IDS 701 UDS Final Report

Team 1

Elisa Y. Chen, Xiaoquan Liu, Genesis Qu

Stakeholder

The stakeholder of our study is the dean of a college who is deciding whether to advocate for the implementation of online instruction in the case of another potential lockdown.

Executive Summary

Understanding the impact online schooling has on the mental health of college students is vital for the university staff to help them take corrective actions to guarantee the well-being of students, especially in situations like a pandemic lockdown. With the rise of online instruction format and the fact that college students are more stressed and anxious than ever before, it is important to assess whether online schooling comes at the cost of students' mental well-being.

In our analysis, we study whether online / hybrid schooling causes higher anxiety levels among students undertaking post-secondary education in the US during the years 2021 to 2022. The aim of this analysis is to gain a deeper understanding of potential causes of higher anxiety levels among students and how US college administrators could address them.

We learned that controlling for differences between individual students, taking online / hybrid classes reduced students' anxiety during the pandemic compared to in-person classes for both full-time and part-time students. Additionally, whether a student is aware of mental health assistance offered by the school plays a large role in determining the anxiety level of the student. With the knowledge of where to access resources offered by their school, students were less anxious on average compared to students who did not have the knowledge. Otherwise, we did not observe any noticeable differences between online / hybrid and in-person students that would raise any practical concerns.

While our study suggests that online learning is adoptable, and even favorable to in-person learning, we note the inherent tradeoff that college administrators must make between reducing disease spread and ensuring student mental health when deciding on the teaching modality during a pandemic. This analysis addresses the latter part of the tradeoff and tries to quantify its impact using causal inference.

Introduction

College students today appear to be more stressed and anxious than ever before. This is a major problem affecting the student's mental and physical well-being, and greatly impacts the student's academic performance. According to a study of over 78,000 students in 19 countries conducted by the National Institutes of Health, 71% of the respondents indicated increased stress and anxiety due to the COVID-19 outbreak. The study identified several stressors, a major one being limited social interactions due to physical distancing, which was an issue for around 86% of respondents[1].

While the concept of distance learning isn't novel, it has never been as widely adopted prior to the COVID-19 outbreak, which abruptly resulted in a nationwide shift to an online instruction format in 2020. This forced many students to adapt to a new form of learning that physically isolated them from their usual support network in school. The benefits of online learning such as accessibility were quickly recognized by many students and academic staff to the extent that online / hybrid learning is becoming more common even in a post-pandemic world. Even though online education could be more engaging and accessible for many students, it is important to assess whether this comes at a cost of the well-being of students before adopting the instructional format.

Problem Statement

In this report, we aim to explore the effects of online schooling on the anxiety levels of college students in the US during the years 2021-2022. More specifically, the question we aim to answer is:

Did online / hybrid schooling cause higher anxiety levels among students in post-secondary education (undergraduate, graduate, Ph.D.) in the US during the years 2021 - 2022?

Answering the proposed causal question would shed more light on whether online schooling will increase the anxiety levels of college students in the US and by how much. In the case that another pandemic occurs, the debate to bring back online school models will need to be guided by statistical evidence answered using causal inference. This study aims to understand the impact of online learning during another potential pandemic on the well-being of students, accounting for the availability of mental health resources. In an ideal world, we would conduct a randomized experiment in which students would be randomly assigned a class instructional format from a controlled pool for us to understand differences in student anxiety levels caused by online schooling. However, this is not plausible to execute practically, so we consider alternative methods to mirror the randomized environment the best we could.

Data

The data we are using to understand the above question comes from the Healthy Mind Study (HMS), which is a web-based survey of questions related to mental health, and utilization of mental health services among higher education students (undergraduate, graduate, Ph.D.) in the U.S. The data is published on an annual basis by the Healthy Mind Network, a team of scholars in the U.S., to understand mental health-related issues among college students.

[1] Xu T, Wang H. High prevalence of anxiety, depression, and stress among remote learning students during the COVID-19 pandemic: Evidence from a meta-analysis.

The HMS reached out to a random sample of students at each participating school via email to invite them to participate in the survey. The 2021-2022 data[2] contained survey responses from 95,860 students from 102 schools nationwide in 34 states. The 2021-2022 survey contains recorded data spanning from September 2021 to the end of May 2022, when most schools are in session. It is important to note that for anonymity purposes, the survey masked the institutions' names in the data, and the only available geographical identifier of the student's institution is the state. HMS did not disclose their non-response rates, which varied by school. Please see Appendix Section 1 for more details.

The survey data had in total 1,712 fields, which include the time of response, respondent ID, demographic variables, measurements of mental health status, and utilization of mental health resources at students' home institutions and beyond.

To assess students' anxiety levels, the survey used GAD-7[3], a screening tool to measure short-term generalized anxiety orders, and summed up the individual scores. The anxiety score variable ranges from 0 to 21, with 0 being no anxiety, and 21 being extreme anxiety. We use this field as our outcome variable.

The survey also recorded class formats during the COVID-19 pandemic, classified into four categories: fully in-person, online, hybrid – a combination of in-person and online instruction, or other formats. We re-encoded this field into a binary variable where students who attended either hybrid or online classes are part of the treatment group, and students who attended classes in-person only are part of the control group. The "other" group, with few samples, was removed from the data.

We conducted a priori variable selection on the 1,712 data columns to filter for fields that we believed empirically would be related to our treatment variable as well as our outcome variable. We selected variables that are highly correlated with the variables of interest resulting in 26 potential candidates for matching. To create a more robust matching process, we identified a parsimonious set of variables by using a correlation matrix to determine whether any correlated features could be left out among the 26 variables. Please see Figure 1 in Appendix for more details. To account for omitted variable bias, we aimed to include variables that could impact a student's stress level that could also play a factor in whether a student is attending classes in-person or online. Please see Table 1 in Appendix for the summary statistics for a number of example variables that we wanted to control for.

Ultimately, we chose 80 variables on students' sex, gender, sexuality, race, socioeconomic background, academic standing, and a suite of variables surveying mental health and access to mental health services.

[2] Healthy Minds Network (2023). Healthy Minds Study among Colleges and Universities, year (2021-2022).
[3] GAD-7 Screening Questionnaire.

To minimize baseline differences between students attending online and hybrid classes, and students who attended in-person classes, we apply a technique called matching to ensure that both treatment and control are comparable.

The resulting matched dataset contained 36,944 rows, which is less than half of the original data but more balanced in the features selected for matching. One of the major factors likely to influence the anxiety levels of students, and what we aimed to control for was the number of COVID-19 cases during the week in the state where the student attended college. We were able to reduce the baseline differences by reducing the disparity between the means and variances of the two groups. Please see Appendix Section 2 for more details on the matching methodology.

It is worth noting that there is always a trade-off between the degree of balance achieved and the number of samples when performing matching. We prioritized having a sufficient number of samples over a strictly balanced dataset for the analysis to obtain statistically meaningful results. Please see Table 2 in Appendix for the summary statistics of the key features after matching is applied.

Model

A regression is necessary to better understand the true effect of different instructional formats on students' anxiety levels because of the many confounding variables that can obscure the true effect of online instruction on anxiety. For example, some students in their final year of a program may have higher anxiety levels due to graduation, and some students may suffer from other mental health problems that may also have an impact on their anxiety levels.

To adjust for those baseline differences, we applied matching on the original dataset on a subset of key covariates and added the rest of the selected variables as controls in the regression. After matching on variables that are correlated with both the students' anxiety levels and instructional format, each observation is assigned a weight proportional to the size of the group that they were matched on, which reflects its relative importance and reliability in the analysis. To extract the true relationship between instructional format and students' anxiety levels, we turned to weighted linear regression models.

We fit four models to the matched data. The first model only considers the effects of instructional format on students' anxiety levels, without any covariates. The second model takes demographic and clinical variables into account. The third and fourth models are fitted with different interaction terms to control for potential heterogeneous treatment effects of online instruction by students' enrollment status and their knowledge of mental health assistance, respectively.

We became interested in these two key variables because, as shown in Figure 1 and Figure 2, the impact of the instructional format on students' anxiety levels seem to vary depending on the students' enrollment status and their knowledge of mental health assistance.

Figure 1. Anxiety level by Enrollment Status & Instructional Format

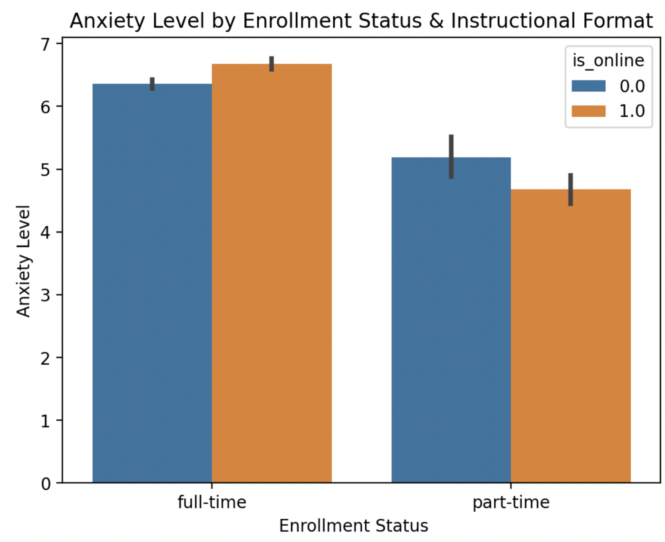
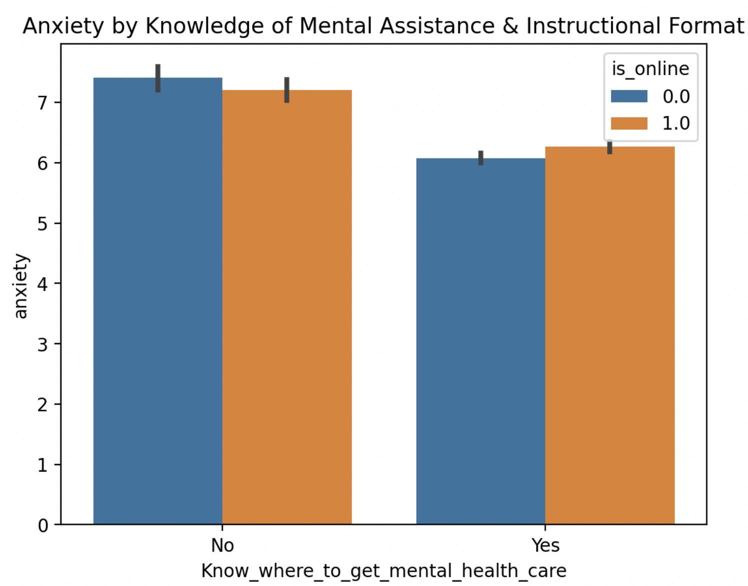


Figure 2. Anxiety level by Knowledge of Mental Assistance & Instructional Format



The resulting coefficients, standard errors, and p-values of the four models are shown below in Table 3.

In Table 3, the simple model shows that without considering covariates, online students have lower anxiety scores compared to in-person students. This effect is not statistically significant at a confidence level of 0.95. However, because our matching process yielded data with some baseline differences still, we will compensate for this bias by fitting additional models with covariates and interaction terms. Comparing across these models allows us to validate the robustness of the effect term in the simple model.

The second model takes students' clinical, demographic, and social variables into account. It validates the simple models' result that students receiving online instruction have lower anxiety scores compared to those receiving in-person instruction, holding all else constant. This difference is not statistically significant. Additionally, the model estimates that part-time students have an anxiety score of 0.22 lower than full-time students, which is statistically significant. Also, holding all else constant, students who know where to access mental health resources have around 0.40 less anxiety score compared to those who don't know where to access mental health resources.

After including different interaction terms to explore potential heterogeneous treatment effects of online instruction by students' enrollment status and their knowledge of mental health assistance in the third and fourth models, we find that, as shown in the third model, the impact of online treatment on anxiety levels for full-time students is different from that for part-time students, and the difference is not statistically significant. In addition, as shown in the fourth model, the impact of online treatment on anxiety levels for students who know where to access mental health resources is different from that for those who don't know where to access mental health resources, and the difference is statistically significant.

Table 3. Weighted Linear Regression Model Results. Variables shown in the table are only those we are interested in.

	Simple Model	Model with Covariates	Model with Interaction Term (is_online: enrollment)	Model with Interaction Term (is_online: know_where)
Intercept	6.6251*** (0.037)	5.2766*** (0.424)	5.2641*** (0.424)	5.3877*** (0.426)
is_online	-0.1359* (0.056)	-0.0556 (0.041)	-0.0428 (0.043)	-0.2535** (0.087)
enrollment status (part-time students)		-0.2234** (0.082)	-0.1623 (0.103)	-0.2263** (0.082)
know_where (to get mental assistance)		-0.3967*** (0.049)	-0.3964*** (0.049)	-0.5163*** (0.067)
is_online: enrollment status (part-time students)			-0.1356 (0.137)	
is_online: know_where (to get mental assistance)				0.2512** (0.096)
R-squared	0.000	0.514	0.514	0.514
Adj. R-squared	0.000	0.512	0.512	0.512

Discussion / Conclusion

The results of the analyses show that online students have lower anxiety levels than in-person students. Controlling for covariates decreases the magnitude of the impact of the online study on students but generally, online learning still decreases students' anxiety levels.

We believe that the result differs from the previous study reported by the National Institutes of Health potentially for the following two reasons: the first is that our study is limited to students in the U.S. while the sample of students taken by the NIH comes from 19 countries. There could be material differences between the U.S. and those countries during the pandemic that affected students' anxiety other than learning modality. Additionally, the implementation of online studies could look different in different countries. The second reason is that the HMS survey included students from a more diverse list of majors compared to the original NIH study, which did not include arts and humanities majors.

We hypothesize that during the COVID-19 pandemic, staying online helped the majority of students maintain peace of mind by reducing the risk of contact with COVID-19. For full-time students, online learning provides more flexibility in terms of time and location as well as a personalized learning environment, which can reduce the stress of commuting and allows students to work at their own pace. However, with anxiety levels being measured on a scale of 0 to 21, a change in the anxiety level of a magnitude of less than 0.5 is not practically significant.

Based on our results, students' enrollment status and knowledge of mental health resources also play important roles in defining students' anxiety levels. Part-time students have 0.22 lower anxiety scores compared to full-time students. With the knowledge of where to access resources offered by their school, their anxiety scores were 0.40 lower compared to students who do not have that knowledge.

Combined with online studying, we see that online treatment increases the anxiety level for students who know where to get mental health assistance, which is different from students who do not know where to get mental health assistance. One possible explanation based on our data is that since those mental health resources are from schools, most of them are onsite services. For students who know where to get mental health services, the reason they have that knowledge is because they know mental health services help them reduce anxiety levels. With online learning and social distancing, they are not able to seek mental help anymore; thus, their anxiety level increases.

For future potential lockdowns, our study suggests that online learning is adoptable, but college deans also need to take the accessibility of mental health resources into consideration. More online mental health resources could help with students' anxiety levels during lockdown periods.

Limitations

As presented throughout our analysis, there were a number of limitations that we want to call out for future studies. Firstly, the survey responses were self-reported meaning that the students could opt out of the survey if they wanted to. This could potentially introduce some degree of participation bias if students with anxiety opted out of the survey due to social stigma around mental health issues. Additionally, we observe a number of variables such as gender and education level that were disproportionately represented in the dataset and hard to balance even with the matching process. This could also be driven by non-response bias. The non-response weights added by HMS were used in their internal studies but were not useful for our analysis. While the responses were randomly sampled, we only obtained results from 34 states potentially leaving some states unrepresented in the results. However, the data does cover a wide range of states including inland and coastal areas.

We also attempted to account for the effects of COVID-19 rates on student's anxiety levels by controlling for the number of covid cases per 10,000 people in the state in which the student goes to school for a given month and week. However, it is possible that online students were completing their classes remotely out-of-state and were impacted by different covid case rates. Having said that, we're assuming that this is not the usual situation and wouldn't change the conclusions dramatically. Because of the context during which the responses were collected, our study is most easily extended to situations in which students are affected by another pandemic lockdown.

References

[1] Xu T, Wang H. High prevalence of anxiety, depression, and stress among remote learning students during the COVID-19 pandemic: Evidence from a meta-analysis. Front Psychol. 2023 Jan 10;13:1103925. doi: 10.3389/fpsyg.2022.1103925. PMID: 36704682; PMCID: PMC9871576.

[2] Healthy Minds Network (2023). Healthy Minds Study among Colleges and Universities, year (2021-2022). Healthy Minds Network, University of Michigan, University of California Los Angeles, Boston University, and Wayne State University. https://healthymindsnetwork.org/reserach/data-for-researchers

[3] GAD-7 Screening Questionnaire. https://adaa.org/sites/default/files/GAD-7_Anxiety-updated_0.pdf

Appendix

Section 1 - Non-Response Analysis
The below statement was released by HMS as part of their survey results to disclose how they addressed non-response bias:

A potential concern in any survey study is that those who respond to the survey will not be fully representative of the population from which they are drawn. In the HMS, we can be confident that those who are invited to fill out the survey are representative of the full student population because these students are randomly selected from the full list of currently enrolled students. However, it is still possible that those who actually complete the survey are different in important ways from those who do not complete the survey. It is important to raise the question of whether the percentage of students who participated are different in important ways from those who did not participate. The HMS address this issue by constructing non-response weights. The non-response weights adjust specifically for the fact that female students have consistently higher response rates than male students in our survey (and in most other survey studies). We construct the weights by comparing the female-male composition of our respondent sample to the reported female-male ratio for the full student population at each institution (which is typically available from basic enrollment statistics). If the respondent sample has a smaller percentage of males and larger percentage of females, as compared to the composition of the full student population, then male students in our sample are assigned a higher non-response weight value than female students. This means that weighted estimates are representative of the female-male distribution in the full student population. For students with nonbinary gender identities, we are not able to use this same process, however, because we are generally not able to obtain accurate statistics on the representation of these groups in the full student population. Therefore, rather than making assumptions, we assign a weight value to students with nonbinary identifiers that leaves their representation in the weighted sample the same as in the unweighted sample. In the future, if and when more reliable information becomes available *at the full student population level, we will be able to incorporate that information into sample weights for groups other than female and male gender identities. Finally, note that these sample weights give equal aggregate weight to each school in the national estimates. An alternative would have been to assign weights in proportion to school size, but we decided that we did not want our overall national estimates to be dominated by schools in our sample with very large enrollments.*

Table 1. The mean and standard deviation (in brackets) between Treatment and Control For Demographic & Mental Health Features on raw data. The p-values for the differences in frequencies and the variance are measured using a chi-squared and Levene's test (in brackets) respectively. We can observe that none of our key features are balanced with all p-values being statistically significant at 0.05 level.

Variable	Treatment	Control	P-Value (delta)
Age	25.00 (8.55)	22.00 (5.29)	0.000 (0.000)
Covid Cases per 10k populations	16.67 (21.75)	19.46 (27.03)	0.000 (0.000)
Sex (at birth)	1.31 (0.49)	1.25 (0.47)	0.000 (0.000)
Is Undergraduate Student	0.51 (0.50)	0.72 (0.45)	0.000 (0.000)
Is Full-Time Enrolled	1.23 (0.40)	1.07 (0.25)	0.000 (0.000)
Has Financial Stress	2.66 (1.09)	2.92 (1.14)	0.000 (0.010)

Section 2 - Matching Methodology

To perform the matching process, we applied the FLAME[4] algorithm on a subset of features (26 out of 80) including students' socioeconomic background, academic standing, enrollment status, existing mental conditions, and COVID-19 rates, which we considered to be most pertinent based on priori variable selection and in terms of correlation (>= 0.2 or <= -0.2) with treatment and outcome. To create a more parsimonious set of features, we also identified variables that could be omitted from the matching process using a correlation matrix. Please see Figure 1 below for more details.

[4] A package in Python for performing exact matching for observational causal inference on datasets containing discrete covariates.

We also coarsened our continuous features (hours_work_paid, age, covid_cases_per_10k, sleep_wknight) to accommodate for the discrete covariate requirement the FLAME algorithm mandates. The binning was performed based on the distribution of the covariate typically assigning a category of ‘high’ and ‘low’ for those values that were ~1.5 standard deviations above and below the median value respectively.

When measuring anxiety during the 2021-2022 pandemic, the number of covid cases is a major confounding factor that we wanted to strictly control for. We thought it was necessary for this feature to be perfectly (or close to perfectly) balanced to eliminate any confounding effects on anxiety levels, which is why we relied on exact matching for this analysis.

By default, FLAME will keep dropping matching variables iteratively until it has been able to match all the treated observations or run out of variables. However, we can control the number of iterations by telling it to stop manually. We determined the cutoff for the iterations by observing how the quality of matches evolves over iterations. Figure 3 shows for each iteration the prediction error (difference in the mean-squared error of regressing) that results from dropping the variables excluded in each step. Setting the cutoff at 8 iterations seemed appropriate as the predictive error increases more dramatically after that point. Also, The cutoff point at 8 iterations is still providing us with a sizable sample size to achieve statistically significant results.

Figure 1. Correlation matrix for the 26 variables considered for matching. The purpose of the correlation matrix is to help us identify the parsimonious set of variables used for matching. Any covariates with a correlation of >=0.5 or <=-0.5 was considered “highly correlated” with one another. We only included the variable which was more correlated with the output variable in the final matching process. In the end, we obtained 21 variables for the matching process.

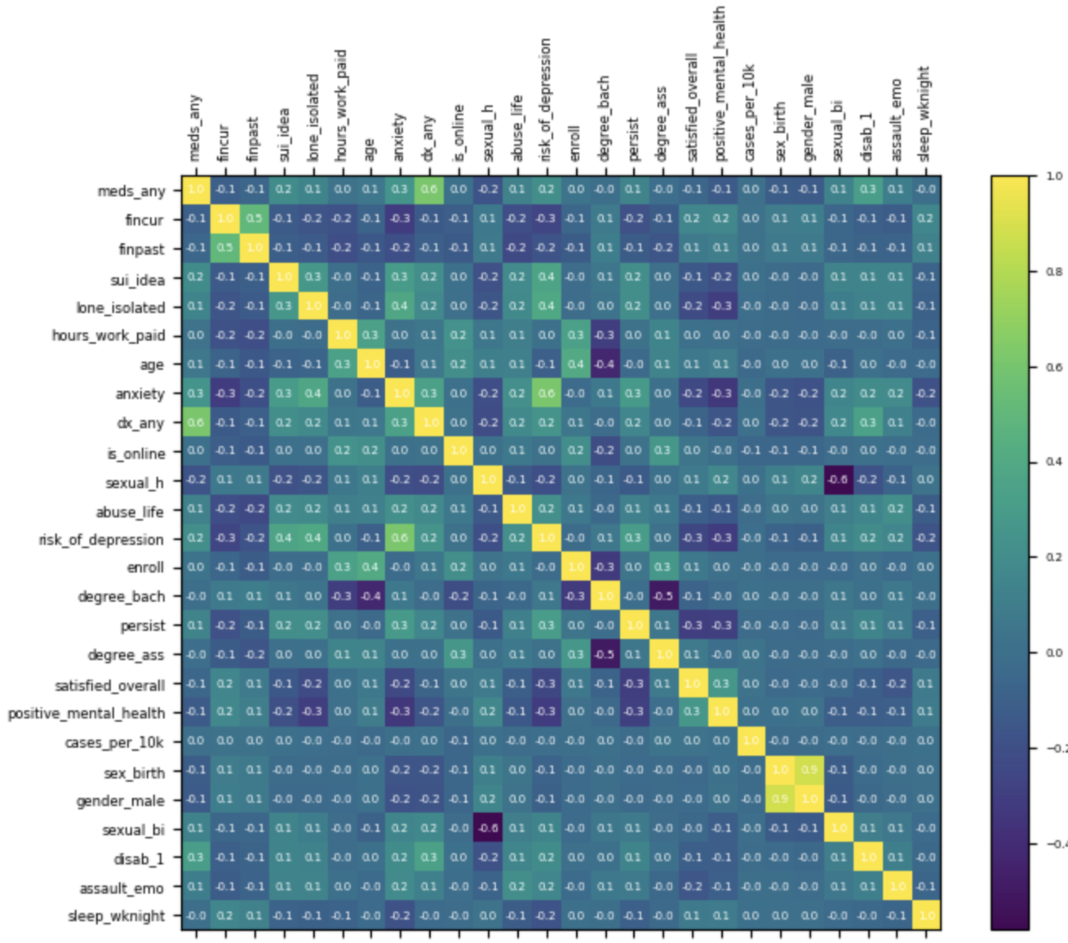


Table 2 displays the mean and standard deviation (in brackets) for a selected number of features that we controlled for during matching. We used chi-squared and Levene's statistics to test whether there is a significant difference between the frequencies and variance of each variable. The difference between the two groups is substantially reduced after matching as the p-values are now statistically insignificant for many of the features. It is good to denote that despite our matching efforts, we still have variables such as student type that are statistically significant meaning that there are differences between our treatment and control for these features likely. However, we see significant improvements in fields such as age, sex, enrollment status, and financial stress. More importantly, we are now more balanced in covid cases as well, which was our primary field to control for.

Table 2. The mean and standard deviation between Treatment and Control For Demographic & Mental Health Features on matched data

Variable	Treatment	Control	P-Value (diff)

Age*	0.93 (0.70)	0.89 (0.65)	0.132 (0.000)
Covid Cases per 10k populations*	0.87 (0.72)	0.98 (0.72)	0.103 (0.072)
Sex at Birth	1.23 (0.42)	1.25 (0.43)	0.405 (0.867)
Is Undergraduate Student	0.68 (0.47)	0.78 (0.41)	0.010 (0.000)
Is Full-Time Enrolled	1.08 (0.47)	1.03 (0.17)	0.230 (0.000)
Has Financial Stress	3.14 (0.90)	3.30 (0.91)	0.127 (0.056)

* *The category has been coarsened into discrete categories. The scale of values might vary from the original raw dataset.*

Figure 3. Predictive Errors at each iteration during the matching process