Spring 5-2024

Relationship between College Graduate Job Placement and Faculty-Led, Team-based, Undergraduate Research Experiences: A Propensity Score Analysis

Julia Sonnenberg-Klein

doi: https://doi.org/10.57709/36962302

ACCEPTANCE

This dissertation, Relationship between College Graduate Job Placement and Faculty-Led, Team-based, Undergraduate Research Experiences: A Propensity Score Analysis, by Julia Sonnenberg-Klein, was prepared under the direction of the candidate's Dissertation Advisory Committee. It is accepted by the committee members in partial fulfillment of the requirements for the degree, Doctor of Philosophy, in the College of Education & Human Development, Georgia State University.

The Dissertation Advisory Committee and the student's Department Chairperson, as representatives of the faculty, certify that this dissertation has met all standards of excellence and scholarship as determined by the faculty.

Hongli Li, PhD Committee Chair	David Johnson, PhD Committee Co-Chair
Michael Frisby, PhD Committee Member	Qiana Lachaud, PhD Committee Member

Date

Jennifer Esposito, PhD
Chairperson, Department of Education Policy Studies

Paul A. Alberto, Ph.D
Dean, College of Education & Human Development

AUTHOR'S STATEMENT

By presenting this dissertation as a partial fulfillment of the requirements for the advanced degree from Georgia State University, I agree that the library of Georgia State University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote, to copy from, or to publish this dissertation may be granted by the professor under whose direction it was written, by the College of Education & Human Development's Director of Graduate Studies, or by me. Such quoting, copying, or publishing must be solely for scholarly purposes and will not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential financial gain will not be allowed without my written permission.

Julia Sonnenberg-Klein

Julia Sonnenberg-Klein

NOTICE TO BORROWERS

All dissertations deposited in the Georgia State University library must be used in accordance with the stipulations prescribed by the author in the preceding statement. The author of this dissertation is:

Julia Sonnenberg-Klein
Educational Policy Studies
College of Education & Human Development
Georgia State University

The directors of this dissertation are:

Dr. Hongli Li
Department of Educational Policy Studies
College of Education & Human Development
Georgia State University
Atlanta, GA 30303

Dr. David Johnson
Department of Educational Policy Studies
College of Education & Human Development
Georgia State University
Atlanta, GA 30303

CURRICULUM VITAE

Julia Sonnenberg-Klein

ADDRESS: 311 Farleigh Trce
Marietta, GA 30068

EDUCATION:

Ph.D.	2024	Georgia State University Educational Policy Studies
M.Ed.	2013	University of Illinois Urbana-Champaign Education Organization and Leadership
B.S.	1999	University of Illinois Urbana-Champaign Engineering Physics

PROFESSIONAL EXPERIENCE:

2017-present	Executive Director, Vertically Integrated Projects Program Georgia Institute of Technology
2013-2017	Academic Program Manager Georgia Institute of Technology
2007-2013	Education Outreach Manager Georgia Institute of Technology

PRESENTATIONS AND PUBLICATIONS:

Sonnenberg-Klein, J., Coyle, E. J. (2024, early access). Leadership Growth Over Multiple Semesters in Project-Based Student Teams Embedded in Faculty Research (Vertically Integrated Projects). *IEEE Transactions on Education, early access article,* 1-10.

Sonnenberg-Klein, J., Coyle, E. J., & Saigal, K. (2023, June). *How "Multidisciplinary" Is It? Measuring the Multidisciplinarity of Classes and Student Teams.* 2023 ASEE Annual Conference & Exposition, Baltimore, MD.

Sonnenberg-Klein, J., Coyle, E. J., Bringardner, J., Kukta, R. V., Llewellyn, D., Ramirez, N., & Zoltowski, C. B. (2023, June). *Dimensions of Equity: Undergraduate Research* *through Vertically Integrated Projects at Five Institutions.* 2023 IEEE Frontiers in Education Conference, College Station, TX.

Bringardner, J., Chao, H.-C., Coyle, E. J., Dalecka, B., Junha, T., Marshall, S., Ramirez, N., Shlayan, N., Smith, L., & Sonnenberg-Klein, J. (2022, October). *Level-up: Expand Undergraduate Research Capacity (and Serve Faculty) through Vertically Integrated Projects*. Frontiers in Education Conference, Uppsala, Sweden.

Sonnenberg-Klein, J., & Coyle, E. J. (2022, October). *Experience vs. Academic Rank: Leadership Growth over Multiple Semesters in Vertically Integrated Project (VIP) Courses*. 2022 IEEE Frontiers in Education Conference, Uppsala, Sweden.

Sonnenberg-Klein, J., & Moore, E. (2022). *Main Course or Buffet? Degree of Student Specialization in Elective Choices.* 2022 IEEE Frontiers in Education Conference, Uppsala, Sweden.

Sonnenberg-Klein, J., Rosales-Salas, J., Aristizábal-Pinzón, A. A., Muriel, A. F. O., Ruiz, P. S., Azevedo, A. T., Inacio, E., & Morini, C. (2022). *Vertically Integrated Project Programs in South America – Profiles of Multidisciplinary Team-Based Project-Based Programs at 4 Institutions.* 2022 IEEE World Engineering Education Conference (EDUNINE), online.

Sonnenberg-Klein, J., & Coyle, E. J. (2021, June). *Structures and Practices that Support Equity in Undergraduate Research: Enrollment Equity in Vertically Integrated Projects (VIP) Programs at 12 Institutions.* Centering Diversity, Equity + Inclusion in Undergraduate Research + Creative Activity, Held by the Council on Undergraduate Research, Online.

Amoo, M. E., Bringardner, J., Chen, J.-Y., Coyle, E. J., Finnegan, J., Kim, C. J., Koman, P. D., Lagoudas, M. Z., Llewellyn, D. C., Logan, L., Sonnenberg-Klein, J., Trent, N. M., Strachan, S. M., & Ward, B. C. (2020, June). *Breaking Down the Silos: Innovations for Multidisciplinary Programs*. 2020 ASEE Virtual Annual Conference.

Kinglsey, G., Coyle, E. J., Melkers, J., & Sonnenberg-Klein, J. (2019, November). *The VIP Consortium & Institutional Change: 35 Institutions & Counting*. AACU Transforming STEM Higher Education, Chicago, IL.

Sonnenberg-Klein, J., & Coyle, E. J. (2019, April). *Equity in Undergraduate Research through Vertically Integrated Projects: Lessons Learned*. 2019 ABET Symposium, Dallas, TX.

Strachan, S. M., Marshall, S., Murray, P., Coyle, E. J., & Sonnenberg-Klein, J. (2019). Using Vertically Integrated Projects to embed research-based education for sustainable development in undergraduate curricula. *International Journal of Sustainability in Higher Education,* 1313-1328.

Sonnenberg-Klein, J., & Coyle, E. (2018, April). *Vertically Integrated Projects: Diverse, Multi-disciplinary and Scalable Project Based Learning*. 2018 ABET Symposium, San Diego, California.

Sonnenberg-Klein, J., Abler, R. T., & Coyle, E. J. (2018, April). *Diversity and Student Persistence in the Vertically Integrated Project (VIP) Course Sequence.* 2018 CoNECD - The Collaborative Network for Engineering and Computing Diversity Conference, Crystal City, Virginia.

Sonnenberg-Klein, J., Abler, R. T., & Coyle, E. J. (2018, June). *Correlation between Academic Credit-use Policies and Student Persistence in Vertically Integrated Project (VIP) Courses*. 2018 ASEE Annual Conference & Exposition, Salt Lake City, Utah.

Sonnenberg-Klein, J., Abler, R. T., & Coyle, E. J. (2018, June). *Social Network Analysis: Peer Support and Peer Management in Multidisciplinary, Vertically Integrated Teams.* 2018 ASEE Annual Conference & Exposition, Salt Lake City, Utah.

Aazhang, B., Abler, R. T., Bost, L. F., Cavallaro, J. R., Chong, E., Coyle, E., Cullers, J., Dennis, S., Enjeti, P., Filippas, A., Froyd, J. E., Garmire, D., George, J., Hohner, G., Hughes, W., Johnson, A., Kim, C., Kim, H., Klenke, R. H., … Allebach, Jan. (2017, June). *Vertically Integrated Projects (VIP) Programs: Multidisciplinary Projects with Homes in Any Discipline.* 2017 ASEE Annual Conference and Exposition, Columbus, Ohio.

Abler, R. T., Coyle, E. J., Juhna, T., Kim, H., Marshall, S., Pardo, M., Percybrooks, W. S., & Sonnenberg-Klein, J. (2017, June). *Vertically Integrated Projects (VIP) Programs at International Institutions: Multidisciplinary Projects with Homes in Any Discipline.* 2017 ASEE International Forum, Columbus, Ohio.

Edward Coyle, Afroditi Filippas, Brian Gilchrist, Will Hughes, Charles Kim, Hakil Kim, Magda Lagoudas, & Julia Sonnenberg-Klein. (2017, June). *Vertically Integrated Projects (VIP) Programs: Variations in Implementation across Multiple Institutions*. 2017 ASEE Annual Conference and Exposition, Columbus, Ohio.

Sonnenberg-Klein, J., Abler, R. T., Coyle, E. J., & Ai, H. H. (2017, June). *Multidisciplinary Vertically Integrated Teams: Social Network Analysis of Peer Evaluations for Vertically Integrated Projects (VIP) Program Teams.* 2017 ASEE Annual Conference & Exposition, Columbus, Ohio.

PROFESSIONAL SOCIETIES AND ORGANIZATIONS

2014-Present	VIP Consortium
2018-Present	American Society for Engineering Education
2021	Association for the Study of Higher Education

RELATIONSHIP BETWEEN COLLEGE GRADUATE JOB PLACEMENT AND FACULTY-LED, TEAM-BASED, UNDERGRADUATE RESEARCH EXPERIENCES: A PROPENSITY SCORE ANALYSIS

by

JULIA SONNENBERG-KLEIN

Under the Direction of Hongli Li and David Johnson

ABSTRACT

While career development is not the sole purpose of higher education, obtaining a better job is a strong motivating factor in college attendance. The relationship between higher education and graduates' career outcomes can be framed through the theories of human capital, social capital, cultural capital, and signaling. In the context of finding a job, capitals are resources and traits valued by employers. Signaling theory describes the process by which job applicants signal their fitness for employment and the process by which employers interpret signals from applicants. Substantial research shows that access to the job market is shaped by race, ethnicity, and social class, with evident discrimination in the screening of applicants by race/ethnicity. Lower socioeconomic status decreases applicants' appeal to employers, and men are more successful than women in finding employment. The American Association of Colleges and Universities identified eleven high-impact experiences, and participation in multiple high-impact experiences is associated with greater learning gains and compensatory gains for marginalized students. However, little research has been done on how these experiences affect career outcomes. Undergraduate research is considered a high-impact experience, and project-based learning involving large teams embedded in faculty research (LT-PBL-EFR) is a special case of undergraduate research. This dissertation presents a framework through which LT-PBL-EFR may support equity in job placement, and employs propensity score analysis to examine the effect of LT-PBL-EFR on job placement and equity in job placement. When other factors are held constant, participation in three semesters of LT-PBL-EFR is associated with triple the odds of having found a job prior to graduation, comparable to odds associated with internships. The positive association is consistent among non-white students and students of lower socioeconomic status. This suggests that participation in multiple semesters of LT-PBL-EFR improves job placement rates for the general population as well as for non-white students and students of lower socioeconomic status.

INDEX WORDS: Job placement, career outcomes, higher education, high-impact practices, high-impact experiences, discrimination, equity, propensity score analysis, inverse propenstiy score weighting, project-based learning, PBL, undergraduate research, vertically integrated projects, VIP

Relationship between College Graduate Job Placement and Faculty-Led, Team-based, Undergraduate Research Experiences: A Propensity Score Analysis

by

Julia Sonnenberg-Klein

A Dissertation

Presented in Partial Fulfillment of Requirements for the

Degree of

Doctor of Philosophy

in

Educational Policy Studies

in

Educational Policy Studies

in

the College of Education & Human Development

Georgia State University

Atlanta, GA
2024

ACKNOWLEDGMENTS

I would like to thank and acknowledge Georgia State University's Department of Educational Policy Studies, which provided excellent instruction in and balance between research methods and social foundations. I am also very appreciative to my dissertation committee chairs, Professors Hongli Li and David Johnson, and committee members, Professors Qiana Lachaud and Michael Frisby, for their expertise and guidance in research methods, higher education, and social foundations.

I am especially grateful to Professor Edward Coyle, Director of the Georgia Tech Vertically Integrated Projects (VIP) Program, Chair of the VIP Consortium, and developer of the VIP Model. Coyle has provided professional support throughout this study, and he has encouraged my personal and professional growth for over ten years.

This study was only possible because large numbers of faculty have voluntarily incorporated large teams of undergraduates into their research portfolios, and because students have been eager to work with them. While I find the results of the study exciting, it is the faculty and students who have done and continue to do the heavy lifting.

The study also would not have been possible without Georgia Tech's College of Engineering, College of Computing, and School of Electrical and Computer Engineering.

An analysis by Joe Ludlum from Georgia Tech's Office of Academic Effectiveness motivated this study. His and his unit's willingness to help programs examine their impact on student outcomes is an invaluable asset to Georgia Tech.

Finally, I am thankful to my husband and children. Working full time and attending school left me with less time for family, but they were understanding and supportive throughout the process.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABREVIATIONS

CoOp	Cooperative Education
GPA	Grade point average
GT 1000	Freshman experience course
GT 2000	Transfer student experience course
LT-PBL-EFR	Large team project-based learning embedded in faculty research
SES	Socioeconomic status
STEM	Science, technology, engineering and math
URM	Underserved minority group
VIP	Vertically Integrated Projects

1 THE PROBLEM

College education conveys a multitude of benefits to graduates, including learning and cognitive changes, higher quality of life, socioeconomic gains, and transmission of benefits from college graduates to their children (Hout, 2012; Mayhew et al., 2016). While not the sole purpose of higher education, socioeconomic gains are strong motivating factors in college attendance. These gains include higher employment rates, higher earnings, increased job status, greater job satisfaction, and positive returns on the investment of college attendance (Hout, 2012; Mayhew et al., 2016). At the juncture between college and the workforce, prospective employers serve as gatekeepers, judging applicants' potential value to their organizations and offering jobs to those they deem most desirable. These judgements are shaped by race, ethnicity, and social class – research on hiring practices shows pervasive racial discrimination in the screening of applicants (Bertrand & Mullainathan, 2004; Gaddis, 2015; S. K. Kang et al., 2016; Milkman et al., 2015; Quillian et al., 2017), that lower socioeconomic status (SES) decreases applicants' appeal to employers (Ingram & Allen, 2019; Rivera, 2011, 2016), and that men are more successful than women in finding employment (Koc, 2014). Once college graduates are in the workforce, when institutional selectivity is controlled for, wages are lower for women and students from families of lower socioeconomic status, but do not differ by race/ethnicity (Perna, 2005; Wolniak et al., 2008; Wolniak & Engberg, 2019). Instead, wage inequity by race/ethnicity is rooted in institutional selectivity, with less selective institutions disproportionately serving people of color, yielding lower graduation rates and lower salaries for graduates.

Substantial research has been done on factors that support student development and graduation rates, on discrimination in hiring practices, and on earnings once in the workforce. However, little research has been done on college experiences that may empower students to overcome discrimination they face in accessing the workforce. At the Georgia Institute of

Technology, assessment of a project-based learning course involving large teams embedded in faculty research showed potential promise in this area. In project-based learning, students tackle real-world projects, collaborate, and solve problems as professionals would (Krajcik & Blumenfeld, 2005). In a comparison group that did not participate in large team project-based learning embedded in faculty research (LT-PBL-EFR), Asian students and historically underserved minority students (non-white, non-Asian, and Hispanic/Latino) reported 6% lower job placement rates prior to graduation than white non-Hispanic/Latino (white) students, with placements of 70.3% and 69.7% for Asian and historically underserved minority students, compared to 76.1% for white students. This inequity aligns with findings of discrimination in resume audit studies, in which resumes with non-white signifiers yield fewer invitations for interviews than equivalent resumes for white-seeming applicants. In contrast, there was no inequity by race/ethnicity among students who participated in the LT-PBL-EFR course. Compared to non-participant groups of the same race/ethnicities, job placement rates were 8.9% higher for historically underserved students, 9% higher for Asian students, and 1.9% higher for white students. While the apparent gains and equity in job placement among course participants is encouraging, the results were only correlational. Social class, participation in other curricular/cocurricular programs, career-related experiences, and use of career center services may underly the differences. To determine what proportion of the seeming gains could be attributed to LT-PBL-EFR, whether there are differential effects on students from different backgrounds, and how effects compare with effects associated with other programs, more advanced analysis was needed.

Research Questions

The study was guided by the following research questions:

1. How does participation in Georgia Tech's large team project-based learning course embedded in faculty research impact job placement for college graduates?

2. To what extent does participation in the course influence equity in job placement by race/ethnicity, gender, and socioeconomic status?

Purpose

The purpose of this study is to determine how participation in an LT-PBL-EFR program may impact college students' job placement, and job placement equity by race/ethnicity, gender, and SES.

Significance of the Study

Extensive research has been done on the impact of internships, work experience, and career development courses on college graduates' career outcomes, but there is nominal research on how other experiences impact job placement. Miller et al. examined correlations between job placement prior to graduation (also the outcome in this study) and participation in high-impact experiences. They found higher job placement rates associated with internships, culminating capstone courses, and service learning, but not with undergraduate research (Miller et al., 2018). Wolniak and Engberg examined correlations between five high-impact experiences and early career earnings. They found nominal correlation with participation in high-impact experiences, but they did not examine initial job placement rates (Wolniak & Engberg, 2019). A shortcoming of the study was that they did not control for salary differences between graduate school and regular employment. As a result, participating in undergraduate research (which is associated with going to graduate school) was correlated with lower earnings, which the authors discussed but did not control for. In two studies Hu and Wolniak examined correlations between early career earnings, academic engagement and social engagement among high-achieving non-white students (Hu & Wolniak, 2010, 2013). Their framework is relevant, because it drew on theories of human capital (academic engagement) and social capital (social engagement), with one study focusing on differences in earnings by engagement and student major, and the other on engagement and demographics. Among STEM students, there was a positive correlation between earnings and social engagement, and a slightly negative correlation between earnings and academic engagement. Among non-STEM students, academic engagement was positively correlated with earnings, and social engagement was not significant (Hu & Wolniak, 2010). In the second study, they found academic engagement associated with higher salaries among men, and social engagement associated with higher salaries among women (Hu & Wolniak, 2013). Because their studies did not include white students, it did not address equity or differential effects (Hu & Wolniak, 2010, 2013). In qualitative studies relevant to the project, Anderson and Tomlinson used a signaling theory framework to explore employer hiring practices, which is a novel approach, but the framework has not yet been employed in studies on college career outcomes (Anderson & Tomlinson, 2020; Tomlinson & Anderson, 2021).

Resume audit studies show pervasive discrimination in hiring practices and benefits of higher socioeconomic status in finding employment, both of which effect students' transition from college to the workforce (Bertrand & Mullainathan, 2004; Gaddis, 2015; S. K. Kang et al., 2016; Milkman et al., 2015; Quillian et al., 2017; Rivera, 2016). In addition to potentially developing human capital, participation in curricular/cocurricular programs may convey social and cultural capital that affect how employers perceive applicants. In the prior study, equity in job placement among LT-PBL-EFR participants and inequity among non-participants may be the product of this process, but the seeming gains cannot be attributed to LT-PBL-EFR participation without more rigorous analysis.

This study fills a gap in the literature by determining to what degree, if any, participation in LT-PBL-EFR supports student job placement, and whether there are differential gains by race/ethnicity, gender, and/or SES. While the project uses a smaller dataset than the two extant studies on the effects of high-impact practices on career outcomes, it makes three significant contributions. First, the study **incorporates two novel theories** into the framework. While Hu and Wolniak's framework included human capital and social capital (Hu & Wolniak, 2010, 2013), it did not include cultural capital or signaling theory. Cultural capital is relevant to project-based learning and student-faculty interaction which occur in LT-PBL-EFR. Signaling theory describes the process by which capitals are converted into employment through signals that are sent by applicants and interpreted by employers. Anderson and Tomlinson used the theory in their studies on employer hiring practices (Anderson & Tomlinson, 2020; Tomlinson & Anderson, 2021), but the framework has not been applied in studies on college career outcomes. Second, unlike prior studies which were correlational, this study **employs a more rigorous analysis method**, propensity score analysis, which emulates experimental designs, removes selection bias, and can be used to draw causal inferences (Guo & Fraser, 2015). Third, the few studies on high-impact experiences and career outcomes show no correlation between job placement and high-impact experiences (Miller et al., 2018; Wolniak & Engberg, 2019), but initial analysis of job-placement among LT-PBL-EFR participants at Georgia Tech **showed significant differences, with seeming compensatory gains among non-white students and elimination of inequity**. Research shows correlations between job placement rates and work-based learning (Koc, 2014). Large team project-based learning embedded in faculty research may emulate work-based learning, an idea that has not been previously considered. The seeming differential gains among non-white students may also be similar to compensatory gains in college persistence and success among marginalized students who participate in multiple high-impact experiences (Kuh, 2008), making the effect of LT-PBL-EFR by race/ethnicity an important question to explore.

Assumptions and Limitations

A limitation of this study is that it is not experimental. Results from experiments employing random assignment are more reliable, but randomized experiments pose multiple challenges including cost, attrition, and the ethics of requiring or denying treatment. The method employed in this study, inverse propensity score weighting, simulates experimental design through data balancing, and it can be used in causal inference. This is a more rigorous method than the initial analysis, but it is not as rigorous as randomized experiments.

The study relies on observational data, so it is limited to data collected by the institution, such as demographics, academic records, financial aid status, participation in co-curricular programs (work-based learning, living learning communities, Greek fraternities and sororities, status as an athlete), and self-reported use of the campus career center. Other experiences that could enhance student employability, such as membership in clubs and holding elected offices, are not tracked by the institution, so they were not included in the analysis.

The study focused on a single institution, so the findings cannot be generalized to other populations. Under-representation of historically underserved students at the institution (approximately 8% black/African American and 8% Hispanic/Latino) affected the potential sample size for the subgroups. High enrollment in the program, multiple years of data, and analysis by subgroup helped mitigate but did not eliminate this effect.

Overview of the Study

The LT-PBL-EFR program examined in the study was Georgia Tech's Vertically Integrated Projects (VIP) Program, which enrolls students from majors across the campus. The career-outcomes of participants cannot be easily compared to those of non-participants, because the students differ in myriad ways: major, grade point average, participation in curricular and co-curricular activities, work-based learning experiences, SES, race/ethnicity, and sex. The variety of variables could be controlled for in a randomized study, but a randomized study would pose ethical challenges, particularly denying services to students in a control group when the services could increase their career success. This study simulates a randomized study with propensity score analysis. Students' likelihood (propensity) of participating in the program for zero, one, two, or three semesters were estimated, and inverse propensity score weights were used to construct treatment and control groups from existing data. Outcomes for the constructed groups were analyzed through logistic regression to estimate the effect of LT-PBL-EFR on job placement. The method controls for self-selection and factors that can affect career outcomes. The effect of LT-PBL-EFR on job placement was expected to be lower than in the prior study. This is because propensity score analysis reduces bias, yielding lower and more accurate estimates of effects.

Context of the Study

Undergraduate Research Large team project-based learning embedded in faculty research is a form of undergraduate research. The American Association of Colleges and Universities identifies undergraduate research as one of eleven high-impact experiences associated with greater learning gains in college (Association of American Colleges and Universities, n.d.; Kuh, 2008; National Academies of Sciences, 2017). While high-impact experiences benefit all students, they are correlated with differentially greater gains for historically underserved students (Kuh, 2008). Hispanic students who participate in multiple high-impact experiences see greater gains in GPA, with increasing numbers of experiences increasing their GPAs above those of white students who participate in the same number of experiences. African American student persistence to the next semester increases with the number of high-impact experiences, exceeding the persistence of white students who participate in the same number of experiences (Kuh, 2008). However, historically underserved students, transfer students, and first-generation college students participate in research with faculty at lower rates than their peers. Whereas 24% of white and Asian students do research with faculty by the time they graduate, the rates are 18% for both black and Hispanic/Latino students (National Survey of Student Engagement, 2020). The difference is more pronounced for transfer students and first-generation students, with 15% for transfer compared to 30% for students who started as freshmen, and 17% for first generation vs. 27% for non-first generation students (National Survey of Student Engagement, 2020). Traditional undergraduate research experiences tend to be exclusive, with a single or a small number of undergraduates working under the direction of faculty and/or graduate students (Russell et al., 2007). The limited number of opportunities serve only a fraction of students, resulting in competition and selective screening (Association of American Colleges and Universities, 2007). While many programs target historically underserved populations (National Academies of Sciences, 2017), this has not changed the imbalance in participation rates (National Survey of Student Engagement, 2020).

Undergraduate research is a form of academic engagement, and academic engagement is central to theories and research on student success in college. Educationally purposeful activities are associated with persistence, academic achievement, satisfaction, and social engagement (Astin, 1999; Mayhew et al., 2016; Trowler, 2010). Fredricks, Blumenfeld and Paris (2004) identified three types of student engagement: behavioral, emotional, and cognitive. Behavioral engagement in school includes positive conduct (following rules, going to class), learning and academic tasks (concentration, effort, contributing in class), and participating in activities related to school (clubs, athletics, student government). Emotional engagement in school can involve affective reactions to school (happiness, interest, enthusiasm) and identification with school (sense of belonging, mattering). Cognitive engagement in school includes self-regulation (flexibility, willingness to work hard, handling failure) and investment in learning (effort invested in mastering knowledge and skills).

Beyond being a form of student engagement, Kuh attributes the benefits of undergraduate research to student-faculty interaction (Kuh, 2008), which appears in multiple theories of college student development. In Astin's (1999) Theory of Student Involvement, students learn by becoming involved, and student-faculty interaction is one of the most influential types of student involvement. In Tinto's Theory of Student Departure, students' degree of academic and social integration determines their persistence in or departure from college, and this integration is shaped largely by their interaction with faculty (Kim & Sax, 2017). Pascarella positions faculty as critical agents in his General Model for Assessing Change in college students, and Weidman's Model for Undergraduate Socialization prominently positions student-faculty interaction in the socialization process (Kim & Sax, 2017; Weidman, 1989).

Large Team Project-Based Learning Embedded in Faculty Research Large team project-based learning embedded in faculty research carries key aspects of project-based learning (PBL) identified by Krajcik & Blumenfeld (Krajcik & Blumenfeld, 2005). In PBL, learning focuses on a problem that is meaningful and important to the students (Krajcik & Blumenfeld, 2005). In LT-PBL-EFR, the problem is based in a faculty member's research, and students join teams that they find interesting. The second key feature of PBL is that "students explore the driving question by participating in authentic, situated inquiry – processes of problem solving that are central to expert performance in the discipline. As students explore the driving question, they learn and apply important ideas in the discipline" (Krajcik & Blumenfeld, 2005, p. 318). In LT-PBL-EFR, faculty recruit students to their teams because they want/need the students' expertise, which may be in the mentor's own field or others. Students on the teams apply knowledge and skills from their disciplines, and they seek out and learn new knowledge and skills as needed.

The third key feature of PBL identified by Krajcik & Blumfeld is that “students, teachers, and community members engage in collaborative activities to find solutions to the driving question. This mirrors the complex social situation of expert problem solving” (Krajcik & Blumenfeld, 2005, p. 318). In the context of LT-PBL-EFR in higher education, community members can include experts outside of the team, stakeholders, sponsors, and end-users. “Complex social situation of expert problem solving” occurs within LT-PBL-EFR, because large teams require collaboration and coordination within and between subteams. Additionally, students work alongside and in community with their instructors. As a former LT-PBL-EFR student explained, “These interactions have a different dynamic than the typical student-teacher relationship, as students are more like collaborators than pupils. The ability to work directly with researchers and graduate students was fantastic” (R. C. Reece, personal communication, May 13, 2014).

Krajcik & Blumfeld’s fourth key feature of PBL is, “While engaged in the inquiry process, students are scaffolded with learning technologies that help them participate in activities normally beyond their ability” (Krajcik & Blumenfeld, 2005, p. 318). This aspect is somewhat dated, as many learning technologies are now ubiquitous, but the use of technology is essential to most PBL teams. The fifth key element is that “students create a set of tangible products that address the driving question. These are shared artifacts, publicly accessible external representations of the class’s learning” (Krajcik & Blumenfeld, 2005, p. 318). The artifacts of LT-PBL-EFR vary by team and by the nature of each project. Deliverables at Georgia Tech have included research posters, presentations to sponsors, wikis of ongoing documentation and design work, prototypes, and deployments. Products are not typically “publicly accessible,” but the products are used by continuing students, faculty mentors, and project stakeholders.

Project-based learning leverages three learning theories: active construction, situated learning, and social interaction (Krajcik & Blumenfeld, 2005). In active construction, learners construct knowledge instead of having instructors identify/define/provide needed knowledge. In LT-PBL-EFR, faculty invite students to join their teams to help advance the project because they do not already have formulaic answers. While faculty provide guidance, students draw on prior knowledge and acquire/build new knowledge and skills. In situated learning, student learning is situated in authentic, real-world contexts (Krajcik & Blumenfeld, 2005). In LT-PBL-EFR, the real-world contexts are faculty research. Project-based learning also leverages social interaction. When students learn in community, they develop deeper understanding by discussing, sharing, and using ideas in community with peers and faculty. Whether a LT-PBL-EFR team leverages social interaction depends on the instructor. If he/she runs the team like a class of students doing independent projects, little social interaction will occur in the learning and application of ideas. This results in more work for the instructor, who would have to track many independent elements of the project. By leveraging social interaction, instructors give students space to collaborate, reducing the number of project elements involved. Instead of receiving updates from individual students, instructors receive updates from subteams, with individuals discussing their contributions. In the LT-PBL-EFR program on which this study focuses, social network analysis of student peer evaluations showed high levels of interaction between students within teams, indicating that social interaction is indeed a part of the process (Melkers et al., 2012; Sonnenberg-Klein et al., 2017, 2018c).

Large team project-based learning embedded in faculty research is a special case of PBL. Two characteristics differentiate it from other PBL experiences. The first is team size. While PBL involves social interaction, PBL student teams can be very small. In contrast, LT-PBL-EFR teams are large, with 10-20 students per team. The second characteristic differentiating LT-PBL-EFR from other PBL experiences is the nature of the projects. Project-based learning problems can take many forms, but LT-PBL-EFR projects are embedded in faculty research. The term "faculty research" includes a wide range of faculty-driven activities including research, design, development, and creative endeavors. Faculty can use LT-PBL-EFR teams to advance sponsored research, to explore new ideas in a low-stakes setting, or to advance community-focused projects. Large team project-based learning embedded in faculty research occurs organically when professors mentor large numbers of students within team contexts (C. L. Leverette, personal communication, February 23, 2023).

Vertically Integrated Projects Georgia Tech's VIP Program is a campus-wide LT-PBL-EFR program. Georgia Tech is a four-year Research I university located in Atlanta, Georgia, in the southeast US (*Carnegie Classifications*, n.d.). In Fall 2022, the campus enrolled approximately 18,500 undergraduate students (Georgia Institute of Technology, 2022a). The student body is 42% white, 27% Asian, 8% Hispanic/Latino, 8% black/African American, 4% two or more races, and approximately 2% of other races/ethnicities (Peterson's, n.d.). While it is a predominantly white institution, Georgia Tech is the largest producer of black engineers in the country (Smothers, 1994).

The VIP model involves institutional elements that support long-term LT-PBL-EFR experiences (VIP Consortium, n.d.). These include offering VIP for academic credit, which makes it accessible to students who might not otherwise participate in undergraduate research; enabling students to participate for multiple years, which is achieved through curricular policies that allow VIP credits to count toward degree requirements; and offering the course for letter grades, which holds students accountable. The model has been implemented at 47 colleges and universities around the world, with department-level, college-level, and camps-wide programs at a variety of institutions including public and private (Boise State University and New York University), large and small (University of Michigan and Rice University), Historically Black Colleges and Universities (Howard and Morehouse), Hispanic-Serving Institutions (Arizona State University and Texas A&M University), and a Native Hawaiian-Serving Institution (University of Hawaii). VIP directors have adapted the model to their institutional contexts. For example, Boise State University allows students to participate pass-fail instead of for a letter grade (contrary to the standard VIP framework), which enables their program to engage a larger proportion of their student population, which is primarily non-residential and nearly 20% nontraditional. The University of Hawaii program allows transfer students to join VIP teams the semester prior to transferring to the University, which engages students and instructors from their feeder community colleges. The program at Howard University has fewer financial resources to support research, so their program uses industry sponsorships (making projects less faculty-centered) to support VIP teams and provide summer internships for VIP students. While the VIP model provides an institutional framework, the VIP model is not required for LT-PBL-EFR. Some faculty maintain large teams of students without VIP Program support. (C. L. Leverette, personal communication, February 23, 2023).

In the study, semesters of participation in VIP represent dosage levels. Curricular policies on how VIP credits count toward degree requirements affect the number of students who participate from each major and the number of semesters of participation (Georgia Institute of Technology, 2022b; Sonnenberg-Klein et al., 2018b). Some policies enable students to fulfill key degree requirements, such as Junior Design (in Computer Science) or Senior Design (in Electrical Engineering, Computer Engineering, and Industrial Engineering). Policies in this category yield higher participation when students can begin fulfilling requirements in their first semester of participation, as in Computer Science, compared to threshold policies that require prior participation before credits can begin counting toward the design sequence. In the 2021-2022 academic year, 46% of Computer Science bachelor's degree recipients had participated in VIP, and 33% had participated for three or more semesters. In contrast, Electrical Engineering and Computer Engineering have threshold policies for VIP in Senior Design, alongside threshold policies that allow VIP to count as in-major electives after a minimum number of credits are earned. While 52% of students in the major had participated (higher than in Computer Science), only 16% participated for three or more semesters (substantially lower than Computer Science). Majors in which VIP counts only as a free elective have lower participation, with 20% in liberal arts majors and 3% in Business Administration. However, participation rates are also influenced by the types of projects offered. While VIP counts only as a free elective in Mechanical Engineering, many VIP projects are related to the major, and 30% of degree recipients in the major had participated for at least one semester. Once involved in the program, although the number of semesters of participation varies by major, it does not vary by race/ethnicity (Sonnenberg-Klein et al., 2018a) nor by match-mismatch between student major and instructor department (Sonnenberg-Klein et al., 2018b).

Theoretical Framework

The benefits of higher education to graduates' career outcomes can be framed through the theories of human capital, social capital, cultural capital, and signaling. Capitals are tangible and intangible resources and traits that can be converted into money or property (Bourdieu, 1986). In the context of this study, they are resources and traits valued by employers, allowing students to convert the resources/traits into employment. Signaling theory describes the process by which job applicants signal their fitness for employment to employers, and the process by which employers interpret them.

Human Capital Human capital theory is a seminal theory in economics, with five Nobel Prizes awarded to researchers for its development and related work (Sweetland, 1996). The theory posits that investments in people yield economic and psychological gains by "increasing the resources in people" (Becker, 1993, p. 10). Investments in human capital improve health, knowledge, or skills that can lead to greater earnings and wellbeing. The concept emerged in the 1600 and 1700s, but garnered greater attention in the 1960s when researchers in economics found correlations between education and wages (Kern, 2009). While investments in human capital can include migration and medical care, education is the most researched aspect of human capital building (Sweetland, 1996). On average, high school graduates earn higher salaries than non-graduates, and college graduates earn more than high school graduates (Mayhew et al., 2016). While college is a costly investment (financial expense, time, and lost wages while in college), it yields a 12-20% rate of return on investment (Mayhew et al., 2016).

Cultural Capital If economic success were closely tied to knowledge and skills, education would yield similar outcomes for all students. However, Bourdieu observed persistent differences in educational-career outcomes by socioeconomic class, and he developed the theory of cultural capital to explain the inequity (Bourdieu, 1986). His theory challenged "commonsense" views that correlated success/failure with aptitude, and challenged human capital models that assigned specific economic benefits to education regardless of socioeconomic class (Bourdieu, 1986, p. 17). Whereas Becker faulted lower classes, describing generation-to-generation legacies of "low education, welfare dependence, early pregnancy, and marital instability" (Becker, 1993, p. 16), Bourdieu maintained that children from low-socioeconomic families were locked out of economic success by a system of codes, symbols, and resources that are transmitted from each generation of the upper-class to the next. He said economists had "let slip the best hidden and socially most determinant educational investment, namely, the domestic transmission of cultural capital" (Bourdieu, 1986, p. 17). A key difference between human capital and cultural capital is intentionality. Human capital building is usually intentional – requiring time and resources – while cultural capital can be accumulated unconsciously. As Bourdieu explained, "Cultural capital can be acquired, to a varying extent, depending on the period, the society, and the social class, in the absence of any deliberate inculcation, and therefore quite unconsciously" (Bourdieu, 1986, p. 18).

Cultural capital can exist in three states: an embodied state, an institutionalized state, and an objectified state. The embodied state includes dispositions, speech patterns, and ways of interacting. These can include body language, mannerisms, form of greeting, tone of voice, ease of conversation, and familiarity/comfort with gendered and classed topics. Within the context of employment, the embodied state can include performance of the employer's culture, both personal and professional. This might include using terminology and discussing practices specific to the field (professional culture) or discussing hobbies associated with privilege (boating, golf, etc.) or sports (an interest more common among males). Speech patterns and ways of interacting associated with affluence and white-maleness are also forms of cultural capital, putting people of color, women, and low-socioeconomic job applicants at a disadvantage in the job market. In his analysis of the Ebonics controversy, which was shaped by both education and employment issues, Collins centers Standard English as a commodity "necessary for market success," a form of cultural capital defined by the dominant class (Collins, 1999, p. 214). As a result, job applicants who use African American English Vernacular, who use pronunciations associated with African

Americans (“ax” vs. “ask”), or have non-American or non-European accents hold less cultural capital than speakers of Standard English (Cole et al., 2022).

The institutionalized state of cultural capital consists of credentials associated with institutions. Knowledge and skills gained in college represent human capital, but a college degree confers cultural capital in the institutionalized state (Bourdieu, 1986). An academic credential “confers on its holder a conventional, constant, legally guaranteed value with respect to culture” (Bourdieu, 1986, p. 20). The cultural capital of a degree is associated with degree level (associate’s degree, bachelor’s degree, etc.), field of study (art history, civil engineering, etc.), and the reputation of the degree-granting institution. Although institutions of higher education are accredited to ensure quality across degree programs, degrees from prestigious and highly selective institutions carry greater cultural capital (Rivera, 2016). This places non-white and non-affluent students at a disadvantage in the job market because institutional selectivity and prestige are shaped by race and social class. Twenty-nine percent of black and 22% of Hispanic college students attend highly or moderately selective four-year institutions, compared to 41% of white and 58% of Asian students (National Center for Education Statistics, 2023b). Elite schools confer greater cultural capital and disproportionately serve the affluent (Rivera, 2016), while for-profit schools confer less cultural capital and disproportionately serve black and Hispanic students (National Center for Education Statistics, 2023a). This study focuses on a single institution, implying equal cultural capital for the credential across degree recipients.

Because cultural capital is defined as symbols, possessions and ways of being that can be converted into money, a key aspect of cultural capital is who defines the symbols, possessions and ways of being that are valued. “Scholars such as Gloria Ladson-Billings and Dolores Delgado Bernal have asked: whose knowledge counts and whose knowledge is discounted?” (Yosso, 2005, p. 69). Cultural capital does not define the *right* way of being. Instead, it defines the economically beneficial way of being, which is decided by the dominant (white male) social class. While Bourdieu's theory shifted fault from low socioeconomic families to a system of capitals, use of the word "cultural" in cultural capital is loaded. The term lends itself to deficit thinking, with some families lacking the culture needed for socioeconomic success. These families possess symbols and ways of being that differ from the dominant white-male class – symbols and indicators of non-whiteness, non-maleness, and non-affluence – making them deficient in cultural capital. Yosso moves away from this deficit mindset by exploring the concept of cultural wealth, "the array of cultural knowledge, skills, abilities and contacts possessed by socially marginalized groups that often go unrecognized and unacknowledged" (Yosso, 2005, p. 69). With her approach, Yosso seeks to develop schools that acknowledge the cultural wealth that students of color bring to the classroom. However, in the transition from college to the workforce, employers decide which symbols and ways of being "count," which places non-white non-male non-affluent students at a disadvantage.

Social Capital Social capital, the third capital in the framework, is derived from an individual's connections with others, and the resources that people in their social network can provide (Bourdieu, 1986). In the years up to and in their transition to the workforce, students leverage their social networks to understand employer culture, to find out about employment opportunities, and as references as they apply for jobs. While this study focuses on the transition from college to the workforce, it is important to recognize that students face different obstacles and draw on different networks as they approach that transition. In a study of obstacles faced by students who sought counseling through their university, Lucas and Berkel found African American students perceived barriers to their vocational goals, and that Asian American students struggled with their vocational identities (Lucas & Berkel, 2005). Ma & Shea found first-generation students perceive greater barriers to finding employment than their non-first-generation peers (Ma & Shea, 2021). Kodama and Huynh recommend expanding advising support for Asian American students, many of whom are first-generation college students who struggle with academic and career-related decisions, lack clarity in their career-preferences, and have low career readiness (Fouad et al., 2008; Kodama & Huynh, 2017). During college, students build social capital by developing connections with peers and instructors. Students who participate in internships and cooperative education (CoOp) also build connections with working professionals, enhancing their ability to find employment after college (Smith & Green, 2021).

Signaling Theory Students who complete college gain human capital through knowledge and skills; cultural capital through educational credentials, their institutions' reputations, and by learning the culture of their fields; and social capital through connections with peers, instructors, and potentially with professionals. These are referred to as capital because they can be converted into money through employment, and signaling theory provides a working model by which this conversion unfolds. The theory was developed in economics by Spence (1973) to describe how signals are used to inform decisions when decisionmakers do not have access to needed information. In his initial presentation of the theory, Spence described the signaling between job applicants and employers. Employers want to hire people with specific skills and attributes. Their ability to observe or confirm applicants' skills and attributes is limited, so they rely on signals that indicate fitness for employment (Spence, 1973). A tangible signal is a credential. Instead of testing applicants' mastery of the field, employers rely on education credentials, which are institutionalized forms of cultural capital. Less tangible signals are dispositions and ways of being, which are embodied forms of cultural capital that may or may not align with employers' work or personal cultures. Between the extremes of tangible credentials and less tangible behaviors are stories that applicants share at job fairs, in interviews, and in their resumes. These stories might demonstrate motivation, resilience, ability to work well with others, etc. In their analysis of employer perspectives, Anderson and Tomlinson found employers look for signals in five areas: qualifications and credentials (institutionalized cultural capital that signals human capital); personal and psychological qualities (cultural capital); work-related experience (another form of credential); person-organization fit (social and cultural capital); and other experiences (Anderson & Tomlinson, 2020).

Relevance to LT-PBL-EFR When students participate in LT-PBL-EFR, they gain human capital, cultural capital, social capital, and comPelling stories that they can use to signal their qualifications to employers. They gain **human capital** through the knowledge and skills that they learn and apply in their projects. In VIP, this includes disciplinary knowledge and skills as well as professional skills such as communication, collaboration, and leadership (VIP Consortium, n.d.). Analysis of university exit surveys showed VIP students more strongly agreed than non-participants that their Georgia Tech educations contributed to their understanding of technologies related to their field, their ability to work in multidisciplinary teams, and their ability to work with people from diverse backgrounds (Ludlum, 2015). Analysis of VIP peer evaluations showed student leadership activity (coordinating teams' work, serving as a technical leader) was not correlated with student academic rank, but with number of semesters of participation in VIP, showing both leadership growth and a dosage effect (Sonnenberg-Klein, 2023).

In LT-PBL-EFR, students also gain **cultural capital**. By applying their disciplinary knowledge to real projects, they learn the culture and practices of their field. For example, a Computer Science student developing code for a multi-semester project that would continue beyond their own involvement would learn to comment their code well enough for teammates to be able to understand, build on, and edit it. They would also learn to use branching and version control in a large-team context, as in industry. Through this process, the student would learn the culture of collaborative code-development. Students also gain recognition for working on larger more complex projects than typical students, and for working closely with faculty. This type of experience is usually reserved for high-achieving students (Association of American Colleges and Universities, 2007; Russell et al., 2007) and represent an informal credential. In LT-PBL-EFR, students also develop/apply dispositions and ways of being that are valued by employers. These include resilience, comfort with collaboration, and ability to navigate conflict.

Within LT-PBL-EFR students also build **social capital**. Because students join teams they are interested in, peer interactions are shaped more by project interests than by student major or demographics. Social network analysis of peer evaluations show that within their VIP teams, students interacted more often with students of other races/ethnicities than their own, and more often with students from other majors than their own, connecting students with a wide range of peers (Sonnenberg-Klein et al., 2017, 2018c). Student-faculty relationships within VIP also differ from those in regular classes, replacing the instructor-pupil dynamic with collaboration (R. C. Reece, personal communication, May 13, 2014). Participating in VIP for multiple semesters may strengthen connections with faculty and peers, further increasing students' social capital.

Finally, through LT-PBL-EFR, students gain comPelling stories that they can use to **signal** their qualifications to employers. These can be considered through the five signaling areas identified by Anderson and Tomlinson: qualifications and credentials; personal and psychological qualities; work-related experience; person-organization fit; and other experiences (Anderson & Tomlinson, 2020). The first signaling area, credentials (education requirements), are used primarily as a screening mechanism. In the second signaling area, employers look for personal and psychological qualities such as confidence, resilience, adaptability, and the ability to manage relationships (Anderson & Tomlinson, 2020). In LT-PBL-EFR, students build resilience as they encounter obstacles and learn from their failures. They learn adaptability when they take on new roles in their teams, and when parameters of the project change. By collaborating with team-mates, students also develop and apply relationship skills. In the third area, employers look for work-related experiences, which includes experience in and knowledge of the field. Large team project-based learning embedded in faculty research exposes students to practical applications in their fields. While not as immersive as CoOps and internships, LT-PBL-EFR would provide more signaling strength than regular classroom experiences. In the fourth area, employers look for signals of person-organization fit, which is mentioned across the literature on employability (Anderson & Tomlinson, 2020; Finlay & Coverdill, 2002; Tomlinson & Anderson, 2021; Williams et al., 2016). Anderson and Tomlinson describe it as mindset, passion, behaviors and personal qualities (which overlaps with the second signaling areas, psychological qualities), but also includes “networks and connections” (social capital), and “appropriate skills and behaviors,” which is enactment of employers’ culture (cultural capital) (Anderson & Tomlinson, 2020, p. 3). Stories from LT-PBL-EFR could signal these traits – having sought-out an interesting project, being passionate about the project on which they worked, working collaboratively, etc. The fifth of Anderson and Tomlinson’s signal areas is other experience that makes applicants stand out, such as clubs, leadership, and notable non-academic activities. In this area LT-PBL-EFR would also provide comPelling stories, because the projects are unusual and interesting, and students who participate for multiple semesters take on increasing leadership responsibilities (Sonnenberg-Klein, 2023).

Marbles in Jars If a child were presented with marble-filled jars and asked to choose the best, he/she would only be able to see the marbles along the walls of each jar. He/she would not have all of the information needed, so they would base their choice on the marbles they could see. If the child were familiar with and liked two kinds of marbles, such as aggies and cat's eyes, and if the child wanted to be confident that they were getting the best option, they would choose the jar showing the most aggies and cat's eyes. Extending this idea to workforce hiring, employers have limited information, and they are drawn to applicants who display qualities and attributes they readily recognize and value. Under this framing, all applicants send signals, but employers do not as readily recognize or value signals of non-whiteness and non-affluence. Anderson and Tomlinson found employers look for credentials; work-related experience; qualities such as resilience, confidence and adaptability; and person-organization fit (shared mindset, common passions, etc.). If LT-PBL-EFR enables students to gain work-related experience and demonstrate/develop qualities desired by employers in contexts related to their majors, students would be better positioned to send signals valued by employers. Students already well-positioned to send these signals (white, male, affluent) would gain useful skills, but the LT-PBL-EFR experiences may or may not improve their odds of finding a job. Employers place less value on signals associated with non-whiteness, non-maleness, and low SES. For these students, LT-PBL-EFR may provide signals that improve their odds of finding a job. If this were true, then LT-PBL-EFR would yield greater job placement gains among non-white, female, and low SES students.

2 REVIEW OF THE LITERATURE

Higher Education and Career Outcomes

Education beyond the 12th grade confers a myriad of benefits including learning and cognitive changes, higher quality of life, socioeconomic gains, and benefits that are transmitted to the children of college graduates (Hout, 2012; Mayhew et al., 2016). The socioeconomic gains include higher employment rates, higher earnings, increased job status, greater job satisfaction, and positive returns on the investment of college attendance (Hout, 2012; Mayhew et al., 2016). These gains align with students' motivations for attending college. In the national Freshman Survey administered by the Higher Education Research Institute, the most common reason students reported for deciding to attend college was "to be able to get a better job," with 84% of students finding it very important (Stolzenberg et al., 2020, p. 42). Students' motivation to be able to get a better job maps to the concept of employability, which is the ability to find a job, maintain a job, or to find another job if desired (Hillage & Pollard, 1998; Suleman, 2018). When assessed across a population, job placement of graduates can be used as a measure of success for an institution or a type of degree (Mayhew et al., 2016; National Center of Education Statistics, 2020).

In the United States, college graduates' job placement has not been consistently reported. Between 2008 and 2022 the United States government enacted, strengthened, rescinded, and then continued to consider requirements for institutions to report and publicly disclose graduate job placement rates. In the 2008 renewal of the Higher Education Opportunity Act, institutions were required to provide current and prospective students with information on job placement for graduates and certificate holders (Carey & Kelly, 2011; Sampson, 2008). In a survey of 152 colleges and universities conducted three years after the policy was put in place, one-third of institutions had not complied with the 2008 mandate (Carey & Kelly, 2011). In 2010, to address predatory practices at some for-profit institutions, the Department of Education (DOE) established a gainful employment reporting requirement (Baker, 2011; U.S. Department of Education, 2014). A reporting methodology was not established until 2014, so from 2011 until 2014, institutions were only required to report employment data if their states or accrediting agencies required it as well (Baker, 2011). In 2015 DOE began publishing post-enrollment earnings in a web-based consumer tool called the College Scorecard (Itzkowitz, 2022). Instead of relying on institutions to report earnings, the government utilized student loan information (which includes information on the institution attended and program of study) alongside borrowers' post-college earnings reported in tax filings. With these, the College Scorecard reports post-enrollment earnings by institution and by individual degree-programs within institutions (Itzkowitz, 2022; U.S. Department of Education, n.d.). In 2019, DOE rescinded the reporting requirement (U.S. Department of Education, 2019), citing flaws in the debt to earnings rate formulas; the failure to account for non-institutional factors that influence student outcomes and earnings; and the unfair targeting of some institutions (for-profits) over others (non-profits). Although the reporting requirement was rescinded, earnings data continued to be reported in the College Scorecard. In 2022 DOE initiated a negotiated rulemaking process under Title IX, seeking consensus among representatives from groups potentially affected by new rules, including civil rights organizations, and each type of institution (Jaschik, 2022; U.S. Department of Education, 2021, 2022).

Theories Framing Research on Career Outcomes Across the literature, research on career outcomes for college graduates has primarily been framed by theories from four areas: human capital theory, career development theories, and status attainment theories (Mayhew et al., 2016), along with recent and novel work in signaling theory. Holmes (2013) posited three overarching perspectives on graduate employability that reflect theories of human, social, and cultural capital that are woven into the aforementioned theories.

Human capital theory is a seminal theory in economics, with five Nobel Prizes awarded to researchers for its development and related work (Sweetland, 1996). The theory posits that investments in people yield economic and psychological gains by “increasing resources in people” (Becker, 1993, p. 11), with education the most researched aspect of human capital building (Sweetland, 1996). Fitting under human capital theory is the idea of “employability as possession” described by Holmes, with employability based on skills and behaviors that can be learned (Holmes, 2013, p. 542), which ties to the conveyance of human capital.

The second set of theories that shape research on graduate employability are career development theories. The theories connect higher education with dimensions of career development, such as career self-efficacy and professional identity (Mayhew et al., 2016). The theories include Holland’s theory of vocational development (J. L. Holland, 1959), the dispositional theory of job attitudes (Staw & Ross, 1985), and social cognitive career theory (Lent et al., 1994; Mayhew et al., 2016). Holland’s theory of vocational development is a theory of personality development in which each student’s interests and aspirations interact with their educational contexts, with person-environment fit driving changes in interests and aspirations (J. L. Holland, 1959; Mayhew et al., 2016; Spokane et al., 2002). The theory is primarily used in research on college major selection, relationship between gender and college major, and congruence between major and profession. The second career development theory, dispositional theory of job attitudes, posits that student (or employee) dispositions toward their careers affect satisfaction, which shapes quality of work and life (Mayhew et al., 2016; Staw & Ross, 1985). The theory is primarily used in studies on job satisfaction. The third career development theory, social cognitive career theory, was developed by Lent, Brown and Hackett to explain relationships between career-related interests, career goals, and career-related actions (Lent et al., 1994). Social cognitive career theory builds on Bandura's social learning theory in which there is mutual interaction between an individual, their environment, and their choices (Bandura, 1977; Lent et al., 1994). The social cognitive career theory model involves ten factors, but in a simplified model, students' backgrounds and experiences shape their self-efficacy and career goals/actions; their career goals/actions shape their self-efficacy and experiences; and their self-efficacy shapes their experiences and career goals/actions (Lent et al., 1994). While employment is included in the full model, research employing social cognitive career theory has widely focused on self-efficacy, not job placement (Lent & Brown, 2019).

Status attainment theory, the third theory shaping research on career outcomes, applies status attainment and social mobility models to student access to higher education, progression through college, and subsequent careers (Blau & Duncan, 1967; Mayhew et al., 2016). Status attainment theory intersects with theories of social capital and cultural capital, because college can yield disparate outcomes for students from different socioeconomic classes (Bourdieu, 1986). Through the lens of status attainment theory, if students of low socioeconomic backgrounds lack access to competitive institutions, and if students from competitive institutions are more employable, then the sorting function of higher education reproduces social inequity. This represents what Holmes calls a positional approach to employment where "education serves to reinforce existing patterns of the way that advantage and disadvantage are distributed within society, to reinforce social position and societal stratification" (Holmes, 2013, p. 547), clearly rooted in cultural capital.

Signaling theory, a relevant but less applied theory, relates to career development theories. The theory was developed by Spence in economics (1973) to describe how signals are used to inform decisions when the decisionmaker does not have access to needed information. In his initial presentation of the theory, Spence described the signaling between job-applicants and employers. Employers want to hire people with specific skills and attributes, but their ability to observe or confirm applicants' skills and attributes is limited. Instead of rigorously assessing their skills, employers rely on signals that indicate applicants' fitness for employment (Spence, 1973). A tangible signal is level or type of education: Instead of testing applicants' mastery of the field, employers consider students' education credentials. Subtle signals are speech pattern and behaviors that aligns with employers' work or personal culture, which has implications for socially and culturally marginalized applicants. Signaling theory fits under Holmes' perspective of "employability as processual," in which students with similar qualifications are differentially effective at persuading employers of their employability (Holmes, 2013, p. 548). Holmes presents the processual model as the most realistic and maintains that students can learn how to signal their qualifications effectively. Holmes' framing does not address equity, though – students who send similar career-related signals may see very different results if they are from different backgrounds.

Research in the US With differing histories of evolution of higher education, researchers in the United States and Europe have taken different approaches to studying college graduate employment and employability. The United States was the first country to realize mass higher education. Social changes in the 1890s supported organic expansion of higher education through the 20th century, with colleges and universities feeding and being shaped by a continually evolving workforce (Goldin & Katz, 1999). Nationally, there is a clear correlation between further education and higher employment rates (National Center of Education Statistics, 2020). Men and women with bachelor's degrees earn 37% and 39% more than peers without degrees (Mayhew et al., 2016). Degree completion carries more benefits than four-years of college attendance without a degree, with 15% and 12% higher wages for men and women holding bachelor's degrees compared to non-degree holders who attended four years of college (Mayhew et al., 2016). This gain is referred to as a credentialing or “sheepskin” effect (Mayhew et al., 2016, p. 433), with a college degree signaling employability more effectively than a comparable number of years in school. In addition to higher wages, graduates also see greater growth in occupational status in their careers and enjoy a 12% to 20% rate of return on their investment (Mayhew et al., 2016, p. 433). Internships and higher grades are associated with higher employments rates, with rates varying by major (Mayhew et al., 2016).

In a synthesis of research on higher education and career outcomes in the United States, Mayhew et al. identified sixty-two rigorous studies published between 2000 and 2013 (Mayhew et al., 2016). The studies examined post-college career outcomes across a multitude of student and institutional factors, examining outcomes such as earnings, job satisfaction, satisfaction with salary, major-job congruence, gender and profession, rate of return, growth in occupational status, and hours worked. Research on employment rates at the time of graduation was limited. Of the sixty-two studies, two examined employment rates (Mayhew et al., 2016). In the earlier of the two studies, Wolniak and Pascarella (2007) studied a variety of employment outcomes by institution type, comparing outcomes for work college graduates with graduates from liberal arts colleges and regional institutions. They found higher full-time employment rates among work college students compared to liberal arts college students. In the second study that examined employment rates, Long (2010) conducted a longitudinal analysis of three cohorts of students from three decades, also examining a variety of career outcomes. In employment he found that labor force participation rates increased with additional years of education, but his study did not focus on employment at the time of graduation.

Career development courses have also been a focus of research, but job placement is rarely examined. Folsom and Reardon (Folsom & Reardon, 2003) conducted a survey of the literature on career development course outputs and outcomes, identifying forty-six studies published between 1976 and 2001. Most of the studies used pre/post-tests to measure changes in traits such as occupational knowledge, career decision-making, and vocational identity. Of the forty-six studies, five examined effects outside of courses, such as retention to the next term and time to graduation, but none examined job placement.

Beyond research on work-based learning and career development courses, there is minimal research on other experiences that support college graduate job placement. Two studies examine career outcomes and high-impact experiences. The study by Wolniak and Engberg (2019) examined correlations between five high-impact experiences and early career earnings. However, the authors only analyzed salaries for graduates who were employed, and they did not examine employment rates. Miller et al. examined correlation between job placement prior to graduation (the outcome in this study) and participation in high-impact experiences, and they found three experiences correlated with job placement: internships, culminating capstone courses, and service learning (Miller et al., 2018). While not focusing on high-impact experiences, in two studies Hu and Wolniak examined the effect of academic engagement and social engagement on early career earnings among graduates who were employed. They found academic engagement associated with higher salaries among men, and social engagement associated with higher salaries among women (Hu & Wolniak, 2013). For social engagement, they found a positive correlation with earnings for students in Science, Technology, Engineering and Math (STEM), and no significant correlation with non-STEM student earnings (Hu & Wolniak, 2010). For academic engagement, they found a positive correlation with non-STEM student earnings, and a slightly negative correlation with STEM student earnings (Hu & Wolniak, 2010).

Research in Europe In contrast to the intertwined expansion of the US workforce and higher education in the 20th century, the expansion of higher education in Europe is more recent (Calderon, 2018), and the establishment of new institutions sparked a "crisis of confidence" in institutional quality (Bassett & Tapper, 2009, p. 130). In 1999, ministers of higher education across Europe signed the Bologna Declaration to standardize educational systems across the continent (European Commission, 2018). Under the Bologna Process, institutions implemented Bachelor's, Master's and Doctoral degrees, with Bachelor's degrees to be "relevant to the European labour market" (Teichler, 2011, p. 5). As a result, European-wide, national, and disciplinary-level qualification frameworks were developed for higher education with a focus on learning outcomes, competences and employability (Teichler, 2011). Whereas research in the US has focused on the social and economic benefits of college, European researchers have primarily focused on dimensions of employability. Governments maintain that specific skills will help students find jobs, so institutions and researchers have worked to identify these skills, producing lists with items such as verbal communication, problem solving, numeracy, and IT skills (Holmes, 2013). The focus on these competencies have led to the vocationalization of higher education (Holmes, 2013; Suleman, 2018). Whereas the United States grapples with accountability efforts at the K12 level (Abendroth & Porfilio, 2015), Europe's Bologna Process has driven this type of scrutiny of higher education (Goldin & Katz, 1999; Suleman, 2018). As Tomlinson (2017) describes:

> One salient approach which has been popular amongst policy makers and within certain quarters of the higher education community has been the promotion of graduates' so called 'employability skills.' The underlying assumption here is that whatever deficits graduates continue to have after acquiring technical or subject-specific knowledge can be plugged by the acquisition of additional sets of skills which add value to their profiles. (p. 16)

Suleman conducted a survey of the literature, compiling a taxonomy of research methods and findings on employability skills (Suleman, 2018). While she found agreement across the studies on some cognitive, technical and relational skills, she concluded that identification of specific employability skills is impossible. While some researchers acknowledge the wide variety of categorizations and lack of agreed-upon classifications, the quest to categorize and build on existing classifications continues. Instead of studying correlations between student traits and educational experiences, most researchers ask students, faculty and employers to review lists of skills and rank their importance (Holmes, 2013). Holmes is especially critical, “Crucially, all such studies investigate the *expressed ‘perceptions’* of the respondents: none attempt to devise some form of objective measure of the purported skills and/or attributes, nor even discuss the theoretical and practical prospect of so doing” (Holmes, 2013, p. 546, emphasis in original).

Williams et al. conducted a broad survey of the literature for operationalizations of employability published between 1960 and 2014, not restricted to higher education (Williams et al., 2016). Terminology preferences differ between the United States and Europe, so Williams et al.’s search terms yielded matches from European authors. Of three papers that involved college students or recent graduates, one focused on career development during college, and two focused on career management after college, but not job placement. Despite the crises of confidence in higher education in Europe, a recent report shows that bachelor’s and master’s degree recipients in Bologna Process nations see lower unemployment rates and higher salaries than peers without college degrees, but rates vary dramatically by nation (Education & European Commission, 2018).

Discrimination and Privilege in Employment

In a meritocratic society, people “get what they deserve based solely on their individual efforts” regardless of race, class, or other attributes (Delgado Bernal & Villalpando, 2009, p. 80).

While this is an ideal for which to strive, it does not reflect society. Distributions of resources and opportunities are not based solely on merit; they are shaped by group membership and social class to such an extent that the privileging of high social classes over lower social classes and the privileging of white men over non-whites and women seems normal. A prime example of this lies in college graduates' transition from higher education to the workplace. In this transition, employers serve as gatekeepers to the workforce, judging the merit and worth of recent graduates (Chugh & Brief, 2008). The seemingly objective process is based on credentials and prior experience, but it privileges white men and the affluent without justification – discrimination by definition. Only by examining and understanding the mechanisms of discrimination and privilege can the processes begin to be dismantled.

Theories and Mechanisms Discrimination occurs when an action, such as a hiring process, disadvantages one group over another (Fibbi et al., 2021; Pager & Shepherd, 2008). While prejudices and stereotypes can lead to unfair treatment, discrimination involves the *act* of unfair treatment. Definitions of discrimination vary by discipline, with the phenomenon studied in a wide variety of fields. The 2004 National Academies Panel on Methods for Assessing Discrimination drew experts from economics, psychology, sociology, public policy, law, criminal justice, and statistics. The panel adopted a social science definition of discrimination with two components: differential treatment, and disparate effect (National Research Council, 2004). In differential treatment, individuals are treated unfairly because of their race. Prior to the 1964 Civil Rights Act, differential treatment in hiring was overt, as in job advertisements that stated "white men only" (Chugh & Brief, 2008, p. 336). Overt differential treatment in hiring has been replaced by subtle forms, sometimes with discriminators not realizing they are treating people differently (Bertrand et al., 2005; Fibbi et al., 2021; National Research Council, 2004). Disparate effects, the second component of discrimination, occur when rules and standards are applied equally to all people, but the standards are inadequately justified and place a racial group at a disadvantage (National Research Council, 2004; Pager & Shepherd, 2008). This is difficult to recognize, because the standards are not explicitly racial (Pager & Shepherd, 2008). The definition of discrimination provides a framework for examining employer perceptions, with additional understanding supported by theories of discrimination: implicit discrimination, aversive racism, white supremacy, social magic, and cumulative disadvantage.

Implicit discrimination is driven by implicit attitudes, which are unconscious associations and attitudes individuals have toward social groups (Bertrand et al., 2005). Implicit attitudes can be measured with a computer-based tool that examines response times and concept associations, a dramatic difference from self-reporting (Greenwald et al., 1998). Individuals are more influenced by implicit attitudes when they are inattentive to tasks, under high cognitive load or time pressure, and when there is ambiguity in the decision being made (Bertrand et al., 2005). An earlier theory related to implicit discrimination is aversive racism, through which individuals consciously support equality but hold negative attitudes toward minorities, which affects their behavior. The individuals develop an aversion to the minority group, avoid interaction, are less friendly, and discriminate in hiring despite their conscious support of equality (Dovidio & Gaertner, 2000; Fibbi et al., 2021).

White supremacy is the centering of whites' perceptions and interests, with whites as the dominant social class defining society's norms (Christian et al., 2019; Gillborn, 2005, 2006). Whiteness and affluence are accepted as inherently good, while color and lower socioeconomic classes are accepted as less-than (Dovidio & Gaertner, 2000; Salter et al., 2018). Related to white supremacy are Bourdieu's theory of social magic (Ingram & Allen, 2019) and cumulative disadvantage (DiPrete & Eirich, 2006), both of which involve gain or loss of privilege. In Bourdieu's theory of social magic, social circumstances are transformed into personal traits. Through social magic, privilege is converted to an inherent property of a person – worthiness and merit. Through the lens of white supremacy, whiteness (and white-maleness) is a privileged social circumstance, which through social magic, conveys worthiness to whites (particularly white men). Where whites are advantaged, minority groups are disadvantaged. In their study of stratification dynamics, Blau and Duncan (1967) introduced the concept of cumulative disadvantage. Under the theory, many small differences in how a minority group is treated combine to produce large differences between groups, which leads to underrepresentation (Milkman et al., 2015). "Being black was a cumulative disadvantage because race had both direct and indirect effects on outcomes at different stages in the life course and because highly educated blacks received lower status returns than did highly educated whites (DiPrete & Eirich, 2006, p. 273)."

Employers' Perceptions Employers serve as gatekeepers, and their perceptions of recent graduates fall into three primary areas: judgement of institutional quality, judgement of applicants' experiences, and bias. Perceptions in these areas are shaped by race and social class and are enacted through the mechanisms above. Discrimination occurs when these perceptions and mechanisms lead to differential treatment and disparate impact.

The first area of employer perceptions of recent graduates is judgment of institutional quality. Employers seek a variety of skills and traits in job applicants. Because they cannot directly assess or observe applicants' skills, employers rely on signals that indicate ability or productivity, with this process described by Spence's signaling theory (Mayhew et al., 2016; Spence, 1973; Williams et al., 2016). Some signals represent minimum credentials, such as a specific degree or licensure. Other signals are subject to value judgements. In an analysis of hiring practices at an influential technology company, Ingram and Allen (2019) found the company requires “strong educational credentials” (p. 731). While this can be interpreted as academic achievement, our society highly values degrees from elite institutions, so a degree from a prestigious university would be more highly valued than a degree from less competitive institution (Rivera, 2016). In an audit study, in which fictitious resumes were submitted for job openings to test the impact of different applicant attributes, Gaddis (2015) found that resumes listing degrees from elite universities (Harvard, Stanford and Duke) yielded 1.7 times as many callbacks as equally matched resumes listing degrees from nationally ranked but less prestigious institutions (University of Massachusetts-Amherst, University of California-Riverside, and University of North Carolina-Greensboro). Applicants holding degrees from elite institutions also received more callbacks by phone, representing higher employer interest. In five of thirteen email callbacks, in which employers mistakenly quoted internal communications, the employer tied institution status to the desirability of the applicant: “Ok, she had me at Stanford. Eat our dust [competitor];” “Forget the others: HARVARD GRAD;” “Kids coming out of Duke are by far the most capable. Push this one to the top of the list;” “We had a real bright app pop up this morning—Stanford grad with great credentials;” “Harvard guy wants to work for us!” (Gaddis, 2015, p. 1471). Employer enthusiasm for applicants from prestigious institutions was clear.

While degrees from ivy league schools are more highly valued than other degrees, the schools’ admissions decisions are not based solely on academic and extracurricular accomplishments. Instead, ivy league schools consider the revenue each applicant would generate from tuition (Rivera, 2016). Although these institutions’ admissions are not based entirely on merit, graduates are viewed as “more” than others, holding more value in the workforce than other college graduates, which is social magic. Affluence gives students access to ivy league schools, this access is perceived as an accomplishment, and it increases the student's worth in employers' eyes. In a broad study of high-paying jobs, Rivera found employers' valuing of prestigious university credentials favors students of high socioeconomic status, because elite schools disproportionately serve the affluent:

> These [elite] schools are overwhelmingly homogeneous socioeconomically. . . For example, at Harvard College, which is by no means unique within its peer group, nearly half of the students come from families in the top 4 percent of household incomes. A mere 4 percent come from the bottom 20 percent (Rivera, 2016, pp. 273–274).

Employer preference for applicants from "good schools" is seemingly objective, but it results in discrimination through differential effects. White students are more likely to be affluent. Affluent students have greater access to prestigious schools, which do not base admissions entirely on merit. Because prestigious schools do not base admission entirely on merit, the hiring standard of having graduated from a "good school" is not justified, and it places minority applicants at a disadvantage, which constitutes differential effects. Cumulative disadvantage also contributes to inequity in hiring. A small proportion of students from prestigious schools are not affluent, but because minority students have less access to high quality K-12 educations, they are less likely to qualify for admission. This leads to even lower group representation in prestigious institutions, which leads to group differences in hiring.

The second area in employer perceptions of college graduates is judgement of applicant experiences. Beyond educational credentials, employers seek qualities such as resilience, confidence, adaptability, and person-organization fit (Anderson & Tomlinson, 2020). Like academic credentials, judgements of these attributes are socially and racially-rooted, with affluent students getting higher-paying jobs than their peers from the same institution (Rivera, 2016). The criteria by which applicants are judged are seemingly objective. Desirable traits such as passion and natural curiosity can be demonstrated through "good" extracurricular activities and "interesting leisure pursuits" (Ingram & Allen, 2019, p. 733). However, judgements of both are based on employers' values and social status, and employers respond more favorably to signals aligned with their own experiences (Anderson & Tomlinson, 2020; Derous & Ryan, 2019). Affluent students can engage in different activities than less affluent students, increasing their appeal to prospective employers. Being a "go-getter" or "self-starter" can be demonstrated by starting-up a company, which requires time and access to financial resources (Ingram & Allen, 2019, p. 733). High-status companies expect applicants to be highly accomplished in elite activities, and affluent applicants are more likely to play expensive sports that require special training or elite spaces, such as polo or dressage (Rivera, 2016). In contrast, less affluent applicants are more likely to participate in low-cost sports like basketball or soccer (Rivera, 2016). In an interview, the captain of the polo team would be more highly regarded than the captain of the soccer team, or a student who worked during college and did not have time for sports. There are also correlations between race and participation in different types of extracurricular activities, as well as student access to resources. White students participate in internships at higher rates than Asian, black/African American, and Hispanic/Latino students (National Survey of Student Engagement, 2020), with these experiences considered prestigious, reserved for highly qualified students. Study abroad, which is well regarded and requires resources, has higher participation among White and Asian students than black/African American and Hispanic/Latino students (National Survey of Student Engagement, 2020). The conversion of social position (access to startup funding, resources for elite sports, time to start a company or be an athlete) to personal traits (leader, go-getter) constitutes social magic. The valuing of activities dominated by affluent whites represents white supremacy. As Ingram and Allen explain, "It is through subtlety and lack of an explicit description of what constitutes the 'ideal' graduate that processes of discrimination are first obscured, and second, transformed into 'objective' criteria which naturalise [sic] privilege as earned and developed skills" (Ingram & Allen, 2019, p. 729). The imbalance of access to resources needed to participate in experiences valued by employers also represents cumulative disadvantage. When the valuing of these experiences affects hiring decisions, it puts lower socioeconomic and minority students at a disadvantage, representing discrimination through disparate effects.

While judgements of institutional prestige and student activities are shaped by race and social class, applicants also contend with employer bias, the third aspect of employer perceptions. In a resume audit study, Bertrand & Mullainathan (2004) found that comparable resumes with white names generated 50% more calls for interviews than resumes with black names. Additionally, resumes with white names and high qualifications yielded greater gains in callbacks than non-white high-qualification resumes, widening the racial gap for highly skilled applicants. Examining the difference between "whitened" first names and experiences for black and Asian applicants, Kang (2016) found whitened first names yielded 30% and 57% more callbacks for black and Asian applicants; whitened experiences (hiking, travel) yielded 80% and 43% more callbacks; and whitening of both yielded 155% and 83% more callbacks (S. K. Kang et al., 2016, p. 491). Examining race and institutional prestige, Gaddis found black applicants from elite institutions did only as well as similarly qualified whites from less competitive institutions (Gaddis, 2015). "These racial differences suggest that even a bachelor's degree from an elite institution cannot fully counteract the importance of race in the labor market" (Gaddis, 2015, p. 1451). Bias in callback rates by race/ethnicity has not shown signs of decreasing. A meta-analysis of twenty-four audit studies published between 1990 and 2015 found no decrease in discrimination against black/African Americans, with a modest decrease in discrimination against Hispanics/Latinos

(Quillian et al., 2017). Whites received 36% more callbacks than blacks/African Americans and 24% more than Hispanic/Latinos. An audit study by Milkman et al. found women and non-whites discriminated against other women and non-whites, showing white supremacy, a centering of white-maleness imposed by the dominant class and accepted by the oppressed (Milkman et al., 2015). These audit studies document differential treatment. The mechanisms for this are implicit discrimination and aversive racism. Implicit discrimination occurs when resume reviewers perceive minority applicants in a negative way, dismissing them at higher rates than white applicants with comparable qualifications. Implicit bias has a stronger effect when people are inattentive to the task, under high cognitive load, or under time pressure, all common in the workplace (Bertrand et al., 2005). Aversive racism moves beyond subconscious attitudes, with resume reviewers having an aversion to the idea of hiring a minority. This is how a minority graduate from an elite institution could be overlooked. The credential is highly desired, but an aversion leads to rejection (Dovidio & Gaertner, 2000).

Summary In the transition from college to the workforce, employers judge the merit and worth of recent graduates. Social magic shapes employer judgements of institutional quality and student experiences, with a person's social status transformed into valued traits inherent to the person. At the same time, white supremacy prioritizes experiences of affluent whites and devalues minority experiences. Employer standards are seemingly objective but some standards are unjustified, putting lower socioeconomic and minority students at a disadvantage, which is discrimination through disparate effects. These effects are exacerbated by cumulative disadvantage, with differences in treatment at multiple stages leading to underrepresentation. Beyond employer judgements of institutional quality and student experiences, employer bias leads to discrimination through differential treatment. Employers reject minority applicants at higher rates than white applicants when their names or activities on their resumes signal minority identity. This can be implicit discrimination, in which subconscious negative attitudes toward minorities affect decisions, even when the decision maker supports equality and when the decision maker is a woman or minority. The rejections can also result from aversive racism, where individuals who are not consciously biased develop aversions to interaction with minorities, which affects their hiring decisions. Employers judge applicants' merit. An alternative meaning of merit is "deservingness" (Sommerland, 2014, p. 2325). Mechanisms of discrimination (implicit discrimination, aversive racism, white supremacy, social magic, and cumulative disadvantage) lead employers to find non-white applicants less deserving of employment.

3 METHODOLOGY

Propensity Score Analysis

The study employs inverse propensity score weighting for multiple treatment levels, with treatment levels of 0, 1, 2, and 3 semesters of VIP participation. Propensity score analysis methods enable researchers to draw causal inferences from observational data (Guo & Fraser, 2015). Although randomized experiments are superior to nonrandomized studies, Rubin maintained that ignoring observational data would be counter-productive, particularly in situations where randomized experiments pose ethical problems, are cost-prohibitive, or would take too long to conduct (Rubin, 1974). Experimental studies on programs that can affect professional development and career outcomes for college students pose ethical problems, because the potential impact of negative outcomes is too great: Assigning students to a treatment condition that could lead to unemployment or underemployment would be unethical. Within the context of higher education, propensity score analysis has been used to study the impact of experiences on academic performance and skill development, including the effect of community college attendance on bachelor’s degree completion (Dietrich & Lichtenberger, 2015); of freshman experience courses on GPAs and retention rates (Clark & Cundiff, 2011); of honors programs on students’ ability, motivation, creativity and academic achievement (Kool et al., 2017); and of undergraduate research experience on retention (Rodenbusch et al., 2016) and professional skill development (Carter et al., 2016). The methods have also been used to study the effect of different student experiences on labor market outcomes, such as the academic and labor market outcomes for transfer students compared to students who began at their institutions as freshmen (Xu et al., 2018); the impact of graduating from an HBCU on labor market outcomes (Price et al., 2011); the effect of earning a master’s degree on earnings (Titus, 2007); the effect of study abroad experiences on student socioeconomic status (Waibel et al., 2018); and the effect of gender and academic achievement on earnings after college for students majoring in science, technology, engineering and math (Olitsky, 2014). Some studies use survey instruments to collect data and use propensity score analysis to control for selection bias (Carter et al., 2016; Clark & Cundiff, 2011; Kool et al., 2017), while other studies rely on existing institutional data (Rodenbusch et al., 2016), state-level data sets (Dietrich & Lichtenberger, 2015), and national surveys (Olitsky, 2014; Titus, 2007; Waibel et al., 2018). This study makes use of existing institutional data from academic records and responses to an institutional survey administered to students prior to graduation.

Propensity score analysis methods were developed by Heckman (1979) in econometrics, a field built around structural equation modeling; and by Rosenbaum & Rubin (1983) in statistics, a field that developed around randomized experiments (Guo & Fraser, 2015; Sobel, 2005). The methods employ the Neyman-Rubin counterfactual framework developed for randomized experiments. (The framework is also referred to as the Rubin causal model and the potential outcomes model.) The foundation of the framework is that for each person in a study (i), there are multiple potential outcomes. In a study with a treatment condition ($W_i = 1$) and a nontreatment condition ($W_i = 0$), observation i would have an outcome of Y_{1i} under the treatment condition, and an outcome of Y_{0i} under the nontreatment condition. A general expression for the outcome is:

$$Y_i = W_i Y_{1i} + (1 - W_i) Y_{0\mathrm{i}} \qquad (1)$$

The framework involves factual and counterfactual outcomes. The observed outcome is the factual outcome. For an individual that received treatment, the counterfactual outcome would be the outcome without treatment. For an untreated individual, the counterfactual outcome would be the outcome with treatment. The effect of the treatment on observation i would be the difference between Y_{0i} and Y_{1i}. However, once the experiment has been conducted, only one outcome is observed for i, making it impossible to observe the effect of the treatment on the individual or group (P. Holland et al., 1985). In randomized experiments employing the Nyeman-Rubin counterfactual framework, the counterfactual outcome for the treatment group is assumed to be equal to the observed outcome of the control group (Guo & Fraser, 2015). The effect of the treatment is estimated by comparing mean outcomes for the treated, $E(Y_1|W = 1)$, to mean outcomes for the untreated, $E(Y_0|W = 0)$. The average treatment effect, τ, can be defined as a difference of the two means.

$$\tau = E(Y_1|W = 1) - E(Y_0|W = 0) \qquad (2)$$

To apply the Neyman-Rubin counterfactual framework to randomized studies, researchers must make fundamental assumptions. A key assumption is the ignorable treatment assignment assumption (Guo & Fraser, 2015; Rosenbaum & Rubin, 1983). Under ignorable treatment assignment, if matrix X consists of factors related to whether individuals are in the treatment or nontreatment group, conditional on covariates X, the assignment of individuals to treatment and nontreatment is independent of potential outcomes. Outcomes may of course depend on treatments, but treatment assignment does not depend on potential outcomes. For example, if terminally ill patients were assigned to an experimental treatment and non-terminal patients were assigned to traditional treatment, treatment assignment would depend on potential outcomes, violating the ignorable treatment assignment assumption. The assumption is also referred to as conditional independence, selection on observables, unconfoundedness, and exogeneity (Guo & Fraser, 2015), and it can be expressed as:

$$(Y_0, Y_1) \perp W|X \qquad (3)$$

Another assumption of the Neyman-Rubin counterfactual model is the stable unit treatment value assumption (SUTVA). Under this assumption, the outcome for an individual under a given condition is independent of the system used to assign treatment and of the treatment conditions assigned to other units. The SUTVA assumption is violated if the ignorable treatment assignment assumption is violated; if multiple versions of treatment are handled as equivalents in the analysis (residential and out-patient treatment); if interaction between participants affects outcomes (a disruptive student affects outcomes for others); or if treatments alter social conditions that then affect outcomes (a job training program affects employer demand for workers) (Guo & Fraser, 2015; Winship & Morgan, 1999).

In randomized experiments, correlation between a treatment and an outcome is not basis for causal conclusions. While the treatment may cause the outcome, the outcome may cause the treatment, or the treatment and outcome may be caused by another factor not included in the analysis. For example, if a career-preparation course were correlated with high job placement rates, the course may lead to high job placement (causation in one direction); the course may attract individuals who already have job offers (causation in the other direction); or another factor such as affluence may lead to both higher job placement and enrollment in the course (spurious correlation). Three criteria for a causal relationship are temporal order, with the cause preceding the effect; correlation; and that the correlation cannot be explained by another variable (Guo & Fraser, 2015). Causal inferences need to be guided by substantive theories, which includes choosing appropriate covariates, choosing an appropriate data analysis model, and basing theories on prior studies (Guo & Fraser, 2015). When analysis results support a causal hypothesis, investigators should discuss all alternative explanations (Cochran & Chambers, 1965).

Methods used to draw causal conclusions in randomized studies cannot be applied in the same ways to observational studies. In randomized experiments, individuals willing to participate in the treatment are randomly assigned to treatment and nontreatment (Heckman & Smith, 1995). In observational studies, a myriad of factors affect whether individuals participate in a program, including their willingness or ability to participate, administrators' conscious and unconscious selection criteria, and participants' ability or willingness to comply with the program. Beyond yielding imbalance in traits among participants and non-participants, factors affecting participation and compliance are often also correlated with potential outcomes. A student struggling in math is more likely to be referred to a tutoring program than a student who easily masters the material. An affluent student who plans to attend college would be more likely to enroll in an expensive college test preparation course than a low socio-economic student who is unsure about attending college. In randomized social experiments, selection bias is controlled by building both treatment and control groups with individuals willing to participate in the program, and denying treatment to a random selection of individuals (Heckman & Smith, 1995). When random selection yields differences between treatment and control groups, imbalances in randomized studies are accounted for with ordinary least squares regression. However, when ignorable treatment assignment is violated, as in observational studies, regression estimates of treatment effects are likely to be biased (Austin, 2011; Guo & Fraser, 2015). One approach to approximating a randomized experiment with observational data is to match treated and untreated cases on covariates. However, this is not possible when the number of covariates is large.

Beyond covariate matching, Rubin extended the use of the counterfactual framework to observational contexts through the use of propensity scores (Guo & Fraser, 2015). In propensity score analysis, all covariates are used to construct a score for each individual that estimates his/her/their probability of receiving treatment and that balances covariates between groups. Propensity score analysis methods rebalance data to resemble results of randomized experiments, so that treatment effects can be estimated. Guo & Fraser identified five types of treatment effects and emphasize the importance of defining which is being estimated. This study involves estimate of the average treatment effect (ATE). The ATE estimates the effect of the treatment for the full population if everyone received treatment. Another relevant effect is the average treatment effect on the treated (ATT), which compares observed outcomes for the treated with estimates of outcomes that would have occurred without treatment. The ATT and ATE are relevant to policy decisions because they focus on individuals willing to participate in the treatment (ATT), and on effects if the treatment was widely used (ATE). Both are discussed because they share similarities. Other estimations of effects are the average treatment effect for the untreated; marginal treatment effect, which relates to individuals indifferent to participating; and local average treatment effect, which estimates effects for participants who complied with the treatment (Björklund & Moffitt, 1987; Guo & Fraser, 2015; Heckman, 2005).

The propensity score (e_i) for observation i ($i = 1...N$) is defined as the conditional probability of being assigned to treatment ($W_i = 1$) instead of nontreatment ($W_i = 0$) given a vector of variables (x_i) that covary with treatment assignment.

$$e(x_i) = pr(W_i = 1 \mid X_i = x_i) \qquad (4)$$

In a study with only two possible conditions (treatment and non-treatment), the probability of being assigned to the non-treatment condition would be $1 - e_i$. The probability that an individual will receive treatment is a function of x_i, but assignment to one group or another is independent of x_i. This represents ignorable treatment assignment for individuals at a given x_i and can be expressed as:

$$x_i \perp w_i | e(x_i) \qquad (5)$$

At a given probability of receiving treatment, the average treatment effect at $e(x_i)$ would be expected to be the mean difference in outcomes for treated and untreated groups (Guo & Fraser, 2015).

Inverse Propensity Score Weighting Students in the VIP Program can participate for multiple semesters, which represents different treatment levels, or dosages. Guo & Fraser describe two main approaches for propensity score analysis with multiple treatment levels: matching with a scalar balancing score, and weighting with generalized propensity scores. In matching with a scalar balancing score, a single propensity score is estimated for each observation. The method is problematic, because the assumption of errors of constant variance in covariates is difficult to satisfy (Guo & Fraser, 2015). The other approach is to use inverses of propensity scores as weights in the estimation of treatment effects, with marginal density of the treatment used as a stabilizing factor (Guo & Fraser, 2015; Leite, 2016). This method can be applied to binary treatment conditions (treated, untreated) with propensity scores, and to multiple treatment conditions with generalized propensity scores. This section discusses weighting and analysis for binary conditions, and then generalized propensity scores for multiple dosages are discussed in the next section.

In survey research, weighting is used to correct for imbalances in response rates among groups and for differences in probabilities of group members being included in the sample (Dillman et al., 2014). In a survey study, if a simple random sample of cell phone numbers was used to select a sample, and if fewer senior citizens owned cell phones compared to other age groups, responses for senior citizens in the sample could be weighted more heavily to account for low probability of being included. Weights for each individual in the sample would be equal to the inverse of the probability of selection (Mercer et al., 2018). In propensity score weighting, weights are based on an inverse of probability of being in the observed condition, with the propensity score used as the probability (Guo & Fraser, 2015).

Under binary conditions, weights are defined differently in the estimation of ATT and ATE. When estimating the ATT, weighting is used to construct a control group that looks like the treatment group (Griffin et al., 2022). The weighting estimator is defined as (Guo & Fraser, 2015):

$$\omega_{ATT}(W, x) = W + (1 - W)\frac{\hat{e}(x)}{1 - \hat{e}(x)} \qquad (6)$$

Treated observations would be unweighted, because for $W = 1$, the weight would be 1. Untreated observations, $W = 0$, would have a weight of $\frac{e(x_i)}{1-e(x_i)}$, which are the odds of receiving treatment (probability of treatment divided by probability of not receiving treatment).

When estimating the ATE, weighting is used to make the treated group look like the control group (Griffin et al., 2022), and the weighting estimator is defined as (Guo & Fraser, 2015):

$$\omega_{ATE}(W, x) = \frac{W}{\hat{e}(x)} + \frac{1 - W}{1 - \hat{e}(x)} \qquad (7)$$

For a treated individual, $W = 1$, and the weight would be $\frac{1}{e(x_i)}$. For an untreated individual, $W = 0$, and the weight would be $\frac{1}{1-e(x_i)}$. A flaw with ω_{ATE} is that large weights can bias the results. A treated case with a low propensity score would have a large weight ($\frac{1}{e(x_i)}$), and an untreated case with a high propensity score would have a large weight ($\frac{1}{1-e(x_i)}$). To address overly influential weights, Robins et al. introduced a stabilizing term. The term is the marginal probability of the individual being in the condition that was observed, which is the sum of probabilities of being in a given condition across all cases in the condition, divided by the number of cases in the condition (Guo & Fraser, 2015; Leite, 2016). Where i represents the i^{th} observation out of n_1 treated cases, and j represents the j^{th} observation in n_0 untreated cases, the stabilized weights for ATE estimates are:

$$\omega_i(W = 1, x) = \frac{\sum_{i=1}^{n_1} \hat{e}(x_i)}{n_1} \cdot \frac{1}{\hat{e}(x_i)} \tag{8}$$

$$\omega_j(W = 0, x) = \frac{\sum_{j=1}^{n_0} \left[1 - \hat{e}(x_j)\right]}{n_0} \cdot \frac{1}{1 - \hat{e}(x_j)} \tag{9}$$

Generalized Propensity Scores Imbens extended propensity score weighting from binary conditions to multiple treatment levels with generalized propensity scores (Imbens, 2000). With the method, a balancing score is estimated for each individual for every possible level of treatment, yielding a vector of scores for each observation. Treatment conditions can represent exact dosages or wide groupings of dosage, such as low and high dosages (Godley et al., 2022). Unlike propensity scores under binary treatment, in which the propensity for treatment and propensity for non-treatment sum to 1, values in the vector of scores do not sum to 1. A strength of the generalized propensity scores approach is that it involves fewer assumptions than the scalar method for multiple treatments (Guo & Fraser, 2015).

Propensity score analysis employing inverse propensity score weighting for multiple dosage levels involves three assumptions: weak unconfoundedness (Imbens, 2000); positivity (also referred to as overlap); and no unmeasured confounders (McCaffrey et al., 2013). The assumption of weak unconfoundedness was developed by Imbens for propensity score weighting for multiple dosage levels. The propensity score analysis methods developed by Rosenbaum & Rubin assume strongly ignorable treatment assignment conditional on X (Rosenbaum & Rubin, 1983). If multiple doses are available, strongly ignorable treatment assignment requires treatment assignment to be independent of all potential outcomes, which is a difficult assumption to satisfy

(Imbens, 2000). In contrast, weak unconfoundedness assumes only pairwise independence of treatment assignment with each potential outcome (Imbens, 2000).

Under the positivity assumption, every individual has a non-zero chance of receiving any level of treatment (McCaffrey et al., 2013). The probability of being assigned to treatment cannot be zero (the basic assumption), but it also cannot be 1, because this would preclude the possibility of any other treatment assignment. A violation of the positivity assumption is implied when there is a lack of overlap in covariates between groups. For example, if a program only served students of low socioeconomic status, and if the non-treatment group included students of high socioeconomic status (precluding them from treatment), positivity would be violated.

The third assumption is that there are no unmeasured confounders. The assumption is not testable, but to reduce the potential of an unmeasured confounder, researchers include a broad set of covariates in the analysis (McCaffrey et al., 2013). Another approach to address potential confounders is to do a sensitivity analysis, wherein researchers determine the strength an unmeasured covariate would need to have in order to affect estimated treatment effects (Rosenbaum, 2005). However, sensitivity analysis methods for propensity score weighting for continuous or multiple treatments have not yet been developed. Rosenbaum's methods were developed for propensity score matching, but have not been extended to inverse propensity score weighting (Guo & Fraser, 2015; Leite, 2016). VanderWeele and Arah developed generalized equations for sensitivity analysis for binary treatments (treatment and control), and it can be used with inverse propensity score weighting, but their approach does not address multiple treatments or dosages (Rudolph & Stuart, 2018; VanderWeele & Arah, 2011). Because methods appropriate for multiple dosages have not yet been developed, the study does not include sensitivity analysis.

Data

The study draws data from two sources. The first is the Georgia Tech data warehouse, which includes information from academic records and participation in co-curricular programs such as living learning communities, fraternities and sororities, and NCAA athletics. The second source is the Georgia Tech Career and Salary Survey, which is administered to students prior to graduation by the Office of Academic Effectiveness.

A limitation of the survey is that it does not differentiate between students who have received but have not accepted a job offer, students who are deciding between multiple offers, and students who began their job search at a later date than their peers (family circumstances, change in plans, etc.). Another limitation is self-selection. Students who had not received job offers or who had their offers rescinded may have been less inclined to complete the survey. In Spring of 2020, when COVID began affecting schools and workplaces, some companies rescinded job offers (Hess, 2020). Salary survey completion rates are reported by academic year. COVID occurred in the second half of the 2019-2020 academic year, and the survey completion rate was 54%. In the following academic year, which includes Fall 2020, the completion rate was 49%. In the 2021-2022 academic year, it rebounded to 57%. Survey response rates for prior years were not readily available.

Participants

The study examined job placement for students who completed the Georgia Tech Career and Salary Survey prior to graduation in the 2017-2022 calendar years. This includes three calendar years before COVID began affecting schools and the economy (2017, 2018 and 2019), and three calendar years including and after COVID began (2020, 2021, 2022). Georgia Tech is a technologically focused institution, so participants primarily majored in computing and engineering. The sample frame was limited to survey respondents who reported planning to enter the job

Table 1

Citizens and Permanent Residents who Planned to Enter the Workforce

	Semesters of VIP									
Major	0	1	2	3	4	5	6	7	8	Total
Aerospace Engineering	320	34	24	6	3					387
Applied Lang/Intercultural St	12	2	1							15
Architecture	56	5								61
Biochemistry	53	2	2		1					58
Biology	63	11	3				1			78
Biomedical Engineering	453	40	13	23	11	5				545
Business Administration	852	23	4	1		1				881
Chemical and Biomolecular Eng	406	33	8	7	1	1				456
Chemistry	32	1	2							35
Civil Engineering	245	17	9	2						273
Computational Media	62	5	1	14	6					88
Computer Engineering	216	47	17	13	8	4				305
Computer Science	985	97	18	244	45	5			1	1395
Earth & Atmospheric Sciences	8									8
Economics	53	3	1	2						59
Electrical Engineering	257	44	20	11	12	2				346
Environmental Engineering	68	11	9	5	1		2			96
History, Technology, & Society	11	1	1							13
Industrial Design	82	16	2							100
Industrial Engineering	813	57	54	6	3	1				934
International Affairs	45	10	1	1	1	1		1		60
Lit., Media, & Communication	60	4	1							65
Materials Science & Engr	147	6	2	5	1					161
Mathematics	34	2	4	2						42
Mechanical Engineering	835	114	52	24	9	4	2			1040
Music Technology	9		1							10
Neuroscience	43	6	2							51
Nuclear & Radiological Engr	30	1								31
Physics	21	5	5	1	1	1				34
Psychology	33	3	4							40
Public Policy	25	6	1	1						33
Total	6329	606	262	368	103	25	5	1	1	7700

market after graduation (N = 8,436), which excluded students planning to go to graduate school, be self-employed, own their own business, and students not seeking employment. Because non-citizen non-permanent residents seeking to enter the workforce face different challenges than citizens and permanent residents, such as securing a work visa or searching for a job in their home country while living in the U.S., noncitizen nonresidents were excluded from the analysis. This left 7,700 survey respondents from 31 majors (Table 1). Of these, 1,371 (18%) had completed one or more semesters of VIP.

Propensity score analysis requires a non-zero chance of being treated at each dosage level. A very small number of students participated in VIP for five to eight semesters, so these treatment levels and cases were excluded. Because separate analysis would be done on white and non-white subgroups, only majors with at least one white and one non-white student at each treatment level were included. If students who participated for four semesters were included in the analysis, applying the restriction of one white and one non-white student at each treatment level would yield a sample of 3,995 students in six majors. Limiting the analysis to students who participated 0-3 semesters yielded a sample of 5,817 students in 11 majors, for a 46% larger sample size. Beyond maximizing the sample size, limiting the analysis to three semesters of participation was supported by a study on leadership development in VIP that found student growth through the third semester of participation, and no difference in leadership ratings in peer evaluations for third and fourth semester students (Sonnenberg-Klein, 2023). Additionally, because the CoOp degree designator requires three semesters of participation, limiting the analysis to zero to three semesters of VIP would allow for a more balanced comparison of the effects of the two programs.

The final sample was 55% white, 23% Asian, 8% Hispanic or Latino, 7% black or African American, 4% two or more races, 2% unknown, and less than 1% American Indian/Alaskan Native and Native Hawaiian or other Pacific Islander (Table 2), with 36% Female (Table 3).

Table 2

Treatment Level Frequencies by Race/Ethnicity, 11 Included Majors

	Semesters of VIP				
	0	1	2	3	Total
American Indian or Alaska Native	1				1
Asian	972	173	76	149	1370
Black or African American	281	43	18	31	373
Hispanic or Latino	412	36	17	17	482
Native Hawaiian or Other Pacific Islander	1				1
Two or more	190	24	9	15	238
Unknown	103	5	9	11	128
White	2785	219	97	123	3224
Total	4745	500	226	346	5817

Table 3

Treatment Level Frequencies for Male and Female Students, 11 Included Majors

Major	Male				Female				Total
	Semesters of VIP				Semesters of VIP				
	0	1	2	3	0	1	2	3	
Aerospace Engineering	255	27	13	4	65	7	11	2	384
Biomedical Engineering	162	16	4	7	291	24	9	16	529
Chemical and Biomolecular Eng	225	14	3	3	181	19	5	4	454
Civil Engineering	137	4	6	1	108	13	3	1	273
Computer Engineering	179	34	14	6	37	13	3	7	293
Computer Science	761	68	16	160	224	29	2	84	1344
Electrical Engineering	205	35	12	8	52	9	8	3	332
Environmental Engineering	13	2	2	1	55	9	7	4	93
Industrial Engineering	403	28	23	2	410	29	31	4	930
Materials Science & Engr	76	3	1	1	71	3	1	4	160
Mechanical Engineering	614	82	41	19	221	32	11	5	1025
Total	3030	313	135	212	1715	187	91	134	5817

Software

The analysis was done in R, a free open source software language for statistical computing, version 4.3.2 (The R Foundation, n.d.). Scripts were written in Rstudio, an integrated development environment that can be used to generate html, Microsoft Word, and pdf output (Posit, n.d.). Generalized propensity scores were estimated with the IPW package (Wal & Geskus, 2011, 2023). Effective sample sizes were computed with the WeightIt package (Greifer, 2023). Estimates of the effect of variables on job placement were estimated with weighted logistic regression through the survey package (Lumley, 2023). Residuals for logistic regressions were plotted with binnedplot from the arm package (Gelman et al., 2022; *R: Binned Residual Plot*, n.d.), and model fit was reported with R^2 and adjusted R^2 values, which were computed with the PoliSciData package (Pollock, 2020).

Predictor Variables

Propensity score analysis involves selection of two sets of variables, with one set expected to affect or show imbalance between treatment assignments, and another set expected to affect outcomes. In a simplistic scenario, variables theorized to affect treatment assignment could be used to predict propensity scores, and variables theorized to affect outcomes could be used to estimate treatment effects, with potential overlap between the two sets. In practice, however, simulation studies and analysis using observational data have shown that including all outcome covariates in the estimation of propensity scores and including all treatment assignment predictors in the estimation of outcomes reduces bias (Brookhart et al., 2006; Drake, 1993; Griffin et al., 2020). When all variables are included in both models, even if one of the two models is misspecified, bias is not increased (Drake, 1993) and treatment estimates are unaffected (Bang & Robins, 2005; J. D. Y. Kang & Schafer, 2007). This is referred to as a doubly robust model, because failure in one of the two models still allows for effective estimates, as long as the misspecified model is not “grossly misspecified” (J. D. Y. Kang & Schafer, 2007, p. 523). While the intention was to use doubly robust models when possible, one outcome predictor (year of graduation) was used in the outcome model but excluded from the propensity score model. This is discussed in the results section.

The number of variables used in propensity score estimation is not limited by sample size (Guo & Fraser, 2015; McCaffrey et al., 2004). The estimation of treatment effects involves regression. To avoid biased estimates in regression models, at least ten observations are needed for each covariate in the regression model (Peduzzi et al., 1996), which can limit the number of variables that can be included in estimates of treatment effects. When sample size limits the number of treatment assignment covariates that can be included in the outcome model, which eliminates the option of using a doubly robust model, Griffin et al. recommend including treatment assignment covariates that remained imbalanced after weighting (Griffin et al., 2020). Criteria used to assess data balance are discussed in the procedure section.

The counterfactual framework is only reliable when guided by substantive knowledge and appropriate theories (Guo & Fraser, 2015). While all selection and outcome covariates are to be included in both models, each included variable must be justified by theories or prior research connecting the variable to treatment selection or outcomes, or by observed imbalances by the variable in treatment selection or treatment levels (Rodenbusch et al., 2016). While importance is placed on understanding mechanisms underlying treatment assignment, if there is a correlation between a variable and treatment assignment but no theoretical explanation, the variable should still be included in the model (Sobel, 2005). The outcome in the study was student employment status prior to graduation, and the treatment was VIP Program participation, with semesters of participation representing levels of treatment. Each variable selected for inclusion in the analysis showed an imbalance in participation in VIP; had the potential of being correlated with participation in VIP, based on the literature; or was expected to contribute to or hinder job placement, based on the literature. Variables can be loosely grouped into four categories: background, academics, student engagement, and career preparation (Table 4).

Table 4

Predictor and Outcome Variables

	Variable	Relationship Expected in VIP Participation*	Relationship Expected in Job Placement**	Possible Values
Background				
1	Citizenship		x	Binary
2	First Generation Student Status	x	x	Binary
3	Gender		x	Binary
4	Pell Grant Status		x	Binary
5	Race/Ethnicity***	x	x	8 categories
Academics				
6	Dual Bachelor's Master's degree		x	Binary
7	GPA	x	x	Continuous
8	Graduation Year	x	x	6 categories
9	GT1000 Freshman Experience Course	x	x	Binary
10	GT2000 Transfer Experience Course	x	x	Binary
11	Major***	x	x	38 categories
12	Transfer Student		x	Binary
13	Undergraduate Research Course (non-VIP)	x		Binary
14	Vertically Integrated Projects	treatment	x	Binary
Student Engagement				
15	Study Abroad		x	Binary
16	Greek Fraternity/Sorority	x		Binary
	Living Learning Communities			
17	- Global Leadership	x		Binary
18	- Grand Challenges	x		Binary
19	- Honors Program	x		Binary
20	- I House	x		Binary
21	- Ignite Summer Launch	x		Binary
22	- Women in Science and Technology	x		Binary
23	NCAA athlete	x		Binary
Career Preparation				
24	CoOperative Education Program	x	x	3 categories
25	Full-time internship		x	Binary
26	Job placement status when surveyed prior to graduation		outcome	Binary
27	Part-time Internship		x	Binary
28	Use of Career Center Services		x	Integer

* Variables from both columns were to be included in propensity score estimates.
** As many variables as possible from both columns were to be included in outcome effect estimates.
*** Categories were combined or adjusted.

Background Variables Early career outcomes for college students vary by socioeconomic status (SES) (Wolniak & Engberg, 2019), but there are many ways in which SES can be measured. The National Center for Education Statistics convened an expert panel to recommend appropriate measures for SES for use in the National Assessment of Educational Progress. The panel recommended inclusion of five SES indicators: family income, parent education, parent occupational status, neighborhood SES, and school SES (Cowan et al., 2012). Indicators for one of the five variables was included in the study. **Pell Grant** recipient status was used as a binary indicator for family income. Parent education was going to be represented by student status as **first-generation college students**, a term that refers to students without a parent/guardian who completed a four-year degree (exclusion of the variable is discussed in the results section). Status as a first-generation college student was expected to correlate with participation in VIP and with career outcomes. Nationally, first-generation students participate in undergraduate research at lower rates than their peers, with 17% of first generation students participating by the time they graduate, compared to 27% for non-first generation students (National Survey of Student Engagement, 2020). A survey by the National Association of Colleges and Employers found lower job placement rates for first-generation college students (Koc, 2014), who perceive greater barriers to finding employment than non-first-generation students (Ma & Shea, 2021). Neighborhood SES was not included in the study, because it is difficult to quantify. While an indicator called the Social Deprivation Index classifies social disadvantage by zip code (Robert Graham Center - Policy Studies in Family Medicine & Primary Care, 2019), it was not used as an indicator, because zip codes can include both affluent and impoverished neighborhoods. An indicator for parent occupational status was not available. School SES was also excluded, because at the college level, all students in the study were attending the same institution, and at the high school level, SES cannot be determined for students from private schools.

Based on the prior study and resume audit studies that show inequity in hiring practices, career outcomes were expected to vary by **race and ethnicity**. Georgia Tech combines race and ethnicity into a single race/ethnicity indicator with nine possible values: Hispanic or Latino, American Indian or Alaska Native, Asian, black or African American, Native Hawaiian or Other Pacific Islander, white, two or more races, and unknown. Hispanic/Latino students of all races are categorized as Hispanic/Latino. Race/ethnicity is also correlated with VIP Participation, with higher participation among Asian students, lower participation among white students, and representative participation among black/African American and Hispanic/Latino students over the six-year period included in the study.

The sample included citizens and permanent residents. Career outcomes were expected to differ by **citizenship** status because noncitizen permanent residents face discrimination. While the US Immigration and Nationality act prohibits discrimination in recruitment, hiring, and firing based on citizenship or national origin (United States Code, 1964; US Citizenship and Immigration Services, 2019), in 2023, the Justice Department found Georgia Tech had violated the act by allowing employers to post advertisements for job openings that discouraged non-citizens from applying, and by allowing employers to screen potential applicants by citizenship status (US Department of Justice, 2023). This confirms discrimination experienced by permanent residents, which may lead to lower job placement rates among non-citizen permanent residents. Studies have also shown job placement to vary by **gender**, with lower placement rates for women in general, and higher rates in engineering, so gender was included as well (Koc, 2014; Sagen et al., 2000). Neither gender nor citizenship were expected to correlate with VIP participation.

Academic Variables Seven academic variables were expected to covary with VIP participation and/or job placement, with six covarying with VIP participation and five covarying with job placement. In preparation for the study, to examine balance in VIP participation by possible co-variates, participation rates were obtained for all students who graduated in 2017 through 2022.

Participation in Georgia Tech's VIP Program has increased nearly linearly since the program's establishment in Spring 2009. Because of growth over time, participation varies by student **semester of graduation**, with 10% of graduates in spring of 2017 having participated, compared to 26% of fall 2022 graduates. Student job placement rates were expected to vary by semester of graduation, particularly for students who graduated after COVID began in 2020. The method used to estimate propensity scores can handle a large number of variables, but the number of variables used to estimate effects is limited to 10 observations per variable. To limit the number of variables in the estimation of effects, graduation year (six levels) was used instead of semester (eight-teen levels).

Grade point average (GPA) was expected to covary with both VIP participation and job placement. Studies have found graduates' success in finding employment varies by GPA (Sagen et al., 2000; Wolniak & Engberg, 2019). Nationally, students who participate in undergraduate research tend to have higher GPAs (Russell et al., 2007), so VIP participation was also expected to vary by GPA.

VIP participation and job placement were expected to vary by **major**. Graduates' success in finding employment varies by major, with higher placement in STEM fields and lower placement in the humanities (Sagen et al., 2000; Wolniak & Engberg, 2019). Enrollment in VIP differs by major, because some degree programs incentivize multiple semesters of participation

(Georgia Institute of Technology, 2022b; Sonnenberg-Klein et al., 2018b). The sample did not include double majors.

Studies on college career outcomes typically group majors into large groups, such as arts & humanities, business, social sciences, etc., with groupings varying by study (Sagen et al., 2000; Wolniak et al., 2008; Wolniak & Engberg, 2019). Sagen et al. grouped natural science, life science, and engineering together in a Specialized Hard group (Sagen et al., 2000). Wolniak et al. grouped math, CS and engineering together, with a separate group for sciences (Wolniak et al., 2008), while Wolniak and Engberg grouped all four areas together as STEM (Wolniak & Engberg, 2019). While colleges within Georgia Tech would provide a useful grouping mechanism, the approach has multiple shortcomings: one major is jointly administered by two colleges, with students in the major equally divided between the two; some colleges are very small; and some majors are closely related to majors in other colleges. Majors were going to be grouped together if the number of variables in the logistic regression needed to be limited.

Status as a **transfer student** is correlated with slightly lower VIP participation rates, and it may correlate with job placement. In the National Survey of Student Engagement, 30% of seniors who began at their college as freshmen had done research with faculty, compared to 15% of seniors who had transferred to their college (National Survey of Student Engagement, 2020). While participation rates for transfer students in Georgia Tech's VIP program are lower than for non-transfer students, the difference is smaller with 14% for transfer students and 20% for non-transfers by the time they graduate. Research on the impact of transfer student status on career outcomes is mixed. Xu et al. found that eight years after graduation, transfer students earned less than non-transfer students (Xu et al., 2018), while Smart and Ethington found no differences by transfer student status on job status, job stability, or job satisfaction (Smart & Ethington, 1985).

VIP participation was 4% lower for students who participated in the **GT1000 freshman experience course or the GT2000 transfer student experience course** (19% between the two), compared to students who did not take a freshman/transfer experience course (23%). VIP participation also differed by whether students **enrolled in university-designated undergraduate research**. Students who did not enroll in an undergraduate research course participated in VIP at a slightly lower rate of 18% compared to 20%. Because VIP is a type of undergraduate research, differences between the effect of the two on job placement was also of interest.

Student Engagement Nine variables involve student engagement in an array of experiences beyond typical coursework. Six of the nine student engagement variables involved living learning communities in which participants live in dorms with fellow community members, enroll in community-linked courses, and participate in activities outside of class (tours, retreats, social events, etc.). Participation in VIP varies by living learning community. This may be because communities attract different majors, attract students who would (or would not) have been interested in VIP regardless of community participation, or because communities cultivate more (or less) student interest in VIP. For example, the **Global Leadership** living learning community involves team-based projects that focus on real-world problems, which are key elements of VIP. Over half of Global Leadership students participate in VIP, compared to 21% of students who do not participate in the Global Leadership community. This may be because the Global Leadership community attracts students interested in team-based projects, or because the community cultivates student interest in this type of course. In the opposite direction, students who participate in the **Women in Science and Technology** living learning community participate in VIP at lower rates than peer women, at 14% compared to 20%, but participate in work-based learning (CoOp and internships) at higher rates, at 55% compared to 39% for women peers. Again, this may be because the community attracts similar students, or because the community cultivates greater interest in work-based learning.

The **Honors Program** living learning community allows VIP credits to fulfill Honors Program requirements. Thirty-one percent of honors program students participated in VIP, compared to 20% for non-Honors Program students. Other living learning communities are **I House** with 9% higher participation in VIP; **Ignite Summer Launch** with 10% higher participation; **Grand Challenges** with 5% higher participation; and **Explore** (careers in science, math and health) with 5% higher participation.

Wolniak & Engberg found a correlation between **study abroad** and early career earnings (Wolniak & Engberg, 2019), so job placement may vary by participation in study abroad. There is no difference in participation in VIP by participation in study abroad.

VIP participation varies by participation in the Greek Fraternity/Sorority system and status as an NCAA athlete. Students in the **Greek system** participate in VIP at an 8% lower rate than students who are not in the Greek system. Like the Women in Science and Technology living learning community, students in the Greek system participate in work-based learning at higher rates, with 16% higher participation. **NCAA athletes** participate in VIP at an 18% lower rate than their peers, with a difference by gender: 2% of NCAA athlete men participate in VIP, compared to 5% of NCAA athlete women. VIP participation for athletes who are not NCAA athletes does not differ from participation rates for non-athletes.

Career Preparation Student job placement is expected to vary by career preparation experiences. Work-based learning improves graduate success in finding employment (Sagen et al., 2000). At Georgia Tech, the **CoOp** program alternates three semesters of work with three semesters of school. Employers commit to employing their CoOp students for three semesters, and students commit to working for the same employer all three semesters. Students enroll in CoOp courses during work semesters, and students who complete all three semesters of CoOp receive a designation on their degrees. CoOp student participation in VIP also varies. Students who receive CoOp degree designations participate in VIP at a lower rate, at 10% compared to students with no CoOp experience.

When students do **internships**, they are encouraged but not required to enroll in a Career Center course with an INTN subject code, with different course numbers for full time and part time internships. The office of institutional research provided data for two binary internship variables, with one for full-time internships, and the other for part-time internships.

The Career Center offers services to support students' job searches, and job placement is expected to covary with **use of Career Center services**. The Georgia Tech Salary and Career survey asks students if they made use of specific services: Career Center career advising; career fairs; workshops/webinars/events; company information sessions; job board/interview scheduling platform; interviews with employers through the center; and other unlisted services. Responses to the seven questions were combined into a single score, ranging from zero to seven.

Outcome Variables

The analysis involved outcome variables at two levels. In the second level of the analysis, which involved logistic regression, student job placement prior to graduation was the outcome variable. In the first level of the analysis, which involved propensity score estimation, participation in VIP was the outcome variable. Students can enroll in the course for multiple semesters, so number of semesters of participation (zero to three) represented different levels of dosage. The analysis method accounts for the dependence between the categories – that a student who participates for two semesters must have first participated for one semester.

In the outcome model, number of semesters of VIP participation was a predictor variable, and a dosage effect was assumed. This means that additional benefits are accrued with additional semesters of participation. The assumption was informed by the cumulative benefits of participation in multiple high-impact experiences (Kuh, 2008), and by a longitudinal analysis of VIP peer evaluations that found leadership growth through the third semester of participation (Sonnenberg-Klein, 2023). With this assumption, the variable was handled as an integer in the logistic regression outcome model. To test the assumption, the variable was treated as a factor in a separate analysis. This was done to determine if all treatment levels were associated with similar or with progressively higher odds of job placement. Progressively higher odds of job placement would imply a dosage effect. Similar odds across all treatment levels would imply no dosage effect.

Assumptions

A key assumption of the Neyman-Rubin counterfactual model that extends to propensity score analysis is the stable unit treatment value assumption (SUTVA). In propensity score analysis, the SUTVA assumption is violated if multiple versions of treatment are handled as equivalents in the analysis; if interaction between participants affects outcomes; or if treatments alter social conditions that then affect outcomes (Guo & Fraser, 2015; Winship & Morgan, 1999). A potential violation of the SUTVA assumption is the variation in number of credit hours students were able to register for in each semester of participation. In VIP, students can register for one or two credit hours, with each credit hours associated with two to three hours of work outside of class. This was not deemed a significant violation for three reasons. First, while one and two-credit hour students would do different amounts of work outside of class, both groups received the same number of contact hours. Peer interaction and faculty mentorship are key aspects of the experience, and because students receive the same amount of interaction/mentorship regardless of number of credit hours, the students would have very similar experiences. Second, while there are guidelines on the number of hours students should work outside of class, student time on task is not carefully monitored. Faculty often forget how many hours students registered for and report that one-credit hour students frequently do more work than is expected. Third, two semesters of one credit are not equivalent to one semester of two credits. This is because student expertise on the team is closely connected with number of semesters on the team, making number of semesters a more effective measure than number of credits.

The SUTVA assumption would also be violated if interaction between participants affected outcomes. While peer mentorship is a key aspect of LT-PBL-EFR, and positive or negative experiences with peers could feasibly affect outcomes, this was not deemed a violation. This is because propensity score analysis has been used to study the effectiveness of addiction treatment programs that involve group therapy, and a study including group therapy is used to teach propensity score analysis methods (Diamond et al., 2002; Griffin et al., 2020), implying that peer interaction is not a strong violation of SUTVA.

Finally, the SUTVA assumption would be violated if treatments altered social conditions that then affect outcomes (Guo & Fraser, 2015; Winship & Morgan, 1999). If the present study were conducted in a closed system in which student outcomes reshaped the job market, then SUTVA would be violated. However, because Georgia Tech is only one institution in a large job market ecosystem, outcomes for a subset of Georgia Tech students would have little effect on the labor market.

Another assumption in propensity score analysis is positivity, in which every individual has a non-zero chance of receiving any level of treatment (McCaffrey et al., 2013). The greatest threat to this assumption was difference in participation in VIP by major. If no students from a given major participated in three semesters of VIP, then students in the major would likely have a zero chance of participating, violating the assumption. Participation in VIP varies by major. To ensure the assumption was not violated, majors were only included when there was at least one white and one non-white case in each treatment level. To further ensure the assumption was not violated, propensity scores were examined by treatment level for each analysis grouping and sub-group. Groupings that did not meet the assumption were excluded from further analysis.

Propensity score analysis also assumes that there are no unmeasured confounders. Because sensitivity analysis methods have not been developed for inverse propensity score weighting for multiple dosage levels (Rudolph & Stuart, 2018; VanderWeele & Arah, 2011), this assumption is not typically tested. However, conducting analyses across multiple subgroups yielded notable differences, implying the presence of unmeasured confounders. These are discussed in the results and discussion sections.

Procedure

Groupings and Subgroups Participation in VIP varies by major, and in initial analyses, inverse propensity score weighting yielded better balance for majors with high proportions of cases in the 2-3 semester dosage levels. For this reason, data was analyzed by groupings, with the first group consisting of majors in which 10% or more of students in the sample participated in 2-3 semesters of VIP. The thresholds for the remaining groups were 8% or more, 6% or more, and then all majors that met the inclusion criteria (Table 5).

Because non-white non-Asian subgroups were comparatively small, they were combined into two groups. Black/African American and Hispanic/Latino students were combined into a single underserved minority group (URM). While American Indian/Alaskan Natives, Native Hawaiians and other Pacific Islanders are also marginalized in education, the literature review for the study consisted of research on discrimination against black/African Americans, Hispanics/Latinos, and Asians seeking to enter the workforce. For this reason, the remaining groups,

Table 5

Groupings

	Number of Semesters				Percent in	Grouping			
Major	0	1	2	3	2-3 Sem	3 Maj	5 Maj	8 Maj	11 Maj
Computer Science	985	97	18	244	19%	X	X	X	X
Environmental Engineering	68	11	9	5	15%	X	X	X	X
Computer Engineering	216	47	17	13	10%	X	X	X	X
Electrical Engineering	257	44	20	11	9%		X	X	X
Aerospace Engineering	320	34	24	6	8%		X	X	X
Mechanical Engineering	835	114	52	24	7%			X	X
Biomedical Engineering	453	40	13	23	7%			X	X
Industrial Engineering	813	57	54	6	6%			X	X
Materials Science & Engr	147	6	2	5	4%				X
Civil Engineering	245	17	9	2	4%				X
Chemical and Biomolecular Eng	406	33	8	7	3%				X
N for Grouping						1730	2446	4930	5817

Table 6

Combined Racial/Ethnic Groupings

	Semesters of VIP					
	0	1	2	3	Total	% of Total
Asian	972	173	76	149	1370	24%
Other or Unknown	295	29	18	26	368	6%
URM	693	79	35	48	855	15%
White	2785	219	97	123	3224	55%

American Indian/Alaskan Natives, Native Hawaiians and other Pacific Islanders, unknown race/ethnicity, and two or more races, were combined into a single “other or unknown” group. This yielded grouping proportions of 55% 24%, 15% and 6% (Table 6).

The prior study implied greater effects of VIP participation on job placement among historically underrepresented students and Asian students. Research implies connections between both gender and SES in job market outcomes. While race/ethnicity, gender, and SES-related variables were included in the analysis, a single analysis would not capture differences in the experiences of students from different backgrounds. For example, undergraduate research and work-based learning may affect job placement for non-white students differently than for the full sample. Doing separate analyses on subgroups centers their experiences in the study. To this end, five analyses were conducted for each analysis grouping: 1) full group, 2) non-white students, 3) white students, 4) female students, and 5) Pell grant recipients.

Procedure The first step in the procedure was to select pre-treatment and outcome covariates, all of which would be included in both propensity score estimation and the outcome model. For each grouping and subgroup, frequencies for variables were examined. Variables were included in the analysis when they represented 5% or more of the sample. One exception was the CoOp variable, which had three levels: No CoOp, Some CoOp with no degree designator, and CoOp Degree Designator. No CoOp was used as the reference category, yielding two binary variables. Because CoOp was expected to be strongly correlated with job placement, even if less than 5% of students had earned the CoOp degree designator, all three levels were still included.

Next, generalized propensity scores were estimated, the scores were examined to ensure they met assumptions, and inverse propensity score weights were generated and examined. An assumption of propensity score analysis is that no case has a 0% or 100% chance of being assigned to a treatment group. If this occurs, the method may not be appropriate for the given grouping, or cases may need to be removed from the sample. Generalized propensity scores were examined through boxplots and summary statistics by dosage level. If scores showed very low or very high likelihoods of receiving an observed treatment (scores near zero or one), data was examined to determine if the grouping was appropriate. If scores met the assumption, weights were examined to check for outlier weights. If groupings were adjusted or cases were excluded, the procedure was repeated from the beginning.

If scores and weights were acceptable, data balance was assessed to determine the degree to which inverse propensity score weighting removed bias between the groups. Balanced data would have a weak correlation between each variable and dosage level. To assess balance for multiple dosages, Leite recommends doing a standardized regression for each covariate on the weighted and unweighted sample, with the single covariate as the independent variable and dosage as the dependent variable (Leite, 2016). A standardized regression coefficient smaller than 0.1 represents sufficient balance in the weighted sample (Leite, 2016). If weighting yielded poor balance, the data was examined to determine if groupings were appropriate or if cases should be excluded. If groupings were adjusted and/or cases were excluded, the procedure was repeated from the beginning until acceptable balance was achieved. Data cannot always be balanced well, but this can be handled by including unbalanced variables in the outcome model (Griffin et al., 2020). This would already be done in a doubly robust model, but if sample size limited the number of variables that could be included in the outcome model, treatment assignment covariates that remained imbalanced were to be included (Griffin et al., 2020).

If acceptable balance was achieved for a grouping/subgroup, effects were estimated with logistic regression, and adjusted log odds ratios for statistically significant variables were examined. Model fit was assessed with binned residual plots, a function in the arm package (R: Binned Residual Plot, n.d.), and fit was reported with R^2 and adjusted R^2 statistics. A binned residual plot groups cases into categories based on their fitted values, and then plots the average residual against the average fitted value for the bin. Gray lines outline boundaries of plus and minus two standard errors. In a well-fitted model, 95% of binned residual markers fall between the gray lines, although Webb describes having a majority of makers within between the gray lines as reasonable fit (Webb, n.d., sec. 8.7). After confirming model fit, odds ratios for statistically significant variables were assessed against expectations, such as higher job placement rates for students with work-based learning experiences, and lower job placement rates for Pell grant recipients or transfer students.

As mentioned, to assess the assumption of a dosage effect, semesters of participation in VIP were handled as factors in a separate analysis, following the same procedure.

Expectations

To balance data, I expected that I would need to **combine groups, exclude variables, and exclude cases** from the analysis. A large number of majors were identified for inclusion, but if not enough students from a major/group of majors participated in VIP, meaningful comparisons would not be possible. I planned to exclude variables with low frequencies (less than 5%) in order to achieve data balance. While sample size can limit the number of variables that can be included in an outcome model, exclusion of variables in previous steps were expected to make sample-size limitation a non-issue.

The overall effect of the VIP Program on job placement was expected to be lower than in the prior study, or not statistically significant. This is because samples matched on a limited number of variables, as in the prior study, yield biased results. Propensity score analysis reduces bias, yielding lower and more accurate estimates of effects. Based on the preliminary analysis and literature review, career outcomes were expected to be higher for white students, higher for students with work-based learning experiences, lower for Pell grant recipients and transfer students, and lower for students with low grade point averages. If VIP had an effect, differential gains were expected for marginalized students who had participated. Kuh found that high-impact experiences benefit all students, but historically underserved students see compensatory gains (Kuh, 2008). For example, Hispanic students who participate in high-impact experiences see greater gains in GPA, with increasing numbers of experiences increasing their GPAs above those of white students who participate in the same number of experiences. Similarly, African American student persistence increases with the number of high-impact experiences, exceeding the persistence of white students who participate in the same number of high-impact experiences (Kuh, 2008). VIP impact on job placement may be compensatory as well.

4 RESULTS

The analysis plan involved twenty analyses, with four groupings (three, five, eight, and eleven majors) and five analyses for each group (full group, non-white, white, female, and Pell grant recipients). Results are summarized in the following order: variables included in the models; evaluation of propensity scores; examination of propensity score weights; data balance; outcome model fit; and estimates of effects.

Variables

Twenty-seven variables were considered for use in the analysis. Of these, ten were excluded (Table 7). Seven were excluded because frequencies were small, with less than 5% in each major grouping and subgroup analysis groups. This applied to part-time internships, status as an NCAA athlete, and all of the living learning communities except the Honors Program (see Appendix A for frequency tables). First-generation student status was excluded because information was missing for 64% of cases. Although this represented loss of an SES indicator, two SES-related variables were retained (status as a Pell grant recipient, and transfer student status).

Table 7

Excluded Variables

	Variable	Frequencies of less than 5%	Missing Data	Other
Background				
1	First Generation Student Status		X	
Academic				
2	Dual Bachelor's Master's degree			No cases, categorized as graduate students
	Living Learning Communities			
3	- Global Leadership	X		
4	- Grand Challenges	X		
5	- I House	X		
6	- Ignite Summer Launch	X		
7	- Women in Science and Tech.	X		
8	NCAA athlete	X		
Career Preparation				
9	Part-time Internship	X		
10	Use of Career Center Services		X	Unexpected correlations

The variable representing student use of the career center was excluded for multiple reasons. First, the data was only available for three of the six years in the study. A separate analysis was considered for cases that included the data, but examination showed unexpectedly low reported rates of career center use, and negative correlation with use of the center and job placement prior to graduation. The low reported rates of career center use may have been the result of question wording and skip-logic in the survey. Students were first asked if they had used the career center. If they indicated they had used the center, they were presented with a list of services provided by the center and asked which they had used. A problem is that students may not have known which services the career center provides. For example, the job board is a key service offered by the center, but in years in which the questions were asked, only 30% reported using the job board. Students may have considered the job board a service of the institute instead of the center, leading them to report not having used the center before seeing the list of services offered. The negative correlation between center use and job placement may reflect higher use of the center among students struggling to find jobs as they neared graduation, or higher use among students with fewer connections and resources. For these reasons, the variable was excluded from the analysis.

Grade point average was a continuous variable. While centering of GPAs did not effect coefficient estimates or effect estimates, GPAs were centered on the mean across all retained cases, which was 3.6. When adjusted odds ratios were calculated, GPA differences were limited to 0.4, representing the difference between the centered mean and the highest possible GPA.

Seventeen variables were used in the analysis, with four background, eight academic, three student engagement, and two career preparation variables (Table 8). When nominal variables were coded, this yielded twenty-three to thirty-seven coded variables that could be used in the outcome models, with the number of coded variables varying by major-grouping and sub-group (Table 9). In each of the final analyses, after variables with small frequencies were excluded, all sample sizes met the ten-cases per variable minimum.

Table 8

Included Variables

	Variable	Type	Number of Variables after Coding	Exclusions and Notes
Background				
1	Citizenship	Nominal	1	
2	Female	Binary	1	
3	Pell	Binary	1	Excluded for Pell subgroup
4	Race/Ethnicity	Nominal	3, 2	2 levels for non-white subgroups; Excluded for white subgroups
Academics				
5	Grade Point Average	Continuous	1	
6	Graduation Year	Nominal	5	Excluded from Propensity Score Model; Included in Output Model
7	GT1000 Freshman Exp. Course	Binary	1	
8	GT2000 Transfer Exp. Course		1	
9	Major	Nominal	2, 4, 7, 10	Varied by major grouping
10	Transfer Student	Binary	1	
11	Undergraduate Research	Binary	1	
12	VIP, Number of Semesters	Integer	1	
Student Engagement				
13	Study Abroad	Binary	1	
14	Greek Fraternity/Sorority	Binary	1	
15	Honors Program	Binary	1	
Career Preparation				
16	CoOperative Education	Nominal	2	
17	Full Time Internship	Binary	1	

Table 9

Maximum Possible Number of Variables in Outcome Models after Coding

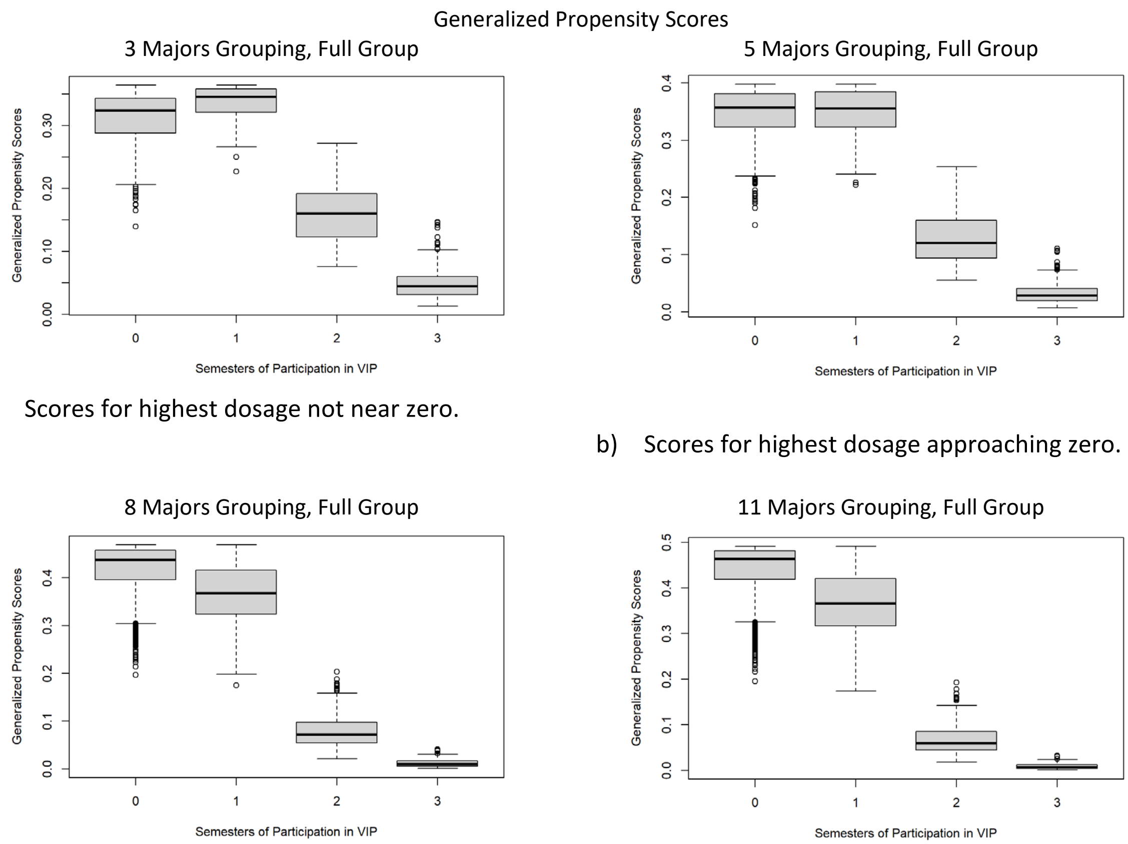
	Full Group		Non-White		White		Female		Pell	
Major Grouping	Variables	N	Variables	N	Variables	N	Variables	N	Variables	N
3 Majors	25	1730	24	951	22	779	25	474	24	434
5 Majors	27	2446	26	1247	24	1199	27	631	26	647
8 Majors	30	4930	29	2268	27	2662	30	1714	29	1162
11 majors	33	5817	32	2593	30	3224	33	2127	32	1381

Note: Effective sample size in each analysis smaller than initial sample size listed above.

Inclusion of graduation year had an unanticipated effect on propensity score estimates and weights. Propensity score analysis is intended to reduce bias in samples by controlling for self-selection and administrative selection of participants. If an administrator chose participants who had work-based learning experience more often than students without the experience, propensity score analysis would control for the selection bias. In this study, student participation increased over time. Students in 2022 were more likely to participate in the program than students in 2017, which yielded lower propensity scores (and higher weights) for 2017 participants. To examine the impact of year on the propensity score model, weights for two major groupings were considered. When year was included in the propensity score model, mean weights for VIP participants who graduated in 2022 were 0.69 in the three-major grouping and 0.65 in the eight-major grouping. In contrast, weights for 2017 graduates in the same groupings were 1.52 and 1.55, more than twice as large. Because the 2017 and 2022 students did not differ in meaningful ways, the difference in weighting was deemed inappropriate. For this reason, year was excluded from the propensity score model, but was still included in the outcome model.

Evaluation of Propensity Scores

An assumption of propensity score analysis is that no individual has a 100% or 0% chance of treatment assignment. This assumption was examined through boxplots and summary statistics of generalized propensity scores by treatment level. Analysis groupings were based on the proportion of students in each major that participated in multiple semesters of VIP (higher dosages), with majors added to the analysis in groups. As majors with lower proportions of high dosage were added, generalized propensity scores for students at higher dosage levels (i.e. the likelihood that a student was in the high dosage group that they were observed in) were expected to decrease. As shown in Figures 1 and 2, as majors with lower high-dosage proportions were added, generalized propensity scores for the second highest dosage decreased, and scores for the highest dosage collapse to nearly zero.



a) Scores for highest dosage not near zero.

b) Scores for highest dosage approaching zero.

c) Scores for highest dossage collapse near zero.

Figure 1. Boxplots of Generalized Propensity Scores by Dosage Level

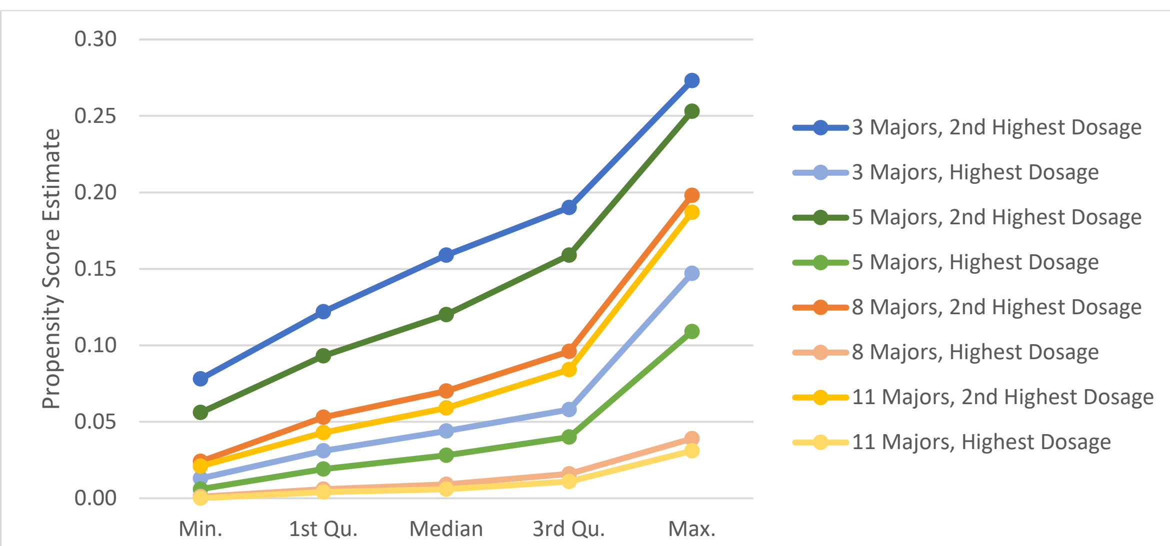


Figure 2. Propensity Scores for Each Major-Grouping, 2 Highest Dosage Levels,

While propensity scores estimate the chance of receiving an observed treatment, by definition, none of the students in the highest dosage levels had zero chance of receiving the highest dosage because it was the observed treatment. However, having score estimates collapse to nearly zero is problematic. It implies that not enough students in the group engaged in enough high-dosage treatment for meaningful analysis. For this reason, further analysis was not done for major groupings in which score estimates for the highest dosage collapsed to nearly zero. Scores were considered to have collapsed to nearly zero when the minimum was less than 0.005, the first quartile was less than 0.015, and the median was less than 0.02 (Table 10). Under these criteria, the eight and eleven major groupings were excluded from further analysis.

Table 10

Summary Statistics for Generalized Propensity Scores, Full Groupings

		Semesters of Participation in VIP				
Grouping	Statistic	0 Sem	1 Sem	2 Sem	3 Sem	Comments
3 Majors	Min.	0.140	0.227	0.076	0.013	
	1st Qu.	0.288	0.321	0.123	0.031	
	Median	0.324	0.346	0.16	0.045	
	Mean	0.314	0.337	0.161	0.048	
	3rd Qu.	0.343	0.358	0.191	0.060	
	Max.	0.364	0.364	0.272	0.147	
5 majors	Min.	0.151	0.222	0.055	0.007	
	1st Qu.	0.323	0.323	0.094	0.019	
	Median	0.357	0.355	0.12	0.029	
	Mean	0.347	0.349	0.13	0.033	
	3rd Qu.	0.381	0.385	0.16	0.041	
	Max.	0.398	0.398	0.254	0.110	
8 Majors	Min.	0.197	0.175	0.021	0.001*	Group excluded
	1st Qu.	0.396	0.324	0.054	0.006*	from further analysis
	Median	0.438	0.367	0.071	0.009*	
	Mean	0.420	0.368	0.079	0.012	
	3rd Qu.	0.458	0.416	0.098	0.016	
	Max.	0.469	0.469	0.204	0.041	
11 Majors	Min.	0.196	0.174	0.018	0.000*	Group excluded
	1st Qu.	0.419	0.317	0.044	0.004*	from further analysis
	Median	0.463	0.366	0.059	0.006*	
	Mean	0.443	0.368	0.068	0.008	
	3rd Qu.	0.482	0.421	0.084	0.012	
	Max.	0.491	0.491	0.193	0.032	

* Propensity scores for the dosage level collapsed to nearly zero.

Propensity scores for subgroups in the two remaining groupings were examined. Both the three majors group and five majors group had at least one female and one Pell grant recipient at each dosage level. As with the eight and eleven major groupings, scores for both white sub-groups and the Pell subgroup in the five majors grouping collapsed to nearly zero, reflecting low participation rates at high dosage levels (Figure 3). Because the three subgroups did not meet the criteria set in the previous step (Table 11), separate analysis were not done for white subgroups or Pell grant recipients in the five majors grouping, but white students and Pell grant recipients were still retained in the sample.

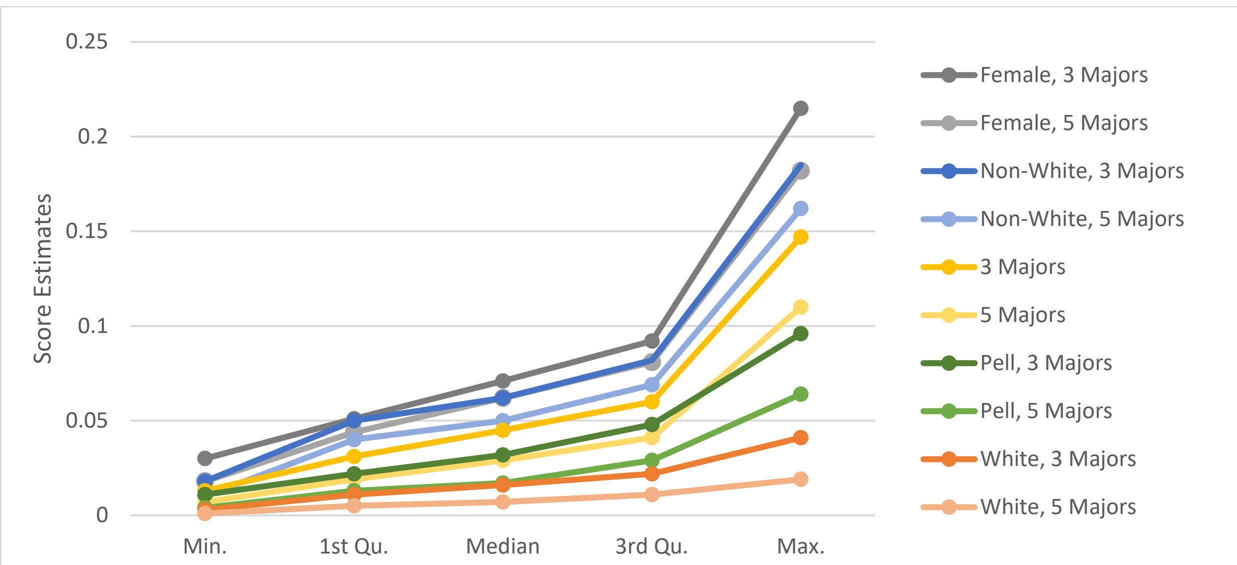


Figure 3. Propensity Scores for Highest Dosage by Subgroup, 3 and 5 Majors

Weights

Weights for the remaining analysis groups were examined (Table 12). In all seven groups, the median weight for each of the three lowest dosages was near one (Figure 4). Weights for the highest dosage level had the widest distributions and lower medians across all seven analysis groupings. The maximum weight across all groups was 3.525.

Table 11

Summary Statistics for Generalized Propensity Scores, 3 Majors and 5 Majors

			Semesters of Participation in VIP				
Grouping			**0**	**1**	**2**	**3**	Comments
3 Majors	Full Group	Min.	0.140	0.227	0.076	0.013	
		1st Qu.	0.288	0.321	0.123	0.031	
		Median	0.324	0.346	0.160	0.045	
		3rd Qu.	0.343	0.358	0.191	0.060	
		Max.	0.364	0.364	0.272	0.147	
	Non-White	Min.	0.128	0.236	0.079	0.018	
		1st Qu.	0.258	0.316	0.138	0.050	
		Median	0.287	0.327	0.180	0.062	
		3rd Qu.	0.308	0.335	0.195	0.082	
		Max.	0.339	0.339	0.291	0.185	
	White	Min.	0.265	0.232	0.043	0.003*	Further analysis not done for subgroup.
		1st Qu.	0.358	0.316	0.096	0.011*	
		Median	0.377	0.351	0.133	0.016*	
		3rd Qu.	0.398	0.380	0.166	0.022	
		Max.	0.415	0.409	0.194	0.041	
	Female	Min.	0.101	0.265	0.061	0.030	
		1st Qu.	0.254	0.307	0.129	0.051	
		Median	0.283	0.326	0.156	0.071	
		3rd Qu.	0.305	0.331	0.195	0.092	
		Max.	0.332	0.332	0.310	0.215	
	Pell	Min.	0.171	0.308	0.101	0.010	
		1st Qu.	0.318	0.336	0.124	0.023	
		Median	0.347	0.345	0.138	0.032	
		3rd Qu.	0.367	0.358	0.167	0.052	
		Max.	0.384	0.382	0.242	0.088	
5 Majors	Full Group	Min.	0.151	0.222	0.055	0.007	
		1st Qu.	0.323	0.323	0.094	0.019	
		Median	0.357	0.355	0.120	0.029	
		3rd Qu.	0.381	0.385	0.160	0.041	
		Max.	0.398	0.398	0.254	0.110	
	Non-White	Min.	0.134	0.226	0.065	0.010	
		1st Qu.	0.278	0.323	0.119	0.040	
		Median	0.311	0.345	0.144	0.050	
		3rd Qu.	0.338	0.356	0.184	0.069	
		Max.	0.363	0.363	0.294	0.162	
	White	Min.	0.269	0.220	0.040	0.001*	Further analysis not done for subgroup.
		1st Qu.	0.396	0.320	0.064	0.005*	
		Median	0.424	0.358	0.085	0.007*	
		3rd Qu.	0.444	0.393	0.102	0.011	
		Max.	0.459	0.447	0.164	0.019	
	Female	Min.	0.114	0.278	0.075	0.018	
		1st Qu.	0.272	0.317	0.121	0.044	
		Median	0.306	0.336	0.142	0.062	
		3rd Qu.	0.330	0.350	0.170	0.081	
		Max.	0.354	0.354	0.310	0.182	
	Pell	Min.	0.193	0.261	0.064	0.004*	Further analysis not done for subgroup.
		1st Qu.	0.352	0.333	0.089	0.013*	
		Median	0.385	0.363	0.101	0.017*	
		3rd Qu.	0.408	0.380	0.136	0.029	
		Max.	0.422	0.422	0.229	0.064	

* Value fell below cutoff, indicating propensity scores for the treatment level collapsed near zero.

Table 12

Inverse Propensity Score Weight Summary Statistics

Subgroup	Min	1st Quartile	Median	Mean	3rd Quartile	Max
3 Majors						
Full Group	0.223	0.904	0.963	0.997	1.086	2.494
Non-White	0.288	0.895	0.978	0.998	1.075	2.921
Female	0.271	0.873	0.981	0.998	1.077	3.199
Pell	0.230	0.909	0.971	0.997	1.061	2.211
5 Majors						
Full Group	0.165	0.903	0.971	0.998	1.082	2.750
Non-White	0.228	0.885	0.976	0.998	1.086	3.538
Female	0.234	0.885	0.969	1.001	1.086	2.596

Data Balance

Data balance was assessed by regressing each variable on dosage before and after weighting. A standardized regression coefficient smaller than 0.1 represents sufficient balance in the weighted sample (Leite, 2016). As shown in Table 13 and Figure 5, weighting improved data balance for all seven analysis groups. Although data remained imbalanced in a few instances, most of the coefficients were close to 0.1, and the unbalanced variables were included in the outcome model as recommended.

Outcome Model Fit

Residuals were assessed with binned plots (Figure 6) (Gelman et al., 2022; *R: Binned Residual Plot*, n.d.). In the plots, 95% of binned residuals are expected to fall within plus or minus two standard errors (between the gray lines), with the majority of markers in the region representing reasonably good fit (Webb, n.d., sec. 8.7). Model fit was deemed reasonably good. Summary statistics for residuals are available in the Appendix C. Model fits are reported with R^2 and adjusted R^2 values in Table 14.

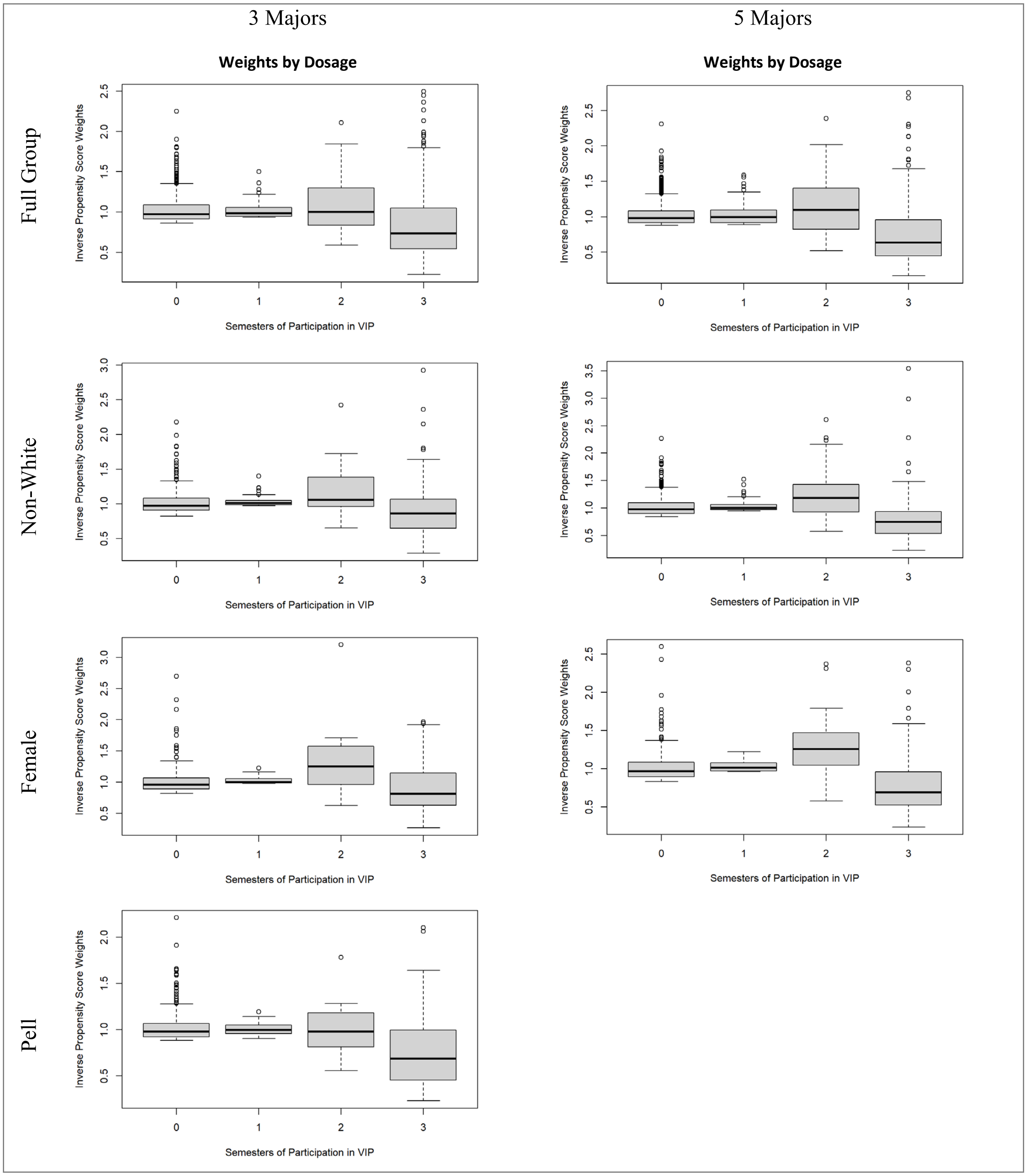


Figure 4. Inverse Propensity Score Weights by Dosage Level

Table 13

Data Balance: Standardized Regression Coefficients before and after Weighting

Majors Grouping	Full Group		Non-White		Female		Pell	
Variable	Baseline	Weighted	Baseline	Weighted	Baseline	Weighted	Baseline	Weighted
3 Majors								
Citizenship	0.00	0.02	0.10	0.00	0.02	0.06	0.02	0.05
CoOperative Education	0.21	0.09	0.02	0.05	0.44	0.28*	0.22	0.18*
Female	0.21	0.02	0.21	0.02			0.17	0.01
Freshman Exp Course	0.01	0.01	0.11	0.01	0.03	0.03	0.11	0.08
Full Time Internship	0.24	0.03	0.21	0.03	0.15	0.05	0.39	0.03
GPA	0.26	0.05	0.25	0.02	0.19	0.01	0.20	0.03
Greek Fraternity/Sorority	0.03	0.02	0.07	0.00	0.07	0.03	0.05	0.02
Honors Program	0.38	0.04	0.50	0.03	0.62	0.04		
Major	0.16	0.10	0.30	0.12*	0.31	0.06	0.22	0.07
Pell Grant	0.09	0.02	0.19	0.03	0.11	0.01		
Race/Ethnicity	0.35	0.05	0.15	0.03	0.37	0.05	0.26	0.25*
Study Abroad	0.01	0.01	0.04	0.04	0.00	0.00	0.14	0.11*
Transfer Exp Course							0.18	0.00
Transfer Student	0.09	0.01	0.13	0.01	0.16	0.01	0.04	0.02
Undergradaute Research	0.01	0.02	0.05	0.04	0.06	0.07	0.01	0.01
5 Majors								
Citizenship	0.02	0.05	0.09	0.02	0.07	0.06		
CoOperative Education	0.26	0.13*	0.20	0.10	0.32	0.10		
Female	0.23	0.02	0.22	0.02				
Freshman Exp Course	0.00	0.01	0.10	0.02	0.02	0.04		
Full Time Internship	0.20	0.02	0.17	0.01	0.12	0.05		
GPA	0.23	0.03	0.24	0.02	0.24	0.00		
Greek Fraternity/Sorority	0.04	0.00	0.05	0.02	0.05	0.03		
Honors Program	0.39	0.03	0.52	0.02	0.59	0.07		
Major	0.21	0.16*	0.33	0.13*	0.33	0.10		
Pell Grant	0.09	0.01	0.17	0.00	0.08	0.01		
Race/Ethnicity	0.38	0.06	0.16	0.02	0.35	0.01		
Study Abroad	0.01	0.01	0.04	0.07	0.01	0.01		
Transfer Exp Course								
Transfer Student	0.13	0.02	0.17	0.04	0.17	0.04		

* Sample remained imbalanced after weighting.

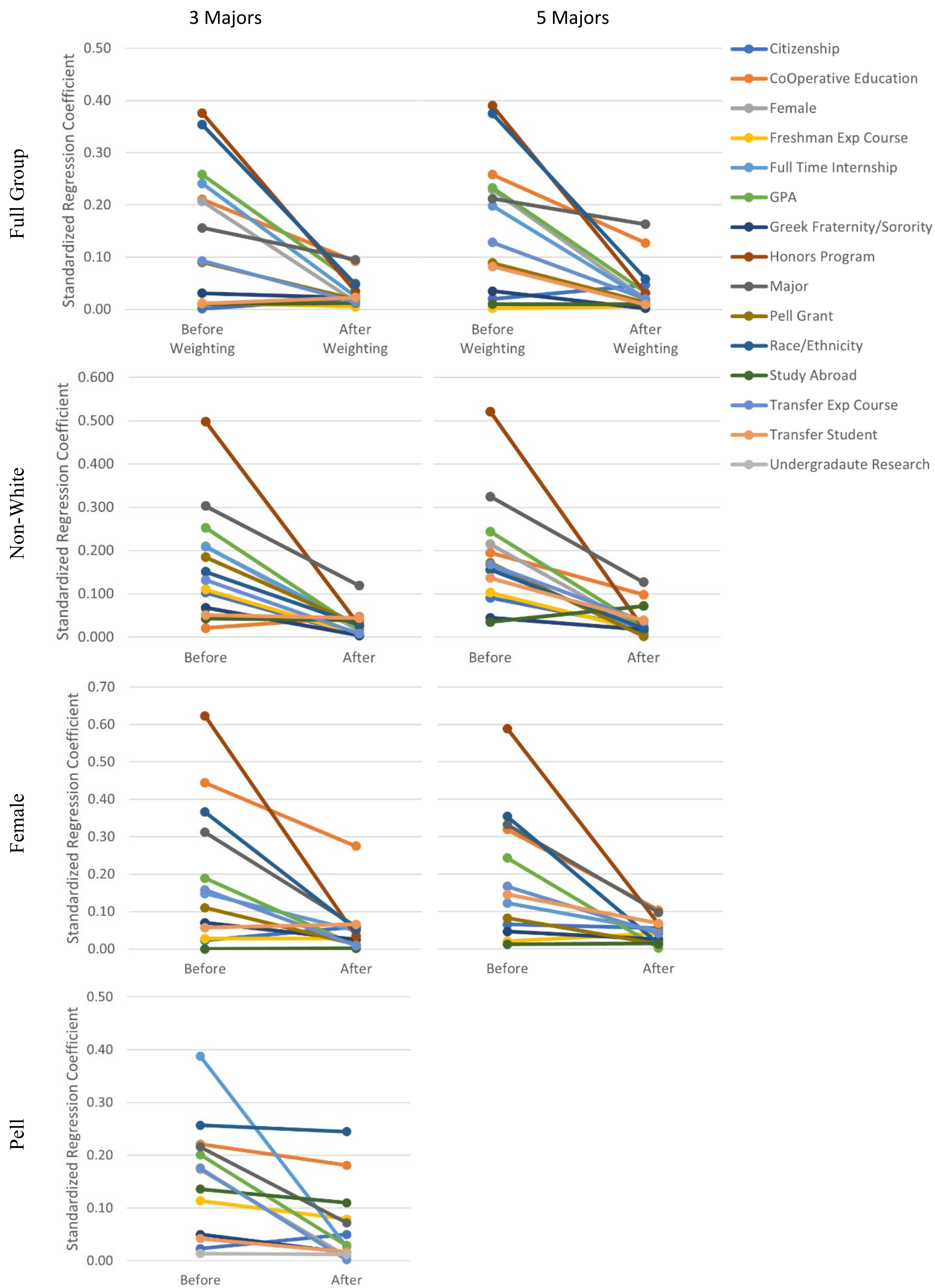


Figure 5. Data Balance before and after Weighting

Table 14

Outcome Model Fit

	Full Group		Non-White		Female		Pell
	3 Majors	5 Majors	3 Majors	5 Majors	3 Majors	5 Majors	3 Majors
N	1730	2446	951	1247	474	631	434
ESS	1642	2315	904	1173	438	587	410
R^2	.146	.152	.151	.170	.143	.116	.170
Adj. R^2	.134	.143	.130	.153	.099	.080	.124

Estimates of Effects

Regression results and 95% confidence intervals for each grouping appear in Tables 15-21. Adjusted odds ratios for statiscially significant variables are summarized in Table 22. All odds ratios and 95% confidence intervals are available for each analysis in Appendix C. To aid in the interpretation of results for each subgroup, variables in Figure 7 are sorted by adjusted odds ratios, placing the most influential variables at the top of each list, and color coded by expected effects or variable category. To focus on student background and experiences, major and graduation year were excluded from the figure. Shades of green represent variables expected to have positive effects on job placement (GPA and work-based learning experiences). Shades of red represent varaibles expected to have negative effects (being Pell grant recipients and transfer students). Variables for the Greek social system and study abroad are student engagement programs and are shaded purple. While the freshman experience course is an academic activity, it is designed to increase student engagement, so it is shaded purple another with the student engagement variables. Semesters of VIP participation are shaded gold. Gender is unshaded, because Koc found women in engineering did now have lower job placement than men (Koc, 2014). As mentioned, variables were sorted from largest to smallest adjusted odds ratio. When ratios were less than one, inverse odds ratios (the odds of not having found a job) were used to sort.

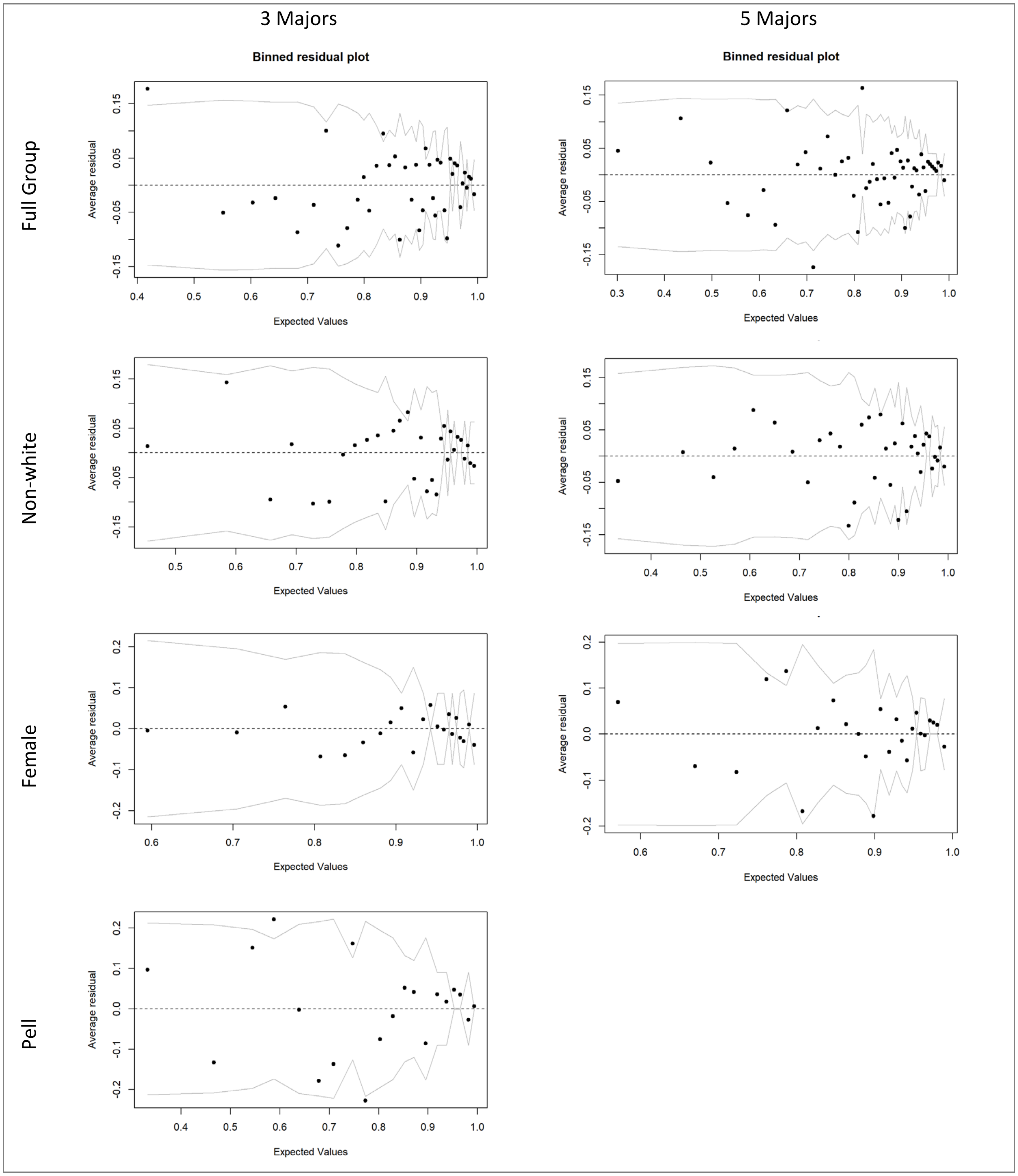


Figure 6. Binned Residual Plots

Table 15

Regression Results, 3 Majors, Full Group

Variable	Estimate	*SE*	*t*	*p*	95% CI	
					LL	UL
(Intercept)	-3.791	0.699	-5.428	<.001***	-5.161	-2.421
VIPSEM	0.382	0.103	3.718	<.001***	0.181	0.584
CITZ Resident NonCitizen	0.174	0.300	0.579	.563	-0.415	0.762
Female	0.443	0.205	2.161	.031*	0.041	0.845
RCETH Asian	0.243	0.186	1.307	.191	-0.122	0.608
RCETH Other or Unknown	0.349	0.328	1.064	.288	-0.295	0.993
RCETH URM	0.453	0.240	1.886	.059	-0.018	0.924
Pell	-0.424	0.164	-2.585	.010**	-0.745	-0.102
Transfer Student	-0.272	0.192	-1.419	.156	-0.647	0.104
Greek System	0.816	0.219	3.721	<.001***	0.386	1.246
Study Abroad	0.534	0.239	2.236	.026*	0.066	1.002
Freshman Exp Course	0.007	0.206	0.036	.971	-0.396	0.410
Honors Program	0.509	0.442	1.153	.249	-0.357	1.375
Major Computer Eng	1.050	0.363	2.890	.004**	0.337	1.762
Major Computer Sci	1.248	0.339	3.682	<.001***	0.583	1.913
Undergrad Research	-0.109	0.205	-0.532	.595	-0.510	0.292
COOP1 Some CoOp	0.110	0.310	0.354	.724	-0.498	0.718
COOP3 CoOpDegDesig	1.358	0.480	2.828	.005**	0.416	2.299
Full-time Internship	1.031	0.229	4.506	<.001***	0.582	1.480
GPA	1.069	0.156	6.866	<.001***	0.764	1.375
YR2018	0.518	0.275	1.886	.060	-0.021	1.056
YR2019	0.076	0.268	0.283	.778	-0.451	0.602
YR2020	-0.364	0.252	-1.448	.148	-0.858	0.129
YR2021	-0.092	0.274	-0.334	.739	-0.629	0.446
YR2022	-0.313	0.252	-1.242	.214	-0.807	0.181

* Significant at the .05 level; ** Significant a the .01 level; *** Significant at the .001 level.

Table 16

Regression Results, 3 Majors, Non-white Students

Variable	Estimate	*SE*	*t*	*p*	95% CI	
					LL	UL
(Intercept)	-2.394	1.003	-2.387	.017*	-4.362	-0.426
VIPSEM	0.414	0.123	3.364	.001***	0.172	0.655
CITZ Resident NonCitizen	0.302	0.312	0.968	.333	-0.310	0.915
Female	0.393	0.264	1.486	.138	-0.126	0.911
RCETH Asian	0.006	0.338	0.016	.987	-0.658	0.669
RCETH URM	0.201	0.370	0.543	.587	-0.525	0.926
Pell	-0.479	0.218	-2.198	.028*	-0.907	-0.051
Transfer Student	-0.487	0.247	-1.971	.049*	-0.973	-0.002
Greek System	0.609	0.360	1.691	.091	-0.098	1.315
Study Abroad	1.016	0.414	2.453	.014*	0.203	1.829
Freshman Exp Course	-0.087	0.300	-0.290	.772	-0.676	0.502
Honors Program	0.773	0.781	0.989	.323	-0.760	2.305
Major Computer Eng	0.112	0.636	0.176	.861	-1.137	1.360
Major Computer Sci	0.514	0.614	0.837	.403	-0.691	1.718
Undergrad Research	-0.240	0.273	-0.879	.379	-0.775	0.295
COOP1 Some CoOp	0.347	0.423	0.822	.411	-0.482	1.177
COOP3 CoOpDegDesig	1.743	1.100	1.584	.114	-0.416	3.902
Full-time Internship	0.798	0.298	2.678	.008**	0.213	1.384
GPA	0.981	0.216	4.550	<.001***	0.558	1.403
YR2018	0.703	0.431	1.631	.103	-0.143	1.548
YR2019	-0.271	0.362	-0.748	.455	-0.980	0.439
YR2020	-0.231	0.362	-0.639	.523	-0.941	0.478
YR2021	0.104	0.370	0.282	.778	-0.621	0.829
YR2022	-0.316	0.348	-0.907	.365	-0.999	0.368

* Significant at the .05 level; ** Significant a the .01 level; *** Significant at the .001 level.

Table 17

Regression Results, 3 Majors, Female Students

Variable	Estimate	*SE*	*t*	*p*	95% CI	
					LL	UL
(Intercept)	-3.752	1.406	-2.668	.008**	-6.515	-0.988
VIPSEM	0.250	0.183	1.369	.172	-0.109	0.610
CITZ Resident NonCitizen	-0.234	0.700	-0.335	.738	-1.610	1.142
RCETH Asian	0.517	0.457	1.130	.259	-0.382	1.415
RCETH Other or Unknown	0.128	0.597	0.215	.830	-1.045	1.301
RCETH URM	0.579	0.501	1.157	.248	-0.405	1.563
Pell	0.310	0.387	0.801	.424	-0.451	1.070
Transfer Student	0.083	0.479	0.173	.863	-0.859	1.024
Greek System	0.721	0.425	1.698	.090	-0.114	1.555
Study Abroad	0.816	0.404	2.019	.044*	0.022	1.611
Freshman Exp Course	0.151	0.400	0.377	.706	-0.635	0.937
Honors Program	1.633	0.869	1.879	.061	-0.075	3.342
Major Computer Eng	0.865	0.582	1.486	.138	-0.279	2.009
Major Computer Sci	1.122	0.432	2.595	.010**	0.272	1.971
Undergrad Research	-0.279	0.384	-0.728	.467	-1.033	0.474
COOP1 Some CoOp	0.479	0.630	0.761	.447	-0.758	1.717
COOP3 CoOpDegDesig	1.529	1.226	1.247	.213	-0.880	3.937
Full-time Internship	0.247	0.407	0.607	.544	-0.553	1.048
GPA	1.078	0.335	3.222	.001**	0.421	1.736
YR2018	1.065	0.805	1.323	.186	-0.516	2.646
YR2019	0.866	0.651	1.331	.184	-0.413	2.145
YR2020	-0.709	0.497	-1.427	.154	-1.686	0.268
YR2021	-0.064	0.581	-0.111	.912	-1.207	1.078
YR2022	0.408	0.564	0.723	.470	-0.701	1.516

* Significant at the .05 level; ** Significant a the .01 level; *** Significant at the .001 level.

Table 18

Regression Results, 3 Majors, Pell Grant Recipients

Variable	Estimate	*SE*	*t*	*p*	95% CI	
					LL	UL
(Intercept)	-4.734	1.314	-3.604	<.001***	-7.316	-2.152
VIPSEM	0.581	0.171	3.406	.001***	0.246	0.917
CITZ Resident NonCitizen	-0.480	0.380	-1.263	.207	-1.226	0.267
Female	0.999	0.385	2.595	.010**	0.242	1.755
RCETH Asian	0.076	0.333	0.229	.819	-0.578	0.731
RCETH Other or Unknown	1.346	0.906	1.486	.138	-0.435	3.126
RCETH URM	0.412	0.384	1.075	.283	-0.342	1.166
Transfer Student	0.028	0.304	0.091	.927	-0.570	0.626
Greek System	0.696	0.523	1.332	.184	-0.332	1.724
Study Abroad	0.148	0.576	0.257	.798	-0.984	1.280
Freshman Exp Course	1.457	0.502	2.902	.004**	0.470	2.444
Transfer Exp Course	-0.659	0.512	-1.286	.199	-1.666	0.348
Major Computer Eng	1.423	0.825	1.724	.086	-0.200	3.045
Major Computer Sci	2.028	0.783	2.590	.010**	0.489	3.567
Undergrad Research	-0.110	0.433	-0.255	.799	-0.962	0.741
COOP1 Some CoOp	0.228	0.615	0.371	.711	-0.981	1.436
COOP3 CoOpDegDesig	2.290	1.064	2.152	.032*	0.198	4.382
Full-time Internship	0.596	0.386	1.543	.124	-0.163	1.356
GPA	0.953	0.273	3.493	.001***	0.417	1.488
YR2018	0.770	0.485	1.587	.113	-0.184	1.724
YR2019	-0.300	0.427	-0.703	.483	-1.140	0.539
YR2020	-0.331	0.434	-0.762	.447	-1.184	0.523
YR2021	0.202	0.495	0.409	.683	-0.771	1.175

* Significant at the .05 level; ** Significant a the .01 level; *** Significant at the .001 level.

Table 19

Regression Results, 5 Majors, Full Group

Variable	Estimate	*SE*	*t*	*p*	95% CI	
					LL	UL
(Intercept)	-3.453	0.556	-6.209	<.001***	-4.544	-2.363
VIPSEM	0.286	0.082	3.508	.001***	0.126	0.447
CITZ Resident NonCitizen	-0.005	0.238	-0.020	.984	-0.471	0.462
Female	0.497	0.161	3.098	.002**	0.182	0.812
RCETH Asian	0.103	0.151	0.680	.497	-0.193	0.398
RCETH Other or Unknown	0.213	0.259	0.824	.410	-0.294	0.720
RCETH URM	0.195	0.173	1.131	.258	-0.143	0.534
Pell	-0.446	0.131	-3.408	.001***	-0.702	-0.189
Transfer Student	-0.188	0.153	-1.226	.220	-0.488	0.113
Greek System	0.736	0.168	4.394	<.001***	0.408	1.065
Study Abroad	0.237	0.155	1.531	.126	-0.067	0.541
Freshman Exp Course	0.002	0.155	0.013	.990	-0.303	0.307
Honors Program	0.536	0.363	1.476	.140	-0.176	1.248
Major Aerospace Eng	0.188	0.315	0.597	.551	-0.429	0.805
Major Computer Eng	1.165	0.340	3.424	.001***	0.498	1.832
Major Computer Sci	1.348	0.312	4.319	<.001***	0.736	1.960
Major Electrical Eng	0.863	0.336	2.570	.010*	0.205	1.522
Undergrad Research	0.101	0.153	0.659	.510	-0.199	0.400
COOP1 Some CoOp	0.242	0.242	0.998	.318	-0.233	0.717
COOP3 CoOpDegDesig	1.153	0.276	4.173	<.001***	0.611	1.694
Full-time Internship	0.982	0.181	5.438	<.001***	0.628	1.335
GPA	0.957	0.123	7.761	<.001***	0.716	1.199
YR2018	0.396	0.200	1.976	.048*	0.003	0.789
YR2019	0.058	0.197	0.296	.767	-0.328	0.445
YR2020	-0.326	0.199	-1.637	.102	-0.717	0.065
YR2021	0.249	0.218	1.139	.255	-0.179	0.677
YR2022	-0.122	0.201	-0.604	.546	-0.516	0.273

* Significant at the .05 level; ** Significant a the .01 level; *** Significant at the .001 level.

Table 20

Regression Results, 5 Majors, Non-white Students

Variable	Estimate	*SE*	*T*	*p*	95% CI	
					LL	UL
(Intercept)	-2.642	0.881	-2.999	.003**	-4.371	-0.914
VIPSEM	0.375	0.103	3.628	<.001***	0.172	0.578
CITZ Resident NonCitizen	0.104	0.259	0.402	.688	-0.405	0.613
Female	0.496	0.225	2.205	.028*	0.055	0.937
RCETH Asian	-0.070	0.281	-0.250	.802	-0.621	0.480
RCETH URM	0.078	0.293	0.267	.789	-0.497	0.653
Pell	-0.477	0.179	-2.668	.008**	-0.828	-0.126
Transfer Student	-0.418	0.207	-2.018	.044*	-0.825	-0.012
Greek System	0.565	0.289	1.953	.051	-0.003	1.132
Study Abroad	0.415	0.250	1.663	.097	-0.075	0.905
Freshman Exp Course	-0.139	0.234	-0.595	.552	-0.597	0.319
Honors Program	0.671	0.652	1.028	.304	-0.609	1.951
Major Aerospace Eng	-0.551	0.633	-0.871	.384	-1.793	0.691
Major Computer Eng	0.367	0.637	0.576	.565	-0.883	1.616
Major Computer Sci	0.720	0.610	1.180	.238	-0.477	1.918
Major Electrical Eng	0.003	0.637	0.004	.997	-1.248	1.253
Undergrad Research	0.188	0.235	0.801	.424	-0.273	0.649
COOP1 Some CoOp	0.403	0.348	1.158	.247	-0.280	1.086
COOP3 CoOpDegDesig	1.393	0.535	2.605	.009**	0.344	2.442
Full-time Internship	0.868	0.244	3.554	<.001***	0.389	1.347
GPA	0.990	0.176	5.630	<.001***	0.645	1.335
YR2018	0.652	0.312	2.089	.037*	0.040	1.264
YR2019	-0.128	0.277	-0.463	.643	-0.671	0.415
YR2020	-0.258	0.293	-0.881	.378	-0.834	0.317
YR2021	0.189	0.300	0.629	.530	-0.400	0.778
YR2022	-0.224	0.286	-0.783	.434	-0.785	0.337

* Significant at the .05 level; ** Significant a the .01 level; *** Significant at the .001 level.

Table 21

Regression Results, 5 Majors, Female Students

Variable	Estimate	*SE*	*T*	*p*	95% CI LL	95% CI UL
(Intercept)	-2.571	1.176	-2.187	.029*	-4.879	-0.263
VIPSEM	0.282	0.151	1.863	.063	-0.015	0.580
CITZ Resident NonCitizen	0.332	0.565	0.587	.557	-0.778	1.441
RCETH Asian	0.134	0.370	0.362	.717	-0.593	0.860
RCETH Other or Unknown	0.323	0.550	0.587	.557	-0.757	1.404
RCETH URM	0.403	0.400	1.006	.315	-0.383	1.189
Pell	-0.155	0.286	-0.543	.587	-0.717	0.406
Transfer Student	-0.231	0.377	-0.614	.540	-0.971	0.509
Greek System	0.563	0.380	1.481	.139	-0.184	1.309
Study Abroad	0.370	0.303	1.221	.223	-0.225	0.965
Freshman Exp Course	0.181	0.324	0.558	.577	-0.456	0.818
Honors Program	0.696	0.742	0.938	.349	-0.761	2.154
Major Aerospace Eng	0.174	0.451	0.386	.700	-0.711	1.059
Major Computer Eng	0.942	0.568	1.659	.098	-0.173	2.056
Major Computer Sci	1.202	0.422	2.848	.005**	0.373	2.031
Major Electrical Eng	0.349	0.473	0.738	.461	-0.581	1.279
Undergrad Research	-0.020	0.318	-0.062	.951	-0.643	0.604
COOP1 Some CoOp	0.059	0.499	0.118	.906	-0.922	1.039
COOP3 CoOpDegDesig	0.857	0.604	1.418	.157	-0.329	2.042
Full-time Internship	0.626	0.351	1.782	.075	-0.064	1.315
GPA	0.848	0.282	3.010	.003**	0.295	1.402
YR2018	0.577	0.512	1.126	.261	-0.429	1.583
YR2019	0.624	0.519	1.201	.230	-0.396	1.644
YR2020	-0.597	0.431	-1.386	.166	-1.442	0.249
YR2021	0.032	0.467	0.068	.946	-0.885	0.948
YR2022	0.022	0.441	0.049	.961	-0.844	0.888

* Significant at the .05 level; ** Significant a the .01 level; *** Significant at the .001 level.

Table 22

Adjusted Odds Ratios for Statistically Significant Background and Experience Variables

	Full Groups		Non-White		Female		Pell
	3 Majors	5 Majors	3 Majors	5 Majors	3 Majors	5 Majors	3 Majors
	AOR	AOR	AOR	AOR	AOR	AOR	AOR
Background							
Female	1.557	1.644		1.642			2.714
SES							
Pell	0.655 (1.53)	0.64 (1.56)	0.619 (1.61)	0.621 (1.61)			
Transfer			0.614 (1.63)	0.658 (1.52)			
Academic							
GPA .4 Higher	1.534	1.467	1.480	1.486	1.539	1.404	1.464
VIP 1 sem	1.465	1.773	2.288	2.117			3.197
VIP 2 sem	2.147	2.361	3.46	3.079			5.716
VIP 3 sem	3.146	1.332	1.513	1.455			1.788
Student Engagement							
Fresh Exp Crse							4.293
Greek	2.262	2.088					
Study Abroad	1.705		2.762		2.262		
Career Related							
CoOp Deg.	3.887	3.166		4.028			9.876
Internship	2.804	2.668	2.222	2.382			

Note: To focus on student background and experiences, major and graduation year excluded.

	3 Majors		5 majors	
Full Group	CoOp Degree Designator	3.9	CoOp Degree Designator	3.2
	VIP 3 Semesters	3.1	Full-Time Internship	2.7
	Full-Time Internship	2.8	VIP 3 Semesters	2.4
	Greek	2.3	Greek	2.1
	VIP 2 Semesters	2.1	VIP 2 Semesters	1.8
	Study Abroad	1.7	Female	1.6
	Female	1.6	Pell Grant	(1.6) [Rev]
	Pell Grant	(1.5)[Rev]	GPA .4 pts Higher	1.5
	GPA .4 pts Higher	1.5	VIP 1 Semester	1.3
	VIP 1 Semester	1.5		
Non-White	VIP 3 Semesters	3.5	CoOp Degree Designator	4
	Study Abroad	2.8	VIP 3 Semesters	3.1
	VIP 2 Semesters	2.3	Full-Time Internship	2.4
	Full-Time Internship	2.2	VIP 2 Semesters	2.1
	Transfer Student	(1.6) [Rev]	Female	1.6
	Pell Grant	(1.6) [Rev]	Pell Grant	(1.6) [Rev]
	VIP 1 Semester	1.5	Transfer Student	(1.5) [Rev]
	GPA .4 pts Higher	1.5	GPA .4 pts Higher	1.5
			VIP 1 Semester	1.5
Female	Study Abroad	2.3	GPA .4 pts Higher	1.4
	GPA .4 pts Higher	1.5		
Pell	CoOp Degree Designator	9.9		
	VIP 3 Semesters	5.7		
	Freshman Exp Course	4.3		
	VIP 2 Semesters	3.2		
	Female	2.7		
	VIP 1 Semester	1.8		
	GPA .4 pts Higher	1.5		

Positive Correlation Expected

Negative Correlation Expected

Semesters of VIP

Student Engagement

[Rev] Odds of not having found a job

Figure 7. Ajusted Odds Ratios by Subgroup, Rounded to the Nearest 10th

Note: To focus on student background and experiences, major and graduation year excluded.

Dosage Assumption

To examine the assumption of a dosage effect, analysis was repeated for the two full major groupings with the VIP semesters of participation variable handled as a factor, with zero semesters of participation as the reference level. Results for the two analyses showed a similar pattern. The coefficients were progressively larger for each factor level, implying stronger effects by semesters of participation (Figure 8). Statistical significance also increased with dosage level, with only the third semester of participation having statistical significance at the .05 level (Table 23). Together, these implied the dosage assumption was reasonable.

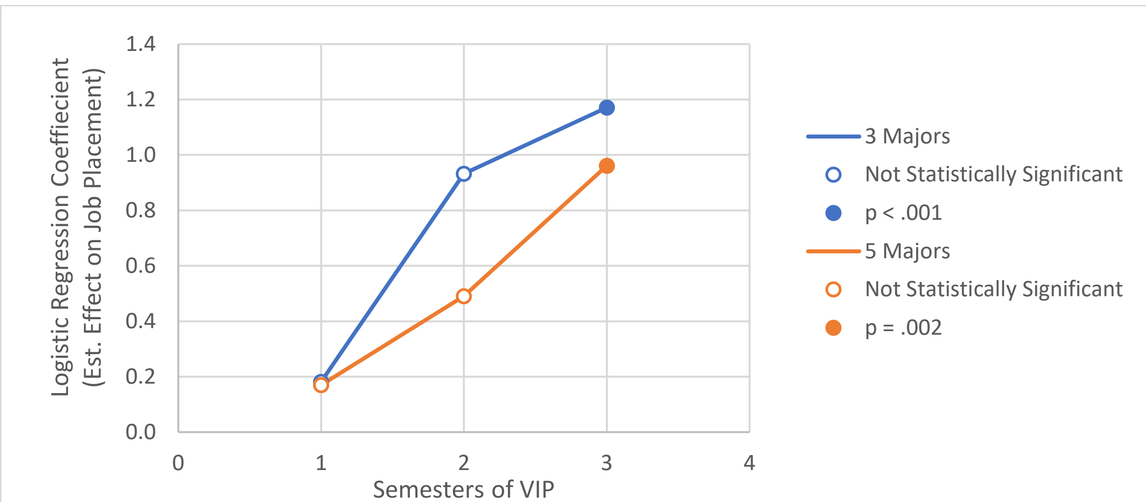


Figure 8. Examining Dosage Assumption: Semesters of VIP Handled as a Factor Variable

Table 23

Logistic Regression Results for Dosage Treated as a Factor

Variable	Estimate	*SE*	*T*	*p*	95% CI	
					LL	UL
3 Majors						
VIPSEM1	0.182	0.267	0.680	0.496	-0.342	0.706
VIPSEM2	0.932	0.663	1.407	0.160	-0.368	2.232
VIPSEM3	1.171	0.342	3.428	0.001	0.501	1.841
5 Majors						
VIPSEM1	0.170	0.210	0.808	0.419	-0.242	0.582
VIPSEM2	0.490	0.332	1.477	0.140	-0.160	1.141
VIPSEM3	0.961	0.312	3.079	0.002	0.349	1.573

Race/Ethnicity and Socioeconomic Status

To provide context for the relationship between race/ethnicity, SES status, and variables associated with higher and lower job placement, representation/participation in statistically significant programs and groups were examined for the five majors grouping (Figures 9-10). As shown in Figure 9, more URM and Asian students were Pell grant recipients and transfer students. With the exception of VIP, URM and Asian students tended to participate in programs associated with higher job placement at lower rates than white students (with the exception of Asian students in internships). Examination by SES showed substantial overlap between Pell status and transfer student status (Figure 10). Pell students and transfer students participated in programs associated with higher job placement at lower rates than non-Pell non-transfer students.

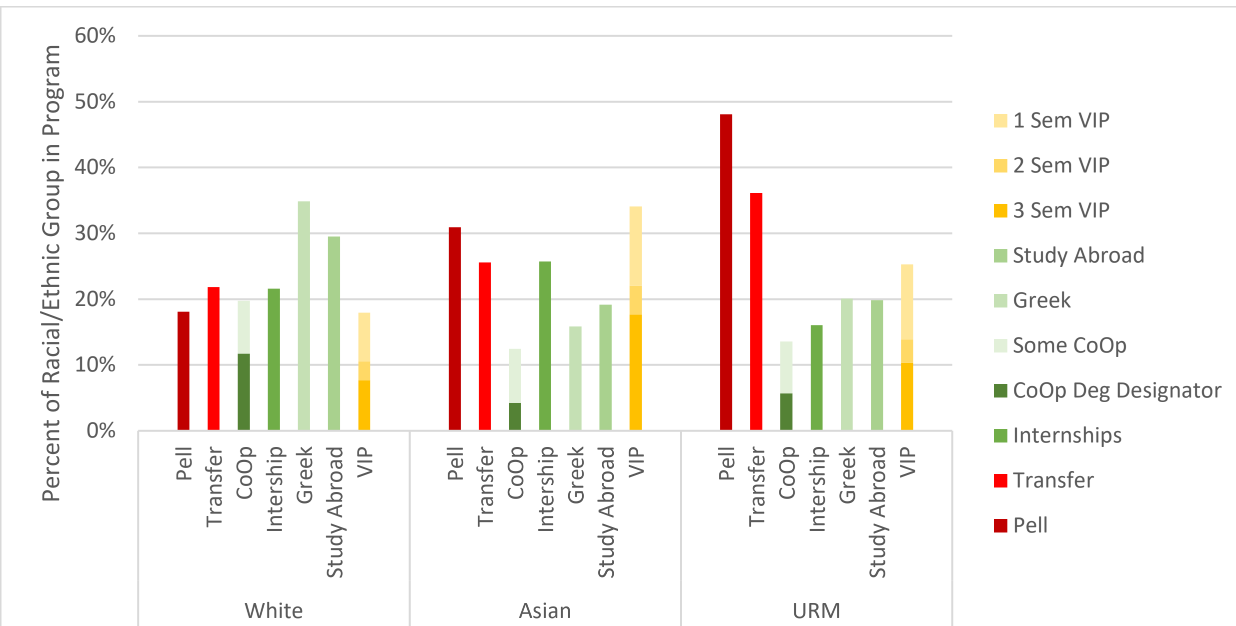


Figure 9. Membership in Programs and Groups Associated with Higher and Lower Job Placement by Race/Ethnicity, 5 Majors

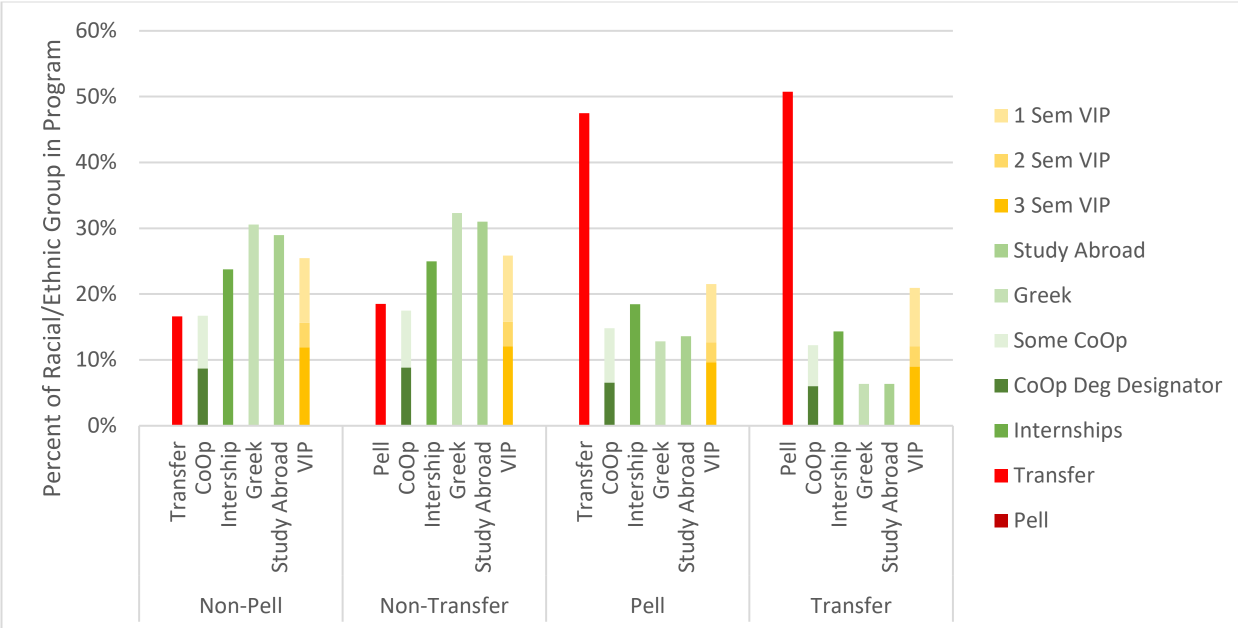


Figure 10. Membership in Programs and Groups Associated with Higher and Lower Job Placement by SES, 5 Majors

A notable result was that among Pell grant recipients, enrollment in the freshman experience course was associated with higher job placement, while it was not statistically significant for any other group. Comparisons between Pell students who did and did not take the course showed marked differences. Students who took the freshman experience course participated in CoOp at higher rates than other Pell grant recipients (20% vs. 11%), did internships at higher rates (24% vs 18%), and participated in study abroad at higher rates (22% vs 8%).

Participation in VIP by Major

While semesters of participation in VIP by major were considered in earlier stages of the analysis, participation was visualized to aid in interpretation of the results (Figure 11).

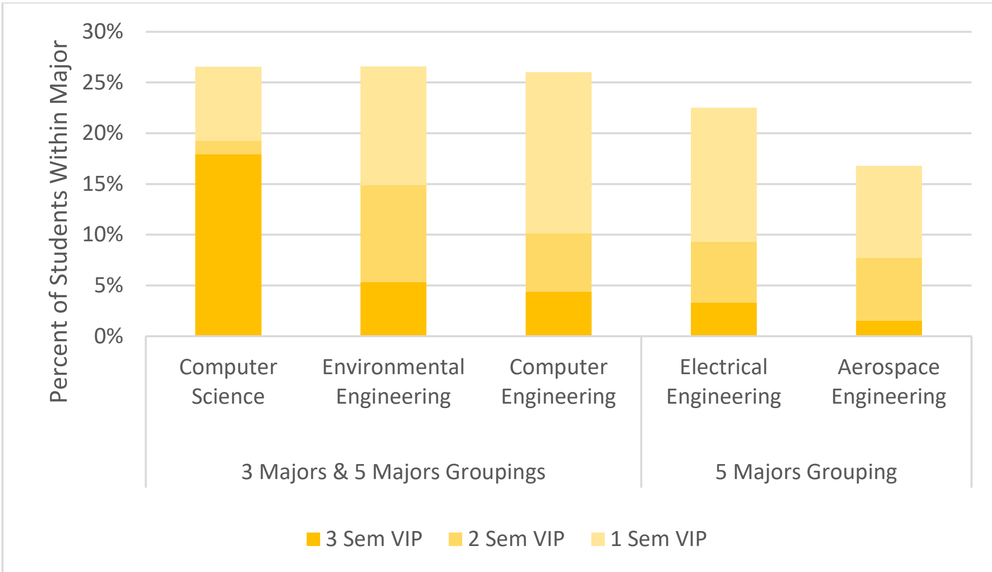


Figure 11. Semesters of Participation in VIP by Degree Program

5 DISCUSSION

Employers have limited information on applicants, so they rely on signals that indicate fitness for employment (Spence, 1973). These signals are rooted in human capital, cultural capital, and social capital, and judgements of signals are shaped by race, ethnicity, gender, and social class (Bertrand & Mullainathan, 2004; Gaddis, 2015; Ingram & Allen, 2019; S. K. Kang et al., 2016; Koc, 2014; Milkman et al., 2015; Quillian et al., 2017; Rivera, 2016). This study sought to answer two research questions: the degree to which LT-PBL-EFR affected job placement; and whether participation influenced equity in job placement.

Limitations

A limitation of this study is that it is not experimental. Inverse propensity score weighting simulates experimental design through data balancing, but it is not as rigorous as randomized experiments. While propensity score analysis can be used in causal inference, the results imply the presence of unmeasured confounders, so causal inferences would be inappropriate.

The study also relied on observational data collected by the institution, which is limited. Institutional data included demographics, academic records, status as Pell grant recipients, status as transfer students, and participation in co-curricular programs such as work-based learning, living learning communities, Greek fraternities and sororities, and status as an athlete. However, beyond the Greek social system, the dataset did not include measures of social engagement such as membership in student clubs or professional organizations.

The job placement data was based on surveys administered by the institute prior to graduation, which carried limitations. The first limitation is student self-selection, with approximately 50% of students responding. Second, the survey does not distinguish between students who received no job offers, students who received undesirable job offers, and students who received but had not yet accepted job offers. The dataset also did not include information on whether CoOp and internship students were hired by their CoOp/internship employers. If this information had been included, it would have allowed for more meaningful comparisons between students who were applying to companies for which they had not yet worked. Finally, using job placement as the outcome did not differentiate between high and low salaries.

The study focused on a single institution, so the findings cannot be generalized to other populations. Under-representation of historically underserved students at the institution (approximately 8% black/African American and 8% Hispanic/Latino) affected the potential sample size for the subgroups. High enrollment in VIP, multiple years of data, and analysis by subgroup helped mitigate but did not eliminate this effect.

The literature review cited discrimination found in resume audit studies. While these informed the theoretical framework, this study did not measure discrimination. Instead, it was designed to measure statistical significance across groups, to determine which programs or background attributes may be associated with higher or lower job placement across the full sample and for subgroups.

The study was also limited by the number and types of majors that were included, because participation in VIP varies by major. Georgia Tech is also a predominately technological institution. While the sample initially included students from thirty-one majors, majors with low participation rates in two and three semesters of VIP did not meet assumptions for the method, reducing the number of included majors to five. Although a large portion of treated students were Computer Science majors, their high likelihood of participating in VIP resulted in smaller weights for Computer Science participants, mitigating the imbalance.

General Job Placement Findings

VIP The first research question asked how participation in LT-PBL-EFR at Georgia Tech impacted job placement for college graduates, with job placement reported in a campus survey prior to graduation. VIP can be framed as a source of social and cultural capital. Students may broaden their social networks by working closely with faculty who may be knowledgeable about and/or have connections in industry, with students from different majors, and with students from different backgrounds. Some teams also interact with industry partners, yielding contacts and potential references for students seeking employment. Cultural capital may be conferred as students learn the culture of their disciplines (and adjacent disciplines) in an applied project. Working for on a large-scale project involves cooperation, dedication, and persistence among students who participate for multiple semesters, which are dispositions valued by employers. While participation may not confer these dispositions – students could already have them –VIP may provide students with signals that represent the dispositions, and which employers positively receive.

If VIP enables students to gain experience, develop qualities desired by employers, and pursue passions – all in contexts related to their fields – participants would be better positioned to send signals valued by employers. If true, this would yield higher job placement among participants compared to non-participants, and findings support this hypothesis. Results showed that across the five majors retained in the analysis, when other factors were held constant, participation in three semesters of VIP was associated with triple the odds of having found a job in the three majors grouping, and more than double the odds in the five majors grouping.

Work-based Learning In the framework, work-based learning is a source of human capital where students learn and apply career-related skills; cultural capital, as students learn the practices and standards of their field; and social capital, with students expanding their networks. In both groupings, work-based learning was associated with the highest adjusted odds ratio, with CoOp the most influential with odds of 3.9 and 3.2. in the three and five majors groupings. There was similarity in adjusted odds ratios associated with internships and three semesters of VIP, with odds of

3.1 and 2.8 for VIP and internships in the three majors grouping, and 2.4 and 2.7 in the five majors grouping. While VIP involves fewer hours than an internship, it may yield human capital development in a work-like context. Student accountability to instructors would mirror accountability to supervisors, collaborations with classmates would resemble collaborations with coworkers, and grades would be equivalent to performance evaluations. Also like the workplace, students are expected to get up to speed on an existing project, work around constraints, and document their work so others can continue after they leave. As a one to two credit course, students attend a one-hour class and are expected to work two to three hours outside of class each week for each credit. Depending on the number of credits taken each semester, three semesters of participation would involve 135 to 315 hours of work, comparable to four to eight weeks of work in a full-time job.

Significant Student Engagement Variables Two variables identified as potential sources and signifiers of social and cultural capital, the Greek social system and study abroad, were associated with higher odds of job placement. Fraternity and sorority membership is exclusive by definition, and membership is built upon signaling. Unlike clubs with open membership, students pledge (petition to join) fraternities and sororities, and acceptances (bids) are based on members' judgement of applicants' fit with their organizations, similar to fit sought by employers (Anderson & Tomlinson, 2020; Finlay & Coverdill, 2002). Beyond the membership process favoring students who have signals that would be valued by employers, members also gain social and cultural capital. This includes academic support, status and socialization in an exclusive organization, and connections with campus and national alumni networks. While the Greek system is not related to career training, in both full-group analyses, membership was associated with double the odds of job placement. Selection for study abroad is not based on signals, but the program is expensive, and it serves higher proportions of white and affluent students. Study abroad is not typically career-oriented, but it was associated with higher job placement in the 3 majors grouping as well as in the non-white and female subgroups. While study abroad and the Greek social system may provide participants with social capital and signals valued by employers, some of the correlation could be spurious. Both serve disproportionately more white and affluent students who are already well-positioned to find jobs. Students' social positions may have led to participation in the programs as well as higher job placement.

Equity

The second research question asked to what extent VIP participation influences equity in job placement by race/ethnicity, gender, and socioeconomic status. This question can be explored by examining the degree to which race/ethnicity, gender, and socioeconomic status were associated with job placement across the full groupings; and by examining differences in which variables were significant for subgroups and the full groupings.

Gender Across the two full groupings, status as a woman was positively associated with job placement, which agrees with Koc's finding that engineering is an exception to the general trend of lower job placement among women (Koc, 2014). This study involved four engineering majors and computer science, but gender differences for computer science (which is outside of engineering) were not examined. As a subgroup, women seemed to neither benefit from nor be hurt by factors as expected – Pell grant status, transfer student status, work-based learning, and sorority membership were not statistically significant predictors of job placement, and VIP was not significant either. In a study involving student engagement and career outcomes, Hu & Wolniak found academic engagement associated with higher salaries among men, while social engagement was associated with higher salaries among women (Hu & Wolniak, 2013). If women's job placement is shaped by social engagement, this study design would have been unable to account for it. This would explain the lack of significance for the female subgroup on key variables and implies the presence of unmeasured confounders.

Socioeconomic Status Based on the literature, a negative association between low SES and job placement was expected. In an ideal system, students earning degrees from the same institution would be equally well prepared to find employment and equally competitive, but this was not the case. Across the two full groupings, Pell grant recipients had odds of 1.5 to 1.6 of *not* having found a job. The lower job placement odds associated with Pell status may stem from two directions: the experiences of low SES students during college, and/or employer reaction to signifiers of low SES. During college, low SES students were under-represented in every program associated with higher job placement. While these variables were included in the model, the pattern may extend to other forms of campus programing, such as student organizations, activities related to their majors, and key services that support student success (tutoring, advising, career counseling, etc.). Just as the significance of the freshman experience course may signify unmeasured engagement in the Pell subgroup, the significance of the Pell variable and observed low participation in other programs imply low engagement in unmeasured areas. If true, then SES may have an indirect effect on job placement, with low SES leading to lower engagement in unmeasured areas, and low engagement in unmeasured areas leading to lower job placement. Lower engagement could stem from family obligations or jobs; stressors such as housing, financial, or food insecurity; or the unspoken expectation that students should participate in a variety of campus programming outside of class. Lower participation in these areas would leave students with fewer signals of what Anderson and Tomlinson refer to as standout employability, putting the students at a disadvantage in their job searches (Anderson & Tomlinson, 2020).

Alternatively, the lower job placement odds may be due in part or in whole to classism in employer reactions to signifiers of low SES. At the lowest level, this may include employer judgements on accents, speech patterns, confidence and comfort in professional settings, and familiarity with or involvement in affluent hobbies. At a higher level, employers seek person-organization fit, which is mentioned across the literature on employability (Anderson & Tomlinson, 2020; Finlay & Coverdill, 2002; Tomlinson & Anderson, 2021; Williams et al., 2016). Person-organization fit can include demonstration of mindset, passion, behaviors and personal qualities valued by employers, enactment of employers' culture, and "networks and connections" (Anderson & Tomlinson, 2020, p. 3). If low SES students have less exposure/access to employer culture and employer networks, even if they were equally capable, employers would judge them as less qualified.

The significance of the VIP variable in both the full grouping and the Pell subgroup implies that, unlike Greek membership and study abroad, the differences in job placement associated with VIP persist for low SES students. Greek membership and study abroad may be signals (not sources) of capital, or they do not confer the same benefit on low SES students (discussed further below). In contrast, differences in job placement associated with VIP are seen across all groups. The benefit could be tied to signals of human capital, because the projects tend to be career related, and students develop/apply soft skills such as communication, collaboration, and conflict resolution; signals of cultural capital, because faculty typically mentor high-achieving students, and students learn the practices and culture associated with their field; or a result of social capital, because students gain connections through the program.

Race/Ethnicity The role of race/ethnicity in student job placement is intertwined with the effects of SES. Resume audit studies show pervasive discrimination in the screening of job applicants, but in this analysis, there was no correlation between race/ethnicity and job placement when other factors were accounted for. The results are similar to findings in studies on career earnings, which found no difference in earnings by race/ethnicity when institutional selectivity *and SES* were controlled for (Perna, 2005; Wolniak et al., 2008; Wolniak & Engberg, 2019). While the lack of significance of race/ethnicity is encouraging, the over-representation of non-white students among Pell grant recipients and transfer students places the typical non-white student (who is more likely to be low-SES) at a disadvantage in the job market.

Subgroup Differences by Race/Ethnicity and SES Analysis of non-white subgroup job placement yielded three key differences from results across the full sample and for the Pell subgroup. First, while CoOp was the most influential variable for the full sample and Pell subgroup, it was not significant for the non-white subgroup in the three-majors grouping (although it was significant in the five majors grouping). The lack of significance was unexpected, but it may have been due to the low proportion of non-white students completing the three semesters of work-based learning required for the CoOp degree designator, with 4.2% of Asian students and 5.7% of URM students. However, the lack of significance may be a sign of a larger issue, because non-white students completed the program at notably lower rates than white students. In the three majors group, only 25% and 31% of Asian and URM students who began the CoOp program earned the degree designator, compared to 55% of white students who began the program. The inequity in completion by race/ethnicity is problematic.

The second notable difference between results for the full group, non-white subgroup, and Pell subgroup was the lack of significance for the Greek social system in the non-white and Pell subgroups, while it was significant across both full groupings. This lack of significance may result from different scenarios. The Greek social system may confer social and/or cultural capital onto white and affluent students, but not onto non-White students and low SES students, if employers place less value on membership in the black Greek system. Resume audit studies found that including black organizations on resumes had a negative effect on invitations for interviews (S. K. Kang et al., 2016), which could explain the lack of significance of the Greek variable for the nonwhite subgroup. The high proportion of non-white students in the Pell group could effect the significance in results for the Pell subgroup as well.

The third notable difference between results for non-white subgroup compared to the full group and Pell subgroup were the negative odds associated with being a transfer student for both non-white subgroups, implying a differentially worse effect on non-white students. It is especially interesting that transfer student status was not significant for the Pell subgroup indicating the effect occurs at the intersection of race/ethnicity and transfer status, and not low SES.

Results for the Pell subgroup differed from the others in three additional ways. First, as previously discussed, the freshman experience course was associated with higher odds of job placement among Pell grant recipients. The course was not statistically significant for any other group, which indicates differential effects. Freshmen experience courses are classified as a high-impact experience (Association of American Colleges and Universities, n.d.). In the general Georgia Tech course, students learn about campus support services, attend events, visit clubs, write resumes, attend career fairs, and write reflections on their experiences. In discipline-specific freshman experience courses, students learn about the campus and are introduced to their majors, and often engage in hands-on projects. Pell grant recipients who took the course participated in CoOp at higher rates than other Pell grant recipients (20% vs. 11%), did internships at higher rates (24% vs 18%), and participated in study abroad at higher rates (22% vs 8%). While these variables were included in the model, the high engagement likely extended to other areas not represented in the study, such as student organizations, career-related clubs, etc., providing the students with signals valued by employers. This again implies the presence of unmeasured confounders.

Second and third notable differences were that internships and study abroad were not associated with higher job placement for the Pell subgroup. The lack of significance for internships was not due to low frequencies, because 19% of Pell recipients in the three majors grouping had done internships. This may indicate that Pell students do not benefit from internships to the same degree, that Pell students are participating in less prestigious internships than their peers (small businesses instead of large companies), or class-based bias against low SES applicants. While study abroad was associated with higher job placement across the three majors grouping and the three majors non-white subgroup, it was not statistically significant among Pell grant recipients. This may mean the higher job placement associated with study abroad in the full group and non-white subgroups is a product of higher SES, or that study abroad does not confer social and/or cultural capital on low SES students to the same degree as more affluent students.

VIP Access & Participation Key differences between VIP and other programs associated with higher odds of job placement are access and participation. While CoOp and internships are curricular and co-curricular programs, students are screened by employers who place generally lower value on signals associated with marginalized groups. Membership screening in the Greek social system similarly disadvantages non-white and non-affluent students, yielding lower participation rates among non-white and low SES students. While the study abroad program has modest grade point average requirements, the program is costly and yields lower participation rates among non-white students, transfer students and Pell grant recipients. In contrast, the VIP program does not screen students by GPA, resumes, or letters of recommendation. Admissions are handled on a rolling basis, which may reduce the amount of side-by-side comparison involved in typical hiring and admissions processes, which may decrease bias in admissions. The lack of screening may also appeal to students who feel marginalized and/or face discrimination, which could explain the higher participation among non-white students (Sonnenberg-Klein et al., 2018a). With the exception of internships, which have higher participation among Asian students, VIP is the only program associated with higher job placement that has higher participation among both URM and Asian students compared to white students. Additionally, Pell and transfer students participate in VIP at higher rates than any other program associated with higher odds of job placement. While adjusted odds ratios could not be compared between analysis groupings, the significance of VIP across all groups and the comparatively high participation rates among Asian, URM, Pell and transfer students shows equity in student participation and outcomes.

Alternative Explanations

Results indicate that participation in VIP is associated with higher odds of job placement prior to graduation. In the framework of the study, students develop human capital, cultural capital and social capital through VIP, and participation enables students to send signals valued by employers. The results imply that on average, signals associated with VIP benefit all students. However, there are multiple alternative explanations.

A simpler explanation for the job placement odds associated with VIP is spurious correlation. The VIP Program may simply attract students who are more engaged in campus life, who in turn find jobs at higher rates. The significance of the freshman experience course in the Pell subgroup is an example of this. Pell grant recipients who participated in the freshman experience course participated in CoOp, internships, and study abroad at higher rates than other Pell grant recipients. The freshman experience course may have influenced their engagement, but enrollment in the course may be the product of a prior orientation toward engagement. If VIP attracts students who are highly engaged in unmeasured aspects of campus life, then VIP participants would have higher job placement rates even if the program conferred no benefits. If the most engaged students participate in VIP for multiple semesters, this would yield a seeming dosage effect, even if no benefits were conferred.

Another explanation for the seeming gains is unmeasured confounders. This alternative explanation is supported by the lack of significance for key variables in the female subgroup. For women, Hu and Wolniak found correlation between social engagement and higher salaries (Hu & Wolniak, 2013). Beyond the Greek system variable, this study did not account for social engagement. Unmeasured social engagement may explain gains across the full groupings and/or differential gains for non-white and Pell students. Because this study relied on existing institutional data, this alternate explanation could only be explored by collecting additional data.

A framework other than signaling such as professional identity development and Lent's social cognitive career theory could explain the higher job placement odds. In social cognitive career theory, students' experiences shape their self-efficacy and career goals/actions; their career goals/actions shape their self-efficacy and experiences; and their self-efficacy shapes their experiences and career goals/actions (Lent et al., 1994). The model is built-out to include proximal affordances (nearby resources and support) and distal affordances (background resources and support). As students develop their professional identities, they become more dedicated to their career goals and take actions to achieve their goals. Participation in VIP may support student professional identity development, which would lead students to take actions that advance their career development. Additionally, under Holmes' idea of employability as processual, if students have similar qualifications, those with stronger professional identities are more effective at persuading employers, which links the professional identify formation and signaling frameworks (Holmes, 2013). Participation in VIP could be the result of and/or support professional identity development, which would support higher job placement.

Conclusions

The results of the analysis are compelling. When factors tracked by the institution are used to estimate the likelihood of participating in VIP, and when this likelihood and the impact of other programs are controlled for, participation in three semesters of VIP is associated with double to triple the odds of job placement prior to graduation, comparable to odds associated with internships. Propensity score analysis emulates experimental design, removes selection bias, and can be used to draw causal inferences (Guo & Fraser, 2015). The presence of expected correlations for work-based learning, GPA, and SES indicators support the validity of the results. However, the potentially spurious correlation for the freshman experience course in the Pell grant subgroup and the lack of significance of key variables in the female subgroups imply the presence of unmeasured confounders and preclude causal inference. While causation cannot be confirmed, the consistency in results across groups strongly imply that VIP confers benefits to students in accessing the job market.

In addition to supporting student job placement, the program also contributes a small degree of equity through more representative student participation. Low SES students see lower odds of job placement than their more affluent peers; non-white students are overrepresented among low SES students; and both groups participate in programs associated with higher odds of job placement at lower rates than their affluent and white peers. In contrast, non-white students participate in VIP at higher rates than white students, and Pell and transfer students participate in VIP at higher rates than any other program associated with higher job placement. The higher odds of job placement associated with one semester of VIP is approximately equal to the odds of

Pell grant recipients not having found a job, potentially counteracting the job-placement disadvantage associated with lower SES.

Implications

Institutional decisions based on assessment of educational programs should account for program effects and program cost (Mayhew et al., 2016). While causation cannot be confirmed, the consistency in results across subgroups strongly imply that participation in VIP supports student entry into the job market, with a dosage effect for at least the second and third semester. Additionally, low SES and non-white students who participate in programs associated with higher job placement at lower rates than their white and affluent peers participate in VIP at high rates. While CoOp is associated with higher job placement odds, and internships are associated with odds comparable to the odds associated with three semesters of VIP, there are limitations to the degree to which CoOp and internship programs can be further scaled at Georgia Tech and other institutions. Georgia Tech has the largest voluntary CoOp program in the nation, with over 700 employers in the US and abroad (Georgia Institute of Technology, n.d.). Despite the large program, only 16% of students in the sample began the CoOp program, with approximately 8% of all students earning the degree designator. In contrast, the VIP program served 24% of students in the sample with modest institutional investments. The prospect of a program that could be scaled to *benefit all students* requires consideration of **curricula, costs** and **faculty incentives** within institutions, and consideration of **national funding models** to support LT-PBL-EFR across the country.

Costs The direct costs of VIP differ by institutional and department characteristics. Georgia Tech is classified as a Research I institution, an indicator of very high research activity (*Carnegie Classifications*, n.d.). This context enables the studied VIP Program to leverage existing resources – VIP teams are embedded in faculty projects, so funding for team start-up and ongoing support were not provided. The main direct cost of the program is salary funding for program administration, with a cost of approximately $125 per enrollment instance per academic year, which includes fringe (insurance, retirement, etc.). VIP staff handle faculty support, student recruiting, and enrollment management; coordinate policies with academic units; conduct data analysis; and assist with fundraising efforts in instructors' home departments. An indirect program administration cost is course release time for the faculty director, who plays a key role in faculty engagement. While most Georgia Tech VIP teams can leverage existing research resources, institutions and departments with fewer resources around faculty research may need to provide seed money for team start-up and/or ongoing funding for team operations. At Boise State University, colleges provide funding to the VIP Program proportional to enrollments from the colleges, and VIP faculty can request support from this pool of funding as needed (Amoo et al., 2020). At Kennesaw State University, teams are given $5,000 in startup funding (Kennesaw State University, n.d.). With donations, Georgia Tech is piloting a seed funding program as well, to support team establishment in non-research focused departments. Expectations around faculty time and compensation also differ by institution. The Cooper Union's faculty is unionized. When the institution established a VIP Program, faculty time spent on VIP had to be incorporated into faculty contracts and approved by the union (Bringardner, Chao, et al., 2022). The New York University VIP Program was institutionalized when COVID began affecting institutions. Because the NYU faculty were teaching heavier loads than in typical semesters, VIP instructors were given overload pay, and this became the norm for the program (Bringardner, Sonnenberg-Klein, et al., 2022).

Balancing the cost of the program, when incorporated into the curriculum, VIP can reduce the number of sections required in other courses. In Georgia Tech's Computer Science program, students can take a conventional project-based course to fulfill their Junior Design requirement, or they can participate in VIP (the most popular choice among a handful of options). In Spring 2024, VIP enrolled over 1,250 Computer Science students, which reduced the number of conventional junior design courses that needed to be offered.

Curricula VIP enrollment and number of semesters of participation differed substantially by degree program (Figure 12). Enrollment patterns are shaped by policies in each department on how VIP credits count toward degree requirements (Sonnenberg-Klein et al., 2018b). In Computer Science, the major with the highest proportion of students who participated for three semesters, VIP is one of multiple pathways through which students can fulfill a design requirement. In the VIP pathway, students must work with the same team for three semesters, giving students who join a VIP team incentive to persist through the third semester. In Environmental Engineering, one to five credits of VIP can be used as in-major technical electives. The policy does not incentivize multiple semesters of participation, yielding lower enrollment in second and third semesters compared to Computer Science. Through 2019, Computer Engineering and Electrical Engineering used a threshold policy in which credits could be used as free electives if five or fewer were earned; if six were earned, then three could count as free electives, and three could count as in-major electives, which incentivized multiple semesters of participation. In 2020 the curriculum was restructured, with less guidance on how VIP credits count. This led to lower enrollment in second and subsequent semesters of VIP. The school of Aerospace Engineering does not have a policy on how VIP credits count toward degree requirements, yielding the lowest participation among the studied majors.

To maximize job placement gains, institutions and departments should consider how multiple semesters of participation could be incorporated into the curriculum. Job placement odds associated with VIP Participation increased with number of semesters of participation through three semesters, the highest dosage included in the analysis. At the institutional level, some colleges and universities require students to participate in experiential learning, and students are able to choose from a variety of options (University of Georgia, n.d.; Virginia Commonwealth University, n.d.). Treating LT-PBL-EFR as an experiential learning option in these models would provide incentive to participate in LT-PBL-EFR while maintaining student autonomy to choose other pathways, as in the Computer Science degree. To support maximal benefit, a minimum number of semesters would need to be incentivized. At the department level, students are more likely to participate if they can count VIP credits count toward their degree requirements, and they are more likely to participate for multiple semesters if a minimum number of credits are required in order for credits to count in a meaningful way. As the variety of projects in the Georgia Tech program expands, more departments have opted to identify which VIP projects can count toward their degree requirements. This requires coordination between the VIP Program and departments, to keep team listings and department approvals up to date, with information easily accessible to students.

Faculty Incentives Of the thirty-six departments with faculty involved in VIP, only two provide teaching/research release time to instructors for leading VIP teams. These two units account for *one third* of all VIP instructors. In the School of Electrical and Computer Engineering, research-active faculty teach three courses per year, and VIP faculty are released from one of the three courses annually. In the Georgia Tech Research Institute, an applied research unit supported by grants, instructor time spent on VIP is charged to a central account, releasing researchers from their sponsored project activities while working with VIP. As mentioned, NYU pays faculty an overload rate for leading VIP teams, and at Cooper Union, time spent on VIP is incorporated into faculty contracts and approved by the union. If departments want to fully scale the program to serve all interested students, they would need to provide teaching release time or other forms of support to enable and incentivize participation. Because of the proactive policies in the School of Electrical and Computer Engineering and the Georgia Tech Research Institute, Georgia Tech's VIP program is largely ECE and computing-focused. For a balanced program, incentives and/or faculty support would need to be implemented across multiple departments.

National Level At the national level, substantial resources are invested in undergraduate research and workforce development in research settings. The National Science Foundation anticipates awarding $76,370,000 for Research Experiences for Undergraduates (REU) grants annually, with approximately 1,600 supplements to existing NSF grants and support for 180 REU sites (National Science Foundation, n.d.-a). REU sites and supplements involve student stipends, which limits the number of students who can be served. REU sites typically operate for three years and serve 10-20 students per year. If the same duration is assumed for supplements, this yields a cost of $4,300 per student across the REU portfolio. NSF also offers Research Experiences and Mentoring (REM) supplements as add-ons to three categories of existing grants (National Science Foundation, n.d.-b). The grant supplement provides up to $110,000 for six or more mentees (a much higher per-mentee cost), and it also involves student stipends. Because the VIP model is built around academic credit, it is more scalable than the REU and REM models which are limited by stipends. If federal agencies provided grants for VIP sites, and if programs were incentivized to both incorporate VIP into the curriculum and to scale-up, it could have a substantial impact on workforce outcomes and equity.

Directions for Further Research

The framework of the study considers dynamics involved in the transition from college to the workforce, but the study did not directly assess whether signals sent by VIP students differed from those sent by non-VIP students, or how signals differed between students who participated for one, two, and three semesters. The results confirmed that VIP participation is associated with higher odds of job placement. To determine if this difference can be attributed to signaling, researchers could examine signals students send to employers. This could be done by presenting students with hypothetical scenarios, or by observing interactions between students and employers. In the real-life context, researchers could record student-employer interactions at job fairs, analyze resumes, and compare resumes and signals with employer decisions on which candidates to interview. However, this would be intrusive for both students and employers, and interference could negatively affect student success. Observations could be done in a lower-stakes setting, such as mock-interviews facilitated by the career center. A weakness with this approach would be that interviewer decisions would be hypothetical and potentially less realistic. Alternatively, students from a variety of programs could be interviewed or surveyed and asked how they would answer hypothetical job interview questions, and to describe the contexts their responses were related to (club, internship, etc.). This could provide insight into stories students would likely share with employers, to determine how stories differ by student experience, and to what degree signals of affluence, performance of the employer's culture, and connections with the industry differ by student background and experience. Beyond determining if differences in signaling underly the job placement differences, findings could enable career counselors, instructors, and VIP staff to further support students as they craft their pitches to employers.

This study also did not differentiate between CoOp and internships students who accepted positions with their CoOp and internship employers, and students who were seeking jobs from employers they had not worked for. These represent very different dynamics. Accounting for this difference may yield additional insight into the degree to which different experiences support job placement.

Another element worthy of analysis is inequitable participation in programs associated with higher job placement rates. The VIP model was not developed as a diversity, equity and inclusion initiative. However, in stark contrast to other programs in the analysis, the program enrolls higher than expected proportions of URM and Asian students. To study why participation rates differ in VIP, CoOp, Internships, and study abroad, a policy and document analysis could be conducted to identify differences in the messages units communicate to students and how applicants are screened. Understanding why VIP has been successful in this area where others have not could enable institutions to increase equity in participation across all programs, which could in turn support equity in job placement.

REFERENCES

Automatic citation updates are disabled. To see the bibliography, click Refresh in the Zotero tab.

APPENDICES

Appendix A. Descriptive Statistics

Variable Frequencies

Variable			11 Majors Grouping	8 Majors Grouping	5 Majors Grouping	3 majors Grouping
Background						
Resident Non-citizen			288	249	140	109
Female			2127	1714	631	474
Pell Grant Recipient			1381	1162	647	434
Race/	American Indian or Alaska Native			1		
Ethnicity	Asian		1370	1244	726	596
	Black or African American		373	299	171	110
	Hispanic or Latino		482	414	192	122
	Native Hawaiian or Other Pacific Islander		1			
	Two or more		238	200	107	78
	Unknown		128	110	51	45
	White		3224	2662	1199	779
Academics						
Graduation	2017		964	814	368	246
Year	2018		998	821	400	265
	2019		1021	861	431	291
	2020		1010	860	426	318
	2021		852	736	359	255
	2022		972	838	462	355
Freshman Experience Course			2788	2268	778	495
Transfer Exp Course			166	140	64	42
Transfer Student			1261	1068	610	382
Undergraduate Research Course			1713	1305	655	351
VIP	1 Semester		500	444	233	155
	2 Semesters		226	207	88	44
	3 Semesters		346	332	279	262
Student Engagement						
Study Abroad			1987	1668	610	366
Greek Social System			2037	1702	631	453
Living Learning	Global Leadership		19	18	11	9
Communities	Grand Challenges		203	177	96	74
	Honors Program		298	259	133	107
	I House		17	14	5	4
	Ignite Summer Launch		58	54	27	21
	Women in Science and Technology		57	48	13	10
NCAA Athlete			66	58	17	9
Career Preparation						
CoOp	Some CoOp		666	572	197	130
	CoOp Degree Designator		746	645	199	96
Full Time Internship			1573	1322	547	408
Part Time Internship			218	192	79	65
Outcome: Job Placement			4652	3958	1998	1476
Total			5817	4930	2446	1730

Variables as Percentages within Groupings

Variable			11 Majors Grouping	8 Majors Grouping	5 Majors Grouping	3 majors Grouping
Background						
Resident Non-citizen			5	5	6	6
Female			37	35	26	27
Pell Grant Recipient			24	24	26	25
Race/ Ethnicity	American Indian or Alaska Native			<1		
	Asian		24	25	30	34
	Black or African American		6	6	7	6
	Hispanic or Latino		8	8	8	7
	Native Hawaiian or Other Pacific Islander		<1			
	Two or more		4	4	4	5
	Unknown		2	2	2	3
	White		55	54	49	45
Academics						
Graduation Year	2017		17	17	15	14
	2018		17	17	16	15
	2019		18	17	18	17
	2020		17	17	17	18
	2021		15	15	15	15
	2022		17	17	19	21
Freshman Experience Course			48	46	32	29
Transfer Exp Course			3	3	3	2
Transfer Student			22	22	25	22
Undergraduate Research Course			29	26	27	20
VIP	1 Semester		9	9	10	9
	2 Semesters		4	4	4	3
	3 Semesters		6	7	11	15
Student Engagement						
Study Abroad			34	34	25	21
Greek Social System			35	35	26	26
Living Learning Communities	Global Leadership		<1	<1	<1	1
	Grand Challenges		3	4	4	4
	Honors Program		5	5	5	6
	I House		<1	<1	<1	<1
	Ignite Summer Launch		1	1	1	1
	Women in Science and Technology		1	1	1	1
NCAA Athlete			1	1	1	1
Career Preparation						
CoOp: Some CoOp			11	12	8	8
CoOp: CoOp Degree Designator			13	13	8	6
Full Time Internship			27	27	22	24
Part Time Internship			4	4	3	4
Outcome: Job Placement			80	80	82	85

Grade Point Averages

	11 Majors	8 Majors	5 Majors	3 Majors
Min	1.900	1.900	1.900	1.900
1st Quartile	3.308	3.333	3.333	3.366
Median	3.667	3.688	3.714	3.750
Mean	3.567	3.577	3.587	3.62
3rd Quartile	4.000	4.000	4.000	4.000
Max	4.000	4.000	4.000	4.000

First Generation Student Status – Missingness by Year

First Gen Status	Year of Graduation						Total
	2017	2018	2019	2020	2021	2022	
Not First Generation	13	135	626	804	679	773	3030
First Generation	108	122	114	127	111	118	700
Missing	843	741	281	79	62	81	2087
Total	964	998	1021	1010	852	972	5817

Appendix B. Sample RStudio Script

````
---
title: "5 Majors"
author: "jsk"
date: "`r Sys.Date()`"
output:
  html_document:
    df_print: paged
  word_document:
---
```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```
```{r MoreSetup, include=FALSE}
# library(twangContinuous)
# library(twang)
library(knitr)
library(rmarkdown)
library(rstudioapi)
library(readxl)
library(survey)
library(ipw)
library(WeightIt)
library(poliscidata)
library(ggplot2)
library(arm)

## Set the seed of R's pseudo random number generator
## so that the results are replicable.
set.seed(1)

```
```{r ImportData, include=FALSE}

## Read data, convert to a dataframe
Data <- read_excel ("Data.xlsx")
class(Data) <- class(as.data.frame(Data))

Data <- subset(Data, GPA >= 1.9)

# Update groupings in MAJREV
Data$MAJREV[Data$MAJ=="Computer Science"] <- "Computer Science"
Data$MAJREV[Data$MAJ=="Computational Media"] <- "Computational Media"
Data$MAJREV[Data$MAJ=="Electrical Engineering"] <- "Electrical Engineering"
Data$MAJREV[Data$MAJ=="Computer Engineering"] <- "Computer Engineering"
Data$MAJREV[Data$MAJ=="Civil Engineering"] <- "Civil Engineering"
Data$MAJREV[Data$MAJ=="Environmental Engineering"] <- "Environmental Engineering"

# Group Black or African American and Hispanic or Latino together as URM
Data$RCETH[Data$RCETH == "Black or African American"] <- "URM"
Data$RCETH[Data$RCETH == "Hispanic or Latino"] <- "URM"

# Create binary URM variable
Data$URM[Data$RCETH == "URM"] <- 1
Data$URM[Data$RCETH != "URM"] <- 0

# Order CoOp Variables
Data$COOP[Data$COOP == "No CoOp"] <- "0 No CoOp"
Data$COOP[Data$COOP == "Some CoOp"] <- "1 Some CoOp"
Data$COOP[Data$COOP == "CoOp Deg Designator"] <- "3 CoOpDegDesig"

```
# Subset Statements
```{r SubsetStatements, include=TRUE}
Data <- subset(Data, CITZ != "Alien, Non-Resident*")
Data <- subset(Data, VIPSEM %in% c(0,1,2,3)) #TREATMENT LEVELS
````
````

````
Data <- subset(Data, MAJREV %in% c("Computer Science",
"Environmental Engineering",
"Computer Engineering",
"Electrical Engineering",
"Aerospace Engineering"))

Data$MAJREV[Data$MAJREV=="Computer Science"] <- "Computer Sci"
Data$MAJREV[Data$MAJREV=="Environmental Engineering"] <- "0 Env Eng"
Data$MAJREV[Data$MAJREV=="Computer Engineering"] <- "Computer Eng"
Data$MAJREV[Data$MAJREV=="Electrical Engineering"] <- "Electrical Eng"
Data$MAJREV[Data$MAJREV=="Mechanical Engineering"] <- "Mechanical Eng"
Data$MAJREV[Data$MAJREV=="Biomedical Engineering"] <- "Biomedical Eng"
Data$MAJREV[Data$MAJREV=="Materials Science & Engr"] <- "Mat Science & Eng"
Data$MAJREV[Data$MAJREV=="Civil Engineering"] <- "Civil Eng"
Data$MAJREV[Data$MAJREV=="Aerospace Engineering"] <- "Aerospace Eng"
Data$MAJREV[Data$MAJREV=="Industrial Engineering"] <- "Industrial Eng"
Data$MAJREV[Data$MAJREV=="Chemical and Biomolecular Eng"] <- "Chem & Biomolec Eng"

```
# GPA - Center around Grand Mean for 5 Majors Grouping
```{r GrandMeanGPA, include=TRUE}

Data$GPA.raw <- Data$GPA
Data$GPA.from.5maj.Mean <- Data$GPA-3.586502
Data$GPA <- Data$GPA.from.5maj.Mean

```
```{r SubsetStatementsToPull, include=FALSE}

##################################
# Data <- subset(Data, RCETH != "White")
##################################
# Unhide lines below when white is excluded to set ref category to "other or Unknown"
# Hide lines below when white included
#
# Data$RCETH[Data$RCETH == "Two or more"] <- "0 Other or Unknown"
# Data$RCETH[Data$RCETH == "American Indian or Alaska Native"] <- "0 Other or Unknown"
# Data$RCETH[Data$RCETH == "Native Hawaiian or Other Pacific Islander"] <- "0 Other or Unknown"
# Data$RCETH[Data$RCETH == "Unknown"] <- "0 Other or Unknown"

##################################
# COOP AND INT12
##################################
# Data <- subset(Data, COOP == "0 No CoOp")
# Data <- subset(Data, INT12 == 0)
```
# Fequency Tables
```{r FeqTable, echo=FALSE}
freqtable <- table(Data$MAJREV,Data$VIPSEM)
freqtable

```
```{r SampleBarChart, echo=FALSE}

ggplot(Data, aes(x = VIPSEM, fill = MAJ)) + geom_bar()

```
```{r TidyUpVariables, include=FALSE}

# Add leading zeros to reference categories (need to be first alphabetically)
Data$RCETH[Data$RCETH == "White"] <- "0 White"

# Make CITZ more readable
Data$CITZ[Data$CITZ == "US Citizen*"] <- "0 US Citizen"
Data$CITZ[Data$CITZ == "Alien, Resident*"] <- "Resident NonCitizen"
Data$CITZ[Data$CITZ == "Alien, Non-Resident*"] <- "Not a US Citizen or Resident"

# Group Black or African American and Hispanic or Latino as URM
Data$RCETH[Data$RCETH == "Black or African American"] <- "URM"
````

```
Data$RCETH[Data$RCETH == "Hispanic or Latino"] <- "URM"

# Create binary URM variable
Data$URM[Data$RCETH == "URM"] <- 1
Data$URM[Data$RCETH != "URM"] <- 0

# Create binary Asian variable
Data$Asian[Data$RCETH == "Asian"] <- 1
Data$Asian[Data$RCETH == "URM"] <- 0
Data$Asian[Data$RCETH == "URM"] <- 0
Data$Asian[Data$RCETH == "Two or more"] <- 0
Data$Asian[Data$RCETH == "0 White"] <- 0
Data$Asian[Data$RCETH == "Unknown"] <- 0

# Create "other or unknown"
Data$RCETH[Data$RCETH == "Two or more"] <- "Other or Unknown"
Data$RCETH[Data$RCETH == "American Indian or Alaska Native"] <- "Other or Unknown"
Data$RCETH[Data$RCETH == "Native Hawaiian or Other Pacific Islander"] <- "Other or Unknown"
Data$RCETH[Data$RCETH == "Unknown"] <- "Other or Unknown"

` ` `
` ` `{r VarsAsNumeric, include=FALSE}
## convert character and categorical numeric variables to factors

Data$VIPSEM <- as.numeric(Data$VIPSEM)
Data$Female <- as.numeric(Data$Female)
Data$URM <- as.numeric(Data$URM)
Data$GT1OR2 <- as.numeric(Data$GT1OR2)
Data$TRAN <- as.numeric(Data$TRAN)
Data$UROP <- as.numeric(Data$UROP)
Data$STAB <- as.numeric(Data$STAB)
Data$LLGL <- as.numeric(Data$LLGL)
Data$LLGRCH <- as.numeric(Data$LLGRCH)
Data$GRK <- as.numeric(Data$GRK)
Data$LLHON <- as.numeric(Data$LLHON)
Data$LLIH <- as.numeric(Data$LLIH)
Data$LLIG <- as.numeric(Data$LLIG)
Data$ATH <- as.numeric(Data$ATH)
Data$LLWST <- as.numeric(Data$LLWST)
Data$INTANY <- as.numeric(Data$INTANY)
Data$INTNOT12 <- as.numeric(Data$INTNOT12)
Data$INT12 <- as.numeric(Data$INT12)
Data$PELL  <- as.numeric(Data$PELL )

` ` `
` ` `{r VarsAsFactors, include=FALSE}
## convert character and categorical numeric variables (like term code) to factors

# Outcome
Data$EMPSTAT     <- as.factor(Data$EMPSTAT)

# Demographics
Data$CITZ <- as.factor(Data$CITZ)
Data$RCETH       <- as.factor(Data$RCETH)

# Majors
Data$MAJ<- as.factor(Data$MAJ)
Data$MAJREV      <- as.factor(Data$MAJREV)
Data$MAJGRP      <- as.factor(Data$MAJGRP)

# Time-related
Data$TRMTXT              <- as.factor(Data$TRMTXT)
Data$TRMCD       <- as.factor(Data$TRMCD)
Data$YR          <- as.factor(Data$YR)

# $COOP
Data$COOP        <- as.factor(Data$COOP)

` ` `
```

````
```{r IPWnote, include=FALSE}
# Page numbers from Leite book, Practical Propensity Score Methods using R
```
# Propensity Score Model
```{r p135, echo=FALSE}

# variables and plus signs go within single set of quotes
# section beings with "VIPSEM ~ ......"

formulaDose <- formula("VIPSEM ~ CITZ + Female + RCETH + PELL  + TRAN + GRK + STAB        + GT1 + LLHON
        + MAJREV + UROP + COOP + INT12 + GPA")

covariateNames <- c("CITZ", "Female", "RCETH", "PELL ", "TRAN", "GRK", "STAB", "GT1", "LLHON",
        "MAJREV", "UROP", "COOP", "INT12", "GPA")


modelDoses <- lm(formula=formulaDose, data=Data)

Data$GPS <- dnorm(Data$VIPSEM,mean=modelDoses$fitted,sd=sd(Data$VIPSEM))

formulaDose


```
# page 137:
designAN <- svydesign(id=~1, weights=~1,data=Data)

```
```{r p140, include=FALSE}

# page 140:

Data$numerator <- with(Data, dnorm(VIPSEM, mean=mean(VIPSEM), sd=sd(VIPSEM)))

Data$IPW <- with(Data, numerator/GPS)

```
```{r p141, include=FALSE}

# page 141

designIPW <- svydesign(ids=~1,weights=~IPW,data=Data)

balanceTableIPW <- data.frame()

for (var in 1:length(covariateNames)) {
  balanceFormula <-paste("VIPSEM~",covariateNames[var],sep="")
  maxEffBaseline <- max(abs(coef(svyglm(balanceFormula,designAN))[-1]))
  maxEffIPW <- max(abs(coef(svyglm(balanceFormula,designIPW))[-1]))
  balanceTableIPW <- rbind(balanceTableIPW,c(var,maxEffBaseline,maxEffIPW)) }

names(balanceTableIPW) <- c("variable","coefBaseline","coefIPW")

balanceTableIPW$variable <- covariateNames

balanceTableIPW$coefBaseline <- round(balanceTableIPW$coefBaseline/
sqrt(coef(svyvar(~VIPSEM,designAN))),3)

balanceTableIPW$coefIPW <- round(balanceTableIPW$coefIPW/ sqrt(coef(svyvar(~VIPSEM,designIPW))),3)

```
# Balance Table
```{r SaveBalanceTable, echo=FALSE}

# Fix balance table output

library(gt) # format html table

gt(balanceTableIPW) %>%
````

```
  tab_style(
    style = list(cell_text(size = "small", stretch = "ultra-condensed")),
    locations = list(cells_body(), cells_column_labels())
  )

```
```{r p142, include=FALSE}
# page 142 - Regression

```
Sample Sizes
```{r SampleSizes, echo=FALSE}

SampleSize <- data.frame("N" = nrow(Data),"ESS"= round(ESS(Data$IPW),0))

SampleSize

```
Generalized Propensity Scores

```{r GPSsummary, echo=FALSE}

GPS.summary <- round(cbind(summary(Data$GPS[Data$VIPSEM==0]), summary(Data$GPS[Data$VIPSEM==1]), sum-
mary(Data$GPS[Data$VIPSEM==2]), summary(Data$GPS[Data$VIPSEM==3])),3)

colnames(GPS.summary) <- c("0 sem", "1 sem", "2 sem", "3 sem")
kable(GPS.summary)

boxplot(GPS ~ VIPSEM, data = Data, xlab = "Semesters of Participation in VIP", ylab = "Generalized Propen-
sity Scores")

```
Inverse Propensity Score Weights

```{r IPWsummary, echo=FALSE}

summary(Data$IPW)

boxplot(IPW ~ VIPSEM, data = Data, xlab = "Semesters of Participation in VIP", ylab = "Inverse Propensity
Score Weights")


hist(Data$IPW)

Data$Semesters.of.VIP <- as.factor(Data$VIPSEM)
Data$Generalized.Propensity.Scores <- Data$GPS

IPW.summary <- round(cbind(summary(Data$IPW[Data$VIPSEM==0]), summary(Data$IPW[Data$VIPSEM==1]), sum-
mary(Data$IPW[Data$VIPSEM==2]), summary(Data$IPW[Data$VIPSEM==3])),3)

colnames(IPW.summary) <- c("0 sem", "1 sem", "2 sem", "3 sem")
kable(IPW.summary)


```
REGRESSION
```{r RegressionVar, include=FALSE}

OutcomeModel <- svyglm(EMPSTAT ~ VIPSEM  + CITZ  + Female+ RCETH + PELL  + TRAN  + GRK   + STAB  + GT1
        + LLHON         + MAJREV+ UROP          + COOP  + INT12 + GPA + YR, design = designIPW, family =
quasibinomial())


```
Residuals
```{r Residuals, echo=FALSE}

# source on how to pull residuals from the svyglm output: https://websites.umich.edu/~survey-
method/asda/R%20svyglm%20with%20Residuals%20Example%207jan2021.pdf
```

````
summary(residuals(OutcomeModel))

binnedplot(fitted(OutcomeModel),
           residuals(OutcomeModel, type = "response"),
           nclass = NULL,
           xlab = "Expected Values",
           ylab = "Average residual",
           main = "Binned residual plot",
           cex.pts = 0.8,
           col.pts = 1,
           col.int = "gray")


# Assessing model fit
# https://rdrr.io/cran/poliscidata/man/fit.svyglm.html

fit.svyglm(OutcomeModel)

# Can manually calculate McFadden R^2, came out the same as with fit.svyglm.
# 1- OutcomeModel$deviance/OutcomeModel$null.deviance

```


#  Regression Results
```{r RegressionResults, echo=FALSE}
summary(OutcomeModel)

# https://bookdown.org/rwnahhas/RMPH/survey-logistic.html

round(
  cbind(
    summary(OutcomeModel, df.resid=degf(OutcomeModel$survey.design))$coef,
    confint(OutcomeModel, ddf=degf(OutcomeModel$survey.design), level = 0.95)
  )
, 4)

```

# Adjusted Odds Ratios with Confidence Intervals
```{r AdjustedOddsRatiosAndConfidenceIntervals, echo=FALSE}
# https://bookdown.org/rwnahhas/RMPH/survey-logistic.html

OR.CI <- cbind("AOR" = exp(coef(OutcomeModel)),
                       exp(confint(OutcomeModel,
                           df.resid=degf(OutcomeModel$survey.design))))[-1,]
round(OR.CI, 3)

```
````
````

Appendix C. R Output

3 Majors

jsk

Subset Statements

```
Data <- subset(Data, CITZ != "Alien, Non-Resident*")
Data <- subset(Data, VIPSEM %in% c(0,1,2,3)) #TREATMENT LEVELS

Data <- subset(Data, MAJREV %in% c("Computer Science",
"Environmental Engineering",
"Computer Engineering"))

# Shorten Major Names after subsetting by major

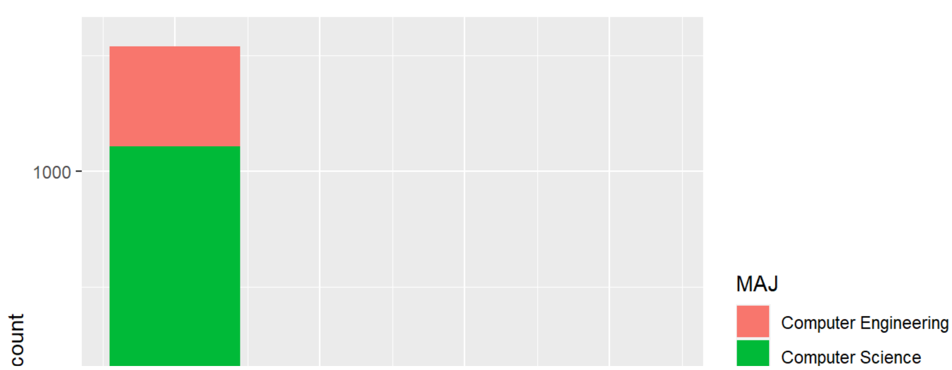
Data$MAJREV[Data$MAJREV=="Computer Science"] <- "Computer Sci"
Data$MAJREV[Data$MAJREV=="Environmental Engineering"] <- "0 Env Eng"
Data$MAJREV[Data$MAJREV=="Computer Engineering"] <- "Computer Eng"
Data$MAJREV[Data$MAJREV=="Electrical Engineering"] <- "Electrical Eng"
Data$MAJREV[Data$MAJREV=="Mechanical Engineering"] <- "Mechanical Eng"
Data$MAJREV[Data$MAJREV=="Biomedical Engineering"] <- "Biomedical Eng"
Data$MAJREV[Data$MAJREV=="Materials Science & Engr"] <- "Mat Science & Eng"
Data$MAJREV[Data$MAJREV=="Civil Engineering"] <- "Civil Eng"
Data$MAJREV[Data$MAJREV=="Aerospace Engineering"] <- "Aerospace Eng"
Data$MAJREV[Data$MAJREV=="Industrial Engineering"] <- "Industrial Eng"
Data$MAJREV[Data$MAJREV=="Chemical and Biomolecular Eng"] <- "Chem & Biomolec Eng"
```

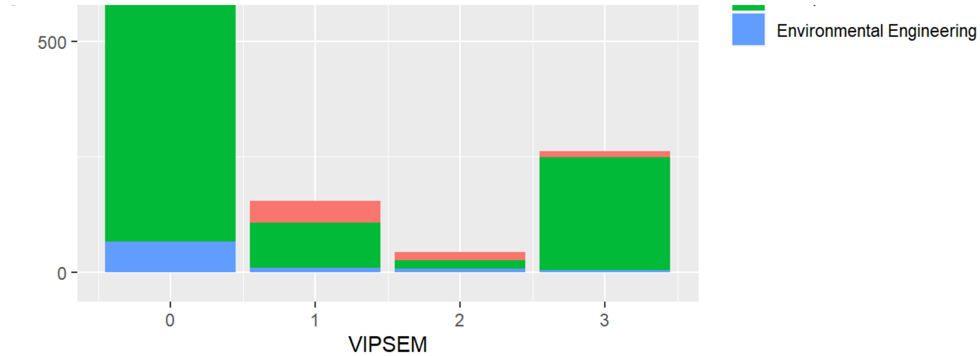
GPA - Center around Mean for 5 Majors Grouping

```
Data$GPA.raw <- Data$GPA
Data$GPA.from.5maj.Mean <- Data$GPA-3.586502
Data$GPA <- Data$GPA.from.5maj.Mean
```

Fequency Tables

```
##
##                  0   1   2   3
##   0 Env Eng     68  11   9   5
##   Computer Eng 216  47  17  13
##   Computer Sci 985  97  18 244
```





Propensity Score Model

```
## VIPSEM ~ CITZ + FEMALE + RCETH + PELL + TRAN + GRK + STAB + GT1 +
##     LLHON + MAJREV + UROP + COOP + INT12 + GPA
```

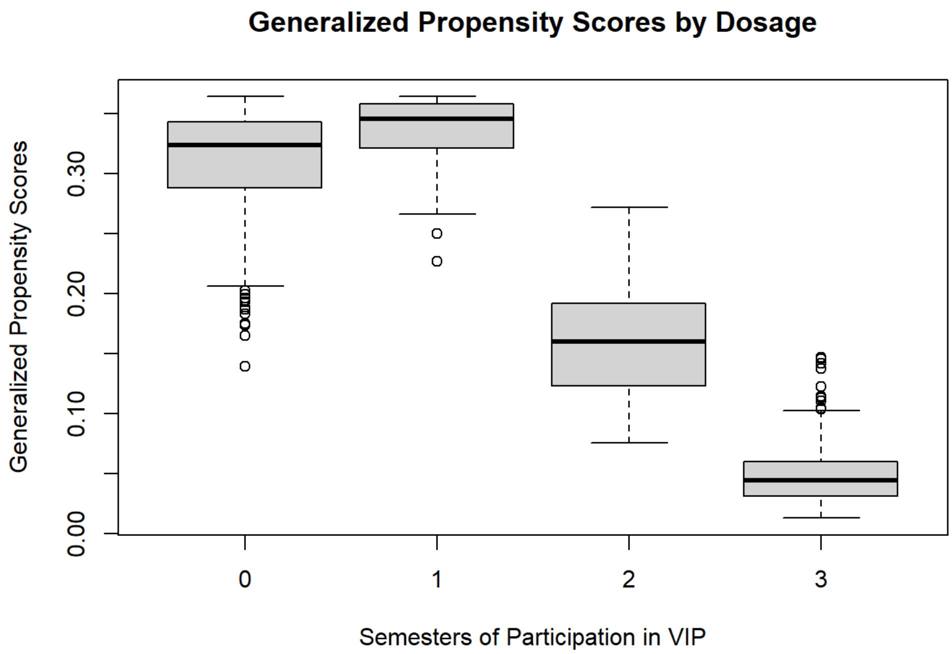
Balance Table

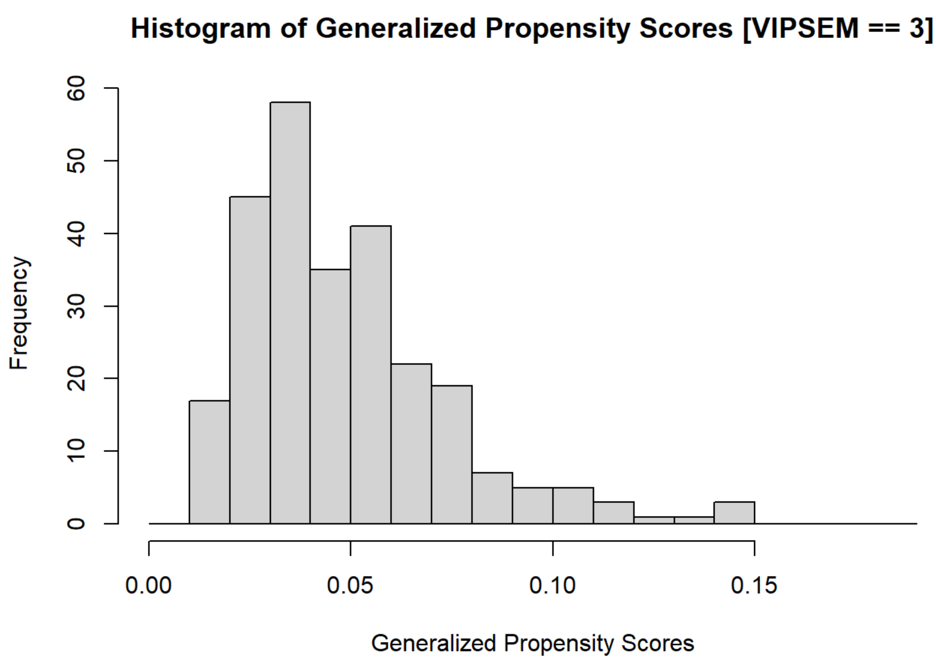
variable	coefBaseline	coefIPW
CITZ	0.001	0.021
FEMALE	0.207	0.015
RCETH	0.354	0.048
PELL	0.090	0.019
TRAN	0.093	0.014
GRK	0.031	0.022
STAB	0.010	0.012
GT1	0.012	0.005
LLHON	0.376	0.035
MAJREV	0.156	0.095
UROP	0.011	0.023
COOP	0.211	0.092
INT12	0.241	0.025
GPA	0.258	0.050

Sample Sizes

Generalized Propensity Scores

	0 sem	1 sem	2 sem	3 sem
Min.	0.140	0.227	0.076	0.013
1st Qu.	0.288	0.321	0.123	0.031
Median	0.324	0.346	0.160	0.045
Mean	0.314	0.337	0.161	0.048
3rd Qu.	0.343	0.358	0.191	0.060
Max.	0.364	0.364	0.272	0.147



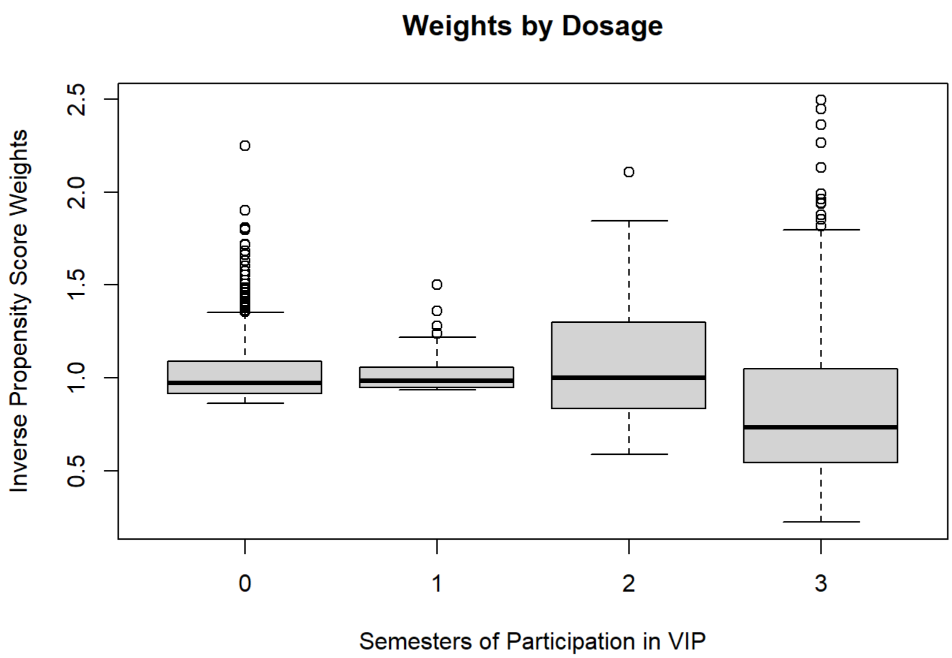


Inverse Propensity Score Weights

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.2230  0.9037  0.9630  0.9974  1.0858  2.4938
```

	0 sem	1 sem	2 sem	3 sem
Min.	0.863	0.934	0.589	0.223
1st Qu.	0.916	0.949	0.837	0.546
Median	0.970	0.983	1.001	0.732
Mean	1.021	1.017	1.087	0.858
3rd Qu.	1.090	1.058	1.298	1.047

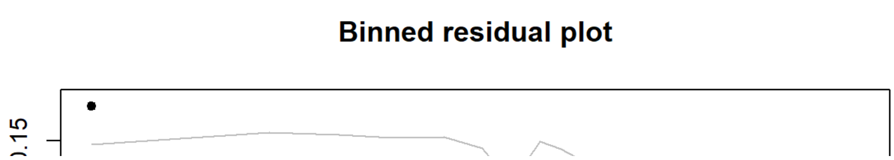
	0 sem	1 sem	2 sem	3 sem
Max.	2.249	1.500	2.106	2.494

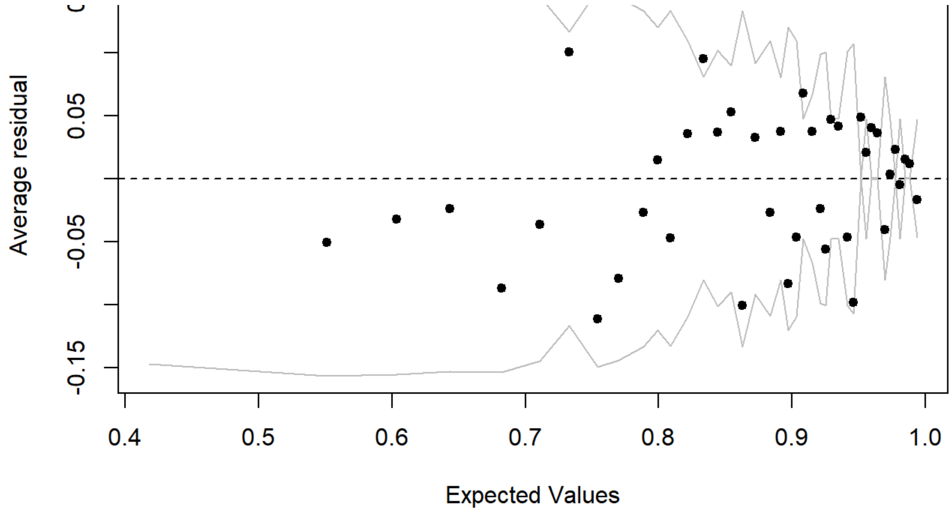


REGRESSION

Residuals

```
##     Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -3.2673  0.1967  0.3951  0.1562  0.5907  1.6936
```






```
##              R-Squared      Adjusted R-Squared
##                  0.146                   0.134
```

Regression Results

```
##
## Call:
## svyglm(formula = ..1, design = ..2, family = ..3)
##
## Survey design:
## survey::svydesign(...)
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)              0.043643   0.417922   0.104 0.916842
## VIPSEM                   0.381982   0.102737   3.718 0.000207 ***
## CITZResident NonCitizen  0.173707   0.299937   0.579 0.562568
## FEMALE                   0.442879   0.204943   2.161 0.030835 *
## RCETHAsian               0.243300   0.186103   1.307 0.191275
## RCETHOther or Unknown    0.348979   0.328144   1.063 0.287710
## RCETHURM                 0.452832   0.240077   1.886 0.059439 .
## PELL                    -0.423571   0.163863  -2.585 0.009823 **
## TRAN                    -0.271640   0.191495  -1.419 0.156221
## GRK                      0.816083   0.219324   3.721 0.000205 ***
## STAB                     0.533688   0.238683   2.236 0.025483 *
## GT1                      0.007406   0.205457   0.036 0.971250
## LLHON                    0.508858   0.441476   1.153 0.249224
## MAJREVComputer Eng       1.049602   0.363200   2.890 0.003903 **
## MAJREVComputer Sci       1.248094   0.339020   3.681 0.000239 ***
## UROP                    -0.108847   0.204455  -0.532 0.594537
## COOP1 Some CoOp          0.109667   0.310005   0.354 0.723563
## COOP3 CoOpDegDesig       1.357665   0.480102   2.828 0.004741 **
## INT12                    1.030875   0.228791   4.506 7.06e-06 ***
## GPA                      1.069159   0.155717   6.866 9.21e-12 ***
## YR2018                   0.517848   0.274584   1.886 0.059473 .
## YR2019                   0.075809   0.268370   0.282 0.777610
```

```
## YR2020                   -0.364390   0.251593  -1.448 0.147708
## YR2021                   -0.091451   0.274153  -0.334 0.738740
## YR2022                   -0.312933   0.251944  -1.242 0.214380
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 1.114832)
##
## Number of Fisher Scoring iterations: 6
```

```
##                          Estimate Std. Error t value Pr(>|t|)   2.5 %  97.5 %
## (Intercept)                0.0436     0.4179  0.1044   0.9168 -0.7760  0.8633
## VIPSEM                     0.3820     0.1027  3.7180   0.0002  0.1805  0.5835
## CITZResident NonCitizen    0.1737     0.2999  0.5791   0.5626 -0.4146  0.7620
## FEMALE                     0.4429     0.2049  2.1610   0.0308  0.0409  0.8448
## RCETHAsian                 0.2433     0.1861  1.3073   0.1913 -0.1217  0.6083
## RCETHOther or Unknown      0.3490     0.3281  1.0635   0.2877 -0.2946  0.9926
## RCETHURM                   0.4528     0.2401  1.8862   0.0594 -0.0180  0.9237
## PELL                      -0.4236     0.1639 -2.5849   0.0098 -0.7450 -0.1022
## TRAN                      -0.2716     0.1915 -1.4185   0.1562 -0.6472  0.1039
## GRK                        0.8161     0.2193  3.7209   0.0002  0.3859  1.2463
## STAB                       0.5337     0.2387  2.2360   0.0255  0.0656  1.0018
## GT1                        0.0074     0.2055  0.0360   0.9713 -0.3956  0.4104
## LLHON                      0.5089     0.4415  1.1526   0.2492 -0.3570  1.3747
## MAJREVComputer Eng         1.0496     0.3632  2.8899   0.0039  0.3372  1.7620
## MAJREVComputer Sci         1.2481     0.3390  3.6815   0.0002  0.5832  1.9130
## UROP                      -0.1088     0.2045 -0.5324   0.5945 -0.5099  0.2922
## COOP1 Some CoOp            0.1097     0.3100  0.3538   0.7236 -0.4984  0.7177
## COOP3 CoOpDegDesig         1.3577     0.4801  2.8279   0.0047  0.4160  2.2993
## INT12                      1.0309     0.2288  4.5058   0.0000  0.5821  1.4796
## GPA                        1.0692     0.1557  6.8660   0.0000  0.7637  1.3746
## YR2018                     0.5178     0.2746  1.8859   0.0595 -0.0207  1.0564
## YR2019                     0.0758     0.2684  0.2825   0.7776 -0.4506  0.6022
## YR2020                    -0.3644     0.2516 -1.4483   0.1477 -0.8578  0.1291
## YR2021                    -0.0915     0.2742 -0.3336   0.7387 -0.6292  0.4463
## YR2022                    -0.3129     0.2519 -1.2421   0.2144 -0.8071  0.1812
```

Adjusted Odds Ratios with Confidence Intervals

```
##                           AOR 2.5 % 97.5 %
## VIPSEM                  1.465 1.198  1.792
## CITZResident NonCitizen 1.190 0.661  2.143
## FEMALE                  1.557 1.042  2.328
## RCETHAsian              1.275 0.885  1.837
## RCETHOther or Unknown   1.418 0.745  2.698
## RCETHURM                1.573 0.982  2.519
## PELL                    0.655 0.475  0.903
## TRAN                    0.762 0.523  1.110
## GRK                     2.262 1.471  3.477
## STAB                    1.705 1.068  2.723
## GT1                     1.007 0.673  1.507
```

```
## LLHON                    1.663 0.700  3.954
## MAJREVComputer Eng       2.857 1.401  5.824
## MAJREVComputer Sci       3.484 1.792  6.774
## UROP                     0.897 0.601  1.339
## COOP1 Some CoOp          1.116 0.608  2.050
## COOP3 CoOpDegDesig       3.887 1.516  9.967
## INT12                    2.804 1.790  4.391
## GPA                      2.913 2.146  3.953
## YR2018                   1.678 0.980  2.876
## YR2019                   1.079 0.637  1.826
## YR2020                   0.695 0.424  1.138
## YR2021                   0.913 0.533  1.562
## YR2022                   0.731 0.446  1.199
```

5 Majors

jsk

Subset Statements

```
Data <- subset(Data, CITZ != "Alien, Non-Resident*")
Data <- subset(Data, VIPSEM %in% c(0,1,2,3)) #TREATMENT LEVELS

Data <- subset(Data, MAJREV %in% c("Computer Science",
"Environmental Engineering",
"Computer Engineering",
"Electrical Engineering",
"Aerospace Engineering"))

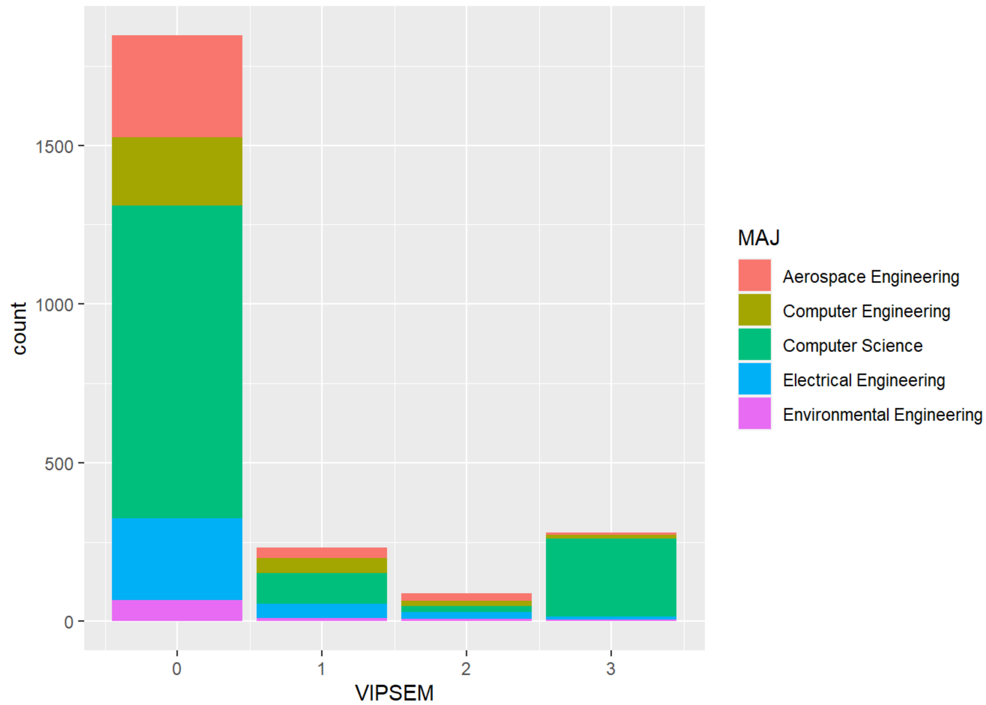
Data$MAJREV[Data$MAJREV=="Computer Science"] <- "Computer Sci"
Data$MAJREV[Data$MAJREV=="Environmental Engineering"] <- "0 Env Eng"
Data$MAJREV[Data$MAJREV=="Computer Engineering"] <- "Computer Eng"
Data$MAJREV[Data$MAJREV=="Electrical Engineering"] <- "Electrical Eng"
Data$MAJREV[Data$MAJREV=="Mechanical Engineering"] <- "Mechanical Eng"
Data$MAJREV[Data$MAJREV=="Biomedical Engineering"] <- "Biomedical Eng"
Data$MAJREV[Data$MAJREV=="Materials Science & Engr"] <- "Mat Science & Eng"
Data$MAJREV[Data$MAJREV=="Civil Engineering"] <- "Civil Eng"
Data$MAJREV[Data$MAJREV=="Aerospace Engineering"] <- "Aerospace Eng"
Data$MAJREV[Data$MAJREV=="Industrial Engineering"] <- "Industrial Eng"
Data$MAJREV[Data$MAJREV=="Chemical and Biomolecular Eng"] <- "Chem & Biomolec Eng"
```

GPA - Center around Mean for 5 Majors Grouping

```
Data$GPA.raw <- Data$GPA
Data$GPA.from.5maj.Mean <- Data$GPA-3.586502
Data$GPA <- Data$GPA.from.5maj.Mean
```

Fequency Tables

```
##
##                  0   1   2   3
##   0 Env Eng      68  11   9   5
##   Aerospace Eng  320  34  24   6
##   Computer Eng   216  47  17  13
##   Computer Sci   985  97  18 244
##   Electrical Eng 257  44  20  11
```



Propensity Score Model

```
## VIPSEM ~ CITZ + FEMALE + RCETH + PELL + TRAN + GRK + STAB + GT1 +
##     LLHON + MAJREV + UROP + COOP + INT12 + GPA
```

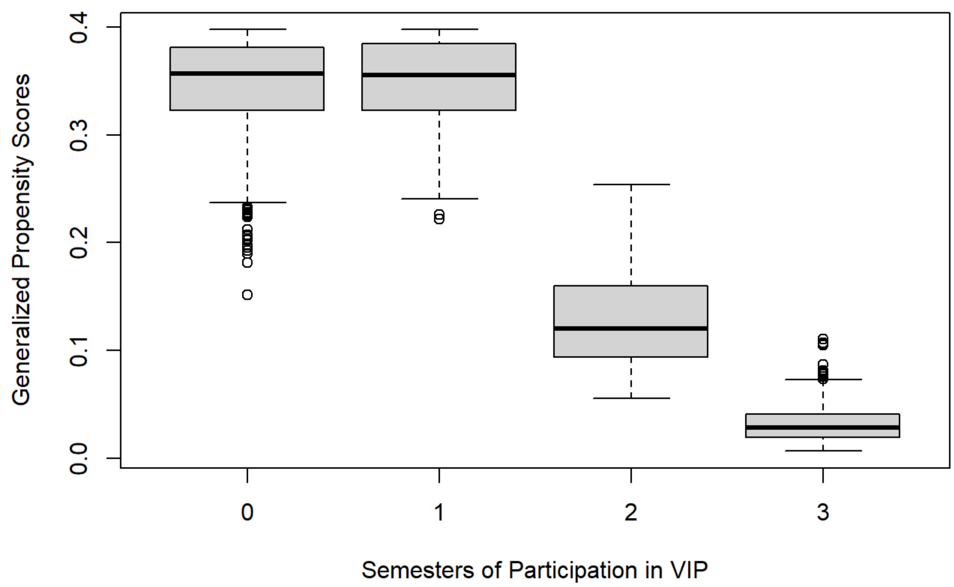
Balance Table

variable	coefBaseline	coefIPW
CITZ	0.020	0.047
FEMALE	0.229	0.020
RCETH	0.375	0.058
PELL	0.089	0.013
TRAN	0.128	0.020
GRK	0.035	0.002
STAB	0.010	0.010
GT1	0.002	0.006
LLHON	0.390	0.031
MAJREV	0.212	0.163
UROP	0.082	0.009
COOP	0.258	0.127
INT12	0.198	0.019
GPA	0.233	0.032

Sample Sizes

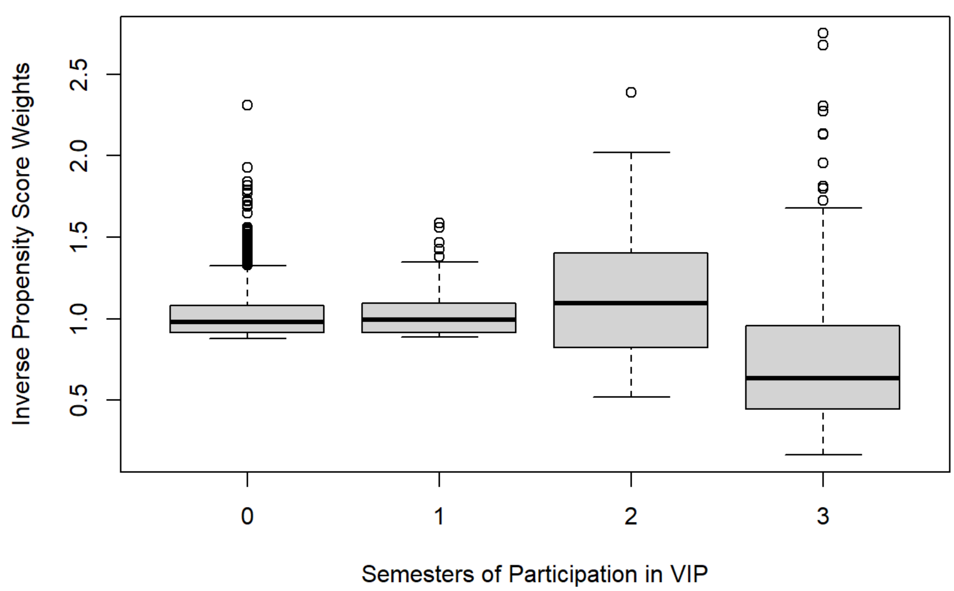
Generalized Propensity Scores

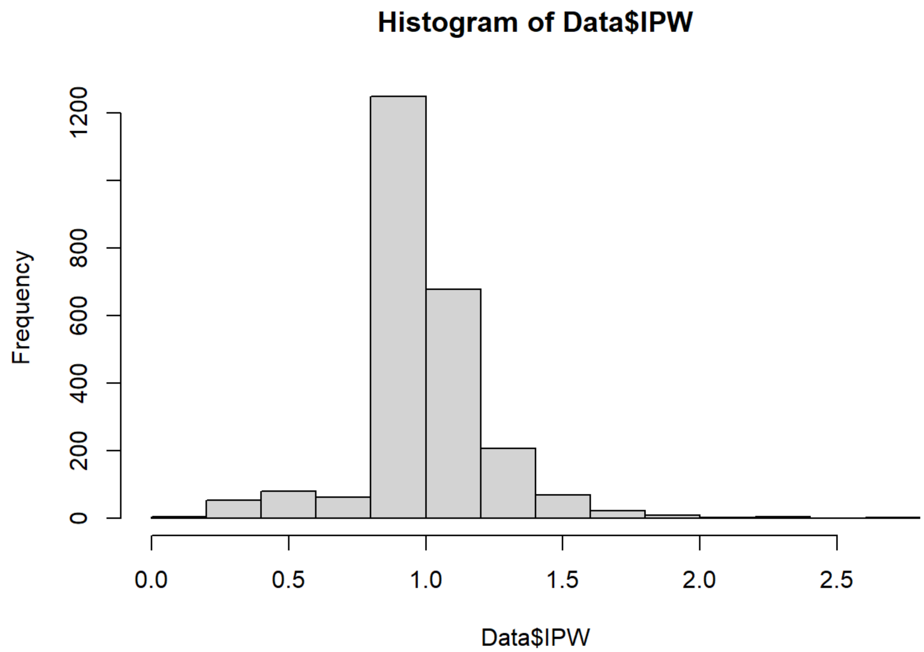
	0 sem	1 sem	2 sem	3 sem
Min.	0.151	0.222	0.055	0.007
1st Qu.	0.323	0.323	0.094	0.019
Median	0.357	0.355	0.120	0.029
Mean	0.347	0.349	0.130	0.033
3rd Qu.	0.381	0.385	0.160	0.041
Max.	0.398	0.398	0.254	0.110



Inverse Propensity Score Weights

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
## 0.1653  0.9033  0.9709  0.9983  1.0824  2.7500
```



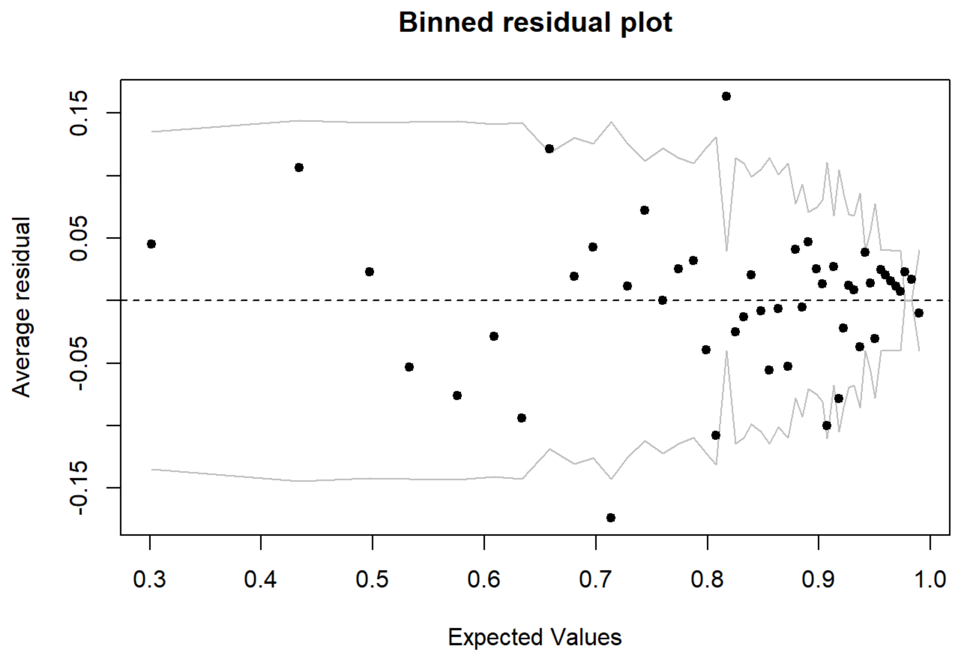


	0 sem	1 sem	2 sem	3 sem
Min.	0.879	0.887	0.519	0.165
1st Qu.	0.917	0.917	0.824	0.448
Median	0.979	0.993	1.095	0.637
Mean	1.026	1.027	1.148	0.746
3rd Qu.	1.082	1.094	1.399	0.956
Max.	2.309	1.588	2.387	2.750

REGRESSION

Residuals

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -2.8910  0.2042  0.4323  0.1499  0.6429  1.8063
```




```
##              R-Squared      Adjusted R-Squared
##                  0.152                   0.143
```

Regression Results

```
##
## Call:
## svyglm(formula = ..1, design = ..2, family = ..3)
##
## Survey design:
## survey::svydesign(...)
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             -0.019709   0.364025  -0.054 0.956828
## VIPSEM                   0.286429   0.081654   3.508 0.000460 ***
## CITZResident NonCitizen -0.004696   0.237800  -0.020 0.984246
## FEMALE                   0.497013   0.160455   3.098 0.001974 **
## RCETHAsian               0.102595   0.150840   0.680 0.496470
## RCETHOther or Unknown    0.212972   0.258535   0.824 0.410156
## RCETHURM                 0.195256   0.172695   1.131 0.258316
## PELL                    -0.445836   0.130803  -3.408 0.000664 ***
## TRAN                    -0.187817   0.153143  -1.226 0.220162
## GRK                      0.736336   0.167588   4.394 1.16e-05 ***
## STAB                     0.237254   0.154974   1.531 0.125918
## GT1                      0.001941   0.155363   0.012 0.990031
## LLHON                    0.535886   0.362973   1.476 0.139973
## MAJREVAerospace Eng      0.187702   0.314578   0.597 0.550778
## MAJREVComputer Eng       1.165074   0.340321   3.423 0.000629 ***
## MAJREVComputer Sci       1.347925   0.312084   4.319 1.63e-05 ***
## MAJREVElectrical Eng     0.862982   0.335807   2.570 0.010233 *
## UROP                     0.100636   0.152825   0.659 0.510276
## COOP1 Some CoOp          0.241687   0.242174   0.998 0.318384
## COOP3 CoOpDegDesig       1.152514   0.276182   4.173 3.11e-05 ***
## INT12                    0.981493   0.180490   5.438 5.93e-08 ***
## GPA                      0.957364   0.123362   7.761 1.24e-14 ***
## YR2018                   0.395920   0.200325   1.976 0.048224 *
## YR2019                   0.058386   0.196990   0.296 0.766959
## YR2020                  -0.326352   0.199313  -1.637 0.101680
## YR2021                   0.248670   0.218237   1.139 0.254629
## YR2022                  -0.121511   0.201140  -0.604 0.545826
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 0.989541)
##
## Number of Fisher Scoring iterations: 5
```

```
##                          Estimate Std. Error t value Pr(>|t|)   2.5 %  97.5 %
```

```
## (Intercept)              -0.0197     0.3640 -0.0541   0.9568 -0.7335  0.6941
## VIPSEM                    0.2864     0.0817  3.5078   0.0005  0.1263  0.4465
## CITZResident NonCitizen  -0.0047     0.2378 -0.0197   0.9842 -0.4710  0.4616
## FEMALE                    0.4970     0.1605  3.0975   0.0020  0.1824  0.8117
## RCETHAsian                0.1026     0.1508  0.6802   0.4965 -0.1932  0.3984
## RCETHOther or Unknown     0.2130     0.2585  0.8238   0.4102 -0.2940  0.7199
## RCETHURM                  0.1953     0.1727  1.1306   0.2583 -0.1434  0.5339
## PELL                     -0.4458     0.1308 -3.4084   0.0007 -0.7023 -0.1893
## TRAN                     -0.1878     0.1531 -1.2264   0.2202 -0.4881  0.1125
## GRK                       0.7363     0.1676  4.3937   0.0000  0.4077  1.0650
## STAB                      0.2373     0.1550  1.5309   0.1259 -0.0666  0.5411
## GT1                       0.0019     0.1554  0.0125   0.9900 -0.3027  0.3066
## LLHON                     0.5359     0.3630  1.4764   0.1400 -0.1759  1.2477
## MAJREVAerospace Eng       0.1877     0.3146  0.5967   0.5508 -0.4292  0.8046
## MAJREVComputer Eng        1.1651     0.3403  3.4235   0.0006  0.4977  1.8324
## MAJREVComputer Sci        1.3479     0.3121  4.3191   0.0000  0.7359  1.9599
## MAJREVElectrical Eng      0.8630     0.3358  2.5699   0.0102  0.2045  1.5215
## UROP                      0.1006     0.1528  0.6585   0.5103 -0.1990  0.4003
## COOP1 Some CoOp           0.2417     0.2422  0.9980   0.3184 -0.2332  0.7166
## COOP3 CoOpDegDesig        1.1525     0.2762  4.1730   0.0000  0.6109  1.6941
## INT12                     0.9815     0.1805  5.4379   0.0000  0.6276  1.3354
## GPA                       0.9574     0.1234  7.7606   0.0000  0.7155  1.1993
## YR2018                    0.3959     0.2003  1.9764   0.0482  0.0031  0.7887
## YR2019                    0.0584     0.1970  0.2964   0.7670 -0.3279  0.4447
## YR2020                   -0.3264     0.1993 -1.6374   0.1017 -0.7172  0.0645
## YR2021                    0.2487     0.2182  1.1394   0.2546 -0.1793  0.6766
## YR2022                   -0.1215     0.2011 -0.6041   0.5458 -0.5159  0.2729
```

Adjusted Odds Ratios with Confidence Intervals

```
##                           AOR 2.5 % 97.5 %
## VIPSEM                  1.332 1.135  1.563
## CITZResident NonCitizen 0.995 0.624  1.587
## FEMALE                  1.644 1.200  2.252
## RCETHAsian              1.108 0.824  1.489
## RCETHOther or Unknown   1.237 0.745  2.054
## RCETHURM                1.216 0.866  1.706
## PELL                    0.640 0.495  0.828
## TRAN                    0.829 0.614  1.119
## GRK                     2.088 1.503  2.901
## STAB                    1.268 0.936  1.718
## GT1                     1.002 0.739  1.359
## LLHON                   1.709 0.839  3.482
## MAJREVAerospace Eng     1.206 0.651  2.236
## MAJREVComputer Eng      3.206 1.645  6.249
## MAJREVComputer Sci      3.849 2.087  7.099
## MAJREVElectrical Eng    2.370 1.227  4.579
## UROP                    1.106 0.820  1.492
## COOP1 Some CoOp         1.273 0.792  2.047
## COOP3 CoOpDegDesig      3.166 1.842  5.442
## INT12                   2.668 1.873  3.802
## GPA                     2.605 2.045  3.318
## YR2018                  1.486 1.003  2.201
## YR2019                  1.060 0.720  1.560
## YR2020                  0.722 0.488  1.067
## YR2021                  1.282 0.836  1.967
## YR2022                  0.886 0.597  1.314
```

8 Majors

jsk

Subset Statements

```
Data <- subset(Data, CITZ != "Alien, Non-Resident*")
Data <- subset(Data, VIPSEM %in% c(0,1,2,3)) #TREATMENT LEVELS

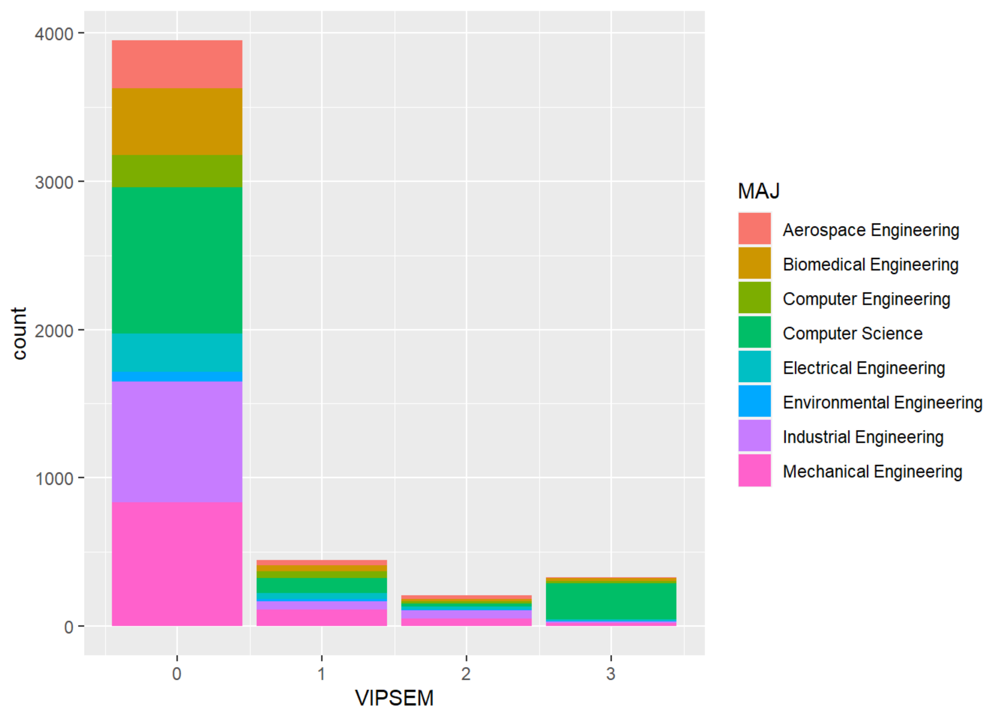
Data <- subset(Data, MAJREV %in% c("Computer Science",
"Environmental Engineering",
"Computer Engineering",
"Electrical Engineering",
"Aerospace Engineering",
"Mechanical Engineering",
"Biomedical Engineering",
"Industrial Engineering"))

# Shorten Major Names after subsetting by major

Data$MAJREV[Data$MAJREV=="Computer Science"] <- "Computer Sci"
Data$MAJREV[Data$MAJREV=="Environmental Engineering"] <- "0 Env Eng"
Data$MAJREV[Data$MAJREV=="Computer Engineering"] <- "Computer Eng"
Data$MAJREV[Data$MAJREV=="Electrical Engineering"] <- "Electrical Eng"
Data$MAJREV[Data$MAJREV=="Mechanical Engineering"] <- "Mechanical Eng"
Data$MAJREV[Data$MAJREV=="Biomedical Engineering"] <- "Biomedical Eng"
Data$MAJREV[Data$MAJREV=="Materials Science & Engr"] <- "Mat Science & Eng"
Data$MAJREV[Data$MAJREV=="Civil Engineering"] <- "Civil Eng"
Data$MAJREV[Data$MAJREV=="Aerospace Engineering"] <- "Aerospace Eng"
Data$MAJREV[Data$MAJREV=="Industrial Engineering"] <- "Industrial Eng"
Data$MAJREV[Data$MAJREV=="Chemical and Biomolecular Eng"] <- "Chem & Biomolec Eng"
```

Fequency Tables

```
##
##                    0   1   2   3
##   0 Env Eng       68  11   9   5
##   Aerospace Eng  320  34  24   6
##   Biomedical Eng 453  40  13  23
##   Computer Eng   216  47  17  13
##   Computer Sci   985  97  18 244
##   Electrical Eng 257  44  20  11
##   Industrial Eng 813  57  54   6
##   Mechanical Eng 835 114  52  24
```



Propensity Score Model

```
## VIPSEM ~ CITZ + FEMALE + RCETH + PELL + TRAN + GRK + STAB + GT1 +
##     LLHON + MAJREV + UROP + COOP + INT12 + GPA
```

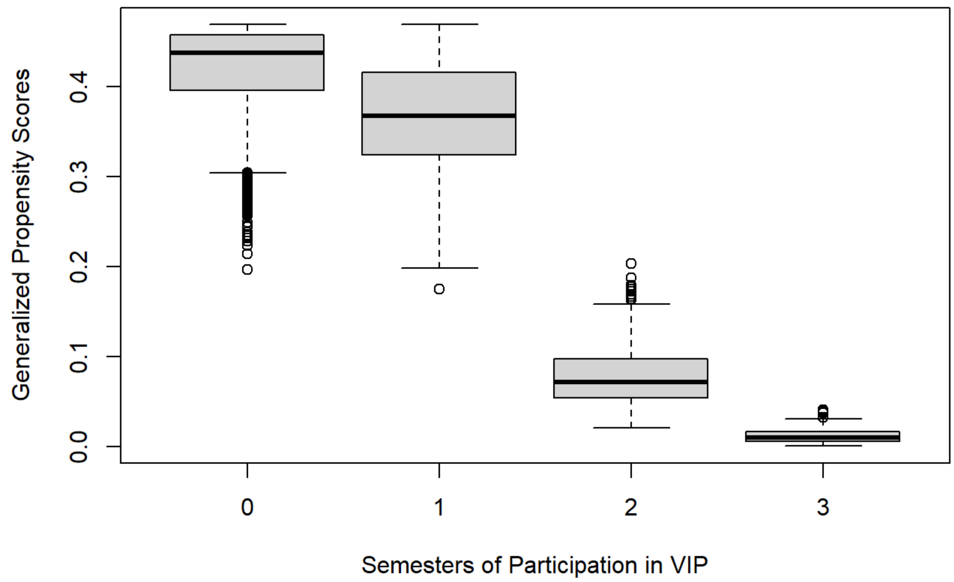
Balance Table

variable	coefBaseline	coefIPW
CITZ	0.071	0.091
FEMALE	0.049	0.003
RCETH	0.386	0.051
PELL	0.034	0.050
TRAN	0.052	0.022
GRK	0.126	0.020
STAB	0.057	0.026
GT1	0.104	0.001
LLHON	0.271	0.029
MAJREV	0.325	0.154
UROP	0.033	0.026
COOP	0.283	0.127
INT12	0.109	0.002
GPA	0.186	0.025

Sample Sizes

Generalized Propensity Scores

	0 sem	1 sem	2 sem	3 sem
Min.	0.197	0.175	0.021	0.001
1st Qu.	0.396	0.324	0.054	0.006
Median	0.438	0.367	0.071	0.009
Mean	0.420	0.368	0.079	0.012
3rd Qu.	0.458	0.416	0.098	0.016
Max.	0.469	0.469	0.204	0.041




```
## [1] 1.12686

## [1] 0.7959824
```

11 Majors

jsk

Subset Statements

```
Data <- subset(Data, CITZ != "Alien, Non-Resident*")
Data <- subset(Data, VIPSEM %in% c(0,1,2,3)) #TREATMENT LEVELS

Data <- subset(Data, MAJREV %in% c("Computer Science",
"Environmental Engineering",
"Computer Engineering",
"Electrical Engineering",
"Aerospace Engineering",
"Mechanical Engineering",
"Biomedical Engineering",
"Industrial Engineering",
"Materials Science & Engr",
"Civil Engineering",
"Chemical and Biomolecular Eng"))

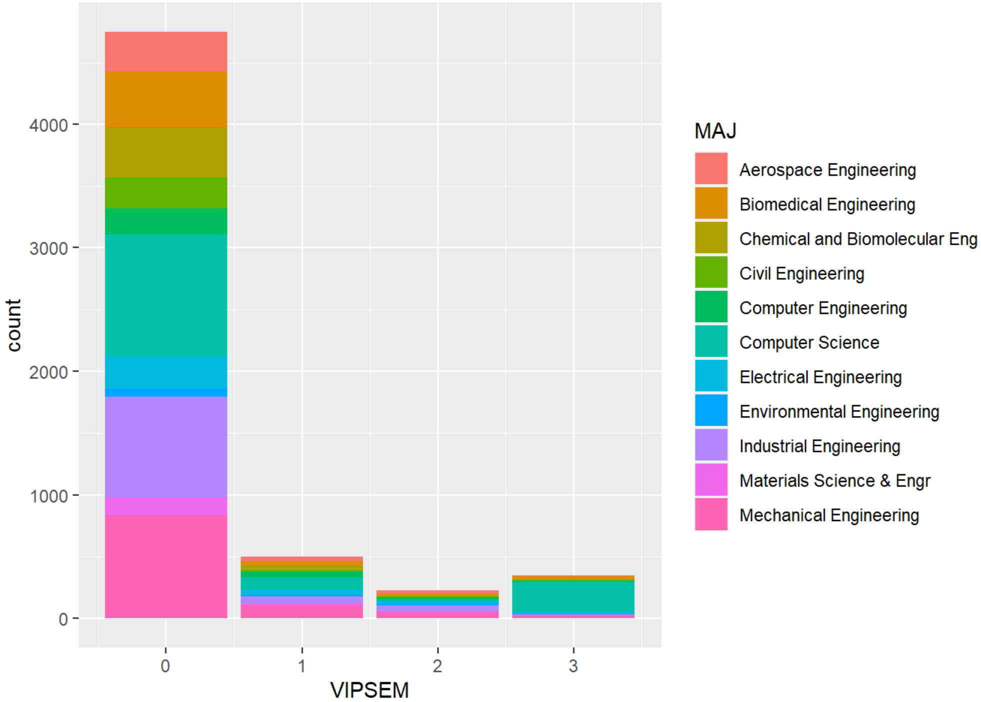
# Shorten Major Names after subsetting by major

Data$MAJREV[Data$MAJREV=="Computer Science"] <- "Computer Sci"
Data$MAJREV[Data$MAJREV=="Environmental Engineering"] <- "0 Env Eng"
Data$MAJREV[Data$MAJREV=="Computer Engineering"] <- "Computer Eng"
Data$MAJREV[Data$MAJREV=="Electrical Engineering"] <- "Electrical Eng"
Data$MAJREV[Data$MAJREV=="Mechanical Engineering"] <- "Mechanical Eng"
Data$MAJREV[Data$MAJREV=="Biomedical Engineering"] <- "Biomedical Eng"
Data$MAJREV[Data$MAJREV=="Materials Science & Engr"] <- "Mat Science & Eng"
Data$MAJREV[Data$MAJREV=="Civil Engineering"] <- "Civil Eng"
Data$MAJREV[Data$MAJREV=="Aerospace Engineering"] <- "Aerospace Eng"
Data$MAJREV[Data$MAJREV=="Industrial Engineering"] <- "Industrial Eng"
Data$MAJREV[Data$MAJREV=="Chemical and Biomolecular Eng"] <- "Chem & Biomolec Eng"
```

Fequency Tables

```
##
##                        0   1   2   3
##    0 Env Eng          68  11   9   5
##    Aerospace Eng     320  34  24   6
##    Biomedical Eng    453  40  13  23
##    Chem & Biomolec Eng 406  33   8   7
##    Civil Eng         245  17   9   2
##    Computer Eng      216  47  17  13
##    Computer Sci      985  97  18 244
##    Electrical Eng    257  44  20  11
##    Industrial Eng    813  57  54   6
##    Mat Science & Eng 147   6   2   5
```

```
##    Mechanical Eng       835 114  52  24
```



Propensity Score Model

```
## VIPSEM ~ CITZ + FEMALE + RCETH + PELL + TRAN + GRK + STAB + GT1 +
##     LLHON + MAJREV + UROP + COOP + INT12 + GPA
```

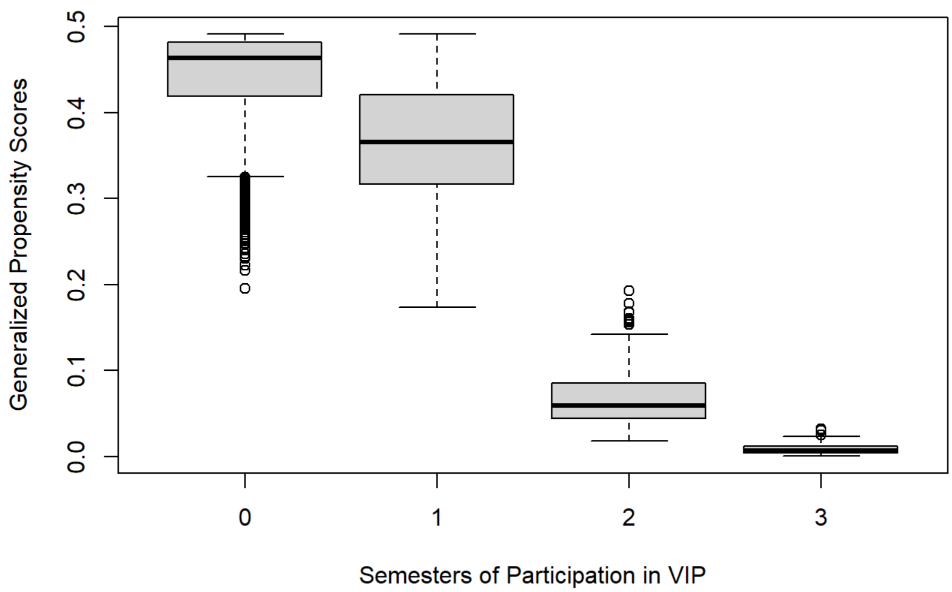
Balance Table

variable	coefBaseline	coefIPW
CITZ	0.074	0.074
FEMALE	0.040	0.011
RCETH	0.395	0.052
PELL	0.045	0.029
TRAN	0.065	0.004
GRK	0.130	0.027
STAB	0.058	0.024
GT1	0.098	0.010
LLHON	0.283	0.003
MAJREV	0.398	0.197
UROP	0.051	0.025
COOP	0.267	0.123
INT12	0.104	0.009
GPA	0.181	0.017

Sample Sizes

Generalized Propensity Scores

	0 sem	1 sem	2 sem	3 sem
Min.	0.196	0.174	0.018	0.000
1st Qu.	0.419	0.317	0.044	0.004
Median	0.463	0.366	0.059	0.006
Mean	0.443	0.368	0.068	0.008
3rd Qu.	0.482	0.421	0.084	0.012
Max.	0.491	0.491	0.193	0.032



3 Majors - Non-White

jsk

Subset Statements

```
Data <- subset(Data, CITZ != "Alien, Non-Resident*")
Data <- subset(Data, VIPSEM %in% c(0,1,2,3)) #TREATMENT LEVELS

Data <- subset(Data, MAJREV %in% c("Computer Science",
"Environmental Engineering",
"Computer Engineering"))

# Shorten Major Names after subsetting by major

Data$MAJREV[Data$MAJREV=="Computer Science"] <- "Computer Sci"
Data$MAJREV[Data$MAJREV=="Environmental Engineering"] <- "0 Env Eng"
Data$MAJREV[Data$MAJREV=="Computer Engineering"] <- "Computer Eng"
Data$MAJREV[Data$MAJREV=="Electrical Engineering"] <- "Electrical Eng"
Data$MAJREV[Data$MAJREV=="Mechanical Engineering"] <- "Mechanical Eng"
Data$MAJREV[Data$MAJREV=="Biomedical Engineering"] <- "Biomedical Eng"
Data$MAJREV[Data$MAJREV=="Materials Science & Engr"] <- "Mat Science & Eng"
Data$MAJREV[Data$MAJREV=="Civil Engineering"] <- "Civil Eng"
Data$MAJREV[Data$MAJREV=="Aerospace Engineering"] <- "Aerospace Eng"
Data$MAJREV[Data$MAJREV=="Industrial Engineering"] <- "Industrial Eng"
Data$MAJREV[Data$MAJREV=="Chemical and Biomolecular Eng"] <- "Chem & Biomolec Eng"


Data <- subset(Data, RCETH != "White")

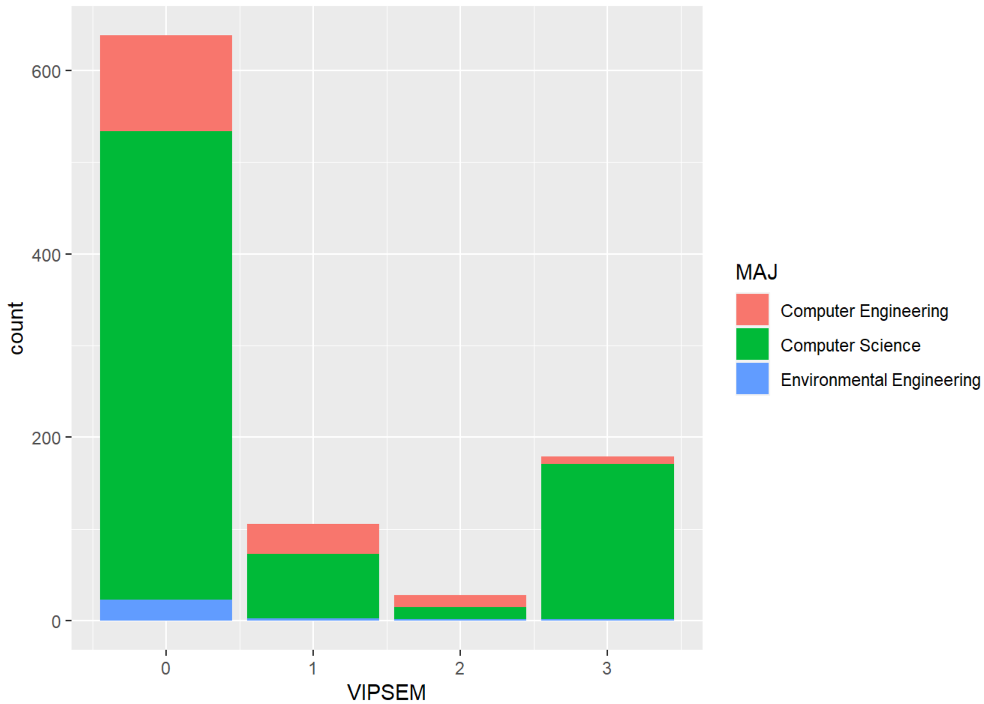
# Set ref category to "other or Unknown"
Data$RCETH[Data$RCETH == "Two or more"] <- "0 Other or Unknown"
Data$RCETH[Data$RCETH == "American Indian or Alaska Native"] <- "0 Other or Unknown"
Data$RCETH[Data$RCETH == "Native Hawaiian or Other Pacific Islander"] <- "0 Other or Unknown"
Data$RCETH[Data$RCETH == "Unknown"] <- "0 Other or Unknown"
```

GPA - Center around Grand Mean for 5 Majors Grouping

```
Data$GPA.raw <- Data$GPA
Data$GPA.from.5maj.Mean <- Data$GPA-3.586502
Data$GPA <- Data$GPA.from.5maj.Mean
```

Fequency Tables

```
##
##                 0   1   2   3
##   0 Env Eng     23   3   2   2
##   Computer Eng 104  33  13   8
##   Computer Sci 511  70  13 169
```



Propensity Score Model

```
## VIPSEM ~ CITZ + FEMALE + RCETH + PELL + TRAN + GRK + STAB + GT1 +
##     LLHON + MAJREV + UROP + COOP + INT12 + GPA
```

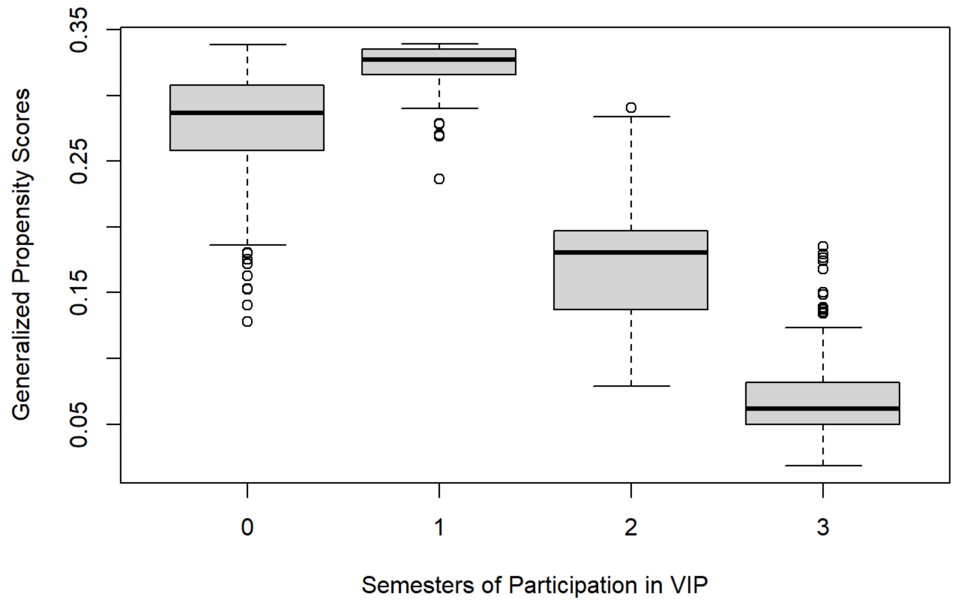
Balance Table

variable	coefBaseline	coefIPW
CITZ	0.103	0.004
FEMALE	0.210	0.015
RCETH	0.151	0.028
PELL	0.185	0.028
TRAN	0.132	0.008
GRK	0.068	0.004
STAB	0.043	0.039
GT1	0.110	0.006
LLHON	0.498	0.027
MAJREV	0.303	0.119
UROP	0.051	0.044
COOP	0.021	0.047
INT12	0.209	0.026
GPA	0.253	0.018

Sample Sizes

Generalized Propensity Scores

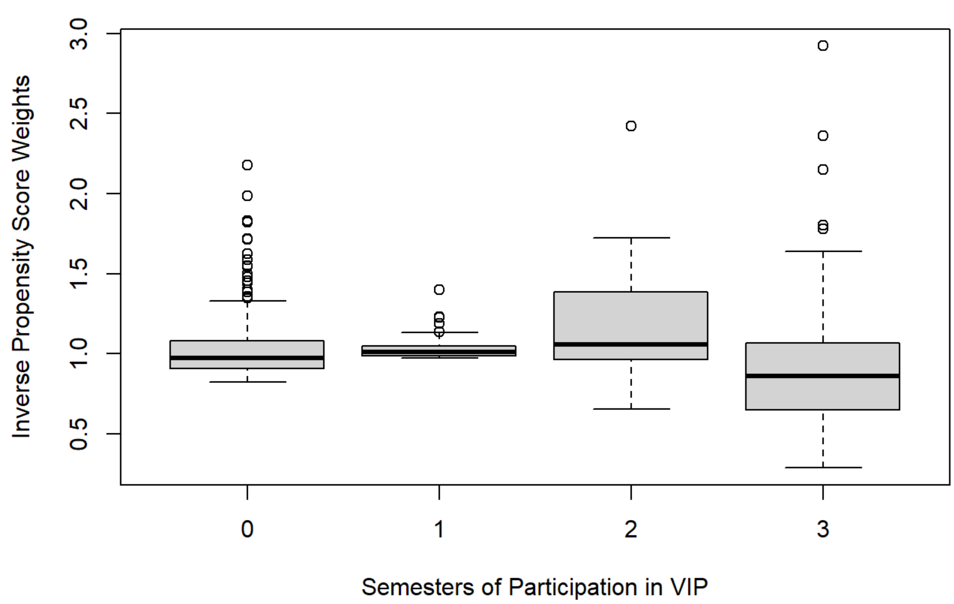
	0 sem	1 sem	2 sem	3 sem
Min.	0.128	0.236	0.079	0.018
1st Qu.	0.258	0.316	0.138	0.050
Median	0.287	0.327	0.180	0.062
Mean	0.280	0.322	0.177	0.071
3rd Qu.	0.308	0.335	0.195	0.082
Max.	0.339	0.339	0.291	0.185

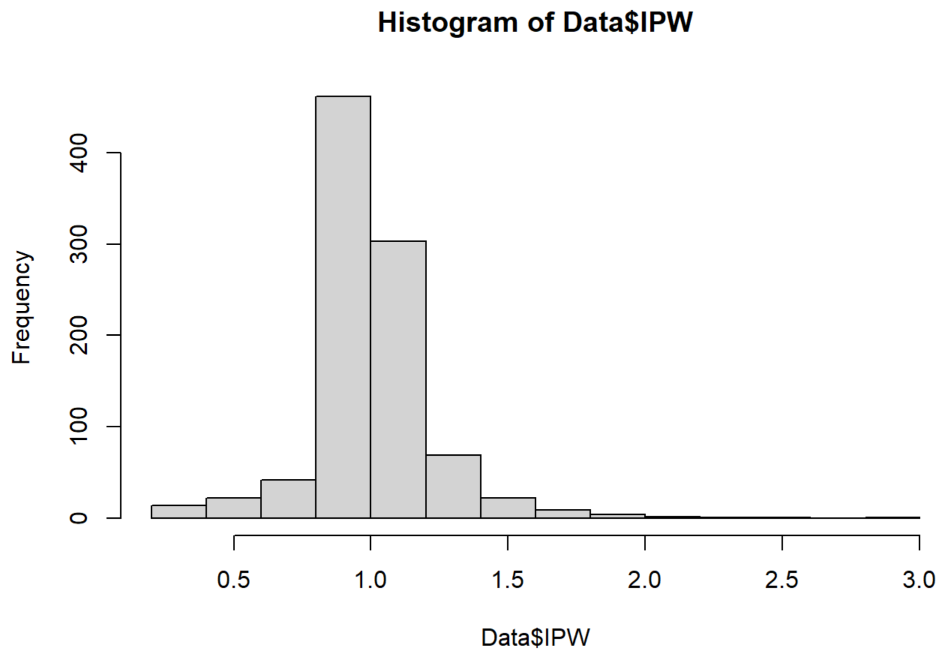


Inverse Propensity Score Weights

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.2877  0.8953  0.9784  0.9982  1.0750  2.9209
```

	0 sem	1 sem	2 sem	3 sem
Min.	0.823	0.975	0.655	0.288
1st Qu.	0.905	0.986	0.975	0.651
Median	0.973	1.010	1.055	0.860
Mean	1.015	1.029	1.164	0.893
3rd Qu.	1.080	1.046	1.379	1.068
Max.	2.180	1.398	2.420	2.921

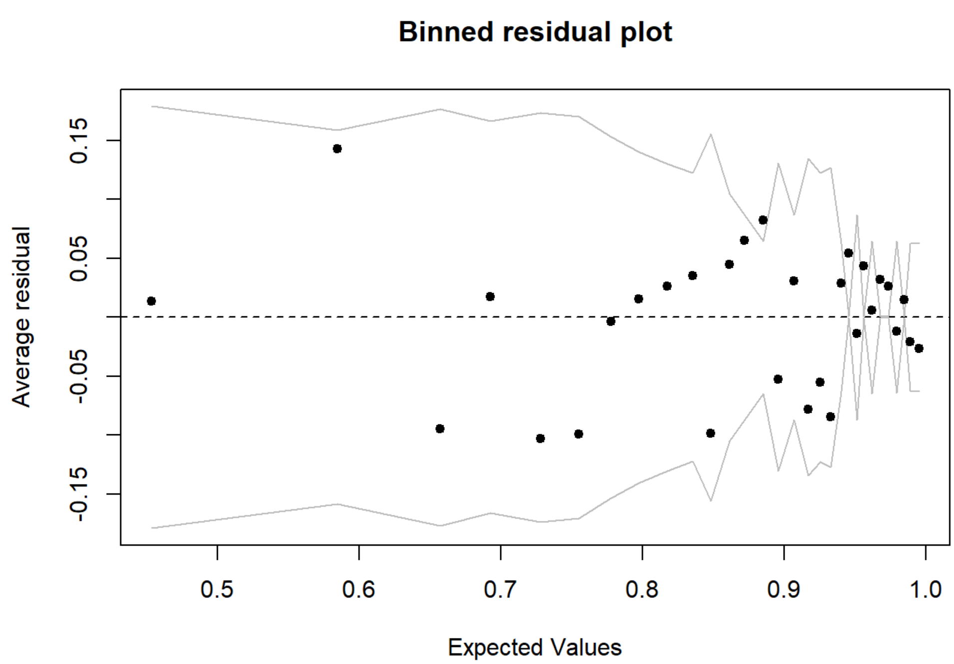




REGRESSION

Residuals

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
## -3.2248  0.1844  0.3660  0.1537  0.5771  1.6435
```




```
##                R-Squared      Adjusted R-Squared
##                    0.151                   0.130
```

Regression Results

```
## 
## Call:
```

```
## svyglm(formula = ..1, design = ..2, family = ..3)
##
## Survey design:
## survey::svydesign(...)
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)              1.122760   0.701958   1.599  0.11006
## VIPSEM                   0.413775   0.123006   3.364  0.00080 ***
## CITZResident NonCitizen  0.302104   0.312078   0.968  0.33328
## FEMALE                   0.392530   0.264200   1.486  0.13769
## RCETHAsian               0.005512   0.337938   0.016  0.98699
## RCETHURM                 0.200686   0.369660   0.543  0.58733
## PELL                    -0.479257   0.218010  -2.198  0.02817 *
## TRAN                    -0.487372   0.247327  -1.971  0.04907 *
## GRK                      0.608605   0.359953   1.691  0.09121 .
## STAB                     1.015992   0.414149   2.453  0.01434 *
## GT1                     -0.087090   0.300211  -0.290  0.77181
## LLHON                    0.772636   0.780981   0.989  0.32277
## MAJREVComputer Eng       0.111591   0.636013   0.175  0.86076
## MAJREVComputer Sci       0.513491   0.613734   0.837  0.40299
## UROP                    -0.239891   0.272780  -0.879  0.37940
## COOP1 Some CoOp          0.347367   0.422667   0.822  0.41138
## COOP3 CoOpDegDesig       1.742782   1.100124   1.584  0.11350
## INT12                    0.798378   0.298184   2.677  0.00755 **
## GPA                      0.980489   0.215511   4.550 6.09e-06 ***
## YR2018                   0.702785   0.430842   1.631  0.10319
## YR2019                  -0.270454   0.361610  -0.748  0.45470
## YR2020                  -0.231054   0.361511  -0.639  0.52289
## YR2021                   0.104040   0.369468   0.282  0.77832
## YR2022                  -0.315689   0.348215  -0.907  0.36486
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 1.367761)
##
## Number of Fisher Scoring iterations: 6
```

```
##                         Estimate Std. Error t value Pr(>|t|)   2.5 %  97.5 %
## (Intercept)               1.1228     0.7020  1.5995   0.1100 -0.2548  2.5003
## VIPSEM                    0.4138     0.1230  3.3639   0.0008  0.1724  0.6552
## CITZResident NonCitizen   0.3021     0.3121  0.9680   0.3333 -0.3103  0.9145
## FEMALE                    0.3925     0.2642  1.4857   0.1377 -0.1260  0.9110
## RCETHAsian                0.0055     0.3379  0.0163   0.9870 -0.6577  0.6687
## RCETHURM                  0.2007     0.3697  0.5429   0.5873 -0.5248  0.9261
## PELL                     -0.4793     0.2180 -2.1983   0.0282 -0.9071 -0.0514
## TRAN                     -0.4874     0.2473 -1.9706   0.0491 -0.9727 -0.0020
## GRK                       0.6086     0.3600  1.6908   0.0912 -0.0978  1.3150
## STAB                      1.0160     0.4141  2.4532   0.0143  0.2032  1.8287
## GT1                      -0.0871     0.3002 -0.2901   0.7718 -0.6762  0.5021
## LLHON                     0.7726     0.7810  0.9893   0.3228 -0.7600  2.3053
## MAJREVComputer Eng        0.1116     0.6360  0.1755   0.8608 -1.1366  1.3597
## MAJREVComputer Sci        0.5135     0.6137  0.8367   0.4030 -0.6909  1.7179
## UROP                     -0.2399     0.2728 -0.8794   0.3794 -0.7752  0.2954
## COOP1 Some CoOp           0.3474     0.4227  0.8218   0.4114 -0.4821  1.1768
## COOP3 CoOpDegDesig        1.7428     1.1001  1.5842   0.1135 -0.4162  3.9017
## INT12                     0.7984     0.2982  2.6775   0.0075  0.2132  1.3836
## GPA                       0.9805     0.2155  4.5496   0.0000  0.5576  1.4034
## YR2018                    0.7028     0.4308  1.6312   0.1032 -0.1427  1.5483
## YR2019                   -0.2705     0.3616 -0.7479   0.4547 -0.9801  0.4392
## YR2020                   -0.2311     0.3615 -0.6391   0.5229 -0.9405  0.4784
## YR2021                    0.1040     0.3695  0.2816   0.7783 -0.6210  0.8291
## YR2022                   -0.3157     0.3482 -0.9066   0.3649 -0.9990  0.3677
```

Adjusted Odds Ratios with Confidence Intervals

```
##                           AOR 2.5 % 97.5 %
## VIPSEM                  1.513 1.188  1.925
## CITZResident NonCitizen 1.353 0.733  2.496
## FEMALE                  1.481 0.882  2.487
## RCETHAsian              1.006 0.518  1.952
## RCETHURM                1.222 0.592  2.525
## PELL                    0.619 0.404  0.950
## TRAN                    0.614 0.378  0.998
## GRK                     1.838 0.907  3.725
## STAB                    2.762 1.225  6.226
## GT1                     0.917 0.509  1.652
## LLHON                   2.165 0.468 10.027
## MAJREVComputer Eng      1.118 0.321  3.895
## MAJREVComputer Sci      1.671 0.501  5.573
## UROP                    0.787 0.461  1.344
## COOP1 Some CoOp         1.415 0.617  3.244
```

```
## COOP3 CoOpDegDesig      5.713 0.660 49.492
## INT12                   2.222 1.238  3.989
## GPA                     2.666 1.746  4.069
## YR2018                  2.019 0.867  4.704
## YR2019                  0.763 0.375  1.551
## YR2020                  0.794 0.390  1.614
## YR2021                  1.110 0.537  2.291
## YR2022                  0.729 0.368  1.444
```

3 Majors - White

jsk

Subset Statements

```
Data <- subset(Data, CITZ != "Alien, Non-Resident*")
Data <- subset(Data, VIPSEM %in% c(0,1,2,3)) #TREATMENT LEVELS

Data <- subset(Data, MAJREV %in% c("Computer Science",
"Environmental Engineering",
"Computer Engineering"))

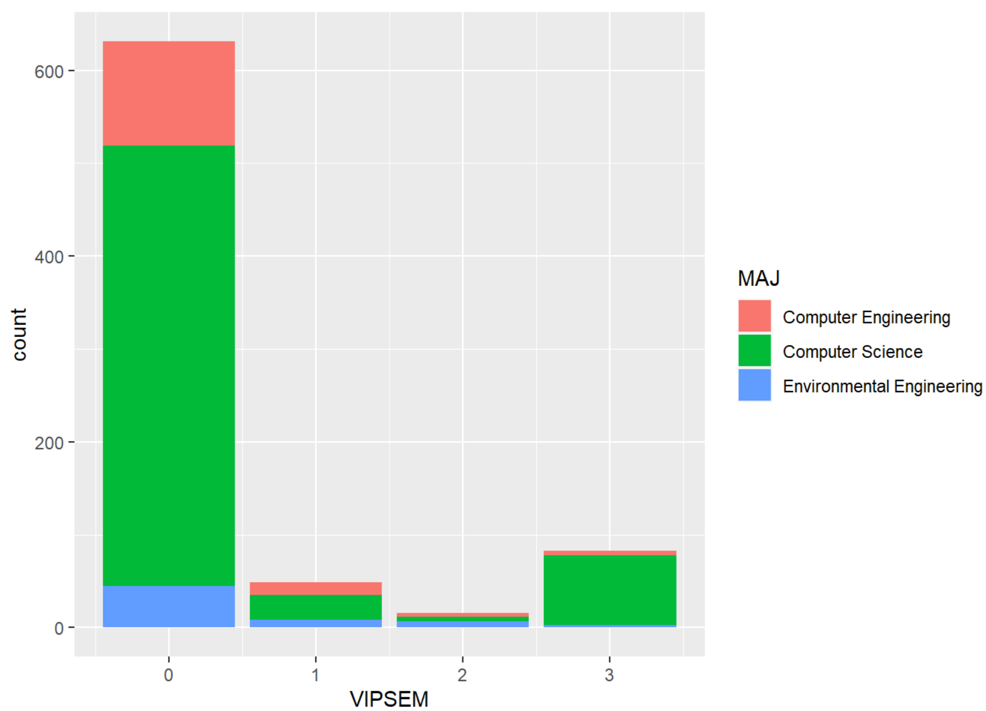
# Shorten Major Names after subsetting by major

Data$MAJREV[Data$MAJREV=="Computer Science"] <- "Computer Sci"
Data$MAJREV[Data$MAJREV=="Environmental Engineering"] <- "0 Env Eng"
Data$MAJREV[Data$MAJREV=="Computer Engineering"] <- "Computer Eng"
Data$MAJREV[Data$MAJREV=="Electrical Engineering"] <- "Electrical Eng"
Data$MAJREV[Data$MAJREV=="Mechanical Engineering"] <- "Mechanical Eng"
Data$MAJREV[Data$MAJREV=="Biomedical Engineering"] <- "Biomedical Eng"
Data$MAJREV[Data$MAJREV=="Materials Science & Engr"] <- "Mat Science & Eng"
Data$MAJREV[Data$MAJREV=="Civil Engineering"] <- "Civil Eng"
Data$MAJREV[Data$MAJREV=="Aerospace Engineering"] <- "Aerospace Eng"
Data$MAJREV[Data$MAJREV=="Industrial Engineering"] <- "Industrial Eng"
Data$MAJREV[Data$MAJREV=="Chemical and Biomolecular Eng"] <- "Chem & Biomolec Eng"

Data <- subset(Data, RCETH == "White")
```

Fequency Tables

```
##
##                  0   1   2   3
##   0 Env Eng     45   8   7   3
##   Computer Eng 112  14   4   5
##   Computer Sci 474  27   5  75
```



Propensity Score Model

```
## VIPSEM ~ FEMALE + PELL + TRAN + GRK + STAB + GT1 + LLHON + MAJREV +
##     UROP + COOP + INT12 + GPA
```

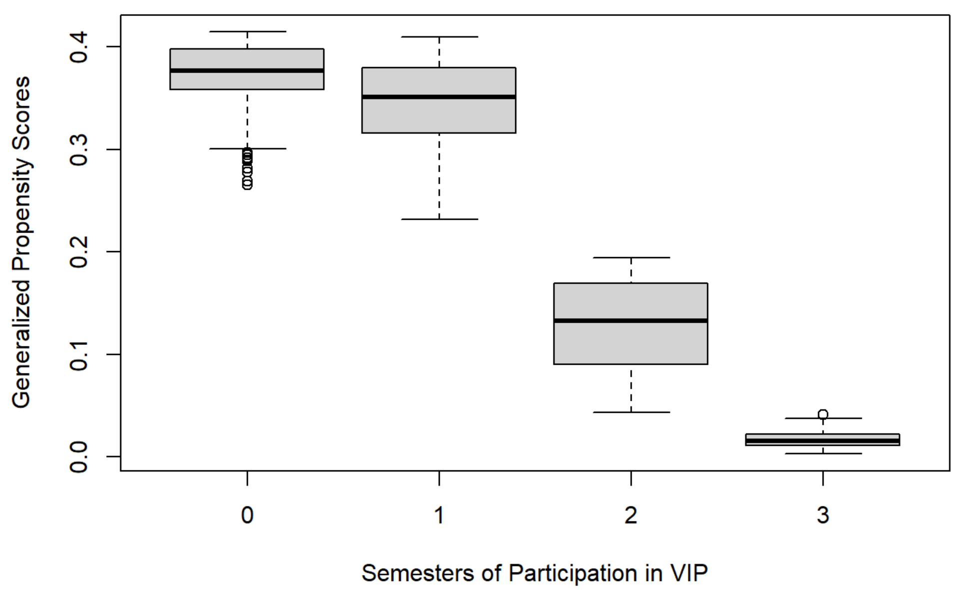
Balance Table

variable	coefBaseline	coefIPW
FEMALE	0.184	0.055
PELL	0.066	0.069
TRAN	0.095	0.048
GRK	0.001	0.110
STAB	0.051	0.010
GT1	0.024	0.074
LLHON	0.122	0.008
MAJREV	0.227	0.045
UROP	0.085	0.023
COOP	0.211	0.046
INT12	0.288	0.034
GPA	0.291	0.025

Sample Sizes

Generalized Propensity Scores

	0 sem	1 sem	2 sem	3 sem
Min.	0.265	0.232	0.043	0.003
1st Qu.	0.358	0.316	0.096	0.011
Median	0.377	0.351	0.133	0.016
Mean	0.373	0.349	0.128	0.017
3rd Qu.	0.398	0.380	0.166	0.022
Max.	0.415	0.409	0.194	0.041



3 Majors - Female

jsk

Subset Statements

```
Data <- subset(Data, CITZ != "Alien, Non-Resident*")
Data <- subset(Data, VIPSEM %in% c(0,1,2,3)) #TREATMENT LEVELS

Data <- subset(Data, MAJREV %in% c("Computer Science",
"Environmental Engineering",
"Computer Engineering"))

Data <- subset(Data, FEMALE == 1)

# Shorten Major Names after subsetting by major

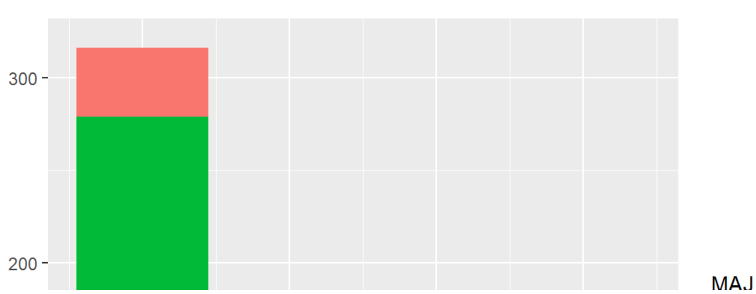
Data$MAJREV[Data$MAJREV=="Computer Science"] <- "Computer Sci"
Data$MAJREV[Data$MAJREV=="Environmental Engineering"] <- "0 Env Eng"
Data$MAJREV[Data$MAJREV=="Computer Engineering"] <- "Computer Eng"
Data$MAJREV[Data$MAJREV=="Electrical Engineering"] <- "Electrical Eng"
Data$MAJREV[Data$MAJREV=="Mechanical Engineering"] <- "Mechanical Eng"
Data$MAJREV[Data$MAJREV=="Biomedical Engineering"] <- "Biomedical Eng"
Data$MAJREV[Data$MAJREV=="Materials Science & Engr"] <- "Mat Science & Eng"
Data$MAJREV[Data$MAJREV=="Civil Engineering"] <- "Civil Eng"
Data$MAJREV[Data$MAJREV=="Aerospace Engineering"] <- "Aerospace Eng"
Data$MAJREV[Data$MAJREV=="Industrial Engineering"] <- "Industrial Eng"
Data$MAJREV[Data$MAJREV=="Chemical and Biomolecular Eng"] <- "Chem & Biomolec Eng"
```

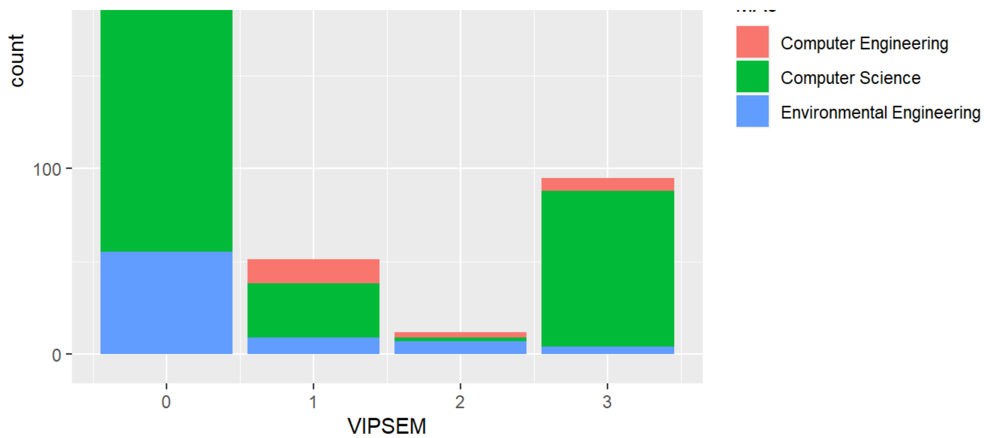
GPA - Center around Grand Mean for 5 Majors Grouping

```
Data$GPA.raw <- Data$GPA
Data$GPA.from.5maj.Mean <- Data$GPA-3.586502
Data$GPA <- Data$GPA.from.5maj.Mean
```

Fequency Tables

```
##
##                  0   1   2   3
##   0 Env Eng     55   9   7   4
##   Computer Eng  37  13   3   7
##   Computer Sci 224  29   2  84
```





Propensity Score Model

```
## VIPSEM ~ CITZ + RCETH + PELL + TRAN + GRK + STAB + GT1 + LLHON +
##     MAJREV + UROP + COOP + INT12 + GPA
```

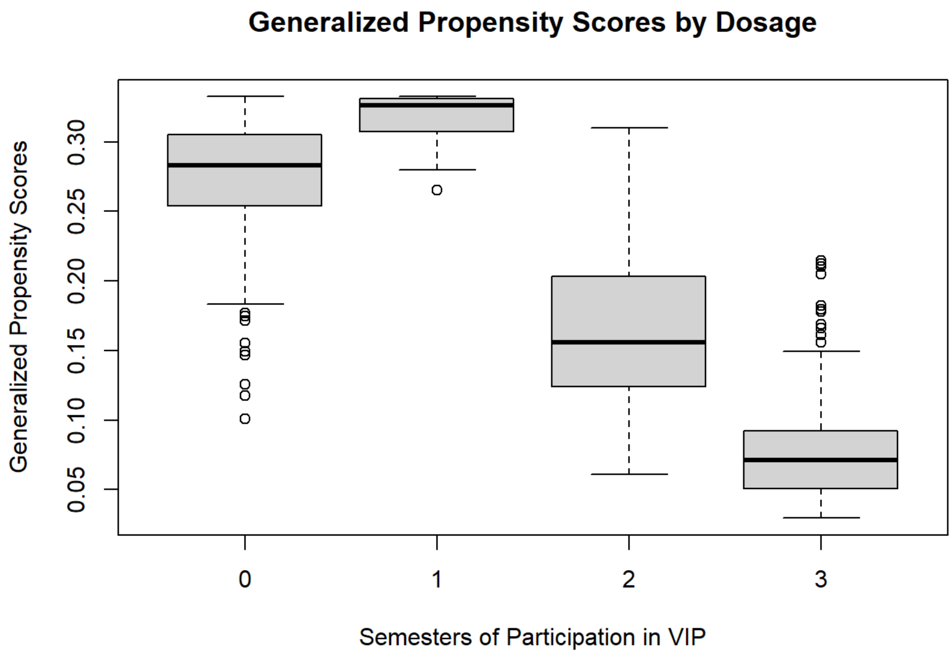
Balance Table

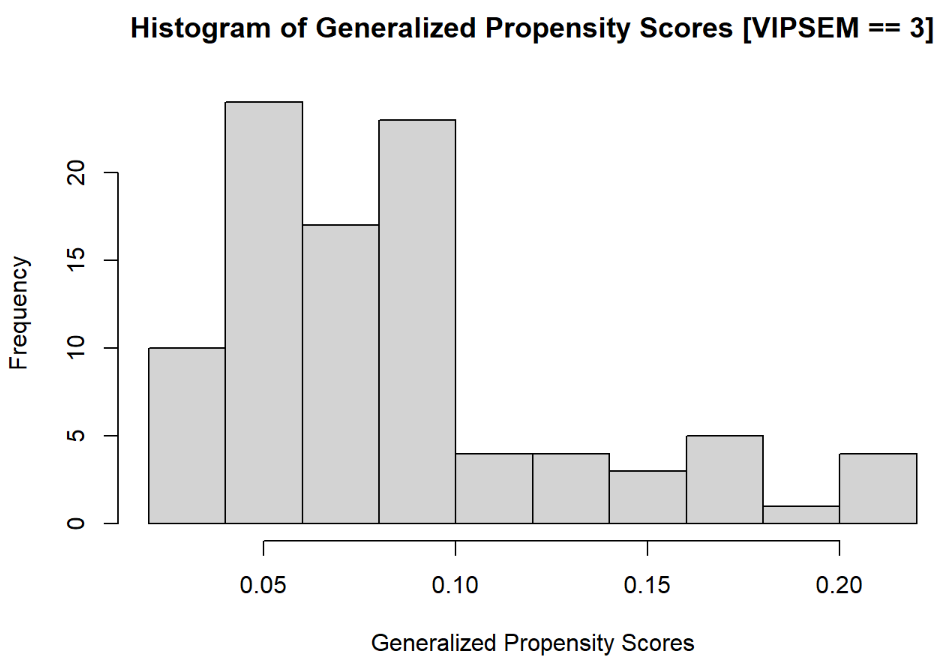
variable	coefBaseline	coefIPW
CITZ	0.023	0.059
RCETH	0.366	0.053
PELL	0.110	0.012
TRAN	0.158	0.009
GRK	0.070	0.025
STAB	0.000	0.002
GT1	0.028	0.028
LLHON	0.623	0.035
MAJREV	0.312	0.060
UROP	0.057	0.066
COOP	0.444	0.275
INT12	0.148	0.049
GPA	0.189	0.009

Sample Sizes

Generalized Propensity Scores

	0 sem	1 sem	2 sem	3 sem
Min.	0.101	0.265	0.061	0.030
1st Qu.	0.254	0.307	0.129	0.051
Median	0.283	0.326	0.156	0.071
Mean	0.276	0.319	0.167	0.084
3rd Qu.	0.305	0.331	0.195	0.092
Max.	0.332	0.332	0.310	0.215



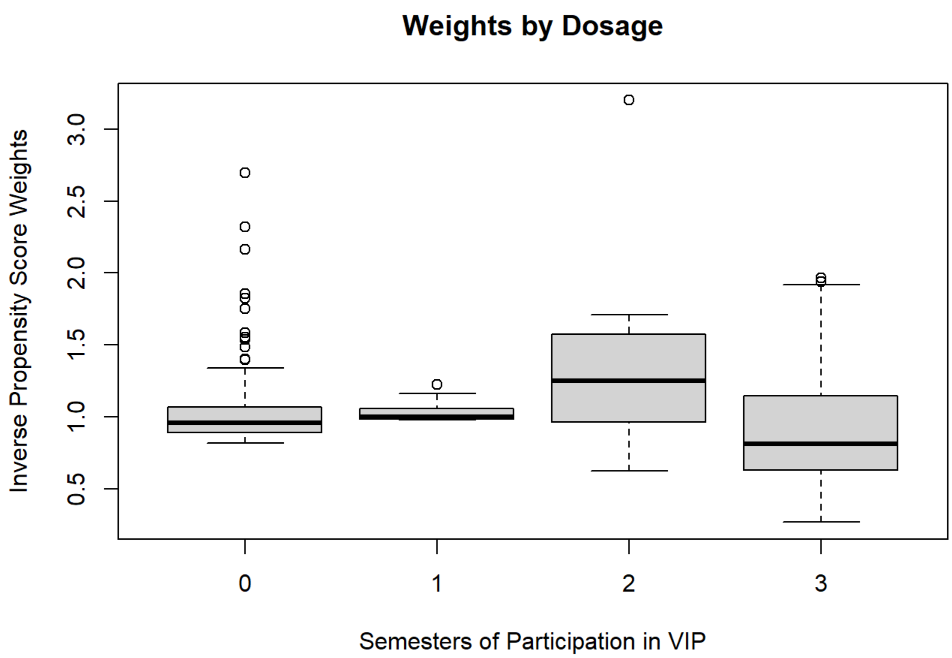


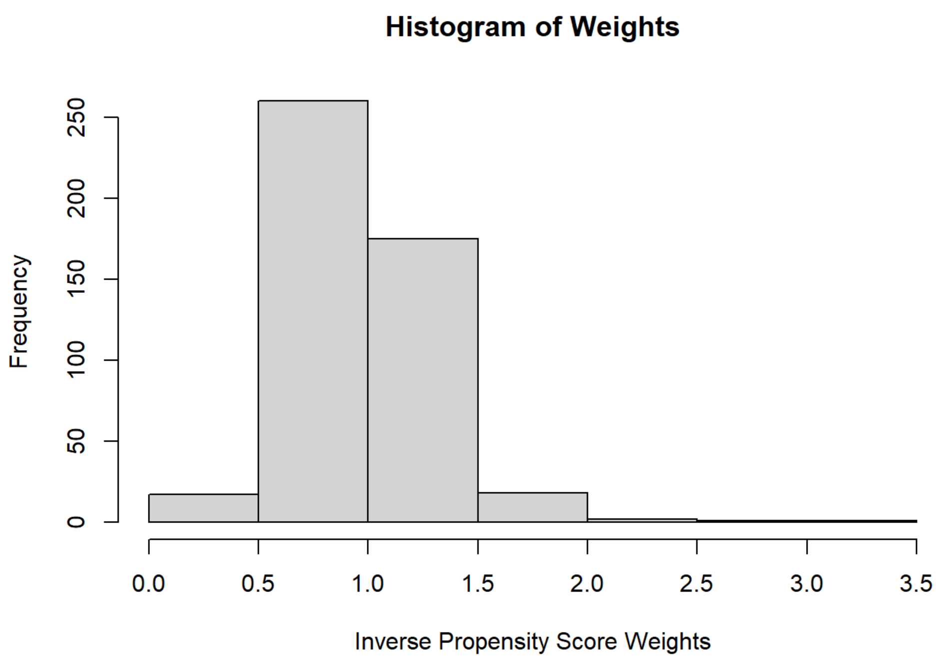
Inverse Propensity Score Weights

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.2709  0.8728  0.9810  0.9983  1.0770  3.1992
```

	0 sem	1 sem	2 sem	3 sem
Min.	0.819	0.980	0.629	0.271
1st Qu.	0.892	0.985	1.005	0.632
Median	0.961	1.000	1.251	0.815
Mean	1.015	1.026	1.369	0.881
3rd Qu.	1.070	1.061	1.516	1.150

	0 sem	1 sem	2 sem	3 sem
Max.	2.696	1.228	3.199	1.965



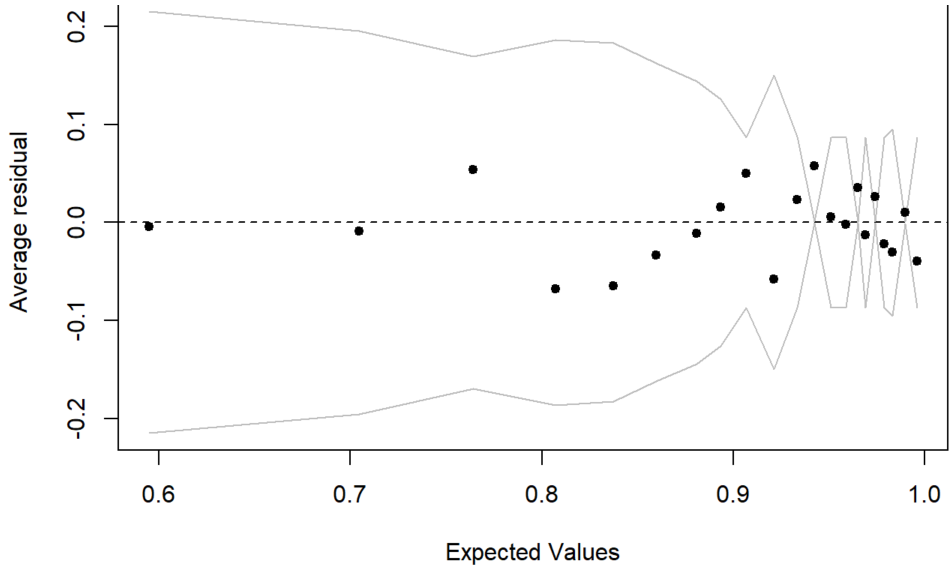


REGRESSION

Residuals

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -3.1331  0.1885  0.3194  0.1545  0.5013  1.1464
```

Binned residual plot




```
##              R-Squared      Adjusted R-Squared
##                  0.143                   0.099
```

Regression Results

```
##
## Call:
## svyglm(formula = ..1, design = ..2, family = ..3)
##
## Survey design:
## survey::svydesign(...)
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)               0.11601    0.73245   0.158  0.87423
## VIPSEM                    0.25041    0.18289   1.369  0.17162
## CITZResident NonCitizen -0.23439    0.70020  -0.335  0.73798
## RCETHAsian                0.51646    0.45703   1.130  0.25907
## RCETHOther or Unknown     0.12829    0.59685   0.215  0.82991
## RCETHURM                  0.57936    0.50074   1.157  0.24788
## PELL                      0.30996    0.38701   0.801  0.42360
## TRAN                      0.08283    0.47912   0.173  0.86283
## GRK                       0.72061    0.42448   1.698  0.09027 .
## STAB                      0.81639    0.40443   2.019  0.04412 *
## GT1                       0.15100    0.40010   0.377  0.70604
## LLHON                     1.63335    0.86934   1.879  0.06091 .
## MAJREVComputer Eng        0.86528    0.58213   1.486  0.13788
## MAJREVComputer Sci        1.12168    0.43222   2.595  0.00976 **
## UROP                     -0.27934    0.38348  -0.728  0.46673
## COOP1 Some CoOp           0.47919    0.62970   0.761  0.44706
## COOP3 CoOpDegDesig        1.52876    1.22570   1.247  0.21295
## INT12                     0.24744    0.40744   0.607  0.54395
## GPA                       1.07839    0.33465   3.222  0.00136 **
## YR2018                    1.06465    0.80463   1.323  0.18646
```

```
## YR2019                    0.86626    0.65096   1.331  0.18395
## YR2020                   -0.70913    0.49708  -1.427  0.15440
## YR2021                   -0.06439    0.58132  -0.111  0.91185
## YR2022                    0.40757    0.56397   0.723  0.47025
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 1.323531)
##
## Number of Fisher Scoring iterations: 6
```

```
##                           Estimate Std. Error t value Pr(>|t|)   2.5 % 97.5 %
## (Intercept)                 0.1160     0.7324  0.1584   0.8742 -1.3232 1.5553
## VIPSEM                      0.2504     0.1829  1.3692   0.1716 -0.1090 0.6098
## CITZResident NonCitizen    -0.2344     0.7002 -0.3347   0.7380 -1.6103 1.1415
## RCETHAsian                  0.5165     0.4570  1.1300   0.2590 -0.3816 1.4145
## RCETHOther or Unknown       0.1283     0.5969  0.2149   0.8299 -1.0445 1.3011
## RCETHURM                    0.5794     0.5007  1.1570   0.2478 -0.4046 1.5633
## PELL                        0.3100     0.3870  0.8009   0.4236 -0.4505 1.0704
## TRAN                        0.0828     0.4791  0.1729   0.8628 -0.8586 1.0243
## GRK                         0.7206     0.4245  1.6976   0.0902 -0.1135 1.5547
## STAB                        0.8164     0.4044  2.0186   0.0441  0.0217 1.6111
## GT1                         0.1510     0.4001  0.3774   0.7060 -0.6352 0.9372
## LLHON                       1.6334     0.8693  1.8788   0.0609 -0.0749 3.3416
## MAJREVComputer Eng          0.8653     0.5821  1.4864   0.1378 -0.2786 2.0092
## MAJREVComputer Sci          1.1217     0.4322  2.5952   0.0097  0.2724 1.9710
## UROP                       -0.2793     0.3835 -0.7284   0.4667 -1.0329 0.4742
## COOP1 Some CoOp             0.4792     0.6297  0.7610   0.4470 -0.7582 1.7165
## COOP3 CoOpDegDesig          1.5288     1.2257  1.2473   0.2129 -0.8797 3.9373
## INT12                       0.2474     0.4074  0.6073   0.5439 -0.5532 1.0481
## GPA                         1.0784     0.3347  3.2224   0.0014  0.4208 1.7360
## YR2018                      1.0647     0.8046  1.3232   0.1864 -0.5164 2.6457
## YR2019                      0.8663     0.6510  1.3307   0.1839 -0.4129 2.1454
## YR2020                     -0.7091     0.4971 -1.4266   0.1544 -1.6859 0.2676
## YR2021                     -0.0644     0.5813 -0.1108   0.9118 -1.2067 1.0779
## YR2022                      0.4076     0.5640  0.7227   0.4702 -0.7006 1.5158
```

Adjusted Odds Ratios with Confidence Intervals

```
##                             AOR 2.5 % 97.5 %
## VIPSEM                    1.285 0.897  1.840
## CITZResident NonCitizen 0.791 0.200  3.132
## RCETHAsian                1.676 0.683  4.115
## RCETHOther or Unknown     1.137 0.352  3.674
## RCETHURM                  1.785 0.667  4.775
## PELL                      1.363 0.637  2.917
## TRAN                      1.086 0.424  2.785
## GRK                       2.056 0.893  4.734
## STAB                      2.262 1.022  5.009
## GT1                       1.163 0.530  2.553
## LLHON                     5.121 0.928 28.271
```

```
## MAJREVComputer Eng      2.376 0.757  7.458
## MAJREVComputer Sci      3.070 1.313  7.179
## UROP                    0.756 0.356  1.607
## COOP1 Some CoOp         1.615 0.468  5.566
## COOP3 CoOpDegDesig      4.612 0.415 51.294
## INT12                   1.281 0.575  2.852
## GPA                     2.940 1.523  5.675
## YR2018                  2.900 0.597 14.097
## YR2019                  2.378 0.662  8.547
## YR2020                  0.492 0.185  1.307
## YR2021                  0.938 0.299  2.939
## YR2022                  1.503 0.496  4.554
```

3 Majors - Pell

jsk

Subset Statements

```
Data <- subset(Data, CITZ != "Alien, Non-Resident*")
Data <- subset(Data, VIPSEM %in% c(0,1,2,3)) #TREATMENT LEVELS

Data <- subset(Data, MAJREV %in% c("Computer Science",
"Environmental Engineering",
"Computer Engineering"))

Data <- subset(Data, PELL == 1)

# Shorten Major Names after subsetting by major

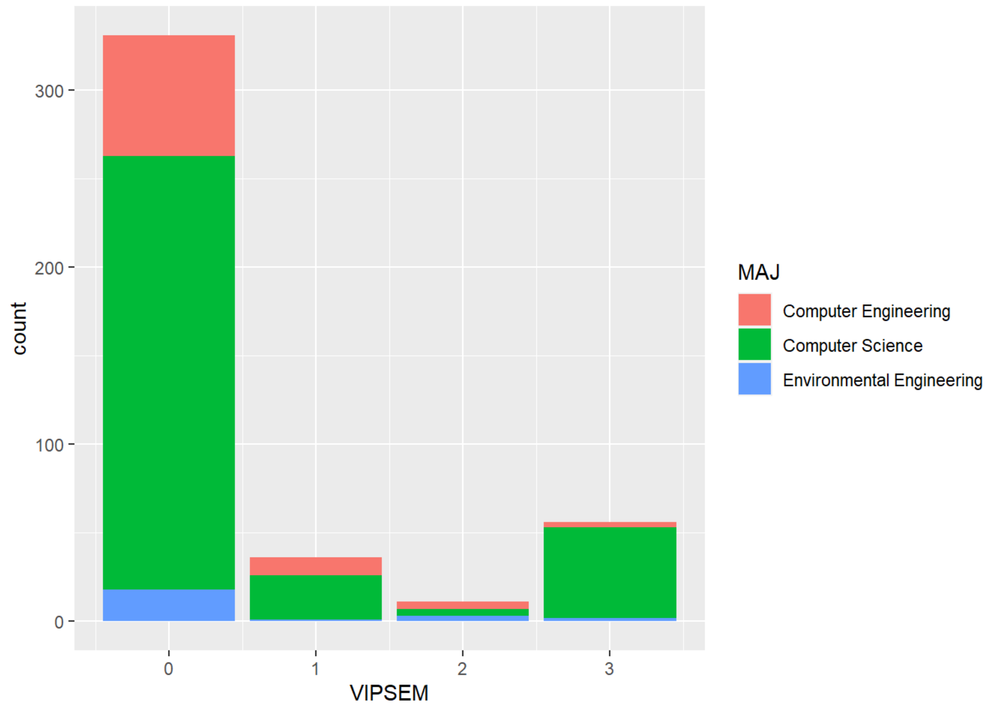
Data$MAJREV[Data$MAJREV=="Computer Science"] <- "Computer Sci"
Data$MAJREV[Data$MAJREV=="Environmental Engineering"] <- "0 Env Eng"
Data$MAJREV[Data$MAJREV=="Computer Engineering"] <- "Computer Eng"
Data$MAJREV[Data$MAJREV=="Electrical Engineering"] <- "Electrical Eng"
Data$MAJREV[Data$MAJREV=="Mechanical Engineering"] <- "Mechanical Eng"
Data$MAJREV[Data$MAJREV=="Biomedical Engineering"] <- "Biomedical Eng"
Data$MAJREV[Data$MAJREV=="Materials Science & Engr"] <- "Mat Science & Eng"
Data$MAJREV[Data$MAJREV=="Civil Engineering"] <- "Civil Eng"
Data$MAJREV[Data$MAJREV=="Aerospace Engineering"] <- "Aerospace Eng"
Data$MAJREV[Data$MAJREV=="Industrial Engineering"] <- "Industrial Eng"
Data$MAJREV[Data$MAJREV=="Chemical and Biomolecular Eng"] <- "Chem & Biomolec Eng"
```

GPA - Center around Grand Mean for 5 Majors Grouping

```
Data$GPA.raw <- Data$GPA
Data$GPA.from.5maj.Mean <- Data$GPA-3.586502
Data$GPA <- Data$GPA.from.5maj.Mean
```

Fequency Tables

```
##
##                 0   1   2   3
##   0 Env Eng     18   1   3   2
##   Computer Eng  68  10   4   3
##   Computer Sci 245  25   4  51
```



Propensity Score Model

```
## VIPSEM ~ CITZ + FEMALE + RCETH + TRAN + GRK + STAB + GT1 + GT2 +
##     MAJREV + UROP + COOP + INT12 + GPA
```

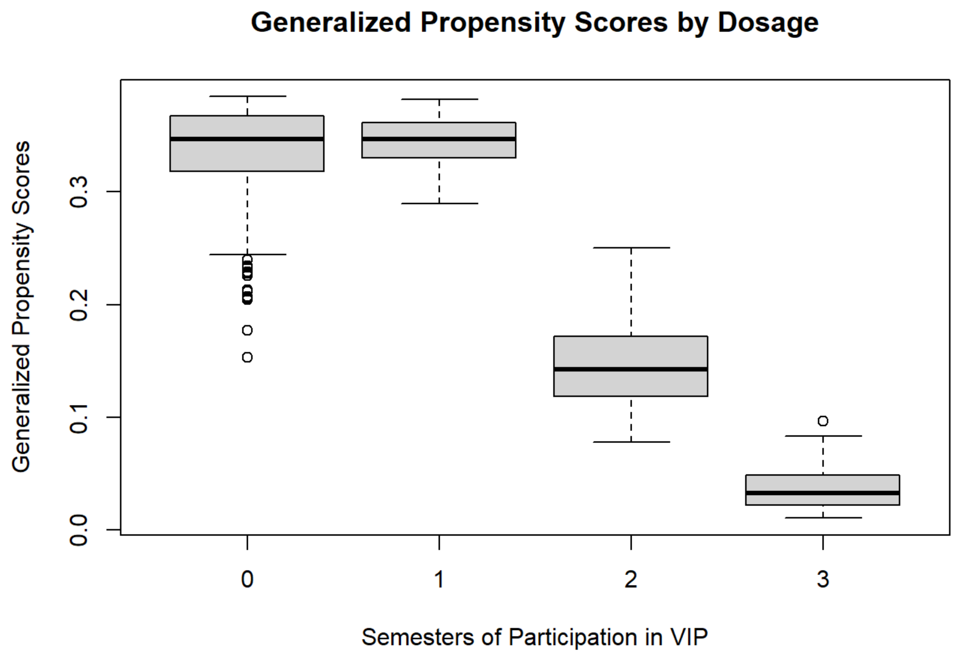
Balance Table

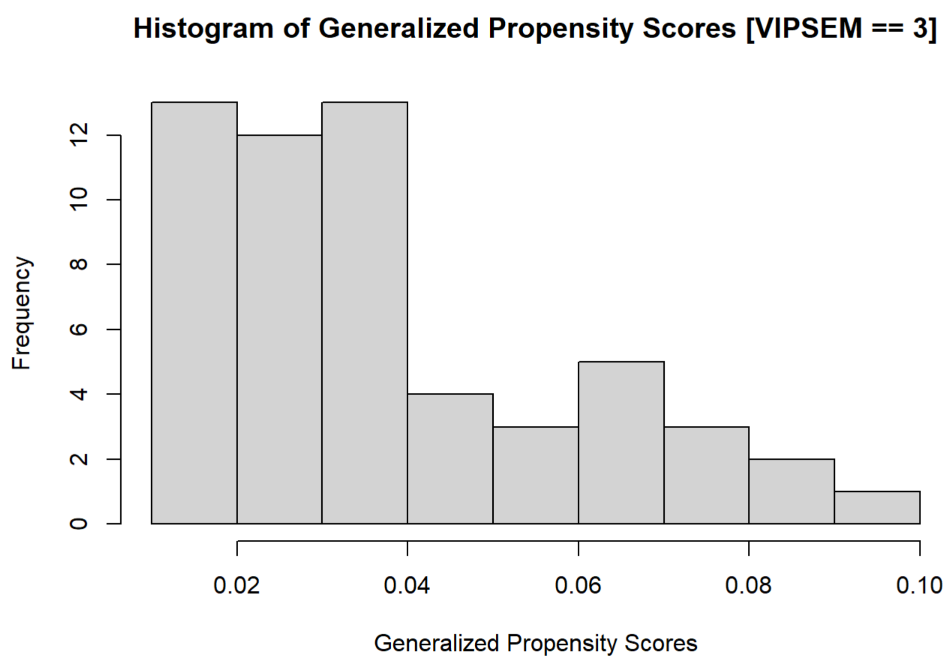
variable	coefBaseline	coefIPW
CITZ	0.023	0.050
FEMALE	0.173	0.009
RCETH	0.257	0.245
TRAN	0.042	0.017
GRK	0.050	0.016
STAB	0.136	0.110
GT1	0.114	0.079
GT2	0.176	0.002
MAJREV	0.216	0.072
UROP	0.014	0.012
COOP	0.221	0.181
INT12	0.388	0.027
GPA	0.201	0.029

Sample Sizes

Generalized Propensity Scores

	0 sem	1 sem	2 sem	3 sem
Min.	0.153	0.289	0.078	0.011
1st Qu.	0.318	0.331	0.118	0.022
Median	0.347	0.347	0.142	0.032
Mean	0.336	0.346	0.152	0.038
3rd Qu.	0.368	0.361	0.172	0.048
Max.	0.384	0.382	0.250	0.096

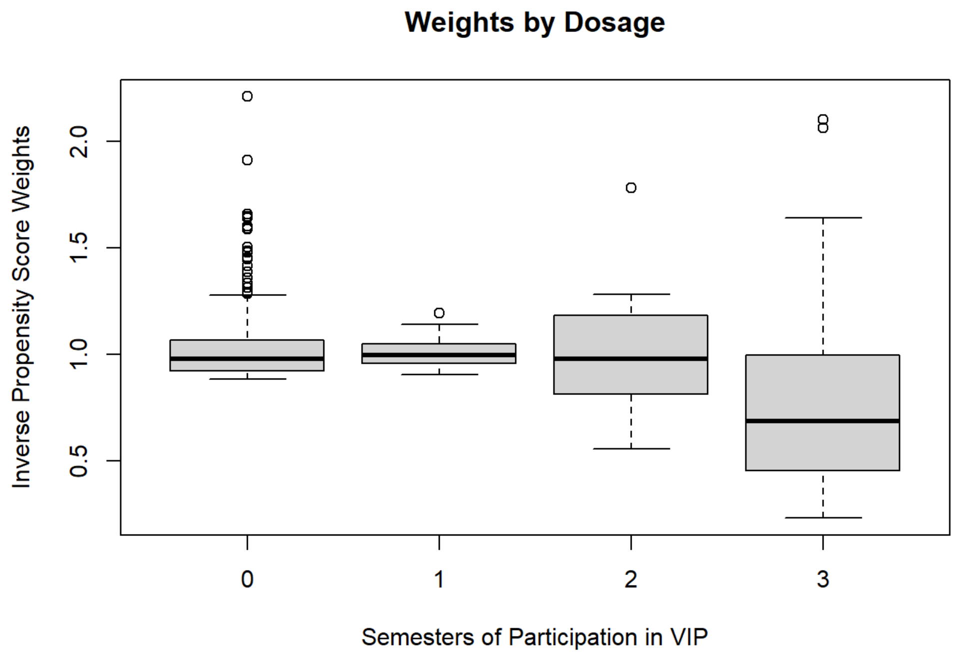




Inverse Propensity Score Weights

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.2298  0.9088  0.9713  0.9965  1.0614  2.2109
```

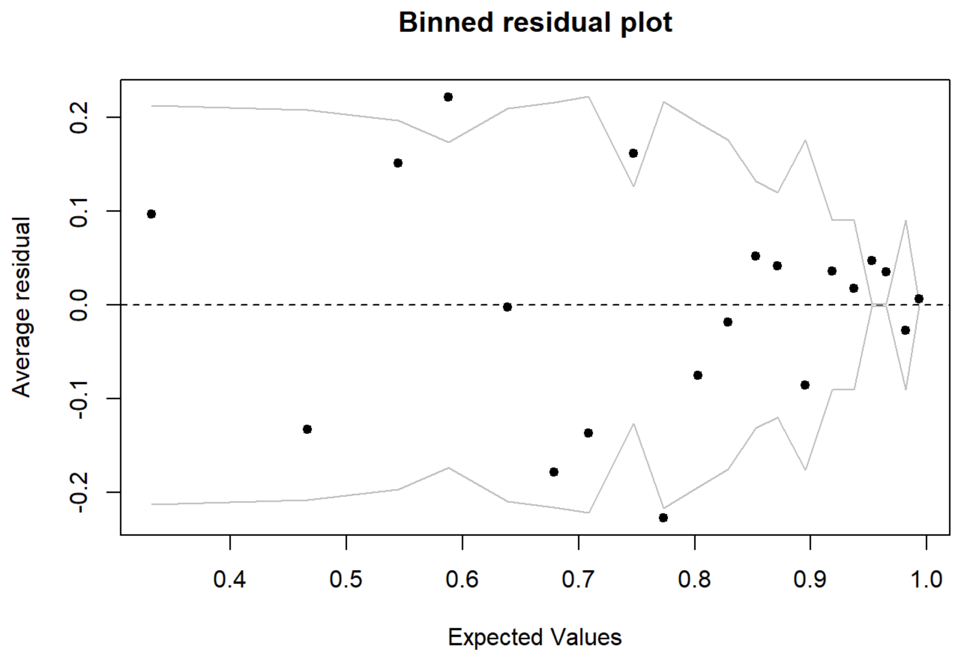
	0 sem	1 sem	2 sem	3 sem
Min.	0.882	0.904	0.556	0.230
1st Qu.	0.922	0.956	0.811	0.459
Median	0.977	0.997	0.978	0.685
Mean	1.032	1.004	1.015	0.781
3rd Qu.	1.066	1.042	1.182	0.986
Max.	2.211	1.194	1.781	2.103



REGRESSION

Residuals

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max. 
## -2.60126  0.09092  0.41826  0.12613  0.73391  1.60769
```




```
##              R-Squared      Adjusted R-Squared
##                  0.170                   0.124
```

Regression Results

```
##
## Call:
## svyglm(formula = ..1, design = ..2, family = ..3)
##
## Survey design:
## survey::svydesign(...)
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)              -1.31785    0.88056  -1.497 0.135264
## VIPSEM                    0.58115    0.17063   3.406 0.000725 ***
## CITZResident NonCitizen -0.47952    0.37955  -1.263 0.207167
## FEMALE                    0.99850    0.38476   2.595 0.009795 **
## RCETHAsian                0.07633    0.33281   0.229 0.818700
## RCETHOther or Unknown     1.34565    0.90584   1.486 0.138171
## RCETHURM                  0.41225    0.38367   1.074 0.283233
## TRAN                      0.02778    0.30411   0.091 0.927254
## GRK                       0.69644    0.52304   1.332 0.183760
## STAB                      0.14779    0.57598   0.257 0.797625
## GT1                       1.45714    0.50208   2.902 0.003906 **
## GT2                      -0.65883    0.51240  -1.286 0.199254
## MAJREVComputer Eng        1.42273    0.82538   1.724 0.085511 .
## MAJREVComputer Sci        2.02821    0.78309   2.590 0.009939 **
## UROP                     -0.11038    0.43303  -0.255 0.798921
## COOP1 Some CoOp           0.22794    0.61482   0.371 0.711024
## COOP3 CoOpDegDesig        2.29006    1.06439   2.152 0.032019 *
## INT12                     0.59632    0.38636   1.543 0.123498
## GPA                       0.95246    0.27267   3.493 0.000529 ***
## YR2018                    0.77008    0.48524   1.587 0.113279
## YR2019                   -0.30016    0.42703  -0.703 0.482512
## YR2020                   -0.33068    0.43409  -0.762 0.446632
## YR2021                    0.20229    0.49503   0.409 0.683016
## YR2022                   -0.28728    0.44520  -0.645 0.519112
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 0.9172511)
##
## Number of Fisher Scoring iterations: 5
```

```
##                          Estimate Std. Error t value Pr(>|t|)   2.5 % 97.5 %
## (Intercept)               -1.3178     0.8806 -1.4966   0.1352 -3.0485 0.4128
## VIPSEM                     0.5811     0.1706  3.4058   0.0007  0.2458 0.9165
## CITZResident NonCitizen   -0.4795     0.3796 -1.2634   0.2071 -1.2255 0.2665
```

```
## FEMALE                   0.9985     0.3848  2.5951   0.0098  0.2423 1.7547
## RCETHAsian               0.0763     0.3328  0.2294   0.8187 -0.5778 0.7305
## RCETHOther or Unknown    1.3457     0.9058  1.4855   0.1381 -0.4347 3.1260
## RCETHURM                 0.4122     0.3837  1.0745   0.2832 -0.3418 1.1663
## TRAN                     0.0278     0.3041  0.0914   0.9273 -0.5699 0.6255
## GRK                      0.6964     0.5230  1.3315   0.1837 -0.3316 1.7244
## STAB                     0.1478     0.5760  0.2566   0.7976 -0.9843 1.2799
## GT1                      1.4571     0.5021  2.9022   0.0039  0.4703 2.4440
## GT2                     -0.6588     0.5124 -1.2858   0.1992 -1.6659 0.3483
## MAJREVComputer Eng       1.4227     0.8254  1.7237   0.0855 -0.1995 3.0450
## MAJREVComputer Sci       2.0282     0.7831  2.5900   0.0099  0.4891 3.5674
## UROP                    -0.1104     0.4330 -0.2549   0.7989 -0.9615 0.7407
## COOP1 Some CoOp          0.2279     0.6148  0.3707   0.7110 -0.9805 1.4363
## COOP3 CoOpDegDesig       2.2901     1.0644  2.1515   0.0320  0.1980 4.3821
## INT12                    0.5963     0.3864  1.5434   0.1235 -0.1631 1.3557
## GPA                      0.9525     0.2727  3.4931   0.0005  0.4165 1.4884
## YR2018                   0.7701     0.4852  1.5870   0.1132 -0.1836 1.7238
## YR2019                  -0.3002     0.4270 -0.7029   0.4825 -1.1395 0.5392
## YR2020                  -0.3307     0.4341 -0.7618   0.4466 -1.1839 0.5225
## YR2021                   0.2023     0.4950  0.4086   0.6830 -0.7707 1.1753
## YR2022                  -0.2873     0.4452 -0.6453   0.5191 -1.1623 0.5878
```

Adjusted Odds Ratios with Confidence Intervals

```
##                          AOR 2.5 % 97.5 %
## VIPSEM                 1.788 1.279  2.501
## CITZResident NonCitizen 0.619 0.294  1.306
## FEMALE                 2.714 1.274  5.783
## RCETHAsian             1.079 0.561  2.076
## RCETHOther or Unknown  3.841 0.647 22.790
## RCETHURM               1.510 0.710  3.211
## TRAN                   1.028 0.566  1.869
## GRK                    2.007 0.718  5.610
## STAB                   1.159 0.374  3.597
## GT1                    4.294 1.600 11.520
## GT2                    0.517 0.189  1.417
## MAJREVComputer Eng     4.148 0.819 21.015
## MAJREVComputer Sci     7.600 1.630 35.431
## UROP                   0.895 0.382  2.098
## COOP1 Some CoOp        1.256 0.375  4.206
## COOP3 CoOpDegDesig     9.876 1.219 80.031
## INT12                  1.815 0.849  3.880
## GPA                    2.592 1.517  4.430
## YR2018                 2.160 0.832  5.607
## YR2019                 0.741 0.320  1.715
## YR2020                 0.718 0.306  1.686
## YR2021                 1.224 0.463  3.239
## YR2022                 0.750 0.313  1.800
```

5 Majors - Non-White

jsk

Subset Statements

```
Data <- subset(Data, CITZ != "Alien, Non-Resident*")
Data <- subset(Data, VIPSEM %in% c(0,1,2,3)) #TREATMENT LEVELS

Data <- subset(Data, MAJREV %in% c("Computer Science",
"Environmental Engineering",
"Computer Engineering",
"Electrical Engineering",
"Aerospace Engineering"))

# Shorten Major Names after subsetting by major

Data$MAJREV[Data$MAJREV=="Computer Science"] <- "Computer Sci"
Data$MAJREV[Data$MAJREV=="Environmental Engineering"] <- "0 Env Eng"
Data$MAJREV[Data$MAJREV=="Computer Engineering"] <- "Computer Eng"
Data$MAJREV[Data$MAJREV=="Electrical Engineering"] <- "Electrical Eng"
Data$MAJREV[Data$MAJREV=="Mechanical Engineering"] <- "Mechanical Eng"
Data$MAJREV[Data$MAJREV=="Biomedical Engineering"] <- "Biomedical Eng"
Data$MAJREV[Data$MAJREV=="Materials Science & Engr"] <- "Mat Science & Eng"
Data$MAJREV[Data$MAJREV=="Civil Engineering"] <- "Civil Eng"
Data$MAJREV[Data$MAJREV=="Aerospace Engineering"] <- "Aerospace Eng"
Data$MAJREV[Data$MAJREV=="Industrial Engineering"] <- "Industrial Eng"
Data$MAJREV[Data$MAJREV=="Chemical and Biomolecular Eng"] <- "Chem & Biomolec Eng"

Data <- subset(Data, RCETH != "White")

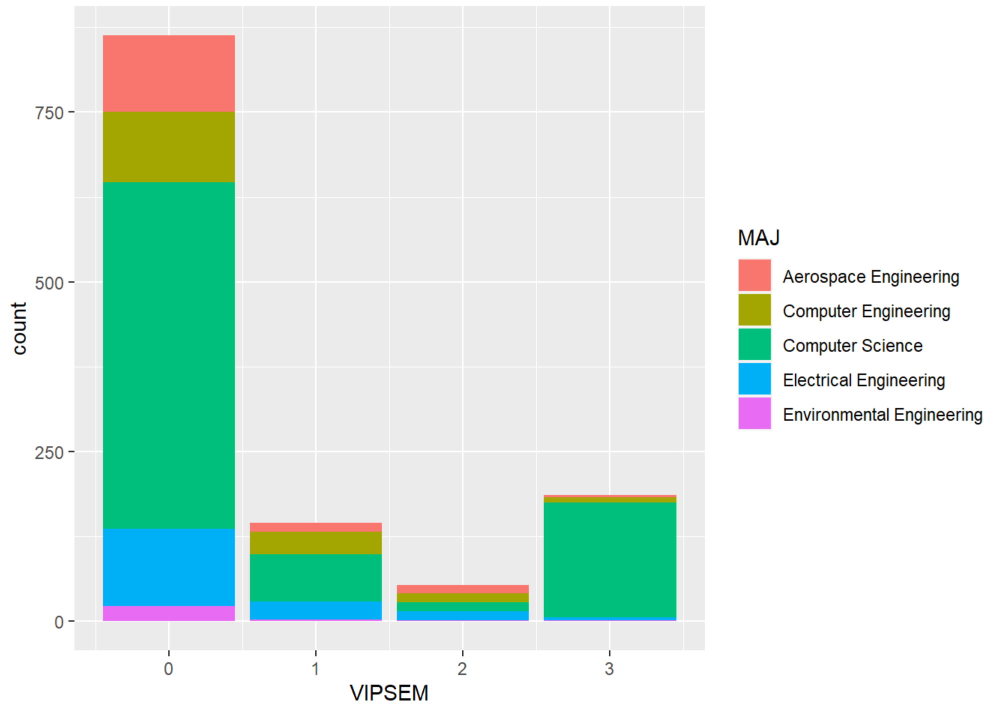
# Set ref category to "other or Unknown"
Data$RCETH[Data$RCETH == "Two or more"] <- "0 Other or Unknown"
Data$RCETH[Data$RCETH == "American Indian or Alaska Native"] <- "0 Other or Unknown"
Data$RCETH[Data$RCETH == "Native Hawaiian or Other Pacific Islander"] <- "0 Other or Unknown"
Data$RCETH[Data$RCETH == "Unknown"] <- "0 Other or Unknown"
```

GPA - Center around Grand Mean for 5 Majors Grouping

```
Data$GPA.raw <- Data$GPA
Data$GPA.from.5maj.Mean <- Data$GPA-3.586502
Data$GPA <- Data$GPA.from.5maj.Mean
```

Fequency Tables

```
##
##                    0   1   2   3
##   0 Env Eng       23   3   2   2
##   Aerospace Eng  112  13  12   3
##   Computer Eng   104  33  13   8
##   Computer Sci   511  70  13 169
##   Electrical Eng 113  26  13   4
```



Propensity Score Model

```
## VIPSEM ~ CITZ + FEMALE + RCETH + PELL + TRAN + GRK + STAB + GT1 +
##     LLHON + MAJREV + UROP + COOP + INT12 + GPA
```

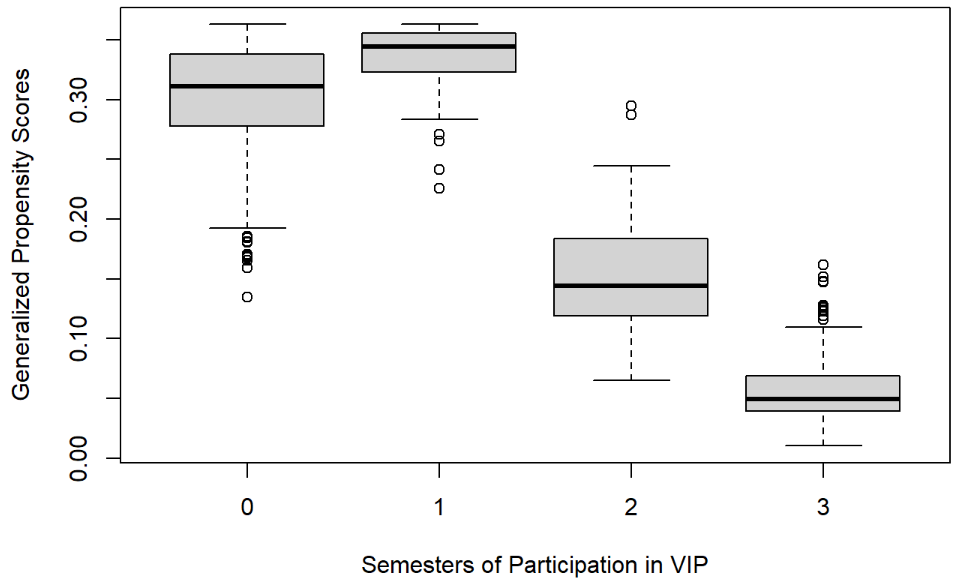
Balance Table

variable	coefBaseline	coefIPW
CITZ	0.091	0.022
FEMALE	0.215	0.017
RCETH	0.156	0.020
PELL	0.172	0.002
TRAN	0.168	0.037
GRK	0.045	0.017
STAB	0.035	0.072
GT1	0.103	0.015
LLHON	0.521	0.020
MAJREV	0.325	0.127
UROP	0.137	0.039
COOP	0.195	0.098
INT12	0.166	0.009
GPA	0.244	0.024

Sample Sizes

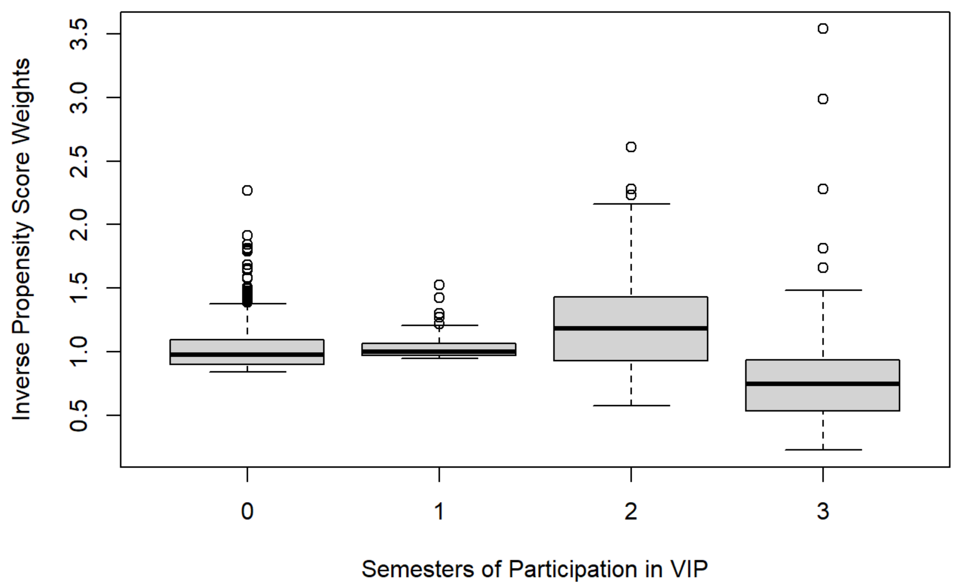
Generalized Propensity Scores

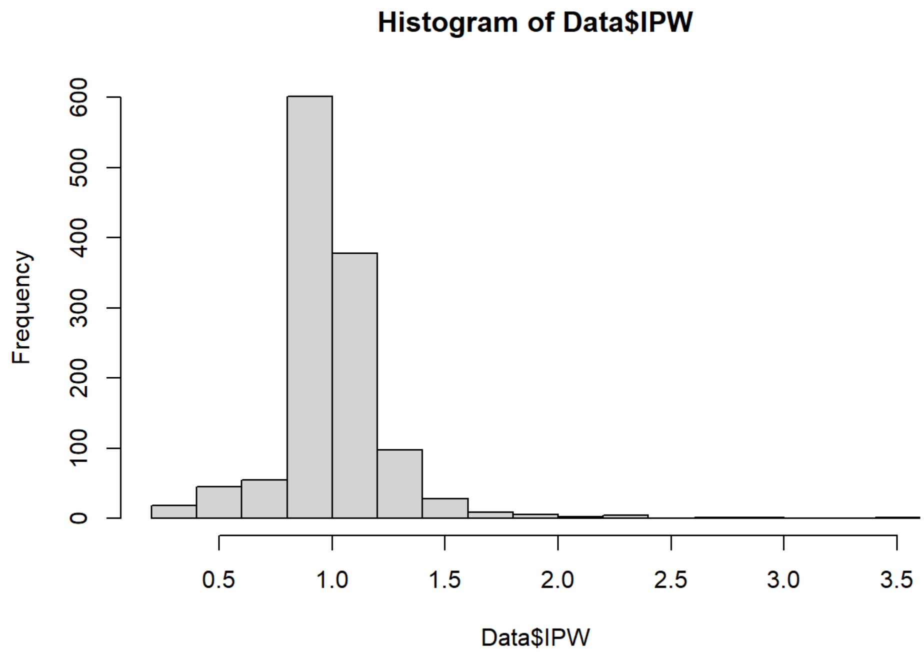
	0 sem	1 sem	2 sem	3 sem
Min.	0.134	0.226	0.065	0.010
1st Qu.	0.278	0.323	0.119	0.040
Median	0.311	0.345	0.144	0.050
Mean	0.305	0.336	0.155	0.057
3rd Qu.	0.338	0.356	0.184	0.069
Max.	0.363	0.363	0.294	0.162



Inverse Propensity Score Weights

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.2279  0.8854  0.9759  0.9981  1.0857  3.5377
```



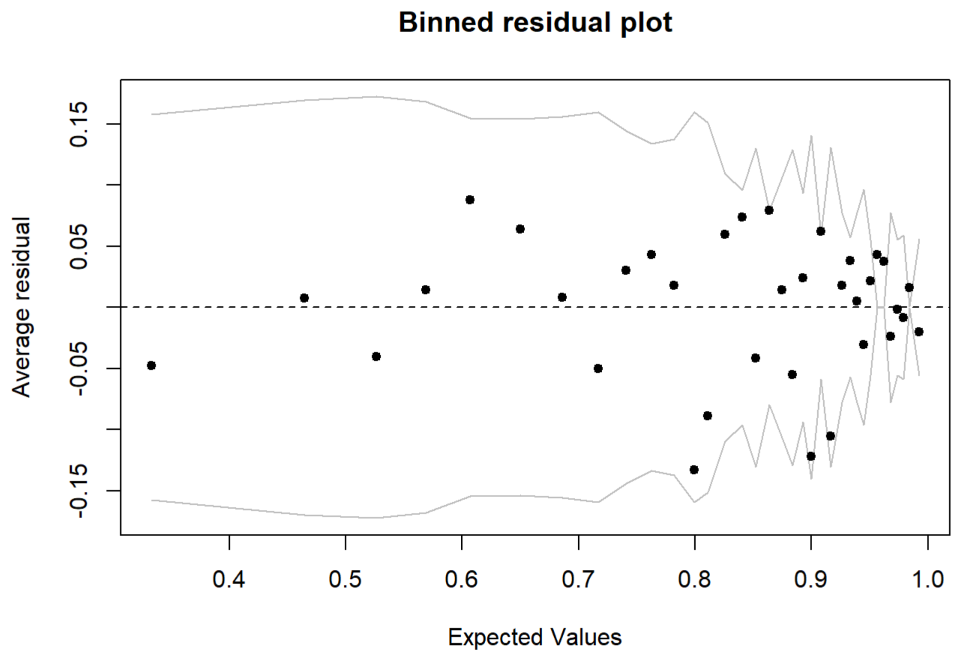


	0 sem	1 sem	2 sem	3 sem
Min.	0.840	0.950	0.579	0.228
1st Qu.	0.901	0.970	0.928	0.536
Median	0.979	1.001	1.181	0.745
Mean	1.022	1.034	1.232	0.793
3rd Qu.	1.096	1.067	1.434	0.931
Max.	2.268	1.526	2.609	3.538

REGRESSION

Residuals

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -3.2226  0.1833  0.4041  0.1453  0.6369  1.8367
```




```
##              R-Squared      Adjusted R-Squared
##                  0.170                   0.153
```

Regression Results

```
##
## Call:
## svyglm(formula = ..1, design = ..2, family = ..3)
##
## Survey design:
## survey::svydesign(...)
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)              0.907316   0.671110   1.352 0.176637
## VIPSEM                   0.374931   0.103345   3.628 0.000297 ***
## CITZResident NonCitizen  0.104349   0.259347   0.402 0.687494
## FEMALE                   0.495910   0.224912   2.205 0.027646 *
## RCETHAsian              -0.070196   0.280499  -0.250 0.802432
## RCETHURM                 0.078342   0.293058   0.267 0.789264
## PELL                    -0.477085   0.178811  -2.668 0.007729 **
## TRAN                    -0.418335   0.207274  -2.018 0.043782 *
## GRK                      0.564558   0.289125   1.953 0.051090 .
## STAB                     0.415094   0.249598   1.663 0.096559 .
## GT1                     -0.138821   0.233469  -0.595 0.552221
## LLHON                    0.670894   0.652348   1.028 0.303952
## MAJREVAerospace Eng     -0.551245   0.633018  -0.871 0.384023
## MAJREVComputer Eng       0.366708   0.636838   0.576 0.564839
## MAJREVComputer Sci       0.720414   0.610314   1.180 0.238072
## MAJREVElectrical Eng     0.002457   0.637277   0.004 0.996924
## UROP                     0.188081   0.234921   0.801 0.423511
## COOP1 Some CoOp          0.403114   0.348252   1.158 0.247279
## COOP3 CoOpDegDesig       1.393224   0.534750   2.605 0.009289 **
## INT12                    0.867828   0.244171   3.554 0.000394 ***
## GPA                      0.989732   0.175800   5.630 2.24e-08 ***
## YR2018                   0.651835   0.311993   2.089 0.036891 *
## YR2019                  -0.128133   0.276733  -0.463 0.643433
## YR2020                  -0.258345   0.293197  -0.881 0.378420
## YR2021                   0.188769   0.300141   0.629 0.529510
## YR2022                  -0.223720   0.285833  -0.783 0.433959
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 1.199333)
##
## Number of Fisher Scoring iterations: 5
```

```
##                         Estimate Std. Error t value Pr(>|t|)   2.5 %  97.5 %
## (Intercept)               0.9073     0.6711  1.3520   0.1766 -0.4093  2.2239
```

```
## VIPSEM                   0.3749     0.1033  3.6279   0.0003  0.1722  0.5777
## CITZResident NonCitizen  0.1043     0.2593  0.4024   0.6875 -0.4045  0.6132
## FEMALE                   0.4959     0.2249  2.2049   0.0276  0.0547  0.9372
## RCETHAsian              -0.0702     0.2805 -0.2503   0.8024 -0.6205  0.4801
## RCETHURM                 0.0783     0.2931  0.2673   0.7893 -0.4966  0.6533
## PELL                    -0.4771     0.1788 -2.6681   0.0077 -0.8279 -0.1263
## TRAN                    -0.4183     0.2073 -2.0183   0.0438 -0.8250 -0.0117
## GRK                      0.5646     0.2891  1.9526   0.0511 -0.0027  1.1318
## STAB                     0.4151     0.2496  1.6630   0.0966 -0.0746  0.9048
## GT1                     -0.1388     0.2335 -0.5946   0.5522 -0.5969  0.3192
## LLHON                    0.6709     0.6523  1.0284   0.3039 -0.6089  1.9507
## MAJREVAerospace Eng     -0.5512     0.6330 -0.8708   0.3840 -1.7931  0.6907
## MAJREVComputer Eng       0.3667     0.6368  0.5758   0.5648 -0.8827  1.6161
## MAJREVComputer Sci       0.7204     0.6103  1.1804   0.2381 -0.4769  1.9178
## MAJREVElectrical Eng     0.0025     0.6373  0.0039   0.9969 -1.2478  1.2527
## UROP                     0.1881     0.2349  0.8006   0.4235 -0.2728  0.6490
## COOP1 Some CoOp          0.4031     0.3483  1.1575   0.2473 -0.2801  1.0863
## COOP3 CoOpDegDesig       1.3932     0.5347  2.6054   0.0093  0.3441  2.4423
## INT12                    0.8678     0.2442  3.5542   0.0004  0.3888  1.3469
## GPA                      0.9897     0.1758  5.6299   0.0000  0.6448  1.3346
## YR2018                   0.6518     0.3120  2.0893   0.0369  0.0397  1.2639
## YR2019                  -0.1281     0.2767 -0.4630   0.6434 -0.6710  0.4148
## YR2020                  -0.2583     0.2932 -0.8811   0.3784 -0.8336  0.3169
## YR2021                   0.1888     0.3001  0.6289   0.5295 -0.4001  0.7776
## YR2022                  -0.2237     0.2858 -0.7827   0.4340 -0.7845  0.3370
```

Adjusted Odds Ratios with Confidence Intervals

```
##                           AOR 2.5 % 97.5 %
## VIPSEM                  1.455 1.188  1.782
## CITZResident NonCitizen 1.110 0.667  1.846
## FEMALE                  1.642 1.056  2.553
## RCETHAsian              0.932 0.538  1.616
## RCETHURM                1.081 0.609  1.922
## PELL                    0.621 0.437  0.881
## TRAN                    0.658 0.438  0.988
## GRK                     1.759 0.997  3.101
## STAB                    1.515 0.928  2.471
## GT1                     0.870 0.551  1.376
## LLHON                   1.956 0.544  7.034
## MAJREVAerospace Eng     0.576 0.166  1.995
## MAJREVComputer Eng      1.443 0.414  5.034
## MAJREVComputer Sci      2.055 0.621  6.806
## MAJREVElectrical Eng    1.002 0.287  3.500
## UROP                    1.207 0.761  1.914
## COOP1 Some CoOp         1.496 0.756  2.963
## COOP3 CoOpDegDesig      4.028 1.411 11.500
## INT12                   2.382 1.475  3.845
## GPA                     2.691 1.906  3.799
## YR2018                  1.919 1.041  3.539
## YR2019                  0.880 0.511  1.514
## YR2020                  0.772 0.434  1.373
## YR2021                  1.208 0.670  2.176
## YR2022                  0.800 0.456  1.401
```

5 Majors - White

jsk

Subset Statements

```
Data <- subset(Data, CITZ != "Alien, Non-Resident*")
Data <- subset(Data, VIPSEM %in% c(0,1,2,3)) #TREATMENT LEVELS

Data <- subset(Data, MAJREV %in% c("Computer Science",
"Environmental Engineering",
"Computer Engineering",
"Electrical Engineering",
"Aerospace Engineering"))

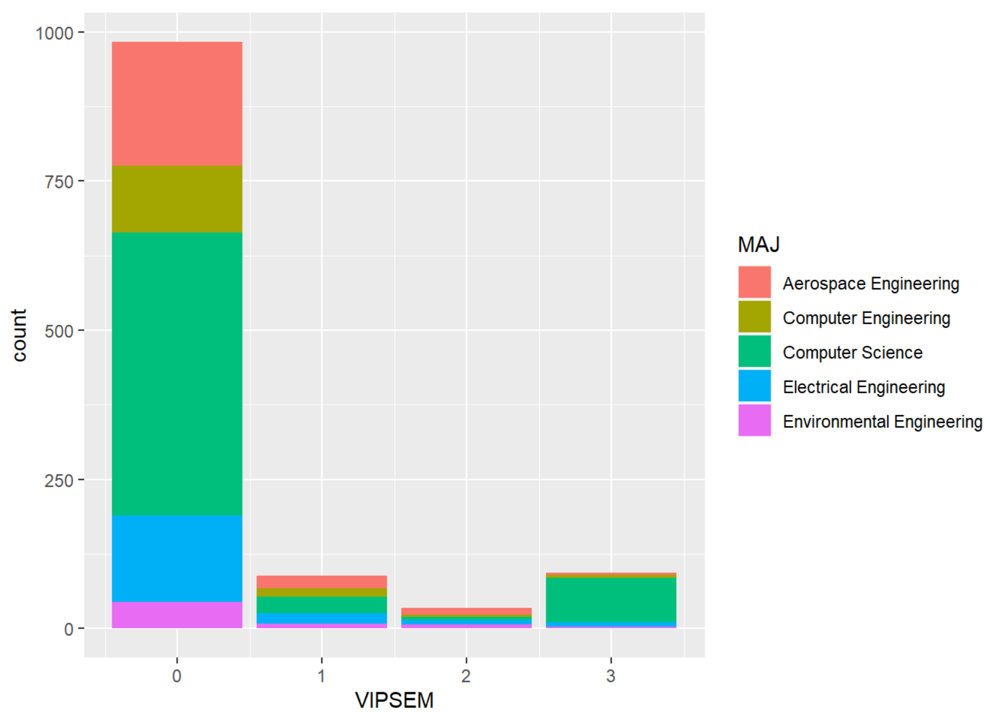
# Shorten Major Names after subsetting by major

Data$MAJREV[Data$MAJREV=="Computer Science"] <- "Computer Sci"
Data$MAJREV[Data$MAJREV=="Environmental Engineering"] <- "0 Env Eng"
Data$MAJREV[Data$MAJREV=="Computer Engineering"] <- "Computer Eng"
Data$MAJREV[Data$MAJREV=="Electrical Engineering"] <- "Electrical Eng"
Data$MAJREV[Data$MAJREV=="Mechanical Engineering"] <- "Mechanical Eng"
Data$MAJREV[Data$MAJREV=="Biomedical Engineering"] <- "Biomedical Eng"
Data$MAJREV[Data$MAJREV=="Materials Science & Engr"] <- "Mat Science & Eng"
Data$MAJREV[Data$MAJREV=="Civil Engineering"] <- "Civil Eng"
Data$MAJREV[Data$MAJREV=="Aerospace Engineering"] <- "Aerospace Eng"
Data$MAJREV[Data$MAJREV=="Industrial Engineering"] <- "Industrial Eng"
Data$MAJREV[Data$MAJREV=="Chemical and Biomolecular Eng"] <- "Chem & Biomolec Eng"

Data <- subset(Data, RCETH == "White")
```

Fequency Tables

```
##
##                    0   1   2   3
##   0 Env Eng       45   8   7   3
##   Aerospace Eng  208  21  12   3
##   Computer Eng   112  14   4   5
##   Computer Sci   474  27   5  75
##   Electrical Eng 144  18   7   7
```



Propensity Score Model

```
## VIPSEM ~ FEMALE + PELL + TRAN + GRK + STAB + GT1 + LLHON + MAJREV +
##     UROP + COOP + INT12 + GPA
```

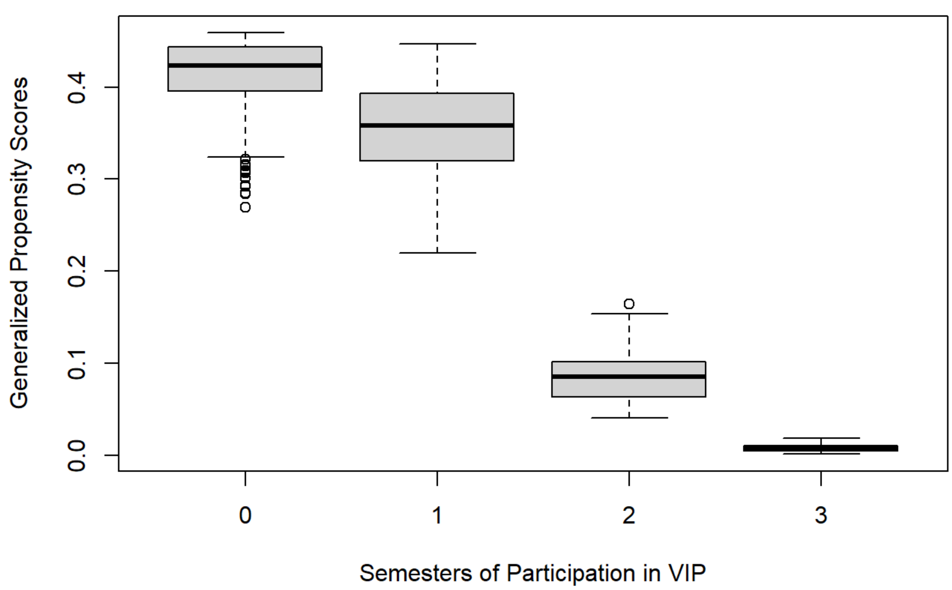
Balance Table

variable	coefBaseline	coefIPW
FEMALE	0.235	0.044
PELL	0.120	0.085
TRAN	0.130	0.038
GRK	0.024	0.067
STAB	0.088	0.015
GT1	0.043	0.045
LLHON	0.143	0.036
MAJREV	0.311	0.112
UROP	0.009	0.004
COOP	0.208	0.034
INT12	0.233	0.010
GPA	0.240	0.032

Sample Sizes

Generalized Propensity Scores

	0 sem	1 sem	2 sem	3 sem
Min.	0.269	0.220	0.040	0.001
1st Qu.	0.396	0.320	0.064	0.005
Median	0.424	0.358	0.085	0.007
Mean	0.416	0.354	0.088	0.008
3rd Qu.	0.444	0.393	0.102	0.011
Max.	0.459	0.447	0.164	0.019



5 Majors - Female

jsk

Subset Statements

```
Data <- subset(Data, CITZ != "Alien, Non-Resident*")
Data <- subset(Data, VIPSEM %in% c(0,1,2,3)) #TREATMENT LEVELS

Data <- subset(Data, FEMALE == 1)

Data <- subset(Data, MAJREV %in% c("Computer Science",
"Environmental Engineering",
"Computer Engineering",
"Electrical Engineering",
"Aerospace Engineering"))

# Data <- subset(Data, COOP == "0 No CoOp")
# Data <- subset(Data, INT12 == 0)

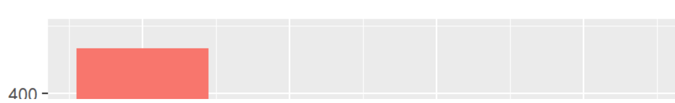
Data$MAJREV[Data$MAJREV=="Computer Science"] <- "Computer Sci"
Data$MAJREV[Data$MAJREV=="Environmental Engineering"] <- "0 Env Eng"
Data$MAJREV[Data$MAJREV=="Computer Engineering"] <- "Computer Eng"
Data$MAJREV[Data$MAJREV=="Electrical Engineering"] <- "Electrical Eng"
Data$MAJREV[Data$MAJREV=="Mechanical Engineering"] <- "Mechanical Eng"
Data$MAJREV[Data$MAJREV=="Biomedical Engineering"] <- "Biomedical Eng"
Data$MAJREV[Data$MAJREV=="Materials Science & Engr"] <- "Mat Science & Eng"
Data$MAJREV[Data$MAJREV=="Civil Engineering"] <- "Civil Eng"
Data$MAJREV[Data$MAJREV=="Aerospace Engineering"] <- "Aerospace Eng"
Data$MAJREV[Data$MAJREV=="Industrial Engineering"] <- "Industrial Eng"
Data$MAJREV[Data$MAJREV=="Chemical and Biomolecular Eng"] <- "Chem & Biomolec Eng"
```

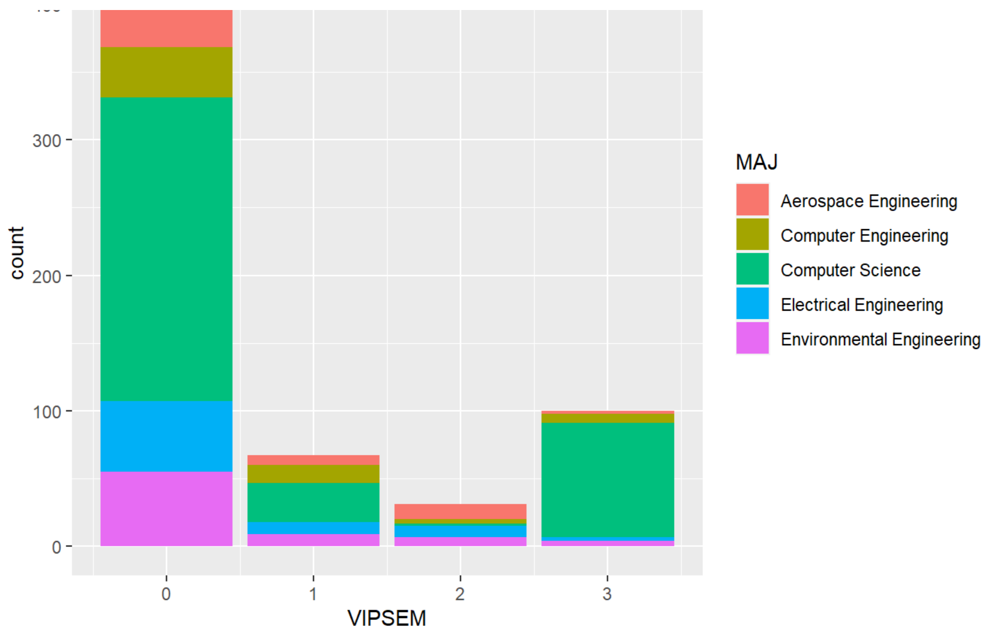
GPA - Center around Grand Mean for 5 Majors Grouping

```
Data$GPA.raw <- Data$GPA
Data$GPA.from.5maj.Mean <- Data$GPA-3.586502
Data$GPA <- Data$GPA.from.5maj.Mean
```

Fequency Tables

```
##
##                    0   1   2   3
##   0 Env Eng       55   9   7   4
##   Aerospace Eng   65   7  11   2
##   Computer Eng    37  13   3   7
##   Computer Sci   224  29   2  84
##   Electrical Eng  52   9   8   3
```





Propensity Score Model

```
## VIPSEM ~ CITZ + RCETH + PELL + TRAN + GRK + STAB + GT1 + LLHON +
##     MAJREV + UROP + COOP + INT12 + GPA
```

Balance Table

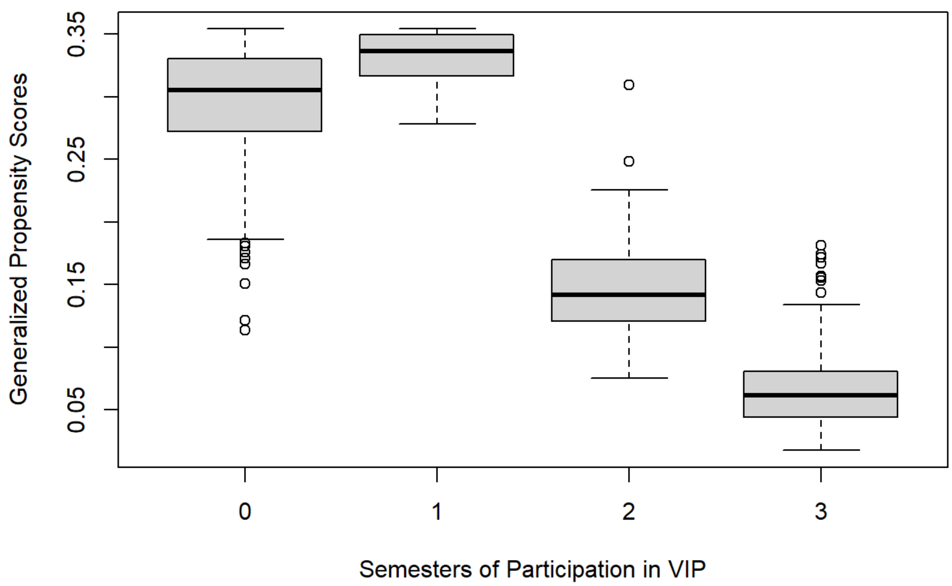
variable	coefBaseline	coefIPW
CITZ	0.065	0.055
RCETH	0.354	0.013
PELL	0.082	0.012
TRAN	0.167	0.041
GRK	0.046	0.025
STAB	0.012	0.014
GT1	0.020	0.041
LLHON	0.589	0.066
MAJREV	0.332	0.097
UROP	0.145	0.068
COOP	0.319	0.103
INT12	0.122	0.048
GPA	0.243	0.002

Sample Sizes

Generalized Propensity Scores

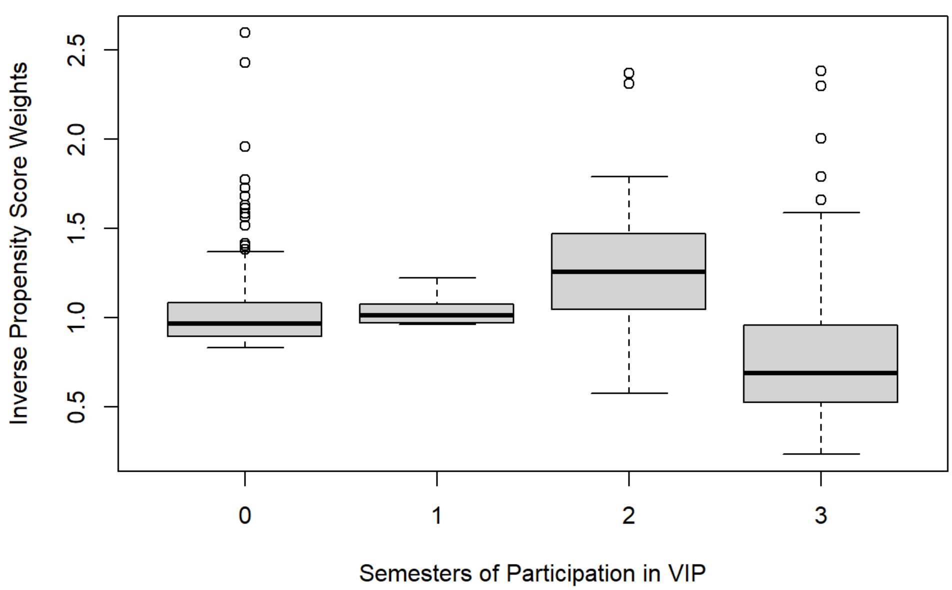
	0 sem	1 sem	2 sem	3 sem
Min.	0.114	0.278	0.075	0.018
1st Qu.	0.272	0.317	0.121	0.044
Median	0.306	0.336	0.142	0.062

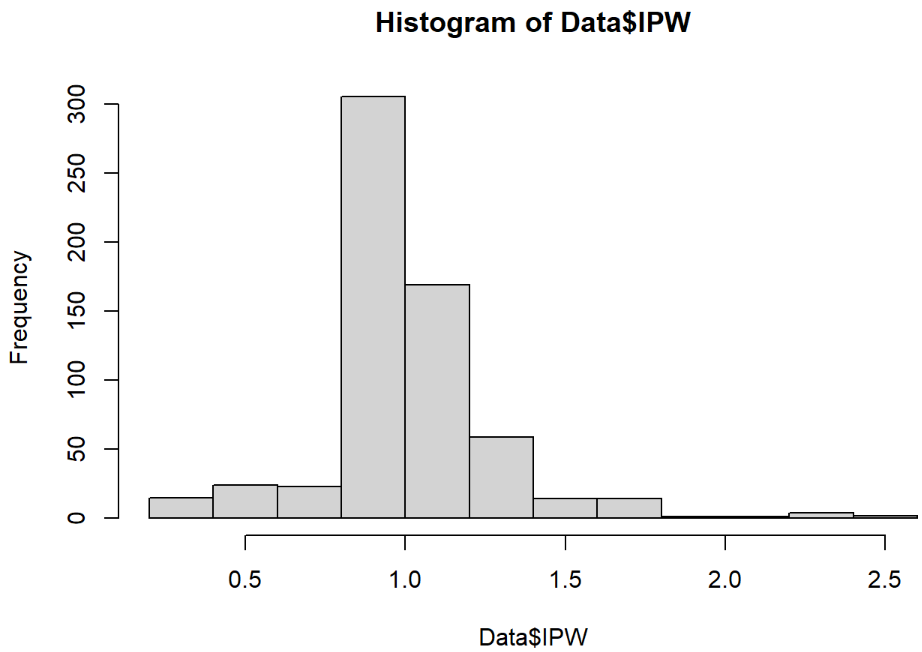
	0 sem	1 sem	2 sem	3 sem
Mean	0.297	0.333	0.153	0.069
3rd Qu.	0.330	0.350	0.170	0.081
Max.	0.354	0.354	0.310	0.182



Inverse Propensity Score Weights

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.2338  0.8845  0.9686  1.0005  1.0864  2.5963
```



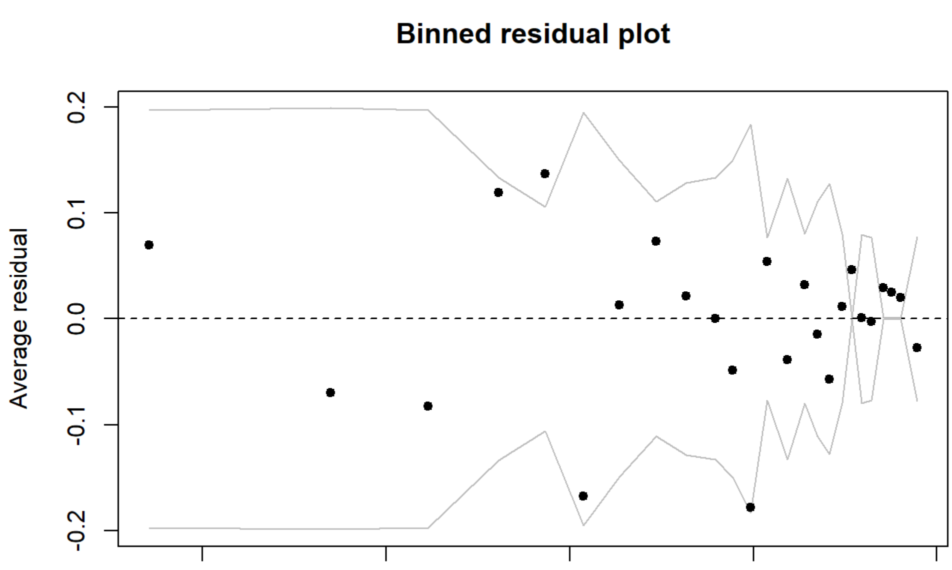


	0 sem	1 sem	2 sem	3 sem
Min.	0.834	0.960	0.575	0.234
1st Qu.	0.893	0.972	1.048	0.526
Median	0.966	1.012	1.258	0.688
Mean	1.021	1.026	1.284	0.805
3rd Qu.	1.084	1.074	1.471	0.958
Max.	2.596	1.223	2.369	2.382

REGRESSION

Residuals

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -3.1863  0.2317  0.3838  0.1657  0.5491  1.2394
```



0.6 0.7 0.8 0.9 1.0

Expected Values

```
##              R-Squared      Adjusted R-Squared
##                  0.116                   0.080
```

Regression Results

```
##
## Call:
## svyglm(formula = ..1, design = ..2, family = ..3)
##
## Survey design:
## survey::svydesign(...)
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)               0.47076    0.64218   0.733  0.46380
## VIPSEM                    0.28214    0.15141   1.863  0.06288 .
## CITZResident NonCitizen  0.33167    0.56502   0.587  0.55742
## RCETHAsian                0.13399    0.36993   0.362  0.71733
## RCETHOther or Unknown     0.32326    0.55030   0.587  0.55714
## RCETHURM                  0.40278    0.40024   1.006  0.31465
## PELL                     -0.15536    0.28598  -0.543  0.58715
## TRAN                     -0.23117    0.37665  -0.614  0.53962
## GRK                       0.56248    0.37990   1.481  0.13923
## STAB                      0.37003    0.30302   1.221  0.22251
## GT1                       0.18103    0.32440   0.558  0.57701
## LLHON                     0.69635    0.74219   0.938  0.34850
## MAJREVAerospace Eng       0.17395    0.45080   0.386  0.69972
## MAJREVComputer Eng        0.94147    0.56763   1.659  0.09772 .
## MAJREVComputer Sci        1.20202    0.42201   2.848  0.00454 **
## MAJREVElectrical Eng      0.34924    0.47344   0.738  0.46100
## UROP                     -0.01957    0.31766  -0.062  0.95089
## COOP1 Some CoOp           0.05871    0.49936   0.118  0.90644
## COOP3 CoOpDegDesig        0.85646    0.60382   1.418  0.15659
## INT12                     0.62554    0.35109   1.782  0.07530 .
## GPA                       0.84808    0.28180   3.009  0.00273 **
## YR2018                    0.57682    0.51237   1.126  0.26070
## YR2019                    0.62376    0.51936   1.201  0.23021
## YR2020                   -0.59684    0.43054  -1.386  0.16618
## YR2021                    0.03149    0.46675   0.067  0.94624
## YR2022                    0.02177    0.44096   0.049  0.96064
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 1.161694)
##
## Number of Fisher Scoring iterations: 5
```

```
##                          Estimate Std. Error t value Pr(>|t|)   2.5 % 97.5 %
## (Intercept)                0.4708     0.6422  0.7331   0.4638 -0.7903 1.7318
## VIPSEM                     0.2821     0.1514  1.8634   0.0629 -0.0152 0.5795
## CITZResident NonCitizen    0.3317     0.5650  0.5870   0.5574 -0.7779 1.4412
## RCETHAsian                 0.1340     0.3699  0.3622   0.7173 -0.5925 0.8604
## RCETHOther or Unknown      0.3233     0.5503  0.5874   0.5571 -0.7574 1.4039
## RCETHURM                   0.4028     0.4002  1.0063   0.3146 -0.3832 1.1887
## PELL                      -0.1554     0.2860 -0.5433   0.5871 -0.7169 0.4062
## TRAN                      -0.2312     0.3767 -0.6137   0.5396 -0.9708 0.5085
## GRK                        0.5625     0.3799  1.4806   0.1392 -0.1835 1.3085
## STAB                       0.3700     0.3030  1.2211   0.2225 -0.2250 0.9651
## GT1                        0.1810     0.3244  0.5581   0.5770 -0.4560 0.8181
## LLHON                      0.6964     0.7422  0.9382   0.3485 -0.7611 2.1538
## MAJREVAerospace Eng        0.1740     0.4508  0.3859   0.6997 -0.7113 1.0592
## MAJREVComputer Eng         0.9415     0.5676  1.6586   0.0977 -0.1732 2.0561
## MAJREVComputer Sci         1.2020     0.4220  2.8484   0.0045  0.3733 2.0307
## MAJREVElectrical Eng       0.3492     0.4734  0.7377   0.4610 -0.5805 1.2790
## UROP                      -0.0196     0.3177 -0.0616   0.9509 -0.6434 0.6042
## COOP1 Some CoOp            0.0587     0.4994  0.1176   0.9064 -0.9219 1.0393
## COOP3 CoOpDegDesig         0.8565     0.6038  1.4184   0.1566 -0.3293 2.0422
## INT12                      0.6255     0.3511  1.7817   0.0753 -0.0639 1.3150
## GPA                        0.8481     0.2818  3.0095   0.0027  0.2947 1.4015
## YR2018                     0.5768     0.5124  1.1258   0.2607 -0.4293 1.5830
## YR2019                     0.6238     0.5194  1.2010   0.2302 -0.3961 1.6436
## YR2020                    -0.5968     0.4305 -1.3863   0.1662 -1.4423 0.2486
## YR2021                     0.0315     0.4668  0.0675   0.9462 -0.8851 0.9481
## YR2022                     0.0218     0.4410  0.0494   0.9606 -0.8442 0.8877
```

Adjusted Odds Ratios with Confidence Intervals

```
##                           AOR 2.5 % 97.5 %
## VIPSEM                  1.326 0.985  1.785
## CITZResident NonCitizen 1.393 0.459  4.226
## RCETHAsian              1.143 0.553  2.364
## RCETHOther or Unknown   1.382 0.469  4.071
## RCETHURM                1.496 0.682  3.283
## PELL                    0.856 0.488  1.501
## TRAN                    0.794 0.379  1.663
## GRK                     1.755 0.832  3.701
## STAB                    1.448 0.798  2.625
## GT1                     1.198 0.634  2.266
## LLHON                   2.006 0.467  8.619
## MAJREVAerospace Eng     1.190 0.491  2.884
## MAJREVComputer Eng      2.564 0.841  7.816
## MAJREVComputer Sci      3.327 1.452  7.620
## MAJREVElectrical Eng    1.418 0.560  3.593
## UROP                    0.981 0.525  1.830
## COOP1 Some CoOp         1.060 0.398  2.828
## COOP3 CoOpDegDesig      2.355 0.719  7.708
## INT12                   1.869 0.938  3.725
```

```
## GPA                      2.335 1.343  4.061
## YR2018                   1.780 0.651  4.870
## YR2019                   1.866 0.673  5.174
## YR2020                   0.551 0.236  1.282
## YR2021                   1.032 0.413  2.581
## YR2022                   1.022 0.430  2.430
```

5 Majors - Pell

jsk

Subset Statements

```
Data <- subset(Data, CITZ != "Alien, Non-Resident*")
Data <- subset(Data, VIPSEM %in% c(0,1,2,3)) #TREATMENT LEVELS

Data <- subset(Data, PELL == 1)

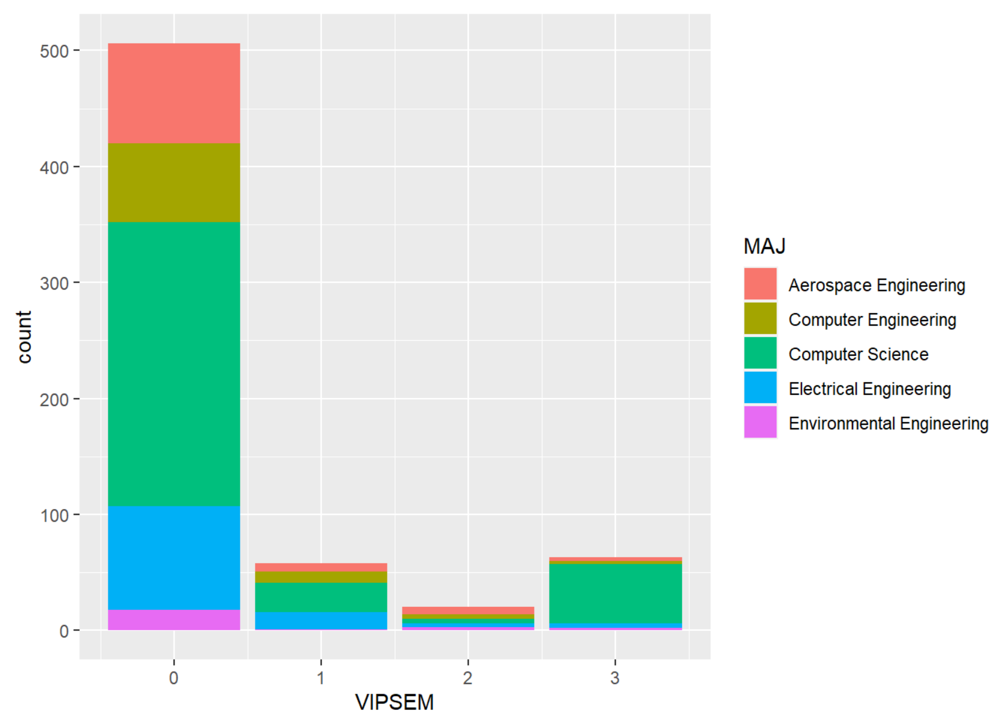
Data <- subset(Data, MAJREV %in% c("Computer Science", "Environmental Engineering", "Computer Engineering", "Electrical Engineering",

# Data <- subset(Data, COOP == "0 No CoOp")
# Data <- subset(Data, INT12 == 0)

Data$MAJREV[Data$MAJREV=="Computer Science"] <- "Computer Sci"
Data$MAJREV[Data$MAJREV=="Environmental Engineering"] <- "0 Env Eng"
Data$MAJREV[Data$MAJREV=="Computer Engineering"] <- "Computer Eng"
Data$MAJREV[Data$MAJREV=="Electrical Engineering"] <- "Electrical Eng"
Data$MAJREV[Data$MAJREV=="Mechanical Engineering"] <- "Mechanical Eng"
Data$MAJREV[Data$MAJREV=="Biomedical Engineering"] <- "Biomedical Eng"
Data$MAJREV[Data$MAJREV=="Materials Science & Engr"] <- "Mat Science & Eng"
Data$MAJREV[Data$MAJREV=="Civil Engineering"] <- "Civil Eng"
Data$MAJREV[Data$MAJREV=="Aerospace Engineering"] <- "Aerospace Eng"
Data$MAJREV[Data$MAJREV=="Industrial Engineering"] <- "Industrial Eng"
Data$MAJREV[Data$MAJREV=="Chemical and Biomolecular Eng"] <- "Chem & Biomolec Eng"
```

Fequency Tables

```
##
##                   0   1   2   3
##    0 Env Eng     18   1   3   2
##    Aerospace Eng  86   7   6   3
##    Computer Eng   68  10   4   3
##    Computer Sci  245  25   4  51
##    Electrical Eng 89  15   3   4
```



Propensity Score Model

```
## VIPSEM ~ CITZ + FEMALE + RCETH + TRAN + GRK + STAB + GT1 + GT2 +
##     MAJREV + UROP + COOP + INT12 + GPA
```

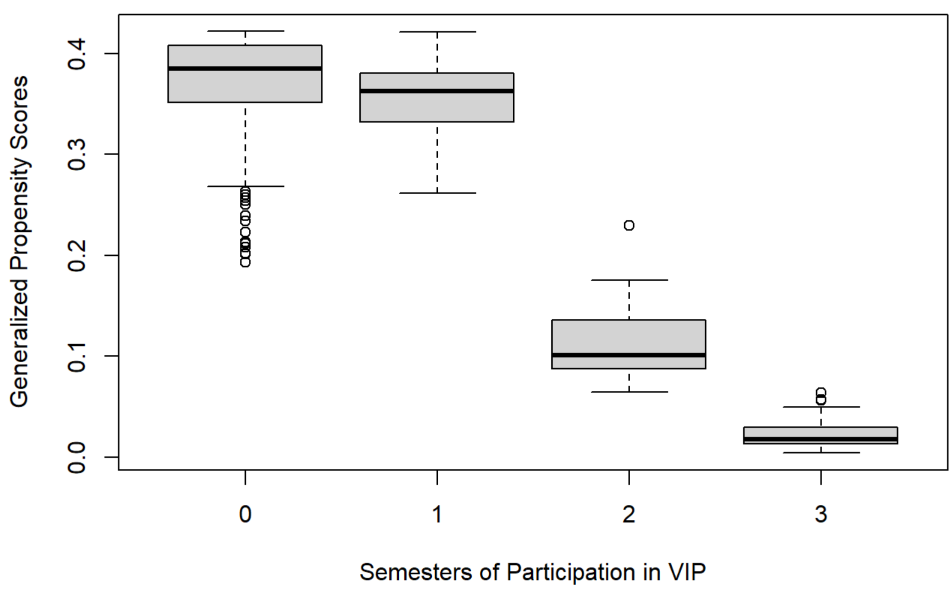
Balance Table

variable	coefBaseline	coefIPW
CITZ	0.052	0.045
FEMALE	0.224	0.003
RCETH	0.318	0.121
TRAN	0.033	0.038
GRK	0.020	0.054
STAB	0.000	0.119
GT1	0.107	0.072
GT2	0.063	0.025
MAJREV	0.283	0.046
UROP	0.043	0.012
COOP	0.282	0.198
INT12	0.348	0.029
GPA	0.164	0.019

Sample Sizes

Generalized Propensity Scores

	0 sem	1 sem	2 sem	3 sem
Min.	0.193	0.261	0.064	0.004
1st Qu.	0.352	0.333	0.089	0.013
Median	0.385	0.363	0.101	0.017
Mean	0.373	0.356	0.115	0.022
3rd Qu.	0.408	0.380	0.136	0.029
Max.	0.422	0.422	0.229	0.064



3 Majors - VIPSEM as Factor

jsk

Subset Statements

```
Data <- subset(Data, CITZ != "Alien, Non-Resident*")
Data <- subset(Data, VIPSEM %in% c(0,1,2,3)) #TREATMENT LEVELS

Data <- subset(Data, MAJREV %in% c("Computer Science",
"Environmental Engineering",
"Computer Engineering"))

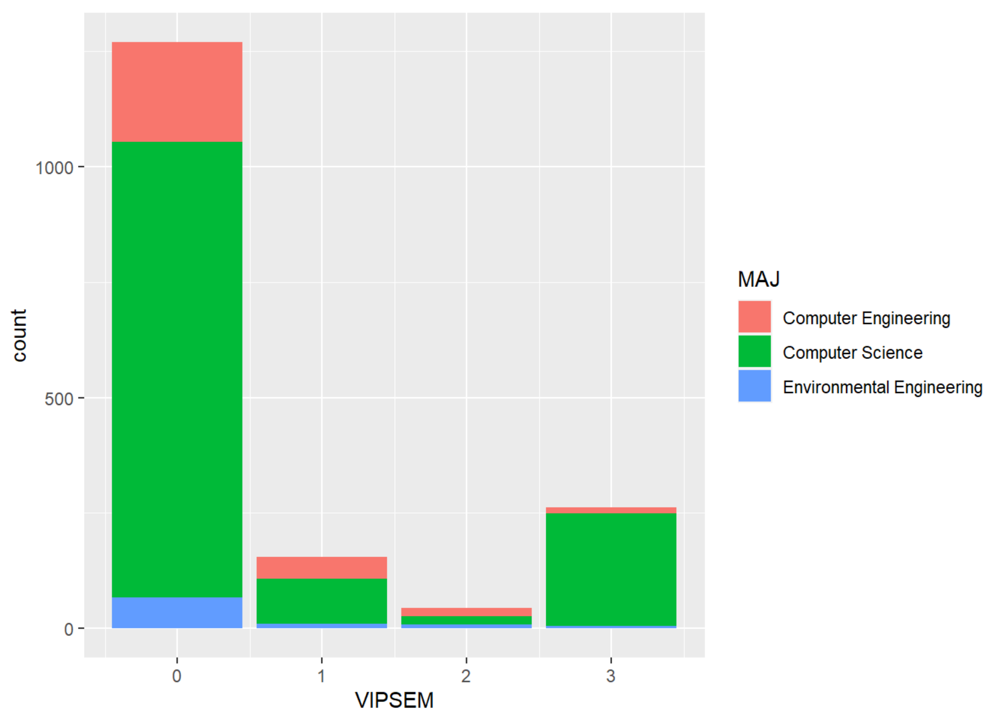
Data <- subset(Data, GPA >= 1.9)

# Shorten Major Names after subsetting by major

Data$MAJREV[Data$MAJREV=="Computer Science"] <- "Computer Sci"
Data$MAJREV[Data$MAJREV=="Environmental Engineering"] <- "0 Env Eng"
Data$MAJREV[Data$MAJREV=="Computer Engineering"] <- "Computer Eng"
Data$MAJREV[Data$MAJREV=="Electrical Engineering"] <- "Electrical Eng"
Data$MAJREV[Data$MAJREV=="Mechanical Engineering"] <- "Mechanical Eng"
Data$MAJREV[Data$MAJREV=="Biomedical Engineering"] <- "Biomedical Eng"
Data$MAJREV[Data$MAJREV=="Materials Science & Engr"] <- "Mat Science & Eng"
Data$MAJREV[Data$MAJREV=="Civil Engineering"] <- "Civil Eng"
Data$MAJREV[Data$MAJREV=="Aerospace Engineering"] <- "Aerospace Eng"
Data$MAJREV[Data$MAJREV=="Industrial Engineering"] <- "Industrial Eng"
Data$MAJREV[Data$MAJREV=="Chemical and Biomolecular Eng"] <- "Chem & Biomolec Eng"
```

Fequency Tables

```
##
##                  0   1   2   3
##   0 Env Eng     68  11   9   5
##   Computer Eng 216  47  17  13
##   Computer Sci 985  97  18 244
```



Propensity Score Model

```
## VIPSEM ~ CITZ + FEMALE + RCETH + PELL + TRAN + GRK + STAB + GT1 +
##     LLHON + MAJREV + UROP + COOP + INT12 + GPA
```

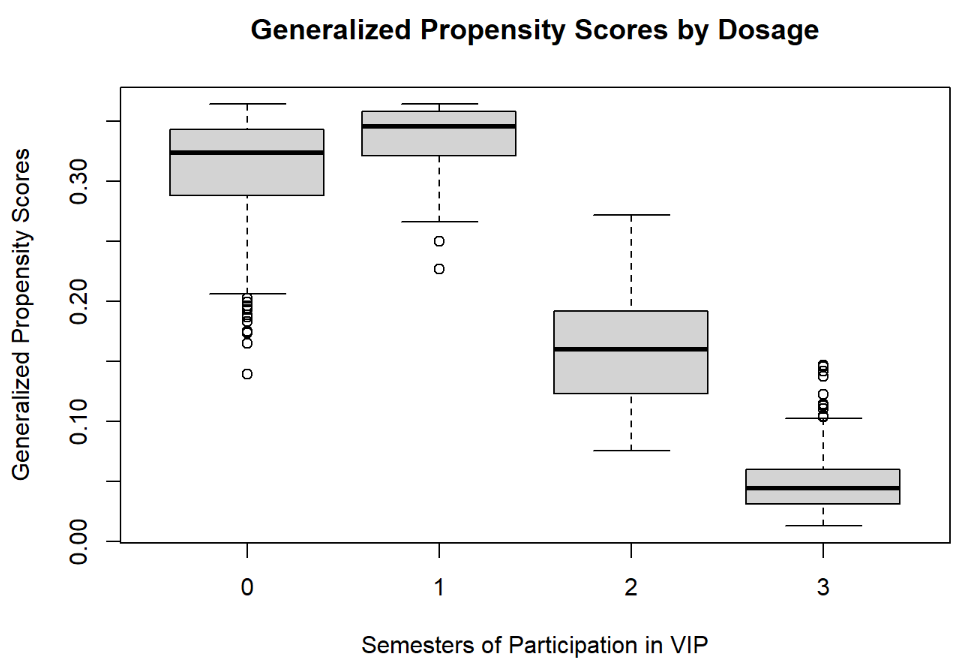
Balance Table

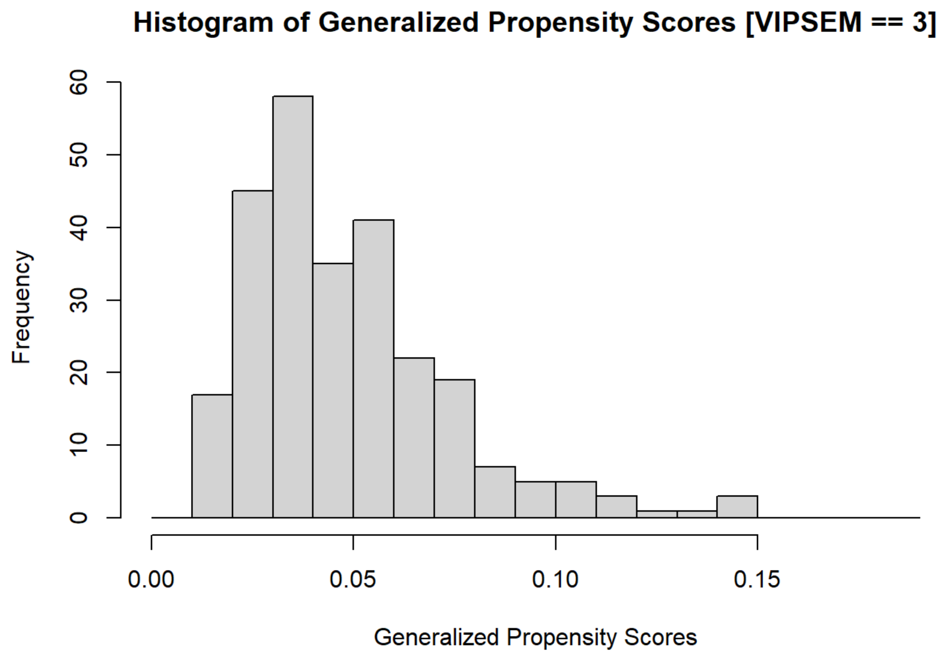
variable	coefBaseline	coefIPW
CITZ	0.001	0.021
FEMALE	0.207	0.015
RCETH	0.354	0.048
PELL	0.090	0.019
TRAN	0.093	0.014
GRK	0.031	0.022
STAB	0.010	0.012
GT1	0.012	0.005
LLHON	0.376	0.035
MAJREV	0.156	0.095
UROP	0.011	0.023
COOP	0.211	0.092
INT12	0.241	0.025
GPA	0.258	0.050

Sample Sizes

Generalized Propensity Scores

	0 sem	1 sem	2 sem	3 sem
Min.	0.140	0.227	0.076	0.013
1st Qu.	0.288	0.321	0.123	0.031
Median	0.324	0.346	0.160	0.045
Mean	0.314	0.337	0.161	0.048
3rd Qu.	0.343	0.358	0.191	0.060
Max.	0.364	0.364	0.272	0.147

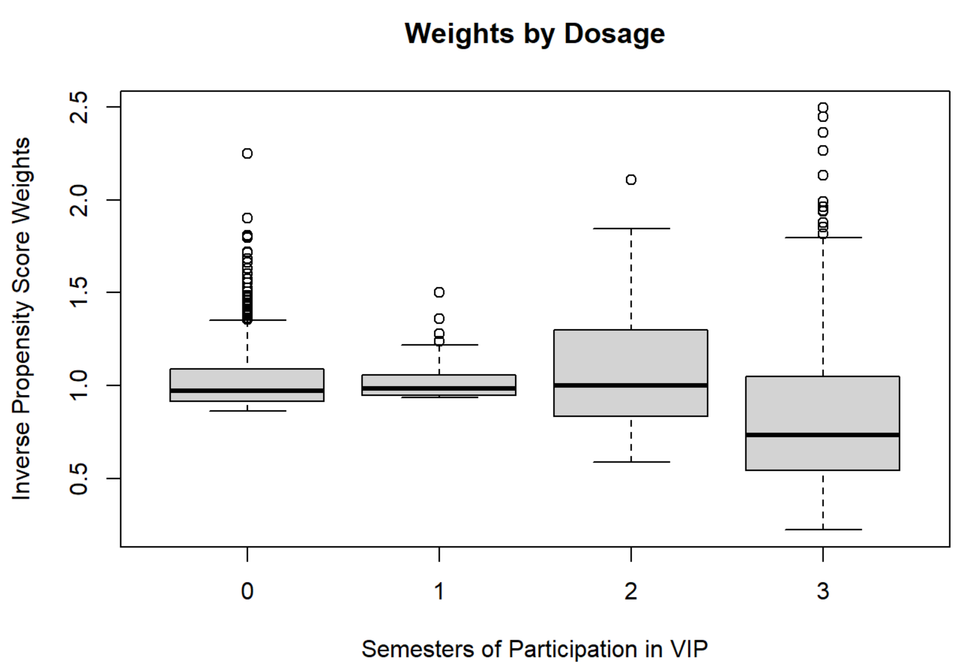




Inverse Propensity Score Weights

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.2230  0.9037  0.9630  0.9974  1.0858  2.4938
```

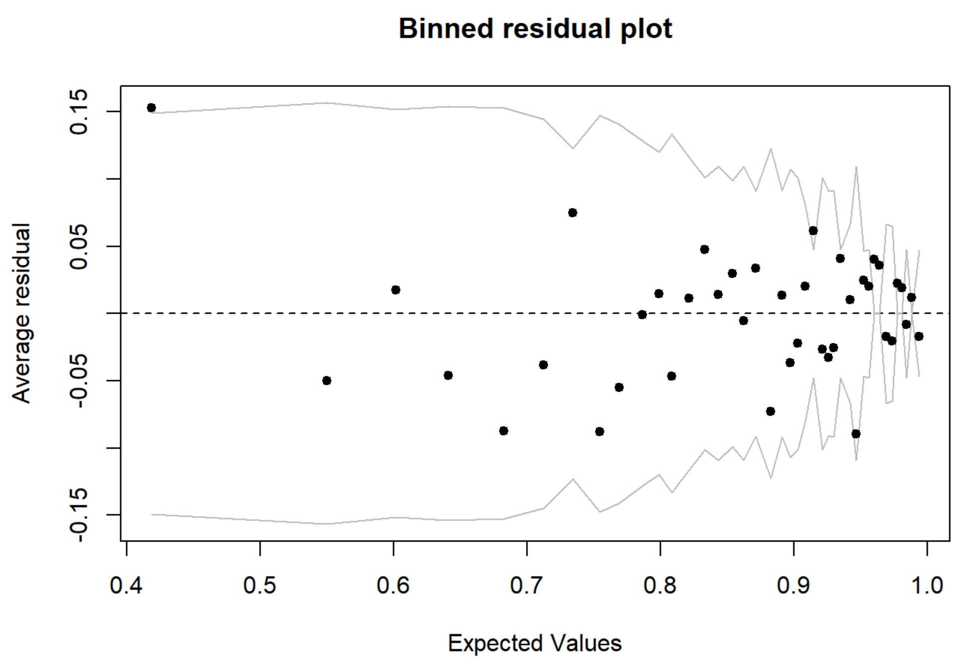
	0 sem	1 sem	2 sem	3 sem
Min.	0.863	0.934	0.589	0.223
1st Qu.	0.916	0.949	0.837	0.546
Median	0.970	0.983	1.001	0.732
Mean	1.021	1.017	1.087	0.858
3rd Qu.	1.090	1.058	1.298	1.047
Max.	2.249	1.500	2.106	2.494



REGRESSION

Residuals

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -3.2793  0.1954  0.3950  0.1560  0.5908  1.6008
```




```
##              R-Squared      Adjusted R-Squared
##                  0.147                   0.134
```

Regression Results

```
##
## Call:
## svyglm(formula = ..1, design = ..2, family = ..3)
##
## Survey design:
```

```
## survey::svydesign(...)
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)              -3.803095   0.698238  -5.447 5.88e-08 ***
## VIPSEM1                   0.181750   0.267118   0.680 0.496338
## VIPSEM2                   0.932014   0.662636   1.407 0.159751
## VIPSEM3                   1.170960   0.341603   3.428 0.000623 ***
## CITZResident NonCitizen  0.192183   0.301397   0.638 0.523794
## FEMALE                    0.445327   0.204713   2.175 0.029740 *
## RCETHAsian                0.248400   0.187174   1.327 0.184651
## RCETHOther or Unknown    0.351574   0.328570   1.070 0.284766
## RCETHURM                  0.457984   0.238723   1.918 0.055218 .
## PELL                     -0.423221   0.163384  -2.590 0.009670 **
## TRAN                     -0.275231   0.191854  -1.435 0.151590
## GRK                       0.825494   0.216807   3.808 0.000145 ***
## STAB                      0.541607   0.239831   2.258 0.024054 *
## GT1                       0.007806   0.205959   0.038 0.969773
## LLHON                     0.515569   0.442519   1.165 0.244151
## MAJREVComputer Eng       1.068006   0.364941   2.927 0.003473 **
## MAJREVComputer Sci       1.254187   0.343228   3.654 0.000266 ***
## UROP                     -0.108745   0.204691  -0.531 0.595307
## COOP1 Some CoOp          0.113581   0.311282   0.365 0.715246
## COOP3 CoOpDegDesig       1.342228   0.480836   2.791 0.005306 **
## INT12                     1.032026   0.228495   4.517 6.71e-06 ***
## GPA                       1.071552   0.156117   6.864 9.36e-12 ***
## YR2018                    0.516460   0.274067   1.884 0.059677 .
## YR2019                    0.065361   0.268669   0.243 0.807821
## YR2020                   -0.359504   0.251387  -1.430 0.152876
## YR2021                   -0.087782   0.275029  -0.319 0.749634
## YR2022                   -0.310522   0.251583  -1.234 0.217272
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 1.132976)
##
## Number of Fisher Scoring iterations: 6
```

```
##                          Estimate Std. Error t value Pr(>|t|)   2.5 %  97.5 %
## (Intercept)               -3.8031     0.6982 -5.4467   0.0000 -5.1726 -2.4336
## VIPSEM1                    0.1817     0.2671  0.6804   0.4963 -0.3422  0.7057
## VIPSEM2                    0.9320     0.6626  1.4065   0.1597 -0.3676  2.2317
## VIPSEM3                    1.1710     0.3416  3.4278   0.0006  0.5010  1.8410
## CITZResident NonCitizen   0.1922     0.3014  0.6376   0.5238 -0.3990  0.7833
## FEMALE                     0.4453     0.2047  2.1754   0.0297  0.0438  0.8468
## RCETHAsian                 0.2484     0.1872  1.3271   0.1846 -0.1187  0.6155
## RCETHOther or Unknown     0.3516     0.3286  1.0700   0.2848 -0.2929  0.9960
## RCETHURM                   0.4580     0.2387  1.9185   0.0552 -0.0102  0.9262
## PELL                      -0.4232     0.1634 -2.5904   0.0097 -0.7437 -0.1028
## TRAN                      -0.2752     0.1919 -1.4346   0.1516 -0.6515  0.1011
## GRK                        0.8255     0.2168  3.8075   0.0001  0.4003  1.2507
## STAB                       0.5416     0.2398  2.2583   0.0241  0.0712  1.0120
## GT1                        0.0078     0.2060  0.0379   0.9698 -0.3961  0.4118
## LLHON                      0.5156     0.4425  1.1651   0.2441 -0.3524  1.3835
## MAJREVComputer Eng        1.0680     0.3649  2.9265   0.0035  0.3522  1.7838
## MAJREVComputer Sci        1.2542     0.3432  3.6541   0.0003  0.5810  1.9274
## UROP                      -0.1087     0.2047 -0.5313   0.5953 -0.5102  0.2927
## COOP1 Some CoOp           0.1136     0.3113  0.3649   0.7152 -0.4969  0.7241
## COOP3 CoOpDegDesig        1.3422     0.4808  2.7914   0.0053  0.3991  2.2853
## INT12                      1.0320     0.2285  4.5166   0.0000  0.5839  1.4802
## GPA                        1.0716     0.1561  6.8638   0.0000  0.7654  1.3778
## YR2018                     0.5165     0.2741  1.8844   0.0597 -0.0211  1.0540
## YR2019                     0.0654     0.2687  0.2433   0.8078 -0.4616  0.5923
## YR2020                    -0.3595     0.2514 -1.4301   0.1529 -0.8526  0.1336
## YR2021                    -0.0878     0.2750 -0.3192   0.7496 -0.6272  0.4516
## YR2022                    -0.3105     0.2516 -1.2343   0.2173 -0.8040  0.1829
```

Adjusted Odds Ratios with Confidence Intervals

```
##                            AOR 2.5 % 97.5 %
## VIPSEM1                  1.199 0.710  2.025
## VIPSEM2                  2.540 0.692  9.316
## VIPSEM3                  3.225 1.650  6.303
## CITZResident NonCitizen 1.212 0.671  2.189
## FEMALE                   1.561 1.045  2.332
## RCETHAsian               1.282 0.888  1.851
## RCETHOther or Unknown   1.421 0.746  2.707
## RCETHURM                 1.581 0.990  2.525
## PELL                     0.655 0.475  0.902
## TRAN                     0.759 0.521  1.106
## GRK                      2.283 1.492  3.493
## STAB                     1.719 1.074  2.751
```

```
## GT1                    1.008 0.673  1.509
## LLHON                  1.675 0.703  3.989
## MAJREVComputer Eng     2.910 1.422  5.952
## MAJREVComputer Sci     3.505 1.788  6.871
## UROP                   0.897 0.600  1.340
## COOP1 Some CoOp        1.120 0.608  2.063
## COOP3 CoOpDegDesig     3.828 1.491  9.829
## INT12                  2.807 1.793  4.394
## GPA                    2.920 2.150  3.966
## YR2018                 1.676 0.979  2.869
## YR2019                 1.068 0.630  1.808
## YR2020                 0.698 0.426  1.143
## YR2021                 0.916 0.534  1.571
## YR2022                 0.733 0.448  1.201
```

5 Majors - VIPSEM as factor

jsk

Subset Statements

```
Data <- subset(Data, CITZ != "Alien, Non-Resident*")
Data <- subset(Data, VIPSEM %in% c(0,1,2,3)) #TREATMENT LEVELS

Data <- subset(Data, MAJREV %in% c("Computer Science",
"Environmental Engineering",
"Computer Engineering",
"Electrical Engineering",
"Aerospace Engineering"))

Data <- subset(Data, GPA >= 1.9)

# Data <- subset(Data, COOP == "0 No CoOp")
# Data <- subset(Data, INT12 == 0)

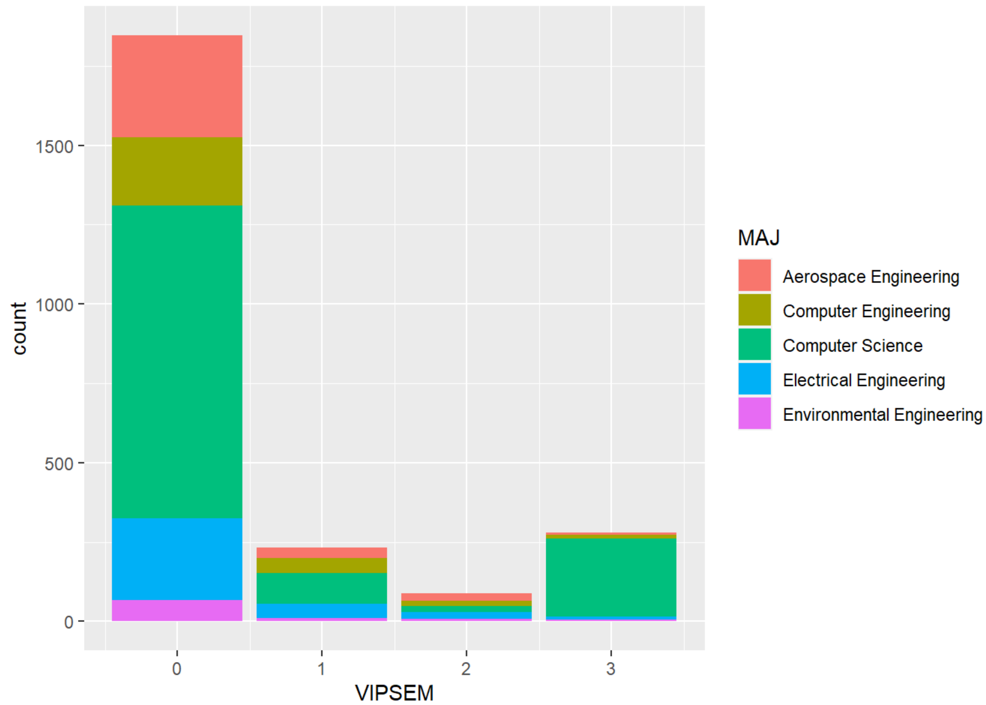
Data$MAJREV[Data$MAJREV=="Computer Science"] <- "Computer Sci"
Data$MAJREV[Data$MAJREV=="Environmental Engineering"] <- "0 Env Eng"
Data$MAJREV[Data$MAJREV=="Computer Engineering"] <- "Computer Eng"
Data$MAJREV[Data$MAJREV=="Electrical Engineering"] <- "Electrical Eng"
Data$MAJREV[Data$MAJREV=="Mechanical Engineering"] <- "Mechanical Eng"
Data$MAJREV[Data$MAJREV=="Biomedical Engineering"] <- "Biomedical Eng"
Data$MAJREV[Data$MAJREV=="Materials Science & Engr"] <- "Mat Science & Eng"
Data$MAJREV[Data$MAJREV=="Civil Engineering"] <- "Civil Eng"
Data$MAJREV[Data$MAJREV=="Aerospace Engineering"] <- "Aerospace Eng"
Data$MAJREV[Data$MAJREV=="Industrial Engineering"] <- "Industrial Eng"
Data$MAJREV[Data$MAJREV=="Chemical and Biomolecular Eng"] <- "Chem & Biomolec Eng"


# Data <- subset(Data, RCETH == "White")

# Set ref category to "other or Unknown"
# Data$RCETH[Data$RCETH == "Two or more"] <- "0 Other or Unknown"
# Data$RCETH[Data$RCETH == "American Indian or Alaska Native"] <- "0 Other or Unknown"
# Data$RCETH[Data$RCETH == "Native Hawaiian or Other Pacific Islander"] <- "0 Other or Unknown"
# Data$RCETH[Data$RCETH == "Unknown"] <- "0 Other or Unknown"
```

Fequency Tables

```
##
##                    0   1   2   3
##   0 Env Eng       68  11   9   5
##   Aerospace Eng  320  34  24   6
##   Computer Eng   216  47  17  13
##   Computer Sci   985  97  18 244
##   Electrical Eng 257  44  20  11
```



Propensity Score Model

```
## VIPSEM ~ CITZ + FEMALE + RCETH + PELL + TRAN + GRK + STAB + GT1 +
##     LLHON + MAJREV + UROP + COOP + INT12 + GPA
```

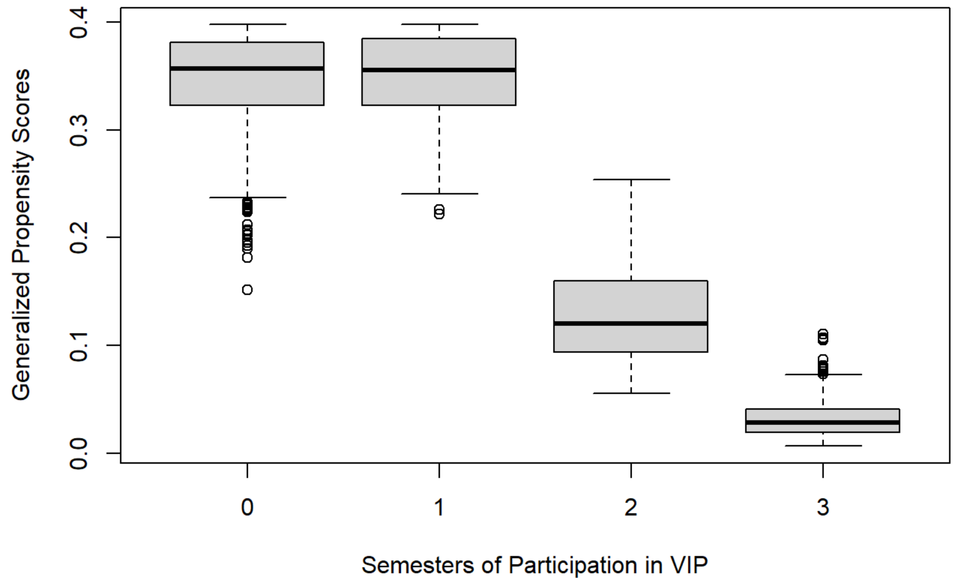
Balance Table

variable	coefBaseline	coefIPW
CITZ	0.020	0.047
FEMALE	0.229	0.020
RCETH	0.375	0.058
PELL	0.089	0.013
TRAN	0.128	0.020
GRK	0.035	0.002
STAB	0.010	0.010
GT1	0.002	0.006
LLHON	0.390	0.031
MAJREV	0.212	0.163
UROP	0.082	0.009
COOP	0.258	0.127
INT12	0.198	0.019
GPA	0.233	0.032

Sample Sizes

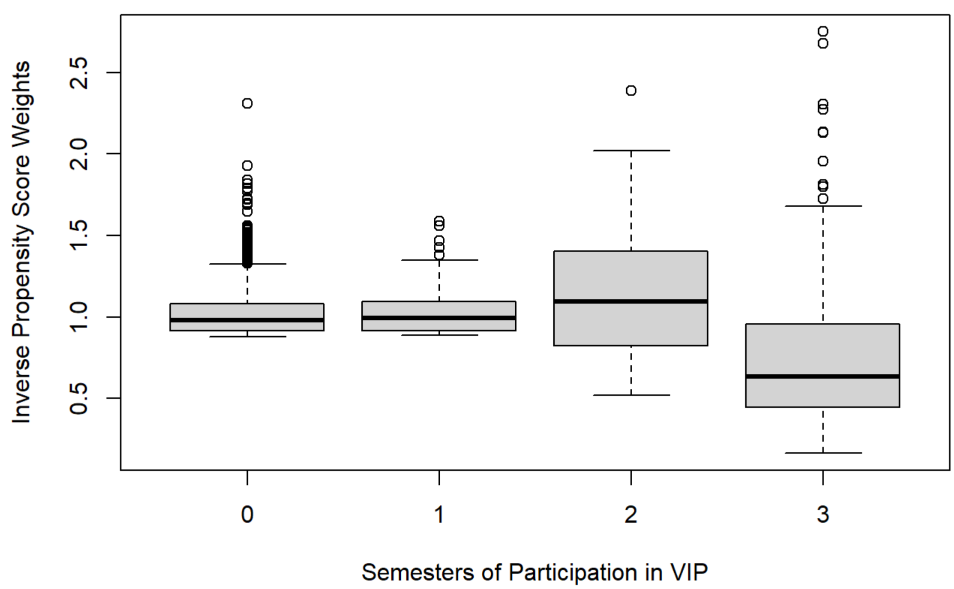
Generalized Propensity Scores

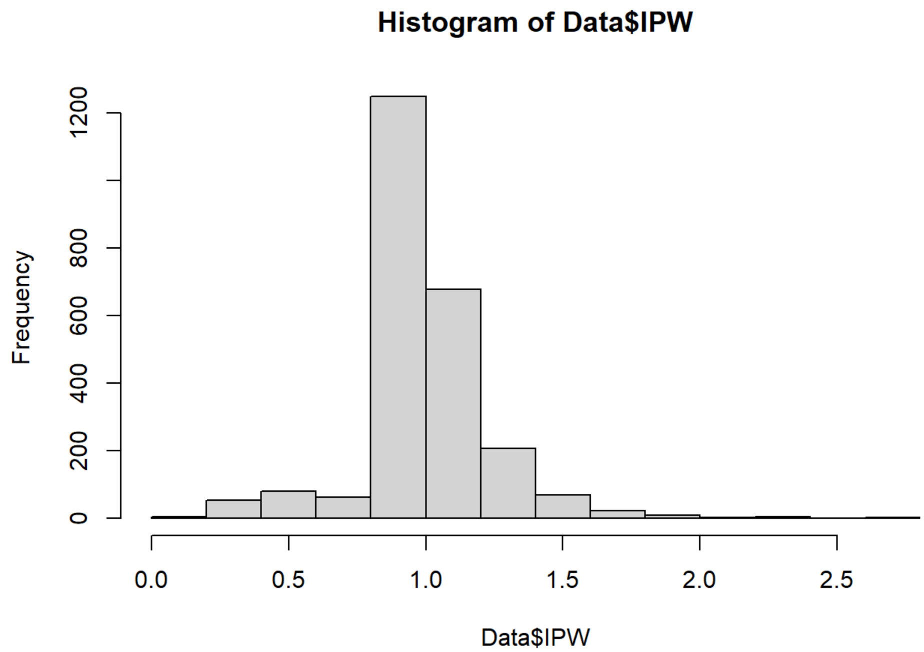
	0 sem	1 sem	2 sem	3 sem
Min.	0.151	0.222	0.055	0.007
1st Qu.	0.323	0.323	0.094	0.019
Median	0.357	0.355	0.120	0.029
Mean	0.347	0.349	0.130	0.033
3rd Qu.	0.381	0.385	0.160	0.041
Max.	0.398	0.398	0.254	0.110



Inverse Propensity Score Weights

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.1653  0.9033  0.9709  0.9983  1.0824  2.7500
```



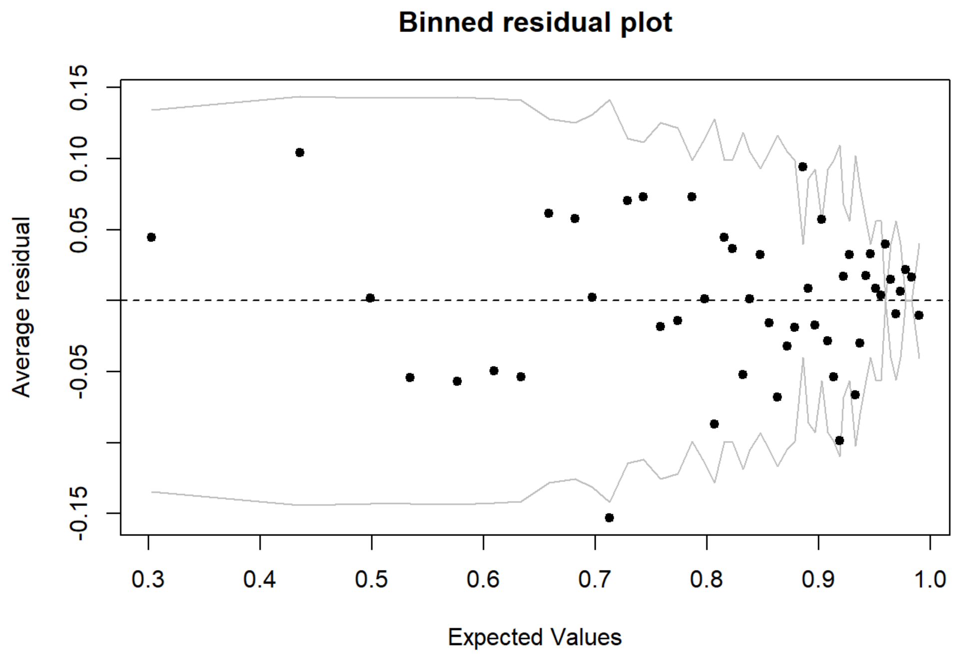


	0 sem	1 sem	2 sem	3 sem
Min.	0.879	0.887	0.519	0.165
1st Qu.	0.917	0.917	0.824	0.448
Median	0.979	0.993	1.095	0.637
Mean	1.026	1.027	1.148	0.746
3rd Qu.	1.082	1.094	1.399	0.956
Max.	2.309	1.588	2.387	2.750

REGRESSION

Residuals

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -2.9636  0.2031  0.4334  0.1494  0.6444  1.7991
```




```
##              R-Squared     Adjusted R-Squared
##                  0.153                  0.143
```

Regression Results

```
##
## Call:
## svyglm(formula = ..1, design = ..2, family = ..3)
##
## Survey design:
## survey::svydesign(...)
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             -3.432656   0.555524  -6.179 7.54e-10 ***
## VIPSEM1                  0.169768   0.210115   0.808 0.419183
## VIPSEM2                  0.490117   0.331738   1.477 0.139692
## VIPSEM3                  0.961207   0.312214   3.079 0.002103 **
## CITZResident NonCitizen  0.001895   0.236244   0.008 0.993600
## FEMALE                   0.495817   0.160323   3.093 0.002007 **
## RCETHAsian               0.109435   0.151860   0.721 0.471206
## RCETHOther or Unknown    0.218460   0.258415   0.845 0.397980
## RCETHURM                 0.200322   0.172533   1.161 0.245731
## PELL                    -0.447829   0.131109  -3.416 0.000647 ***
## TRAN                    -0.189782   0.153221  -1.239 0.215609
## GRK                      0.740931   0.166427   4.452 8.89e-06 ***
## STAB                     0.242903   0.154647   1.571 0.116384
## GT1                     -0.001385   0.154537  -0.009 0.992852
## LLHON                    0.543647   0.363969   1.494 0.135395
## MAJREVAerospace Eng      0.185768   0.313706   0.592 0.553791
## MAJREVComputer Eng       1.167779   0.339519   3.440 0.000593 ***
## MAJREVComputer Sci       1.330388   0.311833   4.266 2.06e-05 ***
## MAJREVElectrical Eng     0.866501   0.334758   2.588 0.009699 **
## UROP                     0.099871   0.152643   0.654 0.512994
## COOP1 Some CoOp          0.248676   0.244737   1.016 0.309685
## COOP3 CoOpDegDesig       1.153296   0.275787   4.182 2.99e-05 ***
## INT12                    0.980032   0.180464   5.431 6.18e-08 ***
## GPA                      0.955091   0.123381   7.741 1.44e-14 ***
## YR2018                   0.391525   0.200444   1.953 0.050901 .
## YR2019                   0.056955   0.197130   0.289 0.772668
## YR2020                  -0.323745   0.198816  -1.628 0.103578
## YR2021                   0.252712   0.219617   1.151 0.249971
## YR2022                  -0.117111   0.201070  -0.582 0.560325
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 0.9997478)
##
## Number of Fisher Scoring iterations: 5
```

```
##                          Estimate Std. Error t value Pr(>|t|)    2.5 %  97.5 %
## (Intercept)               -3.4327     0.5555 -6.1791   0.0000 -4.5220 -2.3433
## VIPSEM1                    0.1698     0.2101  0.8080   0.4192 -0.2423  0.5818
## VIPSEM2                    0.4901     0.3317  1.4774   0.1397 -0.1604  1.1406
## VIPSEM3                    0.9612     0.3122  3.0787   0.0021  0.3490  1.5734
## CITZResident NonCitizen    0.0019     0.2362  0.0080   0.9936 -0.4614  0.4652
## FEMALE                     0.4958     0.1603  3.0926   0.0020  0.1814  0.8102
## RCETHAsian                 0.1094     0.1519  0.7206   0.4712 -0.1884  0.4072
## RCETHOther or Unknown      0.2185     0.2584  0.8454   0.3980 -0.2883  0.7252
## RCETHURM                   0.2003     0.1725  1.1611   0.2457 -0.1380  0.5386
## PELL                      -0.4478     0.1311 -3.4157   0.0006 -0.7049 -0.1907
## TRAN                      -0.1898     0.1532 -1.2386   0.2156 -0.4902  0.1107
## GRK                        0.7409     0.1664  4.4520   0.0000  0.4146  1.0673
## STAB                       0.2429     0.1546  1.5707   0.1164 -0.0603  0.5462
## GT1                       -0.0014     0.1545 -0.0090   0.9929 -0.3044  0.3017
## LLHON                      0.5436     0.3640  1.4937   0.1354 -0.1701  1.2574
## MAJREVAerospace Eng        0.1858     0.3137  0.5922   0.5538 -0.4294  0.8009
## MAJREVComputer Eng         1.1678     0.3395  3.4395   0.0006  0.5020  1.8336
## MAJREVComputer Sci         1.3304     0.3118  4.2663   0.0000  0.7189  1.9419
## MAJREVElectrical Eng       0.8665     0.3348  2.5884   0.0097  0.2101  1.5229
## UROP                       0.0999     0.1526  0.6543   0.5130 -0.1995  0.3992
## COOP1 Some CoOp            0.2487     0.2447  1.0161   0.3097 -0.2312  0.7286
## COOP3 CoOpDegDesig         1.1533     0.2758  4.1818   0.0000  0.6125  1.6941
## INT12                      0.9800     0.1805  5.4306   0.0000  0.6262  1.3339
## GPA                        0.9551     0.1234  7.7410   0.0000  0.7131  1.1970
## YR2018                     0.3915     0.2004  1.9533   0.0509 -0.0015  0.7846
## YR2019                     0.0570     0.1971  0.2889   0.7727 -0.3296  0.4435
## YR2020                    -0.3237     0.1988 -1.6284   0.1036 -0.7136  0.0661
## YR2021                     0.2527     0.2196  1.1507   0.2500 -0.1779  0.6834
## YR2022                    -0.1171     0.2011 -0.5824   0.5603 -0.5114  0.2772
```

Adjusted Odds Ratios with Confidence Intervals

```
##                           AOR 2.5 % 97.5 %
## VIPSEM1                 1.185 0.785  1.789
## VIPSEM2                 1.633 0.852  3.129
## VIPSEM3                 2.615 1.418  4.823
## CITZResident NonCitizen 1.002 0.630  1.592
## FEMALE                  1.642 1.199  2.248
## RCETHAsian              1.116 0.828  1.503
## RCETHOther or Unknown   1.244 0.750  2.065
## RCETHURM                1.222 0.871  1.714
## PELL                    0.639 0.494  0.826
## TRAN                    0.827 0.612  1.117
## GRK                     2.098 1.514  2.907
## STAB                    1.275 0.941  1.727
## GT1                     0.999 0.738  1.352
## LLHON                   1.722 0.844  3.516
## MAJREVAerospace Eng     1.204 0.651  2.228
## MAJREVComputer Eng      3.215 1.652  6.256
## MAJREVComputer Sci      3.783 2.052  6.972
## MAJREVElectrical Eng    2.379 1.234  4.586
## UROP                    1.105 0.819  1.491
## COOP1 Some CoOp         1.282 0.794  2.072
## COOP3 CoOpDegDesig      3.169 1.845  5.442
## INT12                   2.665 1.870  3.796
## GPA                     2.599 2.040  3.310
## YR2018                  1.479 0.998  2.191
## YR2019                  1.059 0.719  1.558
## YR2020                  0.723 0.490  1.068
## YR2021                  1.288 0.837  1.981
## YR2022                  0.889 0.600  1.319
```