2023 Annual Conference & Exposition
Baltimore Convention Center, MD | June 25 - 28, 2023
The Harbor of Engineering Education for 130 Years
Paper ID #37090

Office Hours, Demographic Groups, and COVID

Heidi Huang, University of Michigan, Ann Arbor

Heidi Huang is a fourth-year undergraduate student at the University of Michigan, majoring in Computer Science and minoring in Mathematics. Currently, she is working with Professor Andrew Deorio and graduate student Kevin Yan to investigate office hours data from her own department.

Kevin Yan, University of Michigan, Ann Arbor

I'm a Master's student majoring in Computer Science at the University of Michigan. I am working with Heidi Huang and Dr. Andrew DeOrio on Engineering Education.

Dr. Andrew Deorio, University of Michigan, Ann Arbor

Andrew DeOrio is a teaching faculty member at the University of Michigan and a consultant for web and machine learning projects. His research interests are in engineering education and interdisciplinary computing. His teaching has been recognized with the Provost's Teaching Innovation Prize, and he has twice been named Professor of the Year by the students in his department. Andrew is trying to visit every U.S. National Park.

Office Hours, Demographic Groups and COVID

Heidi Huang*, Kevin Yan*, Andrew DeOrio
Department of Electrical Engineering and Computer Science
University of Michigan

1 Abstract

The COVID-19 pandemic has had a significant impact on student life and academic performance. Moreover, different demographic groups have experienced the pandemic very differently. An important factor in the student learning process and therefore academic performance is student help seeking behaviors in office hours. In this paper, we examine the ways in which the pandemic has affected student use of office hours.

We examined office hours usage trends with data from three large computing courses. We augmented this dataset with demographic information from university records. Each entry in our dataset contains information about the time and date of an office hours encounter, the instructor, the student, and the student's gender and underrepresented minority (URM) status. In total, our dataset comprised 33,136 unique encounters.

Our research question examines whether the onset of the pandemic affected the way that students use office hours and if these patterns varied with demographics.

Our results suggest the impact of the pandemic on office hours usage patterns was small. We were surprised to find that the factor with the largest magnitude was not the onset of the pandemic, but rather a student's gender. Men averaged 31.49% fewer office hours encounters compared to women, regardless of pandemic onset. Additionally, we found that the relationship between underrepresented-minority status and office hours encounters was not statistically significant.

2 Introduction and Related Work

The COVID-19 pandemic has affected many areas of life, and it has disproportionately affected some demographic groups. Racial minorities experience higher mortality rates [1][4]. Women and racial and ethnic minorities are also more likely to report high levels of threat and fear of COVID-19 [9]. In certain professions, Black and Hispanic women are more likely to lose their jobs [6]. The short-term effects of the COVID-19 disproportionately affect low-income,

*equal contribution, name in alphabetical order food-insecure households [11]. These conditions have the potential to affect the mental health and performance of students.

Research has also linked the pandemic to trends in education. After the onset of the pandemic and the shift to online learning, test results generally remained consistent or were only marginally worse [2][5]. However, students' mental health declined [7], and they felt less confident in their studies [5]. In one notable example, women overperformed in the COVID-affected semester when compared to previous terms [5]. The diversity of results suggests that the effects of the pandemic and the shift to online learning are heterogeneous and that further research is necessary.

Office hours are an important aspect of university education. Examining which students use office hours and their help-seeking behavior offers insight into student performance, social dynamics, and the quality of the resources available. Doebling et al. discuss how students with a low sense of belonging were more likely to seek help from TAs [3]. Keefer et al. discuss how opaque autograder feedback was associated with substantially longer wait times in office hours [8]. Smith et al. and Keefer et al. both study similar web-based office hours queues, noting high wait times and unequal distribution of office hours usage [8][10].

Many factors contribute to how and why students use office hours. Examining the interaction between these factors and the pandemic can offer insight into how demographic disparities and social dynamics are affected by online learning and the difficulties imposed by the pandemic itself.

Doebling et al. also examine the behavior of different student demographic groups in office hours [3]. The study examines help-seeking behavior of undergraduate computing students at California Polytechnic State University through a mixed-methods study, collecting 138 student survey responses and 15 one-on-one student interviews. Our dataset contains 33,136 cleaned entries collected from computing courses over 2 years at a large, public research institution. We also examine an expanded 5 year dataset containing 52,473 entries. Both datasets contained gender and ethnicity markers: Doebling et al. obtained markers through voluntary disclosure and we obtained markers through university records. We expand on the current research with this larger sample size and the additional pandemic variable.

2.1 Contributions

This study examines whether the onset of the pandemic affected student help-seeking behavior as measured by their use of office hours. We further investigate possible associations with student demographics. Specifically, our research questions are:

- Did office hours usage patterns change after the onset of the pandemic?
- Did office hours usage patterns differ with gender or URM status?

3 Methods

In this section, we describe our process of data collection, the curriculum of courses included in the final dataset, how we cleaned our data, and the statistical methods used during our analysis.

Our dataset consists of office hours usage data extracted from a web-based office hours queue and demographic information from University records.

3.1 Curriculum

Our study takes place at a large, public research institution. The three courses chosen for analysis are a required intermediate programming course (CS2), a required Data Structures and Algorithms course (CS3), and a popular upper-level elective computing course. In all three courses, the majority of the grade is based on coding projects and exams rather than on homework assignments.

All three courses in our study used a web-based office hours queue both before and during the pandemic. Students added themselves to the queue to request help, and an instructor would help the student once they reached the front of the queue (the vast majority of instructors were undergraduate or graduate student teaching assistants). With the web-based queue, students could be helped fairly and instructors were provided a centralized system for managing encounters. Instructors were paired with students on a first-available basis, and students could not choose a specific instructor.

All the office hours in our data set were scheduled in advance. Instructors held their office hours on a regular schedule using the web-based queue. The courses have consistent data spanning semesters from September 2016 to May 2021.

3.2 Data Filtering

The raw office hours queue data contained 235,988 records. We define an office hours encounter to be a period of time from when an instructor begins helping a student and ending when that instructor stops helping the student and begins helping a different student.

We cleaned the data and augmented it using demographic information obtained through university records. We then filtered the dataset to use data from three courses. We also removed semesters that occurred over a year before the pandemic onset. Our final dataset consists of 33,136 unique student office hours encounters spanning from January 2019 to May 2021.

Records were filtered for several reasons. Entries with extreme wait times or encounter lengths were invalid because they represent people adding themselves by mistake or the instructor forgetting to close the queue at the end of the day. Additionally, many special topics courses shared a course number, rendering them unusable. Entries in which the student removed themselves from the queue were invalid because they did not represent an encounter with an instructor. Finally, incomplete semesters where a course only used the web-based queue for a short time were removed. After removing all of these invalid records, we were left with 170,601 entries.

We then augmented our dataset with university demographic records, appending each encounter from the queue data with the student's gender and underrepresented-minority status. If demographic data for a particular student were missing, we removed that student's entries from the queue data entirely. We removed 208 entries this way, with 170,393 valid entries remaining.

We filtered out entries that were not for the courses CS2, CS3, or our upper-level elective, which were chosen due to their having at least 5 years of complete and high-quality data. After limiting our dataset to include the most recent 5 years of data, we were left with 64,638 valid entries.

We avoided a confounding variable of international vs. domestic students by filtering out students with the "International" classification. We discuss this decision more in section 3.3. This left us with 52,473 valid entries.

Lastly, to minimize the effects of courses changing over time, we limited our scope to encounters that occurred between January 2019 and May 2021 inclusive. This leaves us with a final dataset of 33,136 entries, each representing a unique student office hours encounter along with that student's gender and underrepresented-minority status.

3.3 Variables

Student Encounters. Our dependent variable was the number of office hours encounters per student in a single course and semester. The number of office hours encounters quantifies student help-seeking behavior over time. Each student encounter includes timestamps of when each student joined and was subsequently removed from the queue.

3.3.1 Independent Variables

We examined three independent variables in our analysis: gender, underrepresented-minority status, and pre- vs. post-pandemic occurrence.

Gender. We classify students as Man or Woman based on academic records provided by our institution. Unfortunately, our records do not include a marker for nonbinary students.

Our cleaned dataset includes 2,197 men and 1,087 women students.

Underrepresented-Minority Status (URM). We disaggregated by URM or non-URM based on academic records provided by our institution. Our demographic records define URM as "African Americans, Hispanic Americans, American Indians/Native Alaskans, Native Hawaiians/Pacific Islanders (excluding Asian Americans), and multi-racial students identifying at least one of previously listed URM categories." The academic records provided by our university also included an "International" category. Our institution defines international students as "having a citizenship status of Non-Resident Alien or Alien Under Tax Treaty".

The "International" category includes students with a broad and diverse range of experiences. "URM" and "non-URM" are contextualized terms that reflect the lived experiences of domestic students. Thus, we eliminated the possible confounding variable of international vs. domestic by focusing only on students in the URM and non-URM categories.

Our cleaned dataset includes 334 URM students, and 2,950 non-URM students. The dataset also included 770 international students who were filtered from our final analysis.

Pre- vs. Post-Pandemic Onset. The most significant effects of the COVID-19 pandemic began in March of 2020, during our Winter 2020 semester. We consider all semesters occurring before the Winter 2020 semester to be "prepandemic" semesters, whereas semesters occurring after the Winter 2020 semester to be "postpandemic" onset. We filtered out the Winter 2020 semester because of the chaos caused by the onset of the COVID-19 pandemic during that semester.

Our cleaned dataset includes 1,651 students who were enrolled pre-pandemic and 1,984 students who were enrolled post-pandemic. These categories are not mutually exclusive.

3.4 Statistical Methods

We analyzed our data with a Poisson regression test using the Generalized Estimating Equations (GEE) procedure. GEE accounts for repeated measures, which includes when the same student takes multiple classes (e.g. when a student takes CS2 and goes on to take CS3).

Our goal was to measure how different independent variables may have influenced the number of encounters a student has in a single course and semester. In our analysis, we used our core independent variables: gender, URM status, and pre- vs. post-pandemic onset. Additionally, we controlled for two confounding variables: course and fall/winter semesters. The outcome variable was the number of encounters, which we treated as a continuous variable.

4 Results

In this section, we show our results on the relationship between the average number of encounters per student and their gender, URM status, and whether the semester occurred before or after the onset of the pandemic.

4.1 Summary Statistics

Our final dataset, after filtering, consists of 33,136 unique student encounters, including encounters from 3,284 unique students and 201 unique instructors (the vast majority of instructors were teaching assistants). The data span semesters from January 2019 to May 2021. Figure 1 shows the distribution of student encounters in our finalized dataset. Students had a median of 4 encounters and a mean of 8.1 encounters with a standard deviation of 11.7 encounters. The long tail shows that while some students use office hours significantly more than others, the majority of students only use office hours a handful of times during a semester.

4.2 GEE Test

We ran a Poisson regression test using the GEE procedure with gender, URM status, and pre- vs post-pandemic onset as our independent variables and encounters per student as our dependent variable. We also controlled for two confounding variables: the course and the semester (fall/winter).

All predictor variables were discrete, and the outcome variable was encounters per student, which can be treated as a continuous variable. Table 1 shows our results.

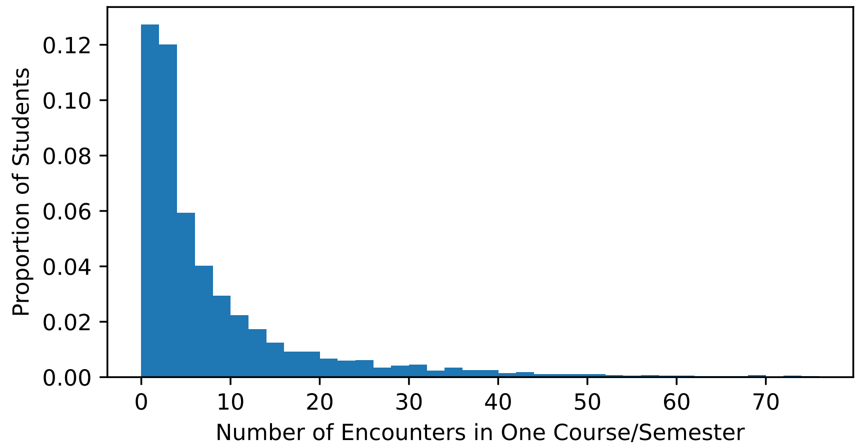

Figure 1: Distribution of student office hours encounters. Each data point represents the number of office hours encounters made by a student during one course in one semester. The data are taken from three major courses spanning multiple semesters.

	coef	% difference vs. baseline	std err	z	P> |z|	[0.025	0.975]
Intercept	2.4658	-	0.059	41.973	0.000	2.351	2.581
Gender: Men (vs. Women Baseline)	-0.3782	-31.49%	0.050	-7.541	0.000	-0.477	-0.280
Pandemic Onset: Post-Pandemic (vs. Pre-Pandemic baseline)	-0.0932	-8.90%	0.047	-1.993	0.046	-0.185	-0.002
Course: CS3 (vs. CS2 baseline)	-0.1143	-10.80%	0.046	-2.466	0.014	-0.205	-0.023
Course: Upper-Level (vs. CS2 baseline)	-0.4665	-37.28%	0.053	-8.747	0.000	-0.571	-0.362
URM Status: URM (vs. non-URM baseline)	0.1190	+12.64%	0.075	1.578	0.114	-0.029	0.267
Semester: Winter (vs. Fall baseline)	0.0161	+1.62%	0.041	0.391	0.696	-0.064	0.097

Table 1: Results from applying the GEE statistical test. The coefficient represents a multiplicative, logarithmic comparison with the baseline. For example, the Gender row shows that men had $e^{-0.3782} = 0.6851$ of the encounters that women did, meaning that men averaged $1 - 0.6851 = 31.49\%$ fewer office hours encounters compared to women. These calculations are provided in the "% difference vs. baseline" column.

When analyzing statistical significance, we used a conservative threshold of $\alpha = 0.01$

4.3 Confounding Variables

Due to the upper-level elective being significantly more advanced than CS2 and CS3, we determined that the course could be a confounding variable. Our data showed possible differences in the average number of encounters between fall and winter semesters, so we added this as an independent variable in our statistical tests.

Our GEE results show that there is no statistically significant relationship between semester (fall vs. winter) and number of encounters. We did find a statistically significant relationship between the course, a confounding variable, and the number of encounters. This is also shown in Figure 3, as the upper-level elective has a consistently lower number of encounters compared to both CS2 and CS3.

We see the CS3 is not statistically significant when compared with CS2, but the upper-level elective is statistically significant (p-value = 0.000). We calculate the coefficient value of -0.467 to a value of 0.627. In other words, students taking the upper-level elective average 37.3% fewer encounters than students taking CS2.

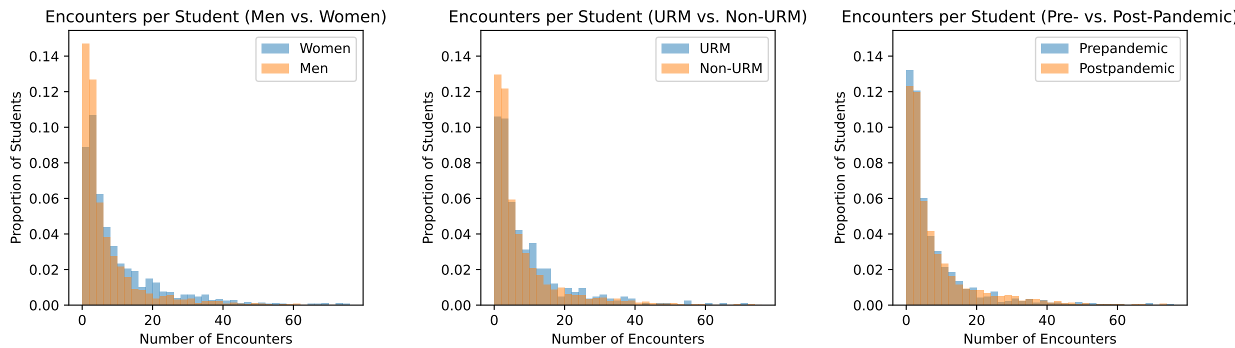

Figure 2: Distribution of the number of office hours encounters disaggregated by different independent variables: gender, URM status, and pre- vs post-pandemic onset. Disaggregating by gender shows the greatest difference, which we confirmed with a statistical test.

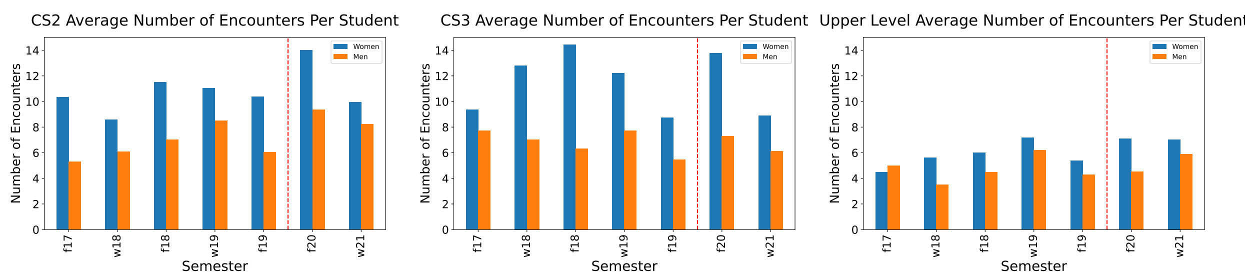

Figure 3: Average number of office hours encounters disaggregated by gender over time. The red dotted line separates semesters before and after the pandemic onset. Women consistently use office hours more than men. We include additional semesters from our cleaned dataset to show that the pattern holds over time.

4.4 Pre- vs. Post-Pandemic Onset and Average Student Encounters

Figure 2 shows the average number of encounters per student disaggregated by pre- vs post-pandemic onset. Our GEE results from Table 1 show that after the onset of the pandemic, students averaged 8.90% fewer encounters in a single course and semester. However, the relationship between pandemic onset and average number of encounters is not statistically significant when using $\alpha = 0.01$ (p-value = 0.046).

4.5 Gender and Average Student Encounters

Figure 2 shows the average number of encounters per student disaggregated by gender. The distribution clearly differs between men and women, with women on average using office hours more often than men.

Our GEE results from Table 1 show that the relationship between gender and average number of encounters is statistically significant (p-value = 0.000). We calculate the coefficient of -0.3782 to a value of 0.6851. In other words, men average 31.49% fewer encounters than women.

Using the expanded dataset spanning 5 years, we took a closer look at how gender influenced the number of encounters across multiple individual semesters. Figure 3 shows that women consistently attended office hours more than men over an extended period of time.

4.6 URM and Average Student Encounters

Figure 2 also shows the average number of encounters per student disaggregated by URM status. Our GEE results from Table 1 show that the relationship between URM status and average number of encounters is not statistically significant when using $\alpha = 0.01$ (p-value = 0.114).

5 Discussion

While many articles describe the pandemic as a major shift in the educational landscape, we did not observe a shift in student use of office hours after its onset. Our data did not provide strong support for statistically significant differences in office hours usage patterns after the onset of the pandemic.

We used a conservative threshold $\alpha = 0.01$ to determine statistical significance. In the case of the pre- vs. post-pandemic onset variable, the p-value was 0.046. It would also be possible to view these results with a $\alpha = 0.05$, which would change the conclusion. When we consider the interpretation of a p-value as a continuum rather than a binary significant or not significant metric, we can see that the data does not provide as strong a support for a decisive conclusion. This contrasts with our data's strong support for the gender difference we observed, with men averaging 31.49% fewer office hours encounters than women (p = 0.000).

5.1 Corroborating Prior Work

Our results corroborate the findings of Doebling et al. [3], who examined the help-seeking behavior of undergraduate computing students using surveys and interviews. They found that men used office hours and other help-seeking resources less frequently than women. They did not observe a statistically significant association between help-seeking frequency and ethnicity.

Doebling et al. received 138 survey responses and 15 interview responses, and their class sizes were mostly under 35 students. Our dataset contained 33,136 clean entries collected over 2 years, and enrollment in each of CS2, CS3, and the upper-level was around 1000, 700, and 450 per semester, respectively. This dataset is a subset of a larger one that contains 52,473 entries collected over 5 years that we used to confirm the trends we observed over a longer period of time.

Our results support the generalizability of Doebling et al.'s findings. We too observed that women students attended office hours more frequently than men. Similarly, we observed that URM status did not have a statistically significant association with office hours usage.

5.2 Limitations

One limitation of our study is that we cannot disentangle remote learning from the pandemic. As a result, the pandemic could have caused interesting student behaviors without affecting usage patterns. Because remote learning was immediately adopted following the onset of the pandemic, we cannot separate how remote learning and pandemic onset interacted in our final results.

An example of possible interference would be if pandemic onset caused widespread mental health issues, potentially lowering office hours attendance. However, remote learning made attending office hours easier (i.e. joining a Zoom meeting instead of physically going to office hours in-person) and thus could have raised office hours attendance. In this way, the pandemic could have caused interesting student behaviors that were not observable in our data.

Another limitation is that our data may not be representative of all courses. Our data only include three courses, two of which are introductory. Additionally, these courses have high-quality data because they have already used office hours queues for many years. Furthermore, all three chosen courses rely heavily on projects and thus may differ from courses that rely more on problem sets. For these reasons, courses that are unlike our chosen courses may not be adequately represented by our data.

A third limitation is that our results may not be entirely independent across semesters. Each semester, course staff might make improvements and changes based on previous semesters, introducing a possible uncontrolled variable. However, because the courses we examined have been very consistent over the semesters that we observed, we believe this limitation is sufficiently mitigated.

6 Conclusion

We examined the relationship between the onset of the COVID-19 pandemic and students' use of office hours in a large computer science program. We found that neither the onset of the pandemic nor URM status were associated with changes in student office hours usage patterns. However, we found that men used office hours less than women in general, averaging 31.5% fewer encounters per student who attended office hours.

Future research on this subject could examine more factors, including grades, family income, and twice-disaggregated demographics, i.e. URM men, URM women, non-URM men, etc. We were unable to perform this analysis due to not having enough data for URM women. Additionally, a more qualitative approach could investigate the pressures experienced by students both before and after the onset of the pandemic. Analysis of survey data or student interviews could shed further light on students' motivations and how they changed during the pandemic. Examining additional confounding factors and students' own perceived struggles could help instructors better understand student circumstances during remote learning and other times of struggle, improving support both inside and outside of the classroom.

References

[1] Carethers, J. Insights into disparities observed with COVID-19. *Journal of Internal Medicine 289*, 4 (2021).

[2] De Santos-Berbel, C., Hernando García, J. I., and De Santos Berbel, L. Undergraduate Student Performance in a Structural Analysis Course: Continuous Assessment before and after the COVID-19 Outbreak. *Education Sciences 12*, 8 (2022).

[3] Doebling, A., and Kazerouni, A. M. Patterns of Academic Help-Seeking in Undergraduate Computing Students. *Koli Calling International Conference on Computing Education Research* (2021).

[4] Doetter, L. F., Preuss, B., and Frisina, P. G. The Intersections of Pandemic, Public Policy and Social Inequality in the United States. *Forum for Social Economics* (2022).

[5] Engelhardt, B., Johnson, M., and Meder, M. E. Learning in the time of Covid-19: Some preliminary findings. *International Review of Economics Education 37* (2021).

[6] Gezici, A., and Ozay, O. An Intersectional Analysis of COVID-19 Unemployment. *Journal of Economics Race and Policy 3* (2020).

[7] Kaparounaki, C. K., Patsali, M. E., Mousa, D.-P. V., Papadopoulou, E. V., Papadopoulou, K. K., and Fountoulakis, K. N. University students' mental health amidst the COVID-19 quarantine in Greece. *Psychiatry Research 290* (2020).

[8] Keefer, C., and Deorio, A. When is Automated Feedback a Barrier to Timely Feedback? *American Society of Engineering Education (ASEE)* (2022).

[9] Niño, M., Harris, C., Drawve, G., and Fitzpatrick, K. M. Race and ethnicity, gender, and age on perceived threats and fear of COVID-19: Evidence from two national data sources. *SSM - Population Health 13* (2021), 100717.

[10] Smith, A. J., Boyer, K. E., Forbes, J., Heckman, S., and Mayer-Patel, K. My Digital Hand: A Tool for Scaling Up One-to-One Peer Teaching in Support of Computer Science Learning. *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education* (2017).

[11] Wolfson, J. A., and Leung, C. W. Food insecurity and COVID-19: disparities in early effects for US adults. *Nutrients* (2020).