,779, 12,986, 12,425, 13,471, 13,557, 12,317, 13,251, 13,349, 12,994, 13,422, 13,134
Positive Headcount: 121, 136, 115, 153, 275, 147, 85, 59, 47, 32, 127, 80, 76, 49, 46

Screening Test Site Visits
Number of Tests by Hour
4/11/2021 - 4/17/2021

Legend:
■ Ace Adams South Field
■ Annenberg Center
■ Du Bois/Rodin Field - West
■ Houston Hall
■ Richards Plaza
■ Gimbel Pottruck (Vaccine Clinic)

4
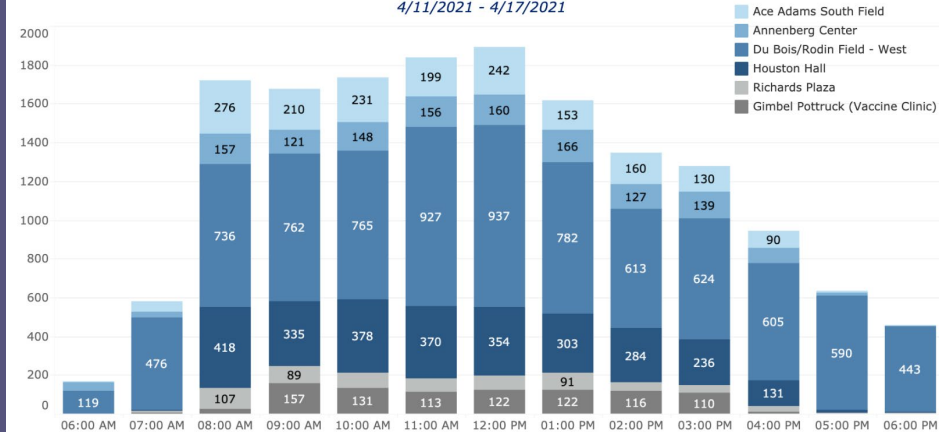
Variation in Dashboards

- Large variation in the information content and quality of disclosure in these covid dashboards across universities

- U Penn's dashboard is quite extensive, but some universities have dashboards with very limited information (or may not possess a dashboard at all)

- Columbus State University's (GA) dashboard: https://www.columbusstate.edu/covid/self-report-tracking.php

As of **April 19** the university has **9 current active positive cases of COVID-19 that have been self-reported by students and employees**.

In addition, and to provide a full snapshot, we have added the number currently in quarantine or isolation, but that have not tested positive for COVID-19. These include:

- **Active Exposure Cases:** 19 students and employees **who have reported being exposed** to someone exhibiting COVID-19 symptoms or have a positive test, but the exposed individuals have not received a positive test themselves.
- **Active Symptomatic Cases:** 6 students and employees **who have reported COVID-19-type symptoms** but have not received a positive test result.

Each of these individuals is working directly with a member of the COVID-19 Response Team to determine the appropriate time for taking a COVID-19 test and their projected return to campus.

Research Questions

1) How do changes made to university covid dashboards affect the spread of covid-19 in the surrounding county?
   → Event study

2) Do universities that track cases on covid dashboards do better (i.e. have fewer cases) than universities that do not once the Fall 2020 semester begins?
   → Difference in differences

Wharton

Our Datasets
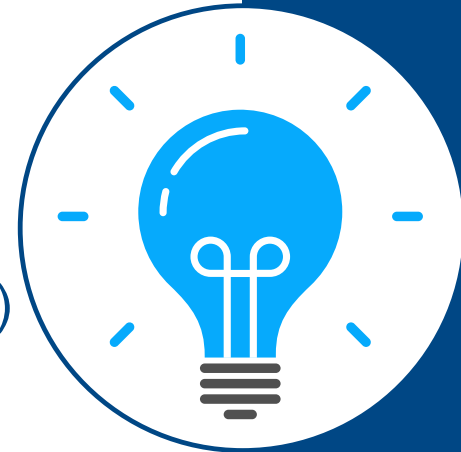
- List of 598 U.S. universities conducting operations in person or using a hybrid model during Fall 2020 indicating the county in which it resides and the school's covid plan || Source: College Crisis Initiative

Wharton

Our Datasets
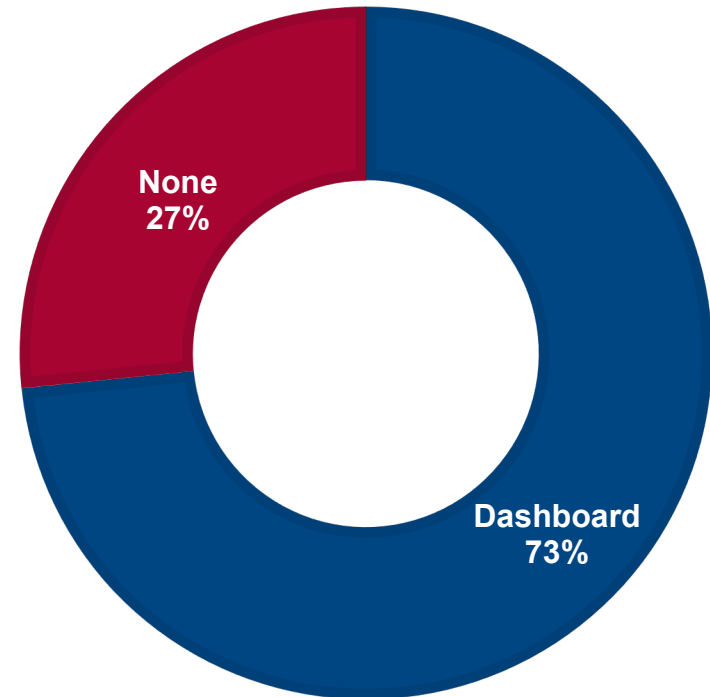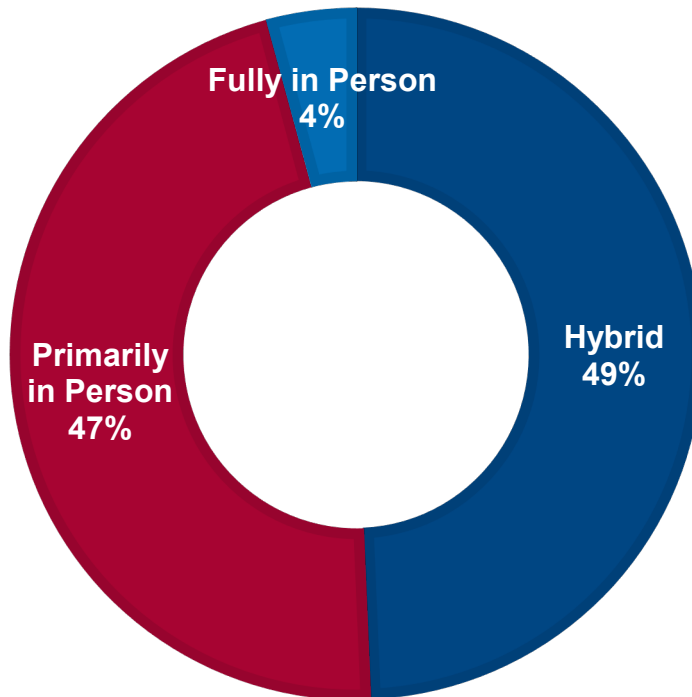
- Website URLs of the online covid dashboards for each university (if applicable) **||** Source: Hand collected

- Date instruction began in Fall 2020 for each university conducting operations in-person **||** Source: Hand collected

- U.S. daily covid cases by county for the period June 1, 2020 to November 24, 2020 **||** Source: New York Times

- University-specific attributes as of Fall 2019 (e.g., graduation rate, enrollment, school type) **||** Source: IPEDS

Wharton

Logit Model

What factors influence a university's decision to provide a covid dashboard?

- We used LASSO to produce a sparse model that identifies important factors related to a university's decision to provide a covid dashboard.

- We then use the 10 variables LASSO identified as our control variables in our subsequent tests (incl. one categorical variable not shown here).

Dependent variable: Indicator variable equal to 1 if the university provides a covid dashboard (0 otherwise)

Independent variables: Various university specific factors available on College Crisis Initiative

Classification Model, Dashboard Adoption

	Dependent variable:
	dashboard_indicator
PLANHybrid	0.594
	(0.496)
PLANPrimarily in person	0.750
	(0.491)
Four_year_school	1.138***
	(0.328)
Private_school	-1.198**
	(0.472)
Academic_programs_offered	15.214
	(619.810)
Secondary_programs_offered	-0.423
	(0.473)
Religious_school	-0.279
	(0.459)
Total_enrollment_2019	0.00003
	(0.00002)
Full_time_retention_rate_2019	0.014
	(0.020)
Graduation_rate_2019	0.057***
	(0.014)
Constant	-18.683
	(619.811)
Observations	558
Log Likelihood	-220.714
Akaike Inf. Crit.	463.428
Note:	*p<0.1;* **p<0.05;** p<0.01

Wharton

Measuring Changes Made to Dashboards

We first used the Wayback Machine Internet Archive (available at archive.org) to retrieve web scrapes of historical HTML versions of each university's covid dashboard.
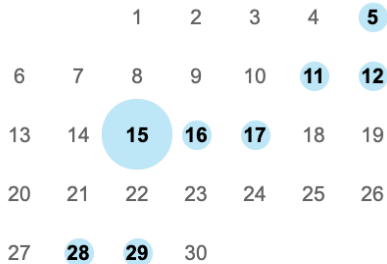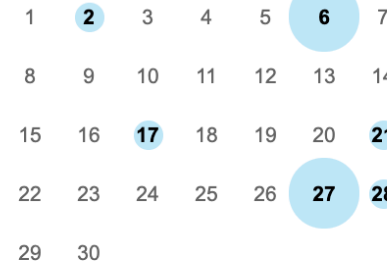

Days on which U Penn's covid dashboard was scraped by the Wayback Machine

Wharton

Measuring Changes Made to Dashboards

- We then created 5 measures of the informativeness (as proxied by the amount of HTML content) of university dashboards:

1	Number of characters
2	Number of hyperlinks (linked webpages)
3	Number of rich content nodes
4	Number of dynamic content nodes
5	Number of formatting nodes

Combined all 5 measures into a single factor using PCA

- We then calculate the change in the PC1 factor between web scrape dates for each university dashboard.

- We then create a series of indicator variables to identify dates representing PC1 factor changes in the top 10%, 20%, 25%, 33%, and 50% quantile of the distribution.

Wharton

7- and 14-Day Event Study

- **Independent variable:** Indicator variable which equals 1 during the [0, +7] day period after a change in the PC1 factor in either the top 10%, 20%, 25%, 33%, or 50% quantile

- **Dependent variable:** Log of daily new county cases per 100k

- **Fixed effects:** County (FIPS), Month

- **Other controls:** Average cases during the [-3, -1] day period & relevant university specific variables identified using LASSO

- **Results:** The [0, +7] day period following dashboard changes above the median (top 50%) and in the top tercile (top 33%) are associated with higher county case rates

7-Day Event Study Models, with FIPS and Month Fixed Effects

	Dependent variable:				
	log_daily_new_cases_per100k				
	(1)	(2)	(3)	(4)	(5)
PC1_top50_7	0.028**				
	(0.013)				
PC1_top33_7		0.024*			
		(0.013)			
PC1_top25_7			0.012		
			(0.014)		
PC1_top20_7				0.016	
				(0.015)	
PC1_top10_7					0.026
					(0.020)
Observations	88,995	88,989	88,963	88,951	88,893
R^2	0.268	0.268	0.268	0.268	0.268
Adjusted R^2	0.264	0.264	0.264	0.264	0.264
F Statistic	901.373*** (df = 36; 88532)	901.113*** (df = 36; 88526)	899.644*** (df = 36; 88500)	898.918*** (df = 36; 88488)	897.559*** (df = 36; 88430)

Interpretation of Results

- Increases in the information content on university covid dashboards are associated with higher case rates in the following 1-2 weeks.

  - ❖ Perhaps a reverse causal mechanism is at play here. We could observe this effect if, in areas where cases are on the rise, universities are motivated to enhance their web disclosure.

  - ❖ More information does not necessarily yield better results; better web disclosure does not lead to lower covid cases.

  - ❖ Universities need to be proactive in managing covid rates through policies and procedures. Information disclosure is not enough.

Wharton

Difference in Differences

Covid cases = α + β_1(Post) + β_2(Treatment) + β_3(Post*Treatment)

- **Dependent variable:** Log of daily new county cases per 100k

- **Post:** Indicator variable equal to 1 for all days after the first day of the Fall 2020 semester

- **Treatment:**

1) Indicator variable equal to 1 if the university has a covid dashboard

2) Indicator variable equal to 1 for universities with high information content dashboards (as proxied by median PC1's above the population median)

3) Indicator variable equal to 1 for universities with low information content dashboards (as proxied by median PC1's below the population median)

4) Indicator variable equal to 1 for dashboards with dynamic HTML charts

- **Fixed effects:** County (FIPS)

- **Other controls:** Relevant university specific variables identified using LASSO

Difference-in-Differences Models, Dashboard Adopters, with FIPS Fixed Effect

	Dependent variable:			
	log_daily_new_cases_per100k			
	(1)	(2)	(3)	(4)
post	0.333***	0.606***	0.622***	0.633***
	(0.015)	(0.008)	(0.008)	(0.007)
Dashboard_Indicator	-0.194***			
	(0.028)			
median_PC1_indicator		-0.052***		
		(0.020)		
below_median_PC1_indicator			-0.028	
			(0.023)	
chart_indicator				-0.035
				(0.027)
post:Dashboard_Indicator	0.394***			
	(0.017)			
post:median_PC1_indicator		0.111***		
		(0.015)		
post:below_median_PC1_indicator			0.054***	
			(0.015)	
post:chart_indicator				0.048**
				(0.022)
Observations	97,465	97,465	97,465	97,465
R^2	0.084	0.080	0.079	0.079
Adjusted R^2	0.080	0.075	0.075	0.075
F Statistic (df = 32; 97006)	279.472***	262.416***	261.081***	260.805***

Note: $p<0.1$; **$p<0.05$;** $p<0.01$

Wharton

Interpretation of Results

- Again, we see that more information does not necessarily yield better results.

- Universities that have covid dashboards (and dynamic HTML content therein) do not do better (in terms of covid cases) than universities that do not have dashboards once the Fall 2020 semester begins (in the post period).

- This could imply that universities need to be proactive in managing covid cases through policies and procedures as information disclosure does not seem to be enough.

Difference-in-Differences Models, Dashboard Adopters, with FIPS Fixed Effect

	Dependent variable:			
	log_daily_new_cases_per100k			
	(1)	(2)	(3)	(4)
post	0.333***	0.606***	0.622***	0.633***
	(0.015)	(0.008)	(0.008)	(0.007)
Dashboard_Indicator	-0.194***			
	(0.028)			
median_PC1_indicator		-0.052***		
		(0.020)		
below_median_PC1_indicator			-0.028	
			(0.023)	
chart_indicator				-0.035
				(0.027)
post:Dashboard_Indicator	0.394***			
	(0.017)			
post:median_PC1_indicator		0.111***		
		(0.015)		
post:below_median_PC1_indicator			0.054***	
			(0.015)	
post:chart_indicator				0.048**
				(0.022)
Observations	97,465	97,465	97,465	97,465
R^2	0.084	0.080	0.079	0.079
Adjusted R^2	0.080	0.075	0.075	0.075
F Statistic (df = 32; 97006)	279.472***	262.416***	261.081***	260.805***

Note: $p<0.1$; **$p<0.05$;** $p<0.01$

15

Wharton

Thank you!

Any questions?