Using institutional records and student survey responses to examine freshmen interest groups (FIGs)

Lovenoor Aulck[1], Joshua Malters[1], Casey Lee[1], Gianni Mancinelli[1], Min Sun[2], Jevin West[1]
[1] *University of Washington Information School*
[2] *University of Washington College of Education*

Background and Context

Freshman orientation seminars (freshman seminars) are courses dedicated to helping incoming students transition to college life both socially and academically. The popularity and ubiquity of these courses has made them among the most studied course genre in American higher education [1, 5, 3]. That said, the existence and effectiveness of these seminars on college campuses across the U.S. continues to be called into question [4]. Though some prior studies have used randomized controlled trials (e.g. [8]), large scale and causally rigorous studies of seminar effectiveness using matched comparison groups are rare [7].

Purpose and Research Question

In this work, we gather data from institutional databases at a large U.S. university (the University of Washington; UW) to examine the impact of freshman interest groups (FIGs, a type of freshman orientation seminar) on graduation rates and first-year retention. We use propensity score matching on nearly 58,000 students across 12 student cohorts (and 18 years of data). We use a suite of variables on students prior to their post-secondary education to match students who enrolled in FIGs with those who did not. We then explore the differences between these groups in terms of educational outcomes while also examining specific ethnic/racial groups (namely, Hispanic and under-represented minority students). Lastly, we use exit surveys from over 12,500 students to better understand which aspects of FIGs are most beneficial.

Data Collection and Analysis

We collected de-identified student data from the UW's data custodians. This data included complete student transcript records (courses taken, grades, etc.), demographics (race, gender, ethnicity, etc.), and entrance application information (high school GPA, entrance exam scores, etc.). Data for analysis was limited to first-time, first-year students who first enrolled at the UW between 1998 to 2010. This totalled 57,979 students, of whom 32,572 enrolled in a FIG (56.2%) and 15,407 did not (43.8%).

We also used data from the Collegeboard's enrollment planning services (EPS) and U.S. Census data. The EPS data was only available on a high-school level and not at an individual level. The EPS data included information on students' major intentions in college, preferences regarding post-secondary campuses, long-term educational attainment goals, and some information on parents. For the U.S. Census data, information on average income, bachelor's degree attainment, and high school completion for each ZIP code was pulled from the U.S. Census' American Fact Finder. We matched the EPS and Census data to students based on their high schools and ZIP codes from their admissions applications.

Students selected whether they would or wouldn't enter a FIG (FIG and non-FIG groups, respectively). To account for this and mitigate selection bias, we used stratified propensity score matching (PSM) [6]. We used students' demographic information and pre-college information from the registrar data as well as information from the EPS and the census data to calculate

propensity scores via a logit model. The model included 197 covariates. After calculating propensity scores, we matched each student in the treatment (FIG) group to students in the control (non-FIG) group using two-levels of stratification and fixed caliper widths. We first matched the students according to year of entry to the University and then by whether they were a STEM-interested student (i.e. a student interested in science, technology, engineering, and/or math (STEM) fields). Using this dual stratification, we paired every FIG student to corresponding non-FIG students from the same entrance year and with the same (binary) indication of STEM interest, after which we used caliper matching. The caliper was kept at one-tenth of the pooled standard deviation of all propensity scores and students were matched one-to-many, with replacement. The distributions of propensity scores can be seen in Figure 1. To further examine specific student groups, we matched each Hispanic and each under-represented FIG student to peers in the same group. We used the same strategy of dual stratification with caliper matching in a one-to-many manner for this round of matching.

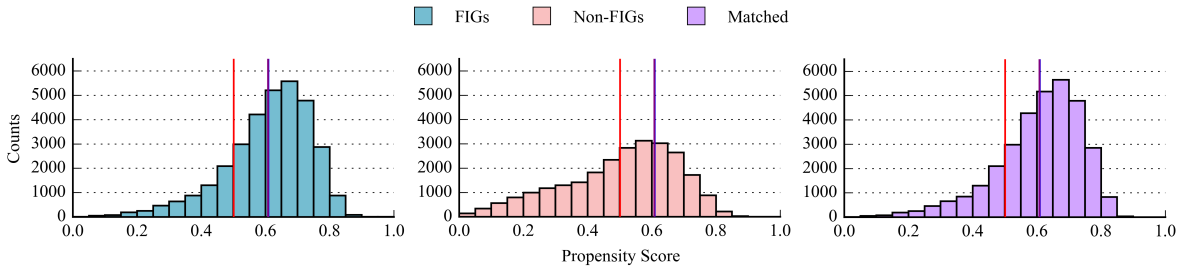

Figure 1: *Propensity score distributions for FIG students (left), non-FIG students (middle), and matched non-FIG students (right). The left and right distributions were used in the analysis*

In addition, the UW's FIG program collected exit surveys from nearly every FIG participant from 2010-2015 (14,514 total; an average of about 2,419 per year). These surveys asked students questions about their FIG experiences and most responses were recorded as open-ended text. We examined answers to a single question in the survey: "What did you find most valuable about the FIG?" for which there were 12,539 non-blank responses. Note that the timeframes for the student registrar data and the survey data do not overlap except for the year 2010. Nevertheless, student demographic data was pulled from the registrar dataset and linked to all survey responses, even for years that were not used when analyzing the transcript data.

To analyze student survey responses, we used a text-mining-based qualitative analysis approach based on grounded theory. First, we developed an initial code book using a form of unsupervised learning called latent Dirichlet allocation [2]. Topic modeling was only used as a starting point in coding the massive text corpus. From there, two coders independently coded 1000 responses and iteratively developed a joint codebook to use. Each tag within the codebook was used independently. After coding these 1000 responses, the coders then coded an additional 1000 responses and discussed consistency in coding. Thereafter, the coders tagged every remaining response individually. After coding all responses, each response that was not coded identically between the two researchers was discussed and a consensus regarding coding was drawn. The final set of codes for each individual response was used in the analysis.

Results and Conclusions

Of the FIG students, 32,512 FIG students (99.8%) had at least one non-FIG student matched and 30,257 FIG students (92.9%) had at least 20 non-FIG students matched. This indicates a high level of common support for PSM. Each FIG student was matched to an average (±SD)

of 66.1 (±34.0) non-FIG students. Of the 1,699 Hispanic students in FIGs, 1,546 (91.0%) were matched to at least one Hispanic non-FIG student; of the 1,578 under-represented minority students in FIGs, 1,428 (90.5%) were matched to at least one under-represented non-FIG student.

Table 1: *Graduation and re-enrollment rates for students after PSM*

Student Group	Measure	FIG	non-FIG	Difference
All Students	Graduation	81.60%	74.57%	+7.03%
	Re-Enrollment	94.18%	90.49%	+3.69%
Hispanic	Graduation	77.96%	69.32%	+8.64%
	Re-Enrollment	93.06%	87.28%	+5.78%
Under-Represented	Graduation	76.81%	62.81%	+14.00%
	Re-Enrollment	94.75%	88.18%	+6.57%

We show graduation and re-enrollment rates for students after PSM in Table 1. FIG students tended to have substantially higher graduation and re-enrollment rates than their matched non-FIG peers (differences of 7.0 and 3.7 percentage points, respectively). It should be noted that for both groups, the re-enrollment rates are greater than 90%, thereby decreasing the maximum relative difference between the two groups in terms of percentage points.

The graduation and re-enrollment rates for Hispanic and under-represented FIG students were substantially higher than non-FIG students. This difference is much more pronounced than the estimated FIG effects for all students in Table 1. It should be noted that the graduation rates for both Hispanic and under-represented students were still below the University's average across all students. These facts could go hand-in-hand as the lower average rates across these groups allows for more differential gains to be realized for students attending FIGs. We know that the FIG curriculum at the University is not tailored to specific student groups based on race, ethnicity, or family backgrounds. Future research is needed to better explain this greater effect of FIGs on Hispanic and under-represented students.

We show the four most commonly applied codes from survey responses in Table 2. Meeting new people (52.4% of all responses), knowing people in classes (23.8%), references to an activity in the FIG (13.7%), and college survivial skills (13.4%) were the most commonly recorded responses. Far and away, students thought social aspects of the FIG were the most valuable. Over half of all responses mentioned meeting new people and/or making friends and about one quarter referenced knowing people in classes and being able to make study groups. Interestingly, when examining the tags across student subgroups, we did not observe salient differences in their rates of use.

Table 2: *Frequency of applied codes (only top 4 are shown). Percentages indicate percent of total responses that were tagged with respective code.*

tag	All students	Under-represented	Non-under-represented	Hispanic	Not Hispanic
meeting new people	52.4%	50.4%	52.5%	53.5%	52.2%
having people in classes	23.8%	23.7%	23.7%	24.5%	23.6%
specific FIG activity	13.7%	16.1%	13.5%	13.1%	13.7%
college survival skills	13.4%	12.6%	13.5%	13.4%	13.4%

References

[1] B. O. Barefoot, C. L. Warnock, M. P. Dickinson, S. E. Richardson, and M. R. Roberts. Exploring the evidence: Reporting outcomes of freshman seminars. Technical Report 2, National Resource Center for the Freshman Year Experience & Students in Transition, 1998.

[2] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3:993–1022, 2003.

[3] J. B. Cuseo. Freshman orientation seminar at community colleges: A research-based rationale for its value, content, and delivery. 1997.

[4] A. M. Griffin and J. Romm. Exploring the evidence: Reporting research on first-year seminars. Technical Report 4, National Resource Center for The First-Year Experience & Students in Transition, 2008.

[5] R. D. Padgett, J. R. Keup, and E. T. Pascarella. The impact of first-year seminars on college students life-long learning orientations. *Journal of Student Affairs Research and Practice*, 50(2):133–151, 2013.

[6] P. R. Rosenbaum and D. B. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.

[7] C. A. Schnell and C. D. Doetkott. First year seminars produce long-term impact. *Journal of College Student Retention: Research, Theory & Practice*, 4(4):377–391, 2003.

[8] G. Strumpf and P. Hunt. The effects of an orientation course on the retention and academic standing of entering freshmen, controlling for the volunteer effect. *Journal of The First-Year Experience & Students in Transition*, 5(1):7–14, 1993.