COVID-19 and College: What Is Driving Decreased College Enrollment During the Coronavirus Pandemic?

Sean McGoey
Philip Merrill College of Journalism
University of Maryland
College Park, Maryland, USA
smcgoey@umd.edu

ABSTRACT

The COVID-19 pandemic has turned the world of education on its head, shutting down campuses worldwide as a precaution against transmission of the novel coronavirus and introducing a generation of students to remote learning.

The effects have extended all the way to the collegiate level, where the out-of-classroom development that comes from many students' first experience living outside the homes they grew up in can have nearly as large an impact as the educational aspect itself.

College enrollments have fallen for nine consecutive years [1], and the National Student Clearinghouse's data demonstrates that undergraduate enrollment fell by 4% in fall 2020 — more than double the size of decreases in previous years [2].

And even though distribution of vaccines has started in a handful of countries, coronavirus cases and deaths are spiking in the United States, forcing even many campuses that had reopened for some amount of in-person instruction to shut their doors again.

I set out to assess whether the pandemic might drive a continued stretch of decreases in enrollment, or whether there were other factors to which the declines could be attributed.

LITERATURE REVIEW

Given the relative newness of the coronavirus pandemic — we are still currently in the midst of the first academic year that has started since the pandemic started to spread in the United States in March 2020 — there is little or no research that has been done specifically into how COVID-19 has impacted the demand for postsecondary education. In fact, fairly little research is approached from the point of view of the students making decisions on whether to enroll.

Lin and Liu's 2019 paper on how delaying college enrollment impacts future earnings for students [3] takes the enrollment decision as the catalyst to examine earnings trajectories rather than seeking to understand the catalysts for a decision to delay enrollment.

Hemelt and Marcotte's paper on how tuition increases at four-year public universities affects students' propensity to enroll [4] looks at the enrollment decision, but specifically through the lens of increasing tuition, not accounting for outside factors that may impact the decision.

Timothy Meyer's 2017 dissertation on enrollment change during the Great Recession of 2008–2009 [5] addresses how at "non-selective" colleges, which would otherwise fit the definition of luxury goods quite well, enrollment actually increased during a period of great economic turmoil, similar to what has taken place during the pandemic.

And Abel and Deitz's 2014 paper on whether the benefit of college still outweighs the rising cost of attendance [6], which Meyer refers to in his dissertation, finds that when economic conditions are poor, the opportunity cost of attending college decreases (since the benefits of working are muted by lower wages, fewer opportunities, etc.).

But neither Meyer nor Abel and Deitz get at the heart of what is happening as part of the COVID-19 pandemic. While the Great Recession was unquestionably an earth-shaking period of economic uncertainty, the pandemic has fundamentally altered what kind of experience colleges can offer, which surely has additional impact on whether or not a student decides to enroll.

METHODOLOGY

Data Gathering and Cleaning

The three primary data sets I used to perform this analysis were: FAFSA application volume reports from the Department of Education's website (with two subsets: applications by state, and applications by high school); The New York Times' COVID-19 tracking data, maintained on the Times' GitHub page; and data from the U.S. Census Bureau's annual American Community Survey.

I downloaded the FAFSA files directly from the Department of Education's studentaid.gov site and did some pre-cleaning to strip out formatting, document title rows, etc. (and, in the case of the high school data, preserving only rows from Maryland high

schools). Then I imported four files into R Studio: one file containing Maryland applications by high school for both the 2019–2020 and 2020–2021 application cycles, one file containing ZIP code and county information for Maryland high schools (also from the Department of Education's website) and one file each containing applications by state (including U.S. territories and foreign countries) for the 2019–2020 and 2020–2021 application cycles.

After the import, some more cleaning was required to standardize column names, join the school-level applications to the ZIP/county data, and combine the two state-level files into one table that would show applications by year for each state.

For the American Community Survey data, I used the **tidycensus** R library to pull in data on racial demographics (specifically, Black, Hispanic and white populations), home ownership vs. renting, and household income (particularly incomes below $50,000 per year or above $200,000 per year), both by state (for the entire United States) and by ZIP code (for Maryland specifically).

I then joined those tables to the application tables to get a table with applications and census data for every ZIP code in Maryland and one for every state in the U.S. At this point, I also used the county information contained in the Maryland ZIP table to make a third object for examination, one rolled up to the county level for Maryland.

Finally, I wrote scripts to read in The New York Times's automatically-updated COVID-19 data from GitHub (which shows the accumulation of cases and deaths by date for each county in the U.S.) and clean it so that there would be one row for each county, containing that county's peak level of COVID-19 cases. I then grouped that data and joined it to the Maryland county table and the state-level table (unfortunately, there was no COVID-19 data available at the ZIP level).

Solution Framework

The Department of Education maintains publicly available reports tracking submissions of the Free Application for Federal Student Aid (FAFSA), which enrolling students must complete in order to receive government financial aid for postsecondary education.

Using those submissions as a proxy for interest in college enrollment, I ran paired t-tests to determine whether there was a meaningful difference in FAFSA applications through October 23, 2019 (for the previous application cycle) and October 23, 2020 (for the current application cycle), looking at both the United States (on a state-by-state basis) and at state of Maryland (by county or by ZIP code).

After finding that there was a significant difference in applications at the state level, I used coronavirus tracking information from The New York Times and data on racial demographics, household income and home ownership from the U.S. Census Bureau to determine whether any of those factors was driving the decrease in applications — using the **cor.test** function in R to perform a linear regression on the change in applications as a function of the variable in question.

EXPERIMENT RESULTS

Testing to determine whether there was a significant change in application volume

After the data was imported and cleaned, the first step I needed to take was to determine whether there was in fact a statistically significant difference in the volume of applications submitted from last year to this year. Without that, any other analysis would have been speculative at best.

I calculated the applications per 10,000 residents for each state/county/ZIP code to remove population-size effects, then used the **t-test()** function in R's **stats** package to test the assumption that the difference in mean value of each list was not 0. For Maryland ZIP codes, the test resulted in

$$\boldsymbol{p = 0.4467},$$

indicating that there was not a statistically significant difference in applications at the ZIP code level in Maryland. For Maryland counties, the test returned a value of

$$\boldsymbol{p = 0.05506},$$

just barely missing the threshold for significance. And at the state level nationwide, the test returned an average decrease of roughly 11 applications per 10,000 residents and a p-value of

$$\boldsymbol{p = 0.04 * 10^{-10}},$$

finding that there was indeed a statistically significant decrease in applications at the state level and making state-level trends the primary area of focus for the rest of the analysis.

3 Testing for correlation between selected variables and application volume

After finding that there was a significant year-over-year difference in application volume by state, I used the other data I had gathered to look for correlation between certain variables and the percentage change in applications in each state. In each case, I

The first variable I tested was a state's peak number of COVID-19 cases per 10,000 residents. But there was no meaningful correlation between COVID-19 cases and a drop in applications.

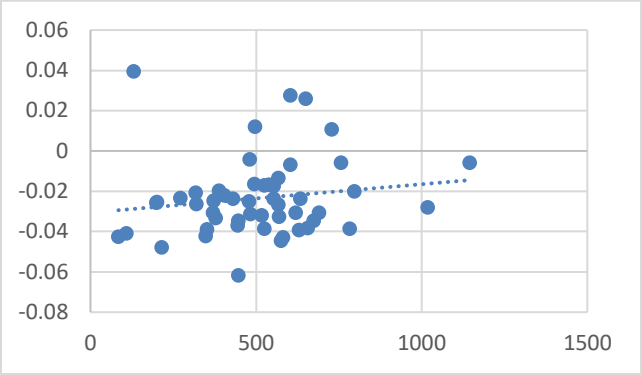


Figure 1: Change in FAFSA applications as a function of a state's peak COVID-19 cases per 10,000 residents (p = 0.29).

With no meaningful relationship between COVID-19 cases and drops in FAFSA applications, I sought to determine whether there were other factors that correlated strongly with applications. The next variable I tested was household income, particularly at the extreme ends of the distribution. I hypothesized that states with higher percentages of households making **less than $25,000 per year** would be likely to see larger drops in applications and that states with larger shares of households making **over $200,000 per year** would see smaller drops — or possibly even increases — in applications.

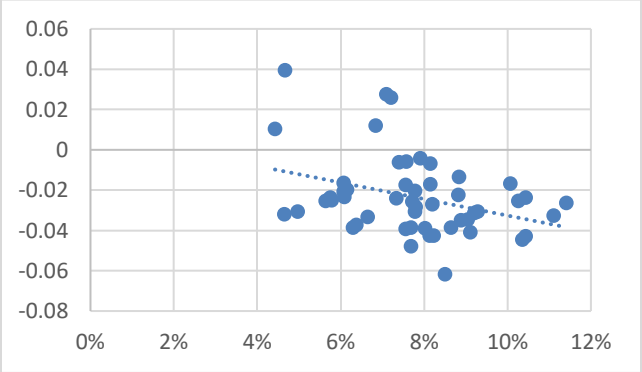
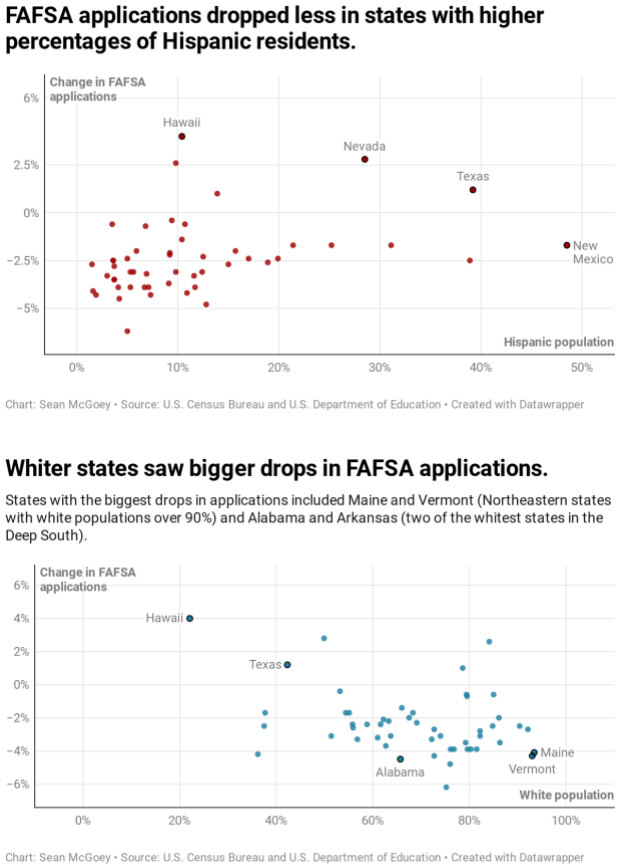


Figure 2: Change in FAFSA applications as a function of a state's percentage of households making < $25,000 (p = 0.012).

There was not a significant correlation between household incomes above $200,000 and change in application volume, but there was a significant (if not terribly strong) relationship between a state's percentage of households making less than $25,000 and larger drops in applications, suggesting that students from families already struggling economically may have been less likely to seek postsecondary education during the pandemic.

Then, I tested the relationship between states' white, Black and Hispanic populations and applications. There has been a great deal of reporting on how the pandemic has disproportionately affected minority communities, so even if there is not a clear statistical link between COVID-19 and applications, we might expect to see bigger drops among these populations.



Figures 3 & 4: Change in FAFSA applications as a function of a state's Hispanic (p = 0.009) and white populations (p = 0.013).

There was no significant connection between a state's Black population and drops in applications, but more heavily Hispanic states actually experienced *smaller* drops in applications. And whiter states saw bigger drops in applications; the list of the states with the 10 biggest decreases in applications includes Maine and Vermont (two far Northeastern states with 93% white populations) and Alabama and Arkansas (two of the whitest states in the Deep South).

Finally, I tested for relationships between home ownership vs. renting and changes in application, but there were no significant correlations.

FINDINGS AND CONCLUSIONS

Findings

My analysis found that there was indeed a statistically significant drop in FAFSA applications at a state level from the 2019–2020 application cycle to the 2020–2021 cycle, though not at a county or ZIP code level for the state of Maryland.

Household income below $25,000 a year was positively correlated with larger drops in FAFSA applications, suggesting that students from already economically strained may be more likely to delay or forgo college enrollment during the pandemic. And in an interesting discovery, states with larger Hispanic populations avoided the largest drops in applications, consistent with data from the National Student Clearinghouse showing that Hispanic students have seen some of the smallest declines in enrollment in 2020 [2].

There were no meaningful correlations between COVID-19 cases, household income > $200,000, white population, black population, home ownership, or home renting and a state's year-over-year change in FAFSA applications.

Limitations

One of the most obvious limitations of this analysis is that FAFSA applications are an imperfect proxy for interest in college enrollment. There are any number of reasons why a student might not fill out a FAFSA form at all — or simply after the window of time that I was able to examine.

Future Work

The primary logical extensions of this work would include replicating the analysis with actual enrollment numbers instead of FAFSA applications and incorporating other variables into the correlation analysis: public vs. private high schools, public vs. private colleges, etc.

REFERENCES

[1] Paul Fain, 2019. College Enrollment Declines Continue. *Inside Higher Ed* (May 30, 2019). https://www.insidehighered.com/quicktakes/2019/05/30/college-enrollment-declines-continue

[2] Todd Sedmak, 2020. Fall 2020 Undergraduate Enrollment Down 4% Compared to Same Time Last Year. National Student Clearinghouse Media Center (Oct. 15, 2020). https://www.studentclearinghouse.org/blog/fall-2020-undergraduate-enrollment-down-4-compared-to-same-time-last-year/

[3] Lin, Yuxin., & Liu, Vivian (2019). Timing Matters: How Delaying College Enrollment Affects Earnings Trajectories. *CCRC Working Paper No. 105 – Community College Research Center* (Feb 2019). https://files.eric.ed.gov/fulltext/ED595238.pdf

[4] Hemelt, S.W., & Marcotte, D.E. (2016). The Changing Landscape of Tuition and Enrollment in American Public Higher Education. *RSF: The Russell Sage Foundation Journal of the Social Sciences 2*(1), 42-68. doi:10.1353/rus.2016.0002

[5] Meyer, Timothy J., "A Quantitative Study of Enrollment Change during the Great Recession at Non-selective Small Private Colleges and Universities" (2017). In *Education Doctoral Dissertations in Leadership*. 94. https://ir.stthomas.edu/caps_ed_lead_docdiss/94

[6] Abel, Jaison R. and Deitz, Richard, Do the Benefits of College Still Outweigh the Costs? (August 1, 2014). *Current Issues in Economics and Finance*, Vol. 20, No. 3, 2014. SSRN: https://ssrn.com/abstract=2477864