UC Irvine
UC Irvine Electronic Theses and Dissertations

Title
How Curriculum Structure, Policy, and Intervention Shape Student Social and Academic Outcomes

Permalink
https://escholarship.org/uc/item/5bf4p3rs

Author
Li, XunFei

Publication Date
2024

Copyright Information
This work is made available under the terms of a Creative Commons Attribution-NonCommercial-ShareAlike License, available at https://creativecommons.org/licenses/by-nc-sa/4.0/

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA,
IRVINE

How Curriculum Structure, Policy, and Intervention Shape Student Social and Academic Outcomes

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Education

by

XunFei Li

Dissertation Committee:
Professor Di Xu, Co-Chair
Professor Richard Arum, Co-Chair
Associate Professor Rachel B. Baker
Distinguished Professor Greg J. Duncan
Professor Andrew Penner

2024

© 2024 XunFei Li

DEDICATION

To

my parents Qin CaiLan and Li ShuZhi

whose bravery and love encourage me to chase my aspirations

TABLE OF CONTENTS

Page

LIST OF FIGURES vi

LIST OF TABLES vii

ACKNOWLEDGEMENTS viii

VITA ix

ABSTRACT OF THE DISSERTATION xi

INTRODUCTION 1

CHAPTER 1 Same Major, Different Peers: Gender Sorting Across and Within Majors 5

Introduction 5

Literature review and theoretical framework 8

Institutional Constraints on Students' Course-Taking Patterns 8

Gender Sorting at Major Level 10

What Do We Know About Gender Sorting Between Classes for Students Enrolled in the Same Majors? 12

Importance of Peer Gender Composition 14

Data and methods 15

Data 15

Outcome: Same-Gender Peer Composition 16

Analytic Strategy 18

Ordinary least squared regression (OLS) model 18

Major fixed-effect models for different courses 19

Findings 20

Variation in the Proportion of Same-Gender Classmates 20

Do We See More Gender Segregation in Certain Demographic Groups? 21

How Does the Proportion of Same-Gender Classmates Vary Across Majors? 22

Within-Major Variation in the Proportion of Same-Gender Classmates 24

Role of Classes Students Take Outside Their Major 25

Conclusion 27

References 33

Appendices 48

CHAPTER 2 Optimizing Student Engagement Through Diverse Leadership Tendencies in Group Composition 60
Introduction 60
Literature Review 63
Sense of Belonging in Online STEM Education 63
Benefits And Challenges of Small Student-Led Study Groups 64
Group Composition And Social Skills 65
Current Study 66
Experiment Context 66
Experiment Design 67
Outcome Measures 69
Behavioral outcomes: participation and engagement in study group activities 69
Student interaction within study groups: Conversational data through the communication platform 70
Study-group perceptions 71
Course social and learning experience, and academic outcomes 72
Analytic Sample 72
Methods 74
Results 74
Study Group Participation And Engagement 74
Course Social And Learning Experience and Course Grades 76
Impact On Underrepresented Minority (URM) Students 77
Discussion 79
Reference 83
Appendix 101
Chapter 3 Fast Track to Success?: Investigating Condensed Course Formats and Student Achievement in Community Colleges 130
INTRODUCTION 130
LITERATURE REVIEW 132
Why Condensed Courses Are Offered and Their Prevalence 133
Impact of condensed courses on student academic outcomes 134
Scope limitations of prior studies on summer-term and single-college 134
Methodological limitations: threats of student sorting 135

Beyond current course performance 137
Potential Mechanisms of the impact of condensed courses 137
BACKGROUND AND DATA 138
Data and Institutional Context 138
Condensed course at VCCS 139
Analytic sample and summary statistics 141
EMPIRICAL STRATEGY 143
Empirical model for concurrent course outcomes 143
Empirical model for downstream outcomes 144
RESULTS 145
Current course outcomes 145
Downstream outcomes 147
Exploratory analysis of student-level outcomes using instrumental variable approach 148
DISCUSSION 150
Reference 153
Appendix 167
CHAPTER 4 Summary and Conclusions 170

LIST OF FIGURES

Figure 1. 1 Kernel Density Distribution of the Proportion of Same-Gender Classmates for Female and Male Students 45
Figure 1. 2 10th, 50th and 90th Percentiles of Proportion of Same-Gender Classmates Across All Courses, by Major 46
Figure 1. 3 10th, 50th and 90th Percentiles of Proportion of Same-Gender Classmates Across In-Department and Out-of-Department Courses 47
Figure 2. 1 Treatment effect on study group participation and engagement outcomes 97
Figure 2. 2 Treatment effect on study group perception (SGP), course social experience (CSE), course learning experience (CLE), and course grade (CG) 98
Figure 2. 3 Treatment effect on study group participation and engagement outcomes for URM and non-URM students 99
Figure 2. 4 Treatment effect on study group perception (SGP), course social experience (CSE), course learning experience (CLE), and course grade (CG) for URM students 100
Figure 3. 1 Distribution of course duration (measured by weeks) 165
Figure 3. 2 Distribution of course duration across fields 166

LIST OF TABLES

Table 1. 1 Summary Statistics of Students' Characteristics and Outcomes (Student-Level) 41
Table 1. 2 Summary Statistics of Outcome Variables 42
Table 1. 3 Results of Ordinary Least Squares Models, Major Fixed-Effect Models on the Proportion of Students' Same-Gender Classmates for All Courses 43
Table 1. 4 Results of Major Fixed-Effect Models on the Proportion of Students' Same-Gender Classmates for In-Department Courses and Out-Of-Department Courses 44
Table 2. 1 Student Characteristics by experimental condition 90
Table 2. 2 Treatment effect on aggregated weekly outcomes without covariates 92
Table 2. 3 Treatment effect on aggregated weekly outcomes with covariates 92
Table 2. 4 Impact of treatment condition on LIWC outcomes 93
Table 2. 5 Impact of treatment condition on student perception and academic performance 94
Table 2. 6 Heterogeneous analysis with covariates (n = 287, URM = 123 non-URM = 167) 95
Table 3. 1 Institutional characteristics of VCCS and National sample in fall 2010 157
Table 3. 2 Student Descriptive Statistics 158
Table 3. 3 Course Descriptive Statistics 159
Table 3. 4 Probability of taking condensed courses 160
Table 3. 5 Impact of Condensed Courses on Introductory Course Performance 161
Table 3. 6 Impact of Condensed Courses in Introductory Courses on Subsequent Enrollment in the Subject 162
Table 3. 7 Impact of Condensed Course in Introductory Courses on Subsequent Course Grade in the Subject 163
Table 3. 8 Exploratory student-level analysis: Relationship between % of credits taken with condensed course formats in initial term on subsequent student outcomes 164

ACKNOWLEDGEMENTS

At the end of this turbulent PhD journey, I would like to express my gratitude to my mentors, family, and friends who voyaged through this journey with me, providing me with courage, wisdom, tools, and advice along the way. I am also grateful to Mellon Foundation (grant numbers: 1806-05902, 2011-09741), National Science Foundation (grant numbers: 2228202, FG20825), and Institute of Education Sciences (grant number: R305A220224) for their funding support, and to the Virginia Community College System, especially Cat and Marina, for being wonderful partners and supporters of my research.

I want to thank my advisor, Di, for her encouraging and candid guidance throughout my PhD career. She has consistently urged me to ask meaningful policy questions and consider their implications. Her passion and knowledge in research have profoundly inspired me. I appreciate that she always encouraged me to reach my full potential. I also want to express my gratitude to my co-chair, Richard, who has supported me throughout my entire PhD career. Talking with him always reminds me of how sociology shapes my thinking. He shares his insights, provides honest feedback on my work, and has given me precious opportunities to collaborate with a wonderful team on the MUST project. My deep gratitude also goes to another de facto advisor, Rachel, who has mentored and worked with me very closely on my early work in my PhD and continues to inspire and support me to further explore this strand of work. She answers all my strange questions and guides me in transforming vague ideas into clear research questions. Thank you, Greg, for those enlightening walks on the campus and for those brilliant questions about my work and career plans. Thank you, Andrew, for your amazing understanding and feedback on my work, and your kind encouragement. I also want to thank my mentors in my master's program in Sociology, without whom I would not have been able to pursue my PhD. Dr. Ping-Yin Kuan, who has always sent kind words to me throughout my PhD career.

Thank you, Qiujie, for mentoring me to pursue rigorous research. Thanks to Sabrina for sharing her abundant experience and career guidance with me. Thank you to my UCI friends for staying with me in the timeless office, exercising with me, especially during the segregated pandemic time, and for debating and answering my questions: Xuehan, Yujia, Yi, Gala, Yannan, Cooper, Ahja, Nikki.

I’m deeply grateful to my parents and my brother who have always supported me. The four years of the pandemic without being able to see them made me understand what is most important in my life. Thank you to my best childhood friends, who do not live with me but let me believe I can live my life far from my hometown because they did it too: Yilin and Amanda. Deepest thanks to my wonderful significant other, Sean, for living an interesting life with me and for the time spent cooking, talking, walking, and playing together.

VITA

XunFei Li

EDUCATION

Ph.D. in School of Education 2024
University of California, Irvine Irvine, USA
Visiting Student Researcher in Sociology Aug 2016 - Jun 2017
University of California, Berkeley Berkeley, USA
MA in Sociology 2017
National Chengchi University Taipei, Taiwan
Thesis title: "The Effect of Single-Sex Schooling on High School Girls' Curriculum Tracking Selection: A Counterfactual Analysis of Taiwan Educational Panel Survey"
BA in Sociology 2014
Sun Yat-sen University Guangzhou, China

PUBLICATIONS

Zhou, X., Li, Q., Xu, D., Li, XF., & Fischer, C. (2023). College online courses have strong design in scaffolding but vary widely in supporting student agency and interactivity. *The Internet and Higher Education, 58,* 100912.

Xu, D., Finnegan, C., Bagreev, M., Li, XF., D'Amico, M. M., & Van Noy, M. (2023) Education and Employment Research Center (EERC). Virginia State Noncredit Data Report. Retrieved from https://sites.rutgers.edu/state-noncredit-data/wp-content/uploads/sites/794/2023/03/Virginia-State-Noncredit-Data-Report_FINAL.pdf

Xu, D., Perez Lopez, J., Liu, Y., Hill, M. S., Li, XF., Bañuelos, M., Pacansky-Brock, M., Smedshammer, M., Vincent-Layton, K., & White, J. W.. (2022). "Humanizing online STEM: Preliminary research findings." HumanizeOL.org.

Li, XF., & Yu, R. (2021, April). Construction of Weighted Course Co-Enrollment Network. In *Proceedings of the NetSciLA21 workshop*.

Li, XF., & Kuan, P. Y. (2018). The effect of single-sex schooling on high school girls' curriculum tracking selection in Taiwan. *International Sociological Association*.

PRESENTATIONS

Li, XF., & Xu, D. (2024, March). Impact of short-course formats on student success: Evidence from a statewide community college system. Paper presented at the *Annual Conference of the Association for Education Finance and Policy (AEFP)*, Baltimore, U.S.

Li, XF., Baker, R., & Arum, R. (2024, March). Same major, different peers: Gender sorting across and within majors. Paper presented at the *Annual Conference of the Association for Education Finance and Policy (AEFP)*, Baltimore, USA.

Li, XF., Li, Q., Xu, D., Holton, A., Zhou, X. (2023, November) Building Effective Small Study Groups in an Online Introductory STEM Course: The Impact and Mechanisms of Study Groups

on Student Engagement, *2023 APPAM Fall Research Conference*, Atlanta, U.S.
Arum, R., Li, XF, Mcdossi, O. (2022, November) Student Grades and College Pathways, *Pathway Seminars*, University of Stanford.
Li, XF., Baker, R.., and Arum, R. (2022, April) Sorting into and within Majors: How Big and Small Choices Affect Classmate Gender Composition, *Research Committee 28 on Social Stratification and Social Mobility of the International Sociological Association (ISA)*, London, U.K.
Li, XF and Yu, R. "Construction of Weighted Course Co-Enrollment Network", *The 11th International Learning Analytics and Knowledge Conference: Using Network Science in Learning Analytics Workshop*, April 2021, U.S.
Li, XF and Kuan, PY. (2018, May) The Effect of Single-Sex Schooling on High School Girls' Curriculum Tracking Selection in Taiwan, *Spring Meeting of the Research Committee 28 on Social Stratification and Mobility of the International Sociological Association (ISA)*, Seoul, South Korea

HONORS AND AWARDS
Public Impact Fellowship, 2023, University of California, Irvine
Third Place Poster Prize, 2023, APPAM Fall Research Conference
Summer Institute in Computational Social Science Participant, 2023, University of California, LosAngelas

- A two-week intensive training program with a competitive selection process.

Gardner Seminar Fellows, 2022-2023, Center for Studies in Higher Education, University of California, Berkeley

- A year-long program with only 11 fellows selected from all University of California campuses.

Provost PhD Fellowship, 2019-2020, University of California, Irvine
Graduate Dean's Recruitment Fellowship, 2019, University of California, Irvine
RC28 Travel Awards, 2018, Research Committee 28 on Social Stratification and Mobility of the International Sociological Association

ABSTRACT OF THE DISSERTATION

How Curriculum Structure, Policy, and Intervention Shape
Student Social and Academic Outcomes
by

XunFei Li

Doctor of Philosophy in Education

University of California, Irvine, 2024

Professor Di Xu, Co-Chair

Professor Richard Arum, Co-Chair

Colleges shape student classroom experiences through macro-level curriculum policies and micro-level learning designs. Understanding the impact of these policies on students' peer environments and learning opportunities, and designing classroom interventions to improve student experiences, is crucial for enhancing educational quality and equity. Study 1 examines gender sorting both across and within majors, documenting sorting at a student-by-course level. We observed considerable sorting within majors. This study suggests that previous research focusing solely on major-level sorting overlooks important opportunities to understand the mechanisms driving gender segregation in higher education. Study 2 demonstrates the impact of leadership diversity in study group compositions on engagement in online STEM courses. Online learning has expanded access to STEM education but is hindered by limited social interaction, leading to reduced satisfaction and success rates. This study introduces a low-cost, scalable method to enhance student engagement by diversifying study group composition based on leadership motivation. Findings indicate that students in modified groups exhibited higher levels of engagement and satisfaction. The benefits were especially pronounced for underrepresented

minority students, with enhanced social networks and a sense of belonging. Study 3 evaluates the impact of condensed courses on student success within a statewide community college system using a quasi-experimental approach. Condensed courses offer flexibility and have the potential to benefit students juggling education and other responsibilities, though concerns exist about potential overburdening. Findings show that condensed courses significantly improve student persistence, completion rates, and grades in current courses. Additionally, students are more likely to succeed in subsequent courses within the same subject. This study contributes to the ongoing dialogue on educational innovation by providing rigorous causal evidence on the effectiveness of condensed courses with large-scale data.

INTRODUCTION

In higher education, classroom experience stands as a pivotal element in shaping the learning and socialization processes of college students, which are key to the accumulation of human and social capital. Central to this experience are two intertwined aspects: learning opportunities and peer environment. Learning opportunities pertain to both the course content and the learning environment to which students are exposed. Students can encounter dramatically different learning opportunities even within a university through their choice of majors and courses, including the peers they are exposed to, which are important components of their learning process. Previous literature on peer effects has shown that peers' preferences and behaviors influence students' academic outcomes, such as major and course choices and performance. Peers' resources and networks, which are crucial to students' social networks, can also affect their labor market outcomes. The peer environment is a key setting for interpersonal interactions, enabling students to perceive the field, compare themselves with others, and identify role models. These experiences can significantly influence their motivation and sense of belonging.

Colleges shape student classroom experiences through macro-level curriculum policies such as degree requirements and course offerings, as well as micro-level classroom learning designs. As such, understanding how curriculum policies are associated with or impact students' peer environments and learning opportunities, and how we can design interventions at the classroom level to improve students’ experiences, becomes crucial for enhancing educational quality and equity.

Paper 1 examines the nuanced dynamics of gender segregation within higher education majors, a critical aspect of classroom experience that shapes the learning and socialization processes in colleges. Prior studies have predominantly focused on gender segregation across majors. However, this perspective overlooks significant within-major segregation, which is crucial given the elective curriculum's central role in American higher education and the classroom's function as a primary site for learning and social interactions. This study extends the existing literature by documenting gender segregation at a student-by-course level within majors at a selective four-year college in the American West. We analyze the variation in the proportion of students' classmates who are the same gender, comparing this across and within majors. Findings reveal that substantial gender sorting persists within majors—61% of the variation for female students and 48% for male students occurs within their chosen fields. Moreover, the variation within majors nearly matches the variation across them. We also observe distinct patterns among specific student groups; for instance, female international students are more likely to have female peers than their domestic counterparts. This paper provides a framework for measuring gender segregation across campus contexts, exploring how major structures and the flexibility of curricular choices influence student sorting. These insights suggest that examining gender segregation solely at the major level may overlook key mechanisms of gender segregation in higher education.

Paper 2 describes an intervention at the classroom level to investigate the impact of leadership diversity in study group compositions on engagement and sense of belonging in online STEM courses, particularly among underrepresented minority (URM) students. As online learning expands access to STEM education, it faces challenges with social interaction, potentially reducing satisfaction and success rates. This study introduces a scalable, low-cost

method to enhance student engagement by diversifying study group composition based on leadership motivation, integrating students with high, medium, and low motivation to lead. A randomized controlled trial within a rigorous online STEM gateway course assessed this approach's effectiveness. Findings indicate that students in the adjusted groups experienced higher levels of participation, engagement, and satisfaction, with conversations showcasing a more positive tone and elevated leadership dynamics. The intervention was particularly beneficial for URM students, improving their social networks and enhancing their sense of belonging in the course. These results suggest that this strategic group organization can significantly boost student engagement and belonging in challenging STEM courses, contributing to greater diversity in STEM degree completion and workforce development. This adaptable strategy offers a promising solution for optimizing online learning outcomes, with important implications for advancing diversity in the STEM workforce.

Paper 3 evaluates a controversial curriculum policy: the impact of condensed courses—completed in 10 weeks or fewer—on student success within a statewide community college system, utilizing a quasi-experimental approach across multiple cohorts and colleges. These courses are debated for their potential flexibility and acceleration of student progress, which may be particularly beneficial for students juggling education, work, and family responsibilities. However, concerns exist that such courses might overburden students and provide insufficient time for learning material synthesis. This study is the first to use a rigorous quasi-experimental design to evaluate the impact of condensed courses on a large scale across a diverse set of subjects. Employing rigorous methods, including student-fixed effects and college-by-course fixed effects, this research mitigates biases related to student self-selection and course variations. The findings reveal that condensed courses significantly improve student persistence,

completion rates, and grades in concurrent courses. These students are also more likely to succeed in subsequent courses within the same subject over three years, although the benefits diminish when considering the specific sections of subsequent courses. This study contributes important insights to the discourse on educational innovation, demonstrating the potential of condensed courses to enhance both immediate and long-term academic outcomes.

CHAPTER 1 SAME MAJOR, DIFFERENT PEERS: GENDER SORTING ACROSS AND WITHIN MAJORS

INTRODUCTION

Since the 1980s, female students in the U.S. have earned a proportionally larger share of college degrees when compared with their male counterparts (DiPrete & Buchmann, 2013), thus narrowing, or even reversing, the gender gap in educational attainment. However, the decrease in vertical stratification has been coupled with a troubling persistence in horizontal sorting. Male and female students are sorted into qualitatively different kinds of education both across and within institutions (Charles & Bradley, 2002; Gerber & Cheung, 2008). In particular, female students are underrepresented in fields with greater labor market returns, including some STEM and technical fields (Barone, 2011; Bobbitt-Zeher, 2007; Davies & Guppy, 1997; Iannelli et al., 2018). Gender segregation across fields of study has persisted for the past three decades in many countries and research has identified multiple sociocultural, institutional, and psychological factors that have contributed to gender segregation across fields of study (Barone, 2011; Barone & Assirelli, 2020; Bobbitt-Zeher, 2007; Charles & Bradley, 2002; Di Prete & Buchmann, 2013; Ganley et al., 2018; Gerber & Cheung, 2008; Penner, 2015; Speer, 2017; Wang et al., 2013; Xie et al., 2015).

However, examining segregation only *across* majors could mask considerable segregation that occurs *within* majors. Even students within the same fields can find themselves in different courses due to various curricular structures, which encompass factors such as elective courses and general education courses (Chaturapruek et al., 2021;

Mann & DiPrete, 2013; McFarland, 2006; Scott-Clayton, 2015). For example, in the Biological Science department in our focal university, "Biology of Aging" has a relatively low proportion of female students (48%), while "Mood Disorders" has an average high proportion (83%). In Sociology, "Social Networks and Social Support" has an average 56% female students, while "Sociology of Education" has an average 79% female students.

Such horizontal sorting is somewhat akin to tracking, a concept that has garnered considerable attention in the K-12 literature for its forms and effects (Domina et al., 2019; Lucas & Berends, 2002; Sorensen, 1970). Tracking in higher education is manifested in both explicit and implicit ways. By dividing students into different majors, colleges channel the instruction, learning opportunities, and peer environments through degree/course requirements. Gender segregation at the major level, including the underrepresentation of women in STEM and more lucrative fields, shows one of the salient results of college tracking. Unlike K-12 education, the majors and curricular experience of college students are widely considered their personal choices. Such conventional understandings ignore the impact of institutional factors (e.g., general education, major requirements, course offerings, prerequisite, etc.) on students' college pathways, and are insufficient to explain why students' sociodemographic characteristics are systematically correlated with their major choices and curricular experiences.

The significance of gender segregation among students in the same major is two-fold. First, given that human capital developed at the course level is consequential in the job market, within-major sorting can lead to both occupational sorting and differential economic outcomes. Different courses involve different instruction, curricular content and student learning. Recent studies have provided evidence of the strong labor market returns

to student coursework in career and technical education (CTE) fields even without a credential or degree (Bahr, 2019; Giani et al., 2020), as well as the differential labor market returns to different specializations in the same field of study (Zölitz & Feld, 2020). Second, within-major sorting can lead to peer sorting and thus differential peer effects. For example, students surrounded mostly by peers of the same gender could develop less diverse networks and more homogeneous social norms, as compared with students in classrooms with a more even gender composition. Such differences could influence crucial individual outcomes such as future career choices, accumulation of social capital, and the development of diverse perspectives (e.g., Raabe et al., 2019; Zölitz & Feld, 2020).

To begin to fill this gap in the literature, this study explores gender sorting both *across* and *within* majors by examining what proportion of each student's classmates are of the same gender over the full trajectory of their college careers. Classroom context, as the basis of potential learning and social interactions, can provide a more accurate depiction of how students experience gender segregation in higher education than overall major context. We accomplish this using a unique data set that captures each class a student takes in and out of their major and their peer in that class. In our main analyses, we use in-department classes as a proxy for major-required courses. This is a somewhat noisy proxy for the major-level curricular requirements that constrain student course choices (e.g., courses in the Chemistry department are usually required for Biology majors). Thus, to examine the validity of our results, in Appendix A3, we include subsample analysis that provides more accurate hand-coded measurement of curricular requirements.

We utilized a selective public university in the American West as a case study to showcase our methodological innovation in analyzing student peer composition within

majors. Although our research is based on a single institution, the methodology is designed for broad applicability. It could easily be applied to examine within-major gender sorting on other campuses or to explore sorting in other significant student demographic groups, such as first-generation college student status or race and ethnicity.

This study extends beyond the analysis of gender segregation at the major level, advancing our understanding of how variations in the gender composition of students' classmates occur both across and within majors.

Specifically, we ask:

RQ1: In five cohorts of students at a selective four-year college in the American West, how much variation is there across students in the proportion of same-gender classmates?

RQ1a: Does the proportion of same-gender classmates vary by sociodemographic background?

RQ1b: How does the median proportion of same-gender classmates vary *across* majors?

RQ2: How much *within*-major variation is there in the proportion of same-gender classmates?

RQ2a: Is *within*-major variation in the proportion of same-gender classmates driven by classes students take outside their home department (e.g., general education requirements)?

LITERATURE REVIEW AND THEORETICAL FRAMEWORK

Institutional Constraints on Students' Course-Taking Patterns

Literature on tracking from the K-12 context, while not perfectly analogous to the

curriculum structure in higher education, provides insights to help us understand course-taking patterns for students of different genders and how institutional constraints can affect gender segregation. The term tracking is used to describe the institutional practice of separating students into groups that experience different instructional exposures. It has received considerable research attention in the K-12 literature. Many schools no longer have official policies to separate students formally into qualitatively distinct tracks (e.g., vocational track vs. academic track), but there are common and binding organizational practices that continue to sort students into qualitatively different individual courses (e.g., academic placement, honors, remedial) (Lucas, 1999).

One reason for the continuous research attention is that researchers consistently find that students from different socio-demographic backgrounds end up in different courses, studying different subjects, and being exposed to different peers (Armstrong & Hamilton, 2013; Bottia et al., 2016; Domina et al., 2019; Lucas & Berends, 2002). The other reason is that these differences in student experiences are associated with variation in student academic and labor market outcomes. In a highly tracked school, students are more likely exposed to peers with similar socio-demographic backgrounds and abilities, and experience greater variation in learning opportunities and instructional content across tracks (Gamoran, 1987; Moller et al., 2013; Moller & Stearns, 2012; Sorensen, 1970).

Although tracking has received less empirical examination in higher education than in K-12 education, it is arguably more explicit and pronounced. In K-12 education, courses under the same subject are usually divided into different levels both vertically (e.g., Introductory Physics vs. AP Physics) and horizontally (e.g., Mandarin vs. Spanish) (Ferrare, 2013). In higher education, tracking is commonly perceived as more student-directed, meritocratic, and

preference-based. By allowing students to make individual choices based on their interests and ability and implementing different requirements (e.g., prerequisites), tracking in higher education is rationalized as providing curricular specialization in response to students' interests, skillsets, and autonomy. However, akin to tracking in K-12, these "personal choices" are influenced by the institutional constraints and curricular structures.

University structures curriculum in multiple ways, including major requirements, general education requirements, and elective course requirements (Scott-Clayton, 2015). Multiple factors systematically play roles in this process which leads to observed gender segregation across and within majors. Mann and DiPrete (2013) have found gender difference in curricular goals even within STEM fields. Female students are more likely than their male counterparts to value a diverse set of courses and liberal arts curriculum, which can be achieved through elective courses and general education. Some fields of study have competitive admissions processes, and some have intentionally challenging introductory coursework, which functions to deter and redirect students to other majors (Astorne-Figari & Speer, 2019). Female students might be more susceptible to this highly competitive environment or may perceive a higher opportunity cost associated with failing certain courses. Majors have different concentrations, levels of difficulty, prerequisite structures, course schedules, instructors, etc. In the following sections, we will discuss how students are sorted at different levels and why it is important to look at gender sorting in higher education both at the major and course level.

Gender Sorting at Major Level

Gender segregation between majors is a long-standing feature of higher education (Barone, 2011; Barone & Assirelli, 2020; Bobbitt-Zeher, 2007; DiPrete & Buchmann, 2013; Ganley et al., 2018; Jacobs, 1986; Penner, 2015; Speer, 2017; Xie et al., 2015). The gap in

college enrollment between genders is narrowing; in certain countries (e.g., the U.S.), female students even outperform male students in overall college attainment. However, gender segregation between majors has persisted (DiPrete & Buchmann, 2013; England & Li, 2006). Female students are underrepresented in most STEM fields such as physical science, engineering, math, and statistics. Specifically, only around 24% of bachelor's degrees in engineering were earned by women in 2020, compared with 19% in 2011 (NCSES 2022). Female students are more likely to choose majors with lower labor market returns such as education, arts, and humanities, compared with more lucrative majors such as economics, and engineering (Bobbitt-Zeher, 2007; Petersen et al., 2014). Major choice has been demonstrated to have lasting consequences for an individual's career, economic success, and peers with which a student interacts (Black et al., 2008; Bobbitt-Zeher, 2007; Leighton & Speer, 2020).

Of the multiple factors that contribute to student choice of major, pre-college academic preparation, prior course-taking experiences, perceived costs and preferences, and future expectations about career and family play the largest roles (Astorne-Figari & Speer, 2019; Xie & Shauman, 2005). Students' pre-college course experience and performance partially inform how likely students are to expect themselves to be interested and successful in relevant fields of study. Gender differences are found in pre-college course taking and course performance. Female students are more likely to pursue high school curriculum tracks in the humanities when compared with their male counterparts (Chen, 2013). These pre-college factors are directly related to student choice of major in the universities they attend.

However, recent research has clearly demonstrated that prior ability is not the only factor in explaining gender differences in choice of major (Chen, 2013; Speer, 2017; Stearns et al., 2020; Xie & Shauman, 2005). Students' broader perceptions of the character of academic

programs is also a basis of student choice. Astorne-Figari and Speer (2019) found that female college students are more likely to switch to majors that "look like them,"–i.e., majors with a higher density of female students. England and Li (2006) documented that the trend of feminization of fields is negatively associated with male entrance. Similarly, the level of perceived gender biases against women, such as discrimination, barriers to success, and sense of belonging, is negatively correlated with the proportion of female students within a major (Ganley et al., 2018). Finally, the choice of major is affected by students' future orientations with respect to career plans, expectancy of success, gender roles, and family responsibilities (Eccles, 1994; Jacobs, 1995; Morgan et al., 2013; Quadlin, 2020; Weeden et al., 2020).

Understanding how multiple factors contribute to gender segregation at the level of major choice helps us to establish the foundational importance of major choice in the observed variation of students' classmates' composition. We note, though, that there is evidence that these factors may also affect sorting by gender into classes within majors.

What Do We Know About Gender Sorting Between Classes for Students Enrolled in the Same Majors?

Majors are, in some ways, simply administrative abstractions, as classes are the true location of instruction and social interaction. But they are important administrative structures, as the way a university designs its major requirements and general education requirements influences course-taking patterns even for different students enrolled in the same major. These factors can exacerbate the gender and racial/ethnic segregation that comes from sorting across majors; given the higher flexibility of course choices in universities (Chaturapruek et al., 2021; McFarland, 2006), students in the same major may end up choosing very different courses. However, compared with examinations of gender sorting across majors, relatively little work has

examined the curricular experiences of students enrolled in the same major.

Both descriptive and causal evidence from K-12 and higher education show gender disparities in course-taking behaviors. Using state-wide high school data in Florida, Conger, Long, and Iatarola (2009) found gender differences in AP course-taking even after controlling for critical factors (e.g., pre-high school characteristics, high school fixed-effects). Drawing on nationally representative longitudinal survey data, Penner and Willer (2019) documented that female college students are less inclined to retake STEM courses after failing them, revealing a gender difference in course-taking patterns. The authors hypothesize that this could be due to gendered beliefs that STEM fields are male domains. Past work has also shown gender differences in course-taking based on course content. For example, Ferrare (2013), studying one United States high school, found that courses relevant to the production of meaning such as literature, art, and household skills are more likely to be taken by female students, while courses related to the transmission of technical knowledge, business preparation, and construction skills are more likely to be taken by male students. A few studies have used randomized controlled trials to provide causal evidence on the effects of various course characteristics on students' educational choices. Female students are more likely to enroll in future STEM courses if they have female instructors as compared with having male instructors (Carrell et al., 2010). Conversely, female students are less likely to persist and get a degree in STEM majors if they have a larger proportion of high-ability classmates. The authors hypothesize that this is potentially due to self-perception about future success (Fischer, 2017).

Apart from studies focusing on course-taking behaviors, a handful of studies use survey and interview data to explore students' perspectives on course selection and their decision-making processes. This body of research finds that students choose courses for a variety of

reasons, including both objective factors such as fulfillment of major requirements and course schedules, and more subjective factors such as perception of potential career benefits and future development, course difficulty, and the instructor-student relationship (Chambliss & Takacs, 2014; Galotti & Umscheid, 2019; Lee et al., 2021; Samara, 2015). Past work has found differences across genders in these decision process. For example, female students tend to have more criteria that they consider when making course decisions (Baker & Orona, 2020). Female students tend to place greater value on whether the courses fulfill certain course requirements and on obtaining good grades when choosing courses, compared with their male counterparts (Galotti & Umscheid, 2019; Lee et al., 2021). These findings provide initial evidence that, even within majors, course-taking preferences associated with demographic characteristics might lead to meaningful segregation. Examining gender segregation only between majors might miss important patterns.

Importance of Peer Gender Composition

Exposure to diverse gender compositions can be influential to student outcomes. First, classroom gender composition has itself been repeatedly found to affect students' educational outcomes, such as academic performances and major choices (Hoxby, 2000; Lavy & Schlosser, 2011; Park et al., 2013; Zölitz & Feld, 2020). Second, female students are more likely influenced by their same-gender peers in terms of course-taking and course preference (Frank et al., 2008; Raabe et al., 2019; Riegle-Crumb et al., 2006; Rosenqvist, 2018). Given the potential influences of the gender composition of students' classmates, it is important to better understand the dynamics of systematic gender sorting of courses within majors.

Prior literature shows that gender can not only play a role in major choices, but also in course-taking patterns for students enrolled in the same major. This study examines the extent to

which within-major sorting contributes to overall gender sorting in university. This study makes a significant contribution to the literature; while previous research has focused on sorting at the major level, our data allow us to examine how student course choices are related to gender composition of classmates and how this varies by majors, course-level features, and individual student characteristics.

DATA AND METHODS

Data

In this study, we use administrative data from a large selective four-year university in the American West. The institution is a Hispanic-Serving Institution with an annual total undergraduate enrollment around 22,000 from 2011 to 2015. The administrative data includes enrollment and performance data for every course taken by each student in our sample. It also contains pre-college academic and background data from student applications, including self-identified sex, race and ethnicity, and low-income status—based on options provided by the university—along with pre-college standardized test scores, transfer and international student status, age, and first-generation college student status.

Our analytical sample includes students who entered the focal university between 2011 and 2015. We set the cohort sample restrictions because the 2011 cohort is the earliest cohort for whom we have complete data and stopping with the 2015 cohort allows students six full years to graduate based on current data availability. Transfer students, students with multiple majors and undeclared students are all included in our sample. Students in our sample are enrolled in a total of 100 majors.

Table 1.1 provides summary statistics describing our sample of the 34,882 students who enrolled at the university from 2011 to 2015. Over half of the students in our sample are

classified as female (54%) or first-generation college students (53%). Furthermore, 35% of students are identified as low-income, and 26% are transfer students. Regarding the ethnic composition of the student sample, 26% are identified as Latino/a, 53% as Asian, 17% as White, and 3% as Black. International students make up 20% of the sample and the majority of them are from Asian countries.

Outcome: Same-Gender Peer Composition

The outcome in our analysis is the proportion of a student's classmates who are of the same gender, measured at the student-by-term-by-course level. Given the restrictions of how the gender data are collected, we only consider students who self-identified as female or male[1]. The focal student is excluded when calculating the total number of same-gender peers and the total number of students in a class. We also exclude specific courses, such as independent study courses, where there is limited potential for peer interaction. Furthermore, when classes include multiple sections (e.g., a lecture and a lab), we only include the main section in the calculation. Our outcome is calculated separately for three categories of courses students are enrolled in: 1) all courses, 2) in-department courses, 3) out-of-department courses. In-department courses are defined as courses offered by the focal student's major home department. For example, for a student majoring in Biological Science, courses offered by this department would be classified as in-department courses.

As shown in Table 1.2, on average, female students have a similar proportion of female classmates in all courses, in-department courses, and out-of-department courses, with the largest variation found in in-department courses. Specifically, they have a proportion of 0.60 (SD =

[1] Since the non-binary gender option was not added until Fall 2019, only 0.33% of students identify as gender non-binary.

0.16) female classmates in all courses, 0.60 (SD = 0.18) in in-department courses, and 0.59 (SD = 0.15) in out-of-department courses.

Male students have the largest proportion and variation of male classmates in in-department courses, suggesting that the variation for their same-gender peers is mainly driven by in-department courses. The larger variation than female students also suggest more pronounced segregation across courses. that Specifically,

on average have a proportion of 0.53 (SD = 0.20) male classmates in all courses, a proportion of 0.57 (SD = 0.22) male classmates in in-department courses, and a proportion of 0.50 (SD = 0.18) male classmates in out-of-department courses.

As shown in Table 1.2, on average, female students have a proportion of 0.60 (SD = 0.16) female classmates in all courses, a proportion of 0.60 (SD = 0.18) female classmates in in-department courses, and a proportion of 0.59 (SD = 0.15) female classmates in out-of-department courses. Male students on average have a proportion of 0.53 (SD = 0.20) male classmates in all courses, a proportion of 0.57 (SD = 0.22) male classmates in in-department courses, and a proportion of 0.50 (SD = 0.18) male classmates in out-of-department courses.

As we noted previously, we use in-department courses as a proxy for courses required by certain majors.[2] The advantage of this approach is that it allows for simple classification of in- and out-of-department courses for all students and strengthens the generalizability of our findings and enhances the potential for replication. However, a limitation of this approach is that it may not fully capture major-required courses since many majors required courses outside of

[2] For example, if a student from Biological Science takes a course called BIOSCI 101, we will classify this course as an in-department course for the student as a proxy for a major requirement. However, if the same student takes MATH 101, we will classify this course as an out-of-department course, even though MATH 101 may also be listed as a major requirement in the catalog.

their home department. To address this concern, we manually coded courses from the general catalog that can be used to fulfill a major, specialization, or school requirement for the 18 largest majors. We then reanalyzed the data based on these manually coded courses and discussed the sample, methods, and results in Appendix A3. The results are generally comparable between these two approaches.

Analytic Strategy

To examine our research questions, we start by using kernel density plots to separately visualize the distribution of the proportion of same-gender classmates for female and male students. These plots illustrate how much variation there is in the proportion of same-gender classmates across students in our sample. Furthermore, to show how such variation differs across and within majors, we plot the 10th percentile, median, and 90th percentile of students' same-gender classmates' proportion for each major for (1) all courses, (2) in-department courses, and (3) out-of-department courses.

Ordinary least squared regression (OLS) model

We use student-by-term-by-course OLS regression models to examine whether the proportion of a student's classmates that share his/her gender is associated with certain demographic characteristics and to examine how much variation can be explained by these demographic factors. We use the following model:

$$PropSameGender_{ijcgt} = \beta X_i + \gamma_g + \upsilon_t + \varepsilon_{ijgt} \quad [1]$$

In equation [1], $PropSameGen_{ijcgt}$ represents the proportion of same gender classmates

for student i in term j, course c, year-of-enrollment[3] g (year 1, year 2, year 3, year 4, year 5 or more), cohort t; X_i is the vector of student-level covariates which include students' gender, ethnicity, age when they enrolled in the university, first-generation status, low-income status, international student status, transfer student status, and standardized pre-college academic performance; γ_g represents year-of-enrollment fixed effect; υ_t represents cohort fixed effect from cohort 2011 to cohort 2015; ε_{ijgt} represents the error term clustered at the student, term, grade, and cohort level.

In the analysis, we estimate models separately for female students and male students. Separating the analyses by gender identity allows us to have a clearer understanding of patterns of sorting. For example, if male students tend to take classes with peers who have higher pre-college academic performance, and female students with peers who have lower pre-college academic performance, the proportion of classes taken with high average pre-college academic performance will be predictive of more same-gender classmates for male students and fewer same-gender classmates for female students. A pooled model might mask both dynamics and fail to show important relationships.

Major fixed-effect models for different courses

Given that the classmates a student encounters are largely decided by the student's major, we next include major fixed effects in the analysis and observe the change to the percent of variation explained (i.e., the r-squared) by student sorting into majors.

[3] Year-of-enrollment represents the accumulated years students have enrolled in the university, it captures the difference in college experience among students from different grades. Cohort indicates the year students firstly enrolled in the university; it captures the time trend across different cohorts.

Equation [2] shows the major-fixed effect model. It includes an additional term as compared with equation [1], θ_m, which represents the major-fixed effect. The standard error is clustered at student, term, year-of-enrollment, cohort, and major level.

$$PropSameGender_{ijcgt} = \beta X_i + \theta_m + \gamma_g + \upsilon_t + \varepsilon_{ijgtm} \quad [2]$$

A major goal of this study is to document how much gender sorting there is within majors and how different types of courses explain this variation. A concern with the analyses presented in Equation [2] is that the remaining variation after major fixed-effect could be due to the classes a student is taking outside-of his/her home department; two students in the same major could have very similar peers in the in-department classes they take, but very different peers in their out-of-department courses for general education or elective courses. To explore the extent to which these out-of-department courses might be contributing to unexplained variation, we repeat the final set of analyses (those presented in Equation [2]), focusing solely on in-department courses. This allows us to examine the level of variation in the proportion of students' same-gender classmates within their major, specifically in these in-department classes. For comparative purposes, we perform the same analysis for out-of-department courses.

FINDINGS

Variation in the Proportion of Same-Gender Classmates

We first examine how much variation exists in the proportion of same-gender classmates across students in our sample. Figure 1.1 shows the kernel density distribution of the proportion of same-gender classmates for both female and male students using the student-by-term-by-course data. The theoretical range for the horizontal axis extends from zero (0% same-gender classmates) to one (100% same-gender classmates). The mean and standard deviation for the proportion of same-gender classmates (see Table 1.2) for female students are 0.60 and 0.16.

Male students have a slightly wider distribution with a mean of 0.53 and a standard deviation of 0.20. From the shape of the distribution, we can observe that the distribution of the proportion of same-gender classmates among male students is less concentrated around the mean, showing a bimodal distribution. We will discuss potential causes for this in the discussion section.

In the sample, 54% of students identify as female. If students were randomly assigned to courses, a male student would be expected to have 46% same-gender peers, while a female student would be expected to have 54% same-gender peers. According to Figure 1.1, female students experience course compositions that approximate a normal distribution around the mean of approximately 60% (6% higher than expected). For male students, we observe a bimodal distribution. A large proportion of male students have around 40% male classmates, while there is another peak of around 80% male classmates. These patterns indicate the existence of gender segregation in students' course-taking patterns. Through enrolling in different courses, driven by a combination of curricular constraints and personal preferences, students are sorted into environments that are different from the overall gender composition of the campus.

This overall figure could mask the differences in the patterns across students' year of study. We further examine these patterns by disaggregating the overall distribution by year of study (e.g., students in their second year of enrollment). Appendix A1 shows how the distribution differs across students' first, second, third, fourth, and fifth years in the university. These distributions show that the bimodal trends for male students become more prominent in their later college years, and a relatively similar trend for female students across college years.

Do We See More Gender Segregation in Certain Demographic Groups?

In this section, we examine if the proportion of a student's classmates who are their same gender can be explained by sociodemographic characteristics. Columns 1 and 2 in Table 1.3

present the results from ordinary least squared (OLS) regression models in which the outcome is a student's proportion of same-gender classmates in a given course. The predictors include students' demographic characteristics: race/ethnicity, age, first-generation college student status, low-income status, international student status, transfer student status, pre-college academic performance, and sets of cohort and year-of-study fixed effects. We present the results for female and male students separately.

Gender segregation differs across demographic groups, and some of these differences persist even after controlling for major fixed-effects. On average, both female and male international students, older students, and students with higher pre-college performance all have a lower proportion of female classmates compared with their counterparts who are non-international students, younger, or have lower pre-college performance. Specifically, female international students have 4.9 percentage points fewer female classmates than domestic female students, while male international students have 2.9 percentage points more male classmates than their non-international counterparts. Regarding differences by ethnic group, Asian female students have 2.4 percentage points fewer female classmates compared with their White counterparts. Similarly, Asian, Latino, and Black male students, all have a lower proportion of male classmates compared with their White counterparts. Furthermore, female transfer students, on average, have 1.5 percentage points fewer female classmates, while no significant difference has been observed for male transfer students.

How Does the Proportion of Same-Gender Classmates Vary Across Majors?

Given that past work has demonstrated significant gender segregation across fields of study in higher education (Barone, 2011), we first descriptively examine the proportion of students' classmates who are the same gender by major. We present the results for the largest

majors in Figure 1.2[4]. In this figure, each bar presents three data points: proportion of students' same-gender classmates at the 10th percentile, the median, and the 90th percentile in that major (each of these statistics is shown in the first three columns in Appendix A2).

Two features stand out in Figure 1.2. First, echoing past work, the figure shows that students in certain majors have substantially different proportions of same-gender classmates compared with other majors. For example, female students in Mechanical Engineering at the median have 19% female peers, compared with 72% for female students at the median in Psychology and Social Behavior. Major choice is indeed consequential in determining students' classmates gender composition. Second, while there are large differences between certain pairs of majors, many majors are tightly clustered and have similar medians. For example, the medians across eight majors, from Chemistry to Biological Science, fall within a 10-percentage-point range. Differences across three-quarters of the majors listed are within a 20-percentage-point range; students from different majors could end up having a similar proportion of same-gender classmates.

We formalize these results by estimating major fixed-effect models. By comparing the percentage of variation being explained by models with only demographic predictors to the percentage explained by major fixed-effect models, we can examine how much of the overall variation is explained by gendered sorting into majors.

Major choice explains much of the gender sorting we observed. From the original OLS model that controls for only student-level demographic characteristics (shown in columns 1 and 2 in Table 1.3) to the major fixed-effect models (shown in columns 3 and 4 in Table 1.3), we explain an additional 35% of the variation for female students and 48% of the variation for male

[4] We indicate STEM fields with triangles, and non-STEM fields with circles.

students.

Within-Major Variation in the Proportion of Same-Gender Classmates

Another important finding presented in Figure 1.2 is that there is considerable variation *within* majors, and the amount of variation differs across majors. The range of distance between the 10th percentile and the 90th percentile ranges from 21% (Pharmaceutical Sciences) to 46% (Computer Science) for female students and ranges from 20% (Pharmaceutical Sciences) to 43% (Computer Science) for male students. Appendix A2 includes detailed statistics for each major. This variation indicates that the structure of majors and the degree of freedom of curricular choice within majors can lead to quite variable experiences for students in the same major.

Significantly, the variation of same-gender classmates within each major is almost commensurate with the differences in median values observed across the majority of fields, except those exhibiting the most severe gender imbalances. This comparison suggests that the student experience regarding classmate gender composition can exhibit substantial heterogeneity, even within a single major.

The results of major fixed-effect models (columns 3 and 4 in Table 1.3) echo that there exists substantial variation remaining within majors. Although the variation in students' same-gender classmates' proportion has been explained substantially by students' major choices, considerable variation remains unexplained (61% variation for female students; 48% variation for male students). This indicates that even within majors, there is still variation in the proportion of same-gender classmates as discussed in the previous section of Figure 1.2.

The major fixed effect models in Columns 3 and 4 in Table 1.3 allow us to examine whether there are still differences across groups of students even after controlling for student major choices. In these models, which compare students in the same major, we still see more

gender segregation in certain groups of students. Some patterns observed in the initial OLS models persist in major fixed-effect models, though with smaller coefficients.

Within majors, students with higher pre-college performance and international female students are still more likely to have a lower proportion of female classmates compared with their counterparts: those with lower pre-college performance and domestic female students, respectively. Unlike the models without major fixed effects, we no longer observe the significant differences for older students and male international students, compared with their respective counterparts. In terms of differences by ethnic group, Asian female students continue to have a lower proportion of female classmates compared with their White counterparts. Unlike the previous results, Latino/a students both have a higher proportion of female peers *within* majors. Another new pattern that emerges is that male students classified as low-income are more likely to have a higher proportion of female classmates.

We further apply restrictions to the major fixed-effect model, focusing solely on student in-department courses, and the corresponding results are presented in the first two columns of Table 1.4. For female students, 72% of the variation in the proportion of their same-gender classmates is accounted for, while this figure rises to 84% for male students. Notably, substantial unexplained variation persists even when focusing solely on in-department courses. For female students, a quarter of the total variation remains unexplained, and for male students, 16% of the total variation remains unexplained.

Role of Classes Students Take Outside Their Major

Figure 1.2 and results from major fixed-effect models show that there is considerable variation *within* majors, the remaining variation in the proportion of same-gender classmates for students in the same majors could be explained by a combination of gender sorting in major-

required courses (particularly in majors with flexible requirements), and gender sorting in courses that fulfill non-major requirements, such as general education or elective courses.

We plot, for each major, three statistics (10th percentile, median, 90th percentile) related to the proportion of same-gender peers for both in-department courses and out-of-department courses in Figure 1.3. By comparing the solid line (in-department courses) and the dashed line (out-of-department courses), we can see that the distribution of same-gender classmates is relatively similar across majors for the out-of-department courses. For highly gender-segregated majors, out-of-department courses serve as a balancing force, bringing the proportion of students' same-gender classmates toward the average campus level.

In-department courses play a more decisive role in the proportion of students' classmates who are their same gender, as compared with out-of-department courses. However, even within in-department courses, there is substantial variation in the proportion of same-gender classmates. For example, among female students in Political Science, some students have 43% female classmates in their in-department courses while some have 65% female classmates. The large within-major variation indicates that students have some degree of choice. Even *within* the courses taken in their home department, students' course choices might lead to sorting along gender lines. Considering the concern that in-department courses may display a pattern that varies significantly from that of actual major-required courses, we also replicated the bar graphs using manually coded major-required data (see Appendix Figure A3.1 and Figure A3.2). The patterns show some differences for majors disproportionately dominated by one gender, detailed discussion can be found in Appendix A3.

In columns three and four of Table 1.4, we present another set of major fixed-effect models, focusing exclusively on students' out-of-department courses. As expected, major fixed-

effects account for a modest proportion of the variation in the proportion of same-gender classmates for both female and male students. This figure is less than 30%, a stark contrast to the over 70% of variation explained for in-department courses (see columns 1-2 in Table 1.4).

CONCLUSION

Horizontal gender segregation in higher education has garnered significant attention in recent decades (Baker et al., 2023; Charles & Bradley, 2002; Ferrare, 2013; Gerber & Schaefer, 2004; Iannelli et al., 2018). Female and male students persistently sort into different fields of study, which leads to differential access to peers, occupational segregation, and income gaps. Prior studies have concentrated on gender segregation across majors. While the role of choice of major as a determinant of segregation is very important, focusing on the major as a level of analysis masks the gender segregation within majors at the course level, the context at which actual learning and social interactions occur. This study extends the horizontal sorting literature by providing evidence about gender segregation at the course level within majors. Specifically, we use data from five cohorts of students at a selective four-year university in the American West to examine how much variation in the proportion of students' same-gender classmates can be explained by students' demographic characteristics, major, as well as the remaining variation that occurs within each major.

Variation in the proportion of students' same-gender classmates exhibits different patterns in courses from students' home department and out-of-department courses. More differences exist in the medians of students' proportion of same-gender classmates across majors in in-department courses. However, for students from the same major, there is more variation in the proportion of same-gender classmates for the classes students take outside their department than inside their department. This indicates that while a student's choice of major is highly

determinative of the peers to whom they are exposed, course choice within the department, and perhaps even more so course choice outside the department (through electives and general education requirements) are additional factors associated with peer composition. The highly overlapping distributions for out-of-department courses across majors shown in Figure 1.3 suggest the role of out-of-department courses in balancing students' same-gender classmates proportion in highly gender-segregated majors. For example, female students in Mathematics have a median of 39% of female classmates in their in-department courses but have a median of 53% of female classmates in their out-of-department courses, which leads to 47% of female classmates in their overall courses.

We examine how different factors account for the observed variation: students' demographic characteristics alone explain 4% of the overall variation for female students and 4% for male students. Echoing past work, we find that major choice explains significant variation in peer composition: 39% of variation in the proportion of same-gender peers is explained for female students and 52% for male students. However, if we simply look at the gender composition at the major level, it could mask the variation around this mean, which is jointly affected by curricular structures and course choices, etc. For example, let's consider students in Computer Science and Mathematics. Around 21% of students in Computer Science are female. This is a statistic we've calculated from student-by-term records, accounting for students who have changed majors. In contrast, Mathematics has a 46% female student population. Interestingly, as shown in Figure 1.2, almost 40% of female students in Computer Science have a proportion of female peers comparable to that of female students in Mathematics. Major choice is predictive of the gender composition of peers but is far from determinative.

We also find that classes taken outside of students' home department do not fully explain this variation. Focusing on only in-department courses, there is still 28% of variation unexplained for female students and 16% for male students. We demonstrate that gender sorting not only exists across majors but also exists across courses within majors.

Gender sorting shows different patterns for students from different demographic groups, even after we control students' major choices. For example, Asian female students (compared with White female students) and international female students (compared with non-international female students) have a lower proportion of female classmates. For male students, Latino students (compared with White students) and low-income students (compared with non-low-income students), are associated with lower proportions of male classmates. Students with higher pre-college performance are associated with lower proportion of female classmates.

We know from past research that a number of factors predict student course choices (e.g., instructor gender, reputation, and classroom climate) (Brown & Kosovich, 2015; Galotti & Umscheid, 2019). Future study should consider incorporating more nuanced course-level features to capture students' course-choice behaviors. This study does not explicitly examine how course content might lead to sorting within majors. Within the requirements for any given major, students have the ability to select courses that might vary substantially in subject matter. Future work could explicitly examine gendered preferences for course content. In-depth interviews with students could examine this with individual students or large-scale text-mining exercises could look for patterns using course descriptions (Harrison et al., 2022).

Also, a major limitation of our study is that it does not include student voices. We are not able with administrative data alone to understand fully the mechanisms underlying students' major and course choices. For example, it is not clear why Asian students, compared with their

White peers in their same major, have fewer female classmates on average. This study provides us information about what predicts students' same-gender classmates' proportion and hopefully will serve as a catalyst for more nuanced qualitative work in the future.

Another interesting result which could be further studied is that in the models with major fixed-effects, more variation is explained in models for male students than in models for female students. This result echoes findings in Mann and DiPrete (2013), which indicate that even within STEM fields, female students exhibit a more diverse range of curricular choices compared with their male counterparts. A relevant explanation could be that female students are concentrated in majors with greater flexibility in major requirements, so students in the same major could take different courses to fulfill major requirements. As shown in Appendix Table A3.1, besides majors that have similar proportions of female and male students, a higher proportion of female students are in majors such as Public Health Sciences with 22% of all courses in in-department courses; while a higher proportion of male students are in majors, such as Mechanical Engineering with 58% of in-department courses. Future analysis can dig deeper into the gender differences in factors explaining student course peer gender composition.

Moreover, the methods presented in this study could be utilized to study how major choices and course choices affect the proportion of students' classmates of other important student characteristics, such as ethnicity/race, first-generation college status, or pre-college academic performance. Similar analyses can also be applied to multiple universities to compare how different predictors of students' peer composition may vary across schools with different cultures and institutional settings. Different student bodies (such as campuses with higher proportion of international or first-generation students) and varying curricular structures (such as

fields of specializations and availability of honors programs across schools) may contribute to students sorting into courses in unique ways.

Demonstrating gender sorting both among and within majors deepens our understanding of horizontal sorting and gender segregation in higher education. By revealing another dimension of gender sorting within majors, this study can draw more research attention to the importance of understanding student experience at a student-by-course level. This perspective more accurately captures the real-time experiences of gender segregation that students face within their selected courses. Even within STEM majors, gender sorting still exists; certain groups of students are more likely concentrated in certain courses with different instruction, curricular content and student learning.

This study can help policymakers and university administrators recognize how the flexibility of course offerings affect student sorting. The flexibility of major requirements in higher education, elective courses, general education, and specialization, creates a broad range of curricular pathways as compared with K-12 education (Chaturapruek et al., 2021; McFarland, 2006). Table A3.1 shows the average percentage of classes students took each term that were from their home departments separately for each of the listed majors. There is substantial variation across majors. For example, 82% of the courses students take each term in Mechanical Engineering are major-specific requirements; while for Education Sciences, only 45% of the courses students took each term are major-specific requirements. The remaining courses students took could fulfill elective requirements, general education requirements, or requirements for an additional major. The differences in the extent of structural constraints imposed by curricular requirements across majors allow different degrees of freedom for students to make their own course choices. Students' preferences, and the role that social and psychological constraints can

play, are affected by the structure of curricular requirements. More flexible structures provide students with greater options in terms of curricular choices, but may also lead to greater sorting. Understanding factors associated with course-level gender segregation could also help us to develop potential interventions to create a more diverse university experience.

REFERENCES

Armstrong, E. A., & Hamilton, L. T. (2013). *Paying for the Party*. Harvard University Press.

Astorne-Figari, C., & Speer, J. D. (2019). Are changes of major major changes? The roles of grades, gender, and preferences in college major switching. *Economics of Education Review*, *70*. https://doi.org/10.1016/j.econedurev.2019.03.005

Bahr, P. R. (2019). The Labor Market Returns to a Community College Education for Noncompleting Students. *Journal of Higher Education*, *90*(2). https://doi.org/10.1080/00221546.2018.1486656

Baker, R. B., Solanki, S. M., & Kang, C. (2023). Horizontal and Vertical Racial Segregation in Higher Education: Examining Trends in California Public Colleges. *Journal of Higher Education*, *94*(1). https://doi.org/10.1080/00221546.2022.2131970

Baker, R., & Orona, G. A. (2020). Gender and Racial Differences in Awareness and Consideration of Curricular Programs: Exploring a Multistage Model of Major Choice. *AERA Open*, *6*(3). https://doi.org/10.1177/2332858420937023

Barone, C. (2011). Some things never change: Gender segregation in higher education across eight nations and three decades. *Sociology of Education*, *84*(2). https://doi.org/10.1177/0038040711402099

Barone, C., & Assirelli, G. (2020). Gender segregation in higher education: an empirical test of seven explanations. *Higher Education*, *79*(1). https://doi.org/10.1007/s10734-019-00396-2

Black, D. A., Haviland, A. M., Sanders, S. G., & Taylor, L. J. (2008). Gender wage disparities among the highly educated. *Journal of Human Resources*, *43*(3). https://doi.org/10.3368/jhr.43.3.630

Bobbitt-Zeher, D. (2007). The gender income gap and the role of education. *Sociology of Education*, *80*(1). https://doi.org/10.1177/003804070708000101

Bottia, M. C., Giersch, J., Mickelson, R. A., Stearns, E., & Moller, S. (2016). Distributive Justice Antecedents of Race and Gender Disparities in First-Year College Performance. *Social Justice Research*, *29*(1). https://doi.org/10.1007/s11211-015-0242-x

Brown, C. L., & Kosovich, S. M. (2015). The Impact of Professor Reputation and Section Attributes on Student Course Selection. *Research in Higher Education*, *56*(5). https://doi.org/10.1007/s11162-014-9356-5

Carrell, S. E., Page, M. E., & West, J. E. (2010). Sex and science: How professor gender perpetuates the gender gap. *Quarterly Journal of Economics*, *125*(3). https://doi.org/10.1162/qjec.2010.125.3.1101

Chambliss, D. F., & Takacs, C. G. (2014). *How College Works*. Harvard University Press.

Charles, M., & Bradley, K. (2002). Equal but Separate? A Cross-National Study of Sex Segregation in Higher Education. *American Sociological Review*, *67*(4), 573. https://doi.org/10.2307/3088946

Chaturapruek, S., Dalberg, T., Thompson, M. E., Giebel, S., Harrison, M. H., Johari, R., Stevens, M. L., & Kizilcec, R. F. (2021). Studying Undergraduate Course Consideration at Scale. *AERA Open*, *7*. https://doi.org/10.1177/2332858421991148

Chen, W. (2013). Causes and Consequences of High School Curriculum - Track Selection: Gender, Belief, Teacher' s Gender, and Cognitive Development. *Taiwanese Sociology*, *25*, 89–123.

Conger, D., Long, M. C., & Iatarola, P. (2009). Explaining race, poverty, and gender disparities in advanced course-taking. *Journal of Policy Analysis and Management*, *28*(4). https://doi.org/10.1002/pam.20455

Davies, S., & Guppy, N. (1997). Fields of study, college selectivity, and student inequalities in higher education. *Social Forces*, *75*(4). https://doi.org/10.1093/sf/75.4.1417

DiPrete, T. A., & Buchmann, C. (2013). The rise of women: The growing gender gap in

education and what it means for American schools. In *The Rise of Women: The Growing Gender Gap in Education and what it Means for American Schools* (Vol. 9781610448000). https://doi.org/10.1177/0094306114562201m

Domina, T., McEachin, A., Hanselman, P., Agarwal, P., Hwang, N. Y., & Lewis, R. W. (2019). Beyond Tracking and Detracking: The Dimensions of Organizational Differentiation in Schools. *Sociology of Education*, *92*(3). https://doi.org/10.1177/0038040719851879

Eccles, J. S. (1994). Understanding women's educational and occupational choices: Applying the eccles et al. model of achievement-related choices. *Psychology of Women Quarterly*, *18*(4). https://doi.org/10.1111/j.1471-6402.1994.tb01049.x

England, P., & Li, S. (2006). Desegregation stalled: The changing gender composition of college majors, 1971-2002. *Gender and Society*, *20*(5). https://doi.org/10.1177/0891243206290753

Ferrare, J. J. (2013). The Duality of Courses and Students: A Field-theoretic Analysis of Secondary School Course-taking. *Sociology of Education*, *86*(2). https://doi.org/10.1177/0038040712456557

Fischer, S. (2017). The downside of good peers: How classroom composition differentially affects men's and women's STEM persistence. *Labour Economics*, *46*. https://doi.org/10.1016/j.labeco.2017.02.003

Frank, K. A., Muller, C., Schiller, K. S., Riegle-Crumb, C., Mueller, A. S., Crosnoe, R., & Pearson, J. (2008). The social dynamics of mathematics coursetaking in high school. *American Journal of Sociology*. https://doi.org/10.1086/587153

Galotti, K. M., & Umscheid, V. A. (2019). Students choosing courses: Real-life academic decision making. *American Journal of Psychology*, *132*(2). https://doi.org/10.5406/amerjpsyc.132.2.0149

Gamoran, A. (1987). The Stratification of High School Learning Opportunities. *Sociology of Education*, *60*(3). https://doi.org/10.2307/2112271

Ganley, C. M., George, C. E., Cimpian, J. R., & Makowski, M. B. (2018). Gender Equity in College Majors: Looking Beyond the STEM/Non-STEM Dichotomy for Answers Regarding Female Participation. *American Educational Research Journal*, *55*(3). https://doi.org/10.3102/0002831217740221

Gerber, T. P., & Cheung, S. Y. (2008). Horizontal stratification in postsecondary education: Forms, explanations, and implications. In *Annual Review of Sociology* (Vol. 34). https://doi.org/10.1146/annurev.soc.34.040507.134604

Gerber, T. P., & Schaefer, D. R. (2004). Horizontal Stratification of Higher Education in Russia: Trends, Gender Differences, and Labor Market Outcomes. *Sociology of Education*, *77*(1). https://doi.org/10.1177/003804070407700102

Giani, M. S., Attewell, P., & Walling, D. (2020). The Value of an Incomplete Degree: Heterogeneity in the Labor Market Benefits of College Non-Completion. *Journal of Higher Education*, *91*(4). https://doi.org/10.1080/00221546.2019.1653122

Harrison, M. H., Hernandez, P. A., & Stevens, M. L. (2022). Should I Start at MATH 101? Content Repetition as an Academic Strategy in Elective Curriculums. *Sociology of Education*, *95*(2). https://doi.org/10.1177/00380407221076490

Hoxby, C. M. (2000). Peer Effects in the Classroom: Learning from Gender and Race Variation. *NBER Working Paper*, *7867*.

Iannelli, C., Gamoran, A., & Paterson, L. (2018). Fields of study: Horizontal or vertical differentiation within higher education sectors? *Research in Social Stratification and Mobility*, *57*. https://doi.org/10.1016/j.rssm.2018.06.004

Jacobs, J. A. (1986). The Sex-Segregation of Fields of Study: Trends during the College Years. *The Journal of Higher Education*, *57*(2). https://doi.org/10.2307/1981478

Jacobs, J. A. (1995). Gender and Academic Specialties: Trends among Recipients of College Degrees in the 1980s. *Sociology of Education*, *68*(2). https://doi.org/10.2307/2112776

Lavy, V., & Schlosser, A. (2011). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics*, *3*(2). https://doi.org/10.1257/app.3.2.1

Lee, H. R., von Keyserlingk, L., Arum, R., & Eccles, J. S. (2021). Why Do They Enroll in This Course? Undergraduates' Course Choice From a Motivational Perspective. *Frontiers in Education*, *6*. https://doi.org/10.3389/feduc.2021.641254

Leighton, M., & Speer, J. D. (2020). Labor market returns to college major specificity. *European Economic Review*, *128*. https://doi.org/10.1016/j.euroecorev.2020.103489

Lucas, S. R., & Berends, M. (2002). Sociodemographic diversity, correlated achievement, and de facto tracking. *Sociology of Education*, *75*(4). https://doi.org/10.2307/3090282

Lucas, S. R. (1999). *Tracking Inequality: Stratification and Mobility in American High Schools*. Teachers College Press.

Mann, A., & DiPrete, T. A. (2013). Trends in gender segregation in the choice of science and engineering majors. *Social Science Research*, *42*(6). https://doi.org/10.1016/j.ssresearch.2013.07.002

McFarland, D. A. (2006). Curricular flows: Trajectories, turning points, and assignment criteria in high school math careers. *Sociology of Education*, *79*(3). https://doi.org/10.1177/003804070607900301

Moller, S., & Stearns, E. (2012). Tracking Success: High School Curricula and Labor Market Outcomes by Race and Gender. *Urban Education*, *47*(6). https://doi.org/10.1177/0042085912454440

Moller, S., Stearns, E., Southworth, S., & Potochnick, S. (2013). Changing course: the gender gap in college selectivity and opportunities to learn in the high school curriculum.

Gender and Education, *25*(7). https://doi.org/10.1080/09540253.2013.853028

Morgan, S. L., Gelbgiser, D., & Weeden, K. A. (2013). Feeding the pipeline: Gender, occupational plans, and college major selection. *Social Science Research*, *42*(4). https://doi.org/10.1016/j.ssresearch.2013.03.008

National Center for Science and Engineering Statistics. 2022. special tabulations (2022, Table Builder) of the National Center for Education Statistics, Integrated Postsecondary Education Data System, Completions Survey, provisional release data.

Park, H., Behrman, J. R., & Choi, J. (2013). Causal Effects of Single-Sex Schools on College Entrance Exams and College Attendance: Random Assignment in Seoul High Schools. *Demography*, *50*(2). https://doi.org/10.1007/s13524-012-0157-1

Penner, A. M. (2015). Gender inequality in science. In *Science* (Vol. 347, Issue 6219). https://doi.org/10.1126/science.aaa3781

Penner, A. M., & Willer, R. (2019). Men's Overpersistence and the Gender Gap in Science and Mathematics. *Socius: Sociological Research for a Dynamic World*, *5*. https://doi.org/10.1177/2378023118821836

Petersen, T., Penner, A. M., & Høgsnes, G. (2014). From motherhood penalties to husband premia: The new challenge for gender equality and family policy, lessons from Norway. *American Journal of Sociology*, *119*(5), 1434-1472.

Quadlin, N. (2020). From Major Preferences to Major Choices: Gender and Logics of Major Choice. *Sociology of Education*, *93*(2). https://doi.org/10.1177/0038040719887971

Raabe, I. J., Boda, Z., & Stadtfeld, C. (2019). The Social Pipeline: How Friend Influence and Peer Exposure Widen the STEM Gender Gap. *Sociology of Education*. https://doi.org/10.1177/0038040718824095

Riegle-Crumb, C., Farkas, G., & Muller, C. (2006). The role of gender and friendship in advanced course taking. *Sociology of Education*, *79*(3).

https://doi.org/10.1177/003804070607900302

Rosenqvist, E. (2018). Two Functions of Peer Influence on Upper-secondary Education Application Behavior. *Sociology of Education*, *91*(1). https://doi.org/10.1177/0038040717746113

Samara, F. (2015). Factors Influencing Students' Choice of Elective Science Courses: A Case Study from the American University of Sharjah. *Open Journal of Social Sciences*, *03*(08). https://doi.org/10.4236/jss.2015.38010

Scott-Clayton, J. (2015). The shapeless river: Does a lack of structure inhibit students' progress at community colleges? In *Decision Making for Student Success: Behavioral Insights to Improve College Access and Persistence*.

Sorensen, A. B. (1970). Organizational Differentiation Of Students And Educational Opportunity. *Sociology of Education*, *43*(4). https://doi.org/10.2307/2111838

Speer, J. D. (2017). The gender gap in college major: Revisiting the role of pre-college factors. *Labour Economics*, *44*. https://doi.org/10.1016/j.labeco.2016.12.004

Stearns, E., Bottia, M. C., Giersch, J., Mickelson, R. A., Moller, S., Jha, N., & Dancy, M. (2020). Do Relative Advantages in STEM Grades Explain the Gender Gap in Selection of a STEM Major in College? A Multimethod Answer. *American Educational Research Journal*, *57*(1). https://doi.org/10.3102/0002831219853533

Wang, M. Te, Eccles, J. S., & Kenny, S. (2013). Not Lack of Ability but More Choice: Individual and Gender Differences in Choice of Careers in Science, Technology, Engineering, and Mathematics. *Psychological Science*, *24*(5). https://doi.org/10.1177/0956797612458937

Weeden, K. A., Gelbgiser, D., & Morgan, S. L. (2020). Pipeline Dreams: Occupational Plans and Gender Differences in STEM Major Persistence and Completion. *Sociology of Education*, *93*(4). https://doi.org/10.1177/0038040720928484

Xie, Y., Fang, M., & Shauman, K. (2015). STEM Education. *Annual Review of Sociology*, *41*. https://doi.org/10.1146/annurev-soc-071312-145659

Xie, Y., & Shauman, K. A. (2005). *Women in Science: Career Processes and Outcomes*. Harvard University Press.

Zölitz, U., & Feld, J. (2020). The Effect of Peer Gender on Major Choice in Business School. *Management Science*. https://doi.org/10.1287/mnsc.2020.3860

Table 1. 1 Summary Statistics of Students' Characteristics and Outcomes (Student-Level)

Student characteristics	Mean	SD
Female student	0.543	0.498
Asian student	0.533	0.499
Latino/a student	0.260	0.438
White student	0.171	0.377
Black student	0.031	0.172
Students from other ethnic groups	0.006	0.075
International student	0.196	0.397
Age when enrolled as freshman	19.416	2.898
First generation status	0.534	0.499
Low-income status	0.354	0.478
Transfer student	0.259	0.438
Pre-college academic performance	223.796	29.046
cohort		
2011	0.192	0.393
2012	0.189	0.391
2013	0.205	0.404
2014	0.205	0.404
2015	0.210	0.407
Total years in college		
1	0.061	0.240
2	0.179	0.383
3	0.139	0.346
4	0.514	0.500
5	0.102	0.303
6	0.004	0.066
Observations	34,882	

Note. This analysis uses student-level records. Other ethnic groups include American Indian or Native American and Native Hawaiian or Pacific Islander.

Table 1. 2 Summary Statistics of Outcome Variables

Outcome variables	Female students			Male students		
	Mean	SD	Observations	Mean	SD	Observations
Same-gender classmate proportion	0.595	0.161	727,782	0.530	0.200	620,245
Same-gender classmate proportion for in-department courses	0.602	0.181	282,432	0.571	0.223	258,385
Same-gender classmate proportion for out-of-department courses	0.590	0.147	445,350	0.500	0.176	361,860

Note. This analysis uses student-by-term-by-course records.

Table 1. 3 Results of Ordinary Least Squares Models, Major Fixed-Effect Models on the Proportion of Students' Same-Gender Classmates for All Courses

	(1)	(2)	(3)	(4)
	For all courses			
	OLS		Major FE	
	Female	Male	Female	Male
Ethnicity (reference group: White)				
Asian	-0.024***	-0.012*	-0.005*	-0.002
	(0.002)	(0.004)	(0.002)	(0.002)
Latino/a	-0.001	-0.024*	0.006**	-0.008**
	(0.005)	(0.008)	(0.001)	(0.002)
Black	0.003	-0.038**	0.002	-0.006
	(0.005)	(0.010)	(0.002)	(0.003)
Others	0.003	-0.015	0.003	-0.006
	(0.010)	(0.013)	(0.005)	(0.005)
International student	-0.049***	0.029***	-0.014*	0.007
	(0.003)	(0.004)	(0.006)	(0.005)
Age (enrolled)	-0.001*	0.003***	-0.000	0.001
	(0.000)	(0.001)	(0.000)	(0.001)
First-generation status	0.004	0.002	0.002	-0.001
	(0.002)	(0.002)	(0.001)	(0.001)
Low-income status	-0.001	-0.001	0.000	-0.003*
	(0.002)	(0.003)	(0.001)	(0.001)
Transfer students	0.015**	-0.014	-0.005	0.011
	(0.003)	(0.011)	(0.003)	(0.006)
Standardized pre-college performance	-0.017***	0.026***	-0.002*	0.005*
	(0.002)	(0.003)	(0.001)	(0.001)
Control cohort	+	+	+	+
Control year of study (1st year, 2nd year, 3rd year, 4th year, 5th year or more in college)	+	+	+	+
Control major			+	+
Observations	727782	620245	727782	620245
Adjusted R-squared	0.043	0.037	0.390	0.518

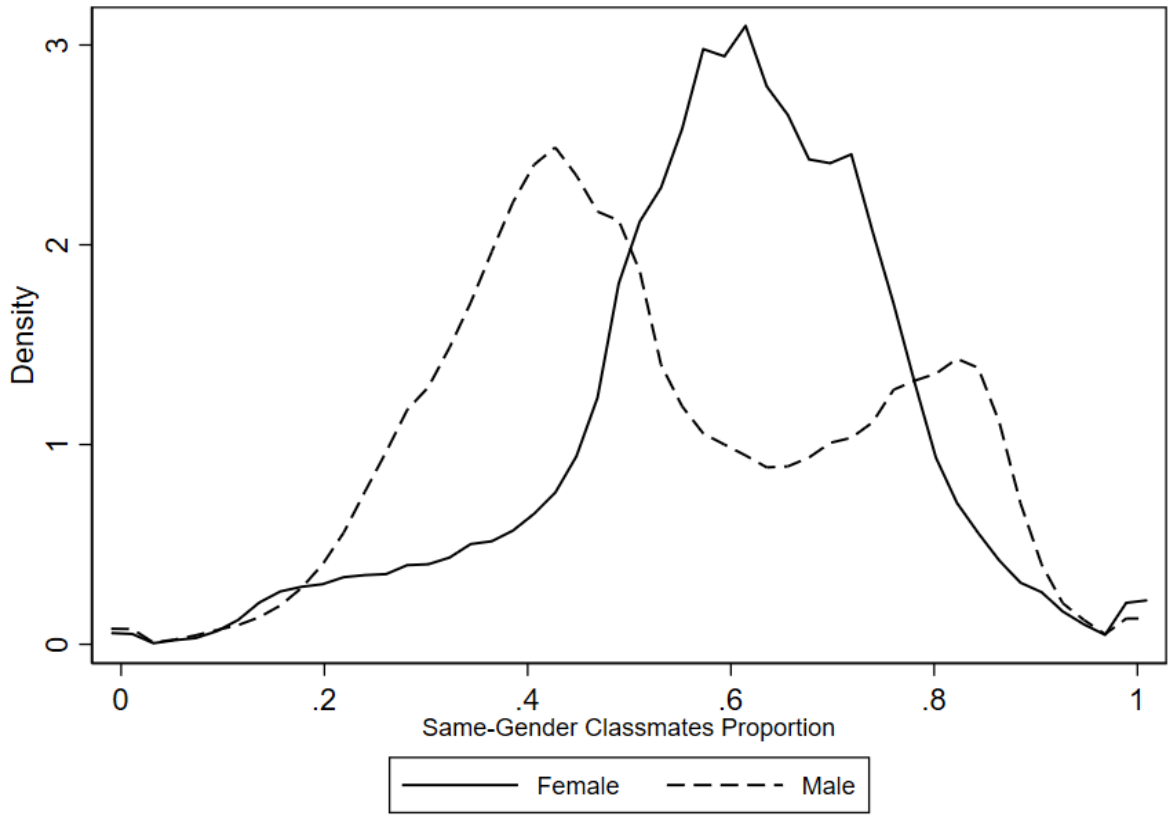
Note. This analysis uses student-by-term-by-course records; Standard errors in parentheses; * p<0.05 ** p<0.01 *** p<0.001

Table 1. 4 Results of Major Fixed-Effect Models on the Proportion of Students' Same-Gender Classmates for In-Department Courses and Out-Of-Department Courses

	(1)	(2)	(3)	(4)
	In-major courses		Out-of-major courses	
	Major FE		Major FE	
	Female	Male	Female	Male
Ethnicity (reference group: White)				
Asian	0.000	-0.004*	-0.007**	-0.000
	(0.001)	(0.002)	(0.002)	(0.003)
Latino/a	0.002	-0.005*	0.008*	-0.007**
	(0.001)	(0.002)	(0.003)	(0.002)
Black	0.000	-0.004*	0.003	-0.007
	(0.004)	(0.002)	(0.003)	(0.003)
Others	-0.004	-0.005	0.007	-0.005
	(0.004)	(0.006)	(0.007)	(0.006)
International student	-0.002	0.001	-0.022**	0.011
	(0.001)	(0.001)	(0.005)	(0.006)
Age (enrolled)	-0.000	0.000	-0.000	0.001
	(0.000)	(0.000)	(0.000)	(0.001)
First-generation status	-0.000	0.001	0.002	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)
Low-income status	-0.001	-0.001	0.001	-0.003
	(0.001)	(0.001)	(0.002)	(0.002)
Transfer students	-0.012***	0.012**	-0.006	0.020*
	(0.002)	(0.003)	(0.006)	(0.007)
Standardized pre-college performance	-0.002*	0.001*	-0.002	0.005*
	(0.001)	(0.000)	(0.001)	(0.002)
Control cohort	+	+	+	+
Control year of study (1st year, 2nd year, 3rd year, 4th year, 5th year in college)	+	+	+	+
Control major			+	+
Observations	282432	258385	445350	361860
Adjusted R-squared	0.724	0.835	0.200	0.292

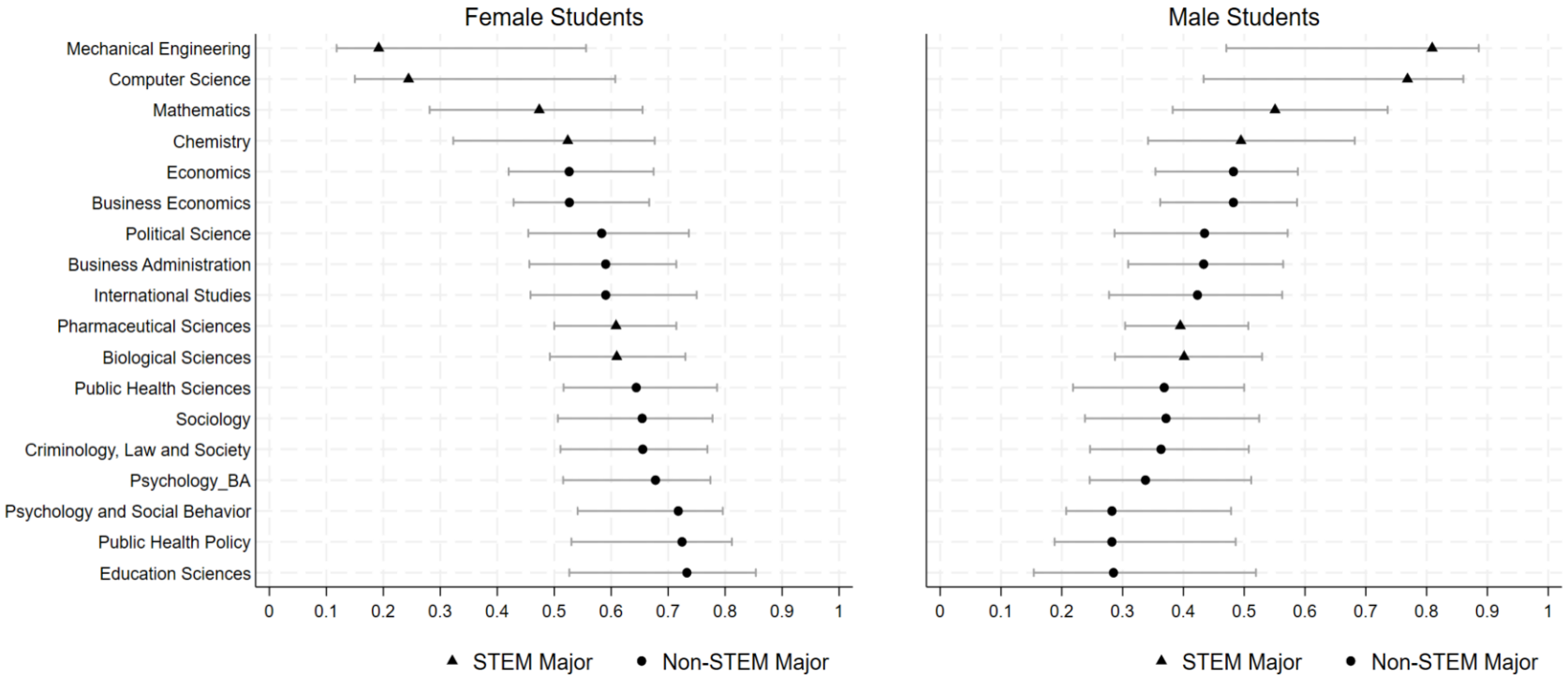
Note. This analysis uses student-by-term-by-course records; Standard errors in parentheses; * p<0.05 ** p<0.01 *** p<0.001

Figure 1. 1 Kernel Density Distribution of the Proportion of Same-Gender Classmates for Female and Male Students



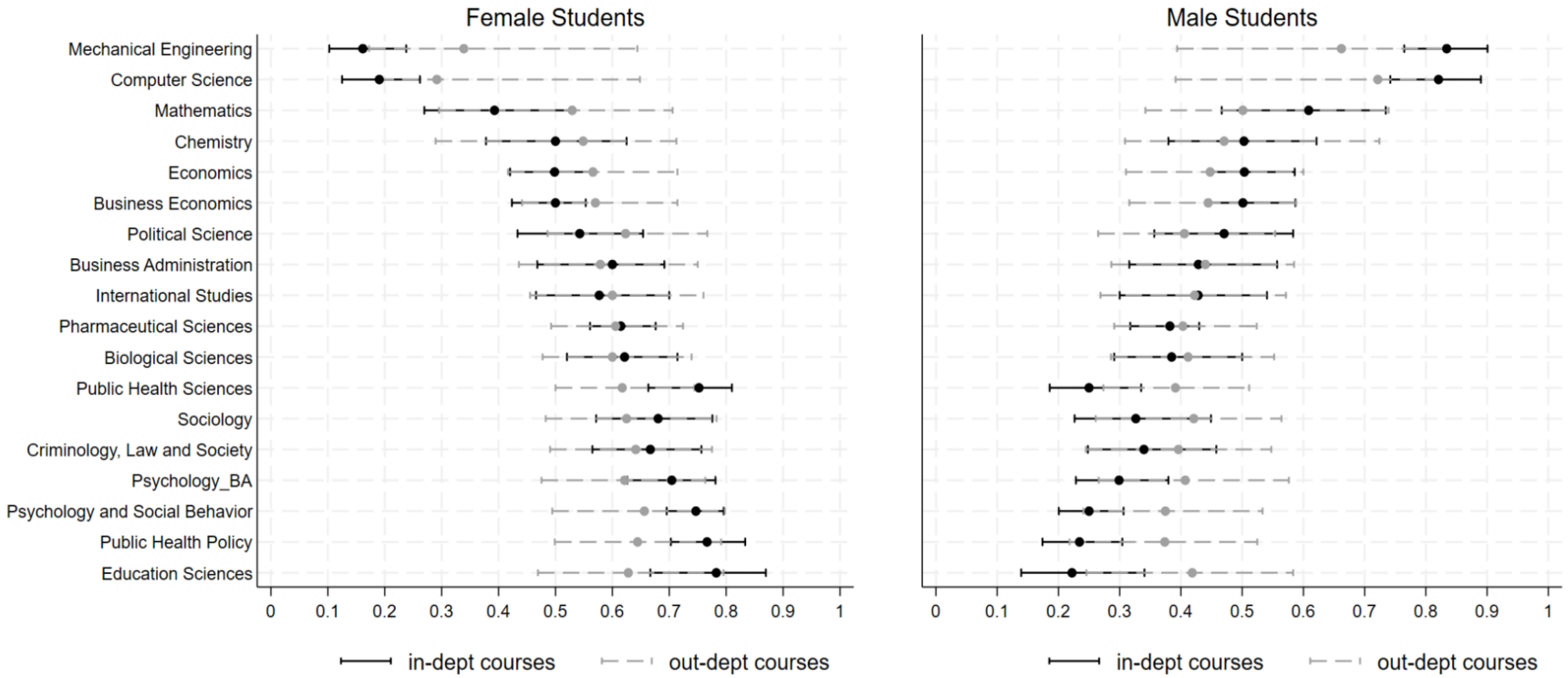
Note. This analysis uses student-by-term-by-course records.

Figure 1. 2 10th, 50th and 90th Percentiles of Proportion of Same-Gender Classmates Across All Courses, by Major



Note. This analysis uses student-by-term-by-course records. The left side of the bar represents the 10th percentiles, while the right side represents the 90th percentile. Triangle markers indicate the median for STEM majors, while circle markers indicate the median for non-STEM majors.

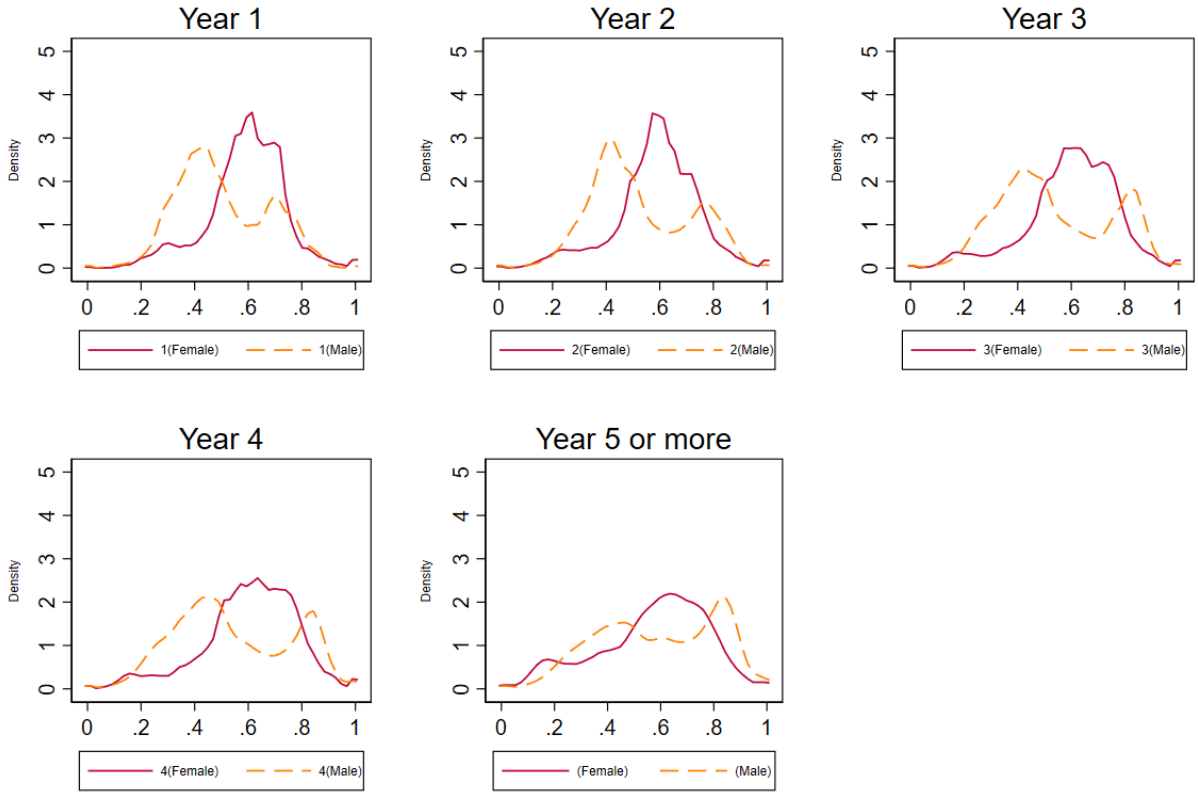
Figure 1. 3 10th, 50th and 90th Percentiles of Proportion of Same-Gender Classmates Across In-Department and Out-of-Department Courses , by Major



Note. This analysis uses student-by-term-by-course records. The left side of the bar represents the 10th percentiles, while the right side represents the 90th percentile. Markers indicate the median. Dash lines indicate out-of-major courses and solid lines indicate in-major courses.

APPENDICES

Appendix A1 Kernel Density Distribution of Same-Gender Classmates' Proportion by Year of Study (1st Year, 2nd Year, 3rd Year, 4th Year, 5th Year Or More In College)



Same-Gender Classmates Proportion Across Year of Study

Note. This figure uses student-by-term-by-course level records. For transfer students, year 1 is recoded as year 3, year 2 is recoded as year 4, and year 3 is recoded as year 5. This adjustment takes their learning experience before transferring into account.

Appendix A2 Statistics of Students' Same-Gender Classmates Proportion across Different Types of Courses

Major	All				In-department courses				Out-of-department courses			
Female students	10th percentile	Median	90th percentile	Difference between 10th and 90th percentiles	10th percentile	Median	90th percentile	Difference between 10th and 90th percentiles	10th percentile	Median	90th percentile	Difference between 10th and 90th percentiles
Mechanical Engineering	0.12	0.19	0.56	0.44	0.10	0.16	0.24	0.14	0.17	0.34	0.64	0.47
Computer Science	0.15	0.24	0.61	0.46	0.13	0.19	0.26	0.14	0.19	0.29	0.65	0.46
Mathematics	0.28	0.47	0.66	0.37	0.27	0.39	0.53	0.26	0.30	0.53	0.71	0.41
Chemistry	0.32	0.52	0.68	0.35	0.38	0.50	0.63	0.25	0.29	0.55	0.71	0.42
Economics	0.42	0.53	0.67	0.25	0.42	0.50	0.56	0.14	0.42	0.57	0.71	0.30
Business Economics	0.43	0.53	0.67	0.24	0.42	0.50	0.55	0.13	0.44	0.57	0.71	0.27
Political Science	0.45	0.58	0.74	0.28	0.43	0.54	0.65	0.22	0.49	0.62	0.77	0.28
Business Administration	0.46	0.59	0.71	0.26	0.47	0.60	0.69	0.22	0.44	0.58	0.75	0.31
International Studies	0.46	0.59	0.75	0.29	0.47	0.58	0.70	0.23	0.46	0.60	0.76	0.30
Pharmaceutical Sciences	0.50	0.61	0.71	0.21	0.56	0.61	0.68	0.12	0.49	0.61	0.72	0.23
Biological Sciences	0.49	0.61	0.73	0.24	0.52	0.62	0.71	0.19	0.48	0.60	0.74	0.26
Public Health Sciences	0.52	0.64	0.79	0.27	0.66	0.75	0.81	0.15	0.50	0.62	0.74	0.24
Sociology	0.51	0.65	0.78	0.27	0.57	0.68	0.78	0.20	0.48	0.63	0.78	0.30
Criminology, Law and Society	0.51	0.66	0.77	0.26	0.56	0.67	0.76	0.19	0.49	0.64	0.78	0.28
Psychology_BA	0.52	0.68	0.77	0.26	0.63	0.70	0.78	0.15	0.48	0.62	0.76	0.29
Psychology and Social Behavior	0.54	0.72	0.80	0.25	0.70	0.75	0.80	0.10	0.49	0.66	0.80	0.30
Public Health Policy	0.53	0.72	0.81	0.28	0.70	0.77	0.83	0.13	0.50	0.64	0.79	0.29
Education Sciences	0.53	0.73	0.85	0.33	0.67	0.78	0.87	0.20	0.47	0.63	0.80	0.33
Male students												
Mechanical Engineering	0.47	0.81	0.89	0.42	0.76	0.83	0.90	0.14	0.39	0.66	0.83	0.44
Computer Science	0.43	0.77	0.86	0.43	0.74	0.82	0.89	0.15	0.39	0.72	0.82	0.43
Mathematics	0.38	0.55	0.74	0.35	0.47	0.61	0.73	0.27	0.34	0.50	0.74	0.40
Chemistry	0.34	0.49	0.68	0.34	0.38	0.50	0.62	0.24	0.31	0.47	0.72	0.42
Economics	0.35	0.48	0.59	0.23	0.45	0.50	0.59	0.14	0.31	0.45	0.60	0.29
Business Economics	0.36	0.48	0.59	0.23	0.45	0.50	0.59	0.14	0.32	0.44	0.59	0.27
Political Science	0.29	0.43	0.57	0.28	0.36	0.47	0.58	0.23	0.26	0.41	0.55	0.29
Business Administration	0.31	0.43	0.56	0.25	0.32	0.43	0.56	0.24	0.29	0.44	0.58	0.30
International Studies	0.28	0.42	0.56	0.28	0.30	0.43	0.54	0.24	0.27	0.42	0.57	0.30
Pharmaceutical Sciences	0.30	0.39	0.51	0.20	0.32	0.38	0.43	0.11	0.29	0.40	0.52	0.23

Biological Sciences	0.29	0.40	0.53	0.24	0.29	0.38	0.50	0.21	0.29	0.41	0.55	0.27
Public Health Sciences	0.22	0.37	0.50	0.28	0.19	0.25	0.34	0.15	0.27	0.39	0.51	0.24
Sociology	0.24	0.37	0.52	0.29	0.23	0.33	0.45	0.22	0.26	0.42	0.56	0.30
Criminology, Law and Society	0.25	0.36	0.51	0.26	0.25	0.34	0.46	0.21	0.24	0.40	0.55	0.30
Psychology_BA	0.25	0.34	0.51	0.27	0.23	0.30	0.38	0.15	0.27	0.41	0.58	0.31
Psychology and Social Behavior	0.21	0.28	0.48	0.27	0.20	0.25	0.31	0.11	0.24	0.37	0.53	0.29
Public Health Policy	0.19	0.28	0.49	0.30	0.17	0.23	0.30	0.13	0.22	0.37	0.52	0.31
Education Sciences	0.15	0.29	0.52	0.37	0.14	0.22	0.34	0.20	0.25	0.42	0.58	0.34

Note. This table uses student-by-term-by-course data.

Appendix A3 Supplemental Analyses of Manually Coded Major-Required Courses for the 18 Largest Majors

Comparison of the main analysis sample and the subsample

In the main analysis, we used in-department courses as a proxy for major-required courses. However, this approach might not comprehensively represent major-required courses since many majors require some courses from outside their home departments. To address this concern, in this appendix we present supplemental analyses that use a more accurate measure of major requirements. We accomplish this by manually coding the course requirements of the 18 largest majors (including major, specialization, and school requirements).

In the main analysis, our sample includes all undergraduate students across all majors (100 majors in total) from the 2011 cohort through the 2015 cohort. This includes students pursuing multiple majors and undeclared students. In this supplemental analysis we use a subsample extracted from the main sample. This subsample is restricted to students who declared only one of the 18 largest majors and who never switched their major.

The rationale for focusing on the 18 largest majors arises from the complexity and labor-intensive process of manually coding the curricular requirements for each major across multiple years. For this same reason, we limited our sample to students who declared only one major and never changed their majors, as incorporating students who have declared multiple majors would demand additional coding beyond these 18 majors. By focusing on "never switchers" we avoid capturing partial academic records of students who might have switched out of these 18 majors, ensuring a comprehensive data representation.

This subsample includes 10,280 students, representing 30% of the unique students from the 2011-2015 cohorts and 70% of those students who never switch majors. Compared with the main analysis sample, this subsample has a higher proportion of transfer students (45% vs. 26%), Asian students (59% vs. 53%), international students (23% vs. 20%), and students with higher pre-college academic scores (227 vs. 224). Conversely, it has a lower proportion of Latino/a students (20% vs. 26%) and first-generation college students (51% vs. 53%).

As the main goal of this appendix is to examine if and how our main findings change with a more accurate coding of the courses required for majors, in this appendix we present results from re-estimations of our main models that restrict the sample to "never switchers" who enrolled in only one of the 18 largest majors. We do this to ensure that the main analysis and the subsample analysis are directly comparable and that the observed differences primarily stem from how we categorize in- and out-of-major courses, and not from sample restrictions. Subsequent comparisons in this appendix section will utilize these two samples.

Enhancing the accuracy of major-required courses identification

In the main analysis, courses offered by a student's major home department were used as an approximation to capture major-required courses. This approach, while inclusive, might not precisely depict the courses required by the declared majors, as most majors require some courses from other departments.

This supplemental analysis uses a unique data set including hand-coded major requirements to identify the specific classes that fulfill each major's requirements, whether as a general requirement for the degree, a specialization within the major, or a school-wide requirement (e.g., the History major within the School of Humanities). Major-required courses can either be classes required for all students within the major or required elective classes (e.g., choose one out of these three classes). Conversely, non-major-required courses fall outside the specific requirements of a student's major, including general education and other elective courses. It's worth mentioning that some courses offered by the home department may also be considered non-major-required courses, as they are designed for general education or students from other departments. To facilitate this, our research team constructed these requirements for each year of the sample using the general catalog for the 18 sampled majors. Notably, these course requirements change often in our sample time period.

To discern any differences in the coding of in-major and out-of-major courses between this method and the one employed in the main analyses, we compare the proportion of in-department courses students enrolled in (our main approach) with the percentage of courses that could satisfy major requirements students enrolled in (our supplemental approach) across the 18 majors (see Appendix Table A3.1). This comparison reveals that the broader metric (contrasting in- and out-of-department courses) on average identifies 18% fewer courses as major-required. The nontrivial difference motivates these supplemental analyses. Specifically, we aim to determine if there is still substantial student sorting within major, using a more accurate definition of major-required courses. If present, we also seek to examine the extent to which this variation changes based on the different definitions of major-required courses.

Table A3.1 Average Percentage of Major-Required and In-Department Courses Taken by Students per Term, and the Proportion of Female and Male Students across 18 Majors

Major	Proportion of major-required courses	Proportion of in-department courses	Proportion of female students	Proportion of male students
Mechanical Engineering	0.82	0.58	0.12	0.88
Computer Science	0.77	0.36	0.17	0.83
Mathematics	0.64	0.52	0.43	0.57
Chemistry	0.64	0.52	0.44	0.56
Economics	0.45	0.42	0.47	0.53
Business Economics	0.54	0.44	0.54	0.46
Political Science	0.54	0.42	0.52	0.48
Business Administration	0.58	0.57	0.60	0.40
International Studies	0.58	0.29	0.63	0.37
Pharmaceutical Sciences	0.69	0.27	0.64	0.36
Biological Sciences	0.64	0.47	0.62	0.38
Public Health Sciences	0.71	0.22	0.73	0.27
Sociology	0.56	0.47	0.70	0.30
Criminology, Law and Society	0.53	0.49	0.59	0.41
Psychology_BA	0.56	0.54	0.66	0.34
Psychology and Social Behavior	0.57	0.56	0.74	0.26
Public Health Policy	0.62	0.45	0.80	0.20
Education Sciences	0.45	0.48	0.90	0.10
Total	0.63	0.45	0.55	0.45

Note. Take the first row as example, it means that in Mechanical Engineering, students on average have 58% of in-department courses, and 82% major-required courses per term,12% of students in that major are female; the data used to generate the last two columns comes from a student-by-major data pool that spans all cohorts. "Major-required courses" refer to courses that fulfill at least one of the three requirements: major, specialization, and school. "In-department courses" refer to courses that are from the main home department. The data is restricted to students who never switched majors and only enrolled in one of the 18 largest majors.

Differences in results: in- & out-of-department courses vs. in- & out-of-major courses

Because of the focus of this appendix – refining the identification of in- and out-of-major courses–the analyses presented in this appendix focus on differences in peer compositions in major-required and non-major-required courses rather than on the overall peer composition from all courses.

In-department courses vs. in-major courses

Similar to the findings from our main analyses, we find notable variation in the same-gender peer composition in major-required courses among students within the same major. Specifically, when we adopt more accurate metrics for major required courses, the unexplained variation in the proportion of students' same-gender peers—after accounting for major FE—increases: from 31% to 39% for female students, and from 14% to 20% for male students. When we integrate more courses from external departments into the in-major category, the explanatory power of the major on students' peer composition diminishes. One potential reason for this is that this method is detecting increased course choices available to students from external departments.

The distribution of students' peer composition in major-required courses within majors widens. Appendix Figure A3.1 compares the 10th, median, and 90th percentiles of students' same-gender peer composition between in-department courses (represented by a solid line and dot) and in-major courses (depicted by a dashed line and hollow dot). By using a more accurate identification of major requirements, there's a noticeable increase in the range (between 1% and 10%) between the 10th and 90th percentiles for 16 out of 18 majors.

These findings confirm the results from the main analysis, which highlighted substantial variation in students' peer compositions even within in-department courses. In fact, these results indicate that the findings in the main analyses generally understate the amount of variation within majors. Moreover, these results indicate that in majors that are disproportionately represented by one gender—such as Mechanical Engineering and Public

Health Policy—including courses from other departments can shift students' same-gender peer compositions closer to the campus-wide gender composition.

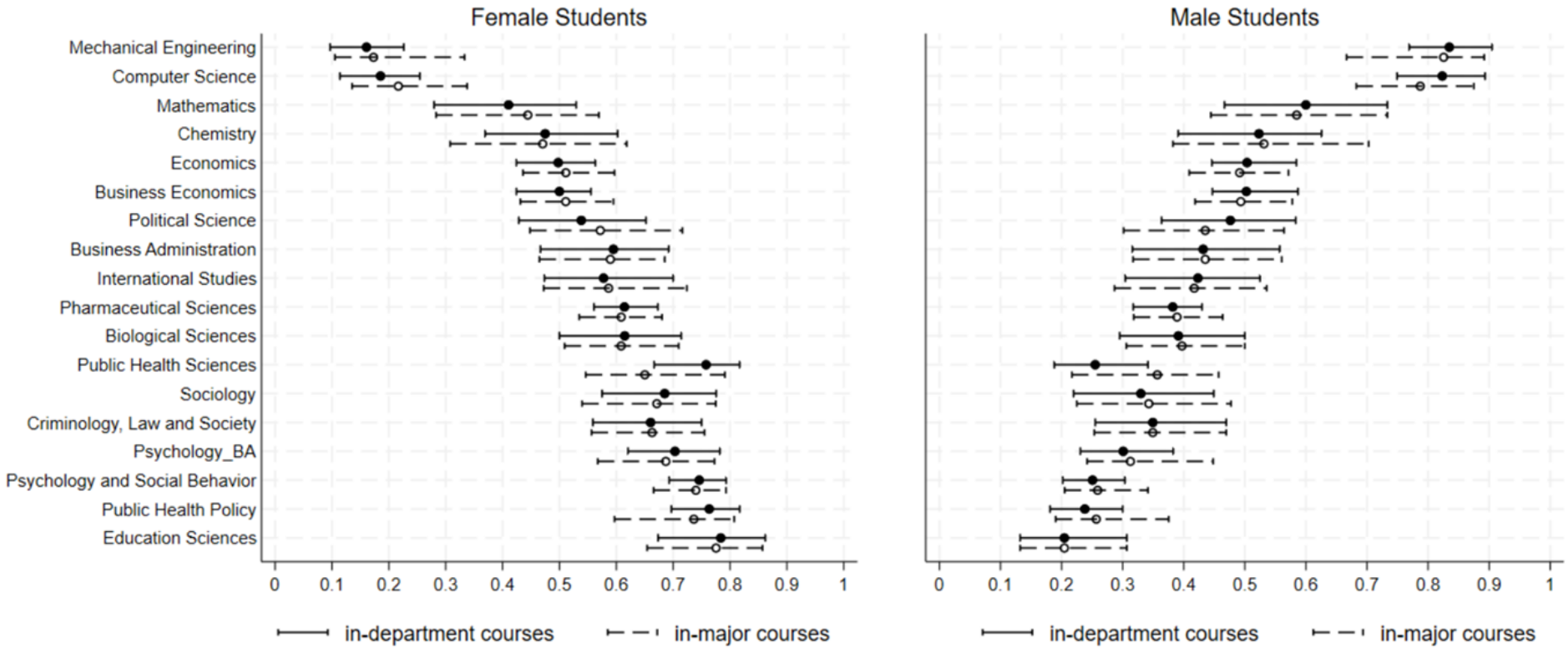
Out-of-department courses vs. out-of-major Courses

For out-of-major courses, the major fixed-effect models explain only a small proportion of the variation: 5% for female students and 8% for male students. This differs substantially from the main analyses, which explain 20% of the variance for female students and 29% for male students. The likely reason for this is that in the main analyses, major-required classes were mistakenly categorized as out-of-major courses. For example, the Biology Science major requires a sequence of courses in the Chemistry department, and the main analyses do not consider these courses to be major-required.

Appendix Figure A3.2 displays distribution bar graphs comparing out-of-department to out-of-major courses. The main finding is that using this more refined definition, the median proportion of same-gender peers is much more similar across majors. This trend suggests that students likely experience more similar peer compositions in non-major-required courses (e.g., general education, elective courses). Such courses can balance students' peer compositions to be more consistent across majors, particularly for majors that are disproportionately represented by one gender.

Appendix Figure A3.1

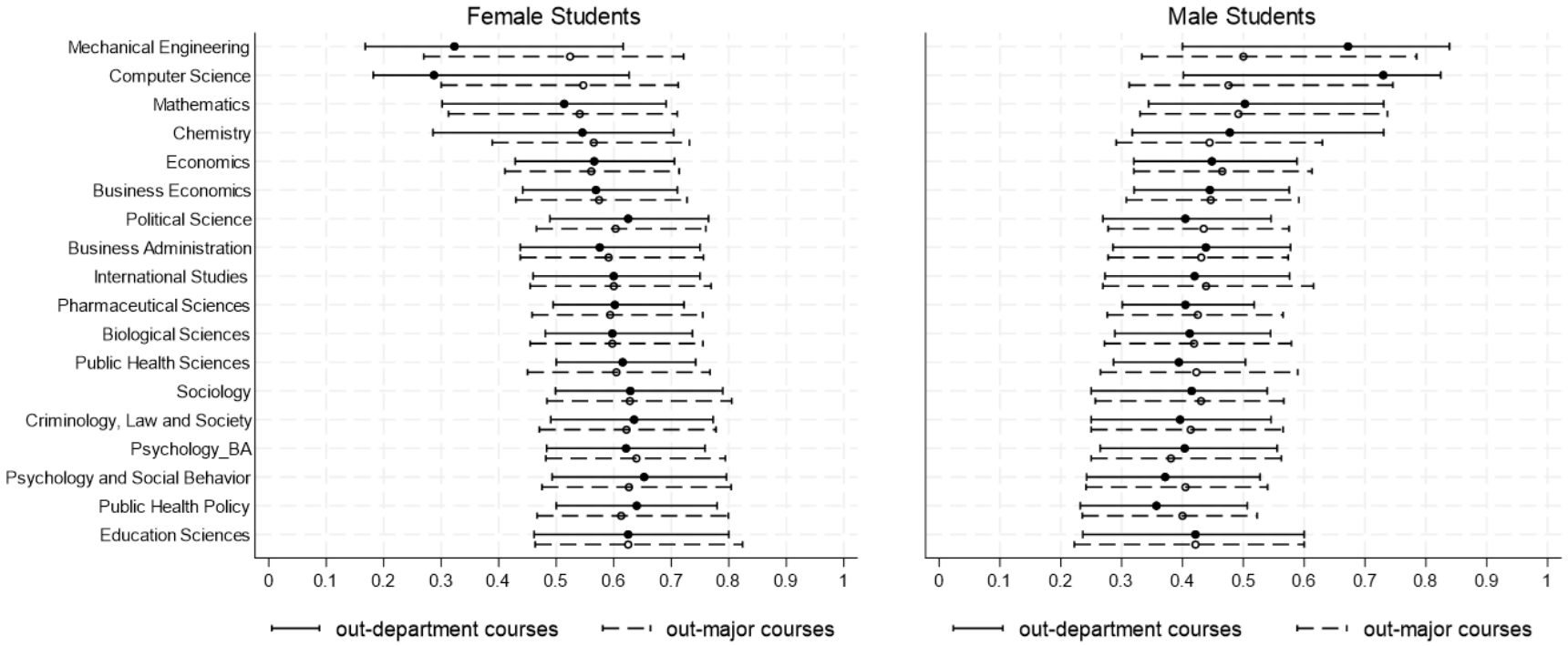
10th, 50th and 90th Percentiles of Proportion of Same-Gender Classmates Across In-Department and In-Major Courses, by Major



Note. This analysis uses student-by-term-by-course data. The left side of the bar represents the 10th percentiles, while the right side represents the 90th percentile. Makers indicate the median. For the in-department courses, the sample is restricted to "never-switchers" and students who only enrolled in one of the 18 largest majors to ensure comparability in the analysis.

Appendix Figure A3.2

10th, 50th and 90th Percentiles of Proportion of Same-Gender Classmates Across Out-of-Department and Out-of-Major Courses, by Major



Note. This analysis uses student-by-term-by-course data. The left side of the bar represents the 10th percentiles, while the right side represents the 90th percentile. Makers indicate the median. For the in-department courses, the sample is restricted to "never-switchers" and students who only enrolled in one of the 18 largest majors to ensure comparability in the analysis.

This supplemental analysis provides a more accurate insight into students' academic experiences within their declared majors. The findings presented in the primary analyses remain similar. When examining students' peer compositions within majors, these improved measurements reveal greater variations within in-major courses. Specifically, after controlling for major fixed effects, we observe an increase in unexplained variation in students' same-gender peer proportions within major-required courses. This reaffirms findings from the main analysis: there is considerable student sorting within majors. Despite being in the same major, students show substantial variation in their peer compositions, potentially due to diverse course choices within the major. For majors heavily skewed towards one gender, out-of-major courses can widen the distribution of students' peer compositions and serve as a counterbalancing factor in their peer composition.

CHAPTER 2 OPTIMIZING STUDENT ENGAGEMENT THROUGH DIVERSE LEADERSHIP TENDENCIES IN GROUP COMPOSITION

INTRODUCTION

The STEM-related industries is often hailed as the future of the economy due to the rapidly expanding influence and the lucrative career opportunities that they offer (Langdon et al., 2013; U.S. Bureau of Labor Statistics, 2023). Yet the STEM workforce faces significant challenges: the supply of skilled professionals is insufficient, and diversity within the STEM workforce is lacking (Miriti, 2020; NCSES, 2023). STEM education is vital for equipping students with the necessary skills to thrive in the STEM industry (Xie et al., 2015). Yet in higher education, the retention rate of students in STEM programs/courses is concerningly low, and there is a notable gap in the persistence of students from diverse backgrounds (Riegle-Crumb et al., 2019). Hispanic and Black students receive a lower percentage of STEM bachelor's degrees compared with the overall rate (U.S. Department of Education, 2016).

The expansion of online STEM courses is one promising response to these challenges, enhancing the accessibility of essential and high-demand introductory STEM courses. This is particularly beneficial for students from underserved communities who often have to manage multiple responsibilities while pursuing their education (Allen & Seaman, 2015; Chirikov et al. 2020; Xu & Xu, 2020). However, the flexibility of the online format also introduces its unique challenges. The limited interaction in online formats not only diminishes student engagement, which is crucial for navigating demanding STEM courses, but also weakens students' sense of belonging—a key factor affecting students' persistence in STEM fields (Baldwin 2009; Bambara et al. 2009; Peacock et al., 2020; Peacock & Cowan, 2019). These challenges are particularly pronounced among students from underserved communities, who experience lower levels of

sense of belonging and satisfaction as well as poorer academic performance in online STEM courses (Ke & Kwak, 2013; Rainey et al., 2018; Xu & Jaggars 2013).

In response to these challenges, the formation of study groups—small groups in which students regularly meet to study and discuss course contents together—have been recognized as a cost-effective and scalable solution to enhance students' sense of belonging, motivation, and academic performance (Arendale, 2014; Rasheed et al., 2021; Sheeran & Cummings, 2018; Thai et al., 2019; Zhou et al., 2023). The scalability and adaptability of study groups across various educational settings are notable, as they do not require specific content-related tasks and are student-oriented. However, despite their proven effectiveness, it is still a significant challenge to maintain student participation and engagement in study groups. Participation tends to wane as courses progress, and overall student engagement remains low (Hara et al., 2000). This situation underscores the need to develop low-cost, effective strategies that can be easily implemented to bolster student engagement in study groups. Successfully identifying and implementing such strategies could allow a wider array of educational settings to utilize study groups more efficiently, specifically addressing the unique challenges presented by the demanding nature of online STEM courses.

Adjusting the composition of study groups to account for student characteristics presents a low-cost and scalable intervention that can improve group dynamics and boost student participation, without burdening students with additional tasks. Prior studies have demonstrated the impact of group members’ characteristics on student outcomes. For example, groups with a higher proportion of female or high-achieving peers positively influence students' academic performance (Hoxby 2000; Zimmerman 2003). However, broadly implementing these strategies presents challenges, particularly in educational contexts most in need. This includes settings

where students are generally less academically prepared or where significant gender segregation exists in certain STEM fields. Additionally, the goal of study groups extends beyond academic performance to include social aspects, such as enhancing students' sense of belonging and engagement. These underscore the need to identify effective ways to construct study groups based on students' tendencies and preferences that can enhance student engagement and sense of belonging.

Leveraging students' relative motivation to lead teamwork offers a cost-effective and scalable strategy. This approach demands minimal additional effort from students and is suitable for a wide range of student demographics across various education settings. Prior literature suggests that peer leaders play an instrumental role in group dynamics and the engagement of group members. Their social skills and group knowledge enable them to facilitate group activities, communicate with members more efficiently, and build more positive group relationships (Arendale, 2004; Barry & Stewart, 1997; Hara et al., 2000; Johnson & Johnson, 2002). Empirical studies have provided evidence that social skills, even when controlling for cognitive ability, lower coordination costs and enhance both collaboration and group performance efficiency (Deming, 2017; Riedl et.al.,2021; Woolley et al., 2010).

This study explores an innovative approach to forming study groups, focusing on students' leadership motivation. This method builds on the proven impact of group composition based on student characteristics (e.g., gender, academic performance) on student outcomes, but focuses on the critical role of leadership and social skills. By intentionally allocating students with a strong inclination toward leading teamwork, we propose a method to enhance the efficiency of study groups. Such groups require shared regulation, where members collaboratively plan and monitor their activities, as evidenced by previous research (Grau &

Whitebread, 2012; Hadwin et al., 2017). Students with higher leadership motivation are more likely to engage actively in these regulatory processes, fostering increased participation and engagement across the group. Our findings present a practical, cost-effective strategy for organizing study groups that bolster student engagement and foster a sense of belonging, with potential significant benefits for underrepresented and minoritized students in various educational contexts.

LITERATURE REVIEW

Sense of Belonging in Online STEM Education

In the realm of online STEM education, a sense of belonging significantly enhances student engagement by affirming their place and active contribution within the educational community. A sense of belonging is integral to student learning due to its correlation with academic self-efficacy, intrinsic motivation, and task value—key factors that enhance learning outcomes (Freeman et al., 2007; Strayhorn, 2018). It is especially vital for underrepresented minority (URM) students, who often experience feelings of isolation and alienation in predominantly white, competitive STEM environments. Such feelings are frequently aggravated by stereotypes and implicit biases that undermine their competence and their right to belong in the field, ultimately influencing STEM retention (Hurtado & Carter, 1997).

Central to fostering this sense of belonging are peer relationships and interactions. One of the major challenges in online STEM education is the lack of peer interactions, which are essential for building a supportive and inclusive educational environment (Lai et al., 2019; Rainey et al., 2018). Addressing this challenge is pivotal for enhancing the educational experience and improving retention rates in online STEM programs.

Benefits And Challenges of Small Student-Led Study Groups

The existing literature extensively discusses various strategies aimed at enhancing peer interaction and engagement among students. Notably, many of these methods incorporate additional tasks or directives, including group projects and peer comments on classmates' work. However, these more structured strategies may prove particularly challenging within the rigorous and competitive environments of STEM courses, potentially exacerbating the difficulties some students face.

Alternatively, a more organic approach to fostering interaction involves building a community where students can openly share concerns and pose questions. This method aims to facilitate interactions in a natural and low-risk manner, thereby potentially increasing the comfort and engagement of students in academic discussions. Study groups are one way to build such a low-risk community. Previous studies have demonstrated the effectiveness of study groups on student outcomes, including learning and psychological outcomes such as interpersonal relationships (Arendale, 2014; Rasheed et al., 2021; Sheeran & Cummings, 2018; Thai et al., 2019; Zhou et al., 2023). In a meta-analysis, Springer et al. (1999) concluded that small-group learning effectively enhances academic achievement, cultivates positive attitudes toward learning, and fosters increased persistence in undergraduate STEM courses.

However, it is less clear how to build an efficient study group to promote students' participation and engagement. Students who may benefit the most from participating in study groups may be less likely to join. Arendale (2014) suggested that students with low academic preparedness are less likely to participate in group-based supplemental instruction. Zhou et al. (2023) found that students who exhibit lower academic preparation and baseline motivation show enhanced academic performance after the study group intervention. This emphasizes the

necessity of devising efficient strategies that can be readily adopted to enhance student engagement in study groups. Successfully identifying and implementing such strategies could enable a broader range of educational environments to utilize study groups more effectively, particularly in addressing the distinct challenges posed by the rigorous nature of online STEM courses.

Group Composition And Social Skills

One low-cost strategy for constructing study groups without adding additional burden on students is to modify the composition of study groups based on student characteristics. Peer composition, as evidenced by peer effect literature, can significantly influence group dynamics and may alter student participation behaviors. Research underscores the pivotal role of group member characteristics in influencing student outcomes. Rigorous quasi-experimental studies suggest that groups with a higher proportion of female or high-achieving peers tend to positively impact students' academic performance (Hoxby, 2000; Zimmerman, 2003).

However, implementing such strategies on a broad scale faces hurdles, especially in educational contexts where students may be less academically prepared or where gender segregation is prevalent in certain STEM fields. Furthermore, while study groups primarily aim to improve academic performance, they also fulfill vital social functions by fostering a sense of belonging and engagement among students. Thus, it is imperative to identify effective methods to construct study groups tailored to students' preferences and tendencies to enhance both academic outcomes and social cohesion.

Existing literature highlights the crucial role of peer leaders in shaping group dynamics and fostering member engagement. Peer leaders possess social skills and group knowledge that

enable them to facilitate group activities, communicate effectively with group members, and cultivate positive group relationships (Arendale, 2004; Barry & Stewart, 1997; Hara et al., 2000; Johnson & Johnson, 2002). Empirical studies further demonstrate that social skills, irrespective of cognitive ability, reduce coordination costs and enhance collaboration and group performance efficiency (Deming, 2017; Riedl et al., 2021; Woolley et al., 2010).

The specific tasks for study groups they involve scheduling and monitoring meetings with peers to provide each other an opportunity to interact. The critical role is to coordinate activities. Leveraging students' motivation to lead teamwork can be directly relevant to social skills while being helpful in study group activities. Motivation to lead is a measurement that considers multiple facets of leadership performance and individual preferences and tendencies. These items assess their leadership preferences, interest in leading, perceived contributions as a follower versus a leader, and tendencies to take charge in group settings (Chan & Drasgow, 2001).

CURRENT STUDY

Experiment Context

This intervention was conducted during the winter quarter of 2023 in an online introductory STEM course. The course is offered at a selective public research university in the American West. The university is a Hispanic-Serving Institution and demonstrates a strong academic student body, with a median high school GPA of 3.95 among first-year students.

This introductory STEM course is a four-credit prerequisite for certain STEM majors and has been in existence for more than ten years. It has its own prerequisites, with students needing to fulfill one of the criteria measured by AP STEM courses, SAT math scores, or math course

experience. It is prerequisite for several STEM majors, including Biological Science, Chemistry, etc. The course is offered in both the fall quarter and the winter quarter. According to the instructor, students who attend the course in the winter quarter are, on average, less academically prepared; they may have failed the class in the fall quarter or delayed taking it for various reasons. Multiple sections taught by different instructors are offered during both the fall and winter quarters, while only this course section offered in the winter quarter is offered online. The course is challenging; previous data revealed that 30% of students in the course received a letter grade of C or lower. Regarding teaching design, the course is offered online, with two synchronous lecture session taught by the instructor and a discussion session led by the TAs each week. It entails a relatively heavy workload, including three midterms evenly distributed across the ten-weeks long, six quizzes, a final exam, and video assignments. No group work is involved except for the study group intervention.

Experiment Design

We conducted a two-arm randomized controlled trial to investigate how modifying the composition of group members' motivation to lead affects students' study group engagement and satisfaction, social and learning experience in the course. The intervention involved weekly study group assignments, which constituted 5% of the course credits. This information is listed in the course syllabus that is accessible to students just before the class started. Additionally, each student were invited to complete two surveys (i.e., pre- and post-course) sent out from the instructor, each offering an extra 5% credit toward the final exam grade (refer to Appendix B1 and B2, for more survey details).

For the weekly study group assignments, each student in the class was allocated to a study group (three students per group) in week 2 and was required to dedicate at least one

uninterrupted hour every week to study with their study group peers on any topic related to the course, either virtually or in person. Students also needed to complete a brief survey on Qualtrics[5], linked through the learning management system Canvas[6], about the group study session on a weekly basis (see Appendix B3). The brief survey is low-risk; students were asked questions such as if they met with their study group peers last week, what they did during the meeting, when they plan to meet next week, and to upload screenshots or relevant proof showing they met last week. Additionally, students were provided with a group-specific Discord[7] link that they could use to communicate with their own study group members and a Zoom link that they could use for their weekly study group meeting. Study group assignments are fixed for each student starting from week 2, except in rare instances where students feel compelled to switch groups due to non-responsive or non-communicative group members.

The participants of the intervention included all students aged 18 years or older enrolled in this online introductory STEM course in the Winter 2023 quarter who completed the pre-course survey by the end of the first week. Starting in week 2, participants were assigned to either the treatment or control conditions via stratified randomization, considering gender, underrepresented minority status, high school GPA, and the number of classmates known prior to the course. This method of stratification was employed to ensure a balanced distribution of student characteristics potentially relevant to the study outcomes. The distinction between the treatment and control conditions was the composition of their study group peers:

[5] Qualtrics is a widely utilized online survey platform that enables users to design, distribute, and analyze surveys and research instruments.
[6] Canvas is a popular learning management system (LMS) utilized by educational institutions worldwide, providing a platform for course management, online learning, and communication between instructors and students.
[7] Discord is an online communication platform frequently used by US college students, it can be replaced with other frequently used platforms.

Treatment condition (Diversified Peers): Students were classified into high, medium, or low categories for motivation to lead based on their relative scores, which were calculated through responses to Motivation to Lead items (see Appendix B1 Q18) in the pre-course survey. These items assess their leadership preferences, interest in leading, perceived contributions as a follower versus a leader, and tendencies to take charge in group settings (Chan & Drasgow, 2001). Subsequently, a student from each category (high, medium, low) was randomly selected to form study groups of three, ensuring that each group consisted of members with varied levels of leadership motivation. A comparison of the group average motivation to lead scores between the treatment and control conditions is displayed in Appendix Figure B1, showing that the distribution for the treatment conditions is less spread than the distribution for the control condition.

Control condition (Randomly Assigned Peers): Students were randomly assigned to study groups of three. Consequently, the composition of each group in the control condition could result from any of the ten possible combinations[8], such as low-low-high or medium-high-high.

Outcome Measures

Behavioral outcomes: participation and engagement in study group activities

To continuously monitor student engagement within study groups, we administered weekly brief surveys (refer to Appendix B3 for an example). These surveys assessed various aspects of participation, including the total hours students spent with study group peers and the frequency of discussions about course-related and non-course-related content. For students who

[8] The detailed group composition distribution for the control and treatment group at the end of the course is shown in the Appendix Table B1.

did not submit their surveys, we code them as "0" to these items.[9] Additionally, we measured study group satisfaction on a weekly basis for most of the weeks. These outcomes are measured by a five-point Likert scale.

Student interaction within study groups: Conversational data through the communication platform

At the beginning of the study group intervention, we provided students with a link to the online communication platform, Discord, to facilitate discussions among group members. Following the intervention, we collected data from each study group's conversations on Discord and employed text analysis techniques to examine the extent to which the conversations across different treatment conditions exhibited specific sentiments and themes. Each message was treated as a separate observation. An example of the Discord data collected is a message like, "Hi, should we find a time to meet?" Additional examples provided in the Appendix B4.

To gain a deeper understanding of the students' interactions within the study groups, we examined the extent to which the conversations reveal different sentiments and themes. We utilized the Linguistic Inquiry and Word Count (LIWC)—a software with a built-in dictionary—to focus on two content-relevant measures and one structural feature among the summary variables provided by LIWC[10]: Leadership, Emotional Tone, and Words Per Sentence (Cohn et al., 2004; Kacewicz et al., 2014; Monsjou et al., 2022). LIWC outputs the percentage indicating to what extend the text message shows the content of the intent measure. Take leadership for instance, 50% shows higher level of leadership-relevant content in the text message than 20%. In the official model analysis, each of the three outcome is standardized. The theme of leadership is

[9] I also tried coding it as missing, and the results are qualitatively similar.
[10] I focused on the summary measures that are relevant to our intervention, but I presented all the summary measures from LIWC to avoid the concerns of fishing.

directly relevant to our experimental design of modified group composition. We hypothesize that treatment conditions—study groups with high, medium, and low motivation to lead peer composition—will exhibit a higher level of leadership in group conversations compared with groups composed of randomly assigned peers. In addition, the high motivation to lead peer in each group is expected to initiate more engaging group dynamics, contributing to group satisfaction and positive interactions.

Study-group perceptions

In the post-course survey, we collected items regarding students' experience and perception regarding study groups (seen in the Appendix B2). These items learn from existing survey items from three aspects: involvement in the group activities, group process, and relationships among group members (Arbaugh et al., 2007; Järvenoja et. al., 2012; Jung & Sosik, 2002). Involvement included questions aimed at assessing the individual contributions of participants within study group settings. Specifically, the items sought to evaluate the extent to which members coordinated activities, facilitated engaging and productive discussions, and maintained focus on assigned tasks. Group process assessed aspects of group dynamics within study groups, focusing on their understanding of group purpose, effectiveness in scheduling activities, productivity in meetings, and flexibility to adapt to changing needs. These items aimed to measure the functional attributes that contribute to the success of group collaboration. Group relationships evaluated the interpersonal dynamics within study groups, examining aspects such as connectivity, listening skills, conflict resolution, cooperation, and mutual respect among members. These items were designed to measure the quality of relationships and collaborative efficacy within the groups.

Course social and learning experience, and academic outcomes

In addition to study group outcomes, we also gathered data on students' social and learning experiences related to the overall course. Social experiences were quantified through variables such as the number of friends students reported, the number of peers from whom they could seek help, and their perceived sense of belonging (Xu et al., 2018). For learning experiences, we assessed students' course motivation (Eccles & Wigfield, 2002), time management skills (Pintrich & De Groot, 1990), and overall satisfaction with the course. Regarding academic performance, we collected students' average quiz grades, average midterm grades, raw final exam scores, and final course grade.

Analytic Sample

In the initial assignment, both the treatment and control conditions each comprised 171 students, distributed across 57 study groups (three students per group), totaling 342 students and 114 study groups. We excluded two students under the age of 18 from our analysis, reducing the total number of students considered from the initial allocation to 340. Beyond the 340 students, a total of 45 students either did not participate in the pre-course survey or enrolled in the class after the survey period had concluded. These students were grouped together to form their own study groups, which are not included in the analytic sample and excluded from the subsequent analysis, thereby averting potential interference with the participant sample.

Attrition can be defined in two ways in this intervention involving the initial 340 students who completed the pre-course survey and were under 18 years old. The first definition pertains to students who drop out of the course, with an attrition rate of 8%. Since the course operates on a quarter system, treatment status and group assignments are established by week 2, coinciding with the period when students are actively adjusting their schedules by adding or dropping

classes. Observing a higher dropout rate in either treatment condition raises concerns that the treatment may cause unintended effects on student attrition, potentially leading to non-comparable characteristics between the groups. The second definition applies to students who did not complete the post-course survey, including those who dropped out of the course, with an attrition rate of 15%. Since some outcome measures are collected from this survey, a higher non-completion rate in either group could compromise the causal interpretation of the treatment's impact on outcomes.

Table B2 in the Appendix presents the attrition rates of the sample categorized by two criteria, with and without the inclusion of additional student controls. Columns (1) and (3) display the relationship between treatment condition and two types of attrition outcomes—students dropping out of the course and students not completing the post-course survey, respectively—without controlling for other variables. The results indicate no statistically significant relationship between treatment conditions and either type of attrition. Columns (2) and (4) incorporate student controls into the model for both attrition outcomes, and the results continue to show no significant relationship between the treatment conditions and attrition rates. However, certain student characteristics, such as the number of friends known in the class and course motivation, are significantly related to attrition. We further conducted an analysis using pre-course survey characteristics of students to predict their treatment status, with results presented in Appendix Table B3. The joint F-test results were not significant, indicating that, collectively, the pre-course student characteristics are not related to students' treatment status. This finding supports the validity of the random assignment.

The final analytic sample include 288 students. Student descriptive statistics for both the treatment and control conditions, along with the differences between them, are detailed in Table

2.1. The sample consisted of 68% female students and 43% underrepresented minorities. No statistically significant differences were observed between the treatment and control conditions regarding the examined student characteristics.

Methods

To assess the impact of eligibility for the treatment condition on various student outcomes, we employed a linear regression model as described in equation [1]. In this model, Y_i represents the various outcomes for student i, as mentioned above. The coefficients β represents our primary estimates of the impact of treatment conditions on student outcomes. We controlled for covariates collected before the course in X_i, including gender, race/ethnicity, high school GPA, and composite scores for course motivation, peer interaction preference, motivation to lead, general stress level, friends known in the course before, and students that they can seek help with in the course. ε_i denotes the error term.

$$Y_i = \alpha + \beta Treat + X_i + \varepsilon_i \quad [1]$$

RESULTS

Study Group Participation And Engagement

In general, we observe significant behavioral differences between the treatment and control conditions. Students in groups with Diversified Peers consistently demonstrated higher levels of participation in group activities and more frequent contact with their group members compared with students in groups with Randomly Assigned Peers. This impact is further corroborated by student interaction dynamics observed through conversational data.

Diversified peers promotes student participation and satisfaction in group activities. We found that students in the treatment condition, characterized by peers with diverse (high, medium, low) motivation to lead teamwork, have significantly higher levels of behavioral engagement compared with students in study groups with randomly assigned peers. We measured participation using three aggregated, weekly-reported behavioral outcomes, derived from weekly surveys: the total time students spent in meetings with their study group members, and the frequency of study group interactions about both course-relevant and non-course-relevant content.

Specifically, as shown in Figure 2.1 of the standardized treatment effect, students in groups with Diversified Peers dedicated more time to studying with their group members (b = 0.277, SE = 0.111, P = 0.013). They also showed a higher frequency of contact for both course-relevant (b = 0.316, SE = 0.110, t = 2.86, P = 0.005) and non-course-relevant content (b = 0.212, SE = 0.113, P = 0.060). Additionally, these students reported greater satisfaction with their study groups (b = 0.327, SE = 0.119, P = 0.006). Table 2.2 and Table 2.3 provides detailed estimates, both without and with covariates. The non-aggregated weekly outcomes, detailed in Appendix Table B4, consistently indicate a positive overall impact on students' weekly participation.

These findings demonstrate the significant effects of modified group composition on students' behavioral outcomes. The magnitude of this effect is substantial, indicating that diversifying study group composition regarding students' motivation to lead teamwork can be an effective tool in educational settings to foster deeper engagement and satisfaction among students in group activities.

Higher level of leadership and positive tone in treatment groups' conversation. To further investigate the impact of Diversified Peers on students' interaction dynamics, we analyzed the

conversational data from each study group. On average, students in the treatment condition sent 120 messages within their group chat, compared with 123 messages sent by students in the control condition. The total word count for messages in the treatment condition was 764, while it was 791 for the control condition. However, no statistically significant differences were observed in these two measures between treatment and control conditions.

We found that the experimental groups exhibited higher levels of leadership and a more positive tone in their conversations. As shown in Table 2.4, conversations in the treatment condition demonstrated a significantly higher leadership tone (b = 0.046, SE = 0.018) and a more positive tone (b = 0.045, SE = 0.018) compared with conversations in the control. Furthermore, messages from the treatment condition had a higher average word count per sentence (b = 0.231, SE = 0.089).

The results from the text analysis suggest that the intervention exerted a significant influence on student interactions within the study groups, offering further support and insights into its positive impact on student participation and satisfaction. The study groups under the treatment condition not only exhibited more leadership and positivity in their communications but also tended to produce more complex messages, as indicated by a higher average word count per sentence. These findings suggest that the experimental design fostered an environment conducive to more engaging and positive interactions.

Course Social And Learning Experience and Course Grades

Limited positive effects on course social and learning experience and course grades. Additionally, we assessed students' overall perceptions of their study groups, along with their social and learning experiences of the course, using post-course surveys (for detailed measures, refer to Appendix B2). We also analyzed their course grades. As shown in Figure 2.2 and Table

2.5, in terms of study group perceptions, students in the treatment condition reported statistically significantly higher satisfaction with their study group (b = 0.240, SE = 0.120) and perceived greater social support from the study group (b = 0.211, SE = 0.121). They also exhibited positive coefficients in study group involvement, perceived group processes, member relations, and perceived academic support, though these were not statistically significant. Regarding course social experiences, students with Diversified Peers showed higher, yet not statistically significant, coefficients for the number of friends, number of help-seeking peers, and sense of belonging. For course learning experiences, students in the treatment condition reported a higher level of course satisfaction (b = 0.211, SE = 0.121). They showed positive, yet not statistically significant, coefficients for motivation and self-regulated learning skills. In terms of course grades, students in the treatment condition had a marginally significant higher average quiz grade (b = 0.190, SE = 0.107), and positive but not significant coefficients for other grades (midterm, raw final exam grade, and final grade).

Impact On Underrepresented Minority (URM) Students

Promising impacts on behavioral outcomes, study group perception, and course perception for underrepresented minority (URM) students. We investigate the heterogeneous effects of the intervention, particularly the treatment effects on URM students. The primary objective is to assess whether the intervention can improve outcomes among students from disadvantaged backgrounds in a high-stakes online STEM introductory course. The results highlight a significant positive impact of the intervention on URM students—including Hispanic, Black, American Indian, Alaska Native, Native Hawaiian, or Other Pacific Islander students—notably in terms of participation and perceptions in group activities, satisfaction, and overall sense of belonging.

Firstly, we found substantial positive treatment effects for URM students on behavioral engagement and weekly study group satisfaction. Specifically, as shown in Figure 2.3 and Panel A in Table 2.6, URM students in the treatment condition have spent longer time studying with their study groups (b = 0.450, SE = 0.171, P = 0.009), higher contact frequency regarding course-relevant content (b = 0.397, SE = 0.169, P = 0.020), and feel more satisfied with the study groups (b = 0.479, SE = 0.179, P = 0.008). No statistically significant interaction effect was observed for the outcomes of behavioral engagement and satisfaction.

Moreover, positive impacts for URM students in the treatment groups extend to overall perceptions of study-group activities as well (see Figure 4 and Panel B in Table 2.6). Among six measures of study group perceptions, five significant impacts were identified, encompassing group process, relationships among group members, perceived academic as well as social support from the study groups, and overall satisfaction. Most of the interaction terms are not statistically significant. Notably, the most pronounced positive treatment effect for URM students was in their overall study group satisfaction (b = 0.441, SE = 0.181, P = 0.015). It follows the perception of the relationships among group members (b = 0.424, SE = 0.178, P = 0.018), which comprehensively assesses the extent to which group members connect, communicate, work through different opinions, and cooperate well with each other. There is also a positive interaction effect for this outcome, showing that URM students benefit significantly more than non-URM students (b = 0.489, SE = 0.239). URM students also have noticeable treatment effects on social supports (b = 0.372, SE = 0.183, P = 0.043) and academic supports (b = 0.353, SE = 0.178, P = 0.048) from their participation in study groups. Academically, study groups with modified group composition based on motivation to lead were perceived as more beneficial in keeping students on tasks, enhancing their understanding of their abilities in relation

to course goals, and increasing their interest in the course. Socially, URM students in the treatment groups perceived a higher level of inclusivity, connectivity, and care from the study groups.

Finally, in addition to the promising findings directly related to the study group, we observed notable positive effects of the treatment on URM students in broader social networks and a higher sense of belonging within the course (see Figure 4 and Panel C in Table 2.6). URM students in the treatment groups reported a higher number of friends (b = 0.360, SE = 0.164, P = 0.029) and more contacts from whom they could seek help (b = 0.376, SE = 0.178, P = 0.035) in the course; they also exhibited an increased sense of belonging. However, there was no significant evidence of treatment effects on URM students' motivation for the course, self-regulated learning, course satisfaction, or academic performance (see Figure 4 and Panel D in Table 2.6).. Meanwhile, non-URM students exhibited higher course satisfaction at the marginally significant level (b = 0.271, SE = 0.159, P = 0.090).

These findings suggest that the intervention holds the potential for scaling up to provide broader support for students in STEM courses, particularly those from more disadvantaged backgrounds. The heterogeneity analysis underscores the intervention's particular efficacy in bridging gaps for underrepresented minority students, pointing towards promising avenues for future educational strategies in high-stakes, online STEM environments.

DISCUSSION

These findings provide strong causal evidence of the significance of modifying group composition based on group members' social skills, impacting student behavioral and psychological outcomes crucial for STEM identity and persistence (Baldwin 2009; Bambara et al. 2009; Peacock et al., 2020; Peacock & Cowan, 2019). In our randomized controlled trial

conducted within a challenging online STEM gateway course, we found substantial differences between treatment and control conditions. Specifically, students with Diversified Peers showed significantly higher levels of participation, engagement, and satisfaction compared with those with Randomly Assigned Peers. This effect was particularly pronounced among URM students, for whom we also observed a positive impact on group member relationships and increased perceived support from the study groups. Additionally, URM students experienced positive treatment effects on a number of friends and help-seeking peers within the course, alongside an enhanced sense of belonging. This study provides the first causal evidence for efficiently constructing student groups in high-stakes online courses, focusing on student engagement and sense of belonging. Specifically, we demonstrated that including at least one student with high motivation to lead in low-stakes, student-led study groups significantly enhances both student engagement and their sense of belonging.

We also explore how the treatment effect is distributed across students with low, medium, and high motivation to lead. Students with low and medium motivation to lead show more positive treatment effects. These results mitigate one possible concern that modifying group composition based on students' motivation to lead may only benefit those with high motivation to lead. The results are included in Appendix Table B5.

This study presents educators with a practical and scalable approach that addresses the critical challenges of fostering a sense of belonging and engagement in challenging gateway STEM online courses. Low engagement and sense of belonging can lead to serious consequences, such as low retention rates in STEM majors, adversely affecting the STEM workforce. The scalability of this modified group composition method across courses of various sizes, subjects, and diverse student bodies, even with limited educational resources, is feasible

for several reasons. First, the group size requirement is minimal, needing only three students. Second, the categorization of students by their motivation level for leading teamwork is relative, not absolute, offering adaptability across different student profiles and respecting students' own preferences. Third, the generic nature of group tasks, focused on course content but independent of specific subject matter, ensures its applicability across disciplines. Additionally, the use of online platforms for facilitating student interactions and communications enables smooth integration into course routines without necessitating extra demanding tasks, particularly beneficial in challenging courses. This study also contributes to the literature by highlighting the potential of our treatment strategy to reduce the equity gap in engagement and sense of belonging in demanding online STEM courses.

Our study highlights the beneficial effects of diversified group composition (high, medium, low) based on students' motivation to lead on engagement and sense of belonging, without identifying the optimal composition. It's possible that compositions such as high-low-low or high-high-low might yield better effects compared with high-medium-low. Existing literature indicates that grouping students with distinct characteristics may reduce interaction levels, partially supporting our approach to emphasize the influence of highly motivated leaders while assembling study groups from each motivation category. Nevertheless, other compositions, such as high-low-low and high-medium-medium, might yield similar outcomes. Additionally, employing randomly assigned groups as a control condition for practical classroom implications might underestimate the treatment effect, as these groups can inadvertently include a mix of high-medium-low motivation levels. Future research should investigate various compositions and adjust control groups to ascertain the optimal arrangement. Another limitation of this study is the absence of tracking students' long-term outcomes to further examine if the positive

treatment effect extends and further improves students' persistence in STEM fields or their social networks in college.

Our work contributes to the growing literature on developing supportive learning communities aimed at enhancing social belonging and satisfaction among students, particularly those from underserved backgrounds, in historically challenging gateway STEM courses. This study underscores the significance of leveraging insights from the well-established literature on peer effects (Hoxby 2000; Zimmerman 2003), and the critical role of social skills (Arendale, 2004; Barry & Stewart, 1997; Deming, 2017; Hara et al., 2000; Johnson & Johnson, 2002; Riedl et.al.,2021; Woolley et al., 2010), elucidating how these factors contribute to the dynamics of student interaction within group processes. We advocate for future research to further investigate how group composition, with respect to student characteristics influences, influences these processes in ways that are advantageous to students' social belonging and learning within demanding educational contexts.

REFERENCE

Allen, I. Elaine;, and Jeff Seaman. 2015. “Grade Level: Tracking Online Education in the United States.” *Babson Survey Research Group; Quahog Research Group, LLC*.

Arbaugh, J. B., Cleveland-Innes, M., Díaz, S. R., Garrison, D. R., Ice, P., Richardson, J. C., ... & Swan, K. P. (2007). Community of inquiry survey instrument (draft v14). *Community of Inquiry Survey*.

Arendale, David. 2014. “Understanding the Peer Assisted Learning Model: Student Study Groups in Challenging College Courses.” *International Journal of Higher Education* 3(2). doi: 10.5430/ijhe.v3n2p1.

Arendale, David R. 2004. “Pathway to Persistence: A Review of Postsecondary Peer Cooperative Learning Programs.” *Best Practices for Access and Retention in Higher Education*.

Baldwin, Roger G. 2009. “The Climate for Undergraduate Teaching and Learning in STEM Fields.” *New Directions for Teaching and Learning* (117). doi: 10.1002/tl.340.

Bambara, Cynthia S., Clifford P. Harbour, Timothy Gray Davies, and Susan Athey. 2009. “Delicate Engagement: The Lived Experience of Community College Students Enrolled in High-Risk Online Courses.” *Community College Review* 36(3). doi: 10.1177/0091552108327187.

Barry, Bruce, and Greg L. Stewart. 1997. “Composition, Process, and Performance in Self-Managed Groups: The Role of Personality.” *Journal of Applied Psychology* 82(1). doi: 10.1037/0021-9010.82.1.62.

Chan, K. Y., & Drasgow, F. (2001). Toward a theory of individual differences and leadership: understanding the motivation to lead. Journal of applied psychology, 86(3), 481.

Chirikov, Igor, Tatiana Semenova, Natalia Maloshonok, Eric Bettinger, and René F. Kizilcec. 2020. “Online Education Platforms Scale College STEM Instruction with Equivalent Learning Outcomes at Lower Cost.” *Science Advances* 6(15). doi: 10.1126/sciadv.aay5324.

Cohn, Michael A., Matthias R. Mehl, and James W. Pennebaker. 2004. “Linguistic Markers of Psychological Change Surrounding September 11, 2001.” *Psychological Science* 15(10). doi: 10.1111/j.0956-7976.2004.00741.x.

Deming, D. J. (2017). The growing importance of social skills in the labor market. The Quarterly Journal of Economics, 132(4), 1593-1640.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. Annual review of psychology, 53(1), 109-132.

Freeman, T. M., Anderman, L. H., & Jensen, J. M. (2007). Sense of belonging in college freshmen at the classroom and campus levels. *The Journal of Experimental Education*, *75*(3), 203-220.

Grau, Valeska, and David Whitebread. 2012. “Self and Social Regulation of Learning during Collaborative Activities in the Classroom: The Interplay of Individual and Group Cognition.” *Learning and Instruction* 22(6). doi: 10.1016/j.learninstruc.2012.03.003.

Hadwin, Allyson, Sanna Järvelä, and Mariel Miller. 2017. “SELF-REGULATION, CO-REGULATION, AND SHARED REGULATION IN COLLABORATIVE LEARNING ENVIRONMENTS.” in *Handbook of Self-Regulation of Learning and Performance, Second Edition*.

Hara, Noriko, Curtis Jay Bonk, and Charoula Angeli. 2000. “Content Analysis of Online Discussion in an Applied Educational Psychology Course.” *Instructional Science* 28(2). doi: 10.1023/A:1003764722829.

Hoxby, Caroline M. (Harvard University). 2000. “Peer Effects in the Classroom: Learning from Gender and Race Variation.” *NBER Working Paper* (7867).

Hurtado, S., & Carter, D. F. (1997). Effects of college transition and perceptions of the campus racial climate on Latino college students' sense of belonging. *Sociology of education*, 324-345.

Järvenoja, H., Volet, S., & Järvelä, S. (2013). Regulation of emotions in socially challenging learning situations: An instrument to measure the adaptive and social nature of the regulation process. *Educational Psychology*, *33*(1), 31-58.

Johnson, David W., and Roger T. Johnson. 2002. “Learning Together and Alone: Overview and Meta-analysis.” *Asia Pacific Journal of Education* 22(1). doi: 10.1080/0218879020220110.

Jung, D. I., & Sosik, J. J. (2002). Transformational leadership in work groups: The role of empowerment, cohesiveness, and collective-efficacy on perceived group performance. *Small group research*, *33*(3), 313-336.

Kacewicz, Ewa, James W. Pennebaker, Matthew Davis, Moongee Jeon, and Arthur C. Graesser. 2014. “Pronoun Use Reflects Standings in Social Hierarchies.” *Journal of Language and Social Psychology* 33(2). doi: 10.1177/0261927X13502654.

Ke, Fengfeng, and Dean Kwak. 2013. “Online Learning across Ethnicity and Age: A Study on Learning Interaction Participation, Perception, and Learning Satisfaction.” *Computers and Education* 61(1). doi: 10.1016/j.compedu.2012.09.003.

Lai, C. H., Lin, H. W., Lin, R. M., & Tho, P. D. (2019). Effect of peer interaction among online learning community on learning engagement and achievement. *International Journal of Distance Education Technologies (IJDET)*, *17*(1), 66-77.

Langdon, D., McKittrick, G., Beede, D., Khan, B., & Doms, M. (2013). STEM: Good jobs now and for the future. In *STEM (Science, Technology, Engineering and Mathematics) Workforce Trends and Policy Considerations*.

Miriti, M. N. (2020). The elephant in the room: race and STEM diversity. BioScience, 70(3), 237-242.

Monsjou, Elizabeth Van, Muise Amy, Karen Fergus, and Charles Ward Struthers. 2022. "The Development and Psychometric Properties of LIWC-22." *The University of Texas at Austin* 1(2).

National Center for Science and Engineering Statistics (NCSES). 2023. Diversity and STEM: Women, Minorities, and Persons with Disabilities 2023. Special Report NSF 23-315. Alexandria, VA: National Science Foundation. Available at https://ncses.nsf.gov/wmpd.

Peacock, S., & Cowan, J. (2019). Promoting sense of belonging in online learning communities of inquiry in accredited courses. *Online Learning*, 23(2), 67-81.

Peacock, S., Cowan, J., Irvine, L., & Williams, J. (2020). An exploration into the importance of a sense of belonging for online learners. *International Review of Research in Open and Distributed Learning*, 21(2), 18-35.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 33–40. https://doi.org/10.1037/0022-0663.82.1.33

Rainey, Katherine, Melissa Dancy, Roslyn Mickelson, Elizabeth Stearns, and Stephanie Moller. 2018. "Race and Gender Differences in How Sense of Belonging Influences Decisions to Major in STEM." *International Journal of STEM Education* 5(1). doi: 10.1186/s40594-018-0115-6.

Rasheed, R. A., Kamsin, A., & Abdullah, N. A. (2021). An approach for scaffolding students peer-learning self-regulation strategy in the online component of blended learning. IEEE Access, 9, 30721-30738.

Riedl, Christoph, Young Ji Kim, Pranav Gupta, Thomas W. Malone, and Anita Williams Woolley. 2021. "Quantifying Collective Intelligence in Human Groups." *Proceedings of the National Academy of Sciences of the United States of America* 118(21). doi: 10.1073/pnas.2005737118.

Riegle-Crumb, C., King, B., & Irizarry, Y. (2019). Does STEM Stand Out? Examining Racial/Ethnic Gaps in Persistence Across Postsecondary Fields. *Educational Researcher, 48*(3). https://doi.org/10.3102/0013189X19831006

Sheeran, Nicola, and Daniel J. Cummings. 2018. "An Examination of the Relationship between Facebook Groups Attached to University Courses and Student Engagement." *Higher Education* 76(6). doi: 10.1007/s10734-018-0253-2.

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. Review of educational research, 69(1), 21-51.

Strayhorn, T. L. (2018). *College students' sense of belonging: A key to educational success for all students*. Routledge.

Thai, Michael, Nicola Sheeran, and Daniel J. Cummings. 2019. "We're All in This Together: The Impact of Facebook Groups on Social Connectedness and Other Outcomes in Higher Education." *Internet and Higher Education* 40. doi: 10.1016/j.iheduc.2018.10.001.

U.S. Bureau of Labor Statistics. (2023). Table 1.11 Employment in STEM occupations, 2022 and projected 2032. Employment Projections program. https://www.bls.gov/emp/tables/stem-employment.htm

U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Fall 2016, Completions component. See Digest of Education Statistics 2017, tables 318.45 and 322.30. https://nces.ed.gov/programs/raceindicators/indicator_reg.asp#:~:text=In%20contrast%2C%20the%20percentages%20of,degrees%20awarded%20in%20STEM%20fields.

Woolley, Anita Williams, Christopher F. Chabris, Alex Pentland, Nada Hashmi, and Thomas W. Malone. 2010. "Evidence for a Collective Intelligence Factor in the Performance of Human Groups." *Science* 330(6004). doi: 10.1126/science.1193147.

Xie, Y., Fang, M., & Shauman, K. (2015). STEM Education. *Annual Review of Sociology*, *41*. https://doi.org/10.1146/annurev-soc-071312-145659

Xu, D., & Jaggars, S. S. (2013). The impact of online learning on students' course outcomes: Evidence from a large community and technical college system. Economics of Education Review, 37, 46-57.

Xu, D., Solanki, S., McPartlan, P., & Sato, B. (2018). EASEing students into college: The impact of multidimensional support for underprepared students. Educational Researcher, 47(7), 435–450. https://doi.org/10.3102/0013189X18778559

Xu, D., & Xu, Y. (2020). The ambivalence about distance learning in higher education: Challenges, opportunities, and policy implications. Higher Education: Handbook of Theory and Research: Volume 35, 351-401.

Zhou, Xuehan, Qiujie Li, Di Xu, Amanda Holton, and Brian K. Sato. 2023. “The Promise of Using Study-Together Groups to Promote Engagement and Performance in Online Courses: Experimental Evidence on Academic and Non-Cognitive Outcomes.” *Internet and Higher Education* 59. doi: 10.1016/j.iheduc.2023.100922.

Zimmerman, David J. 2003. “Peer Effects in Academic Outcomes: Evidence from a Natural Experiment.” *Review of Economics and Statistics* 85(1).

Table 2. 1 Student Characteristics by experimental condition

	Treatment group	Control group	p-value
Gender			
Male	0.286	0.318	0.558
	(0.038)	(0.038)	
Female	0.693	0.676	0.755
	(0.039)	(0.039)	
Non-binary/third gender	0.014	0.000	0.146
	(0.010)	(0.000)	
Prefer not to say	0.007	0.007	0.969
	(0.007)	(0.007)	
Race/ethnicity			
American Indian, Alaska Native, Native Hawaiian, or Other Pacific Islander	0.043	0.041	0.922
	(0.017)	(0.016)	
Black or African American	0.043	0.047	0.857
	(0.017)	(0.018)	
Hispanic, Latino, or Spanish origin	0.357	0.324	0.558
	(0.041)	(0.039)	
Asian	0.443	0.466	0.692
	(0.042)	(0.041)	
White	0.114	0.115	0.988
	(0.027)	(0.026)	
missing	0.000	0.007	0.332
	(0.000)	(0.007)	
Low income status	0.429	0.453	0.681
	(0.042)	(0.041)	
First-generation college student status	0.500	0.500	1.000
	(0.042)	(0.041)	
High school GPA	3.840	3.82	0.531
	(0.019)	(0.025)	
Composite score of course motivation	3.524	3.59	0.367
	(0.054)	(0.05)	
Composite score of peer interaction preference	3.331	3.173	0.094
	(0.068)	(0.065)	
Composite score of motivation to lead	3.109	3.155	0.600
	(0.063)	(0.062)	
Composite score of stress level	2.822	2.763	0.356
	(0.047)	(0.044)	
Number of classmates known in the class	1.671	1.642	0.778
	(0.079)	(0.069)	
Number of friends known in the class	1.386	1.311	0.221
	(0.046)	(0.041)	
Number of students that can ask for help in the class	1.407	1.324	0.191
	(0.049)	(0.041)	
Entry term			
Fall 2016	0.007	0.000	0.305
	(0.007)	(0.000)	
Fall 2017	0.007	0.000	0.305
	(0.007)	(0.000)	
Fall 2019	0.014	0.007	0.531
	(0.010)	(0.007)	
Fall 2020	0.043	0.061	0.495
	(0.017)	(0.020)	
Fall 2021	0.221	0.243	0.663

	(0.035)	(0.035)	
Fall 2022	0.707	0.689	0.741
	(0.039)	(0.038)	
Observations	140	148	Total: 288

Note: Standard errors in parentheses; significant level: + p<0.10 * p<0.05 ** p<0.01 *** p<0.001

Table 2. 2 Treatment effect on aggregated weekly outcomes without covariates

	Study group interaction frequency (course-related)		Study group interaction frequency (non-course-related)		Total study group study times		Study group satisfaction (weekly)	
	No Covariates	With Covariates	No Covariates	With Covariates	No Covariates	With Covariates	No Covariates	With Covariates
Treatment group	0.318**	0.316**	0.228*	0.212+	0.247*	0.277*	0.300*	0.327**
	(0.109)	(0.110)	(0.112)	(0.113)	(0.112)	(0.111)	(0.117)	(0.119)
Observations	288	288	288	288	288	288	283	283
R-squared	0.029	0.126	0.014	0.136	0.017	0.152	0.023	0.127

Note: Standard errors in parentheses; significant level: + p<0.10 * p<0.05 ** p<0.01 *** p<0.001

Table 2. 3 Treatment effect on aggregated weekly outcomes with covariates

	Study group interaction frequency (course-related)	Study group interaction frequency (non-course-related)	Total study group study times	Study group satisfaction (weekly)
Treatment group	0.316**	0.212+	0.277*	0.327**
	(0.110)	(0.113)	(0.111)	(0.119)
Observations	288	288	288	283
R-squared	0.126	0.136	0.152	0.127

Note: Standard errors in parentheses; significant level: + p<0.10 * p<0.05 ** p<0.01 *** p<0.001

Table 2. 4 Impact of treatment condition on LIWC outcomes

	Language of leadership, status	Degree or positive (negative) tone	Average words per sentence
Treatment	0.046**	0.045*	0.231**
	(0.018)	(0.018)	(0.089)
Constant	-0.023+	-0.022+	6.037***
	(0.012)	(0.012)	(0.063)
Mean	45.481	39.639	6.151
SD	38.683	34.544	5.068
Observations	12,854	12,854	12,854

Note: All outcomes are standardized; Standard errors in parentheses; significant level: + $p<0.10$ * $p<0.05$ ** $p<0.01$ *** $p<0.001$

Table 2. 5 Impact of treatment condition on student perception and academic performance

	No Covariates	With Covariates
Study group perception		
Study group satisfaction	0.224+	0.240*
	(0.117)	(0.120)
Composite score on involvement	0.110	0.113
	(0.118)	(0.115)
Composite score on group process	0.116	0.118
	(0.118)	(0.120)
Composite score on member relations	0.139	0.151
	(0.118)	(0.119)
Composite score on academic support	0.122	0.150
	(0.118)	(0.118)
Composite score on social support	0.223+	0.211+
	(0.117)	(0.121)
Course perception		
Number of Friends in Course	0.181	0.073
	(0.118)	(0.109)
Number of Advisers in Course	0.098	0.058
	(0.118)	(0.118)
Composite Score of Motivation	0.014	0.107
	(0.118)	(0.101)
Composite Score of Sense of Belonging	0.131	0.117
	(0.118)	(0.118)
Composite Score of Self-Regulated Learning	0.073	0.063
	(0.118)	(0.114)
Course Satisfaction	0.136	0.209+
	(0.118)	(0.119)
Course grade		
Average quiz grade	0.157	0.190+
	(0.109)	(0.107)
Average midterm grade	0.037	0.032
	(0.101)	(0.095)
Raw final grade	0.043	0.044
	(0.107)	(0.100)
Final grade	0.049	0.043
	(0.096)	(0.091)
Observations	288	288

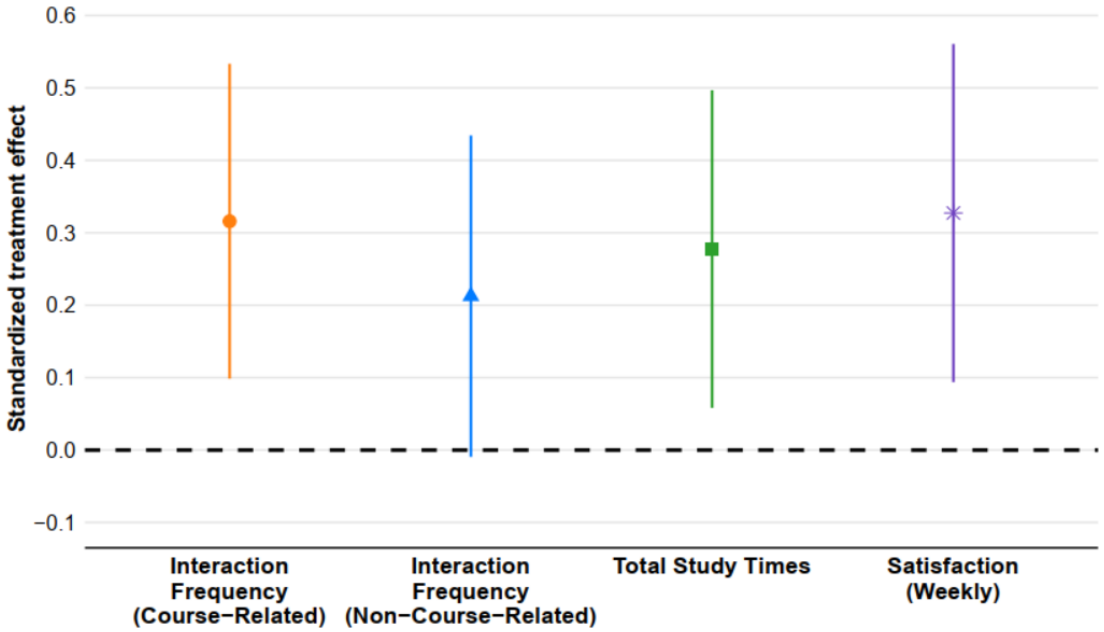
Note: Standard errors in parentheses; significant level: + p<0.10 * p<0.05 ** p<0.01 *** p<0.001; the left column lists the different outcomes for each model.

Table 2. 6 Heterogeneous analysis with covariates (n = 287, URM = 123 non-URM = 167)

	URM	Non-URM	Diff.
Panel A: Behavioral outcomes			
Study group interaction frequency (course-related)	0.397*	0.247	0.149
	(0.169)	(0.150)	(0.227)
Study group interaction frequency (non-course-related)	0.222	0.197	0.024
	(0.173)	(0.153)	(0.232)
Total study group study times	0.450**	0.132	0.319
	(0.171)	(0.152)	(0.230)
Study group satisfaction (weekly)	0.479**	0.211	0.268
	(0.179)	(0.158)	(0.240)
Panel B: Study group perception			
Study group satisfaction	0.441*	0.083	0.358
	(0.181)	(0.160)	(0.243)
Composite score on involvement	0.283	-0.029	0.312
	(0.174)	(0.154)	(0.234)
Composite score on group process	0.318+	-0.042	0.359
	(0.180)	(0.159)	(0.242)
Composite score on member relations	0.424*	-0.065	0.489*
	(0.178)	(0.158)	(0.239)
Composite score on academic support	0.353*	-0.011	0.364
	(0.178)	(0.157)	(0.239)
Composite score on social support	0.372*	0.085	0.286
	(0.183)	(0.161)	(0.245)
Panel C: Course social experience			
Number of Friends in Course	0.360*	-0.152	0.511*
	(0.164)	(0.145)	(0.220)
Number of help-seeking peers	0.376*	-0.188	0.564*
	(0.178)	(0.157)	(0.239)
Composite score of sense of belonging	0.427*	-0.135	0.562*
	(0.177)	(0.157)	(0.238)
Panel D: Course learning experience			
Composite score of motivation	0.024	0.165	-0.142
	(0.155)	(0.137)	(0.208)
Composite score of self-regulated learning	0.128	0.011	0.117
	(0.173)	(0.153)	(0.233)
Course satisfaction	0.133	0.271+	-0.139
	(0.180)	(0.159)	(0.242)
Average quiz grade	0.120	0.235	-0.116
	(0.163)	(0.144)	(0.218)
Average midterm grade	-0.065	0.103	-0.169
	(0.144)	(0.128)	(0.194)
Raw final grade	-0.007	0.074	-0.081
	(0.153)	(0.136)	(0.206)
Final grade	-0.024	0.089	-0.113
	(0.139)	(0.123)	(0.187)
Covariates	Yes	Yes	Yes

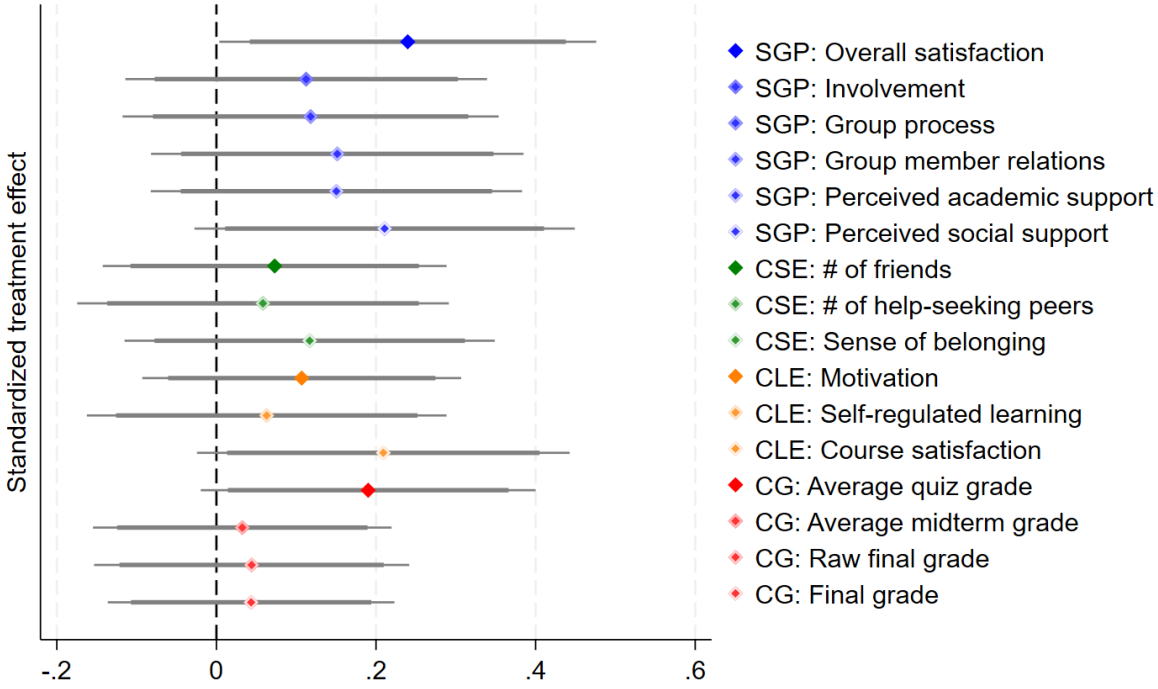
Note: Standard errors in parentheses; significant level: + p<0.10 * p<0.05 ** p<0.01 *** p<0.001; the left column lists the different outcomes for each model; the estimates for URM and Non-URM are derived from the main effect, which includes an interaction term between URM status and the treatment condition; the “Diff.” represents the interaction term.

Figure 2. 1 Treatment effect on study group participation and engagement outcomes



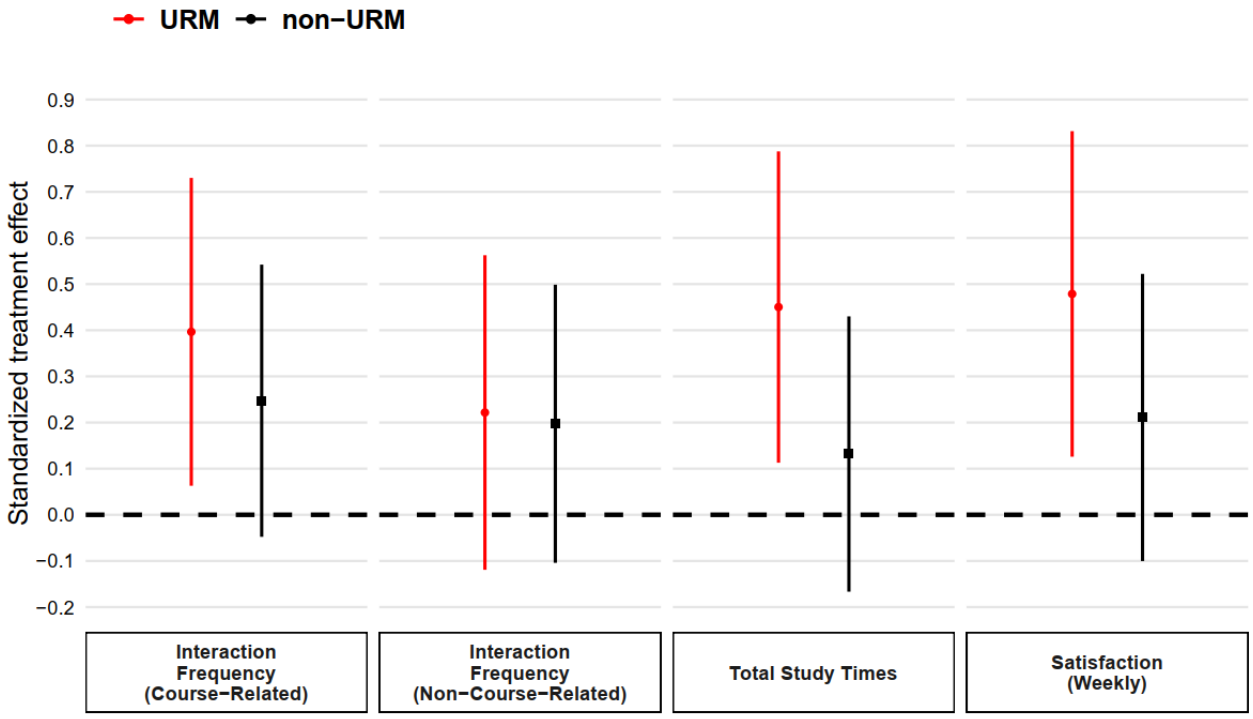
Note: This figure uses estimates at a significance level of 0.05; the outcome of total study times is marginally significant at a significance level of 0.1.

Figure 2. 2 Treatment effect on study group perception (SGP), course social experience (CSE), course learning experience (CLE), and course grade (CG)



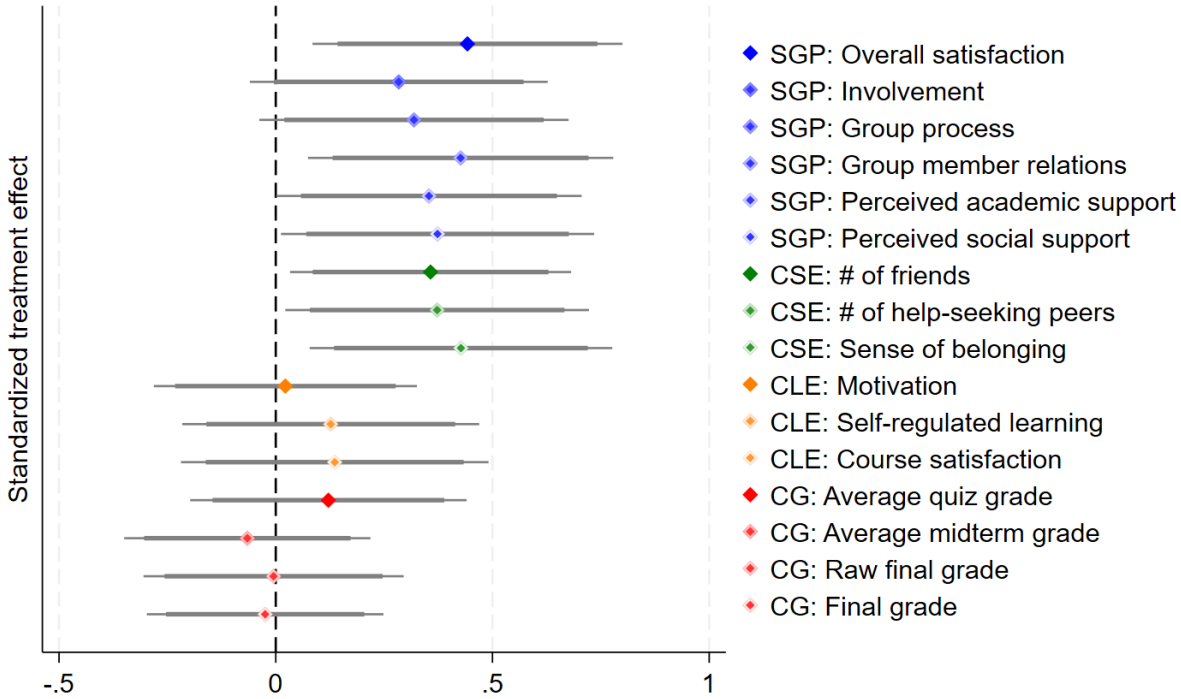
Note: The 95% confidence interval is represented by the thicker line, while the 90% confidence interval is represented by the thinner line.

Figure 2. 3 Treatment effect on study group participation and engagement outcomes for URM and non-URM students



Note: This figure uses estimates at a significance level of 0.05.

Figure 2. 4 Treatment effect on study group perception (SGP), course social experience (CSE), course learning experience (CLE), and course grade (CG) for URM students



Note: The 95% confidence interval is represented by the thicker line, while the 90% confidence interval is represented by the thinner line.

APPENDIX

Appendix Table B1 Group composition of group members' motivation to lead categories for control condition and treatment condition

Group composition of motivation to lead	Control Condition	
	Freq.	**Percent**
high, medium, low	10	18%
high, high, low	7	13%
high, high, medium	7	13%
high, high, high	1	2%
medium, medium, medium	2	4%
medium, medium, high	3	5%
medium, medium, low	8	14%
low, low, low	2	4%
low, low, high	2	4%
low, low, medium	3	5%
high, high	3	5%
high, medium	1	2%
high, low	1	2%
low, medium	2	4%
low, low	2	4%
medium, medium	2	4%
Total	56	100%

Group composition of motivation to lead	Treatment condition	
	Freq.	**Percent**
high, medium, low	47	84%
high, high, low	1	2%
high, medium	5	9%
high, low	3	5%
Total	56	100%

Note: The two people groups in both treatment and control conditions are due to the original group member drop or withdraw from the class.

Appendix Table B2

	Attrition (dropped out)		Attrition (No post-course survey)	
	(1)	(2)	(3)	(4)
Treatment group	-0.005	-0.001	0.037	0.030
	(0.029)	(0.029)	(0.039)	(0.038)
Gender (reference group: male)				
Female		0.026		0.017
		(0.032)		(0.042)
Non-binary/third gender		0.014		0.457**
		(0.114)		(0.150)
Prefer not to say		0.495***		0.344+
		(0.140)		(0.184)
Race/ethnicity (reference group: White)				
American Indian, Alaska Native, Native Hawaiian, or Other Pacific Islander		-0.009		0.097
		(0.082)		(0.108)
Black or African American		0.044		0.071
		(0.080)		(0.105)
Hispanic, Latino, or Spanish origin		0.095+		0.055
		(0.052)		(0.069)
Asian		0.065		-0.005
		(0.050)		(0.066)
missing		-0.051		-0.167
		(0.269)		(0.355)
High school GPA		-0.087+		-0.088
		(0.048)		(0.064)
Composite score of course motivation		-0.013		-0.072*
		(0.023)		(0.031)
Composite score of peer interaction preference		-0.031+		-0.012
		(0.018)		(0.024)
Composite score of motivation to lead		-0.022		-0.024
		(0.020)		(0.026)
Composite score of stress level		-0.013		-0.018
		(0.027)		(0.035)
Number of classmates known in the class		0.005		-0.019
		(0.019)		(0.025)
Number of friends known in the class		-0.117*		-0.140*
		(0.046)		(0.060)
Number of students that can ask for help in the class		0.071+		0.030
		(0.040)		(0.053)

Observations	340	340	340	340
R-squared	0.000	0.099	0.003	0.118

Note: Standard errors in parentheses; significant level: + p<0.10 * p<0.05 ** p<0.01 *** p<0.001

Appendix Table B3 Balance Test of OLS Regression Results

	Treatment group	Treatment group	Treatment group
	Sample A: all students stay in the class	*Sample B: all students with post-course survey*	*Sample C: all students with pre-course survey*
Gender (reference group: male)			
Female	0.046	0.020	0.005
	(0.065)	(0.068)	(0.062)
Non-binary/third gender	0.384	0.695	0.178
	(0.242)	(0.375)	(0.218)
Prefer not to say	-0.014	0.005	-0.052
	(0.396)	(0.403)	(0.268)
Race/ethnicity (reference group: White)			
American Indian, Alaska Native, Native Hawaiian, or Other Pacific Islander	0.054	0.011	0.103
	(0.174)	(0.193)	(0.157)
Black or African American	0.059	-0.000	0.082
	(0.158)	(0.169)	(0.153)
Hispanic, Latino, or Spanish origin	0.017	0.022	-0.005
	(0.112)	(0.116)	(0.100)
Asian	-0.029	-0.043	-0.003
	(0.099)	(0.105)	(0.097)
missing	-0.459	-0.446	-0.440
	(0.519)	(0.521)	(0.516)
Low income status	-0.059	-0.026	
	(0.068)	(0.071)	
First-generation college student status	0.005	0.001	
	(0.075)	(0.078)	
High school GPA	0.054	0.055	0.009
	(0.113)	(0.117)	(0.093)
Composite score of course motivation	-0.066	-0.056	-0.055
	(0.048)	(0.052)	(0.045)
Composite score of peer interaction preference	0.058	0.064	0.051
	(0.037)	(0.040)	(0.034)
Composite score of motivation to lead	-0.017	-0.026	-0.015
	(0.041)	(0.044)	(0.038)
Composite score of stress level	0.055	0.056	0.040
	(0.056)	(0.059)	(0.052)
Number of classmates known in the class	-0.009	-0.013	-0.028
	(0.039)	(0.040)	(0.036)
Number of friends known in the class	0.019	0.025	0.002
	(0.094)	(0.095)	(0.088)
Number of students that can ask for help in the class	0.046	0.052	0.078
	(0.084)	(0.086)	(0.077)
Entry term control	+	+	+
Observations	313	288	340
R-squared	0.049	0.052	0.028

Note: Standard errors in parentheses; significant level: + p<0.10 * p<0.05 ** p<0.01 *** p<0.001

Appendix Table B4 Treatment impact on weekly outcomes with covariates

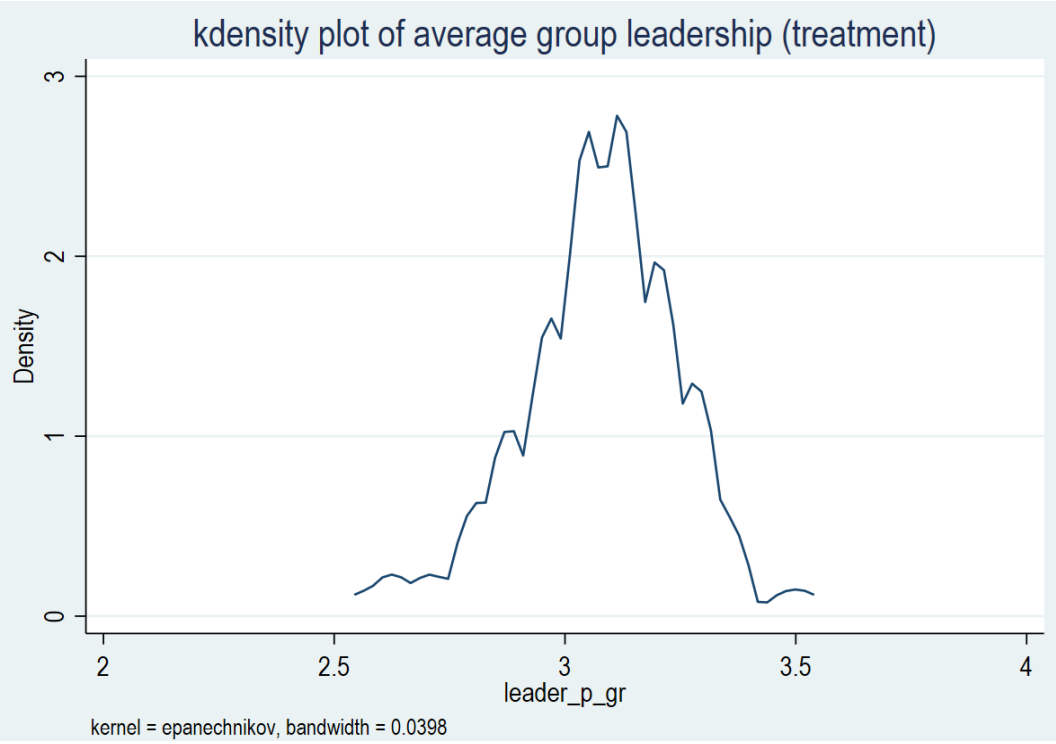
Week	Study group interaction frequency (course-related)	Study group interaction frequency (non-course-related)	Total study group study times	Observations	Study group satisfaction	Observations
2	0.326**	0.155	0.169	288		
	(0.109)	(0.111)	(0.116)			
3	0.289*	0.139	0.113	288		
	(0.112)	(0.116)	(0.114)			
4	0.176	0.160	0.131	288		
	(0.110)	(0.114)	(0.115)			
5	0.275*	0.326**	0.266*	288	0.222+	263
	(0.110)	(0.111)	(0.112)		(0.123)	
6	0.212+	0.209+	0.169	288	0.332**	268
	(0.112)	(0.114)	(0.113)		(0.119)	
7	0.213+	0.150	0.311**	288	0.245*	263
	(0.111)	(0.114)	(0.111)		(0.123)	
8	0.092	0.062	0.102	288	0.279*	259
	(0.114)	(0.114)	(0.113)		(0.124)	
9	0.279*	0.183	0.236*	288	0.272*	257
	(0.111)	(0.113)	(0.113)		(0.124)	

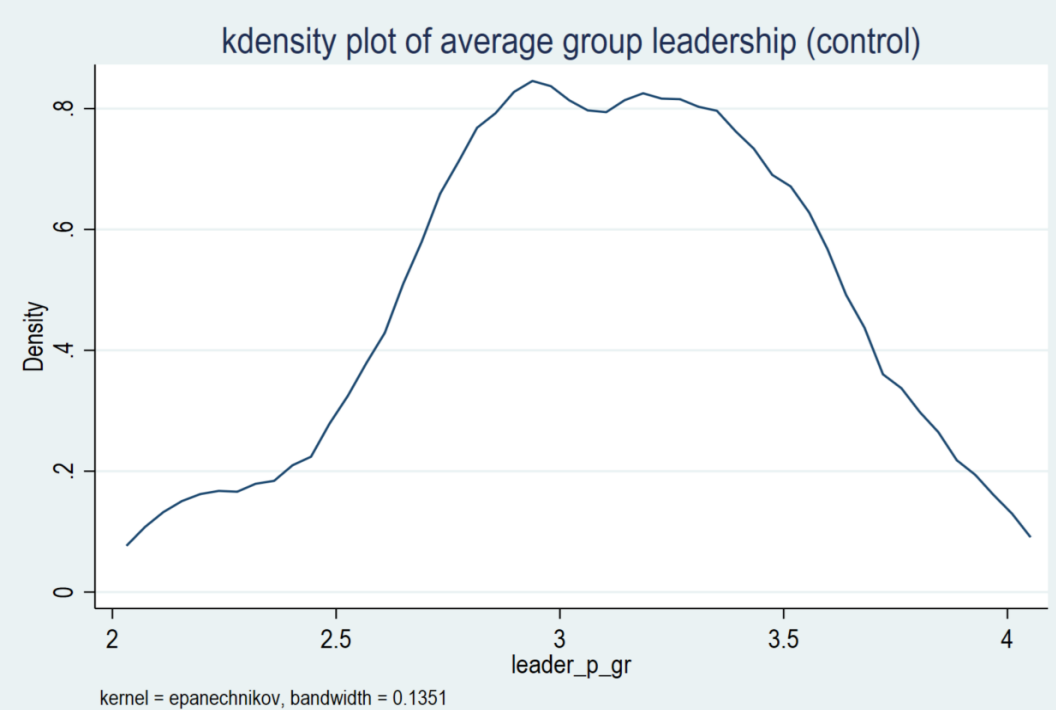
Note: Standard errors in parentheses; significant level: + p<0.10 * p<0.05 ** p<0.01 *** p<0.001; Sample size varies as missing satisfaction data weren't replaced with zeros.

Appendix Table B5 Treatment impact on students with low, medium, and high motivation to lead

	Motivation to Lead				
	Low	Medium	High	Diff. (Low and High)	Diff. (Low and Medium)
Behavioral outcomes					
Study group interaction frequency (course-related)	0.430*	0.264	0.264	0.166	0.167
	(0.195)	(0.188)	(0.185)	(0.269)	(0.271)
Study group interaction frequency (non-course-related)	0.218	0.158	0.312	-0.094	0.061
	(0.202)	(0.194)	(0.191)	(0.278)	(0.280)
Total study group study times	0.341+	0.232	0.186	0.155	0.109
	(0.201)	(0.193)	(0.191)	(0.277)	(0.279)
Study group satisfaction (weekly)	0.176	0.541**	0.184	-0.009	-0.365
	(0.209)	(0.202)	(0.198)	(0.288)	(0.291)
Study group perception					
Study group satisfaction	0.251	0.307	0.111	0.140	-0.056
	(0.211)	(0.202)	(0.200)	(0.291)	(0.292)
Composite score on involvement	0.132	0.222	-0.014	0.147	-0.090
	(0.210)	(0.202)	(0.199)	(0.290)	(0.291)
Composite score on group process	0.297	-0.084	0.149	0.149	0.382
	(0.211)	(0.203)	(0.200)	(0.291)	(0.292)
Composite score on member relations	0.190	0.070	0.159	0.031	0.120
	(0.211)	(0.203)	(0.200)	(0.291)	(0.293)
Composite score on academic support	0.050	0.343+	-0.027	0.077	-0.293
	(0.211)	(0.203)	(0.200)	(0.291)	(0.292)
Composite score on social support	0.120	0.473*	0.061	0.059	-0.353
	(0.209)	(0.201)	(0.199)	(0.289)	(0.290)
Course perception					
Number of Friends in Course	0.069	0.187	0.263	-0.194	-0.119
	(0.211)	(0.202)	(0.200)	(0.291)	(0.292)
Number of Advisers in Course	0.288	0.147	-0.148	0.437	0.141
	(0.208)	(0.200)	(0.197)	(0.286)	(0.288)
Composite Score of Motivation	-0.048	-0.083	0.170	-0.217	0.035
	(0.212)	(0.203)	(0.201)	(0.292)	(0.293)
Composite Score of Sense of Belonging	-0.040	0.299	0.120	-0.160	-0.339
	(0.212)	(0.203)	(0.201)	(0.292)	(0.293)
Composite Score of Self-Regulated Learning	-0.198	0.077	0.318	-0.516+	-0.275
	(0.211)	(0.203)	(0.200)	(0.291)	(0.292)
Course Satisfaction	-0.133	0.047	0.464*	-0.597*	-0.180
	(0.210)	(0.202)	(0.200)	(0.290)	(0.292)
Course grade					
Average quiz grade	0.360+	0.041	0.101	0.259	0.319
	(0.196)	(0.188)	(0.186)	(0.270)	(0.271)
Average midterm grade	-0.005	-0.018	0.139	-0.144	0.013
	(0.179)	(0.172)	(0.170)	(0.247)	(0.249)
Raw final grade	0.125	-0.103	0.130	-0.005	0.228
	(0.189)	(0.182)	(0.180)	(0.261)	(0.262)
Final grade	0.167	-0.120	0.123	0.044	0.287
	(0.170)	(0.163)	(0.161)	(0.235)	(0.236)
Observations	90	98	100	288	288

Appendix Figure B1 Kdensity plots of average group motivation to lead composite scores between treatment and control conditions





Appendix B1 Pre-course survey

Pre-Course Survey (2023 Winter)

Q0 Welcome to the course X Pre-Course Survey. This survey is part of a study exploring strategies that can increase student success in online courses. It is intended to help us understand your feelings about this course and your learning approaches. You will receive **1 extra credit point** (each extra credit point will add 2.5% to your final exam score) by completing this survey. The survey should only take you 5-8 minutes to complete. Your personal information will be kept confidential and will not be shared with the instructor. If you are willing to participate, please log in with your UCI email address to complete the survey by **January 13th at 11:59 PM**.

Alternatively, if you choose not to complete the survey, you will have the opportunity to earn the extra credit point by completing an alternate essay assignment (3-4 paragraphs each).

If you have any questions or comments and/or wish to complete the alternative essay, please contact **XunFei Li (xunfeil@uci.edu)**.

Q1 What is your gender?

- ○ Male
- ○ Female
- ○ Non-binary / third gender
- ○ Prefer not to say

Q2 Are you of Hispanic, Latino, or Spanish origin?

- ○ Yes
- ○ No

Q3 What is your race?

- [ ] White
- [ ] Black or African American
- [ ] American Indian or Alaska Native
- [ ] Asian
- [ ] Native Hawaiian or Other Pacific Islander
- [ ] Some Other Race _____

Q4 Your high school GPA (4.0-scale)?
Note: Please fill in a number.

_____

Q5 How much fun is learning Chemistry to you?

- ○ None at all
- ○ A little
- ○ A moderate amount
- ○ A lot
- ○ A great deal

Q6 How important is performing well in this course (COURSE X) to you?

- Not at all important
- Slightly important
- Moderately important
- Very important
- Extremely important

Q7 How much does being good at the material taught in this course (CHEM 1A) mean to you for your life?

- Not at all
- A little
- A moderate amount
- A lot
- A great deal

Q8 How many students enrolled in this course (COURSE X) did you **know** prior to this course?
Note: This link *will show you the roster of COURSE X.*

- 0
- 1-3
- 4-7
- 7-9
- 10 or more

Q9 Friendship Network Questions How many **friends** do you have in this course (COURSE X)?

- 0
- 1-3

○ 4-7

○ 7-9

○ 10 or more

Q10 Please select students in this course (COURSE X) from the dropdown list that you consider your **friends**.
Note: You can easily locate your friend by typing their first name.

Friend 1	▼
Friend 2	▼
Friend 3	▼
Friend 4	▼
Friend 5	▼
Friend 6	▼
Friend 7	▼
Friend 8	▼
Friend 9	▼
Friend 10	▼

Q11 If you have more than 10 **friends**, or friends in this course (COURSE X) that you cannot find their name, please do your best to list their first and last names here.

_____

Q12 If you did not list any names in the previous question, please tell us why.

○ I already listed their name.

○ I accidentally skipped the question. (If so, you can go back to answer it)

○ I have friends in this course (COURSE X), but I do not know their names.

○ I have friends in this course (COURSE X), but I do not feel comfortable giving their names.

○ I listed their names manually.

Q13 Help-Seeking Network Questions How many students do you have in this course (COURSE X) you would **ask for advice or help**?

- ○ 0
- ○ 1-3
- ○ 4-7
- ○ 7-9
- ○ 10 or more

Q14 Please select students in this course (COURSE X) from the dropdown list that you would **ask for advice or help**.
Note: You can easily locate your friend by typing their first name.

Person 1	▼
Person 2	▼
Person 3	▼
Person 4	▼
Person 5	▼
Person 6	▼
Person 7	▼
Person 8	▼
Person 9	▼
Person 10	▼

Q15 If you have more than 10 students that you would **ask for advice or help**, or students in this course (COURSE X) that you cannot find their name, please do your best to list their first and last names here.

_____

Q16 If you did not list any names in the previous question, please tell us why.

- ○ I already listed their name.
- ○ I accidentally skipped the question. (If so, you can go back to answer it)
- ○ I have students that I can ask for advice or help in this course (COURSE X), but I do not know their names.
- ○ I have students that I can ask for advice or help in this course (COURSE X), but I do not feel comfortable giving their names.
- ○ I listed their names manually.

Q17 The following questions ask about your peer interactions preference in terms of completing coursework. Please select the response that best describes how you feel.

	Very untrue of me	Somewhat untrue of me	Neutral	Somewhat true of me	Very true of me
I prefer to study with peers in the same course rather than studying individually.	○	○	○	○	○
I think that frequently interacting with my classmates would help me succeed in the course.	○	○	○	○	○
I can learn better when studying individually compare to studying with my peers in the same course.	○	○	○	○	○

Q18 Please select the response that best describes how you feel.

	Very untrue of me	Somewhat untrue of me	Neutral	Somewhat true of me	Very true of me
Most of the time, I prefer being a leader rather than a follower when working in a group.	○	○	○	○	○
I am the type of person who is not interested to lead others.	○	○	○	○	○
I believe I can contribute more to a group if I am a follower rather than a leader.	○	○	○	○	○
I have a tendency to take charge in most groups or teams that I work in.	○	○	○	○	○

Q19 The following questions ask you about your general stress levels with regard to various aspects of university life.
Please select the response that best describes how you feel.

	None of the time	A little of the time	Some of the time	Most of the time	All of the time
Academic / coursework demands	○	○	○	○	○
Study / life balance	○	○	○	○	○
Staying on top of coursework	○	○	○	○	○
University environment	○	○	○	○	○
Daily responsibilities of university life	○	○	○	○	○
Friendships	○	○	○	○	○
Please select "Most of the time" for this particular question	○	○	○	○	○
Family relationships	○	○	○	○	○
Romantic relationships	○	○	○	○	○
Relationship break-down	○	○	○	○	○
Discrimination	○	○	○	○	○
Language / cultural issues	○	○	○	○	○
Sexual orientation issues	○	○	○	○	○

Appendix B2 Post-course survey

Post-Course Survey (2023 Winter)

Q0 Welcome to the COURSE X Post-Course Survey. This survey is part of a study exploring strategies to increase student success in online courses. It will help us understand your feelings about this course and learning approaches.

Completing this survey will give you one extra credit point (each extra credit point will add 2.5% to your final exam score). The survey should only take you around 10 minutes to complete. Your personal information will be kept confidential and not shared with the instructor. If you are willing to participate, please log in with your UCI email address to complete the survey by **Sunday, March 19th at 11:59 PM**.
Alternatively, if you choose not to complete the survey, you will have the opportunity to earn the extra credit point by completing an alternative essay assignment (3-4 paragraphs each) by Sunday, March 19th at 11:59 PM.

If you have any questions or comments and/or wish to complete the alternative essay, please contact the peer assistant **XunFei Li (xunfeil@uci.edu)**.

Q0 From weeks 2 to 10, you were assigned a study group, studied with your peers, and completed study group assignments. We would like to hear about your personal experiences and feelings regarding your time in the study group:

Q1 The following questions ask about your involvement in the study group activities. Please select the response that best describes your behavior in the study group.

	Never	Rarely	Sometimes	Frequently	Very Frequently
I spent time coordinating the study group activities.	○	○	○	○	○
I helped keep the study group members engaged and participating in productive dialogue.	○	○	○	○	○
I helped keep the study group members on task.	○	○	○	○	○

Q2 The following questions ask about the group process of your study group. Please select the response that best describes your group.

	strongly disagree	disagree	neutral	agree	strongly agree
My group has a clear understanding of the study group purpose.	○	○	○	○	○
My group is effective in scheduling the study group activities.	○	○	○	○	○
Study group meetings are very productive towards achieving our goals.	○	○	○	○	○
My group works with a great deal of flexibility so that we can adapt to changing or different needs.	○	○	○	○	○

Q3 The following questions ask about the relationships among your group members. Please select the response that best describes your group.

	strongly disagree	disagree	neutral	agree	strongly agree
Group members do not connect very well with one another.	○	○	○	○	○
Group members are effective listeners.	○	○	○	○	○
Group members are able to work through differences of opinion without damaging relationships.	○	○	○	○	○
Group members display high levels of cooperation and mutual support.	○	○	○	○	○
Everyone values what each member contributes to the study group.	○	○	○	○	○

Q4 The following questions ask about the academic and social support you get from the study group experience. Please select the response that best describes your feelings.

	strongly disagree	disagree	neutral	agree	strongly agree
The study group helped keep me on task in a way that helped me to learn.	○	○	○	○	○
Study group members provided feedback that helped me understand my strengths and weaknesses relative to the course's goals and objectives.	○	○	○	○	○
Study group discussions increased my interest in course issues.	○	○	○	○	○
I feel connected with study group members.	○	○	○	○	○
When I am with my study group, I feel included.	○	○	○	○	○
I feel as if study group members do not care about me.	○	○	○	○	○

Q5 How satisfied are you with the study group experience this quarter?

- ○ Very Satisfied
- ○ Satisfied
- ○ Neutral
- ○ Dissatisfied
- ○ Very Dissatisfied

Q6 How much fun is learning Chemistry to you?

- ○ None at all
- ○ A little
- ○ A moderate amount
- ○ A lot
- ○ A great deal

Q7 How important is performing well in this course (COURSE X) to you?

- ○ Not at all important
- ○ Slightly important
- ○ Moderately important
- ○ Very important
- ○ Extremely important

Q8 How much does being good at the material taught in this course (COURSE X) mean to you for your life?

- ○ Not at all
- ○ A little
- ○ A moderate amount
- ○ A lot
- ○ A great deal

Q9-1 Friendship Network Questions How many **friends** do you have in this course (COURSE X)?

- ○ 0
- ○ 1-3
- ○ 4-6
- ○ 7-9
- ○ 10 or more

Q9-2 Please select students in this course (COURSE X) from the dropdown list that you consider your **friends**.
Note: You can easily locate your friend by typing their first name.

Friend 1	▼
Friend 2	▼
Friend 3	▼
Friend 4	▼
Friend 5	▼
Friend 6	▼
Friend 7	▼
Friend 8	▼
Friend 9	▼
Friend 10	▼

Q9-3 If you have more than 10 **friends**, or friends in this course (COURSE X) that you cannot find their name, please do your best to list their first and last names here.

_____

Q9-4 If you did not list any names in the previous question, please tell us why.

- ○ I already listed their name.
- ○ I accidentally skipped the question. (If so, you can go back to answer it)
- ○ I have friends in this course (COURSE X), but I do not know their names.
- ○ I have friends in this course (COURSE X), but I do not feel comfortable giving their names.
- ○ I listed their names manually.

Q10-1 Help-Seeking Network Questions How many students do you have in this course (COURSE X) you would **ask for advice or help**?

- ○ 0
- ○ 1-3
- ○ 4-6
- ○ 7-9
- ○ 10 or more

Q10-2 Please select students in this course (COURSE X) from the dropdown list that you would **ask for advice or help**.
Note: You can easily locate the student by typing their first name.

Person 1	▼
Person 2	▼
Person 3	▼
Person 4	▼
Person 5	▼
Person 6	▼
Person 7	▼
Person 8	▼
Person 9	▼
Person 10	▼

Q10-3 If you have more than 10 students that you would **ask for advice or help**, or students in this course (COURSE X) that you cannot find their name, please do your best to list their first and last names here.

_____

Q10-4 If you did not list any names in the previous question, please tell us why.

- ○ I already listed their name.
- ○ I accidentally skipped the question. (If so, you can go back to answer it)
- ○ I have students that I can ask for advice or help in this course (COURSE X), but I do not know their names.
- ○ I have students that I can ask for advice or help in this course (COURSE X), but I do not feel comfortable giving their names.
- ○ I listed their names manually.

Q00 Below, we would like to know your feelings and experience in COURSE X throughout the quarter.

Q11 Please indicate how true each statement is for you.

	Very untrue of me	Somewhat untrue of me	Neutral	Somewhat true of me	Very true of me
I felt that students in this course cared about each other.	○	○	○	○	○
I felt connected to others in this course.	○	○	○	○	○
I felt a spirit of community.	○	○	○	○	○
I felt isolated in this course.	○	○	○	○	○
I felt confident that others would support me.	○	○	○	○	○

Q12 Please indicate how true each statement is for you.

	Very untrue of me	Somewhat untrue of me	Neutral	Somewhat true of me	Very true of me
I made good use of my study time for COURSE X.	○	○	○	○	○
I kept a record of what my COURSE X assignments were and when they were due.	○	○	○	○	○
I made sure I kept up with the weekly readings and assignments for COURSE X.	○	○	○	○	○
I usually plan my work in advance so that I can turn in my COURSE X assignments on time.	○	○	○	○	○
I worked hard to do well in COURSE X even if I didn't like what I was doing.	○	○	○	○	○
I often felt so lazy or bored when I studied that I quit before I finished what I planned to do.	○	○	○	○	○
When I did not fully understand something, I asked other classmates in COURSE X for ideas.	○	○	○	○	○
I shared my problems with my classmates in COURSE X so we knew what we were struggling with and how to solve our problems.	○	○	○	○	○

Q18 How satisfied are you with the COURSE X course experience?

- ○ Very Satisfied
- ○ Satisfied
- ○ Neutral
- ○ Dissatisfied
- ○ Very Dissatisfied

Appendix B3 Example of weekly survey

Study Group Homework 4

Q0
Study Group Homework 4
What do you need to do?
To get the full credit, you need to:
(1) Answer questions about your study group experience you completed for *Study Group Homework 3*.
(2) Make a plan to study with your group member(s) during Week 5. (By "Week 5", we mean anytime from Wednesday, Feb 8 to Tuesday, Feb 14.)

How to find my group members?
(1) Here are your group members in Group ${e://Field/group_id_new}:
${e://Field/recipientfirstname1} ${e://Field/recipientlastname1} ${e://Field/recipientemail1}
${e://Field/recipientfirstname2} ${e://Field/recipientlastname2} ${e://Field/recipientemail2}
${e://Field/recipientfirstname3} ${e://Field/recipientlastname3} ${e://Field/recipientemail3}
${e://Field/recipientfirstname4} ${e://Field/recipientlastname4} ${e://Field/recipientemail4}

(2) Here is the link to your Discord group chat (if it doesn't work after you click it, try to copy and paste it to your browser):
${e://Field/discord}">${e://Field/discord}

(3) Here is the Zoom link for your group to meet virtually:
${e://Field/zoom}">${e://Field/zoom}

Due Date & Grades
Please complete this homework by **Tuesday, Feb 7 at 11:59 PM.**
For questions about this assignment, contact peer assistant: xunfeil@uci.edu

Q1 **Please answer the questions below about your study group experience you completed for Study Group Homework 3 (6 points)**

Q2 Did your group study together during Week 4 (from Wednesday, Feb 1 to Tuesday, Feb 7)? (3 points)
It is ok if it is not the same time period as you planned.

- ○ No, we didn't (0 point)
- ○ Yes, we just met with each other briefly (1 point)
- ○ Yes, we studied for one hour (2 points)

Q2.1 Please upload a **screenshot** of your virtual meeting as proof. You could provide other proof if your group met in person.
Here are some examples:

Zoom Discord

Q2.2 Did your group meet virtually or in person?

○ meet virtually

○ in person

Q2.3 What did your group do during the Week 4 meeting?
You can get full points for this question as long as you answer this question.

☐ Watching lecture videos/Reading course materials

☐ Doing homework

☐ Having discussions

☐ Others _____

Q3 In the past seven days, how many days did you ever contact your **study group peers** from this class to discuss **course-related content**? (1 point)
You can get full points for this question as long as you answer this question.

○ 0 (1 point)

○ 1 - 2 days (1 point)

○ 3 -4 days (1 point)

○ 5 -6 days (1 point)

○ Everyday (1 point)

Q4 In the past seven days, how many days did you ever contact your **study group peers** from this class to discuss **non-course-related content** (e.g., chit-chat, student organizations, campus life)? (1 point)
You can get full points for this question as long as you answer this question.

○ 0 (1 point)

○ 1 - 2 days (1 point)

○ 3 -4 days (1 point)

○ 5 - 6 days (1 point)

○ Everyday (1 point)

Q5 In the past seven days, how many hours in total did you spend studying with your **study group peers** from this class? (1 point)
You can get full points for this question as long as you answer this question.

- ○ 0 - 1 hour (1 point)
- ○ 1 - 3 hours (1 point)
- ○ 3 - 6 hours (1 point)
- ○ 6 - 10 hours (1 point)
- ○ More than 10 hours (1 point)

Q6 **Please answer the questions below to make a plan with your group members to study together for one uninterrupted hour during week 5 (from Wednesday, Feb 8 to Tuesday, Feb 14) (4 points)**

Q7 Did your group schedule a time to study together during Week 5? (2 points)
It is ok if your plan changes during the week; You can get full points as long as your group scheduled a meeting time.

- ○ No, we haven't scheduled a meeting time and will come back later (0 point)
- ○ Yes, we scheduled a meeting time (2 points)

Q7.1 What time do your group plan to study together during Week 5? (e.g., Monday 8-9 PM)
It is ok if your plan changes during the week.

_____

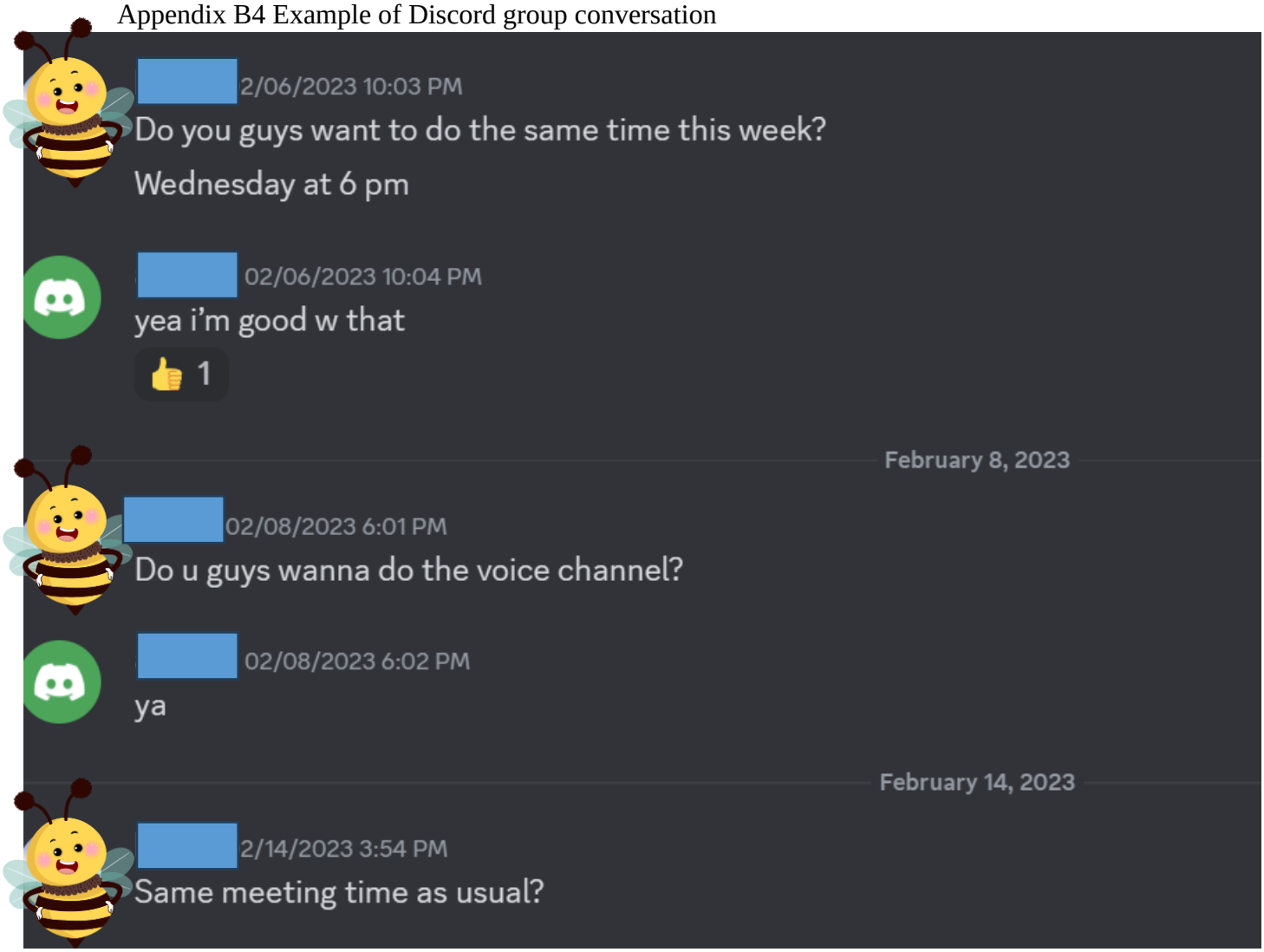
Q8 What do you plan to do during this time period? (2 points)
Your plan can be the same or different from the other group members' plans.

- ○ We didn't make a plan (0 point)
- ○ Watching lecture videos (2 points)
- ○ Doing homework (2 points)
- ○ Having discussions (2 points)
- ○ Others (2 points) _____

Q10 **Reminder for next week's Study Group Homework 4:**

Your group needs to study for one uninterrupted hour together in order to earn the credit. In the Study Group homework 4 next week, you will also be asked to upload a screenshot/picture of your virtual/in person meeting as a proof (e.g., Zoom, Discord). Here are some examples：

Appendix B4 Example of Discord group conversation



CHAPTER 3 FAST TRACK TO SUCCESS?: INVESTIGATING CONDENSED COURSE FORMATS AND STUDENT ACHIEVEMENT IN COMMUNITY COLLEGES

INTRODUCTION

Community colleges are vital in offering accessible postsecondary education in the U.S., with a substantial portion of the enrollment consisting of adult learners and working students. In the fall 2021, 43% of the students enrolled for-credit in U.S. community colleges were aged 22 or older, and nearly three-quarters of students enrolled at community colleges were balancing employment with their studies (AACC, 2023). Yet, juggling work and family responsibilities alongside academic commitments may lead to unique scheduling challenges (Bahr et al., 2021), which underscores the need for effective policies tailored to the specific needs of this population.

Condensed course formats, which compress courses to about half the length of traditional semesters, have emerged as a promising solution to address these challenges.[11] By compressing course durations while maintaining the same credit hours and content, this approach offers an accelerated pathway for students to earn college credits. Additionally, it enhances flexibility by providing multiple entry points within a term for students to start their courses. This format can be particularly appealing to students with significant family and work responsibilities, such as adult learners and working students, enabling them to complete degree requirements in a more flexible and potentially cost-effective manner (Austin & Gustafson, 2006; Brenner, 2024; Daniel, 2015; Stith, 2023).

[11] Condensed course formats, also known in the literature as short-course or compressed courses, cover the same content and credit hours as traditional semester-length courses but are completed typically within 10 weeks (Daniel, 2000; Walsh et al. 2019).

Condensed courses, with their ability to accommodate varied student schedules and accelerate academic progress, have drawn the attention of educators, college administrators, and policymakers nationwide. In particular, condensed courses have been nominated as a potential way to address the significant enrollment declines community colleges faced post-pandemic (Tamburin, 2021). However, despite the increasing popularity of and high hopes around condensed courses, concerns have arisen regarding their potential to intensify students' stress levels. The condensed timeline and increased workload may limit opportunities for thorough comprehension and reflection, potentially increasing the risk of burnout and detracting from the overall educational experience (Anastasi, 2007; Holzweiss et al., 2019; Lutes & Davies, 2018).

The heightened policy attention and conflicting mechanisms surrounding the condensed courses highlight the need for a thorough examination of its impact on students' academic outcomes. Despite a burgeoning research literature in education on this topic, much of the existing studies are descriptive in nature and narrow in scope, often concentrating on a limited selection of courses within one subject area at a single institution. (e.g. Austin & Gustafson, 2006; Harlow et al., 2015). Consequently, the available quantitative evidence on the causal impact of condensed courses on student academic outcomes is limited, particularly for large-scale field settings. Furthermore, existing studies almost exclusively focus on describing the immediate outcomes of the current course, overlooking students' enrollment patterns and performance in subsequent courses. This limitation is critical for a holistic understanding of the condensed courses's impact, considering that concurrent course grades may be influenced by other factors such as grading leniency and thus may not accurately reflect students' learning outcomes. Our study aims to address this research gap by examining the impact of condensed courses using college administrative data from the Virginia Community College System

(VCCS). Specifically, using data on 179,146 students enrolled at any of the 23 institutions at VCCS between the 2010-2011 and 2014-2015 academic years, we examine the effect of taking condensed courses during a student's initial semester in college on both current and downstream outcomes. To minimize bias from student self-selection into condensed courses, we employ a two-way fixed effects model, which controls for both student-fixed effects and college-by-course fixed effects, therefore addressing biases arising from unobserved student characteristics as long as they are consistent within an individual. We also focus exclusively on the first course students took in their first semester to mitigate student selection into courses with different formats. For subsequent course performance, we further include next-class fixed effects into the model to enable comparisons between students enrolled in exactly the same next class after their initial exposure to a subject area in their first semester.

Our results indicate that condensed courses have significant positive impacts on students' current course outcomes, including course persistence, completion rates, and grades. This positive impact extends to key downstream outcomes, where students are more likely to enroll in and pass subsequent courses within the same subjects. These findings add valuable insights to the ongoing conversation surrounding educational innovation by offering robust causal evidence of the effectiveness of condensed courses. To our knowledge, our research represents the first large-scale study to provide causal evidence on the impact of condensed courses across multiple institutions and diverse subjects. Our results underscore the capacity of condensed courses to improve immediate and downstream academic outcomes.

LITERATURE REVIEW

Why Condensed Courses Are Offered and Their Prevalence

Condensed courses have been discussed in U.S. higher education for several decades (Daniel, 2015; Mena-Guacas et al., 2023; Scott & Conrad, 1992). By design, they deliver the same content as traditional-length courses but within a compressed timeframe and can be offered during both summer sessions and regular academic terms. Institutions have offered condensed courses to increase enrollment and address the needs of students with limited time and complex schedules, particularly adult learners and working individuals (Austin & Gustafson, 2006; Brenner, 2024; Daniel, 2015; Stith, 2023). In recent years, community colleges have experienced a substantial decrease in student enrollment (NSCRC, 2024). As a response, condensed courses are increasingly being considered as a strategy to boost enrollment by potentially reducing both the opportunity and financial costs associated with college-level courses (e.g., Tamburin, 2021). Besides boosting enrollment, condensed courses have also been used to improve students' success rates in gateway courses. The developmental education reforms often include an acceleration component that compresses the curriculum and expedites students' progress toward degree attainments and transferring to four-year universities (Barnett et al., 2012; Cho et al., 2012; Guy et al., 2015; Hayward et al., 2014; Jaggars et al., 2015).

Condensed courses are primarily available during summer sessions but are also offered during regular academic terms. Individual colleges usually schedule them based on the junctures of their academic calendar. For example, students can take either the first or the second 8-week session in a traditional semester term or an 8-week session during the summer term. This format is applied across multiple subjects and is accessible in both four-year and two-year colleges. There is limited knowledge about how the course design of condensed courses systematically differs from that of traditional-length courses (Mena-Guacas et al., 2023). From the students'

perspective, these courses appear to be more time-intensive. Although the overall course design remains similar, some studies suggest that condensed courses may have fewer requirements, such as reduced numbers of exams and assignments (e.g., Lutes & Davies, 2018).

Impact of condensed courses on student academic outcomes

Over the past two decades, there has been consistent research interest in the impact of condensed courses—alternatively referred to as compressed, short-term, accelerated, or time-shortened courses—on students' academic success in higher education at both four-year and two-year institutions (Daniel, 2015; Holzweiss et al., 2019; Walsh et al., 2019). In general, the literature has found a positive impact of condensed courses on students' current course performance, primarily measured by course grades; however, this is subject to both sample restrictions and methodological limitations.

Scope limitations of prior studies on summer-term and single-college

Although these studies span a diverse context from public universities, private universities, and community colleges, the majority of the literature focuses exclusively on condensed courses offered within individual colleges (Austin & Gustafson, 2006; Brenner, 2024; Guillory, 2018; Paasch, 2020; Walsh et al., 2019) or specific subjects and courses (Anastasi, 2007; Guillory, 2018; Harlow et al., 2015), with a few exceptions (Bishop et al., 2018; Stith, 2023). For example, Austin and Gustafson examined the impact of the condensed courses on students' course performance using four courses at a public university and found a positive impact. Harlow et al. compared student performance between 6-week condensed courses and 12-week courses in physics and found a negative impact of condensed courses. Although focusing on a single institution and subject may help control for contextual differences and variations in student populations, it also hinders the generalizability of the findings. This limitation is critical

for policy implications, as course offerings and instruction vary widely across institutions and subjects.

Moreover, most studies predominantly examine courses provided during the summer period (Anastasi, 2007; Austin & Gustafson, 2006; Guillory, 2018; Harlow et al., 2015; Walsh et al., 2019). with only a few investigating the condensed courses during the traditional term (Bishop et al., 2018; Brenner, 2024; Holzweiss et al., 2019). While summer is a common period for offering condensed courses, students who choose these courses, as well as the learning environments, might systematically differ from those in traditional terms (Walsh et al., 2019). Such differences could potentially bias the estimation of the impact of condensed courses and may not offer complementary policy insights for courses offered during traditional terms. Some prior studies have controlled for commonly observed student characteristics, course titles, and instructors. However, these controls are unsatisfactory in addressing student selection bias in summer courses and other course features such as motivation and course size.

Methodological limitations: threats of student sorting

Methodologically, no study has applied a rigorous quasi-experimental design to explore the impact of condensed courses.[12] This may stem from the lack of a large-scale, comprehensive dataset that would allow researchers to implement more rigorous models. Existing studies employed Chi-squared tests, ANOVA, or regression analysis to compare student performance in condensed courses with that in traditional-length courses, either without controlling or only

[12] Bostwick et al. (2022) conducted an event study analysis and found that switching from the quarter system to the semester system negatively impacts student on-time graduation rates. Brenner (2024) discovered that transitioning from a 16-week to an 8-week calendar, student achievement gaps were reduced within a two-year institution. However, these studies compare two different academic calendar systems where financial support and instruction may systematically differ across years or institutions. We did not include this strand of studies in our review.

partially controlling for students' self-selection in condensed courses and differences in course-section characteristics such as classroom size and instructor attributes. Overall, a significant proportion of prior studies reported a positive relationship between enrolling in condensed courses and students' current course grades compared with those who enrolled in traditional-length courses (Anastasi, 2007; Austin & Gustafson, 2006; Ferguson & DeFelice, 2010; Stith, 2023; Walsh et al., 2019). Few studies found a negative correlation between taking condensed courses and student performance (Harlow et al., 2015). One study found no correlation between course length and students' current course outcomes, utilizing course-specific pretest and posttest (Walsh et al., 2019). A more detailed summary of the relationships between course length and student academic outcomes can be found in Walsh et al., (2019).

The methodological limitation leads to the threat of students selecting into condensed courses and the concern of classroom level differences between traditional length courses and condensed courses which will also impact students' outcomes (e.g., motivation, classroom size, instructor employment status). Existing empirical studies evaluating the impact of condensed courses use different student controls to mitigate students' selection bias. For example, Walsh et al. (2019) controlled for gender, prior GPA, and citizenship and found that students received 0.35 points higher on a 4-point grading scale in a summer 6-week course when compared with students who took the same course in a 14-week semester.

Some studies have effectively controlled for course differences by focusing only on a small set of courses within a subject at the same institution, using the same instructors and course components. However, these studies face limitations due to small sample sizes and unclear sorting mechanisms regarding students who choose condensed courses versus traditional-length courses. For example, Anastasi (2007) found that although students attending summer condensed

courses had lower academic preparedness, they achieved higher course grades compared with students taking traditional-length courses. This comparison used matched psychology courses to ensure consistency in instructors and content. Similarly, Walsh et al. (2019) conducted course-specific pretests and found no statistical difference in students' posttests across four courses at a private university.

Beyond current course performance

Current course performance may not accurately reflect true learning, as instructors might have different expectations for students in condensed courses or may grade more leniently. It is important to look at other academic performance indicators beyond current course grades. Guillory (2018) found no difference in retention rates between 5-week and 16-week courses. Few studies have examined the impact of condensed courses on students' subsequent outcomes, with those that did generally finding no significant effect on performance in subsequent courses (Austin & Gustafson, 2006; Bishop et al., 2018).

Potential Mechanisms of the impact of condensed courses

Prior studies using surveys and interviews provide insights into the potential mechanisms explaining the impact of condensed courses on student academic outcomes, though none of these studies employ rigorous causal research designs. The existing findings on how course structures differ between condensed courses and traditional-length courses are mixed. Lutes and Davies (2018) found that students in traditional-length courses, on average, spend 21 minutes more per credit per week than those taking condensed courses, according to a student survey. They also reported that instructors might skip certain topics that are not covered in exams. However, some studies have reported that students feel more challenged by the workload in condensed courses and find these courses more demanding (Anastasi, 2007; Holzweiss et al., 2019).

Generally, students reported positive perceptions of condensed courses (Paasch, 2020; Walsh et al., 2019; White, 2023). Students found it less challenging to retain knowledge for shorter periods and maintained more consistent engagement with the course, both of which may help them perform well on exams (Walsh et al., 2019). Students also reported a sense of accomplishment and stronger personal responsibility in condensed courses, which may contribute to their positive perceptions (White, 2023). At the same time, students feel the need for better time management and awareness of the more intense deadlines in condensed courses (Holzweiss et al., 2019).

However, since most studies compare students in condensed courses offered during summer terms with those enrolled in traditional semesters, the perception could be influenced by the unique summer session learning environment, where students tend to take fewer courses and may be able to focus better, instead of reflecting perceptions solely from condensed courses (Walsh et al., 2019). Moreover, suggestive studies showing instructors skipping content not related to exams, and students perceiving a lower workload in condensed courses, partially verify educators' concerns about the quality of learning that may not be observed through current course components. This necessitates a look beyond current course outcomes.

BACKGROUND AND DATA

Data and Institutional Context

We use administrative data from five cohorts of first-time-in-college students within the Virginia Community College System (VCCS), spanning the academic years from 2010-2011 to 2014-2015. For each cohort, we track outcomes for six years following students' initial enrollment. The administrative dataset encompasses individual-level student information, including demographic details such as sex, race/ethnicity, and age. It also includes indicators for

first-generation college student status and eligibility for Pell Grants. Additionally, the dataset incorporates transcript data, capturing student course enrollments and corresponding academic outcomes. Course-section level attributes are provided, including class duration, instructional modality (online versus face-to-face), and instructor characteristics like gender, race/ethnicity, and employment status. Moreover, we link this data with National Student Clearinghouse (NSC) data to monitor students' transfer and degree completion outcomes. Furthermore, we supplement our dataset with unemployment insurance (UI) data, offering insights into individuals' seasonal income and employment status.

VCCS includes twenty-three individual colleges of varying sizes, located in both rural and urban areas. Table 3.1 provides the institutional characteristics of VCCS, compared with the characteristics of the national sample of all public two-year institutions in fall 2010 using the data from the Integrated Postsecondary Education Data System (IPEDS). The comparison indicates that VCCS serves a larger proportion of Black students. Additionally, more than half of the students receive financial aid.

Condensed course at VCCS

VCCS operates on a semester system, where semester-length courses run for 15/16 weeks. However, they also offer condensed courses within or between semesters. For instance, some colleges offer a "First 8-Week Course" and a "Second 8-Week Course" alongside the standard "16-Week Course". Within a particular semester, colleges may offer condensed courses to adapt to the needs of non-traditional students or to increase enrollment, and each college has the autonomy to decide how to incorporate condensed courses. Instructors can also choose to offer condensed courses, subject to prior approval from higher authorities like deans or department heads. Once a course is offered, Students can find information about course lengths

on the college website. College advisors may offer suggestions if a student enrolls in too many courses, reminding them of the potential workload. However, based on conversations with administrators at VCCS, there is no evidence indicating a consistent, systematic approach to either prohibiting or encouraging students to enroll in condensed courses.

Figure 3.1 shows the distribution of course durations spanning from the 2010-2011 to the 2014-2015 academic years for college-level content-based courses (i.e. excluding developmental coursework and student success courses). We define condensed courses as course sections lasting 10 weeks or fewer, based on the existing literature as well as the observed distribution of course lengths across VCCS colleges. Roughly twenty-five percent of course sections fall into condensed courses, with the majority being either 8-week or 10-week courses. Appendix Figure C1 provides more detailed information regarding the distribution of condensed courses, where about two-thirds of the condensed courses are 8-week and 10-week, followed by approximately 10 percent for 5-week and 6-week courses respectively.

Condensed courses are widely adopted throughout VCCS, spanning academic years, colleges, and fields of study. From the 2010-2011 to 2014-2015 academic years, about a quarter of courses at VCCS were offered in this format. Figure 3.2 shows that all major subject areas have a significant portion of courses in short formats, with Social Sciences and Applied subjects having the highest at around thirty percent. In contrast, Humanities and STEM (Science, Technology, Engineering, and Mathematics) subjects have relatively lower proportions, each with about twenty percent. Appendix Figure C2 provides a breakdown of more detailed subjects. Furthermore, every college within the VCCS offers condensed courses, though the extent of their offering varies, ranging from around ten percent to thirty percent. This between-college variation

is illustrated in Appendix Figure C3, which presents a bar graph of the distribution of condensed courses across the colleges.

Analytic sample and summary statistics

To assess the impact of condensed course format on student academic outcomes, we focus on each student's initial enrollment in a subject during their first term in college. These early course choices are less likely to be influenced by previous academic performance or adjustments in study habits that might occur after the students have adapted to the college environment and gained more insight into the demands of different course formats. Accordingly, this sample selection criteria helps to mitigate the potential bias of students' strategic course selection decisions based on their prior experiences or expectations. In cases where students are enrolled in multiple courses within the same subject during their first term, we randomly select one course for analysis. Similarly, if students are enrolled in several courses across subsequent terms at the same time, we select one course from the earliest of these subsequent terms for inclusion in our analysis on downstream outcomes.

The analytic sample includes 447,573 initial course enrollments among 179,146 unique students. Descriptive statistics at the student level are shown in Table 3.2. Half of the sample comprises female students, with an average age at entry term of 22 years old. The sample is predominantly White students (56%), followed by Black students (23%), Hispanic students (10%), and Asian students (5%). Nearly half are Pell Grant eligible (49%), while 37% are first-generation students. High school completion is prevalent (97%), with three-quarters of students intending to pursue an Associate's degree (76%). 72% of students' first enrolled term is in Fall, with 55% enrolled full-time.

Table 3.3 provides summary statistics at the course-section level, comparing condensed course formats with traditional-length course formats, based on aggregated student-by-course-section data. Panel A of Table 3.3 focuses on course-section characteristics, such as course modality, class size, course schedule (weekend course, daytime course), and course credits. Notably, online sections are more common in condensed courses, accounting for 26% compared with 17% in traditional-length courses. Furthermore, condensed course formats have a higher proportion of weekend classes, with 4% compared with only 1% in traditional-length courses. In terms of instructor demographics, Black instructors are slightly more represented in condensed courses (14%) than in traditional-length courses (12%), and there is a lower proportion of full-time instructors in condensed courses (39%) as opposed to traditional-length courses (45%).

Panel B of Table 3.3 sheds light on student performance in initial courses. Students in condensed courses appear to outperform those in traditional-length courses. Specifically, 93% of students in condensed courses persist until the end of the course, compared with 89% in traditional courses. Additionally, 77% of students in condensed courses achieve a grade of C or higher, which exceeds the 66% pass rate observed in traditional-length courses. Moreover, students in condensed courses also attain higher average grades, with a mean grade of 2.7 compared with 2.2 in traditional-length courses.

Finally, Panel C of Table 3.3 shows downstream outcomes, including subsequent course enrollment in the same field of study and their performance in these subsequent classes. It appears that 42% of students initially enrolled in condensed courses during their initial exposure to a subject area go on to enroll in subsequent courses within the same subject area, which is a higher rate compared with the 35% observed for students in traditional-length courses. This suggests that the condensed courses may be more effective at encouraging continued study in the

same field. Additionally, among all students who do enroll in subsequent courses within the subject, both the persistence rate and the course passing rate are similar, regardless of whether their initial courses were in short-course or traditional-length formats. Yet, these descriptive gaps might be partially driven by students' systematic sorting into the condensed course formats, which we address in the next section.

EMPIRICAL STRATEGY

Empirical model for concurrent course outcomes

Evaluating the impact of condensed courses with observational data presents multiple challenges. First, students who enroll in condensed courses may have distinct characteristics that are associated with their academic outcomes. For instance, as shown in Table 3.4, students with substantial family responsibilities, including adult learners and part-time students, are more likely to enroll in condensed courses. Alongside student-level sorting, the availability of condensed courses also varies across subjects and colleges as shown in Figure 3.2 and Figure C3. To address these concerns, we incorporate student fixed effects into the model to account for individual-specific preferences that are constant across different subject areas, as well as course-by-college fixed effects that enable us to compare students enrolled in the same course at the same college but in different course sections. Specifically, the model writes as follows: $Y_{iscjt} = \alpha + \beta Short_{iscjt} + \tau_i + \gamma_{cj} + \rho_t + X_{scjt} + \mu_{iscjt}$ [1]

which relates student i's outcome in section s of course c at college j during time t to the format of the course that the student had during their initial exposure to this subject area in their first semester enrolled at VCCS ($Short_{iscjt}$, and is equal to one if the course is condensed and is zero if it is traditional-length). τ_i indicates student fixed effects to account for constant student preferences for condensed courses. ρ_t is time fixed effects, and γ_{cj} represents college-by-course

fixed effects, thus enabling comparisons among different sections of the same course taught within a particular college while controlling for overall over-time variations and possible shocks that might impact student grades.

While college-by-course fixed effects control for between-course variations, disparities may still exist across course sections within the same college and course, potentially affecting student outcomes. These disparities include course-section features such as class modality and instructor characteristics like employment status and teaching experience. For instance, descriptive statistics presented in Table 3.3 reveal that condensed classes are more frequently delivered online and taught by part-time instructors. However, both the online delivery format and instructors' employment status may affect student performance outcomes (e.g. Xu & Jaggars, 2013; Xu, 2019). To mitigate the influence of these factors, we control for course-section-level and instructor-level confounders in the model (denoted as X_{scjt}), including the mode of delivery, class size, class schedule, as well as instructor's gender, race/ethnicity, employment status (full-time versus part-time), and teaching experience.

Empirical model for downstream outcomes

Students' experience with condensed courses in their initial course within a subject area may affect their subsequent course enrollment behaviors and performance in multiple ways. To capture downstream impacts, we construct three measures using a three-year time window following students' initial enrollment: 1) whether a student enrolls in a subsequent course within the same subject area after their initial exposure, 2) whether a student enrolls and also completes a subsequent course within the subject area, and 3) conditional on enrollment, their academic performance in the next class in the same subject area. For the first two outcomes, we evaluate the impact of condensed courses using the fixed effects models described in equation [1]. To

assess the impact on subsequent course performance conditional on course enrollment (the third outcome), we use a separate model that further includes next-class fixed effect to control for students' strategic selection of subsequent classes based on their initial experience:

$$Y_{is+1c+1jt+1} = \alpha + \beta Short_{iscjt} + \tau_i + \gamma_{cj} + \rho_t + S_{s+1c+1jt+1} + X_{scjt} + \mu_{is+1c+1jt+1}$$

[2]

which relates student i's outcome in the next class section s+1 of the subsequent course c+1 at college j during time t+1 to the format of the course that the student had during their initial exposure to this subject area ($Short_{iscjt}$, and is equal to one if the course is condensed). Equation [2] includes all variables from equation [1], but further includes a new term $S_{s+1c+1jt+1}$, which incorporates information about the college, course, time, and specific section of the next class a student chooses to enroll in after initial exposure to a subject area. As a result, this model specification compares student performance in exactly the same subsequent course section. By doing so, it controls for the potential influence of a student's initial experience in a field on their choice of the next course, as well as their preference for condensed course format subsequently.

RESULTS

Current course outcomes

Table 3.5 presents the estimated effects of condensed course format on the performance of a student's first course within a field of study based on five outcome measures: (i) whether students persist to the end of the course (versus withdrawing), which intends to offer insights into retention and engagement; (ii) whether students pass the course with a grade of D or higher (versus failing or withdrawing), which aims to provide a basic benchmark for academic success and continuation; (iii) whether students achieve a grade of C or higher—as opposed to earning a

D, failing, or withdrawing—which distinguishes between minimal achievement and the more robust performance often required for degree attainment; (iv) students' grades, ranging from 0 to 4, which allow for a more nuanced evaluation of course performance; and (v) whether students retake the course later.

Overall, the findings indicate a positive impact of enrolling in a condensed format for an initial course on students' outcomes in current courses compared with those enrolled in traditional-length courses. Specifically, students who enrolled in condensed course formats experienced a 0.7 percentage point increase in course persistence, a 2.1 percentage point increase in passing the course with a grade of D or higher, a 3.1 percentage point increase in passing the course with a grade of C or higher, an improvement of 0.1 points in their course grade, and a decrease of retaking the course by 1 percentage points, compared with those enrolled in traditional-length courses.

To provide context for these coefficients, Table 3.5 includes baseline averages for each evaluated outcome, which serve as reference points to better understand the magnitude of the observed changes. For example, the baseline average for students achieving a grade of C or higher across all course sections is 68%. With a 3 percentage point increase observed in condensed courses, the number would improve to 71%. This change translates to a 4% relative improvement in the number of students achieving a grade of C or higher. To put this into perspective, consider a college with 1,000 first-time-in-college (FTIC) students each year, with each student taking one course. The adoption of condensed courses could prevent an additional 40 students from receiving a grade of D or failing.

Downstream outcomes

Table 3.6 presents results of taking a condensed course format during students' initial exposure to a field of study on the probability of taking another course in the same field, and their chances of both enrolling in and passing these courses within three years after their initial enrollment.[13] The results consistently demonstrate a positive influence. Students who initially enrolled in a condensed course format within a subject, on average, experienced a 3.2 percentage point increase in the likelihood of enrolling in subsequent courses. The size of the effects is fairly similar for both enrolling in and passing subsequent courses with a grade of D or higher (3.1 percentage point increase) and for enrolling in and passing with a grade of C or higher (2.9 percentage point increase). Given that only 41% of students typically enroll in subsequent courses within a subject, the impact of the condensed course format is notably substantial, boosting enrollment rates by 7%. The consistency across these outcomes—enrollment, passing with a grade of D or higher, and achieving a grade of C or higher—suggests that the positive effects on downstream outcomes are primarily driven by the increased chance of taking another course in the same field of study.

As mentioned previously, assessing the impact of the initial condensed course format on students' later academic performance is challenging due to changes in students' enrollment preferences after their first course experience. For example, if students who start with a condensed course format are more likely to choose easier or more challenging courses subsequently based on their initial experiences in a given field, this could lead to bias in our estimation. To directly provide insights into this possibility, we analyze whether students who

[13] Subsequent enrollment can occur in two ways: either by enrolling in courses during the terms that follow, or by choosing to take another condensed course later in the same term after completing an initial condensed course earlier in that term.

initially enrolled in condensed courses are more likely to choose courses with historically higher average grades, which are calculated from records from the same course spanning the prior three years at the same institution. Results presented in Column (4) of Table 3.6 indicate that the difference in course selection is not statistically significant.

Nevertheless, we estimate a separate model (equation 2) that controls for the specific classes in which students enroll. This allows us to compare only students who enrolled in the identical course section at the same college during the same term, focusing on their academic outcomes in that specific classroom. According to the findings presented in Table 3.7, students whose initial course was in a condensed course format show a 1.94 percentage point increase in the likelihood of completing the subsequent course, after controlling for the exact course section they enrolled in. However, no significant differences were found in passing rates with a grade of D or higher, or in overall course grades. Taken together, these results suggest that taking an initial course in a condensed course format does not yield concerning outcomes on students' subsequent course performance.

Exploratory analysis of student-level outcomes using instrumental variable approach

Our findings thus far indicate that a condensed course format during a student's initial exposure to a field of study significantly increases the likelihood of continuing in subsequent courses within the same field. Such benefits may manifest in two distinct ways: First, a positive initial experience in a condensed course could boost students' interest in continuing studies in the same subject area. This might lead to an increased enrollment in subsequent courses within that field. Second, and perhaps more crucially, the benefits of a condensed course format—such as increased flexibility—may also bolster students' overall commitment to their college education. The latter is especially important as persistent enrollment is key to accessing broader economic

opportunities, which is particularly vital for students from disadvantaged backgrounds (Bettinger, 2014; Herbaut & Geven, 2020).

To examine the possible influence of condensed formats in students' early stage of college career on their subsequent college persistence and performance, we conduct a student-level exploratory analysis that relates the proportion of course credits taken through condensed formats during a student's initial semester in college to their probability of persisting into the second semester in college. To control for students' different course selection, we employ a college-course-set fixed-effect model that compares the results of students who take exactly the same set of courses within the same college during their first semester of enrollment, along with fixed effects for the initial term and available student characteristics. We consider five outcomes in measuring students subsequent college success: (i) whether they persist to the next term in college, (ii) total credits earned within three years after initial enrollment at VCCS, (iii) whether the student received any long-term or short-term degrees from VCCS-- within three years after initial enrollment, and (iv) bachelor's degree attainments within six years after initial enrollment at VCCS.

Table 3.8 presents the results of our exploratory analysis. For easier interpretation, we multiply the proportions of credits taken through condensed format by 10 (such as converting 2 percent to 20 percent). The coefficients hence measure the estimated effects on the outcome measure given a ten percentage point increase in condensed format credits during the initial term. Our findings suggest that students who take 10% more credits in condensed formats in their first

term tend to earn approximately 0.07 more credits within three years of initial enrollment; they are also 0.07 percentage points more likely to earn a credential within three years of starting.[14]

DISCUSSION

Condensed course formats in community colleges have been proposed and increasingly adopted as a strategy to improve educational outcomes by aligning academic offerings with students' scheduling needs and learning preferences. This study leverages administrative data from a statewide community college system to evaluate whether condensed courses—defined as courses lasting no more than 10 weeks—relative to traditional-length courses indeed offer an effective alternative to traditional semester-long courses in terms of student concurrent performance and downstream academic engagement. Based on a two-way fixed effects model, our results indicate that students enrolled in condensed courses during their initial exposure to a field of study generally exhibit improved performance in concurrent courses, evidenced by higher rates of course completion and better grades. Importantly, these students also demonstrate increased likelihood of enrolling in subsequent courses within the same field. After controlling for the specific classes that students enroll in next, we find no evidence that initial enrollment in a condensed course format negatively impacts subsequent class performance. On the contrary, we have identified a small but significant positive effect on students' likelihood of persisting until the end of the subsequent course. Additional exploratory student-level analysis suggests that the initial benefits of condensed formats would translate to a higher number of total credits

[14] However, even with controls for the initial term course portfolio, our analysis does not fully address potential student sorting into condensed courses. For instance, as shown in the sorting test (Table 3.32), part-time students and adult learners are more inclined to enroll in condensed courses, which could also influence their academic outcomes.

earned and an increased probability of earning a college-awarded certificate, although the size of the effects is fairly small.

Our findings regarding the positive impact of condensed courses on students' concurrent outcomes largely align with prior literature (Anastasi, 2007; Austin & Gustafson, 2006; Ferguson & DeFelice, 2010; Stith, 2023; Walsh et al., 2019). For those who report the effect size, our estimate of student current course grades (0.12) is smaller than the 0.35 reported by Walsh et al. (2019), who used a hierarchical model with limited student controls (gender, citizenship, and prior GPA) to compare 6-week summer courses with 14-week courses. It is also smaller than the 0.40 reported by Austin and Gustafson (2006) for summer courses. Our study has found a more pronounced positive impact on students' subsequent course enrollment, while prior studies generally found a null impact of condensed courses on students' subsequent outcomes. The difference might be due to methodological differences and the fact that prior studies focused on limited subjects such as math (Bishop et al., 2018; Guillory, 2018). The presence of null effects in previous studies across various subjects suggests a need to further investigate heterogeneous effects to understand how impacts vary across fields.

The policy implications of this study suggest that condensed courses may be a viable option to better meet students' needs. The promising outcomes of condensed courses on students' current academic performance and subsequent course enrollment underscore the potential of these formats to enhance academic engagement. The potential for condensed course formats to improve immediate academic outcomes suggests that these could be strategically implemented to reduce dropout rates and improve completion rates. However, while the initial benefits are evident, the student-level exploratory analysis suggests that the long-term impact on students' academic success may not be as substantial as hoped. This may be because individual course-

level improvements are not strong enough to impact longer-term outcomes, such as degree attainment. Policymakers should consider strategies to enable the initial positive effects to translate into more significant long-term academic advantages. Promoting condensed formats should be part of broader strategies to enhance educational attainment and foster economic mobility for community college students.

While our study has made significant strides in understanding the impact of condensed course formats on student outcomes, several promising areas for further research have emerged. The mechanisms underlying the effects of these condensed formats relative to traditional-length courses remain partially understood, suggesting a rich field for deeper exploration. Prior research utilizing survey data has illuminated differences in student perceptions across various course formats. Although educators anecdotally express concerns that increased workload might lead to higher stress levels, prior studies suggest mixed findings about students' perceived workload and stress levels in condensed courses (Lutes & Davies, 2013, 2018; Walsh et al., 2019). The variation in teaching techniques across different course formats presents another intriguing avenue for research. Studies have suggested that different pedagogical approaches may yield diverse outcomes in student learning and retention (Mena-Guacas et al., 2023; Vreven & McFadden, 2007). Investigating how specific teaching strategies are more efficient in condensed courses and how these strategies affect learning in condensed versus traditional courses could provide valuable guidelines for curriculum design and instructor training.

REFERENCE

AACC. (March 9, 2023). Share of students enrolled in community colleges for credit in the United States in 2021, by age [Graph]. In Statista. Retrieved May 14, 2024, from https://www.statista.com/statistics/421288/percentage-of-for-credit-students-enrolled-in-us-community-colleges-by-age/

Anastasi, J. S. (2007). Full-Semester and Abbreviated Summer Courses: An Evaluation of Student Performance. *Teaching of Psychology, 34*(1). https://doi.org/10.1080/00986280709336643

Austin, A. M., & Gustafson, L. (2006). Impact of Course Length on Student Learning. *Journal of Economics and Finance Education, 5*(1).

Bahr, P. R., Boeck, C. A., & Cummins, P. A. (2021). *Strengthening Outcomes of Adult Students in Community Colleges*. https://doi.org/10.1007/978-3-030-44007-7_3

Barnett, E. A., Bork, R. H., Mayer, A. K., Pretlow, J., Wathington, H. D., & Weiss, M. J. (2012). Bridging the Gap: An Impact Study of Eight Developmental Summer Bridge Programs in Texas. National Center for Postsecondary Research.

Bettinger, E. (2004). How financial aid affects persistence. In College choices: The economics of where to go, when to go, and how to pay for it (pp. 207-238). University of Chicago Press.

Bishop, T. J., Martirosyan, N., Saxon, D. P., & Lane, F. (2018). The Effect of Acceleration on Subsequent Course Success Rates Based on the North Carolina Developmental Mathematics Redesign. *Journal of College Academic Support Programs, 1*(Volume 1, Issue 1). https://doi.org/10.36896/1.1fa5

Brenner, V. (2024). Course-Level Evaluation of an 8-Week Calendar Implementation at a Two-

Year College. *Community College Journal of Research and Practice*. https://doi.org/10.1080/10668926.2024.2307055

Bostwick, V., Fischer, S., & Lang, M. (2022). Semesters or quarters? The effect of the academic calendar on postsecondary student outcomes. American Economic Journal: Economic Policy, 14(1), 40-80.

Cho, S.-W., Kopko, E., Jenkins, D., & Jaggars, S. S. (2012). *New Evidence of Success for Community College Remedial English Students: Tracking the Outcomes of Students in the Accelerated Learning Program (ALP).*

Daniel, E. L. (2015). A Review Of Time Shortened Courses Across Disciplines. *College Student Journal*, *34*(2).

Ferguson, J. M., & DeFelice, A. E. (2010). Length of online course and student satisfaction, perceived learning, and academic performance. *International Review of Research in Open and Distance Learning*, *11*(2). https://doi.org/10.19173/irrodl.v11i2.772

Guillory, R. M. (2018). *Do 5-Week, Time-Compressed, Face-to-Face Summer Courses Increase Course Retention Rates of Students at a 2-Year College?* Texas A\&M University-Commerce.

Guy, G., Cornick, J., Holt, R., & Russell, A. (2015). Accelerated Developmental Arithmetic Using Problem Solving. *Journal of Developmental Education*, *39*(1).

Harlow, J. J. B., Harrison, D. M., & Honig, E. (2015). Compressed-format compared to regular-format in a first-year university physics course. *American Journal of Physics*, *83*(3). https://doi.org/10.1119/1.4905171

Hayward, C., Willett, T., & Senior Researchers, M. (2014). *Curricular Redesign and Gatekeeper*

Completion: A Multi-College Evaluation of the California Acceleration Project.

Herbaut, E., & Geven, K. (2020). What works to reduce inequalities in higher education? A systematic review of the (quasi-) experimental literature on outreach and financial aid. Research in Social Stratification and Mobility, 65, 100442.

Holzweiss, P. C., Polnick, B., & Lunenburg, F. C. (2019). Online in half the Time: a Case Study with Online Compressed Courses. *Innovative Higher Education*, *44*(4). https://doi.org/10.1007/s10755-019-09476-8

Jaggars, S. S., Hodara, M., Cho, S. W., & Xu, D. (2015). Three Accelerated Developmental Education Programs: Features, Student Outcomes, and Implications. *Community College Review*, *43*(1). https://doi.org/10.1177/0091552114551752

Lutes, L., & Davies, R. (2013). Comparing the Rigor of Compressed Format Courses to Their Regular Semester Counterparts. *Innovative Higher Education*, *38*(1). https://doi.org/10.1007/s10755-012-9226-z

Lutes, L., & Davies, R. (2018). Comparison of workload for university core courses taught in regular semester and time-compressed term formats. *Education Sciences*, *8*(1). https://doi.org/10.3390/educsci8010034

Mena-Guacas, A. F., Chacón, M. F., Munar, A. P., Ospina, M., & Agudelo, M. (2023). Evolution of teaching in short-term courses: A systematic review. *Heliyon*, *9*(6). https://doi.org/10.1016/j.heliyon.2023.e16933

National Student Clearinghouse Research Center. (2024). Current Term Enrollment Estimates: Fall 2023. Retrieved from https://nscresearchcenter.org/current-term-enrollment-estimates/

Paasch, A. J. (2020). *Students Perceptions of Accelerated Program Learning: A Comparative*

Study [University of Wisconsin--Stout]. www.uwstout.edu/

Scott, P. A., & Conrad, C. F. (1992). A Critique of Intensive Courses and an Agenda for Research. In J. C. Smart (Ed.), *Higher Education: Handbook of Theory and Research* (Vol. 8, pp. 411–459). Agathon Press.

Stith, M. (2023). *The Impact of Term Length on Community College Student Success*. University of the Cumberlands.

Tamburin, A. (2021, December 17). New strategy for Tennessee community colleges. Axios. https://www.axios.com/local/nashville/2021/12/17/new-strategy-for-tennessee-community-colleges

Vreven, D., & McFadden, S. (2007). An empirical assessment of cooperative groups in large, time-compressed, introductory courses. *Innovative Higher Education*, *32*(2). https://doi.org/10.1007/s10755-007-9040-1

Walsh, K. P., Sanders, M., & Gadgil, S. (2019). Equivalent but not the Same: Teaching and Learning in Full Semester and Condensed Summer Courses. *College Teaching*, *67*(2). https://doi.org/10.1080/87567555.2019.1579702

White, T. (2023). *Intensive Courses: A Mixed-Methods Approach to Examining Perceptions of Low SES Community College Students*. Oklahoma State University.

Xu, D., & Jaggars, S. S. (2013). The impact of online learning on students' course outcomes: Evidence from a large community and technical college system. Economics of Education Review, 37, 46-57.

Xu, D. (2019). Academic performance in community colleges: The influences of part-time and full-time instructors. American Educational Research Journal, 56(2), 368-406.

Table 3. 1 Institutional characteristics of VCCS and National sample in fall 2010

	National Sample	VCCS
Panel A: Enrollment		
Full-time enrollment	2,990	3,026
Percent of GRS[1] cohort	39%	45%
12-month undergraduate: 2010-11	10,954	12,475
Graduation rate, total cohort	23	22
Panel B: Student Demographics and SES		
Percent of White	60%	72%
Percent of Black or African American	14%	19%
Percent of Hispanic/Latino	12%	3%
Percent of Asian/Native Hawaiian/Pacific Islander	4%	2%
Percent of American Indian or Alaska Native	2%	0%
Percent of Race/ethnicity unknown	5%	2%
Percent of Nonresident Alien	1%	0%
Percent of female	58%	59%
Percent of receiving any financial aid	75%	69%
Panel C: Institution Finance		
Tuition and fees, 2010-11	2,708	3,157
Expenses per full-time equivalent student:		
Instruction expenses per FTE	4,756	4,184
Research expenses per FTE	13	0
Public service expenses per FTE	211	73
Academic support expenses per FTE	932	871
Student service expenses per FTE	1,194	650
Institutional support expenses per FTE	1,734	1,318
All other expenses per FTE	1,964	1,210
Percent of core revenues		
Tuition and fees	17%	23%
State appropriations	27%	36%
Local appropriations	14%	0%
N	808	23

Note: The author derived the data from the IPEDS 2010 Fall data; GRS represents full-time, first-time degree- or certificate-seeking students.

Table 3. 2 Student Descriptive Statistics

Female	50%
Age at entry term	21.73
	(7.82)
White	56%
Black	23%
Hispanic	10%
Asian	5%
Other race	5%
Pell Grant eligible	49%
First-generation student	37%
High school completion	97%
Fall enrollment	72%
Associate's degree intent	76%
Full-time enrollment indicator	55%
N	179,146

Note: Standard errors are in parentheses.

Table 3. 3 Course Descriptive Statistics

	Condensed Courses	Traditional-Length Courses
Panel A: Course-Section Characteristics		
Face-to-face section	61.48%	76.81%
Online section	25.73%	16.56%
Hybrid section	11.65%	4.98%
Other mode section	1.14%	1.65%
Class size	19.48	22.74
	(8.99)	(9.62)
Course on weekend	4.09%	1.45%
Course during the day	76.21%	79.82%
Course credit hours	2.84	3.09
	(0.92)	(0.65)
Female instructor	49.12%	51.25%
White instructor	79.52%	81.07%
Black instructor	14.35%	11.99%
Hispanic instructor	1.86%	2.08%
Asian instructor	3.13%	3.52%
Instructor of other race	1.14%	1.34%
Full-time instructor	38.75%	45.46%
VCCS teaching experience (month)	56.99	58.46
	(32.46)	(33.23)
N	24,163	97,244
Panel B: Current Course Outcomes		
Persist to the end of the course	93.46%	89.15%
Pass the course (D or higher)	81.41%	72.50%
Pass the course (C or higher)	76.71%	65.95%
Course grade	2.65	2.21
	(1.31)	(1.24)
N	24,163	97,244
Panel C: Subsequent Outcomes		
Student field outcomes		
Take additional course	41.91%	34.85%
Take additional course and pass (D or higher)	34.74%	28.37%
N	24,163	97,244
Student next-class outcomes (conditional on subsequent enrollment in the same subject area)		
Persist to the end of the course	92.71%	91.83%
Pass the course (D or higher)	81.82%	80.86%
Course grade	2.63	2.51
	(1.35)	(1.23)
N	13,022	54,837

Note: This table uses data aggregated at the course-section level from student-by-course-section data; standard errors are in parentheses.

Table 3. 4 Probability of taking condensed courses

	Condensed course enrollment
Female	-0.0048***
	(0.0013)
Adult learner	0.0257***
	(0.0063)
Whether employed	-0.0102**
	(0.0047)
Race/ethnicity (reference: White)	
Black	0.0189***
	(0.0043)
Hispanic	0.0049
	(0.0045)
Asian	0.0097***
	(0.0017)
Other race	0.0040
	(0.0026)
Pell Grant eligible	0.0056
	(0.0061)
First-generation student	-0.0032***
	(0.0007)
High school completion	-0.0030
	(0.0058)
Associate's degree intent	-0.0123**
	(0.0044)
Full-time enrollment indicator	-0.0416**
	(0.0150)
Online section	0.0806**
	(0.0299)
Hybrid section	0.1583***
	(0.0230)
Other mode section	0.0138
	(0.0175)
Class size	-0.0022***
	(0.0003)
Course credit hours	-0.0812***
	(0.0211)
Constant	0.4502***
	(0.0695)
Missing Control (Female, employment status)	YES
College-course-term FE	YES
Declared major control	YES
Observations	447,573
R-squared	0.4375

Note: Robust standard errors in parentheses. Standard errors are clustered at the college level. *** p<0.01, **p<0.05, *p<0.1.

Table 3. 5 Impact of Condensed Courses on Introductory Course Performance

	Persist to the end of the course	**Pass the course (D or higher)**	**Pass the course (C or higher)**	**Course grade**	**Course retaking**
	(1)	(2)	(3)	(4)	(5)
Courses less than 10 weeks	0.0070**	0.0212***	0.0308***	0.1224***	-0.0091***
	(0.0035)	(0.0048)	(0.0052)	(0.0188)	(0.0034)
Course section features control	YES	YES	YES	YES	YES
Student FE	YES	YES	YES	YES	YES
*College*Intro-course FE*	YES	YES	YES	YES	YES
Sample Mean	0.907	0.748	0.679	2.255	0.100
Observations	447,573	447,573	447,573	447,573	447,573
Adjusted R-squared	0.6325	0.7285	0.7220	0.7808	0.5470

Note: Course grades are measured on a scale from zero to four. Each model incorporates controls for student individual fixed-effects and college-by-course fixed-effects. Furthermore, the models account for course section characteristics, including course modality (online vs. face-to-face), the total number of students in the course section, and the course schedule (weekend or weekday, day or evening). Instructor attributes are also included, encompassing gender, race/ethnicity, employment status (full-time vs. part-time), teaching experience. Classes below the college level, student developmental courses, and courses graded on a pass/fail basis are excluded. Robust standard errors in parentheses. Standard errors are clustered at the student and college-by-course level. *** $p<0.01$, **$p<0.05$, *$p<0.1$.

Table 3. 6 Impact of Condensed Courses in Introductory Courses on Subsequent Enrollment in the Subject

	Take additional course	**Take additional course and pass (D or higher)**	**Take additional course and pass (C or higher)**	**Average grade of the next course**
	(1)	(2)	(3)	(4)
Courses less than 10 weeks	0.0321***	0.0314***	0.0294***	0.0051
	(0.0053)	(0.0051)	(0.0048)	(0.0058)
Course section features control	YES	YES	YES	YES
Student FE	YES	YES	YES	YES
*College*Intro-course FE*	YES	YES	YES	YES
Sample Mean	0.405	0.330	0.304	2.403
Observations	447,573	447,573	447,573	179,735
Adjusted R-squared	0.6494	0.6405	0.6329	0.9375

Note: Course grades are measured on a scale from zero to four. Each model incorporates controls for student individual fixed-effects and college-by-course fixed-effects. Furthermore, the models account for course section characteristics, including course modality (online vs. face-to-face), the total number of students in the course section, and the course schedule (weekend or weekday, day or evening). Instructor attributes are also included, encompassing gender, race/ethnicity, employment status (full-time vs. part-time), teaching experience. Classes below the college level, student developmental courses, and courses graded on a pass/fail basis are excluded. The observations in column (4) are fewer than the observations in Table 3.6 because some courses were not offered before thus no average historical grade can be used. Robust standard errors in parentheses. Standard errors are clustered at the student and college-by-course level. *** p<0.01, **p<0.05, *p<0.1.

Table 3. 7 Impact of Condensed Course in Introductory Courses on Subsequent Course Grade in the Subject

	Persist to the end of the course	**Pass the course (D or higher)**	**Course grade**
	(1)	(2)	(3)
Courses less than 10 weeks	0.0194***	0.0017	0.0112
	(0.0074)	(0.0091)	(0.0318)
Course section features control	YES	YES	YES
Student FE	YES	YES	YES
*College*Intro-course FE*	YES	YES	YES
Subsequent-course-section FE	YES	YES	YES
Sample Mean	0.931	0.820	2.533
Observations	181,487	181,487	181,487
Adjusted R-squared	0.9242	0.9364	0.9532

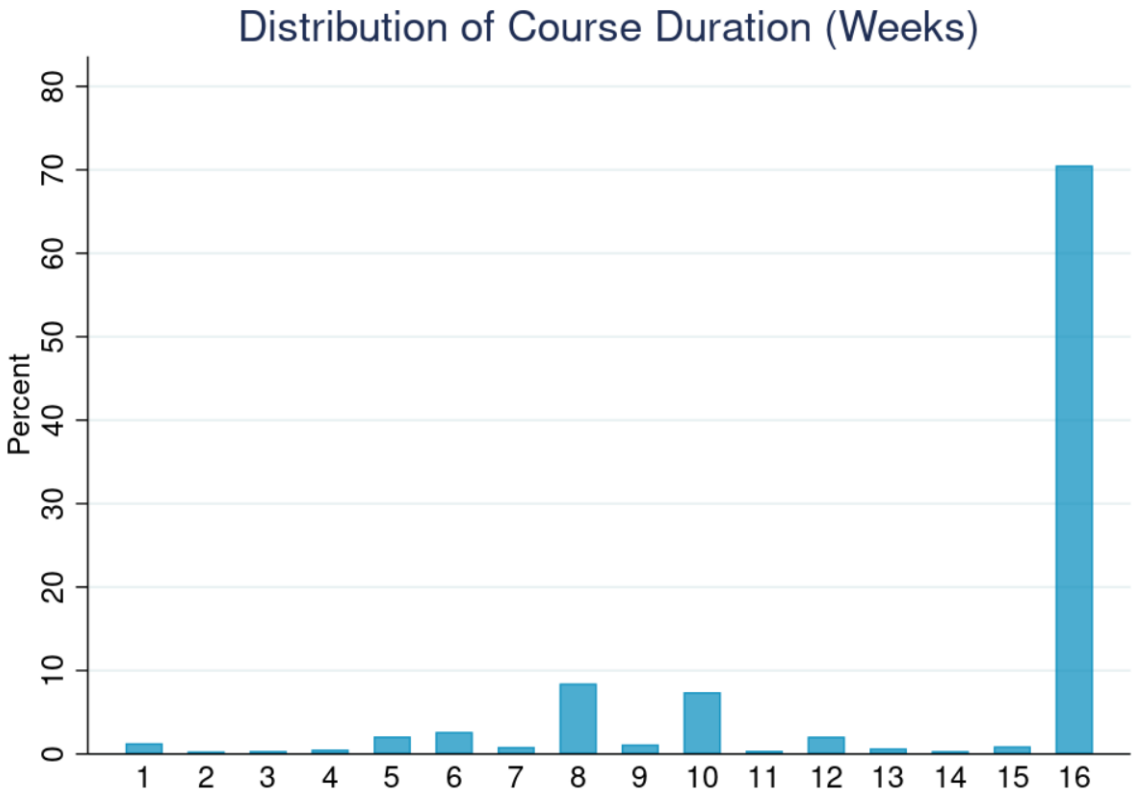
Note: Course grades are measured on a scale from zero to four. Each model incorporates controls for student individual fixed-effects, college-by-course fixed-effects, and subsequent course-section fixed-effects. Furthermore, the models account for course section characteristics, including course modality (online vs. face-to-face), the total number of students in the course section, and the course schedule (weekend or weekday, day or evening). Instructor attributes are also included, encompassing gender, race/ethnicity, employment status (full-time vs. part-time), teaching experience. Classes below the college level, student developmental courses, and courses graded on a pass/fail basis are excluded. Robust standard errors in parentheses. Standard errors are clustered at the student and college-by-course level. *** p<0.01, **p<0.05, *p<0.1.

Table 3. 8 Exploratory student-level analysis: Relationship between % of credits taken with condensed course formats in initial term on subsequent student outcomes

	Persistence into the next term	Attempted credits within three years	Any degree/credential attainment within three years	Bachelor's degree attainment within six years
	(1)	(2)	(3)	(4)
% of Credits Taken with Condensed Course Formats (Multiplied by 10)	0.0005	0.0651*	0.0007**	0.0006
	(0.0005)	(0.0379)	(0.0003)	(0.0006)
First Term Control	YES	YES	YES	YES
Student Characteristics	YES	YES	YES	YES
College-Course-Sets FE	YES	YES	YES	YES
Sample Mean	0.720	33.756	0.371	0.326
Observations	179,200	179,200	179,200	179,200
Adjusted R-squared	0.5165	0.6154	0.5822	0.5460

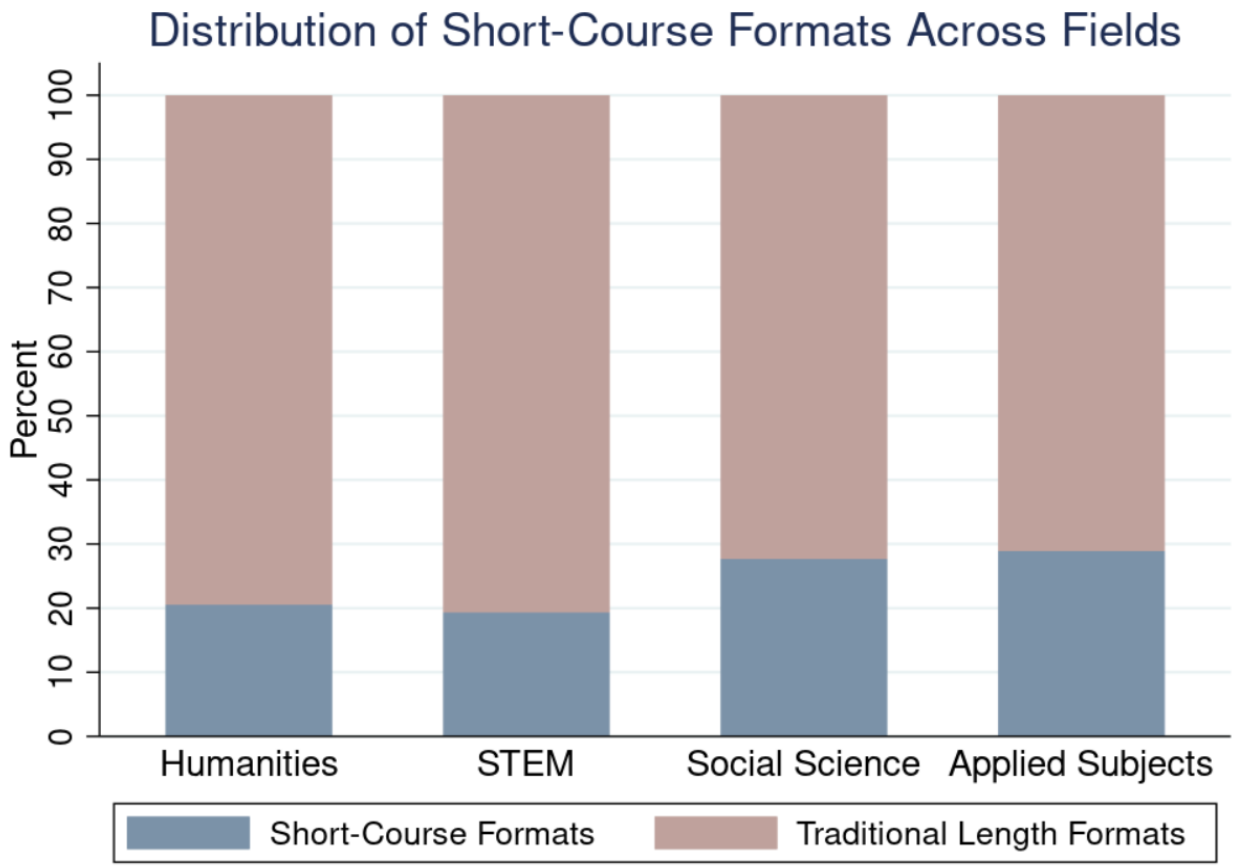
Note: Each model incorporates controls for college fixed-effects, time fixed-effects, college course portfolio fixed-effects, and student features control. Robust standard errors are in parentheses. Standard errors are clustered at the student and college-by-subject level. *** p<0.01, **p<0.05, *p<0.1.

Figure 3. 1 Distribution of course duration (measured by weeks)



Note: We use administrative course-level data from VCCS for the 2010-2011 to 2014-2015 academic years.

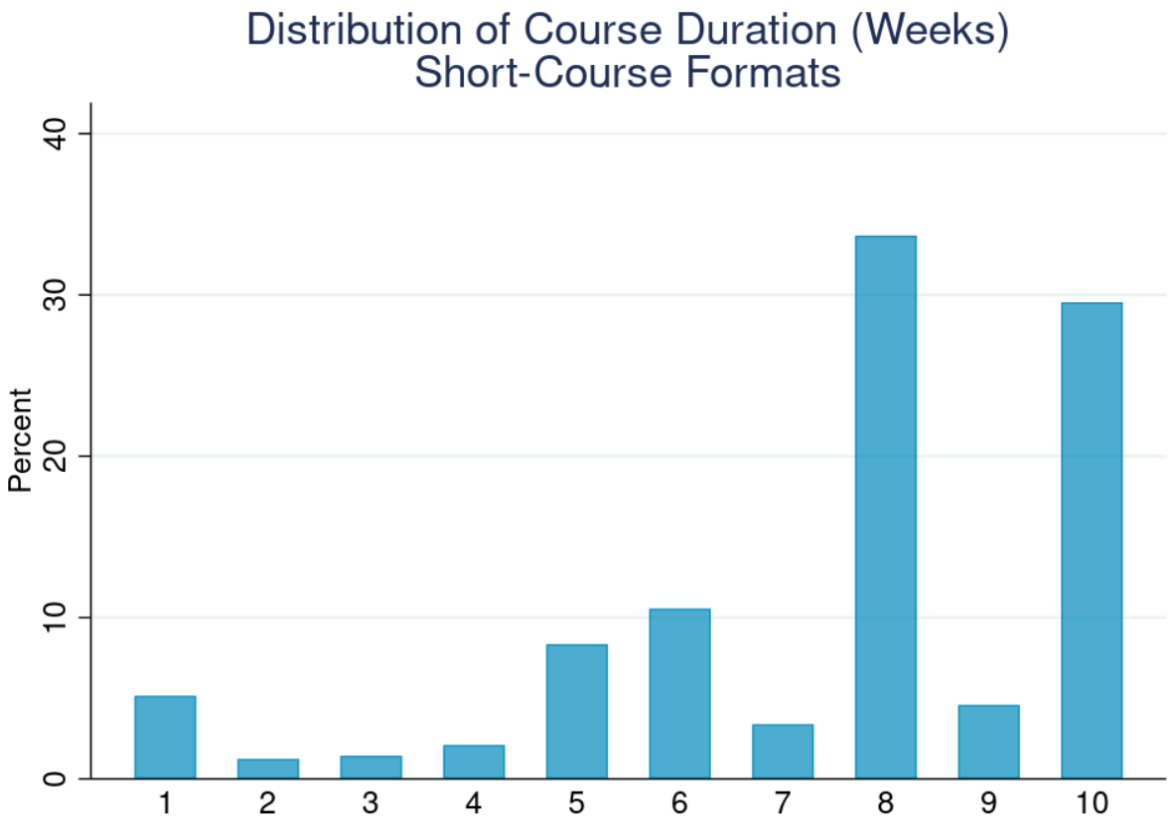
Figure 3. 2 Distribution of course duration across fields



Note: We use administrative course-level data from VCCS for the 2010-2011 to 2014-2015 academic years.

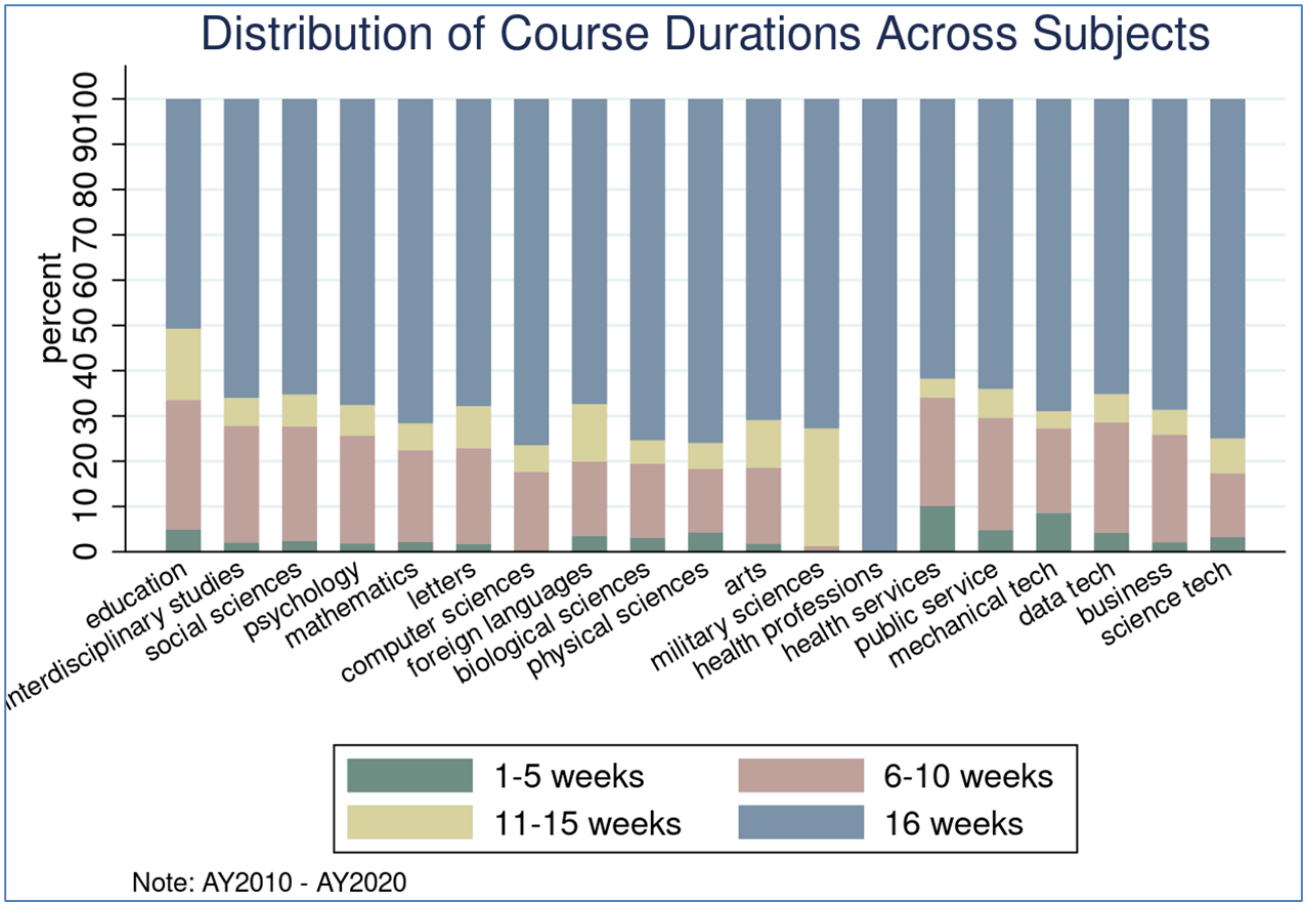
APPENDIX

Appendix Figure C1 Distribution of course duration (measured by weeks) within condensed course formats



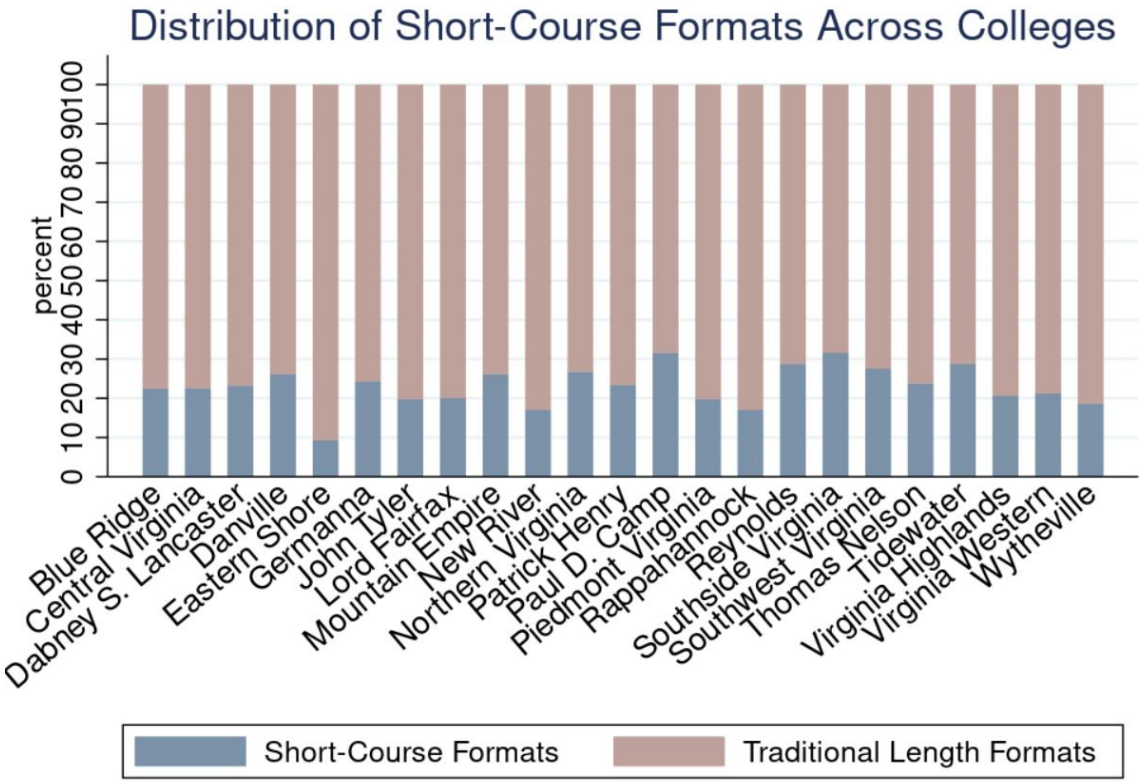
Note: We use administrative course-level data from VCCS for the 2010-2011 to 2014-2015 academic years.

Appendix Figure C2 Distribution of course duration across detailed fields



Note: We use administrative course-level data from VCCS for the 2010-2011 to 2019-2020 academic years.

Appendix Figure C3 Distribution of condensed course formats across fields



Note: We use administrative course-level data from VCCS for the 2010-2011 to 2014-2015 academic years.

CHAPTER 4 SUMMARY AND CONCLUSIONS

By exploring how curriculum structure is associated with students' peer composition, how the curriculum policy—condensed courses—impacts student outcomes, and how to design efficient study groups that promote student engagement by modifying group composition, I demonstrate how curriculum policy and practices influence students' peer environment and learning opportunities through their classroom experience. These studies showcase how rigorous quantitative analysis can shed light on identifying and designing beneficial curriculum policies and practices to improve college students' experiences and learning outcomes.

Specifically, for Paper 1, I found that considerable gender sorting persists even within majors: 61% of the variation in the proportion of same-gender peers arises from within-major variations for female students, and 48% for male students. The within-major variation in students' proportions of same-gender classmates is almost equivalent to the variation observed across the majority of fields. We also find that sorting by gender is more prevalent for certain groups of students. For instance, female international students are more likely to have female peers than female domestic students. In addition to these specific findings, our study provides a guide for how researchers and practitioners can measure this phenomenon across campus contexts to examine how the structure of majors and the degree of freedom of curricular choice within majors can affect student sorting within majors.

For Paper 2, I identified a strategy to construct study groups in online STEM courses to promote student engagement and sense of belonging. The findings indicate that students in study groups with diverse leadership motivation exhibited higher levels of participation, engagement,

and satisfaction. Additionally, the positive impact is corroborated by their group conversations, which demonstrated a more positive tone and a higher level of leadership. The benefits were particularly pronounced for URM students, extending to their broader social networks within the class and enhancing their sense of belonging. These results suggest a promising strategy for facilitating study groups that can bolster student engagement and sense of belonging in challenging STEM courses, with the potential to contribute to diversity in STEM degree completion and the STEM workforce.

For Paper 3, I found that condensed courses positively impact students' current course outcomes, including course persistence, completion rates, and grades. Students enrolled in condensed courses are more likely to enroll in, persist through, and successfully complete subsequent courses within the same subject area over a three-year follow-up period. However, when controlling for the specific course section in which students enroll for subsequent courses, the positive impact on subsequent course performance diminishes, suggesting nuanced dynamics in how condensed courses influence long-term academic trajectories. This study contributes to the ongoing dialogue on educational innovation by providing rigorous causal evidence on the effectiveness of condensed courses with large-scale data. It highlights the potential of condensed courses to enhance immediate academic outcomes and encourage continued engagement in higher education.

Future Research

For Study 1 on student sorting within majors:

Future studies should incorporate more nuanced course-level features to better capture students' course-choice behaviors. This study does not explicitly examine how course content might lead to sorting within majors. Given that students within any major can choose courses that vary substantially in subject matter, future work could explicitly explore gendered preferences for course content. In-depth interviews with individual students or large-scale text-mining exercises using course descriptions could identify patterns. Furthermore, although our research is based on a single institution, the methodology is designed for broad applicability. It could easily be applied to examine within-major gender sorting on other campuses or to explore sorting among other significant demographic groups, such as first-generation college students or those from various racial and ethnic backgrounds.

For Study 2 on constructing efficient study groups:

Future research should aim to replicate the experiment in different educational contexts to determine if similar positive impacts can be observed. Additionally, it should investigate various group compositions and adjust control groups to ascertain the optimal arrangement. Another limitation of this study is the absence of tracking students' long-term outcomes to further examine if the positive treatment effect extends and further improves students' persistence in STEM fields or their social networks in college. This study underscores the significance of leveraging insights from the well-established literature on peer effects (Hoxby 2000; Zimmerman 2003), and the critical role of social skills (Arendale, 2004; Barry & Stewart, 1997; Deming, 2017; Hara et al., 2000; Johnson & Johnson, 2002; Riedl et al., 2021; Woolley et al., 2010), elucidating how these factors contribute to the dynamics of student interaction within

group processes. We advocate for future research to further investigate how group composition, with respect to student characteristics, influences these processes in ways that are advantageous to students' social belonging and learning within demanding educational contexts.

For Study 3 on the impact of condensed courses:

One limitation of our study is the ambiguity surrounding the mechanisms behind the impact of the condensed course format. The distinctions between condensed course formats and traditional-length courses are not thoroughly understood. Previous research using survey data has highlighted differences in student perceptions across various course formats, with some studies suggesting that students may experience a lower workload and feel less stressed in condensed courses (Lutes & Davies, 2013, 2018; Walsh et al., 2019). Differences in teaching techniques used in various course formats may also affect student outcomes (Mena-Guacas et al., 2023; Vreven & McFadden, 2007). Further investigation is needed to explore the detailed mechanisms leading to the impact of the condensed course format.