

Rooming with Joe vs. Zhou: Differential Impacts of International Roommates on College Outcomes

Hsin-Ta Tsai and Jose Eos Trinidad

Working paper no. 112

April 2024



Published by the Centre for Global Higher Education,
Department of Education, University of Oxford
15 Norham Gardens, Oxford, OX2 6PY
www.researchcghe.org

© the authors 2024

ISSN 2398-564X

The Centre for Global Higher Education (CGHE) is an international research centre focused on higher education and its future development. Our research aims to inform and improve higher education policy and practice.

CGHE is a research partnership of 10 UK and international universities, funded by the Economic and Social Research Council, with support from Research England. ESRC award number ES/T014768/1

Rooming with Joe vs. Zhou: Differential Impacts of International Roommates on College Outcomes

Hsin-Ta Tsai and Jose Eos Trinidad

Contents

Peer Effects in Higher Education	9
Current Study	11
Data and Methods.....	14
Findings	19
Robustness Checks	24
Discussion and Conclusion.....	28
Data availability statement.....	32
Declaration of Interest Statement	32
Endnotes.....	33
References.....	34
Appendix A.....	39
Appendix B.....	40

Rooming with Joe vs. Zhou: Differential Impacts of International Roommates on College Outcomes

Hsin-Ta Tsai and Jose Eos Trinidad

Hsin-Ta (Andre) Tsai is a Joint Ph.D. student in Public Policy at Georgia Institute of Technology and Georgia State University. He studies policy analysis and evaluation to assess the impacts of social and educational policies on marginalized communities in the United States and globally, such as people from lower socioeconomic backgrounds and the LGBTQ+ community. He holds a B.A. in Economics and Education Studies from Berea College, an MSc in Comparative and International Education from the University of Oxford, and an MA in Policy Studies from the University of Chicago. His interdisciplinary background and commitment to social justice drive his work to promote equity and inclusion through policy research and data analysis. <https://spp.gatech.edu/people/person/hsinta-andre-tsai>

Jose Eos Trinidad is an Assistant Professor in the Policy, Politics, and Leadership cluster of the University of California Berkeley School of Education. He is a sociologist with expertise in the study of organizations outside schools and the study of schools as organizations. He received his Joint PhD in Sociology and Comparative Human Development from The University of Chicago. His research explores how organizations affect school improvement and inequalities, education policy, and institutional change. Conceptually, he focuses on organizational networks and institutional theory. Substantively, he studies policies on organizational data use, teacher wellbeing, and student-centered learning. Methodologically, he integrates quantitative causal inference and qualitative interview approaches to understand education policy and systems holistically. <https://bse.berkeley.edu/jose-eos-trinidad>

Abstract

This study examines the influence of roommate pairings between international and domestic students on college outcomes. We employ quasi-experimental methods, using institutional data (7,000+ students spanning 15 years) from a US liberal arts college serving low-income students. Results show positive effects on first-year GPA for domestic students, with effects persisting but diminishing over time. A significant increase in second-year retention is also observed. However, no graduation effect and only a small global outlook effect are found. For international students, rooming with domestic US students showed no difference compared to rooming with other international students. This points to the potential benefits of intentional policies to promote interactions among individuals from diverse backgrounds and cultures.

Keywords: Peer effects, International students, Disadvantaged students, College outcomes, Diversity and inclusion

Rooming with Joe vs. Zhou: Differential Impacts of International Roommates on College Outcomes

Hsin-Ta Tsai and Jose Eos Trinidad

Various social scientists have tried to understand how social settings influence individuals' decisions, behaviors, and outcomes (Manski, 1993; Neighbors et al., 2007; Ross & Nisbett, 2011). Such studies often involve the understanding of "peer effects," or how one's colleagues, schoolmates, classmates, and roommates may impact one's actions and outcomes (Manski, 1993; Winston & Zimmerman, 2004; Wolniak & Ballerini, 2020). In the field of education, researchers have investigated the effects of peers on student learning, school choice, college admissions, and school quality, all of which have substantive policy relevance (Carrell et al., 2013; Lyle, 2007; Sacerdote, 2001; Stinebrickner & Stinebrickner, 2006).

Despite the widespread interest, three primary concerns have emerged over the decades of peer effects research, particularly in higher education. First, multiple studies have demonstrated the difficulty of using observational data to estimate the causal effects of peer influence which includes self-selection bias that makes it difficult to isolate the endogenous selection effect from the peer effect, the difficulty in attributing the peer effects to the peer's backgrounds like class and demographics, and the difficulty of addressing simultaneity¹ (Brunello et al., 2010; Sacerdote, 2001; Zimmerman, 2003). Second, studies have concentrated on short-term outcomes like term grades, but few have investigated long-term college outcomes such as multi-year retention and graduation (Lyle, 2007, 2009; Sacerdote, 2001; Shook & Clay, 2012; Zimmerman, 2003). Retention and graduation are especially pressing issues for higher education in the U.S., as national data reveal troublingly low persistence and degree completion rates across all racial/ethnic groups (Museus et al., 2018). Lastly, the concern over the composition and context of groups matters, particularly as peer effects may be more or less salient in subpopulations that have historically been marginalized or economically disadvantaged (Brady et al., 2017; Sacerdote, 2014).

The challenges and concerns of measuring peer effects and the mixed findings of peer influence on short-term college academic outcomes complicate how these data-driven evidence can be used by policymakers (Brady et al., 2017) or be translated into actions by practitioners. This paper aims to address some of these concerns. We employ a variety of methods which minimize bias when estimating the treatment effect (Huntington-Klein, 2021; Lunceford & Davidian, 2004; Stuart, 2010). Our dataset also allows for estimations of short- and long-term college outcomes. More importantly, this paper is based on a college context with a predominantly disadvantaged student body in contrast to most other studies on higher education peer effects that investigate traditionally elite schools, e.g., Dartmouth College and Williams College (Sacerdote, 2001; Zimmerman, 2003).

A core aspect of this study is its assessment of the impact of intercultural roommate pairing (i.e., international and domestic student) on a variety of college outcomes— a study motivated by the importance of diversity, equity, and inclusion in higher education (Bowman, 2010; Gurin et al., 2002; Museus et al., 2018; Tienda, 2013). As Museus et al. (2018) found, culturally engaging campus environments explained 50-60 percent of the variance in sense of belonging for both white students and students of color, a crucial element for college student success. Indeed, study by Shook and Clay (2012) found that sense of belonging in college had a partial mediation effect on the relationship between having a majority-race roommate and higher semester GPAs for minority students during first year. Given the importance of fostering a sense of belonging, we note how the intentional design of pairing domestic and international students can be taken up as a college policy lever, not only to promote diversity and inclusion, but also to lead to more traditional measures of success like increased grades and graduation rates. In this paper, we define “international students” as non-immigrant alien students on a student visa and “domestic students” as US citizens or immigrant permanent residents.

Our institutional dataset comes from Berea College, a small liberal arts college in the rural Appalachian region of eastern Kentucky, traditionally identified as one of the poorest regions of the United States (e.g., Glenn, 1970; Vazzana & Rudi-Poloshka, 2019; Voss et al., 2006). It was the first institute of higher education in the South to welcome and educate non-white and female students in 1855 and, since 1892, it has

been providing free tuition to its students. Its current student population is around 1,500 who are predominantly low-income. Both domestic and international students must demonstrate financial need to be considered for admission (98% of its domestic students are Pell Grant² eligible according to Work Colleges Consortium’s webpage as of 2023). International students are required to complete financial forms at the time of application for admission. Internal data from Berea (Office of Institutional Research and Assessment, 2014, 2015, 2022) show its international students have consistently outperformed their domestic American peers in terms of first-to-second year retention rates and six-year graduation rates (see Figures 1 and Figure 2). This notable difference prompts an investigation into the potential benefits for domestic students through their interaction with international counterparts.

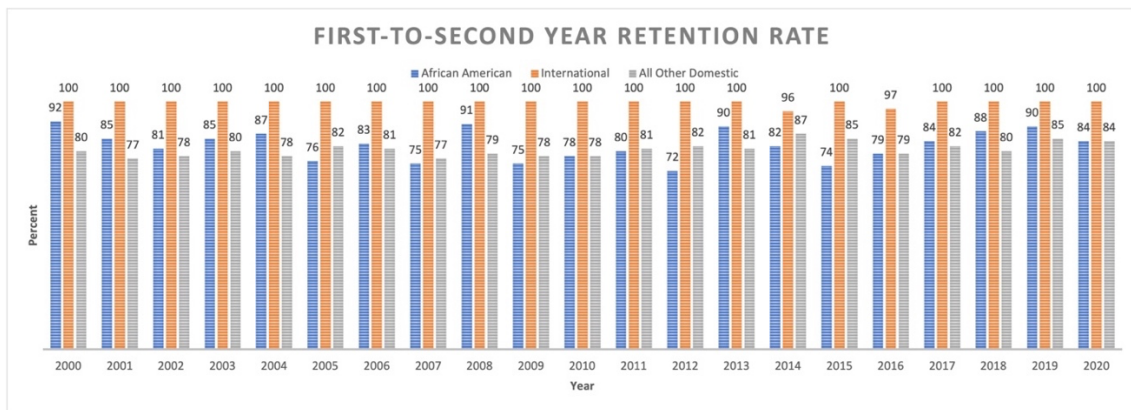


Figure 1 First-to-second year retention rates, first-year students by cohort type. Data from Berea College.

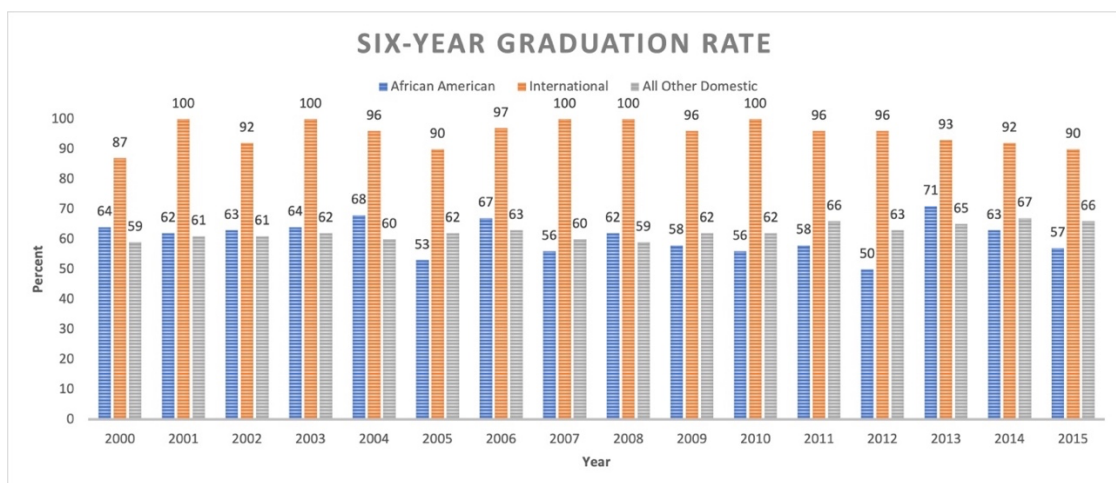


Figure 2 Six-year graduation rates for first-year students, by cohort type. Data from Berea College.

This paper asks if domestic-international roommate pairing during freshman year affects short- and long-term college outcomes, which include academic performance, college persistence, and global outlook. Our findings show that intercultural roommate pairing has positive effects for domestic students and no effects for international students. In particular, for domestic students, we find significant positive effects on academic performance, as measured through GPAs in each of the first four years of college. However, these positive effects diminish in magnitude and statistical significance over time. The positive peer effects also have an impact on domestic students' retention and global outlook. Such positive effects, however, are not observed in the graduation rates for those who had international student roommates, since they did not graduate at higher levels than those who did not have international roommates. Despite the generally positive results for domestic students, these effects did not translate to international students as there are no statistically significant findings for internationals with domestic US roommates (as compared to those with other international roommates).

Peer Effects in Higher Education

Studies on peer effects have documented the significant influence of peers on personal decisions and career choices, particularly for students. For instance, one study discovered that male students' number of binge-drinking episodes increased due to their peers' drinking habits prior to college (Duncan et al., 2005). Another study investigating the relationship between health behavior and academic performance found that male students with roommates who had a history of drinking before college experienced a negative impact on their GPA compared to male students with non-drinking roommates (Kremer & Levy, 2008). Additional studies have explored peer influences on physical fitness scores, academic cheating behavior, likelihood of joining fraternities and sororities, career choices, and perceptions on racial diversity (Boisjoly et al., 2006; Carrell et al., 2008, 2011; Marmaros & Sacerdote, 2002; Sacerdote, 2001).

Studies focusing specifically on academic outcomes in higher education have reported mixed evidence. Zimmerman (2003) conducted a natural experiment at Williams College and found a positive impact on GPA for students who roomed with peers with

higher SAT verbal scores. At Dartmouth College, Sacerdote (2001) found that a roommate's academic ability significantly affected one's own GPA, with higher-ability roommates benefiting each other the most and lower-ability roommates having an adverse effect on each other. Carrell et al. (2009) used the random assignment of students to squadrons at the Air Force Academy and found modest positive effects on freshman GPA from peers' SAT scores, though the effects diminished over time. In contrast, studies using institutional data from the University of Maryland, Wellesley college, and West Point saw little evidence of peer effects on academic performances, prompting further investigations into how context matters for peer effects to take root (Foster, 2006; Lyle, 2007; McEwan & Soderberg, 2006).

Three studies examining peer effects at the current institutional context provide valuable insights into various influences observed among peers. One revealed that students benefit from having roommates with higher high school GPAs which the authors used as a proxy for the student's study effort; however, this effect is significant only for females and not for males (Stinebrickner & Stinebrickner, 2006). In a separate study, (Stinebrickner & Stinebrickner, 2008) found a negative influence on one's GPA when roommates bring video games into the school dorm room. Mehta et al. (2019) showed how students' first-year GPA is influenced by the amount of time their peers spent studying in high school, which the authors referred to as "study propensity." None, however, have studied how sustained interactions between international and domestic roommates may have academic impacts on students.

The presence of and interactions with individuals from a different culture can have a number of important outcomes. Across various scientific studies, the interactions between diverse groups of individuals have been shown to reduce intergroup prejudice (see metaanalysis by Pettigrew & Tropp (2006)). In the case of higher education, studies have shown the benefit of interactions between domestic and international students. For example, Sakurai et al. (2010) found that an intercultural intervention program improved intergroup attitudes and ties. Geelhoed et al. (2003) qualitatively investigated a culturally integrated peer program and found evidence the program reduced intergroup anxiety and promoted intercultural understanding. Sias et al. (2008) showed intergroup anxiety reduction facilitated intercultural friendship development. Shook and Clay (2012) found that interracial roommate relationships

increased minority students' sense of belonging and improved academic performance, highlighting important benefits of intergroup contact through university housing assignment. Collectively, these studies highlight the benefits that come from intergroup relations and interventions. However, less is known about how sustained contact—as manifested in being cross-cultural college roommates—can have implications for students' college outcomes.

This paper aims to contribute to the literature on peer effects in higher education by examining peer effects on various college outcomes, in the context of Berea College's unique international-domestic freshman year roommate pairings. Our findings seek to provide theoretical and practical evidence to further illuminate the impacts of increased diversity and intercultural integration within institutions of higher education.

Current Study

This paper contributes to the growing literature on peer effects and higher education policies by examining the influence of international-domestic roommate pairings on college outcomes. In this section, we highlight the hypotheses we make about the role of sustained intergroup contact—i.e., the pairing between domestic and international students as first-year roommates—on college outcomes.

Studies like by Zhao et al. (2005) revealed that international students overall spent more time on and were more engaged in educationally purposeful activities compared to domestic U.S. students during freshmen year in college. The authors found that first-year international students showed higher levels of engagement specifically related to academic challenge and student-faculty interactions when compared to their US peers. It is possible that such contact may be beneficial for domestic American students as international students positively influence their roommates' study habits and academic motivation. Because of this, we hypothesize that,

Hypothesis 1A: Controlling for confounders, domestic students who room with international students will receive higher GPAs, on average.

Hypothesis 1B: Controlling for confounders, domestic students who room with international students will have higher odds of being retained in school.

On the flip side, it may be the case that the positive effects on domestic students may come at the expense of international students. Studies like the one by Mamiseishvili (2012) found a negative association between international students' social integration on campus and their college persistence in the U.S. In an environment with a decreased emphasis on academics and increased focus on social activities—often influenced by one's peers—international students may reduce their academic efforts (e.g., Duncan et al. (2005)). Thus, we conjecture that,

Hypothesis 2A: Controlling for confounders, international students who room with domestic students³ will receive lower GPAs, on average.

Hypothesis 2B: Controlling for confounders, international students who room with domestic students will have lower odds of being retained in school.

A key facet of research on interventions is the fact that effects “fade out,” where initial positive or negative impacts eventually reduce in magnitude as well as decrease in statistical and substantive significance (Bailey et al., 2017). Moreover, the random assignment to international roommates in our sample happened in the first year and students were subsequently allowed to select their own roommates afterwards. Thus, the hypothesized positive impact for domestic students and hypothesized negative impact for international students may reduce over time:

Hypothesis 3A: Controlling for confounders, the positive influence of peers on GPA for domestic students diminishes in magnitude and significance over subsequent college years (from 2nd to 4th year).

Hypothesis 3B: Controlling for confounders, the negative influence of peers on GPA for international students diminishes in magnitude and significance over subsequent college years (from 2nd to 4th year).

One outcome that may be positively affected by such intercultural pairing is global outlook since these paired international-domestic students are in contact with people from a culture different from their own. As part of Berea College's curriculum, all students must take two classes to fulfill the “International Perspectives” (IP)

requirement.⁴ Given the direct relevance of cross-cultural learning through a foreign roommate for domestic students, we hypothesize that,

Hypothesis 4A: Controlling for confounders, domestic students who room with international students will receive higher IP grades, on average.

On the other hand, it is worth noting that for a given year, Berea College rarely accepts more than one candidate from the same country. Thus, an international student often rooms with another of a different nationality during the first year. For international students, having a domestic US roommate or international roommate from a different country is deemed having “intercultural roommates.” Thus, we hypothesize that,

Hypothesis 4B: Controlling for confounders, international students who room with domestic students will receive no higher IP grades, on average.

Finally, while it would be ideal for graduation to be influenced by one’s peers, graduation can be quite a distant outcome that will be affected by many other factors aside from one’s roommate during the first year (which one can switch afterwards). Thus, we hypothesize that,

Hypothesis 5: Controlling for confounders, domestic-international pairing will have no significant effect on graduation for both domestic and international students.

We test these hypotheses, attempting to understand the role of intercultural roommate pairs on short- and long-term academic outcomes for college students. More importantly, we test this in an environment with many more disadvantaged students—highlighting how diversity can lead to significant gains in personal outcomes for traditionally underserved student populations.

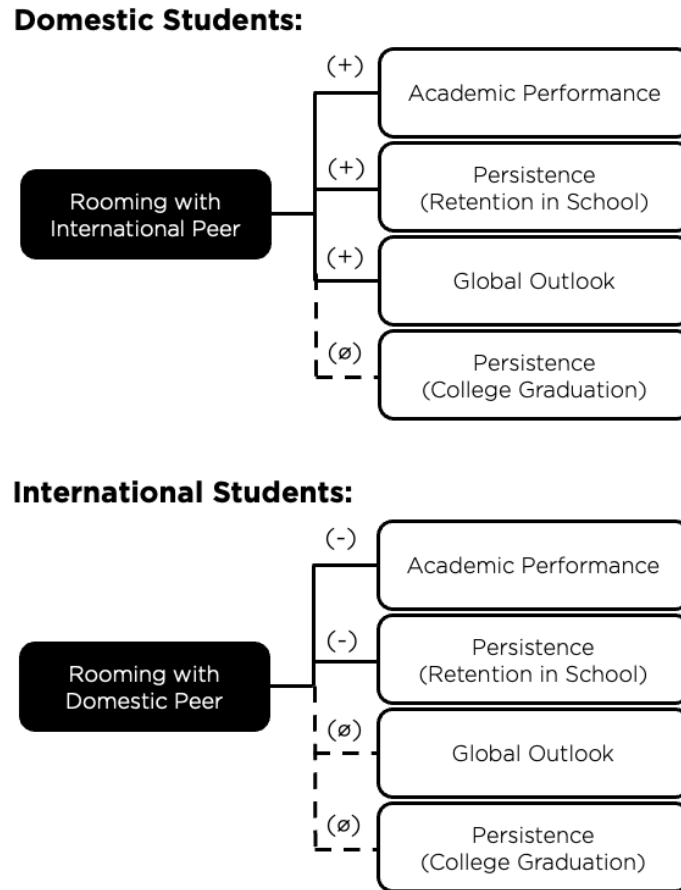


Figure 3 *Illustrated Summary of the Hypothesized Relationships by Subgroup*

Figure 3 illustrates the hypothesized relationships regarding the differential effects of intercultural roommate pairings on the various college outcomes we study.

Data and Methods

This study uses institutional data from Berea College obtained through a direct IRB-approved data sharing agreement. The dataset includes entering student cohorts from 2000 to 2015. The original dataset used in this analysis comprises a total of 7,100 students. We dropped 5 observations with empty data entries and 2 additional observations with entry errors. The final dataset includes 6,665 domestic students and 428 international students. The dataset contains information on students' demographic characteristics, grades, retention, and graduation outcomes. To prevent identification of individual students, the College did not provide certain details like the

students' graduation year and degree program, in keeping with efforts to prevent individual student identification.

Variables

We focus on measuring three key outcome variables. First, *academic performance* is measured by students' grade point averages (GPAs) throughout their first to fourth year in the college (Tables 2 and 4). Second, *college persistence* is captured through retention rates and graduation rates (Table 3 and 5). Third, an indication of *global outlook* is measured by students' mean grades in the two compulsory international perspective (IP) courses (Tables 3 and 5). Although an imperfect proxy, IP grades provide some signal of students' openness to diverse cultures and worldviews.

We use GPA to measure academic performance since prior research supports its use as a reliable and valid predictor in higher education contexts (Bacon & Bean, 2006; Beatty et al., 2015). Retention and graduation are commonly used in higher education research as a measure for persistence. For example, Braunstein et al. (2000) as well as Card & Solis (2022) used graduation and retention to measure students' persistence in college, particularly about how financial factors such as student loans and financial aid can impact both retention and graduation. Using the international perspective course grades to represent global outlook may suffer from issues regarding reliability since students take different IP courses. However, in the absence of alternative measures, we use the mean IP grade as an imperfect approximation of global outlook.

Table 1 Domestic-international Roommate Pairing Status

	<u>Control</u>	<u>Treated</u>	<u>Difference</u>
<i>Panel A. Transfer Student Status</i>			
Transfer Student	0.09 (0.00)	0.04 (0.01)	-0.053*** (0.011)
Obs (n)	6363	730	7093
<i>Panel B. Birth Sex</i>			
Male	0.43 (0.01)	0.53 (0.02)	0.105*** (0.019)
Obs (n)	6362	730	7092
<i>Panel C. Race and Ethnicity</i>			
White	0.72 (0.01)	0.34 (0.02)	-0.381*** (0.018)
Black	0.17 (0.00)	0.12 (0.01)	-0.054*** (0.015)
Hispanic	0.03 (0.00)	0.02 (0.01)	-0.014* (0.007)
AAPI	0.02 (0.00)	0.01 (0.00)	-0.012** (0.005)
Other Race	0.05 (0.00)	0.02 (0.00)	-0.029*** (0.008)
Obs (n)	6363	730	7093
<i>Panel D. Region of State of Origin</i>			
Northeast	0.02 (0.00)	0.02 (0.01)	-0.003 (0.006)
Midwest	0.13 (0.00)	0.07 (0.01)	-0.062*** (0.013)
South	0.82 (0.00)	0.4 (0.02)	-0.415*** (0.016)
West	0.02 (0.00)	0.01 (0.00)	-0.010** (0.005)
Territories and Other States	0 (0.00)	0 (0.00)	-0.001 (0.001)
Foreign	0.01 (0.00)	0.5 (0.02)	0.491*** (0.007)
Obs(n)	6330	728	7058

Notes. Robust standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. Treated = international-domestic roommate pair and Control = Otherwise.

The key independent variable in this study is a binary indicator of whether students were paired with an international or domestic roommate in their first year. Specifically, this variable takes a value of 1 if a domestic student was paired with an international roommate or vice versa, and 0 otherwise. First-year roommate assignments were supposedly random at the College during the sample period from 2000-2015. However, the statistically significant demographic differences observed between those paired and those not paired, as indicated in Table 1, suggest that some students changed roommates despite the intended random assignment. This renders the treated and control groups incomparable, deviating from the assumption of completely random roommate assignment in the three studies conducted at Berea (Mehta et al., 2019; Stinebrickner & Stinebrickner, 2006, 2008). All three studies used a unique dataset collected by the authors, in contrast to the institutional data employed in the current study. While it would have been useful to establish stronger causal effects using an instrumental variable design for the subgroup of compliers (e.g., Carrell et al. (2009) and Foster (2006)), data on students' initial randomized roommate assignments were unfortunately unavailable. Thus, we rely on observed covariates to account for factors predicting roommate pairing and college outcomes in our models.

Our covariates include measures of transfer student status, race and ethnicity, sex at birth, and the states the students are from. The state variable is recoded into the four regions based on the U.S. census classification: Northeast, South, Midwest, and West. Table 1 presents the list of available control variables and the descriptive statistics for the treatment and the control groups in the original raw data. The difference-in-means results reveal statistically significant differences between the treatment and control groups across the demographic characteristics. This indicates the two groups are not inherently comparable, so a simple OLS approach could lead to biased estimation of the treatment effect due to confounding.

Analytic Strategy

Our analytic strategy uses quasi-experimental methods to control for confounding factors. Specifically, we use inverse probability weighting, or IPW, as our primary estimation approach (Chesnaye et al., 2022; Huntington-Klein, 2021; Lunceford & Davidian, 2004). First, we did separate analyses for domestic and international

students. We then estimated the propensity score by modeling treatment assignment as a function of all observed covariates, denoted by X :

$$p(x) = P(T = 1 | X = x)$$

where p is the propensity score, T is an indicator variable denoting the international-domestic roommate pairing status, and X the vector of covariates.

Although some international students reported that Berea College previously required international-domestic roommate pairing (potentially violating the common support/overlap assumption), our data showed exceptions, with a number of international students rooming together. When contacted, Berea College stated pairing was only encouraged, suggesting the policy may have changed over time. To test the overlap assumption, we plotted propensity score distributions, which showed substantial overlap between the treated and control groups, with no zero-density gaps (See Appendix A). This provides evidence that the common support assumption is reasonably satisfied, allowing us to proceed with estimating treatment effects for the international student subpopulation.

Although our data include domestic students' region of origin, we excluded this variable when estimating propensity scores for three reasons. First, around 82 percent of domestic students came from the South, creating a skewed distribution. Second, international students lacked region data, so excluding it allowed consistent covariate sets between the domestic and international student subgroups when estimating the propensity scores. Third, prior work on peer effects at the same institution also did not use state or region variables (Mehta et al., 2019; Stinebrickner & Stinebrickner, 2006). To validate that excluding the "region" variable did not violate the common support and covariate balance assumptions, we conducted overlap as well as formal covariate balance tests with and without this covariate. In both cases, we achieved reasonable common support and balance (see Appendix B). While not perfect, these diagnostic tests provide reasonable evidence that the IPW weighting created treatment and control groups of sufficient comparability to proceed with estimating the treatment effect, which we also followed up with a sensitivity analysis using other model specifications (see the robustness checks section).

Our study estimates both the average treatment effect (ATE) and the average treatment effect on the treated (ATET) using the IPW approach, which constructs synthetic control groups by reweighting the sample observations using the estimated propensity score, p . Specifically, IPW upweights untreated units with characteristics similar to the treated group while treated units resembling the untreated group receive higher weights. That is, a weight of $\frac{1}{p}$ is assigned to the treatment group and $\frac{1}{(1-p)}$ is assigned to the untreated group. This aims to achieve covariate balance between the weighted groups to estimate the ATE. In estimating ATET, the goal is to make the untreated group synthetically comparable to the treated group. IPW does this by first assigning a constant weight of 1 to all treated units and assigning $\frac{p}{(1-p)}$ to the untreated units to create a control group comparable to the treatment group (Huntington-Klein, 2021).

We draw on a battery of robustness checks by comparing our IPW estimates to results obtained through augmented inverse-probability weighting (AIPW), propensity score matching (PSM), and nearest neighbor matching (NNM) models to assess consistency across different model specifications. We particularly take advantage of the "double robustness" property of AIPW estimators (Huntington-Klein, 2021; Lunceford & Davidian, 2004) to account for IPW estimators' relying solely on the correct propensity score model. Lunceford & Davidian (2004), in particular, provide rationale to use AIPW routinely. We note, however, that AIPW only estimates ATE. These matching and weighting methods also help us address missing data. IPW, NNM, and PSM essentially leverage observed outcomes from one treatment group to fill in the missing potential outcomes for subjects in the other group in a careful manner that maintains balance across groups being compared.

Findings

This paper tests several hypotheses related to the effects of freshman year domestic-international roommate pairings on short- and long-term college outcomes, separately for domestic and international students.

Impact on Domestic Students

Academic Outcomes.

The ATE estimates in Table 2 (Panel A) provide evidence supporting Hypothesis 1A pertaining to the positive effect on GPAs. If all domestic students were to room with international roommates, their average 1st and 2nd year GPAs would increase by 0.14 ($p < 0.05$) and 0.10 ($p < 0.1$) points, respectively, compared to the average of 2.72 and 2.69 that would occur if none of the students had done so. When focusing on the treated students, the ATET estimates (Panel B) also support Hypothesis 1A. Specifically, the average 1st year GPA is 0.15 points higher when all the students who room with international students during the first year do so compared to the control mean outcome (of 2.65) that would have occurred if none of these students had roomed with international students ($p < 0.01$). Furthermore, the declining magnitudes and significance levels of these positive ATE and ATET estimates on GPA from 2nd to 4th year suggest the benefits diminish over time, providing evidence for Hypothesis 3A.

Table 2 Effects of International Student Roommate on Domestic Students: Multiyear GPA

<i>Panel A. Average Treatment Effect (ATE)</i>				
	1st Year GPA	2nd Year GPA	3rd Year GPA	4th Year GPA
ATE	0.136** (0.054)	0.090* (0.053)	0.081 (0.053)	0.073 (0.053)
Control	2.72*** (0.012)	2.69*** (0.012)	2.69*** (0.012)	2.69*** (0.012)
<i>Panel B. Average Treatment Effect on the Treated (ATET)</i>				
	1st Year GPA	2nd Year GPA	3rd Year GPA	4th Year GPA
ATET	0.149*** (0.049)	0.109** (0.048)	0.100** (0.048)	0.092* (0.048)
Control	2.65*** (0.017)	2.62*** (0.018)	2.62*** (0.018)	2.62*** (0.018)
Obs (n)	6664	6664	6664	6664

Notes. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Control reports the potential outcome means.

Table 3 Effect of International Student Roommate on Domestic Students: College Persistence and Global Outlook Measures

<i>Panel A. Average Treatment Effect (ATE)</i>			
	2nd Year Retention	6-year Graduation	Mean IP Grade
ATE	0.033 (0.021)	-0.002 (0.027)	0.102* (0.058)
Control	0.803*** (0.005)	0.638*** (0.006)	2.821*** (0.014)
<i>Panel B. Average Treatment Effect on the Treated (ATET)</i>			
	2nd Year Retention	6-year Graduation	Mean IP Grade
ATET	0.038* (0.020)	0.001 (0.026)	0.084 (0.059)
Control	0.795*** (0.006)	0.621*** (0.007)	2.764*** (0.020)
Obs (n)	6663	6664	4854

Notes. Robust standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. ATE and ATET expressed in percentage points. Control reports the potential outcome means.

College Persistence (Retention).

In line with Hypothesis 1B, the ATET estimate indicates a marginally statistically significant increase of 3.80 percentage points in 2nd year retention rates for treated domestic students ($p < 0.1$) relative to the control mean outcome of 79.50 percent (Table 3, Panel B). This points to a positive impact on persistence odds for domestic students directly exposed to an international roommate. However, the non-significant ATE estimate suggests these benefits may not extend more broadly when the roommate assignment is applied to the overall freshman class. While directly treated domestic students see gains, the domestic-international pairing mechanism does not appear to significantly improve average retention across all domestic students based on the ATE.

Global Outlook.

The positive and marginally significant ATE estimate for mean grades in international perspective courses provides some evidence supporting Hypothesis 4A. Domestic students paired with international roommates exhibit a higher level of global outlook

proxied by earning moderately higher IP course grades on average (Table 3, Panel A). The observed difference is an increase of 0.10 grade points ($p < 0.1$), to be exact.

Impact on International Students

Table 4 shows no statistically significant effects for international students with domestic student roommates, countering Hypothesis 2A that international students would achieve lower average grades from domestic roommate exposure. The ATE estimates across all four years show no significant difference. Similarly, the ATET estimates reveal no negative impacts on GPA relative to the untreated group. Regarding retention, neither the ATET nor ATE results indicate statistically significant change in 2nd year retention rates for international students relative to the potential-outcome mean of 98 percent (Table 5). This fails to support Hypothesis 2B that domestic roommate pairing lowers international students' odds of persistence.

Table 4 *Effects of Domestic Student Roommate on International Students: Multiyear GPA*

<i>Panel A. Average Treatment Effect (ATE)</i>				
	1st Year GPA	2nd Year GPA	3rd Year GPA	4th Year GPA
ATE	0.111 (0.088)	0.111 (0.077)	0.098 (0.069)	0.087 (0.068)
Control	3.346*** (0.085)	3.304*** (0.074)	3.289*** (0.066)	3.280*** (0.064)
<i>Panel B. Average Treatment Effect on the Treated (ATET)</i>				
	1st Year GPA	2nd Year GPA	3rd Year GPA	4th Year GPA
ATET	0.109 (0.089)	0.108 (0.078)	0.094 (0.070)	0.084 (0.069)
Control	3.348*** (0.086)	3.306*** (0.075)	3.291*** (0.066)	3.281*** (0.065)
Obs (n)	428	428	428	428

Notes. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Control reports the potential outcome means.

Table 5 Effect of Domestic Student Roommate on International Students: College Persistence and Global Outlook Measures

<i>Panel A. Average Treatment Effect (ATE)</i>			
	2nd Year Retention	6-year Graduation	Mean IP Grade
ATE	0.020 (0.019)	0.042 (0.037)	0.006 (0.120)
Control	0.980*** (0.019)	0.920*** (0.036)	3.406*** (0.115)
<i>Panel B. Average Treatment Effect on the Treated (ATET)</i>			
	2nd Year Retention	6-year Graduation	Mean IP Grade
ATET	0.021 (0.020)	0.042 (0.038)	0.007 (0.122)
Control	0.979*** (0.020)	0.920*** (0.037)	3.402*** (0.117)
Obs (n)	428	428	396

Notes. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. ATE and ATET expressed in percentage points. Control reports the potential outcome means.

The absence of any negative estimated effects on grades across years also provides no substantive evidence for Hypothesis 3B. We do not find declining magnitudes or significance over time as proposed for potential negative peer influences. Nonetheless, the non-significant estimates for performance in IP courses (Table 5) provide evidence for Hypothesis 4B that international students paired with domestic roommates during freshmen year have IP grades that are not statistically significantly different from those paired with another international student from a different country⁵.

Other Impacts

For both domestic and international students, no significant effects are found on 6-year graduation rates in either model (Table 3 and Table 5). This substantiates Hypothesis 5 that first-year intercultural roommate pairing had no lasting effect on eventual graduation outcomes. The null results on this distal outcome indicate that despite some initial benefits, particularly for domestic students, the effects of these

first-year roommate assignments did not persist to influence long-term outcomes like college completion.

It is notable that no significant impacts on retention are found beyond the second year for either domestic or international students. While not shown in the tables, models estimating treatment effects on retention in years three through five⁶ uniformly showed statistically insignificant results. This indicates the effects of freshman roommate assignments on persistence diminish quickly, with no lasting impact on retention to graduation. By the third year and beyond, the initial intercultural exposure appears insufficient to continue improving domestic students' retention or negatively influencing international students' persistence.

Robustness Checks

To assess the robustness of our main findings and to reduce model dependence, we compared the inverse probability weighting model results to estimates obtained using three alternative approaches: augmented inverse-probability weighting, propensity score matching, and nearest neighbor matching. The AIPW method, in particular, provides “doubly robust” estimates by leveraging both propensity score and outcome models (Huntington-Klein, 2021; Lunceford & Davidian, 2004), thereby overcoming IPW's reliance only on the propensity score model. Tables 6 to 8 present ATE and ATET estimates on key outcomes for domestic and international students across these models.

Table 6 Effects of International Student Roommate on Domestic Students: Multiyear GPA

Panel A. Average Treatment Effect (ATE)

	1st Year GPA	2nd Year GPA	3rd Year GPA	4th Year GPA
AIPW	0.136** (0.055)	0.091* (0.054)	0.081 (0.054)	0.074 (0.054)
IPW	0.136** (0.054)	0.090* (0.053)	0.081 (0.053)	0.073 (0.053)
NNM	0.136** (0.056)	0.092* (0.056)	0.083 (0.056)	0.075 (0.056)
PSM	0.139** (0.059)	0.095* (0.057)	0.086 (0.058)	0.078 (0.058)

Panel B. Average Treatment Effect on the Treated (ATET)

	1st Year GPA	2nd Year GPA	3rd Year GPA	4th Year GPA
IPW	0.149*** (0.049)	0.109** (0.048)	0.100** (0.048)	0.092* (0.048)
NNM	0.146*** (0.049)	0.105*** (0.049)	0.097** (0.048)	0.088* (0.048)
PSM	0.146*** (0.049)	0.105*** (0.049)	0.097** (0.048)	0.088* (0.048)

Obs (n)	6664	6664	6664	6664
---------	------	------	------	------

Notes. AIPW = Augmented Inverse Probability Weighting. IPW = Inverse Probability Weighting. NNM = Nearest Neighbor Matching. PSM = Propensity Score Matching. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7 Effects of Domestic Student Roommate on International Students' Grades:
Multiyear GPA

<i>Panel A. Average Treatment Effect (ATE)</i>				
	1st Year GPA	2nd Year GPA	3rd Year GPA	4th Year GPA
AIPW	0.110 (0.088)	0.110 (0.077)	0.096 (0.070)	0.086 (0.068)
IPW	0.111 (0.088)	0.111 (0.077)	0.098 (0.069)	0.087 (0.068)
NNM	0.116 (0.089)	0.115 (0.079)	0.101 (0.071)	0.091 (0.069)
PSM	0.116 (0.089)	0.115 (0.079)	0.101 (0.071)	0.091 (0.069)
<i>Panel B. Average Treatment Effect on the Treated (ATET)</i>				
	1st Year GPA	2nd Year GPA	3rd Year GPA	4th Year GPA
IPW	0.109 (0.089)	0.108 (0.078)	0.094 (0.070)	0.084 (0.069)
NNM	0.113 (0.091)	0.112 (0.080)	0.098 (0.072)	0.089 (0.070)
PSM	0.113 (0.091)	0.112 (0.080)	0.098 (0.072)	0.089 (0.070)
Obs (n)	428	428	428	428

Notes. AIPW = Augmented Inverse Probability Weighting. IPW = Inverse Probability Weighting. NNM = Nearest Neighbor Matching. PSM = Propensity Score Matching. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 8 College Persistence and Global Outlook Measures

	<u>Domestic Student Subsample</u>			<u>International Student Subsample</u>		
<i>Panel A. Average Treatment Effect (ATE)</i>						
	2nd Year Retention	6-year Graduation	Mean IP Grade	2nd Year Retention	6-year Graduation	Mean IP Grade
AIPW	0.034 (0.021)	-0.001 (0.027)	0.108* (0.059)	0.020 (0.020)	0.042 (0.038)	0.007 (0.121)
IPW	0.033 (0.021)	-0.002 (0.027)	0.102* (0.058)	0.020 (0.019)	0.042 (0.037)	0.006 (0.120)
NNM	0.034 (0.021)	0.000 (0.027)	0.106* (0.060)	0.020 (0.020)	0.047 (0.038)	0.006 (0.123)
PSM	0.036* (0.022)	0.002 (0.028)	0.111** (0.056)	0.020 (0.020)	0.047 (0.038)	0.006 (0.123)
<i>Panel B. Average Treatment Effect on the Treated (ATET)</i>						
	2nd Year Retention	6-year Graduation	Mean IP Grade	2nd Year Retention	6-year Graduation	Mean IP Grade
IPW	0.038* (0.020)	0.001 (0.026)	0.084 (0.059)	0.021 (0.020)	0.042 (0.038)	0.007 (0.122)
NNM	0.037* (0.020)	0.001 (0.026)	0.082 (0.059)	0.021 (0.021)	0.048 (0.039)	0.007 (0.126)
PSM	0.037* (0.020)	0.001 (0.026)	0.082 (0.059)	0.021 (0.021)	0.048 (0.039)	0.007 (0.126)
Obs (n)	6663	6664	4854	428	428	396

Notes. Robust standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. ATE and ATET expressed in percentage points. Control reports the potential outcome means.

For domestic students' GPA over four years (Table 6), the positive and significant ATE and ATET hold steady in magnitude and significance across all alternative models. The negligible differences in the point estimates across models lend further support to the finding that international roommate exposure improves domestic students' academic performance. Turning to international students' GPA (Table 7), the ATE and ATET estimates follow a similar pattern to the main IPW results. None of the coefficients are statistically significant under AIPW, NNM, and PSM approaches,

which combined with the remarkable similarities in the point estimates between models reaffirm the robustness of our findings.

For the college persistence and global outlook outcomes in Table 8, the estimates from alternative models largely align with the IPW results. One exception is the ATE for 2nd year retention among domestic students, which only shows marginal significance under PSM ($p < 0.1$). Given this sensitivity to model specification, the evidence for domestic student retention seems inconclusive. We, however, note that past research cautions against the use of the PSM approach given that it is prone to biased results (King & Nielsen, 2019).

The consistency demonstrated across these alternative modeling approaches serves to validate the reliability of our main IPW model findings. The fact that we observe consistent treatment effect estimates across all models suggests the core findings do not seem to be driven solely by the IPW model specifications. The alignment of results across different models with different methodological assumptions provides relatively greater confidence that the estimates reflect true differences arising from the intercultural roommate pairing, rather than biases associated with any single modeling approach. Thus, we can have greater confidence in the reported effects suggesting benefits of intercultural exposure through first-year roommate pairings for domestic students⁷, but no effects for their international student counterparts.

Discussion and Conclusion

This study provides unique evidence that the assignment of domestic and international students as freshman year roommates leads to a number of college benefits for domestic students at a college serving primarily economically disadvantaged students. While the estimated ATE on GPA of having an international roommate was a modest increase of between 0.09 and 0.14 grade points for domestic students, this seemingly small effect could translate into meaningful impacts for a substantial subset of students. As Gelman (2020) notes, one way to interpret this is that it is equivalent to a large GPA boost of between 0.9 and 1.4 points, out of 4.0, for 10 percent of students if the intervention had no effect on the other 90 percent. Therefore, the roommate intervention likely substantially benefited some domestic students even though the overall estimated effect across the population was small.

While the ATET model estimated a statistically significant four percentage point increase in 2nd year retention associated with international roommate exposure, the ATE for retention was not significant. This divergence suggests the retention benefits are concentrated among the treated domestic students rather than more broadly across the overall freshman cohort. At the same time, impacts across all measures are most visible in the first two years of college, suggesting that intercultural exposure during the initial transition to university may have provided the bulk of the benefits. While not closing achievement gaps completely, intentional integration via diverse rooming assignments appears to give domestic students an early boost. This is consistent with prior research where findings also showed academic peer effects that persisted at a diminishing rate into later college years (Carrell et al., 2009).

Our analysis indicates that purposeful exposure to cultural diversity and different academic practices through rooming with an international student can potentially accelerate domestic freshmen's college adjustment. This aligns with prior higher education research showing modest positive peer effects on college grades (Carrell et al., 2009; Lyle, 2007; Sacerdote, 2001; Stinebrickner & Stinebrickner, 2006; Zimmerman, 2003). Plausible mechanisms for these effects may include domestic students broadening their worldviews and improving academic expectations by learning from their international roommates' unique cultural backgrounds and study habits. Cross-cultural pairs may also facilitate greater openness, inclusiveness, and integration across the university's diverse student body. Domestic students with international roommates likely had more opportunities to engage with students outside their familiar contexts. Contemporary theories (Burdett & Crossman, 2012; Geelhoed et al., 2003; Pettigrew & Tropp, 2006; Sakurai et al., 2010; Shook & Clay, 2012; Sias et al., 2008) could explain the modest spill-over improvements observed more widely among the college's freshman population, as seen in the ATE estimates.

These findings carry important practical implications for an institution like Berea College and others focused on supporting students from lower socioeconomic backgrounds. Intercultural roommate pairing is one avenue for facilitating positive peer influences and accelerated learning, especially for domestic students during their critical transition and adjustment to college. Our findings extend past studies such as Dynarski et al. (2021) whose work concluded "a low-cost, low-touch intervention can

strongly affect student application and enrollment at selective colleges.” We demonstrated an equally low-cost and low-touch roommate pairing intervention that can effectively boost students’ early college outcomes at a selective liberal arts college. Moreover, our current study complements and extends the earlier Berea College studies by examining a novel question of growing importance in higher education: the cross-cultural peer dynamics between international and domestic students. While built on the methodological foundations of earlier studies, we take a new estimation approach and utilize a multi-year institutional dataset to analyze this critical but understudied area. This intentional mixing of student subgroups is worth considering for higher education administrators and policymakers as an affordable strategy for cultivating an inclusive educational environment and harnessing demographic diversity for academic gains.

When reading our findings alongside research on culturally engaging campus environments like Museus et al. (2018), our results point to the need for a more comprehensive approach to optimize the benefits of intercultural exposure. While rooming with international students provided some early academic boosts for domestic students, these gains diminished over time. This indicates that the cultural familiarity and validation from this experience alone was insufficient to support persistence through college graduation and retention in later years. In addition, we must develop strategies to ensure meaningful benefits for international students as well. Our current study offers preliminary evidence that diversity initiatives like intercultural roommate pairing can positively impact disadvantaged students when intentionally designed. Yet, work remains to create equitable and empowering campus environments that fully leverage cultural diversity as an asset. Our findings provide insights into one piece of this complex puzzle; further research is needed to develop comprehensive solutions that support *all* students from their first year through graduation.

While we highlight the contributions of the paper, we also note some of its limitations. First, we relied entirely on weighting and matching methods to correct threats to randomization and to attempt causal inference. While these techniques can help minimize selection bias, findings in this paper are not completely causal given the inability to truly randomly assign all students. We do, however, achieve balance across the covariates in our estimation models, making it more plausible that there would be

balance across the unobserved covariates. It is important to note that unobserved differences between the treated and control groups may still confound the results. In particular, we have access to only a limited set of covariates to conduct our analysis, including sex at birth, transfer student status, as well as race and ethnicity. The precise characteristics of international students are also unknown, e.g., the region of the world from which they originate. With few baseline characteristics, our ability to achieve true balanced treated and control groups is constrained. More robust covariate data on factors such as parental education and pre-college experiences could enable better-matched samples and strengthen causal conclusions. Additional limitations include the focus on a single institution and reliance on administrative data.

Despite these limitations, the consistency of results across multiple model specifications provides some reassurance for the direction and significance of our findings. We suggest that this research be used as starting evidence of the benefits of intercultural roommate pairing, which can then be analyzed with multi-site randomized trials. In future studies, researchers must also qualitatively uncover how students conceive of peer effects mechanisms work in international-domestic roommate pairs. Some questions may include: Why do we see such effects for domestic students but not for international students during the early years in college? How exactly do low-income students in the U.S. experience and perceive value from intercultural engagement on campus? What forms of support might further empower them to capitalize on diversity?

From a practical standpoint, schools may understand how the collaboration and close connection between international and domestic students can promote not only positive individual outcomes but also collective outcomes. One of the ways for this to happen is for schools to capture students' ideas, insights, and voices. University administration could consider conducting interviews, focus groups, or ethnographic research to capture student voices. Such lived experiences would enable more targeted enhancements to policy and practice. Overall, this study provides initial evidence that diversity and cultural integration initiatives have the potential to support disadvantaged students' learning and success. It paves the way for further research to better understand how to enact such high-impact practices meaningfully and equitably.

Data Availability Statement

The data that support the findings of this study are available from Berea College. Restrictions apply to the availability of these data, which were used under license for this study. The sharing of the original data must comply with Institutional Review Board (IRB) regulations. Researchers seeking access to data from an institution must obtain approval from that institution's IRB. IRB applications should be submitted directly to the institution housing the requested data (i.e., Berea College in our case).

Declaration of Interest Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The authors report there are no competing interests to declare.

Endnotes

¹ Manski (1993) terms this the “reflection problem,” where the peers simultaneously influence each other.

² Pell Grant is a U.S. federal need-based grant for undergraduate students with exceptional financial need. For more information, please visit studentaid.gov.

³ The reference group here is the international students rooming with other international students.

⁴ For example, students can take classes in foreign languages, different religions, art, or the history of a country. There is a requirement for one of the two IP classes to be non-Western.

⁵ This suggests that the effect is more plausibly attributed to international exposure. From a non-US student perspective, exposure to a US roommate appears to have neither positive nor negative effects compared to exposure to students from other countries. That is “Joe and Zhou” pair is no different from “Ahmad and Zhou” pair.

⁶ We estimate year 5 to account for certain majors that take more than 4 years to complete, e.g., nursing.

⁷ Within the subgroup of international students, the absence of significant differences in roommate experiences between those paired with a US student and those paired with a student of another nationality helps alleviate, to some extent, concerns about comparing US students with comparably low-income international students presumed to have higher academic abilities. If the international students in our sample are inherently more academically prepared relative to the US students, we would expect variations in outcomes between an international student rooming with a domestic US student and one rooming with another international student (see, for example, Sacerdote (2001) and Zimmerman (2003) for peer effect differences based on ability grouping). However, such differentiation is not evident in our findings. While acknowledging the lack of pre-college data to fully account for students’ abilities, the absence of outcome differentiation between the groups in the international student subsample further validates our findings. This is also consistent with prior work at Berea College by Stinebrickner & Stinebrickner (2006), indicating the fixed nature of students’ ability in the short run.

References

- Bacon, D. R., & Bean, B. (2006). GPA in Research Studies: An Invaluable but Neglected Opportunity. *Journal of Marketing Education*, 28(1), 35–42. <https://doi.org/10.1177/0273475305284638>
- Bailey, D., Duncan, G. J., Odgers, C. L., & Yu, W. (2017). Persistence and Fadeout in the Impacts of Child and Adolescent Interventions. *Journal of Research on Educational Effectiveness*, 10(1), 7–39. <https://doi.org/10.1080/19345747.2016.1232459>
- Beatty, A. S., Walmsley, P. T., Sackett, P. R., Kuncel, N. R., & Koch, A. J. (2015). The Reliability of College Grades. *Educational Measurement: Issues and Practice*, 34(4), 31–40. <https://doi.org/10.1111/emip.12096>
- Boisjoly, J., Duncan, G. J., Kremer, M., Levy, D. M., & Eccles, J. (2006). Empathy or Antipathy? The Impact of Diversity. *American Economic Review*, 96(5), 1890–1905. <https://doi.org/10.1257/aer.96.5.1890>
- Bowman, N. A. (2010). College Diversity Experiences and Cognitive Development: A Meta-Analysis. *Review of Educational Research*, 80(1), 4–33. <https://doi.org/10.3102/0034654309352495>
- Brady, R. R., Insler, M. A., & Rahman, A. S. (2017). Bad Company: Understanding negative peer effects in college achievement. *European Economic Review*, 98, 144–168. <https://doi.org/10.1016/j.euroecorev.2017.06.013>
- Braunstein, A., McGrath, M., & Pescatrice, D. (2000). Measuring the Impact of Financial Factors on College Persistence. *Journal of College Student Retention: Research, Theory & Practice*, 2(3), 191–203. <https://doi.org/10.2190/0TTM-U8RA-V8FX-FYVA>
- Brunello, G., de Paola, M., & Scoppa, V. (2010). Peer Effects in Higher Education: Does the Field of Study Matter? *Economic Inquiry*. <https://doi.org/10.1111/j.1465-7295.2009.00235.x>
- Burdett, J., & Crossman, J. (2012). Engaging international students. *Quality Assurance in Education*, 20(3), 207–222. <https://doi.org/10.1108/09684881211240286>
- Card, D., & Solis, A. (2022). Measuring the Effect of Student Loans on College Persistence. *Education Finance and Policy*, 17(2), 335–366. https://doi.org/10.1162/edfp_a_00342
- Carrell, S. E., Fullerton, R. L., & West, J. E. (2009). Does Your Cohort Matter? Measuring Peer Effects in College Achievement. *Journal of Labor Economics*, 27(3), 439–464. <https://doi.org/10.1086/600143>

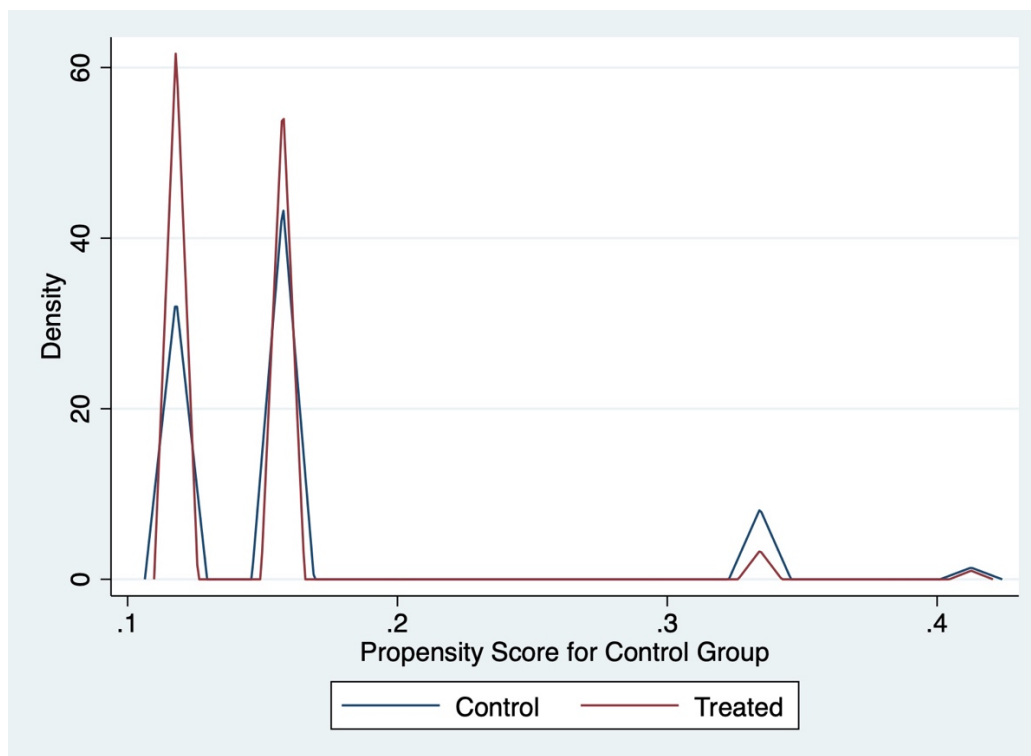
- Carrell, S. E., Hoekstra, M., & West, J. E. (2011). Is poor fitness contagious? *Journal of Public Economics*, 95(7–8), 657–663. <https://doi.org/10.1016/j.jpubeco.2010.12.005>
- Carrell, S. E., Malmstrom, F. v., & West, J. E. (2008). Peer Effects in Academic Cheating. *Journal of Human Resources*, 43(1), 173–207. <https://doi.org/10.3368/jhr.43.1.173>
- Carrell, S. E., Sacerdote, B. I., & West, J. E. (2013). From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation. *Econometrica*, 81(3), 855–882. <https://doi.org/10.3982/ECTA10168>
- Chesnaye, N. C., Stel, V. S., Tripepi, G., Dekker, F. W., Fu, E. L., Zoccali, C., & Jager, K. J. (2022). An introduction to inverse probability of treatment weighting in observational research. *Clinical Kidney Journal*, 15(1), 14–20. <https://doi.org/10.1093/ckj/sfab158>
- Duncan, G. J., Boisjoly, J., Kremer, M., Levy, D. M., & Eccles, J. (2005). Peer Effects in Drug Use and Sex Among College Students. *Journal of Abnormal Child Psychology*, 33(3), 375–385. <https://doi.org/10.1007/s10802-005-3576-2>
- Dynarski, S., Libassi, C., Michelmore, K., & Owen, S. (2021). Closing the Gap: The Effect of Reducing Complexity and Uncertainty in College Pricing on the Choices of Low-Income Students. *American Economic Review*, 111(6), 1721–1756. <https://doi.org/10.1257/aer.20200451>
- Foster, G. (2006). It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism. *Journal of Public Economics*, 90(8–9), 1455–1475. <https://doi.org/10.1016/j.jpubeco.2005.12.001>
- Geelhoed, R. J., Abe, J., & Talbot, D. M. (2003). A Qualitative Investigation of U.S. Students' Experiences in an International Peer Program. *Journal of College Student Development*, 44(1), 5–17. <https://doi.org/10.1353/csd.2003.0004>
- Gelman, A. (2020). *Understanding the “average treatment effect” number*. Statistical Modeling, Causal Inference, and Social Science. <https://statmodeling.stat.columbia.edu/2020/06/30/ate/>
- Glenn, M. E. (1970). *Appalachia in transition* [Book]. Bethany Press.
- Gurin, P., Dey, E., Hurtado, S., & Gurin, G. (2002). Diversity and Higher Education: Theory and Impact on Educational Outcomes. *Harvard Educational Review*, 72(3), 330–367. <https://doi.org/10.17763/haer.72.3.01151786u134n051>
- Huntington-Klein, N. (2021). *The Effect: An Introduction to Research Design and Causality*. Chapman and Hall/CRC. <https://doi.org/10.1201/9781003226055>
- King, G., & Nielsen, R. (2019). Why Propensity Scores Should Not Be Used for Matching. *Political Analysis*, 27(4), 435–454. <https://doi.org/10.1017/pan.2019.11>

- Kremer, M., & Levy, D. (2008). Peer Effects and Alcohol Use among College Students. *Journal of Economic Perspectives*, 22(3), 189–206. <https://doi.org/10.1257/jep.22.3.189>
- Lunceford, J. K., & Davidian, M. (2004). Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *Statistics in Medicine*, 23(19), 2937–2960. <https://doi.org/10.1002/sim.1903>
- Lyle, D. S. (2007). *Estimating and Interpreting Peer and Role Model Effects from Randomly Assigned Social Groups at West Point*. 89(2), 289–299.
- Lyle, D. S. (2009). The Effects of Peer Group Heterogeneity on the Production of Human Capital at West Point. *American Economic Journal: Applied Economics*, 1(4), 69–84.
- Mamiseishvili, K. (2012). International student persistence in U.S. postsecondary institutions. *Higher Education*, 64(1), 1–17. <https://doi.org/10.1007/s10734-011-9477-0>
- Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, 60(3), 531–542. <https://doi.org/10.2307/2298123>
- Marmaros, D., & Sacerdote, B. (2002). Peer and social networks in job search. *European Economic Review*, 46(4–5), 870–879. [https://doi.org/10.1016/S0014-2921\(01\)00221-5](https://doi.org/10.1016/S0014-2921(01)00221-5)
- McEwan, P. J., & Soderberg, K. A. (2006). Roommate Effects on Grades: Evidence from First-Year Housing Assignments. *Research in Higher Education*, 47(3), 347–370. <https://doi.org/10.1007/s11162-005-9392-2>
- Mehta, N., Stinebrickner, R., & Stinebrickner, T. (2019). Time-Use and Academic Peer Effects in College. *Economic Inquiry*, 57(1), 162–171. <https://doi.org/10.1111/ecin.12730>
- Museus, S. D., Yi, V., & Saelua, N. (2018). How culturally engaging campus environments influence sense of belonging in college: An examination of differences between White students and students of color. *Journal of Diversity in Higher Education*, 11(4), 467–483. <https://doi.org/10.1037/dhe0000069>
- Neighbors, C., Lee, C. M., Lewis, M. A., Fossos, N., & Larimer, M. E. (2007). Are Social Norms the Best Predictor of Outcomes Among Heavy-Drinking College Students? *Journal of Studies on Alcohol and Drugs*, 68(4), 556–565. <https://doi.org/10.15288/jsad.2007.68.556>
- Office of Institutional Research and Assessment. (2014). *Berea College Fact Book 2013-2014*.
- Office of Institutional Research and Assessment. (2015). *Berea College Fact Book 2014-2015*.
- Office of Institutional Research and Assessment. (2022). *Berea College Fact Book*.

- Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, 90(5), 751–783. <https://doi.org/10.1037/0022-3514.90.5.751>
- Ross, L., & Nisbett, R. E. (2011). *The Person and the Situation: Perspectives of Social Psychology* (2nd ed.). Pinter & Martin.
- Sacerdote, B. (2001). Peer Effects with Random Assignment: Results for Dartmouth Roommates. *The Quarterly Journal of Economics*, 116(2), 681–704. <https://doi.org/10.1162/00335530151144131>
- Sacerdote, B. (2014). Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward? *Annual Review of Economics*, 6(1), 253–272. <https://doi.org/10.1146/annurev-economics-071813-104217>
- Sakurai, T., McCall-Wolf, F., & Kashima, E. S. (2010). Building intercultural links: The impact of a multicultural intervention programme on social ties of international students in Australia. *International Journal of Intercultural Relations*, 34(2), 176–185. <https://doi.org/10.1016/j.ijintrel.2009.11.002>
- Shook, N. J., & Clay, R. (2012). Interracial roommate relationships: A mechanism for promoting sense of belonging at university and academic performance. *Journal of Experimental Social Psychology*, 48(5), 1168–1172. <https://doi.org/10.1016/j.jesp.2012.05.005>
- Sias, P. M., Drzewiecka, J. A., Meares, M., Bent, R., Konomi, Y., Ortega, M., & White, C. (2008). Intercultural Friendship Development. *Communication Reports*, 21(1), 1–13. <https://doi.org/10.1080/08934210701643750>
- Stinebrickner, R., & Stinebrickner, T. R. (2006). What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds. *Journal of Public Economics*, 90(8–9), 1435–1454. <https://doi.org/10.1016/j.jpubeco.2006.03.002>
- Stinebrickner, R., & Stinebrickner, T. R. (2008). The Causal Effect of Studying on Academic Performance. *The B.E. Journal of Economic Analysis & Policy*, 8(1). <https://doi.org/10.2202/1935-1682.1868>
- Stuart, E. A. (2010). Matching Methods for Causal Inference: A Review and a Look Forward. *Statistical Science*, 25(1). <https://doi.org/10.1214/09-STS313>
- Tienda, M. (2013). Diversity ≠ Inclusion. *Educational Researcher*, 42(9), 467–475. <https://doi.org/10.3102/0013189X13516164>
- Vazzana, C. M., & Rudi-Polloshka, J. (2019). Appalachia Has Got Talent, But Why Does It Flow Away? A Study on the Determinants of Brain Drain From Rural USA. *Economic Development Quarterly*, 33(3), 220–233. <https://doi.org/10.1177/0891242419844320>

- Voss, P. R., Long, D. D., Hammer, R. B., & Friedman, S. (2006). County child poverty rates in the US: a spatial regression approach. *Population Research and Policy Review*, 25(4). <https://doi.org/10.1007/s11113-006-9007-4>
- Winston, G., & Zimmerman, D. (2004). Peer Effects in Higher Education. In C. M. Hoxby (Ed.), *College Choices: The Economics of Where to Go, When to Go, and How to Pay for It* (1st ed., pp. 395–424). University of Chicago Press.
- Wolniak, G., & Ballerini, V. (2020). Peer Effects, Higher Education. In P. N. Teixeira & J. C. Shin (Eds.), *The International Encyclopedia of Higher Education Systems and Institutions*. Springer. https://doi.org/10.1007/978-94-017-8905-9_84
- Zhao, C.-M., Kuh, G. D., & Carini, R. M. (2005). A Comparison of International Student and American Student Engagement in Effective Educational Practices. *Source: The Journal of Higher Education*, 76(2), 209–231.
- Zimmerman, D. J. (2003). Peer Effects in Academic Outcomes: Evidence from a Natural Experiment. *Review of Economics and Statistics*, 85(1), 9–23. <https://doi.org/10.1162/003465303762687677>

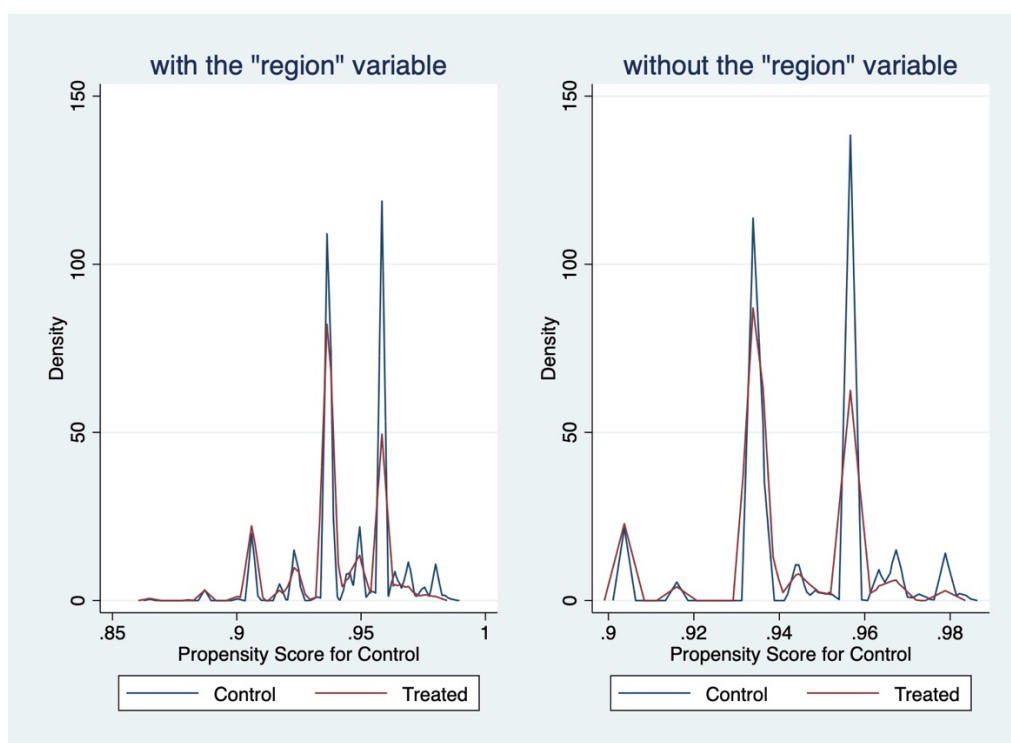
Appendix A



Note. Propensity Score Overlap Plot for the Treatment and Control Groups: International Student Subsample.

The graph displays the estimated density of the predicted probabilities that an untreated international student is indeed untreated and the estimated density of the predicted probabilities that a treated international student is untreated. The propensity score density plots for the treatment and control groups do not show excessive mass concentrated near the extremes of 0 or 1. Additionally, the main distribution mass for each group's density curve lies in areas where there is considerable overlap with the other group's density. Therefore, based on visually examining the estimated propensity score densities, there is no clear evidence to suggest the overlap assumption has been violated, since the plots demonstrate substantial common support between the two groups.

Appendix B



Note. Propensity Score Overlap Plots for the Treatment and Control Groups with and without the Region Variable: Domestic Student Subsample.

The results of the two overlap/common support plot assessment indicate no clear evidence that the overlap assumption has been violated. Formal covariate balance test results using STATA also indicate that reasonable balance was achieved for the observed covariates through inverse probability weighting. However, it is important to note that these diagnostics are only sensitive to imbalance in the observed covariates. It is possible that there is imbalance in unobserved covariates, which could lead to bias in the results.

Furthermore, while the balance diagnostics are reassuring, the fact that the propensity score distribution is quite narrow (i.e., 0.9 to 0.98 in the plot without the region variable) suggests the propensity score model may not be capturing all factors influencing treatment assignment. This limited discrimination between treatment and control groups based on the available covariates implies there may be unobserved confounders that are imbalanced between groups. To assess robustness to potential unmeasured confounding, we

conducted a sensitivity analysis to check for robustness using NNM, PSM, and AIPW models. Reassuringly, these alternative approaches yielded treatment effect estimates very similar to the original IPW results. The consistency in estimated effects across multiple models provides greater assurance that the conclusions are not highly sensitive to model specification or unobserved biases.

Therefore, that is, our observed covariate diagnostics, in terms of balance and overlap, are generally positive, and robustness checks also demonstrate consistency across models. This provides increased confidence in making causal inferences based on the IPW results while acknowledging the potential for unobserved biases.