Backlash or Bonding: Effect of an Integration Policy in the Context of Affirmative Action*

Balasai Vanukuri†

August 10, 2026

Latest version: Click here

Abstract

This paper examines how intergroup contact in elite higher education affects academic performance, psychological well-being, and political attitudes. Using randomized first-year roommate assignments at an elite engineering university in India, I exploit exogenous variation in cross-caste exposure to estimate its causal effects on grades, internship placement, mental health, and attitudes toward affirmative action. Lower-caste students exposed to upper-caste roommates experience improvements in academic performance and internship placement and become less likely to view affirmative action as unfair to upper-caste students, but they also incur mental health costs. Upper-caste students experience declines in academic performance and modestly worse mental health, alongside increased support for affirmative action.

I develop a model that identifies two mechanisms underlying these effects. First, roommate exposure provides information about whether differences in entrance-exam performance reflect merit or circumstances, leading students to update beliefs

*I am grateful to my supervisors Karen Macours and Oliver Vanden Eynde for their invaluable advice and constant support, and to Liam Wren-Lewis and Suanna Oh for their invaluable feedback throughout the project, and to Abhijit Banerjee and Frank Schilbach for their guidance during my visit at MIT and their continued support thereafter. I also thank David Atkin, Luc Behaghel, Anwesha Bhattacharya, Emily Brezza, Denis Cogneau, Andrea Cornejo, Denis Cogneau, Oscar Diaz, Kyle Emerick, Marcello Gantier, Julien Grenet, Nina Guyon, Eliana La Ferrara, Sylvie Lambert, Paola Mallia, David Margolis, Ben Olken, Hector Paredes, Nishith Prakash, Vrinda Sharma, Manpreet Singh, Anmol Somanchi, Akiko Suwa-Eisenmann and seminar participants at PSE, MIT, Tufts, North Eastern university for helpful comments and feedback. Aditya Gupta, Ananya, Vikash Meena, and Divyanshu Verma were invaluable in coordinating the fieldwork. I also thank Sumit Parashar and Quest for their invaluable support in managing survey logistics. This work was generously supported by EUR-PgSE, CEPREMAP, the Chaire PEMS at PSE, the Development Research Group at PSE, and the Weiss Fund for Research in Development Economics. This study was pre-registered in the AEA RCT Registry under the unique identifying number AEARCTR-0013944. The study was approved by the Institutional Review Board at the Paris School of Economics. All errors are my own.

†Paris School of Economics. Email: balasai.vanukuri@psemail.eu.

about the determinants of achievement. Second, caste distance creates social frictions and induces an asymmetric substitution between roommate interactions and student-club participation, particularly for lower-caste students. Consistent with these mechanisms, greater caste distance worsens mental health for both groups and reduces academic performance among upper-caste students, while also increasing support for affirmative action. Academic assistance, in contrast, flows primarily from higher- to lower-performing students. Variation in room size and differences in entrance-exam performance provides further evidence for these mechanisms. The findings reveal that integration can generate economic gains and belief change while imposing asymmetric psychological costs, and suggest that larger rooms, safe spaces, and reductions in non-academic distance can improve the effectiveness of integration policies.

Keywords: Affirmative action; caste integration; intergroup contact; academic outcomes; mental health
JEL Classification: I23; J15; Z13

1 Introduction

India has a long history of running one of the world's most extensive affirmative action (AA) programs, implemented primarily through caste-based quotas in public universities, government employment, and political representation. A substantial body of evidence demonstrates that these quotas have not only increased representation of disadvantaged groups (Deshpande, 2013) but also delivered benefits to recipients through educational attainment (Bagde et al., 2016; Bertrand et al., 2010; Cassan, 2019), earnings, and political empowerment (Chattopadhyay and Duflo, 2004; Gulzar et al., 2020; Pande, 2003). However, they may impose costs on relatively advantaged groups. Beyond these distributional effects, such costs may also induce behavioral responses like discrimination (Glover et al., 2017; Hjort, 2014) and stereotype threat (Hoff and Pandey, 2006) that undermine policy effectiveness. These behavioral frictions could be particularly pronounced in highly competitive settings such as admission to elite universities and coveted public sector jobs, especially when there are substantial premiums for attending elite institutions (Sekhri, 2020).

A growing body of evidence suggests that integration policies, which facilitate direct contact between groups, can reduce prejudice and improve intergroup relations under certain conditions (Allport, 1954; Pettigrew and Tropp, 2006).[1] However, the efficacy of

[1] A wide range of interventions have been shown to reduce bias against minorities, including leadership

such policies in highly competitive environments remains theoretically ambiguous. On one hand, disadvantaged students may benefit from studying collaboratively with advantaged peers and gaining access to their social networks and cultural capital, while advantaged students may reduce misperceptions and stereotypes about disadvantaged groups.[23] On the other hand, when students compete intensely for scarce, high-stakes resources—such as prestigious internships, leadership positions in student organizations, and top academic rankings—intergroup contact may trigger competitive threat rather than cooperation (Lowe, 2021; Tabellini, 2020). Under such conditions, diversity could exacerbate mental health costs, reinforce negative stereotypes, and reduce support for affirmative action policies, leading to "negative contact" (Festinger, 1954; Schäfer et al., 2021).

This paper examines one such integration policy—roommate assignment aimed at increasing caste diversity—and addresses three key questions: First, what is the effect of a roommate's caste identity on academic performance and job market outcomes? Second, do students face a trade-off between performance and mental health due to increased diversity? Third, do social interactions with students from different castes change attitudes toward affirmative action policies and caste-based stereotypes, particularly among advantaged caste students? Understanding the efficacy of integration policies has consequences not only for mitigating behavioral frictions related to affirmative action, but also for creating socially cohesive communities. This is particularly relevant in Indian society, which remains deeply divided along caste lines, with potential implications for economic interactions and market efficiency across caste groups (Munshi, 2019).

I exploit a natural experiment in which first-year students at an elite public engineering university are randomly assigned to hostel rooms, generating plausibly exogenous

and role models (Beaman et al., 2009; Breda et al., 2023), intergroup contact (Boisjoly et al., 2006; Corno et al., 2022; Rao, 2019), and sociocognitive debiasing (Alan et al., 2021 or through norms-based persuationWebb, 2025).

[2]A large number of studies examine inter-ethnic contact through students as roommates (Boisjoly et al., 2006; Carrell et al., 2019; Corno et al., 2022) or classmates (Rao, 2019); see Lowe, 2025 for a meta-analysis on effects of contact experiments.

[3]Two conditions required by Allport's contact theory are contentious in this setting. First, groups must enjoy equal status in the relationship, but caste hierarchies pose a challenge to this assumption. Second, groups must benefit from cooperation. Inside the university, there is no explicit competition among groups and they could potentially benefit from cooperation. However, the elite university setting with competition for high-stakes jobs and perceived losses from affirmative action policies poses challenges to this condition. Two other assumptions are more likely to be applicable in the university setting: sharing common goals and interactions supported by a commonly recognized authority. For the latter, this integration is socially engineered, as the university follows a complex algorithm to ensure mixing occurs, and this algorithm also explicitly stratifies by caste when allocating students to hostels. Qualitative interviews with hostel administrators revealed they do not entertain room changes, particularly when they are suspicious of changes that would increase caste-based segregation.

variation in exposure to roommates from different caste backgrounds[4]. I provide two checks to support the random nature of roommate allocation. First, I show that the observed proportion of mixed-caste rooms closely matches the proportion predicted under random assignment. Second, I document balance on observable socioeconomic characteristics and pre-college attributes across mixed and same-caste rooms. My empirical strategy compares outcomes for students assigned to mixed-caste rooms (the "treatment" group) with those assigned to same-caste rooms (the "control" group). I estimate effects separately for advantaged (upper-caste) and disadvantaged (lower-caste) students, and also report results for the pooled sample. My main specification includes hostel-by-cohort fixed effects and specialization-by-cohort fixed effects.

I collected original survey data on 2,400 students across two cohorts, which I complement with administrative records on roommate assignments, entrance exam scores, and caste status to construct treatment measures. My primary outcomes—academic performance, internship placements, attitudes toward affirmative action, and mental health—are measured through the survey.

First I document key differences between upper and lower caste students in the sample. First, students in the sample differ sharply along socioeconomic, academic, and attitudinal dimensions at entry. Upper-caste students come from substantially more advantaged family backgrounds: they are more likely to have highly educated parents, prior exposure to elite engineering institutions through social networks, and greater household income and assets, and are less likely to rely on education loans. Second, consistent with these advantages, upper-caste students score markedly higher on the college entrance examination, with a gap of roughly one standard deviation that persists within academic specializations. These differences do not appear to reflect lower effort among lower-caste students, who report similar study hours, comparable prior academic performance in school exams, and similar spending on coaching for entrance exam prep. Third lower-caste students enter college with modestly worse mental health, driven by greater anxiety, depression, and concerns about cultural fit and discrimination, while upper-caste students report higher academic and time-management stress. Finally, the two groups hold sharply divergent views on affirmative action: upper-caste students express substantially stronger meritocratic concerns and weaker support for the social justice and diversity rationales underlying reservation policies. Together, these baseline differences

[4]This study doesn't exploit variation in affirmative action policies, instead relies on affirmative action policy to generate sufficient representation of disadvantaged groups, and hence can generate inter-caste contact. In the absence of affirmative action policies, representation of disadvantaged students would be very low in these elite colleges. In the Indian context, affirmative action is through fixed quotas for different groups, primarily caste, in each specialization in all public universities.

underscore the highly stratified social and academic environment into which students are randomly assigned roommates[5].

In the first set of results, I document that assignment to mixed-caste rooms has substantially different effects on lower- and upper-caste students' academic performance and internship outcomes. For lower-caste students, exposure to upper-caste roommates improves academic and early labor market outcomes: overall performance increases significantly by 0.132 s.d., with positive point estimates across GPA (0.091 s.d., $p < 0.1$), math and programming grades (0.104 s.d, $p < 0.1$), and a statistically significant increase in completion of summer internships or academic projects (0.065 percentage points). In contrast, upper-caste students assigned to mixed rooms experience a decline in academic performance, driven by a 0.103 s.d. reduction in GPA, while effects on math and programming grades (0.054 s.d.) are not significant and effect on internship outcome is precise null. When pooling students across caste groups, these offsetting responses produce no statistically detectable difference between mixed and same-caste rooms.

In the second set of results, I document effects on students' mental health and attitudes toward affirmative action, which provide insight into how roommate exposure shapes well-being and views on key policy issues. I find that lower-caste students assigned to mixed rooms experience worse mental health, with the composite mental health index declining by 0.081 s.d. ($p < 0.1$). This index aggregates clinical measures of anxiety and depression (PHQ-4), subjective well-being relative to the pre-college period, and college-environment-related worries. The deterioration is driven primarily by higher PHQ-4 scores and increased worry about college issues among lower-caste students, primarily discrimination and cultural adaptation. In contrast, assignment to mixed rooms has no statistically significant effect on overall mental health for upper-caste students (0.054 s.d.), though the point estimate is negative. However, the effect on anxiety and depression is positive and comparable in magnitude to that for lower-caste students: PHQ-4 increases by 0.090 s.d. for upper-caste students. Upper-caste students are also more likely to report worries related to time management. When pooling students across caste groups, the effect of mixed rooms on the mental health index is 0.079 s.d. and 0.096 s.d. on the PHQ-4 index, and both effects are statistically significant, indicating that mixed-caste exposure entails mental health costs that are borne by both groups, though the channels through which these costs arise differ as indicated by components of the worry index.

In the third set of results, I examine effects on attitudes toward affirmative action policy. For lower-caste students, mixed-room exposure has no statistically significant effect on overall support for affirmative action; the point estimate is 0.046 s.d. and is not statisti-

[5]Corno et al., 2022 also study racial integration in a similar

cally different from zero. Attitudes toward the benefits of affirmative action for their own groups do not improve in mixed rooms. However, lower-caste students become significantly more likely to disagree with perceived costs of affirmative action for non-reserved groups (0.11 s.d.). Upper-caste students assigned to mixed rooms exhibit greater overall support for affirmative action, with an effect of 0.088 s.d. ($p < 0.10$), primarily driven by a significant increase in agreement regarding the benefits of affirmative action for reserved groups (0.115 s.d.). Upper-caste students in mixed rooms are no more likely to disagree with the perceived costs of affirmative action for non-reserved groups, although component-level results indicate they are less likely to agree that affirmative action compromises academic quality (0.142 s.d., $p < 0.10$).

To understand the mechanisms underlying the asymmetric effects on academic performance, internship outcomes, mental health, and attitudes toward affirmative action, I exploit three additional sources of variation. The baseline results reveal that mixed-caste exposure generates substantial academic gains for lower-caste students but mental health costs for both groups, political attitude shifts, primarily about costs and benefits of AA policies to the other group - yet the channels driving these effects remain unclear. Disentangling "standard" ability-based peer effects from identity-based frictions, and distinguishing between learning spillovers and relative deprivation, requires examining how effects vary with the specific structure of inter-caste contact.

First, I exploit heterogeneity in entrance exam score gaps with roommates to isolate ability-based peer learning from caste-mediated costs. If academic gains for lower-caste students operate purely through exposure to higher-ability peers, effects should increase monotonically with roommate performance in both same-caste and mixed-caste rooms. If mental health costs for upper-caste students reflect exposure to lower-performing peers admitted through affirmative action, academic and psychological costs should increase as the score gap widens. Second, I examine variation in the intensity and structure of mixed-caste exposure—comparing partially mixed rooms (where students share with both same-caste and opposite-caste peers) to fully mixed rooms, and comparing two-person rooms (intensive dyadic contact) to three-person rooms (broader exposure with potential for social buffering). This variation reveals whether benefits require diverse peer groups or operate through dyadic interactions, and whether psychological costs stem from unavoidable intensive contact or scale with the number of opposite-caste roommates. Third, I estimate effects using continuous measures of caste distance based on traditional hierarchical rankings while controlling for roommate score differences, and then decompose which specific caste pairings drive academic, psychological, and attitudinal effects. This approach distinguishes identity-based interaction frictions that scale with social distance

from costs linked to affirmative action policies and relative performance gaps.

First, I examine heterogeneity along the direction of entrance exam score gaps with roommates. I calculate the difference between each roommate's score and the student's score, average this across roommates, and then standardize with respect to the control group[6]. Among lower-caste students, larger score gaps predict significantly better academic outcomes in same-caste rooms, consistent with peer learning. However, these gains are substantially attenuated in mixed-caste rooms: the interaction between score distance and mixed-room assignment is negative and significant at the 5% level for both overall performance and GPA. However, internship outcomes tell a different story—they improve in mixed-caste rooms independently of roommate ability, suggesting that lower-caste students benefit from access to upper-caste social networks and information rather than pure ability spillovers. For upper-caste students, neither score gap nor its interaction with mixed-room assignment predicts academic performance, ruling out the possibility that their academic declines are driven by exposure to lower-performing roommates. Instead, the evidence points to caste-mediated interaction costs that operate independently of ability differences.

Mental health costs and change in attitudes of AA policies reveal distinct mechanisms for each group. Lower-caste students experience two types of psychological costs: identity-linked stress that operates in mixed rooms regardless of score gaps (reflected in elevated PHQ-4 scores), and competitive stress that increases with exposure to higher-performing upper-caste roommates (reflected in heightened worry). For upper-caste students, mental health deteriorates and support for affirmative action declines as the score gap with lower-caste roommates widens, consistent with relative deprivation: upper-caste students perceive unfairness when observing substantially lower-scoring peers gain admission to the same or better specializations through reservation policies. As a robustness check, I verify that these patterns depend on the direction rather than magnitude of score gaps. Using absolute score differences eliminates all heterogeneity effects on academic performance, mental health, and affirmative action attitudes, confirming that relative positioning—not simply performance heterogeneity—drives the observed patterns. Absolute gap matters only for relative well-being: lower-caste students report worse well-being when gap with higher caste roommates is higher, while upper-caste students report better well-being when gap with lower caste roommates is higher, underscoring the salience of roommate comparisons in psychological costs.

Second, I examine the structure and intensity of inter-caste exposure, exploiting

[6]This difference is direction and not based on absolute gap with roommate. All standardizations in this paper are done with respect to the observed distribution of the control group unless mentioned otherwise.

variation in room composition and size. Lower-caste students in partially mixed rooms—sharing with one same-caste and one upper-caste student—experience a 0.192 s.d. increase in academic performance, compared to 0.097 s.d. when sharing rooms exclusively with upper-caste students, and this difference is significant at the 10% level. For upper-caste students, effects on academic and internship outcomes do not differ between partially mixed rooms—sharing with one upper-caste and one lower-caste student—and fully mixed rooms—sharing exclusively with lower-caste students. If peer ability were the sole mechanism, effects should be larger in fully mixed rooms where exposure to upper-caste roommates is maximized. Instead, the presence of at least one same-caste roommate appears to mitigate social frictions while preserving learning benefits. In partial mixed rooms, due to presence of upper caste student, mitigates mental health costs for upper caste students and no such declines were observed for lower caste students revealing psychological costs for upper and lower castes arrive from different frictions and presence of co-ethnic peer doesn't eliminate the frictions for lower caste student.

To explore this further, I unpack room composition through subsample analyses by room size, which reveal distinct exposure patterns: three-person rooms provide either broader diversity or balanced diversity with built-in co-ethnic support, while two-person rooms create intensive, unavoidable contact.[7] Academic effects for lower-caste students do not differ by room size, though internship gains are stronger in three-person rooms, consistent with learning spillovers operating more effectively in larger groups. Mental health costs are higher in magnitude in three-person rooms, but the patterns are similar across room sizes. For upper-caste students, mental health deteriorates more sharply in two-person mixed rooms than in any three-person configuration, and these costs are largely mitigated when sharing three-person rooms with at least one same-caste peer. These patterns indicate that intensive dyadic contact amplifies psychological costs through unavoidable exposure and relative deprivation, while broader contact in larger rooms allows social buffering and may improve attitudes toward affirmative action.

To further decompose identity-based costs from ability-based peer effects, I estimate effects using continuous measures of caste distance—an ordinal variable ranking upper-caste students as 1, OBC students as 2, and SC/ST students as 3. For lower-caste students, caste distance shows no significant relationship with academic performance, reinforcing that benefits operate primarily through exposure to higher-performing peers rather than

[7] I do not run statistical tests to check differences between room sizes. I control for room size in all specifications to compare students in the same room sizes. Rather, I use comparative statistics to provide suggestive evidence on how contact intensity with roommates affects outcomes. I do test differences between partially and fully mixed rooms both in the full sample and in the subsample of three-person rooms.

caste distance per se. For upper-caste students, performance declines significantly as caste distance increases (0.090 s.d. for the performance index, 0.129 s.d. for GPA when moving from OBC to SC/ST exposure), even after controlling for entrance score gaps. Mental health deteriorates and support for affirmative action increases with caste distance for both groups, though the mechanisms differ: lower-caste students' costs reflect stereotype salience, while upper-caste students' costs reflect awareness of inequality alongside social distance-linked psychological costs. Decomposing by specific caste pairings reveals that upper-caste students' academic declines are driven entirely by exposure to SC/ST roommates, with no effects from OBC exposure. OBC students exhibit asymmetric responses: they benefit academically from upward exposure to upper-caste peers (mirroring lower-caste patterns) but experience negative effects from downward exposure to SC/ST peers (mirroring upper-caste patterns). SC/ST students gain academically from exposure to both upper-caste and OBC roommates but show no internship gains, suggesting additional frictions such as weaker networks or labor market discrimination. Mental health for SC/ST students deteriorates only through exposure to upper-caste roommates, not OBC roommates, indicating that psychological costs from stereotype salience operate only through contact with the dominant caste group, not through exposure to other socially disadvantaged groups.

Revealed-preference data on roommate choice and social interactions illuminate mechanisms. Students assigned to mixed rooms in their first year—both upper- and lower-caste—are more likely to live in mixed rooms in their second year, though this persistence may partly reflect switching costs. Conditional on switching, lower-caste students are less likely to choose mixed rooms, while upper-caste students show no preference. Upper-caste students adjust by living fewer years with the same roommate when initially assigned to mixed rooms.[8] Social interactions are lower in mixed rooms for both groups: lower-caste students report less eating together, playing sports, and going out, while upper-caste students show similar patterns. Help transactions reveal friction alongside learning: lower-caste students report higher academic support but lower emotional and financial support from upper-caste roommates compared to same-caste peers. Despite strong internship outcomes, lower-caste students report lower explicit internship-related help in mixed rooms, suggesting gains operate through implicit learning or exposure rather than direct assistance. Upper-caste students form more cross-caste friendships when assigned to mixed rooms, reflecting integration from the perspective of upper-caste students, who report very low shares of other-caste friends in control rooms. These patterns confirm that mixed-caste exposure generates learning

[8]Students can only reshuffle with students in the same cohort and same hostel.

and reduces prejudice, but through friction-laden interactions characterized by social distance, reduced emotional support, and selective cooperation on academic rather than personal dimensions.

A large body of evidence examines inter-racial or ethnic contact and its effects on attitudes and outcomes. Social psychologists have emphasized the role of inter-group contact in reducing prejudice, with contact theory predicting that interaction under conditions of equal status reduces bias.[9] In this literature, the main mechanism operates through attitudinal change rather than skill acquisition or network formation.[10] Economists, on the other hand, have widely studied affirmative action policies,[11] focusing on whether academic mismatch occurs (Arcidiacono and Lovenheim, 2016) or not (Bagde et al., 2016), human capital investments (Akhtari et al., 2024), and labor market outcomes (Bleemer, 2022). In this literature, diversity interventions following increased representation through affirmative action are evaluated through the lens of efficiency and distributional effects, with less attention to psychological costs borne by both beneficiaries and non-beneficiaries or to social frictions that arise when hierarchically distant groups interact, particularly in elite institutions. Most prior work on student contact—in the US and South Africa—examines settings where either admissions are not determined by transparent test scores or performance differences between students are not large.[12] My paper brings together these perspectives by examining how students in a high-stakes, real-life setting manage competition alongside caste-based social distance when socially engineered to live together. I document that inter-caste contact in higher education generates simultaneous academic gains and psychological costs, in addition to shifts in attitudes toward affirmative action. The results show how perceived losses from pre-contact competition and redistributive policies (admissions quotas) shape both the psychological costs of diversity and convergence (or divergence) in support for the policies that ensured representation of disadvantaged groups and thereby created diversity.

[9]For meta-analyses in the social psychology literature, see Paluck et al., 2019 and Pettigrew and Tropp, 2006; for recent experimental work on contact theory, see Lowe, 2025.

[10]Corno et al., 2022 is an exception that examines both attitude change and effects on academic performance.

[11]In this paper, I refer to affirmative action policies as preferential admissions provided to historically disadvantaged groups, but not roommate contact, which is designed with the objective to better benefit from diversity.

[12]Socioeconomic differences are similar in size, and in some cases larger in the South African context (Corno et al., 2022), but what distinguishes my setting is the large and transparent gap in test scores due to affirmative action policy. This could lead to perceived unfairness and possible resentment stemming from admissions. This is a very high-stakes setting where students from all groups spend almost 8-9 hours a day for more than two years and approximately $1,500 per year (approximately 56% of per capita nominal GDP of India as of 2024 estimates) preparing for an exam where success rates are less than 1%.

My work contributes to the literature on integration policies and nation building (e.g., Bazzi et al., 2019; Okunogbe, 2024; Alan et al., 2021) and to research on intergroup contact and discrimination (e.g., Boisjoly et al., 2006; Enos, 2014; Scacco and Warren, 2018; Rao, 2019; Mousa, 2020; Lowe, 2021; Dahl et al., 2021; Finseraas and Kotsadam, 2017). While this literature primarily examines prejudice and social cohesion, I study how contact in a competitive elite environment shapes academic performance, early labor market outcomes, and policy preferences simultaneously.[13] Exploiting random roommate assignment, I show that exposure to lower-caste roommates increases upper-caste students' support for affirmative action when ability gaps are small, but reduces support when ability gaps are large. Moreover, the political effects of contact vary with caste distance: greater social distance increases awareness of structural inequality, but also amplifies perceptions of unfairness when performance differences are salient. These findings demonstrate that the consequences of integration depend critically on the joint distribution of identity and perceived merit within integrated settings.

Second, this paper contributes to the literature on ethnic diversity and economic performance (e.g., Alesina and La Ferrara, 2005; Hjort, 2014; Ghosh, 2025) and trade efficiency (Anderson, 2011) by identifying the micro-level mechanisms through which identity shapes productivity in teams. In the context of elite student housing, caste identity operates as a friction on peer interactions, generating higher psychological costs and weaker cooperation in diverse rooms. I decompose these frictions into two distinct components. The first is identity-based: costs arising from stereotype salience and social distance. The second is meritocracy-based: costs that emerge when performance differences interact with perceptions of affirmative action and fairness. By varying both the intensity (one versus two opposite-caste roommates) and the composition (dyadic versus group exposure) of diversity, I show that the magnitude and nature of these frictions depend on the structure of interaction. Consistent with these mechanisms, help exchange and informal academic support decline in diverse rooms even when learning spillovers are present, indicating that diversity affects productivity through both social and psychological channels.

Third, this paper advances the peer effects literature (Sacerdote, 2001; Duflo et al., 2011; see Sacerdote, 2011) by disentangling ability-based learning spillovers from identity-based interaction costs. Leveraging random roommate assignment, I exploit independent variation in entrance exam score gaps and caste distance between roommates.

[13] Existing work rarely studies performance effects jointly with attitudes. An exception is Corno et al., 2022, which finds no negative academic effects for majority students, and Rao, 2019, which documents modest negative effects for high-income students.

Lower-caste students benefit from exposure to higher-performing peers in same-caste rooms, consistent with standard learning spillovers. However, these gains attenuate in mixed-caste rooms, indicating that identity-based frictions dampen ability spillovers. In contrast, upper-caste students' performance declines are unrelated to roommate ability and instead increase with caste distance, isolating identity-driven interaction costs. These asymmetries demonstrate that peer effects depend not only on academic complementarity but also on social identity alignment.

Fourth, this paper contributes to the literature on elite educational institutions and mobility (Chetty et al., 2020; Zimmerman, 2019; Michelman et al., 2022) by showing how roommate networks shape early career trajectories and aspirations within elite settings. Lower-caste students assigned to upper-caste roommates gain access to internships, consistent with network spillovers. However, they do not exhibit increased participation in student leadership or stronger preferences for top jobs. Conversely, upper-caste students exposed to lower-caste roommates reduce their pursuit of top jobs and leadership positions. At the same time, lower-caste students display greater social connectedness in same-caste rooms, suggesting that within-group networks provide support but may also limit exposure to elite opportunity channels. Together, these findings highlight the asymmetric role of networks in shaping mobility within elite institutions.

The remainder of this paper is organized as follows. Section 2 introduces the context in more detail and describes the data. Section 3 provides descriptive statistics. Section 4 presents the empirical specification. Section 5 reports effects on academic outcomes, internships, mental health, and attitudes toward affirmative action. Section 6 tests mechanisms based on heterogeneity in roommate entrance scores relative to the student's score, caste distance, and intensity and composition of exposure. Section 7 concludes.

2 Context and Data

The university is a elite public engineering university that enrolls approximately 1200 under graduate students every year. Admission to this institute, as well as similar institutes, is through a common national entrance exam taken at the end of grade 12 in Mathematics, Physics, and Chemistry, with admission rates ranging from 0.5% to 2% among about 1 million applicants.[14]. Affirmative Action in India is primarily caste-based, reserving a fixed proportion of seats through quotas in each public institute across all specializations. Three broad caste groups benefit from quotas in public higher educational institu-

[14]I focus only on undergraduates in this study and total enrollment in a given academic year is approximately 12000 students in the institution.

tions through Constitutional Provisions: 27% for Other Backward Castes (OBCs), 15% for Scheduled Castes (SC), and 7.5% for Scheduled Tribes (ST or Adivasis)[15][16]. I classify OBC, SC and ST students in my sample as disadvantaged students.

2.1 Room Allocation and the Natural Experiment

The university is fully residential, with all undergraduate students officially assigned to one of 13 single-gender hostels (residence halls)[17]. Students remain in their assigned hostel for the duration of their four-year undergraduate degree. The vast majority of students (97%) are assigned to either two-person or three-person rooms; I exclude the remaining 3% in single or four-person rooms from the analysis.

Upon admission, each student receives a unique college ID composed of their cohort year, academic specialization, and a randomly assigned four-digit number. Room assignments vary by year: in the first year, the institute randomly assigns roommates; in the second year and third year, students select their own roommates within their hostel; in the third and fourth years, students typically live in single rooms.

First-year room allocation process: The first-year allocation follows a two-stage process. First, the central administration assigns students to hostels; second, hostel administrators assign students to specific rooms within each hostel. Hostel assignment itself involves two steps. For each academic specialization, the administration creates an allocation matrix based on the number of first-year vacancies[18] in each hostel and the number of available mentors.[19] Using this matrix and the number of incoming students, the administration randomly allocates students to hostels within each specialization, stratified by caste category. Each hostel then receives a list of incoming students sorted by college ID.

Within each hostel, administrators allocate rooms using one of two methods: either by sequentially assigning students from the sorted college ID list to rooms[20], or by assign-

[15]Other Backward Caste (OBC's) is a collection of castes who are educationally and socially backward. Second group, Scheduled Castes(SC or Dalits) who are the most among discriminated and disadvantaged castes and the third group is Scheduled Tribes (ST or adivasis) are indigenous people who are also among the most disadvantaged groups

[16]10% additional reservation is provided to economically backward students from advantaged caste groups and if the number of women in the incoming cohort are less than 20%, new seats will be added to make it to 20%.

[17]Less than 1% of students in the survey reported commuting from home daily.

[18]Rooms allocated to first year undergraduate students are fixed in a hostel and allocated only to first year students.

[19]Each first-year student is assigned a mentor from the same hostel and specialization, typically a student from two or three cohorts earlier, to facilitate the transition to college life. Mentors are selected independently by a student body at the end of the academic year, before the new cohort arrives.

[20]Students with disabilities are always assigned ground-floor rooms and may not follow the standard

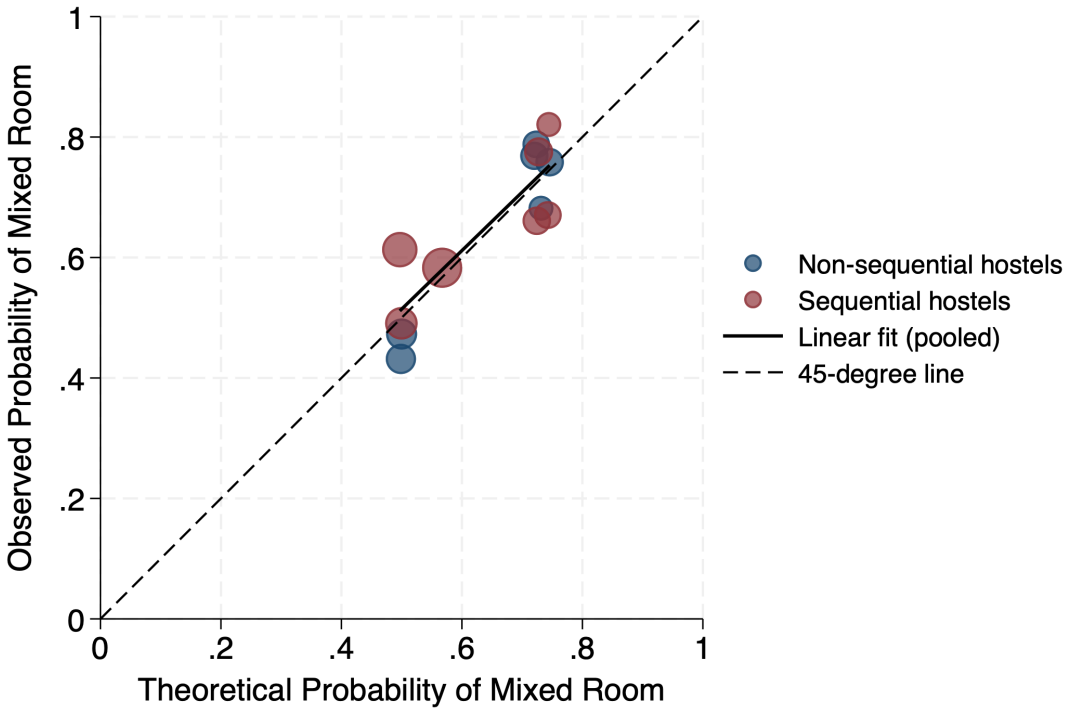


Figure 1: Observed and Theoretical Cross-Caste Mixing Across Hostels

ing rooms based on the order of student arrival during move-in The allocation method is consistent within each hostel across cohorts. Crucially, these sorted lists do not include caste information, eliminating the possibility of sorting by administrators based on personal preferences. Only 1% of students surveyed in the 2024 cohort reported knowing their roommate prior to college. Figure 1 demonstrates that the theoretical share of mixed-caste rooms closely matches the actual share in each hostel, and this alignment holds regardless of which room allocation method administrators use, sequential allocation based on college ID or otherwise. While there is no formal mechanism to change rooms or hostels, approximately 10% of students at follow-up reported changing rooms during their first year. I use the initial room allocation and report intention-to-treat (ITT) estimates throughout my analysis.

2.2 Data and Outcomes

I collected data on two cohorts that entered the institute in 2023 and 2024.[21] Data collection occurred in three phases. In the first phase, I collected baseline data from the 2024

list.

[21] I also collected data for two additional cohorts that entered the college in 2021 and 2022. I do not use their data in this analysis for reasons explained in the appendix.

cohort in the last week of July 2024. In the second phase, I surveyed the 2023 cohort between August and October 2024. In the third phase, I conducted a one-year follow-up survey with the 2024 cohort in August 2025.

First-year students survey (2024 Cohort). Baseline data were collected from first-year students during college orientation week, when they had minimal contact with their assigned roommates[22]. The baseline survey achieved an 89.7% response rate (1,087 out of 1,212 students responded), with 980 students (80.9%) completing the full survey. Students were recruited primarily through hostel orientation sessions. Student representatives conduct these sessions during orientation week to introduce students to hostel life. I, along with enumerators, obtained permission from student representatives and hostel wardens to administer surveys during these sessions. The survey was described to students as examining "student life and well-being in college," without explicitly mentioning caste. The questionnaire was self-administered: survey links were sent to students' hostel WhatsApp groups, and students completed the surveys during the orientation sessions. We took measures to prevent communication between students during the survey. For students who did not attend the sessions, we followed up individually through college emails and WhatsApp messages in hostel cohort groups. A follow-up survey was conducted with the same students approximately one year later at the start of their second academic year.[23] The follow-up survey achieved a response rate of 91.09% (1120 out of 1212 students responded to the survey), with 89.44% students completing the full survey.

I followed the same data collection protocol: students were gathered in common halls and informed that the survey aimed to understand their college life experience and how their expectations compared with reality. Students were compensated with 200 INR and 300 INR (2.5 USD and 3.75 USD respectively).

Second-year students survey (2023 Cohort). I surveyed students from the 2023 cohort between August and October 2024, approximately one year after their initial roommate assignment. This cross-sectional survey captured outcomes after one full year of exposure to randomly assigned roommates. The data collection approach for this cohort differed from the 2024 cohort due to logistical constraints. I recruited two student volunteers from each hostel, along with four coordinators to oversee the data collection process. Students

[22]Students move into hostels one week before classes begin. Surveys were conducted during this orientation week, with 95% of responses received within the first week (2-7 days of roommate exposure). The remaining 5% of respondents had one additional week of exposure.

[23]Students typically have a 2.5-month break after classes finish in early May. I chose to survey students after this break to capture outcomes from their first year, including GPA, summer projects or internships, and roommate choices for their second year.

were contacted through multiple channels: they received emails about the survey, student volunteers conducted door-to-door outreach to explain the survey purpose, posters were displayed at visible locations in each hostel, and survey information was distributed through hostel cohort WhatsApp groups. As with the 2024 cohort, the survey was described as examining student life and college experiences, without explicitly mentioning caste. Students who completed the full survey received compensation of 300 INR (3.75 USD). The survey achieved ane effective response rate of 93.15%, with 88.4% students fully completing the survey[24].

Administrative Data. I complement the survey data with entrance exam records published by the university consortium that administers the admissions exam[25]. These rank lists contain student exam IDs, ranks, assigned institute, specialization, and caste category. These lists also includes information on the exam center location where students took the exam.

I use a multi channel matching procedure to link survey respondents to their entrance exam records. First, I use self-reported ranks from the survey to match students with their exact records in the official rank lists[26], verifying matches using reported locations when available. For students who did not complete the survey, did not report their ranks, or reported them incorrectly, I use additional sources: rank lists published by coaching institutes, self-reported ranks on public LinkedIn profiles, and geographic information to identify exact matches in the official lists.[27] Using this multi-source matching strategy, I successfully matched 99% of students in the 2024 cohort and 95% of students in the 2023 cohort to their official entrance exam records. From these records, I obtain entrance exam scores, ranks, and official caste classifications for analysis.

I also obtained first-year room allocation lists from 9 out of 13 hostels. These administrative records contain room numbers and roommate names, allowing me to directly verify roommate assignments and construct the treatment variable (exposure to a different-caste roommate). I match students to their roommates using room numbers and names, then link roommates to their caste categories using the entrance exam records described above. The concordance between self-reported room numbers in the survey and adminis-

[24]Initial response rate was about 73.4%, I randomly selected 25% subsample of the remaining students and did an intensive tracking where 68.2% of students responded to the survey.

[25]The exam is conducted by a consortium of universities, with responsibility for organizing the exam rotating annually among member institutions. Each year, the organizing university publishes detailed results reports.

[26]Matching rate is quite high for for both the cohorts, 72.67% in 2024 cohort and 50.58% for 2023 cohort where I match with exact rank. Every student is assigned a unique rank that is used for institution and specialization allocation

[27]More details on this matching procedure is provided in the appendix ..

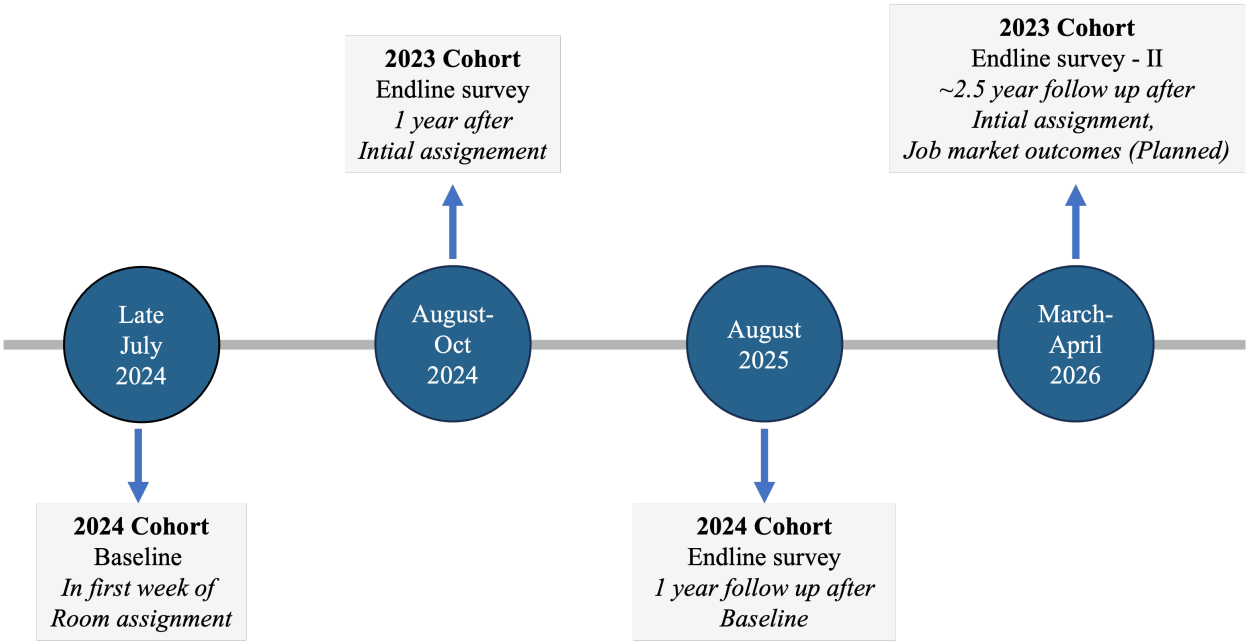


Figure 2: Timeline of Surveys: Baseline for 2024, Endline for 2023, 2024 cohorts and Endline-II for 2023 cohort

trative allocation lists is very high, providing strong validation of treatment assignment[28]. For students in the four hostels without administrative allocation data, I construct treatment status using self-reported roommate information from the survey.

2.3 Primary outcomes:

2.3.1 Academic Outcomes and Job Market Outcomes

I primarily use data collected through surveys described above for the analysis. Academic outcomes are evaluated through GPA at the end of first year, objective measures of skills, such as programming and math through grades in first year calculus and algebra courses and introduction to programming course. I created an index of math and programming skill by taking equal weighted average of 3 courses. For Job market outcomes, I use a dummy variable if the student has done an internship or supervised academic project during the summer after first year of college. I combine GPA, average of math and programming grades in first year courses and the job market outcome to create a performance index[29].

[28] Any discrepancies in these lists where administrative data was not available was resolved through recontacting respondents through student volunteers in the same hostel

[29] Index is created by equal weighted Z-Scores of the three components and standardized with respect to same caste rooms distribution

2.3.2 Mental Health Outcomes and Attitudes:

Mental Health: I measure mental health using three complementary modules. These three modules together provide a comprehensive assessment of mental health, capturing clinical symptoms (PHQ-4), self-assessed well-being changes (relative well-being), and environmental stressors (college life worries). PHQ-4 scale, which combines two items from the GAD-2 (anxiety) and two items from the PHQ-2 (depression). relative well-being module asks students to compare their current well-being to their pre-college period, providing a measure of perceived changes in overall well-being since entering college. Students responded to 5 statements about happiness, joy in daily activities, support system, control in life and stress in life on 5 point Likert scale.

College-related worries, measures students' level of worry across multiple dimensions of college life, including academics, internships, discrimination, differences in cultural and language, being away from home. This module captures ongoing stressors specific to the college environment. For each module, I construct an index by equally weighting the z-scores of its components. I then construct a composite mental health index by equally weighting the z-scores of the three module-level indices. Higher the value of the index represents worse mental health[30].

Support for Affirmative Action (AA): I also collect detailed measures of attitudes toward affirmative action[31]. These measures two dimensions: Costs and benefits of AA polices to the groups who benefit from it and costs and benefits of AA policies to the groups who doesn't benefit form it. I construct indices for the AA beneficiaries and AA policy costs to non-beneficiaries, as well as a composite affirmative action support index combining both indices. Higher values indicate greater support for affirmative action. In addition, I also asked them about their opinion on equality of opportunity in the university and report it as an outcome in poltical attitudes.

Revealed preference measures: Students choose their roommates in the second year, and I observe these choices using administrative data obtained from hostel authorities. I use this information to examine whether students from same-caste and mixed-caste rooms differ in their subsequent roommate choices along caste lines.

Aspirations, Self Worth and Networks: I also measure students' aspirations using their expected salary at graduation. I use a shortened version of the General Self-Efficacy Scale (GSE-6; Romppel et al., 2013) to capture resilience and coping with stressful situations. In addition, I measure the number of close friends within the college, at other elite

[30]Details of the exact statements and scales of the each module are provided in the appendix

[31]Students responded to a series of statements using a five-point Likert scale ranging from "strongly disagree" to "strongly agree."

colleges, and the share of close friends from other caste groups.

In addition to these measures, the survey collects a rich set of supplementary outcomes. These include detailed measures of interactions with roommates, dietary preferences, self-perceptions of skills and aspirations, and socioeconomic characteristics. I use these supplementary outcomes to provide suggestive evidence on the underlying mechanisms. Full details on variable construction and index definitions are provided in the appendix.

3 Descriptive Stats

Table 1 presents summary statistics on students' socioeconomic background, college preparation, pre-college test scores, mental health, and attitudes toward affirmative action policies, poverty, and inequality.

Panel A documents that upper-caste students, unsurprisingly, come from more advantaged family backgrounds. They are more likely to have parents with higher levels of education, to know someone who studied at an elite engineering college, and to come from wealthier households in terms of income and assets. They are also less likely to take education loans. Students who benefit from affirmative action are disadvantaged along all measured socioeconomic dimensions.[32]

Panel B shows substantial differences in college entrance exam scores between upper- and lower-caste students. On average, the gap is about 60 points (1.08 standard deviations). Because quotas operate within academic specializations, the score gap within specialization is similar in magnitude (59.7 points, or 1.07 standard deviations). These differences are large, and cutoff scores by caste category are publicly available to students.[33] Upper-caste students begin exam preparation earlier (by approximately 0.3 years), but spending on coaching—largely standardized across students—is similar across groups. Lower-caste students do not appear to exert less effort in terms of study hours, and differences in prior academic performance (Class 10 and Class 12 scores) are relatively small.

Panel C documents that mental health outcomes are modestly worse, on average, for lower-caste students. These differences are primarily driven by higher levels of anxiety and depression (PHQ-4) and greater college-environment-related worries. Appendix

[32] Approximately 10 percent of upper-caste students belong to the Economically Weaker Sections (EWS) category, which provides access to affirmative action within the upper-caste group. Restricting the comparison to upper-caste students who do not benefit from any form of affirmative action yields even larger differences.

[33] Cutoff scores are published annually after the exam and are widely disseminated through online platforms. Students routinely use this information when ranking universities and specializations.

Table 1: Summary statistics on pre college characteristics

	Lower Caste			Upper Caste			Lower Caste - Upper Caste
	N	Mean	Sd	N	Mean	Sd	Difference
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Socio Economic Indicators							
Mother's education (Bachelors or higher)	881	0.409	0.492	902	0.634	0.482	-0.225
Father's education (Bachelors or higher)	876	0.477	0.500	896	0.624	0.485	-0.147
Knows someone with IIT degree	907	0.184	0.388	912	0.265	0.442	-0.081
Assets Index	809	-0.180	0.647	824	0.150	0.547	-0.330
Taken a loan	890	0.230	0.421	885	0.148	0.355	0.082
Family income (9L or above)	895	0.415	0.493	859	0.655	0.476	-0.240
Owns a laptop/computer	809	0.759	0.428	824	0.900	0.300	-0.141
Consumption Index	503	-0.097	0.740	510	0.010	0.871	-0.107
B. JEE Preperation and Academic Background							
Entrance Exam score (out of 360)	1091	141.166	47.111	1112	201.611	47.275	-60.445
Took more than 1 attempt	934	0.281	0.449	945	0.145	0.352	0.136
Years: JEE preparation	964	2.175	1.365	964	2.473	1.386	-0.298
Average daily study hours JEE	962	8.874	3.472	963	8.443	3.375	0.431
JEE Prep Cost (in lakhs INR)	884	1.276	1.154	875	1.332	1.142	-0.056
Class 12 score (%)	934	90.338	6.494	942	93.286	6.537	-2.948
Class 10 score (%)	935	92.877	8.109	939	95.044	7.282	-2.167
C. Mental Health at Baseline							
Mental Health: BL	490	0.023	0.522	502	-0.011	0.540	0.034
PHQ4: Anxiety and Depression	438	0.039	0.763	455	-0.017	0.813	0.056
Relative wellbeing Index	426	-0.026	0.549	450	-0.002	0.659	-0.024
Worry Index	490	0.064	0.687	502	-0.011	0.586	0.075
D. Attitudes towards Affirmative Action and Redistribution							
AA Merit Index (rev)	441	0.437	0.773	454	-0.339	0.696	0.776
AA Social Justice	439	0.369	0.660	454	-0.309	0.850	0.678
Poverty: unfair society	471	0.104	1.042	477	-0.044	1.058	0.148
Fairness: Similar Living std	471	0.125	1.022	477	0.004	1.054	0.121
Support Caste based AA	471	0.255	0.837	476	-0.211	1.071	0.466
Opportunity inside university equal	429	0.013	0.937	453	0.008	1.092	0.005

Table A1 decomposes these measures and shows that lower-caste students report substantially higher concerns related to cultural and language differences (0.51 standard deviations) and discrimination (0.31 standard deviations). In contrast, upper-caste students are more likely to report worries related to academics and time management. Differences in other mental health measures are smaller and not in same sign.

Panel D reports students' views on affirmative action policies[34]. There are large differences in both meritocratic concerns (0.75 standard deviations) and views about the social justice rationale for affirmative action (0.67 standard deviations). Upper-caste students are more likely to view quotas as unfair relative to merit-based selection (0.91 standard deviations), to believe that quotas compromise academic quality (0.83 standard deviations), and to perceive that they reduce incentives to work hard (0.73 standard deviations) and ultimately harm students by placing them in more challenging environments (0.54 standard deviations). Upper-caste students are also less likely to believe that quotas promote social justice (0.87 standard deviations), foster diversity in colleges and universities (0.49 standard deviations), or benefit non-beneficiary students through increased diversity (0.66 standard deviations). These attitudinal differences align with lower-caste students' greater worry about cultural fit and discrimination.

4 Empirical Specification

I estimate the effect of mixed-caste room assignment on a range of outcomes of interest the following baseline specification on full sample, and on subsample of upper and lower caste respondents:

$$Y_{ikcst} = \beta_0 + \beta_1 \text{Mixed}_{ikc0} + \beta_2 Y_{ikcs0} + \gamma_{kc} + \delta_{sc} + X_{ikcs0} + \varepsilon_{ikcst} \tag{1}$$

where Y_{ikst} is the outcome for student i, in hostel k, cohort c, in specialization s, in the follow-up survey time (time t) and Y_{iks0} is the baseline value of the same variable[35]. Treatment variable Mixed_{ikc0} is a dummy variable equals one if the student is assigned to a room with atleast one different-caste roommate[36]. γ_{kc} is a vector of dummies for hostel k and cohort c, and δ_{sc} is a vector of dummies for specialization cohort fixed effects. I use hostel-cohort to control for systematic differences across hostel like room size, gender

[34] All attitude measures are collected on five-point Likert scales and standardized for comparability.

[35] I control for baseline measures when available and I have baseline measure of mental health and attitudes on affirmative action only for 2024 cohort, I replace baseline measure with 0 when missing and add a dummy for missingness of the variable

[36] In a two room sharing room, if the other roommate is from different caste or if in a three sharing room, atleast one of the other roommates is of different caste, I classify that room as mixed caste room

and hostel size, and any differences across cohorts in terms of student composition in the hostel and I control for specialization-cohort fixed effects to account for differences in students' academic backgrounds. X_{iks0} a vector of controls, which includes college number of roommates. I use ANCOVA specification when ever I have baseline value of the outcome. Standard errors are clustered at the room level[37].

The regression is estimated separately for upper-caste and lower-caste students, allowing the effect of roommate assignment to vary across social groups. The coefficient of interest, β_1, measures the causal effect of being assigned a roommate from a different caste on the outcome Y_{ijkc}. Fewer than 10% of students changed roommates during their first year, and I reconstruct the initial assignment using room numbers reported in the survey and administrative records. Given the random assignment of roommates within hostels and across caste groups, β_1 can be interpreted as an intent-to-treat (ITT) estimate of the effect of inter-caste mixing.

I use additional variation in the treatment coming from in three-sharing rooms, I exploit within-room variation by distinguishing between partially mixed and fully mixed room assignments. I estimate:

$$Y_{ikcst} = \beta_0 + \beta_1 \text{Fully Mixed}_{ikc0} + \beta_2 \text{Partial Mixed}_{ikc0} + \gamma_{kc} + \delta_{sc} + X_{ikcs0} + \varepsilon_{ikcst} \quad (2)$$

where $\text{Partial Mixed}_{ikc0}$ equals one if the student is assigned to a room with exactly one same-caste roommate and one different-caste roommate in a three sharing room, and $\text{Fully Mixed}_{ikc0}$ equals one if all roommates are from different caste categories[38]. The omitted category is a same-caste room. This specification allows me to identify whether having at least one same-caste peer moderates any negative effects of diversity coming from mixed caste rooms.

4.1 Balance Tables

Appendix tables A2 and A3 examine balance on pre-treatment characteristics. Table A2 presents balance on observable socioeconomic indicators and pre-college characteristics. There are no systematic differences between mixed and non-mixed rooms among either

[37]One additional specification will be used in further analysis. First, while the main specification classifies students from OBC, SC, and ST categories as "lower caste" and all others as "upper caste", an alternative approach will consider the hierarchical structure within caste categories and examine the effects of mixing accordingly, rather than relying on the current dichotomous classification.

[38]In a two room sharing room, by definition its a fully mixed room, and in three sharing rooms, if two other roommates are from different caste, I classify that room as fully mixed caste room

upper- or lower-caste students, with joint F-tests showing p-values of 0.633 and 0.465 for upper- and lower-caste students, respectively. Table A3 presents balance on pre-college networks, measured using baseline data for the 2024 cohort, and similarly shows no systematic differences between mixed and non-mixed rooms for either group. Taken together with Figure 1—which shows that the observed proportion of mixed rooms closely matches the theoretical probability under random assignment these results support the plausibility of random roommate allocation.

4.2 Early effect of exposure to mixed room:

In table A4 I report overall mental health indices, attitudes towards affirmative action and beliefs on skills are not different between mixed and non-mixed rooms. There are no systematic differences in these indices between mixed and non-mixed rooms among either upper- or lower-caste students, with joint F-tests showing p-values of 0.479 and 0.271 for upper- and lower-caste students, respectively. I unpack these indices and report estimates of assignment to mixed rooms for both upper and lower caste students in table A5.

Among upper-caste students, I find evidence of early frictions associated with assignment to mixed rooms . Upper-caste students in mixed rooms report lower expected friendship with their roommates (p = 0.039) and higher concerns related to cultural and language differences. Attitudes toward affirmative action policies—particularly those related to diversity and social justice show negative sign for upper-caste students in mixed rooms, although these estimates are not statistically significant. Overall, these patterns suggest early discomfort among upper-caste students when assigned to mixed rooms. I reject null for joint test of insignificance using an F-test (p = 0.031)

For lower-caste students, exposure to mixed rooms does not affect attitudes or overall mental health in the first week. One exception is that lower-caste students assigned to mixed rooms report higher concerns about discrimination. I fail to reject the joint null for lower-caste students. Although the survey was administered within the first week of college, students already knew their assigned roommates and had interacted with them for 2-7 days. This initial period may be salient in shaping expectations and early concerns about roommate, particularly for upper caste students. These differences in early reactions could be reasonable given upper students had lower exposure to lower caste friends pre-college compared to exposure of lower caste students to upper caste friends as reported in A3.

5 Main Results

5.1 Results on Academic Performance and Internships:

Table 2 reports the first set of results on students' academic performance and summer internships. Panel A presents estimates for lower-caste students, Panel B for upper-caste students, and Panel C for the full sample. The same pattern is followed in all the tables with main specification. Standard errors, reported in parentheses, are clustered at the room level. The coefficient on mixed room captures the effect of being assigned to roommates from different caste groups, relative to assignment to same-caste rooms.

Overall performance—measured as an index combining GPA, average mathematics and programming grades, and completion of a summer internship at the end of the first year—increases by 0.132 standard deviations ($p < 0.05$) for lower-caste students assigned to mixed rooms. GPA, mathematics, and programming grades[39] increase by 0.091 and 0.104 standard deviations, respectively, for lower-caste students exposed to different-caste roommates; these estimates are marginally significant at the 10% level. In addition, lower-caste students in mixed rooms are 6.5 percentage points ($p < 0.01$) more likely to complete a summer internship or academic project by the end of the first year—representing a 6.5 percentage point increase (67% increase relative to control group) compared to other lower caste students in same caste rooms.

For upper-caste students, there is no statistically significant decline in overall performance, although the point estimate is negative at 0.058 standard deviations when assigned to mixed rooms. GPA declines by 0.103 standard deviations ($p < 0.05$), while mathematics and programming grades and internship completion are not statistically different form zero. Internship effects are precise null where as math and programming grades are 0.054 standard deviations and imprecisely estimates. In pooled estimates (Panel C), outcomes for students assigned to mixed rooms are not statistically different from those in same-caste rooms in terms of either academic performance or internship completion. Thus, increasing diversity in hostel rooms does not affect average academic or early labor-market outcomes. However, this null average effect masks distributional consequences: mixed rooms improve academic performance and early labor-market outcomes for lower-

[39] First-year students take two mathematics sequences covering calculus, algebra, and differential equations, along with an introductory programming course. Calculus and Algebra create crucial foundation for subsequent coursework in all engineering fields. Programming is included as an outcome for three reasons: (i) a 15-percentage-point gap (0.3 standard deviations) in prior programming exposure between upper- and lower-caste students; (ii) access to laptops—required for coding—creates an additional barrier to learning, although the institute provides a computer center access to students; and (iii) programming proficiency is an important labor-market skill and has been shown to improve job-market outcomes for disadvantaged students (Krill et al., 2019; Liu et al., 2024).

caste students while reducing academic performance—particularly GPA—among upper-caste students.

Table 2: Effect on GPA, Internships, Maths and Programming performance

	Performance Index (1)	GPA (2)	Maths & Programming (3)	Summer Project or Internship (4)
Panel A: Lower Caste Sample				
Mixed Room	0.132***	0.091*	0.104*	0.065***
	(0.045)	(0.052)	(0.059)	(0.024)
DV Mean in same caste rm	-0.357	-0.466	-0.411	0.097
R-squared	0.337	0.434	0.373	0.070
N	1059	968	1052	1047
Panel B: Upper Caste Sample				
Mixed Room	-0.058	-0.103**	-0.054	-0.006
	(0.046)	(0.050)	(0.045)	(0.033)
DV Mean in same caste rm	0.348	0.453	0.458	0.221
R-squared	0.344	0.474	0.421	0.100
N	1078	1043	1076	1069
Panel C: Full Sample				
Mixed Room	0.034	-0.021	0.027	0.028
	(0.032)	(0.038)	(0.038)	(0.020)
DV Mean in same caste rm	-0.009	0.000	0.018	0.158
R-squared	0.409	0.514	0.456	0.063
N	2137	2011	2128	2116

Notes: The table presents results for academic performance outcomes, including GPA, mathematics and programming grades, and summer internship or summer academic project completion. Column 1 reports estimates for a composite performance index constructed using standardized scores of cumulative GPA, mathematics and programming grades, and summer internship or project completion. Column 2 reports results for cumulative GPA, defined as the standardized cumulative GPA up to the end of the first year. Column 3 reports results for mathematics and programming grades, measured as an index constructed using standardized scores from mathematics and programming courses. Column 4 reports results for summer internship or academic project completion, defined as an indicator for whether a student completed a summer internship or academic project during the summer following the first year. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective college entrance exam scores and number of roommates.

5.2 Results on Mental Health

Table 3 reports the effects of mixed-room assignment on four measures of psychological well-being: a composite Mental Health Index (constructed from all three components), the PHQ-4 anxiety–depression scale, a relative well-being index capturing changes since pre-college, and a worry index reflecting concerns about academics, discrimination, and cultural adjustment. Higher values on all indices indicate worse mental health.

Lower-caste students experience a 0.081 standard deviation decline in the composite mental health index when assigned to mixed rooms ($p < 0.10$). This effect is primarily driven by increases in the PHQ-4 and worry indices, both of which are imprecisely estimated but indicating worse mental health, with magnitudes of 0.080 and 0.034 standard deviations, respectively. Decomposition of PHQ-4 responses (in Table A6) indicates lower-caste students in mixed rooms are more likely to report increases in nervousness and anxiety. They also report greater concern about discrimination and cultural adjustment (Table A8). Effects on relative well-being are small and statistically insignificant, although lower-caste students in mixed rooms are more likely to report a reduced support system after entering college (Table A7). Overall these findings suggest that diversity imposes meaningful short run mental health costs on lower caste students, even as they experience academic and internship gains documented earlier.

In contrast, mixed-room assignment has no statistically significant effect on overall mental health for upper-caste students, although the point estimate is positive at 0.066 standard deviations. The PHQ-4 index increases by 0.090 standard deviations, comparable in magnitude to the effect for lower caste students. Upper caste students in mixed rooms are also more likely to report worries related to time management (Table A8).

Pooling across caste groups and comparing diverse versus homogeneous caste rooms, the average effect of mixed-room exposure is a 0.079 standard deviation increase in the composite mental health index ($p < 0.05$), driven mainly by the PHQ-4 component (0.096 standard deviations, $p < 0.05$). Overall, diversity appears to impose meaningful mental-health costs on both upper- and lower-caste students, though the drivers of these costs may differ across groups. While components of the worry index are imprecisely estimated, their patterns provide suggestive evidence on channels of mental-health costs. For upper castes, they stem from competitive pressures (such as time-management concerns) and possibly social frictions associated with interacting with lower-caste students among upper-caste students. Among lower-caste students, the corresponding frictions seem to reflect concerns about discrimination, adjustment to elite institutional environments, and academic competition. The mechanisms section provides additional evidence on these specific channels.

Table 3: Effect on Mental Health: PHQ4, Relative Happiness, Worry Index

	Mental Health Index (1)	PHQ4 Index (2)	Relative well being Index (3)	Worry Index (4)
Panel A: Lower Caste Sample				
Mixed Room	0.081*	0.080	0.014	0.036
	(0.044)	(0.063)	(0.041)	(0.045)
DV Mean in same caste rm	0.014	0.010	0.036	0.046
R-squared	0.144	0.121	0.116	0.142
N	985	934	850	892
Panel B: Upper Caste Sample				
Mixed Room	0.066	0.090	0.016	0.015
	(0.049)	(0.068)	(0.051)	(0.046)
DV Mean in same caste rm	-0.046	-0.016	-0.055	-0.052
R-squared	0.136	0.144	0.091	0.112
N	1031	971	846	896
Panel C: Full Sample				
Mixed Room	0.079**	0.096**	0.023	0.027
	(0.034)	(0.047)	(0.032)	(0.033)
DV Mean in same caste rm	-0.016	-0.003	-0.009	-0.003
R-squared	0.110	0.098	0.074	0.095
N	2016	1905	1696	1788

Notes: The table presents results for mental health outcomes, including the Mental Health Index, PHQ-4 scores, relative happiness, and worry index scores. Column 1 reports estimates for a composite Mental Health Index constructed using standardized scores of the individual mental health measures. Column 2 reports results for PHQ-4 scores, defined as the standardized PHQ-4 score measuring anxiety and depression symptoms. Column 3 reports results for relative well being index, defined as the index of 5 components, with higher magnitude represents worse relative well being score. Column 4 reports results for worry index scores, defined as the standardized worries about various college aspects score. In all the columns, higher values of these indices indicate worse mental health outcomes. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective dependent variables and a missing indicator for the baseline controls and number of roommates.

5.3 Results on Attitudes towards Affirmative Action

Table 4 examines the effect of mixed-caste roommate assignment on students' attitudes toward affirmative action (AA). I construct a composite index of support for affirmative action by combining two components: perceived costs and benefits of AA for reserved groups, and perceived costs and benefits for non-reserved groups. Columns 2 and 3 report results for each component separately. Column 4 reports effects on perceptions about equality of opportunity. All indices are coded so that higher values indicate greater support for affirmative action, stronger agreement with the benefits of AA for beneficiaries, and lower disagreement with the costs of AA policies for those who do not benefit from them. For lower-caste students, mixed-room exposure has no statistically significant effect on overall support for affirmative action; the point estimate is 0.046 standard deviations and is not statistically different from zero. Attitudes toward the benefits of AA for their own groups do not improve in mixed rooms. However, lower-caste students become more likely to disagree with perceived costs of AA for non-reserved groups (0.11 standard deviations, $p < 0.05$), driven in part by increased disagreement with the view that AA is unfair relative to a meritocratic system. Although individual components (reported in Tables A9 and A10) are imprecisely estimated, patterns within the reserved-group dimension suggest that lower-caste students express greater support for AA on social justice and diversity grounds and are less likely to believe AA harms reserved groups by placing them in challenging environments. At the same time, they are more likely to agree that AA may reduce incentives to work hard. Upper-caste students assigned to mixed rooms exhibit greater overall support for affirmative action, with an effect of 0.088 standard deviations ($p < 0.10$), primarily driven by increased agreement regarding the benefits of AA for reserved groups (0.115 standard deviations, $p < 0.05$). Upper-caste students in mixed rooms are no more likely to disagree with the perceived costs of AA for non-reserved groups, although component-level results indicate they are less likely to agree that AA compromises academic quality (0.142 standard deviations, $p < 0.10$). However, they are equally likely to view AA policies as unfair relative to meritocratic systems and to perceive no diversity-related benefits for upper-caste students. Similarly, lower-caste students in mixed rooms are no more likely to believe in diversity-based spillover benefits for non-reserved groups.

In the pooled sample, mixed-caste rooms increase overall AA support by 0.074 standard deviations. The pattern across components reveals convergence in beliefs about AA's benefits for reserved groups, alongside divergence in perceptions of its costs for non-reserved groups. This divergence in views on AA's costs aligns with the higher mental-health costs observed in diverse rooms.

Column 4 examines a broader political attitude—whether opportunities within the university are equal for all students. Both upper- and lower-caste students in mixed rooms are more likely to agree that opportunities are equal across caste groups.[40] Beliefs in equal opportunity rise by 0.20 standard deviations ($p < 0.10$) for upper-caste students, with a similar but imprecisely estimated increase for lower-caste students. In the full sample, this produces a statistically significant improvement in perceived equality of opportunity. Thus, diverse rooms not only shift attitudes toward affirmative action but also promote convergence in beliefs about equality within the university, potentially reflecting increased awareness of—or support for—complementary institutional policies that benefit disadvantaged students within the university, such as fee reductions and early access to library textbooks.

6 Model

Building on the empirical findings, I develop two complementary models of roommate integration. The first model explains why cross-caste exposure changes attitudes toward affirmative action. The second explains why the same roommate assignment can simultaneously improve academic outcomes while reducing wellbeing and altering labour-market outcomes.

For AA attitudes, I build on Bénabou and Tirole's framework of self-image, students hold beliefs about the extent to which caste differences in academic achievement reflect merit rather than unequal opportunities. Cross-caste interaction provides new information, which may update these beliefs and, in turn, attitudes toward affirmative action.

In the second part, I use a model of social capital formation based on two primitives: roommate alignment, λ_{ij}, and arena alignment, ι_c. The model explains how grades, wellbeing, internships, and leadership positions are jointly determined following cross-caste roommate assignment, and why these outcomes may either reinforce or offset one another.

The central idea is that social time is scarce. Students allocate it between their roommate and the broader campus environment, producing two forms of social capital. Time with socially aligned peers generates bonding capital through belonging, whereas interaction across social groups generates bridging capital through access to new information and networks, albeit at an asymmetric psychological cost. Proofs are collected in Appendix A.

[40] This measure is available only for the 2024 cohort, leading to smaller sample sizes.

Table 4: Effect on Support Affirmative action: Benefits and Costs to different groups

	Supports AA Index (1)	Benefits of AA to reserved groups (2)	Costs of AA to non reserved groups (rev) (3)	Opportunity inside university equal (4)
Panel A: Lower Caste Sample				
Mixed Room	0.046	-0.022	0.110**	0.174
	(0.044)	(0.044)	(0.055)	(0.109)
DV Mean in same caste rm	0.240	0.239	0.237	0.011
R-squared	0.117	0.110	0.131	0.076
N	888	883	874	432
Panel B: Upper Caste Sample				
Mixed Room	0.088*	0.115**	0.048	0.191*
	(0.053)	(0.057)	(0.058)	(0.105)
DV Mean in same caste rm	-0.256	-0.251	-0.265	-0.021
R-squared	0.196	0.172	0.184	0.099
N	911	907	901	449
Panel C: Full Sample				
Mixed Room	0.074**	0.046	0.097**	0.176**
	(0.037)	(0.037)	(0.045)	(0.077)
DV Mean in same caste rm	-0.005	-0.002	-0.012	-0.006
R-squared	0.228	0.205	0.196	0.062
N	1799	1790	1775	881

Notes: The table presents results for attitudes towards Affirmative Action, including benefits to the group who benefit from AA policy and perceptions on costs to those who doesn't benefit from AA policy. Column 1 reports estimates for a composite Affirmative Action support Index constructed by mean of dependent variables in columns 2 and 3. Column 2 reports estimates for a index of attitudes that are related to benefits and costs (reverse coded) of AA policies for groups that receive reservation. Column 3 reports estimates for a index of attitudes that are related to benfits and cost (reverse coded) of AA policies for groups that don't benefit from AA. Column 4 reports if students believe opportunities inside university are equal for all. Higher values of in 1,2,3 columns indicate stronger support for Affirmative Action. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective dependent variables and a missing indicator for the baseline controls and number of roommates.

6.1 Environment and Social Capital

Let $\mathcal{N} = \{1, \ldots, n\}$ index an entering cohort. Student i has caste type $c_i \in \{U, L\}$, academic preparation θ_i (entrance score), and social advantage A_i (family income, parental education, schooling quality), distinct from but correlated with academic preparation θ_i. A room assignment $\alpha \colon \mathcal{N} \rightrightarrows \mathcal{N}$ is imposed, $\alpha_i \subseteq \mathcal{N} \setminus \{i\}$, with random assignment used for identification.

Assumption 1 (Endowment gap)**.** *Affirmative action admits lower-caste students at a lower entrance threshold: $\theta \mid U$ first-order stochastically dominates $\theta \mid L$, and likewise $A \mid U$ dominates $A \mid L$. Write $\Delta\theta \equiv \mathbb{E}[\theta \mid U] - \mathbb{E}[\theta \mid L] > 0$ and $\Delta A \equiv \mathbb{E}[A \mid U] - \mathbb{E}[A \mid L] > 0$. Let*

$$p \;\equiv\; \Pr(\theta_j > \theta_i \mid i \textit{ lower-caste, room mixed}) \tag{3}$$

denote the probability that a mixed pairing runs in the expected direction. FOSD implies $p > 1/2$; the empirical magnitude is reported in the proof of Proposition 1.

Assumption 2 (Two aligned relationships)**.** *Roommate alignment $\lambda_{ij} = \lambda(d_{ij}) \in (0, 1]$ is decreasing in social distance d_{ij}, with $\lambda(0) = 1$ (same-caste roommates). In a mixed-caste room let d_M denote the cross-caste social distance and write $\lambda_M \equiv \lambda(d_M)$ for the corresponding alignment level; λ_M is a specialization of λ_{ij}, not a second parameter. Arena alignment ι_c is ordered $\iota_U > \iota_L$, with $\iota_U > \lambda_M$; normalize $\iota_U = 1$.*

Assumption 3 (Directional friction)**.** *Cross-caste roommate time carries psychic friction $\chi_{c_i}(d_{ij}) \geq 0$, evaluated at the pair's social distance d_{ij}. In a mixed room $d_{ij} = d_M$ and I write $\chi_c(d_M)$ or simply $\chi_c(d)$ where unambiguous. The slope $\gamma_c \equiv \chi_c'(d_M)$ satisfies $\gamma_L > 0$; γ_U is left unrestricted, so the model does not presuppose whether upper-caste responses reflect discomfort or competition. In a same-caste room $d_{ij} = 0$ and $\chi_c(0) = 0$.*

Assumption 4 (Arena access)**.** *Arena participation requires skills shaped by prior advantage: $A_i \geq A_{c_i}$. Since $\iota_L < 1$ I assume $A_L > A_U$, so lower-caste students are disproportionately arena-excluded and, when excluded, hold $s_i = 1$.*

Each student allocates one unit of social time between roommate and arena, $s_i + a_i = 1$. Belonging, backing, and bridging are

$$B_i = \lambda_{ij}\, s_i + \iota_{c_i}\, a_i, \qquad K_i = \lambda_{ij}\, s_i, \qquad R_i = 1 - B_i. \tag{4}$$

6.2 Grades and Attitudes

Grades are determined first and enter the placement decision (§6.3); attitudes are treated purely as an outcome and, by Assumption 5 below, feed back into nothing.

6.2.1 Academic transmission

Learning is directional: a student benefits only from a better-prepared roommate. Grades are

$$g_i = \theta_i + \beta \, \lambda_{ij} \, (\theta_j - \theta_i)_+ - \mu_{c_i} \, d_{ij} + \varepsilon_i, \qquad \mathbb{E}[\varepsilon_i \mid \theta_i, \theta_j, A_i, A_j] = 0, \tag{5}$$

with $\beta > 0$ the transmission rate and $\mu_{c_i} \geq 0$ the academic cost of social distance. Transmission depends on the assignment (λ_{ij}), not on realized contact intensity $\tilde{s}_{ij}$, which is what makes grades predetermined with respect to the social-time problem of §6.3.

Proposition 1 (Academic transmission). *Let $\delta \equiv \mathbb{E}[\theta_j - \theta_i \mid \theta_j > \theta_i$, mixed$]$ and $\delta' \equiv \mathbb{E}[\theta_i - \theta_j \mid \theta_i > \theta_j$, mixed$]$ denote the cross-caste score gaps conditional on each ordering, and let $\delta_c \equiv \mathbb{E}[\theta_j - \theta_i \mid \theta_j > \theta_i$, both caste $c]$ denote the within-caste upward gap. Under Assumption 1, in a same-caste room each student receives baseline transmission $\beta \cdot \frac{1}{2} \cdot \delta_c$ (since within caste $\lambda_{ij} = 1$, $d_{ij} = 0$, and the probability that the roommate scores higher is $1/2$ by symmetry). Relative to this baseline:*

(i) *A lower-caste student's expected grade change is $\beta[\lambda_{ij} \, p \, \delta - \delta_L/2] - \mu_L \, d_{ij}$.*

(ii) *An upper-caste student's expected grade change is $\beta[\lambda_{ij} \, (1-p) \, \delta' - \delta_U/2] - \mu_U \, d_{ij}$.*

The asymmetry follows from two objects: p weights cross-caste transmission up for lower-caste students and down for upper-caste students, while the forgone within-caste term $\delta_c/2$ is subtracted from both. For upper-caste students, both terms in the bracket work against them.

6.2.2 A self-image model of deservedness

Let $\psi \in [0, 1]$ be the true share of the caste preparation gap attributable to circumstance rather than ability. Each student holds an ego-protective prior b_i (baseline belief that admission reflects merit) in the spirit of Bénabou and Tirole (2002,2006) and students does not simply adopt the Bayesian posterior ψ_i^B but instead chooses in the style of Alesina and Angeletos(2005)

$$\hat{\psi}_i = \arg\max_{\psi} \left\{ -v(r_i) \, \psi - \frac{1}{2\kappa} \big(\psi - \psi_i^B\big)^2 \right\}, \tag{6}$$

where r_i is own academic standing, $v(r_i) \geq 0$ with $v' > 0$ is the ego stake in a low-ψ ("I deserve my place") belief, and κ scales malleability. This yields the equilibrium distortion

$$\hat{\psi}_i = \psi_i^B - \kappa\, v(r_i), \qquad \alpha_i = \alpha_0 + \varphi\, \hat{\psi}_i + \zeta_i. \tag{7}$$

The posterior is a precision-weighted average of the prior and the co-residence signal z_i:

$$\psi_i^B = \big(1 - \ell(d_i)\big)\, \psi(b_i) + \ell(d_i)\, z_i, \qquad \ell(d_i) = \frac{\tau(d_i)}{\tau(d_i) + \tau_0}, \tag{8}$$

with $\psi'(b_i) < 0$ and ℓ increasing in caste distance d_i (a same-caste roommate reveals nothing about the caste gap, $\ell(0) = 0$).

Assumption 5 (Beliefs are not payoff-relevant)**.** *ψ_i and α_i enter none of the payoff objects in §6.3: neither wellbeing, placement, positions, nor utility. Support for AA is something a student comes to believe, not something traded off against social time. This severs the beliefs→behavior link: no belief feeds back into chosen social time, so no fixed point arises between what a student believes and how they behave.*

Assumption 6 (Beliefs update off exposure, not contact)**.** *The signal precision $\tau(d_i)$ and the prior-precision objects $\tau_0^k(c_i)$ depend on the identity of the roommate (caste distance d_i) and the student's own caste, not on realized interaction. Informativeness is a property of* who *the roommate is, not of* how the pair spend their time.

Proposition 2 (Correction of mis-anchored beliefs)**.**

(i) *If standing r_i is pre-assignment ($r_i = \theta_i$), then under (6)–(8) the mixed- versus same-caste contrast in support is*

$$\alpha_i\big|_{d_i>0} - \alpha_i\big|_{d_i=0} = \varphi\, \ell(d_i)\big(z_i - \psi(b_i)\big). \tag{9}$$

The ego wedge $-\varphi\kappa\, v(r_i)$ is common to both room types and cancels; the contrast is identified purely off the correction channel, and is larger the more strongly anchored the prior (larger b_i) and the more caste-distant the roommate.

(ii) *If standing responds to assignment (r_i includes g_i), the contrast acquires a residual,*

$$\Delta\alpha_i = \varphi\, \ell(d_i)\big(z_i - \psi(b_i)\big) - \varphi\kappa\big[v(r_i\big|_{\textit{mixed}}) - v(r_i\big|_{\textit{same}})\big]. \tag{10}$$

By Proposition 1 and $v' > 0$: for lower-caste students, mixing raises grades so the residual is negative and the correction is attenuated; for upper-caste students, mixing lowers grades

so the residual is positive and the correction is amplified. The clean contrast (9) *therefore understates the upper-caste attitude shift and overstates the lower-caste one.*

Part (ii) implies that the identifying assumption behind (9) is conservative for the paper's main identified result (upper-caste benefits correction): relaxing it strengthens the predicted shift. For the lower-caste fairness correction, which is already flagged as consistent but not separately identified, (9) is an upper bound.

6.2.3 Which dimension moves

ψ is a single dimensional parameter, but deservedness has two sides a student forms views about: whether the policy corrects a real disadvantage (benefits) and whether one's own place is deserved (fairness). Co-residence delivers a single signal z_i about ψ; the question is if students with different starting positions update on different dimensions. Before co-residence, each student already holds stronger priors about their lived experiences or equilibrium beliefs that their group would hold. Co-residence reveals the roommate's circumstances A_j—family resources, schooling quality, daily life behind the score—which, alongside the already-known θ_j, lets the student update the production function $\theta = f(\text{ability}, A)$ and hence ψ. But this update moves beliefs only on the dimension where the prior was imprecise.

Let $\tau_0^k(c_i)$ denote student i's prior precision on dimension k, reflecting accumulated knowledge from living on side c_i of the gap. The posterior weight co-residence receives on dimension k is

$$\ell^k(d_i, c_i) = \frac{\tau(d_i)}{\tau(d_i) + \tau_0^k(c_i)}, \tag{11}$$

and the update on dimension k is

$$\Delta\alpha_i^k = \varphi^k \, \ell^k(d_i, c_i)\big(z_i - \psi(b_i)\big). \tag{12}$$

The prior-precision ordering follows from which side of the gap each group has lived:

$\tau_0^{ben}(L) > \tau_0^{ben}(U)$: lower-caste students have direct experience of disadvantage.

$\tau_0^{fair}(U) > \tau_0^{fair}(L)$: upper-caste students were not admitted through AA, so the fairness of their own place is settled by the admissions process itself.

These are informational endowments implied by the institutional structure and Assumption 1, not behavioral assumptions about motivated reasoning.

Proposition 3 (Dimension-specific updating). *Under* (11)–(12) *and the prior-precision ordering above, holding $z_i - \psi(b_i)$ and d_i fixed:*

(i) $\Delta\alpha_U^{ben} > \Delta\alpha_L^{ben}$*: the benefits correction is larger for upper-caste students.*

(ii) $\Delta\alpha_L^{fair} > \Delta\alpha_U^{fair}$*: the fairness correction is larger for lower-caste students.*

(iii) *The two corrections move the between-caste support gap in opposite directions on the two dimensions. Define the dimension-specific gap as $G^k \equiv \alpha_U^k - \alpha_L^k$. Then:*

$$\Delta G^{ben} = \Delta\alpha_U^{ben} - \Delta\alpha_L^{ben} > 0, \qquad \Delta G^{fair} = \Delta\alpha_U^{fair} - \Delta\alpha_L^{fair} < 0.$$

The benefits gap closes (upper-caste students, who were less supportive, revise upward by more); the fairness gap widens (lower-caste students, who were already more supportive on fairness grounds, revise upward by more). The model predicts opposite signs on the two dimensions; the net effect on overall support is not signed and depends on the relative magnitudes $\varphi^{ben}\left[\ell^{ben}(d,U) - \ell^{ben}(d,L)\right]$ versus $\varphi^{fair}\left[\ell^{fair}(d,L) - \ell^{fair}(d,U)\right]$.

The opposite-sign prediction of part (iii) is the sharpest testable restriction the model produces on attitudes: an exposure-only specification predicts a common shift across dimensions and cannot generate $\Delta G^{ben} > 0 > \Delta G^{fair}$. The split vanishes when prior knowledge is symmetric ($\tau_0^k(U) = \tau_0^k(L)$), students with prior cross-caste contact could have relatively symmetric priors, and the dimensional asymmetry should attenuate for them.

Remark (General ψ). *The belief ψ is modeled as learning about the the production function $\theta = f(\text{ability}, A)$, not just caste gap per se. Any co-residence pairing with $d_{ij} > 0$ delivers information about f, including within caste.*

Corollary 1 (Costly contact is more informative). *If both the academic distance cost $\mu_c d$ and signal precision $\tau(d)$ increase in social distance, then $\partial|\Delta\alpha_i|/\partial(\mu_c d) > 0$: students paying a larger academic cost of cross-caste co-residence also show greater attitude updating. This operates on the overall precision $\tau(d_i)$ in* (11)*, not on the dimensional split.*

Remark (Flag, not a result). *A companion conjecture ties the unexplained upper-caste grade drag ($-\mu_U\, d_{ij}$ in* (5)*) to the same belief revision. An upper-caste student's benefits update is large precisely because $\tau_0^{ben}(U)$ is low. But ψ is one object: revising it upward on benefits grounds simultaneously revises how much of the specific margin Δ_i over this roommate reflects circumstance rather than ability. If the private revision carries a performance cost, the grade drag and the benefits-dimension attitude shift should scale together (fixed ratio γ_U/φ^{ben}) across upper-caste students. This correlation is testable, but the direction of causality between belief and grades is not identified.*

6.3 The Social Life of Integration

This block shares the environment of §6.1 but is otherwise separate from §6.2: grades enter utility (through placement); beliefs, by Assumptions 5–6, neither enter payoffs nor respond to chosen contact.

$$S_i = v(B_i) - c(R_i;\, \Delta_i) - \chi_{c_i}(d_{ij})\, s_i, \qquad Y_i = \pi_{ij}\, R_i + \pi_g\, g_i, \qquad U_i = S_i + Y_i, \tag{13}$$

where $\Delta_i \equiv (\theta_j - \theta_i)_+$ is the preparation gap between roommates (zero in same-caste rooms by convention, positive for a lower-caste student paired above). Throughout the paper Δ_i refers exclusively to this object; the belief correction size $z_i - \psi(b_i)$ from §6.2.2 is never abbreviated. The primitives satisfy: v increasing concave; c increasing convex in R_i and in Δ_i (the psychological burden of bridging rises in the status distance bridged); $\pi_{ij} = \pi(A_j) \geq 0$ the labor-market return to bridging, increasing in the roommate's advantage; $\pi_g \geq 0$ the return to grades.

Timing. (i) α is drawn at random. (ii) Each student chooses desired roommate time s_i^* to maximize (13), taking the partner's choice as given. (iii) Realized interaction is $\tilde{s}_{ij} = \max\{\underline{s},\, \min\{s_i^*, s_j^*\}\}$, $\underline{s} \in [0, 1)$: the min captures the need for mutual participation, the floor $\underline{s}$ the unavoidable contact of co-residence. (iv) Two independent realizations occur: beliefs update off the assignment (d_i, by Assumption 6), while social and academic outcomes are evaluated at $\tilde{s}_{ij}$. These are not a causal chain (Assumptions 5–6). Equation (13) is solved as two single-agent problems rather than a bilateral fixed point: the design runs off random variation in *whom* a student is paired with, not off within-pair equilibrium selection.

$\eta \in (0, 1)$ the elasticity of positions in arena participation.

6.3.1 Optimal allocation

With $a_i = 1 - s_i$, belonging is affine in s: $B_i(s) = \iota_{c_i} + \tau_{c_i}\, s$, where $\tau_c \equiv \lambda_{ij} - \iota_c$ is the bonding pass-through. Since $\iota_U = 1 > \lambda_M$, $\tau_U < 0$: roommate time costs an upper-caste student belonging. τ_L may take either sign.

Lemma 1. *Choosing s is equivalent to choosing B on $[\min\{\lambda_{ij}, \iota_c\},\, \bar{B}_c]$, $\bar{B}_c \equiv \max\{\lambda_{ij}, \iota_c\}$, solving*

$$\max_{s \in [0,1]} \; \Phi_{ij}(B) - \chi_c(d)\, s, \qquad \Phi_{ij}(B) \equiv v(B) + \pi_{ij}(1 - B) - c(1 - B;\, \Delta_i). \tag{14}$$

Φ_{ij} *is strictly concave; the unconstrained optimum* $B^\dagger(\Delta_i, \pi_{ij})$ *solves* $v'(B^\dagger) + c'(1 - B^\dagger; \Delta_i) = \pi_{ij}$*. The optimum is* $B^* = \bar{B}_c$ *whenever* $\Phi'_{ij}(\bar{B}_c) > 0$*, and* $B^* = B^\dagger$ *otherwise. A student with* $\bar{B}_c < B^\dagger$ *is* ***bonding-constrained****: they would prefer more belonging than is attainable.*

Proposition 4 (Who is bonding-constrained)**.** *Define the attainable ceiling* $\bar{B}_i = \max\{\lambda_{ij}, \iota_{c_i}\}$ *if arena-capable (*$A_i \geq A_{c_i}$*),* $\bar{B}_i = \lambda_{ij}$ *if arena-excluded. Student* i *is constrained iff* $\bar{B}_i < B^\dagger$*. Then:*

(i) *Maintaining* $c'(0; \Delta) = 0$*,* $B^\dagger < 1$ *iff* $\pi_{ij} > v'(1)$*, with* $\partial B^\dagger / \partial \Delta_i = c_{R\Delta}/|\Phi''_{ij}| \geq 0$ *and* $\partial B^\dagger / \partial \pi_{ij} = 1/\Phi''_{ij} < 0$*.*

(ii) *Defining* $\underline{\pi}_c(\Delta) \equiv v'(\bar{B}_c) + c'(1 - \bar{B}_c; \Delta)$*, an arena-capable student is unconstrained iff* $\pi_{ij} \geq \underline{\pi}_c(\Delta_i)$*, and* $\underline{\pi}_U < \underline{\pi}_L(\Delta)$ *for all* $\Delta \geq 0$*.*

(iii) *If* $\pi_{ij} \leq v'(1)$ *every student is constrained; if* $\pi_{ij} \geq \underline{\pi}_L(\Delta)$ *none is; between the two thresholds, upper-caste students are unconstrained and lower-caste students remain constrained.*

Corollary 2 (Constraint status)**.** *Constraint status is pinned down by* $(A_i, \iota_{c_i}, \Delta_i)$ *— arena access, arena alignment, realized score gaprefas:arena — shift these jointly. An arena-excluded upper-caste student (*$\bar{B}_i = \lambda_M < 1$*) can be constrained; an arena-capable lower-caste student is unconstrained whenever* $\Delta_i \leq \Delta_{\max}$*,* $B^\dagger(\Delta_{\max}) = \bar{B}_L$*. Constrained students bunch at* $s = 1$ *(*$\tau_c > 0$*) or* $s = 0$ *(*$\tau_c < 0$*); unconstrained students hold interior allocations*

6.3.2 Wellbeing, placement, and positions

Proposition 5. *Achievement–wellbeing divergence*

(i) *For lower-caste students, mixed assignment simultaneously raises expected grades (Proposition 1): academic and social outcomes move in opposite directions from the same assignment, a recomposition from belonging into bridging, not a cost of learning.*

(ii) *Mixed assignment lowers social wellbeing for both students; the loss is larger for the lower-caste student (whose ceiling* $\bar{B}_L < \bar{B}_U$*) and smaller wherever arena alignment or access is greater.*

Corollary 3 (Worry rises in the roommate's score, lower-caste only)**.** *Since the bridging cost rises in* Δ_i*, the wellbeing loss for a lower-caste student rises in the roommate's entrance score and is flat for upper-caste students (*$\Delta_i \approx 0$ *under Assumption 1).*

Corollary 4 (Joint gradient)**.** *For a lower-caste student the roommate's score raises grades, leaves placement flat through the bridging channel (*$\partial R_i / \partial \theta_j = 0$*), and raises the wellbeing cost.*

Corollary 5 (Constraint falls in roommate advantage). *Since interiority requires $\pi_{ij} = \pi(A_j) \geq \underline{\pi}_c(\Delta_i)$, a more advantaged roommate raises the return to bridging and is less likely to leave a student constrained: holding Δ_i fixed, the wellbeing loss attenuates and placement rises in A_j.*

Proposition 6 (Positions). *With $P_i = \iota_{c_i} K_i^{1-\eta} a_i^{\eta}$, $K_i = \lambda_{ij} s_i$:*

(i) *Positions concentrate among upper-caste students through arena alignment and attainment, not through advantage A_i directly.*

(ii) *For upper-caste students, mixing moves the two inputs in opposite directions, $\Delta \log P = (1-\eta)(\log \lambda_M + \Delta \log s) + \eta \, \Delta \log a$, negative iff $\eta < \bar{\eta}$.*

(iii) *Lower-caste students, gated by $\iota_L < 1$ and typically bonding-constrained, do not gain positions despite higher grades—their gains from integration run through bridging and grades into placement (Corollary 4), not through the backing that the position technology requires.*

Proposition 7 (Leadership Positions: backing and arena as complements). *Positions $P_i = \iota_{c_i} K_i^{1-\eta} a_i^{\eta}$ require both backing ($K_i = \lambda_{ij} s_i$, from roommate loyalty) and arena participation (a_i), but the time budget ($s_i + a_i = 1$) forces a trade-off. Positions concentrate among upper-caste students through $\iota_U > \iota_L$ and attainment. Mixing moves the two inputs in opposite directions for upper-caste students—backing down, arena up—so $\Delta \log P = (1-\eta)(\log \lambda_M + \Delta \log s) + \eta \, \Delta \log a$, negative iff $\eta < \bar{\eta}$. Lower-caste students, gated by $\iota_L < 1$ and typically bonding-constrained, do not gain positions despite higher grades; their returns run through placement (Corollary 4).*

6.4 Welfare

Proposition 8 (Sufficient statistic for the unconstrained). *Let $V_c(\lambda) \equiv \max_{s \in [0,1]} \{\Phi_{ij}(\iota_c + (\lambda - \iota_c)\, s) - \chi_c(d)\, s\}$. For any student with $s_c^* < 1$,*

$$\frac{dV_c}{d\lambda} = \Phi'_{ij}(B^*)\, s_c^* = \frac{\chi_c(d)\, s_c^*(\lambda)}{\lambda - \iota_c}, \tag{15}$$

with the primitives v, c, π_{ij} dropping out via the first-order condition. The formula is exact for $s_c^* \in [0, 1)$ and understates *the welfare change (in absolute value) when $s_c^* = 1$—precisely where a bonding-constrained student sits. Welfare is thus exactly measurable for the group whose loss is second-order (unconstrained) and only bounded for the group whose loss is first-order (constrained)—an asymmetry generated by the same institutional feature that produces the inequality being measured.*

Proposition 9 (The marginal switcher prices wellbeing)**.** *For lower-caste students choosing room type ex post, since $\Delta U_i = \Delta S_i + \pi_{ij}\,\Delta R_i + \pi_g\,\Delta g_i$: a student who leaves a mixed room reveals $-\Delta S_i > \pi_{ij}\,\Delta R_i + \pi_g\,\Delta g_i$ (a lower bound on the wellbeing loss in placement units); a stayer reveals the reverse; a marginal (indifferent) student equates the two, point-identifying the wellbeing loss without a functional form for c or χ_c, using only estimable π_g and π_{ij}.*

Theorem 1 (Break-even weight on belief convergence)**.** *Let the planner value $W = \sum_i U_i + \omega_W \sum_i \alpha_i$, with α_i entering W but not U_i (Assumption 5). For a policy with per-student effects $\Delta U < 0 < \Delta\alpha$ (ΔU in placement units, Proposition 9), it is worthwhile iff $\omega_W \geq \bar{\omega}_W \equiv -\Delta U/\Delta\alpha$. Comparing two-person against partially-mixed three-person rooms (near-identical academic gain, differing mainly in the wellbeing/attitude trade),*

$$\bar{\omega}_W^{2v3} = \frac{\Delta U^{(3)} - \Delta U^{(2)}}{\Delta\alpha^{(2)} - \Delta\alpha^{(3)}}.$$

The threshold is informative without committing to a value of ω_W: if implausibly large, two-person rooms are not worth their wellbeing cost regardless of one's view on the value of the attitude shift; if small, the reverse.

Policy: safe spaces. Raising arena alignment ι_L—through caste-specific clubs, mentoring networks, or affinity spaces—relaxes the bonding constraint (Proposition 4), raises the belonging ceiling $\bar{B}_L$, and lowers the wellbeing cost of mixed assignment (Proposition 5). Under Assumption 6, the attitude externality is preserved: belief updating depends on who the roommate is, not on how social time is allocated, so the break-even threshold $\bar{\omega}_W$ falls. The model therefore treats safe spaces and integration as complements rather than substitutes. The limitation is that Assumption 6 is strained when safe spaces enable full disengagement: if the effective signal precision $\tau(d_i)$ degrades with reduced contact, the attitude gain shrinks alongside the wellbeing cost, and the complementarity may not survive.

Policy: assortative matching on non-academic dimensions. Reducing social distance d_{ij} while holding the score gap Δ_i fixed—matching cross-caste roommates who share a language, region, or cultural background—raises roommate alignment λ_{ij}, improving transmission efficiency and grades (equation (5)), while lowering friction $\chi_c(d_{ij})$ and raising the belonging ceiling, reducing the wellbeing cost (Proposition 5). The trade-off is on attitudes: Corollary 1 implies that a less distant roommate generates a less informative signal ($\tau(d)$ falls in d), which could attenuate the correction. Whether

grades and wellbeing gains outweigh any attitude attenuation is an empirical question Theorem 1 is designed to evaluate.

6.5 Discussion

Proposition 5 explains why the same lower-caste student can gain academically while reporting lower wellbeing: the score gap affects grades through academic transmission but affects wellbeing through the cost of bridging status differences, and both effects increase with the roommate's score (Corollary 4). Propositions 4–5 further show that the wellbeing asymmetry originates outside the room, in the outside option each caste has in the social arena. As a result, caste predicts the wellbeing loss only insofar as it predicts arena access and the realized score gap (Corollary 2). The self-image model in §6.2.2 also yields predictions that an exposure-only framework cannot. It identifies both *where* attitudes change—among students with stronger baseline meritocratic beliefs—and *which dimension* changes for each group (Proposition 3). A further prediction, confirmed in the data, is that mixed assignment does not reduce differences in overall support for affirmative action across castes but instead widens them: each group updates most on the dimension about which it was initially least informed, and the lower-caste fairness correction dominates (Proposition 3(iii)). Finally, Propositions 8–9 and Theorem 1 provide a framework for evaluating counterfactual policy interventions by quantifying when the academic and attitudinal benefits of integration outweigh its wellbeing costs. This framework can be used to assess policies such as increasing room size, creating safe spaces for lower-caste students, or matching roommates assortatively on non-academic characteristics.

7 Mechanisms and Discussion

What explains the finding that lower-caste students assigned to mixed rooms experience better academic performance and internship outcomes but worse mental health, while upper-caste students assigned to mixed rooms experience declines in both academic performance and mental health alongside modest improvements in attitudes toward affirmative action? I begin by showing that caste-based frictions persist within diverse rooms, potentially undermining the benefits of inter-caste exposure. To investigate this mechanism, I exploit three additional sources of variation in roommate composition.

First, I analyze how heterogeneity in baseline relative academic distance between roommates, measured using entry-exam performance differences affects roommate performance in mixed and non mixed rooms. If caste frictions interact with academic per-

formance differences, mixed-caste rooms should generate adverse effects, particularly for lower-caste students, relative to similar gaps in same-caste rooms. Second, I use variation in intensity of exposure among students assigned to three-person mixed rooms. Some students share with two roommates of another caste group, while others share with one roommate of their own caste and one of another caste. In the absence of caste-based frictions—if benefits arise purely through learning, peer ability, or access to higher-status social networks—an additional upper-caste roommate should monotonically increase lower-caste students' academic and internship outcomes. However, if caste frictions are present, the marginal effect of an additional upper-caste roommate becomes ambiguous, reflecting a trade-off between learning benefits and social or psychological costs. I also show suggestive evidence of deep vs board interactions with other groups by looking at two sharing and three sharing rooms. Third, I examine heterogeneity by caste distance, showing that treatment effects on how primary outcomes systematically with the social distance between roommates' caste categories and unpack lower castes into OBC and SC/ST show marginal effects of each group.

Finally, I provide evidence on roommate's behavioral outcomes like roommate;s choice in second year, interactions with roommate, including academic help received, financial and emotional support provided and received from roommates.

7.1 Heterogeneity by difference in Baseline Test Scores:

I construct the difference in baseline entrance exam scores between each student and their roommate, average this difference across roommates, and standardize it relative to the distribution of same-caste rooms. The measure is directional—defined as the roommate's score minus the student's score—so that positive values indicate exposure to a relatively higher-performing roommate and negative values indicate exposure to a relatively lower-performing one.

Table 5 exploits heterogeneity in this standardized score gap to distinguish ability-based peer learning from caste-mediated frictions. Among lower-caste students, larger score gaps are associated with significantly better academic outcomes in same-caste rooms, consistent with peer learning. These gains are substantially attenuated in mixed-caste rooms (but not fully): the interaction between score distance and mixed-room assignment is negative and significant at the 5% level for both overall performance and GPA. Effects on GPA and math and programming grades in mixed caste, when there is no difference in test scores between roommates, are about half-size of estimates in table 2 and statistically not different from zero.

First-year internship outcomes tell a different story. Internship receipt is neither increasing in roommate ability in same-caste rooms nor attenuated by score distance in mixed-caste rooms. Instead, lower-caste students appear to benefit from mixed-caste exposure on this margin independently of roommate's ability. Examining reported channels, lower-caste students in mixed rooms are more likely to report effort-based channels—social media, contacting faculty, and job portals, and social capital channel like - help from friends, relative to those in same-caste rooms, this is suggestive evidence as effect on each channel is estimated imprecisely. Overall the internship results point to diversity-driven mechanisms: improved information about application opportunities and access to upper-caste social networks, rather than ability spillovers.

For upper-caste students, neither entrance score distance nor its interaction with mixed-room assignment predicts academic performance. Crucially, the near-zero coefficients reflect small effects rather than imprecise estimation, ruling out the possibility that the academic declines observed for upper-caste students in mixed rooms are driven by exposure to lower-performing lower-caste roommates. Overall these results suggest that mixed-caste exposure generates learning opportunities and informational gains while simultaneously constraining a part of this realization through caste-mediated interaction costs.

Table 6 examines how baseline score differences between roommates shape mental health and attitudes toward affirmative action. The results suggest two distinct channels through which mental health costs accrue to lower-caste students. The first is an caste identity-linked psychological cost that operates in mixed-caste rooms irrespective of the score gap: the PHQ-4 index is elevated in mixed rooms but is neither related to score distance in same-caste rooms nor further affected by score distance in mixed-caste rooms. The second channel operates through perceived competition: the worry index increases by 0.073 standard deviations ($p < 0.1$) when lower-caste students are paired with upper-caste roommates who scored one standard deviation higher at baseline, consistent with stress induced by exposure to higher-ability upper-caste peers. I interpret this pattern as reflecting a combination of competitive pressure and the threat of affirmative-action-linked backlash from upper-caste peers.

For upper-caste students, mixed-caste assignment alone has little direct effect on mental health; all coefficients are small and statistically indistinguishable from zero. Within same-caste rooms, pairing with a higher-performing roommate is associated with worse mental health, consistent with competitive pressure among caste peers. This pattern reverses in mixed-caste rooms, where the positive interaction between score distance and mixed-room assignment suggests that exposure to a higher-performing lower-caste room-

mate may reduce competitive stress. However, this finding warrants cautious interpretation: upper-caste students' roommates in mixed rooms score on average 0.73 standard deviations lower on the entrance exam (with approximately 90% of lower-caste roommates scoring below the upper-caste student in mixed rooms). The relevant variation in mixed rooms is therefore concentrated among upper-caste students paired with considerably lower-performing lower-caste roommates. For the PHQ-4 and worry indices, the interaction term coefficients have higher magnitudes than the score gap coefficients, indicating that exposure to lower-performing roommates leads to relative deprivation[41] effects that dominate any reduced stress from competition. One interpretation centers on relative deprivation linked to affirmative action: upper-caste students may experience resentment or perceived unfairness when observing peers with substantially lower entrance scores gaining admission to the same or better specialization through caste based quotas.

Turning to attitudes toward affirmative action, support among lower-caste students increases across multiple dimensions when they are paired with relatively higher-performing same-caste roommates. High-ability same-caste peers appear to consolidate existing pro-affirmative-action beliefs—both strengthening perceived benefits to beneficiary groups and weakening agreement that affirmative action is unfair to non-beneficiaries. In contrast, exposure to higher-performing roommates in mixed-caste rooms significantly reduces support across all dimensions, suggesting that high-baseline upper-caste peers attenuate belief consolidation. This attenuation is smaller in magnitude than the consolidation effect observed with high-performing same-caste peers, indicating that within-caste role models remain more influential in shaping attitudes. Notably, mixed-caste exposure alone has little independent effect on beliefs of lower caste students; attitude formation operates primarily through ability differences with roommates.

Among upper-caste students, mixed-caste exposure increases support for affirmative action overall. In same-caste rooms, pairing with a relatively higher-performing roommate raises support—equivalently, exposure to a lower-ability same-caste roommate reduces it. Mixed-caste assignment attenuates this pattern: exposure to a relatively higher-ability lower-caste roommate partially offsets the diversity-induced increase in support.

[41] Relative deprivation, as used here, refers to the perceived loss experienced by upper-caste students compared to an admissions system without affirmative action policies. The interpretation is that this sense of loss is heightened when students use their roommates as reference points, comparing where their entrance exam scores would have placed them—in terms of both college admission and specialization—under a merit-only system versus the current affirmative action regime. The direction of score gaps is important for this interpretation, and I preserve the direction of score gaps in the analysis.

Mirroring the mental health results, the score gap between upper-caste students and their lower-caste roommates is large on average, and as this gap widens, overall support declines—the negative effect of score distance dominates the mixed-room interaction coefficient in net. Overall, lower-caste roommates who outscore their upper-caste peers reinforce support for affirmative action, while lower-scoring lower-caste roommates erode it relative to the mixed-room effect at zero score gap.

To show that the direction of entrance exam score gaps matters, I present results using absolute entrance exam score gaps with roommates in Tables A21 and A22. I show that absolute differences with roommates and their interaction with mixed-room assignment do not significantly affect academic outcomes for either lower- or upper-caste students. Similarly, mental health outcomes and support for affirmative action follow the same pattern. These findings support the directional interpretation of mental health costs and attitudes toward affirmative action linked to relative deprivation due to affirmative action policies in earlier tables. Two notable exceptions are internship outcomes for upper-caste students and relative well-being for both groups. For lower-caste students, larger score gaps with roommates indicate worse relative well-being, and for upper-caste students, larger score gaps with roommates indicate better well-being, suggesting that students' comparison points with roommates—and whether they are above (for lower-caste students) or below (for upper-caste students) their roommates—determine their relative well-being.

A plausible channel for both the mental health costs and the reduction in affirmative action support is relative deprivation arising from failure to obtain a preferred specialization. Specialization dissatisfaction is strongly negatively correlated with both the PHQ-4 and worry indices, and affirmative action is commonly perceived in this setting as conferring an unfair advantage that allows lower-caste students with lower baseline scores to access preferred specializations such as computer science. The mental health and attitude results—particularly the costs borne by upper-caste students—are consistent with this channel, suggesting that perceived losses from affirmative action in high-stakes assignment decisions, alongside more modest cultural-distance costs, may be a primary driver of the identity-based burdens documented throughout. Exposure to a relatively higher-performing lower-caste roommate attenuates academic performance, increases worry about college-related issues, and reduces support for affirmative action among upper-caste students relative to exposure to an equally performing same-caste roommate—further implicating score-gap-based relative deprivation as a distinct and reinforcing mechanism.

Table 5: Effect on GPA, Internships, Maths and Programming performance: Hetrogeneity by difference of entrance scores with Roommate

	Performance Index (1)	GPA (2)	Maths & Programming (3)	Internship (4)
Panel A: Lower Caste Sample				
Mixed Room	0.120**	0.055	0.059	0.057***
	(0.053)	(0.062)	(0.069)	(0.020)
Entrance Exam Score diff with RM (z-score)	0.091***	0.117***	0.094**	0.008
	(0.029)	(0.036)	(0.044)	(0.011)
Entrance Exam Score diff with RM x Mixed Room	-0.083**	-0.092**	-0.057	-0.014
	(0.032)	(0.041)	(0.048)	(0.013)
DV Mean in same caste rm	-0.357	-0.466	-0.411	0.032
R-squared	0.343	0.441	0.377	0.083
N	1059	968	1052	1047
Panel B: Upper Caste Sample				
Mixed Room	-0.073	-0.116*	-0.061	0.020
	(0.055)	(0.061)	(0.057)	(0.023)
Entrance Exam Score diff with RM (z-score)	-0.013	0.020	-0.002	0.004
	(0.040)	(0.047)	(0.043)	(0.009)
Entrance Exam Score diff with RM x Mixed Room	0.000	-0.029	-0.003	-0.009
	(0.044)	(0.051)	(0.047)	(0.014)
DV Mean in same caste rm	0.348	0.453	0.458	0.064
R-squared	0.344	0.474	0.421	0.073
N	1078	1043	1076	1069
Panel C: Full Sample				
Mixed Room	0.034	-0.020	0.029	0.036***
	(0.032)	(0.037)	(0.038)	(0.013)
Entrance Exam Score diff with RM (z-score)	0.056**	0.093***	0.054	0.007
	(0.023)	(0.030)	(0.034)	(0.007)
Entrance Exam Score diff with RM x Mixed Room	-0.034	-0.065**	-0.022	-0.008
	(0.025)	(0.032)	(0.037)	(0.008)
DV Mean in same caste rm	-0.009	0.000	0.018	0.048
R-squared	0.411	0.518	0.458	0.048
N	2137	2011	2128	2116

Notes: The table presents results for academic performance outcomes, including GPA, Mathematics and Programming grades, and summer internship or summer academic project completion. I look at heterogeneity by including an interaction of mixed room with standardized difference with roommates entrance exam score. I took average difference difference of extrance exam score with roommates and then standardized it. Column 1 reports estimates for a composite Performance Index constructed using standardized scores of cumulative GPA, Mathematics and Programming grades, and summer internship or project completion. Column 2 reports results for cumulative GPA, defined as the standardized cumulative GPA at the end of the first year. Column 3 reports results for Mathematics and Programming grades, measured as an index constructed using standardized scores from Mathematics and Programming courses. Column 4 reports results for summer internship or academic project completion, defined as an indicator for whether a student completed a summer internship or academic project during the summer following the first year. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective college entrance exam scores and number of roommates.

Table 6: Effect on Mental Health and Affirmative action: Hetrogeneity by difference of entrance scores with Roommate

	Mental Health Index (1)	PHQ4 Index (2)	Relative well being Index (3)	Worry Index (4)	Supports AA Index (5)	Benefits of AA to reserved groups (6)	Costs of AA to non reserved groups (rev) (7)
Panel A: Lower Caste Sample							
Mixed Room	0.078	0.087	0.023	0.001	-0.015	-0.075	0.039
	(0.051)	(0.077)	(0.052)	(0.055)	(0.052)	(0.052)	(0.064)
Entrance Exam Score diff with RM (z-score)	-0.046	0.003	-0.040	-0.036	0.116***	0.121***	0.113***
	(0.033)	(0.047)	(0.033)	(0.035)	(0.027)	(0.026)	(0.042)
Entrance Exam Score diff with RM x Mixed Room	0.052	-0.008	0.032	0.073*	-0.065*	-0.075**	-0.054
	(0.038)	(0.055)	(0.036)	(0.039)	(0.035)	(0.034)	(0.049)
DV Mean in same caste rm	0.014	0.010	0.036	0.046	0.240	0.239	0.237
R-squared	0.117	0.099	0.089	0.117	0.101	0.089	0.121
N	985	934	850	892	888	883	874
Panel B: Upper Caste Sample							
Mixed Room	0.031	0.040	0.034	-0.066	0.112*	0.140**	0.080
	(0.056)	(0.083)	(0.060)	(0.061)	(0.064)	(0.068)	(0.071)
Entrance Exam Score diff with RM (z-score)	0.090**	0.136**	0.062	0.048	0.113**	0.103**	0.127**
	(0.040)	(0.054)	(0.050)	(0.043)	(0.046)	(0.047)	(0.053)
Entrance Exam Score diff with RM x Mixed Room	-0.108**	-0.162**	-0.050	-0.102**	-0.082	-0.066	-0.095
	(0.047)	(0.065)	(0.056)	(0.049)	(0.054)	(0.055)	(0.064)
DV Mean in same caste rm	-0.046	-0.016	-0.055	-0.052	-0.256	-0.251	-0.265
R-squared	0.104	0.102	0.055	0.108	0.122	0.131	0.116
N	1031	971	846	896	911	907	901
Panel C: Full Sample							
Mixed Room	0.076**	0.091*	0.018	0.025	0.074*	0.043	0.099**
	(0.034)	(0.047)	(0.032)	(0.033)	(0.038)	(0.037)	(0.047)
Entrance Exam Score diff with RM (z-score)	-0.011	0.031	-0.006	-0.011	0.098***	0.104***	0.094**
	(0.027)	(0.037)	(0.028)	(0.027)	(0.029)	(0.022)	(0.046)
Entrance Exam Score diff with RM x Mixed Room	0.011	-0.038	0.003	0.012	-0.070**	-0.085***	-0.050
	(0.029)	(0.040)	(0.030)	(0.029)	(0.034)	(0.029)	(0.051)
DV Mean in same caste rm	-0.016	-0.003	-0.009	-0.003	-0.005	-0.002	-0.012
R-squared	0.074	0.063	0.036	0.077	0.187	0.175	0.162
N	2016	1905	1696	1788	1799	1790	1775

Notes: The table presents results for mental health outcomes, including the Mental Health Index, PHQ-4 scores, relative happiness, and worry index scores. I look at heterogeneity by including an interaction of mixed room with score difference with roommates entrance exam score (roommate - student). I took average absolute difference of entrance exam score with roommates and then standardized it. Column 1 reports estimates for a composite Mental Health Index constructed using standardized scores of the individual mental health measures. Column 2 reports results for PHQ-4 scores, defined as the standardized PHQ-4 score measuring anxiety and depression symptoms. Column 3 reports results for worry index scores, defined as the standardized worries about various college aspects score. Column 4 reports estimates for a composite Affirmative Action Support Index constructed by mean of dependent variables in columns 5 and 6. Column 5 reports results for metritocratic concerns of affirmative action policies, and are reverse coded for easier interpretation. Column 6 reports results for social justice aspect of affirmative action. In all the columns, higher values of these indices indicate worse mental health outcomes. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective dependent variables and a missing indicator for the baseline controls and number of roommates.

7.2 Effects by intensity of exposure:

Table 7 examines how the intensity and degree of mixed-caste exposure shapes academic outcomes. Lower-caste students in partially mixed rooms—sharing with one lower-caste student and one upper-caste student—experience a 0.192 standard deviation increase in overall academic performance, compared to a 0.097 standard deviation increase when sharing rooms with all upper-caste roommates. These differences are driven primarily by improvements in mathematics and programming grades, as well as internship outcomes. These findings provide additional evidence that peer ability alone does not explain the better performance of lower-caste students in mixed-caste rooms. If peer ability were the sole mechanism, effects should be weakly larger in fully mixed rooms where exposure to upper-caste roommates is maximized. For upper-caste students, academic declines are similar across partially and fully mixed rooms, and in some specifications are slightly more negative in partially mixed settings.

Table 8 examines mental health and attitudes toward affirmative action by exposure intensity. For lower-caste students, mental health outcomes are broadly similar in partially and fully mixed rooms. AA Support is slightly higher in partial mixed rooms, driven by difference being that students in partial mixed rooms are more likely to believe affirmative action policies don't impose costs on upper-caste students. In contrast, upper-caste students experience significantly worse mental health in fully mixed rooms—where both roommates are from lower-caste groups—with the mental health index increasing by 0.101 standard deviations relative to near-zero effects in partially mixed rooms (significantly different at the 10% level), driven primarily by the PHQ-4 measure. Affirmative-action attitudes among upper-caste students remain broadly similar across exposure levels, with slightly more positive point estimates—particularly regarding disagreement about the costs of affirmative action—in partially mixed rooms.

To examine the robustness of these intensity-of-exposure results, I conduct subsample analyses restricting attention to three-person rooms (Tables A11, A12) and two-person rooms (Tables A13, A14). Room size generates theoretically distinct forms of exposure. Three-person rooms provide either broader diversity—exposure to two opposite-caste roommates in fully mixed configurations—or balanced diversity with built-in support, where students have one same-caste and one opposite-caste roommate in partially mixed configurations. In contrast, two-person mixed rooms create deeper and more intensive interaction with a single opposite-caste roommate, where contact is nearly unavoidable and undiluted by the presence of other roommates.

Among lower-caste students, academic gains are positive and similar in magnitude across both room sizes. Within three-person rooms, partially mixed configurations con-

sistently yield larger point estimates (though not statistically different) than fully mixed rooms. But internship gains are stronger in three-person rooms, while effects in two-person mixed rooms are small and slightly negative, aligning with earlier findings that network formation and peer learning channels operate more effectively in larger rooms. Mental health costs are higher in three-person rooms relative to two-person rooms, with similar effects across partially and fully mixed configurations, driven by PHQ-4 effects. Worry index, though imprecisely estimated, are greater in fully mixed rooms—consistent with earlier finding on exposure to higher relative roommate performance gap. For upper-caste students, academic performance declines and mental health worsens more sharply in two-person mixed rooms than in any three-person configuration. Mental health costs are largely mitigated when upper-caste students share a three-person mixed room with at least one same-caste peer. Changes in attitudes toward affirmative action show slightly more positive views in three-person configurations, supporting the view that broader contact with diverse roommates may improve support for affirmative action.

These intensity-of-exposure results reveal that the costs and benefits of inter-caste contact vary systematically with the structure of exposure. For lower-caste students, mixed rooms generate academic benefits through improved access to networks and higher-performing upper-caste peers, but these gains are partially offset by two forms of psychological cost: increased stereotype salience linked to the presence of upper-caste peers, and heightened worry arising from exposure to upper caste roommates, more so when entrance score gap is higher. The presence of at least one same-caste roommate in partially mixed configurations appears to mitigate caste-based frictions and improve performance, potentially through reduced social isolation or increased emotional support.

For upper-caste students, the psychological and academic costs are not driven solely by pairing with lower-performing roommates. Instead, different structures of exposure activate distinct mechanisms. Intensive contact in two-person mixed rooms—where interaction cannot be avoided—produces the largest costs, consistent with relative deprivation intensifying through deeper, unavoidable contact with a single lower-caste roommate. In contrast, three-person rooms allow social buffering through same-caste support and more selective interaction. Exposure to multiple diverse roommates in three-person configurations may also broaden perspectives, potentially mitigating some mental health costs and improving attitudes toward affirmative action. Overall, the intensity-of-exposure results reinforce the distinct costs and benefits of inter-caste contact for both upper- and lower-caste students, suggesting that contact quality and room composition matter as much as

contact itself.

Table 7: Effect on GPA, Internships, Maths and Programming performance

	Performance Index (1)	GPA (2)	Maths & Programming (3)	Summer Project or Internship (4)
Panel A: Lower Caste Sample				
Partial Mixed Room	0.192***	0.108	0.151**	0.102***
	(0.057)	(0.070)	(0.074)	(0.033)
Full Mixed Room	0.097*	0.081	0.077	0.043
	(0.050)	(0.058)	(0.065)	(0.028)
P-val (Partial = Full)	0.091	0.720	0.298	0.105
DV Mean in same caste rm	-0.357	-0.466	-0.411	0.097
R-squared	0.339	0.435	0.374	0.074
N	1059	968	1052	1047
Panel B: Upper Caste Sample				
Partial Mixed Room	-0.079	-0.086	-0.046	-0.030
	(0.058)	(0.062)	(0.057)	(0.044)
Full Mixed Room	-0.047	-0.112**	-0.059	0.007
	(0.049)	(0.057)	(0.049)	(0.034)
P-val (Partial = Full)	0.555	0.688	0.822	0.367
DV Mean in same caste rm	0.348	0.453	0.458	0.221
R-squared	0.344	0.474	0.421	0.101
N	1078	1043	1076	1069

Notes: The table presents results for academic performance outcomes, including GPA, Mathematics and Programming grades, and summer internship or summer academic project completion. Column 1 reports estimates for a composite Performance Index constructed using standardized scores of cumulative GPA, Mathematics and Programming grades, and summer internship or project completion. Column 2 reports results for cumulative GPA, defined as the standardized cumulative GPA up to the end of the first year. Column 3 reports results for Mathematics and Programming grades, measured as an index constructed using standardized scores from Mathematics and Programming courses. Column 4 reports results for summer internship or academic project completion, defined as an indicator for whether a student completed a summer internship or academic project during the summer following the first year. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective college entrance exam scores and number of roommates.

Table 8: Effect on Mental Health & Affirmative Action Support

	Mental Health Index (1)	PHQ4 Index (2)	Relative well being Index (3)	Worry Index (4)	Supports AA Index (5)	Benefits of AA to reserved groups (6)	Costs of AA to non reserved groups (rev) (7)
Panel A: Lower Caste Sample							
Partial Mixed Room	0.076	0.130	0.001	0.013	0.079	-0.001	0.157**
	(0.059)	(0.081)	(0.055)	(0.059)	(0.052)	(0.052)	(0.067)
Full Mixed Room	0.085*	0.049	0.022	0.049	0.026	-0.035	0.081
	(0.047)	(0.071)	(0.045)	(0.051)	(0.050)	(0.051)	(0.062)
P-val (Partial = Full)	0.873	0.330	0.710	0.566	0.312	0.510	0.253
DV Mean in same caste rm	0.014	0.010	0.036	0.046	0.240	0.239	0.237
R-squared	0.145	0.122	0.116	0.143	0.118	0.111	0.132
N	985	934	850	892	888	883	874
Panel B: Upper Caste Sample							
Partial Mixed Room	-0.005	-0.045	0.001	-0.010	0.103	0.136*	0.081
	(0.066)	(0.091)	(0.064)	(0.063)	(0.073)	(0.077)	(0.086)
Full Mixed Room	0.101**	0.156**	0.023	0.028	0.081	0.104*	0.031
	(0.050)	(0.073)	(0.056)	(0.050)	(0.057)	(0.059)	(0.063)
P-val (Partial = Full)	0.074	0.023	0.728	0.544	0.756	0.648	0.573
DV Mean in same caste rm	-0.046	-0.016	-0.055	-0.052	-0.256	-0.251	-0.265
R-squared	0.139	0.150	0.091	0.112	0.196	0.173	0.185
N	1031	971	846	896	911	907	901

Notes: The table presents results for mental health outcomes, including the Mental Health Index, PHQ-4 scores, relative happiness, and worry index scores. Column 1 reports estimates for a composite Mental Health Index constructed using standardized scores of the individual mental health measures. Column 2 reports results for PHQ-4 scores, defined as the standardized PHQ-4 score measuring anxiety and depression symptoms. Column 3 reports results for worry index scores, defined as the standardized worries about various college aspects score. Column 4 reports estimates for a composite Affirmative Action Support Index constructed by mean of dependent variables in columns 5 and 6. Column 5 reports results for metritocratic concerns of affirmative action policies, and are reverse coded for easier interpretation. Column 6 reports results for social justice aspect of affirmative action. In all the columns, higher values of these indices indicate worse mental health outcomes. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective dependent variables and a missing indicator for the baseline controls and number of roommates.

7.3 Effects by Caste Distance:

Both tests above provide evidence of identity-linked costs[42]: psychological costs, academic performance declines not explained by score gaps with roommates, and shifts in affirmative action attitudes linked to exposure diversity. I use additional variation in exposure linked to students' positions in the traditional caste hierarchical structure and estimate the effect of a continuous measure of caste distance between a student and their roommate in Tables 9 and 10, rather than a binary indicator of mixed versus non-mixed rooms. I define caste order as an ordinal variable with upper-caste students assigned 1, OBC students 2, and SC/ST students 3. To define caste distance with roommates, I calculate the absolute caste distance between each student and each roommate, then take the mean distance across all roommates. The academic results in Table 9 show little evidence that caste distance benefits lower-caste students' performance: coefficients are small and statistically insignificant across GPA, mathematics and programming grades, internships, and the overall performance index. These null effects are not driven by noisy estimates—the point estimates are small, negative in sign, and statistically indistinguishable from zero. This is consistent with lower-caste students deriving academic benefits primarily through assignment to relatively high-performing students rather than through caste distance per se. In contrast, upper-caste students experience significant performance declines as caste distance increases (a reduction of 0.090 standard deviations in the performance index, $p < 0.05$), driven particularly by GPA (a reduction of 0.129 standard deviations, $p < 0.05$, for a one-unit increase in caste distance, corresponding to moving from exposure to OBC roommates to SC/ST roommates). Similar declines appear in mathematics and programming performance (0.094 standard deviations, $p < 0.05$). Consistent with earlier results, there is no negative effect on internships. All regressions control for entrance score gaps between roommates. I also present findings without entrance score gap controls in Table A23, which align with this interpretation. These findings reinforce that the academic costs of inter-caste exposure are concentrated among upper-caste students and are unlikely to be driven purely by peer-ability differences, but instead reflect interaction costs or frictions associated with exposure to socially distant caste groups.

Table 10 examines effects on mental health and attitudes toward affirmative action. Greater caste distance is associated with worse mental health for both lower- and upper-caste students, with positive and statistically significant effects on the mental health and

[42]By identity-linked costs, I refer to the component of the mixed-room effect that remains after accounting for heterogeneity in score gaps between roommates. This component may also be correlated with other socioeconomic indicators associated with caste that are not partialed out in the analysis. However, systemic caste inequalities encompass these differences as part of caste identity itself. Attributing residual costs and analyzing them through a caste identity lens is not unreasonable.

PHQ-4 indices for both groups. Relative well-being indices are also positive, indicating worse well-being, but are imprecisely estimated. The worry index for lower-caste students is sizable but imprecisely estimated. At the same time, caste distance increases support for affirmative action, particularly among upper-caste groups, suggesting that exposure to socially distant peers may raise awareness of inequality and strengthen support for redistributive policy. Among lower-caste students, these effects are driven by reduced agreement with the perceived costs of affirmative action for non-beneficiary groups: greater social distance makes lower-caste students less likely to view affirmative action as unfair to upper-caste students, though this effect is imprecisely estimated. Overall, these results indicate that both mental health costs and belief updating are linked to caste distance. I also present findings without entrance score gap controls in Table A24, which align with this interpretation.

For lower-caste students, point estimates for mental health and affirmative action attitudes using the continuous caste-distance measure are comparable in magnitude to those obtained using the binary mixed-caste exposure indicator—in fact, they are larger after controlling for score gaps. For upper-caste students, caste-distance estimates are similar in size to the corresponding binary-exposure effects. Combined with earlier findings on heterogeneity by roommate exam score difference, this pattern suggests that upper-caste students' costs arise from two distinct components: identity-based interaction frictions that scale with caste distance, and psychological costs associated with exposure to lower-performing lower-caste peers and perceived losses from affirmative action policies.

To further unpack caste distance and identify which hierarchical pairings drive the effects, I estimate the marginal effect of exposure to each caste group separately—for example, the effect on an upper-caste student of exposure to OBC versus SC/ST roommates, and analogous comparisons for the other caste groups. These estimates are presented in Tables A15 and A16. The academic declines for upper-caste students in the caste-distance specification are driven primarily by exposure to SC/ST roommates. Exposure to OBC students does not reduce academic outcomes for upper-caste students—effects are small in magnitude and statistically insignificant—consistent with the interpretation that social distance with roommates underlies academic performance costs. Results for OBC students show an interesting asymmetry based on whether they are exposed to upper-caste students or SC/ST students. They show improved overall academic performance when exposed to upper-caste roommates: small and positive effects on GPA and mathematics/programming grades, and strong positive effects on internships. This pattern mirrors earlier findings: exposure to upper-caste peers generates benefits similar to those observed for lower-caste students in mixed rooms, while exposure to SC/ST peers mirrors

the negative academic effects experienced by upper-caste students exposed to lower-caste roommates. The internship gains observed for lower-caste students in Table 2 appear largely driven by positive effects for OBC students when exposed to upper-caste peers. P-values on comparisons between marginal exposure to upper-caste versus marginal exposure to SC/ST students on the performance index, GPA, and internships are significantly different from zero ($p < 0.05$). SC/ST students experience academic gains when paired with higher-caste roommates (both General and OBC), reinforcing the interpretation that upward peer exposure facilitates learning and skill accumulation. However, the absence of corresponding internship gains suggests the presence of additional frictions—such as weaker networks, relative performance gaps, labor-market discrimination, or effort reallocation toward academic rather than professional outcomes.

Upper-caste students become more supportive of affirmative action primarily when exposed to SC/ST roommates, suggesting belief updating in response to visible disadvantage and greater exposure to inequality. Mental health effects for upper-caste students remain statistically indistinguishable from zero and similar in magnitude across exposure to OBC and SC/ST roommates. For OBC students, mental health effects are not statistically significant, although PHQ-4 estimates are sizable and imprecisely estimated across both upward and downward caste exposure. Attitudinal responses are more complex: exposure to upper-caste roommates increases support for affirmative action, mainly by reducing agreement with its perceived costs to upper-caste groups, whereas exposure to SC/ST roommates shifts attitudes in the opposite direction, consistent with potential backlash among relatively higher-status lower-caste students. The opposite direction of attitudinal shifts and performance effects, combined with similar-sized mental health effects when exposed to upper-castes versus lower-castes, provides suggestive evidence that mental health costs may arise from different sources depending on exposure type and the associated benefits or burdens. OBCs exposed to upper-caste students may experience stress from academic competition, while OBCs exposed to SC/ST students may face costs similar to those upper-caste students experience when exposed to lower-caste peers. Conversely, SC/ST students experience worse mental health only when exposed to upper-caste roommates, not when exposed to OBC roommates, and they also reduce support for affirmative action when exposed to both upper-caste and OBC peers. This pattern may reflect backlash, discrimination, or convergence of beliefs toward dominant-group views through interaction.

Overall, this decomposition highlights status-hierarchical exposure dynamics: upward exposure generates learning benefits but also imposes psychological costs that scale with caste distance for the most disadvantaged students. Psychological costs are not lin-

ear in caste distance for upper-caste and OBC students, highlighting the dual role of affirmative action policies and caste-based social distance in shaping the cost structure. Attitudes toward affirmative action policies are similarly updated based on recognition of disadvantage, but are also mediated by perceived costs linked to affirmative action.

Table 9: Effect on GPA, Internships, Maths and Programming performance: Effects by caste distance

	Performance Index	GPA	Maths & Programming	Internship
	(1)	(2)	(3)	(4)
Panel A: Lower Caste Sample				
Caste distance with roommates	-0.039	-0.039	-0.037	-0.001
	(0.038)	(0.051)	(0.056)	(0.016)
Entrance Exam Score diff with RM (z-score)	0.066***	0.072***	0.074***	0.011
	(0.019)	(0.023)	(0.026)	(0.009)
DV Mean in same caste rm	-0.284	-0.474	-0.307	0.025
R-squared	0.338	0.438	0.376	0.075
N	1059	968	1052	1047
Panel B: Upper Caste Sample				
Caste distance with roommates	-0.090**	-0.129**	-0.094**	0.009
	(0.044)	(0.059)	(0.044)	(0.016)
Entrance Exam Score diff with RM (z-score)	-0.036	-0.036	-0.034	-0.004
	(0.026)	(0.032)	(0.027)	(0.010)
DV Mean in same caste rm	0.353	0.459	0.464	0.064
R-squared	0.345	0.476	0.422	0.072
N	1076	1041	1074	1067
Panel C: Full Sample				
Caste distance with roommates	-0.022	-0.035	-0.025	0.009
	(0.024)	(0.034)	(0.030)	(0.009)
Entrance Exam Score diff with RM (z-score)	0.027**	0.039**	0.033*	0.002
	(0.013)	(0.017)	(0.017)	(0.006)
DV Mean in same caste rm	-0.009	0.000	0.018	0.048
R-squared	0.410	0.517	0.458	0.044
N	2135	2009	2126	2114

Notes: The table presents results for academic performance outcomes, including GPA, Mathematics and Programming grades, and summer internship or summer academic project completion. Instead of looking at the effect of mixed room, I look at the effect of caste distance with roommates on these outcomes. I took average distance of caste (absolute) with roommates and then standardized it. I assigned upper caste as 1, OBC as 2, SC and ST as 3 for this purpose and took absolute distance with roommates and took average with roommates. Column 1 reports estimates for a composite Performance Index constructed using standardized scores of cumulative GPA, Mathematics and Programming grades, and summer internship or project completion. Column 2 reports results for cumulative GPA, defined as the standardized cumulative GPA at the end of the first year. Column 3 reports results for Mathematics and Programming grades, measured as an index constructed using standardized scores from Mathematics and Programming courses. Column 4 reports results for summer internship or academic project completion, defined as an indicator for whether a student completed a summer internship or academic project during the summer following the first year. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective college entrance exam scores and number of roommates and also control for the standardized difference in entrance exam scores with roommates.

Table 10: Effect on Mental Health and Affirmative action: Effects by caste distance

	Mental Health Index (1)	PHQ4 Index (2)	Relative well being Index (3)	Worry Index (4)	Supports AA Index (5)	Benefits of AA to reserved groups (6)	Costs of AA to non reserved groups (rev) (7)
Panel A: Lower Caste Sample							
Caste distance with roommates	0.112***	0.112*	0.053	0.052	0.013	-0.032	0.054
	(0.041)	(0.059)	(0.042)	(0.047)	(0.041)	(0.045)	(0.050)
Entrance Exam Score diff with RM (z-score)	-0.012	-0.004	-0.022	0.008	0.067***	0.066***	0.069***
	(0.022)	(0.031)	(0.022)	(0.023)	(0.019)	(0.020)	(0.025)
DV Mean in same caste rm	0.010	0.008	0.049	0.059	0.250	0.251	0.246
R-squared	0.118	0.102	0.090	0.115	0.096	0.078	0.120
N	985	934	850	892	888	883	874
Panel B: Upper Caste Sample							
Caste distance with roommates	0.077*	0.135**	0.060	-0.009	0.123***	0.137***	0.098*
	(0.040)	(0.061)	(0.052)	(0.044)	(0.046)	(0.048)	(0.054)
Entrance Exam Score diff with RM (z-score)	0.027	0.053	0.040	-0.026	0.071**	0.073**	0.071**
	(0.027)	(0.041)	(0.035)	(0.030)	(0.029)	(0.029)	(0.033)
DV Mean in same caste rm	-0.049	-0.021	-0.057	-0.055	-0.262	-0.256	-0.273
R-squared	0.100	0.099	0.055	0.104	0.119	0.130	0.111
N	1029	969	844	894	909	905	899
Panel C: Full Sample							
Caste distance with roommates	0.061**	0.087**	0.019	0.028	0.077***	0.057**	0.089***
	(0.025)	(0.036)	(0.025)	(0.027)	(0.026)	(0.026)	(0.031)
Entrance Exam Score diff with RM (z-score)	0.005	0.012	-0.001	0.002	0.048***	0.042***	0.058***
	(0.015)	(0.021)	(0.015)	(0.015)	(0.014)	(0.014)	(0.018)
DV Mean in same caste rm	-0.016	-0.003	-0.009	-0.003	-0.005	-0.002	-0.012
R-squared	0.075	0.063	0.036	0.077	0.187	0.173	0.163
N	2014	1903	1694	1786	1797	1788	1773

Notes: The table presents results for mental health outcomes, including the Mental Health Index, PHQ-4 scores, relative happiness, and worry index scores. Instead of looking at the effect of mixed room, I look at the effect of caste distance with roommates on these outcomes. I took average distance of caste (absolute) with roommates and then standardized it. I assigned upper caste as 1, OBC as 2, SC and ST as 3 for this purpose and took absolute distance with roommates and took average with roommates. Column 1 reports estimates for a composite Mental Health Index constructed using standardized scores of the individual mental health measures. Column 2 reports results for PHQ-4 scores, defined as the standardized PHQ-4 score measuring anxiety and depression symptoms. Column 3 reports results for worry index scores, defined as the standardized worries about various college aspects score. Column 4 reports estimates for a composite Affirmative Action Support Index constructed by mean of dependent variables in columns 5 and 6. Column 5 reports results for metritocratic concerns of affirmative action policies, and are reverse coded for easier interpretation. Column 6 reports results for social justice aspect of affirmative action. In all the columns, higher values of these indices indicate worse mental health outcomes. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective dependent variables and a missing indicator for the baseline controls and number of roommates. also control for the standardized difference in entrance exam scores with roommates.

7.4 Friendship Networks and Interactions

In Tables 11–14, I examine revealed-preference outcomes related to roommate choice, interpersonal relationships with roommates, types of interactions, and transactions of help between roommates. These results help unpack what happens within rooms and how each group optimizes interactions in mixed- and non-mixed-caste settings. First (Table 11), students assigned to mixed rooms in their first year—both upper- and lower-caste—are more likely to live in mixed rooms again in their second year. Because students are allowed to choose their own roommates at the end of the first year, subject to remaining within the same hostel and cohort, this persistence may partly reflect mechanical switching costs: changing roommates requires coordination with others who are also willing to switch. I also find that lower- and upper-caste students are similarly likely to continue living with the same roommate.[43] Although upper-caste students are approximately six percentage points less likely to remain with the same roommate, this difference is not statistically significant. Conditional on switching roommates, lower-caste students are less likely to choose mixed rooms in the second year, whereas upper-caste students are about equally likely to select mixed or same-caste rooms. Descriptive means from the control group further suggest that upper-caste students who move away from their original roommates tend to pair with lower-caste students who previously lived in non-mixed rooms. Consistent with this pattern, upper-caste students also report living fewer total years (across the first and second years combined) with their original roommate in survey responses, a finding that aligns with the administrative data.

Second, I examine interactions between roommates and how they differ between mixed- and non-mixed-caste rooms (Table 12). In same-caste rooms, lower-caste students tend to interact more with their roommates than upper-caste students do with same-caste roommates. One possible interpretation for this higher social connectedness among lower-caste roommates is that difficulties navigating elite institutional spaces may increase co-dependence on in-group peers. Lower-caste students in mixed rooms are overall less likely to interact with their roommate, particularly in social interactions such as eating together[44], playing sports together, and going out. Upper-caste students in

[43] In the second year, most two-person rooms remained two-person rooms and most three-person rooms remained three-person rooms. However, in some hostels, students who shared two-person rooms in the first year moved to three-person rooms in the second year, and vice versa. I construct a dummy for each initially assigned roommate indicating whether that roommate was still living with the student in the second year; for three-person rooms, I average across the two roommates. The data show that roughly 50% of students continue living with at least one of their first-year roommates.

[44] Students have common dining rooms in the hostels, and about 50% report eating together more than a few times per week. Anecdotal evidence suggests hostel menus are typically vegetarian, with non-vegetarian food provided once or twice a week on a pay-per-meal basis. Part of this pattern may reflect

mixed rooms—despite having lower mean interactions even in same-caste rooms—also consistently report lower interactions with their roommate in mixed rooms, although these estimates are imprecisely estimated.

Third, I examine patterns of help received from and provided to roommates (Table 13 and Table 14). Lower-caste students report lower financial and emotional support, as well as less support in dealing with stressful situations, but higher academic support from their upper-caste roommates. Interestingly, despite stronger internship outcomes for lower-caste students, they report lower explicit help from roommates in mixed rooms compared with same-caste rooms. Upper-caste students report providing some internship-related help to lower-caste roommates in mixed rooms, although these estimates are imprecisely estimated. One possibility is that lower-caste students provide more internship-related support within same-caste rooms, while the quality of explicit help in mixed rooms may be higher due to greater social-capital networks among upper-caste students. Alternatively, this pattern may reflect exposure effects, where lower-caste students learn implicitly or get information on how to obtain internships. Upper-caste students report receiving less academic- and internship-related help when assigned to mixed rooms, as well as lower emotional support in dealing with stressful situations. Patterns in help provided mirror earlier results on help received: upper-caste students report providing financial help to lower-caste students. The negative estimate for financial help received by lower-caste students in mixed rooms (column 4, Table 13) can be rationalized by higher support within same-caste rooms; even when upper-caste roommates provide help, it remains lower than the support lower-caste students receive from their own group.

Fourth, I report estimates on friendship networks and types of friendship networks in Table 15. Lower-caste students report fewer friends at other elite universities, significant at the 5% level. These other elite universities are considered lower-ranking than the university in this study. If mental health costs—possibly due to backlash from the benefits of affirmative action policies—are salient in mixed rooms, hostile attitudes from friends at other elite universities who could not gain admission to more prestigious institutions may contribute to reduced friendship formation. Upper-caste students form more other-caste friendships within the university. Upper-caste students in non-mixed-caste rooms report a very low share of other-caste friends compared to lower-caste students in non-mixed rooms. Despite social frictions, mixed-caste roommate assignment increased their

food preferences across caste groups, as some upper-caste groups are culturally less likely to consume non-vegetarian food. Students sometimes eat in eateries inside the university premises where both vegetarian and non-vegetarian options are available.

share of other-caste friends.

Overall, the combined evidence supports the presence of social frictions through lower social interactions and reduced emotional support from roommates in mixed rooms. However, consistent with positive academic and internship results, patterns of student interactions and help provision provide strong evidence for positive effects on academic outcomes alongside negative effects on mental health.

7.5 Additional measures to interpret mechanism:

Negative mental health effects do not appear to be driven by isolation among students in mixed-caste rooms. I asked students how informed they felt about various aspects of college life, and they do not appear to be less informed than students in same-caste rooms. Estimates reported in Table A17 are almost uniformly indistinguishable from zero.

Negative academic outcomes for upper-caste students also do not seem to arise from crowding out of their own study time due to helping lower-caste peers. Upper-caste students in mixed-caste rooms report more study hours during normal weeks and similar study hours during exam weeks (Table A19). Both upper- and lower-caste students report lower class attendance when assigned to mixed rooms. This pattern could be driven by worse mental health outcomes, with both groups skipping classes relative to their counterparts in same-caste rooms. Upper-caste students additionally report lower self-ratings in math, programming, and English—consistent with objective grade-based assessments of these skills—but not in soft skills. Lower-caste students do not report higher self-assessed skills in math, programming, or English despite gains in objective math and programming scores, which may reflect comparison with upper-caste peers.

Finally, Table A20 reports effects on self-efficacy, aspirations, leadership, and preferences for jobs in finance and consulting. These results shed light into possible directions for future outcomes. Higher Self efficacy is considered important to vicinage obstacles, set up goals, exert effort to achieve them (Bandura, 1995), occupational aspirations could be important determinants of their future success in the jobs and leadership position in the student clubs play an important role for being part of the elite networks. Upper-caste students in mixed rooms hold fewer leadership positions in student clubs and report lower preferences for finance and consulting careers. Because both outcomes are influenced by access to elite networks, mixed rooms may reduce access to networks that upper-caste students more readily obtain when rooming with same-caste peers. Lower-caste students in mixed rooms do not experience increases in leadership positions or preferences for finance and consulting jobs. However, they report higher expected salaries when assigned

Table 11: Effect on roommate choices at the end of first year

	Mixed Room Y2 (1)	Same roommate in Year 2 (2)	Mixed room in Year 2 (conditional on switching roommates) (3)	Years with First year Roommate (4)
Panel A: Lower Caste Sample				
Mixed Room	0.365***	-0.003	-0.087*	0.041
	(0.036)	(0.035)	(0.047)	(0.038)
DV Mean in same caste rm	0.273	0.471	0.489	1.367
R-squared	0.233	0.233	0.186	0.250
N	1119	1109	550	882
Panel B: Upper Caste Sample				
Mixed Room	0.425***	-0.059	-0.001	-0.122***
	(0.034)	(0.040)	(0.048)	(0.045)
DV Mean in same caste rm	0.230	0.465	0.428	1.487
R-squared	0.254	0.165	0.156	0.260
N	1145	1132	571	841
Panel C: Full Sample				
Mixed Room	0.398***	-0.034	-0.031	-0.034
	(0.026)	(0.030)	(0.033)	(0.031)
DV Mean in same caste rm	0.251	0.468	0.457	1.422
R-squared	0.229	0.175	0.125	0.224
N	2264	2241	1121	1723

Notes: The table presents results for room choices of students. Column 1 reports estimates if students are in a mixed room in the second year. Column 2 reports results for if students are in the same room as their first year roommate in the second year. Column 3 reports results for if students switched from a mixed room to a non-mixed room, conditional on switching the roommate. Column 4 reports results for number of years students lived with their first year roommate (self reported data) All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include number of roommates as control.

Table 12: Effect on relation and interactions with roommate

	Friendly Relation with roommate (1)	Interactions index (2)	Study Interactions (3)	Eating together (4)	Sports (5)	Going out together (6)
Panel A: Lower Caste Sample						
Mixed Room	-0.066	-0.159**	-0.042	-0.192**	-0.209***	-0.150*
	(0.076)	(0.067)	(0.057)	(0.075)	(0.080)	(0.080)
DV Mean in same caste rm	0.061	0.126	0.047	0.125	0.156	0.088
R-squared	0.091	0.116	0.116	0.112	0.134	0.105
N	890	883	938	876	875	873
Panel B: Upper Caste Sample						
Mixed Room	0.022	-0.083	-0.016	-0.074	-0.058	-0.116
	(0.086)	(0.073)	(0.060)	(0.087)	(0.081)	(0.089)
DV Mean in same caste rm	-0.065	-0.111	-0.080	-0.094	-0.140	-0.074
R-squared	0.080	0.141	0.173	0.125	0.158	0.133
N	885	871	915	870	866	867
Panel C: Full Sample						
Mixed Room	-0.015	-0.112**	-0.035	-0.127**	-0.121**	-0.123*
	(0.060)	(0.052)	(0.043)	(0.062)	(0.059)	(0.063)
DV Mean in same caste rm	-0.001	0.010	-0.014	0.017	0.011	0.008
R-squared	0.052	0.094	0.110	0.086	0.112	0.077
N	1775	1754	1853	1746	1741	1740

Notes: The table presents results friendship quality, study interactions, and social interactions. Column 1 reports results for self reported friendship quality (z-score). Column 2 reports results for interactions index, defined as the standardized interactions index of eating, studying, sports and going out. Column 3 reports results for study interactions with roommate (z-score). Column 4 reports results for eating together with roommate (z-score). Column 5 reports results for playing sports togegether with roommate (z-score). Column 6 reports results for going out together with roommate (z-score). All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include number of roommates as control.

Table 13: Effect on Help from roommates: Acads, Intern, Financial, Emotional Support

	Got help index (1)	Got exam prep (2)	Got internship help (3)	Got financial support (4)	Got emotional support (5)
Panel A: Lower Caste Sample					
Mixed Room	-0.028	0.120	0.005	-0.179**	-0.108
	(0.063)	(0.074)	(0.086)	(0.083)	(0.078)
DV Mean in same caste rm	0.085	0.048	0.064	0.155	0.056
R-squared	0.084	0.081	0.079	0.103	0.100
N	955	947	904	919	937
Panel B: Upper Caste Sample					
Mixed Room	-0.102	-0.170*	-0.126	0.043	-0.133
	(0.068)	(0.090)	(0.084)	(0.089)	(0.085)
DV Mean in same caste rm	-0.065	-0.040	-0.052	-0.154	-0.083
R-squared	0.100	0.097	0.106	0.099	0.093
N	982	976	960	961	968
Panel C: Full Sample					
Mixed Room	-0.063	-0.016	-0.066	-0.061	-0.114*
	(0.050)	(0.062)	(0.064)	(0.063)	(0.059)
DV Mean in same caste rm	0.011	0.005	0.006	0.002	-0.012
R-squared	0.065	0.059	0.070	0.067	0.059
N	1937	1923	1864	1880	1905

Notes: The table presents results for type of help received from roommates, including exam preparation, internships, financial help, and emotional support. Column 1 reports estimates for overall help index, an index of different types of help received from first year roommate. Column 2 reports results for exam preparation help. Column 3 reports results for internship help. Column 4 reports results for financial help. Column 5 reports results for emotional support. All the variables are standardized (z-scores) with respect to the full sample. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include number of roommates as control.

Table 14: Effect on Help Provided to roommates: Acads, Intern, Financial, Emotional Support

	Provided help index (1)	Provided exam prep (2)	Provided internship help (3)	Provided financial support (4)	Provided emotional support (5)
Panel A: Lower Caste Sample					
Mixed Room	-0.116**	-0.143**	-0.032	-0.216***	-0.119
	(0.058)	(0.071)	(0.078)	(0.080)	(0.075)
DV Mean in same caste rm	0.116	0.089	0.106	0.177	0.125
R-squared	0.090	0.087	0.084	0.087	0.101
N	847	841	810	825	825
Panel B: Upper Caste Sample					
Mixed Room	0.086	0.141*	0.067	0.168**	0.036
	(0.064)	(0.082)	(0.078)	(0.081)	(0.083)
DV Mean in same caste rm	-0.113	-0.081	-0.102	-0.190	-0.132
R-squared	0.118	0.093	0.117	0.128	0.098
N	844	841	823	831	838
Panel C: Full Sample					
Mixed Room	-0.014	-0.005	0.030	-0.027	-0.055
	(0.044)	(0.055)	(0.057)	(0.057)	(0.056)
DV Mean in same caste rm	0.004	0.006	0.004	-0.003	-0.002
R-squared	0.059	0.050	0.066	0.064	0.067
N	1691	1682	1633	1656	1663

Notes: The table presents results for type of help provided to roommates, including exam preparation, internships, financial help, and emotional support. Column 1 reports estimates for overall help index, an index of different types of help provided to first year roommate. Column 2 reports results for exam preparation help. Column 3 reports results for internship help. Column 4 reports results for financial help. Column 5 reports results for emotional support. All the variables are standardized (z-scores) with respect to the full sample. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include number of roommates as control.

Table 15: Effect on Friendships: Close friends inside and outside college

	Friends Index (1)	Friends inside university (2)	Friends in other elite universities (3)	Share of Other Caste Friends within University (4)	Share of Other Caste Friends outside University (5)
Panel A: Lower Caste Sample					
Mixed Room	-0.075	-0.006	-0.144**	-0.021	0.027
	(0.064)	(0.076)	(0.067)	(0.073)	(0.072)
DV Mean in same caste rm	0.020	-0.026	0.065	0.344	0.403
R-squared	0.096	0.072	0.117	0.094	0.068
N	944	944	944	710	714
Panel B: Upper Caste Sample					
Mixed Room	-0.004	0.022	-0.030	0.181**	0.058
	(0.072)	(0.080)	(0.078)	(0.086)	(0.088)
DV Mean in same caste rm	-0.016	0.035	-0.067	-0.357	-0.369
R-squared	0.098	0.093	0.091	0.074	0.075
N	931	931	931	754	744
Panel C: Full Sample					
Mixed Room	-0.030	0.011	-0.070	0.072	0.030
	(0.047)	(0.054)	(0.051)	(0.057)	(0.055)
DV Mean in same caste rm	0.002	0.004	0.000	-0.014	0.013
R-squared	0.067	0.061	0.067	0.123	0.172
N	1875	1875	1875	1464	1458

Notes: The table presents results for friendship networks, other caste friends inside and outside college. Column 1 reports estimates for friendship index, an index of number of close friends. Column 2 reports results for number of close friends inside university. Column 3 reports results for number of close friends outside university. Column 4 reports results for number of other caste friends within university. Column 5 reports results for number of other caste friends outside university. All the variables are standardized (z-scores) with respect to the full sample. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include number of roommates as control. First three columns have baseline controls for the respective dependent variables and a missing indicator for the baseline controls.

to mixed-caste rooms[45]. Reported expected first-job salaries are substantially higher than students' realized first-job salaries after graduation. These elevated expectations may also reflect comparison with upper-caste reference groups.

8 Conclusion

In this paper, I exploit a natural experiment in social integration policy in India's elite higher educational institutions to estimate how caste identity of roommates in residence halls affects academic performance, mental health, and attitudes toward affirmative action. I use a combination of survey data and administrative data on hostel allocation and entrance exam score ranks to make four main findings. First, lower-caste students have better academic performance and internship outcomes when assigned to mixed-caste rooms, and conversely, upper-caste students have worse academic outcomes but not internships. Second, both students in mixed-caste rooms overall have worse mental health, and two types of channels drive these mental health costs: one identity-linked psychological costs and the other relative deprivation for upper castes, and stress or possible backlash for lower-caste students, linked to differences in entrance exam score gap with roommate. Third, attitudes toward affirmative action change in response to inter-caste exposure, with effects mediated by entrance score differences with roommates. For lower-caste students, positive score gaps with same-caste roommates strengthen support for affirmative action policies. For upper-caste students, support increases when lower-caste roommates have relatively similar entrance scores, but declines when lower-caste roommates score comparatively lower than the student. Fourth, intensity of exposure also matters and exposure to to more diverse lower-caste students (broad contact) also helps in shaping attitudes in support of affirmative action policies, residing with co-ethnic peers in mixed rooms acts as a social buffer to manager stress (for both upper and lower caste students), and improve academic performance for lower caste students. Fifth, I also observe social frictions in diverse rooms are reflected in behavioral outcomes of students like interactions, help in studies, managing stress and financial stress and these can have consequences for friendships, aspirations, network formation. Thus, overall implementing integration policies in competitive and elite settings can have complex implications

One important policy implication of this study is that policies like affirmative action, which aim to increase equality of opportunity in societies with high socio-economic inequalities, have been shown to benefit the disadvantaged, but these benefits could be

[45]Salaries are top-coded at the 95th percentile to reduce the influence of outliers.

larger if diversity were better managed—particularly when backlash against affirmative action policies exists.

If these frictions are not well managed, they could threaten the repeal of affirmative action policies despite persistent inequality of opportunity, similar to what happened to race based affirmative case in the united states *Students for Fair Admissions v. Harvard*. In countries like India, it is important to understand the social engineering required beyond increasing representation to reduce friction between groups when they actually start interacting with each other and possibly benefiting from diversity.

One limitation of this work is that, due to surveying students while still enrolled at the institute, it does not study the long-term economic implications of inter-caste contact, such as job market outcomes—particularly important given that short-term results show both benefits and costs to both exposed groups. Second, effects of similar integration could be different in other non-elite schools where the competitive environment may be less intense, or where initial social distance or performance gaps could be smaller and have fewer negative consequences. However, upper-caste students in such settings may experience stronger backlash against affirmative action policies if they perceive greater relative deprivation through these policies. Third, recent work has shown that caste discrimination remains important in hiring from elite institutions in the Indian context Shukla, n.d. What are the long-term implications of inter-caste exposure for economic attitudes like hiring decisions of upper- and lower-caste students? This is particularly important when most graduates from these elite institutions occupy influential positions in the economy as managers and executives of multinational firms, top bureaucrats, and start-up founders. Long-run effects on social outcomes like marriage, residential choices, or political beliefs are also unknown. Recent work shows evidence on this in developed country contexts where higher childhood contact increased hiring of immigrants Wren-Lewis et al., 2025 and affected residential choices Merlino et al., 2024. Developing country contexts and exposure at college age could have different consequences. Fourth, I focus only on integration of diverse groups through peer groups as roommates. There are other important peer contexts where students interact that could mitigate some of these social frictions, as well as vertical networks—such as mentorship from seniors or juniors—where significant cross-cohort interaction through hostel activities and student club activities can also shape human capital, job preferences, and socio-political outcomes. Fifth, this study does not yet examine alternative allocation mechanisms that could potentially reduce misallocation of human capital gains and foster convergence in political attitudes.

Following the current student sample over time will answer questions on long-term

outcomes, as a significant proportion are on LinkedIn. Conducting similar studies in other residential institutions that employ random allocation of roommates and mentors, or running experiments in partnership with institutions, can help address these questions.

References

Akhtari, M., Bau, N., & Laliberté, J.-W. (2024). Affirmative action and precollege human capital. *American Economic Journal: Applied Economics*, *16*(1), 1–32. https://doi.org/10.1257/app.20210640

Alan, S., Baysan, C., Gumren, M., & Kubilay, E. (2021). Building social cohesion in ethnically mixed schools: An intervention on perspective taking. *The Quarterly Journal of Economics*, *136*(4), 2147–2194. https://doi.org/10.1093/qje/qjab009

Alesina, A., & La Ferrara, E. (2005). Ethnic diversity and economic performance. *Journal of Economic Literature*, *43*(3), 762–800. https://doi.org/10.1257/002205105774431243

Allport, G. W. (1954). *The nature of prejudice*. Addison-Wesley.

Anderson, S. (2011). Caste as an impediment to trade. *American Economic Journal: Applied Economics*, *3*(1), 239–263. https://doi.org/10.1257/app.3.1.239

Arcidiacono, P., & Lovenheim, M. (2016). Affirmative action and the quality–fit trade-off. *Journal of Economic Literature*, *54*(1), 3–51. https://doi.org/10.1257/jel.54.1.3

Bagde, S., Epple, D., & Taylor, L. (2016). Does affirmative action work? caste, gender, college quality, and academic success in india. *American Economic Review*, *106*(6), 1495–1521.

Bandura, A. (Ed.). (1995). *Self-efficacy in changing societies*. Cambridge University Press. https://doi.org/10.1017/CBO9780511527692

Bazzi, S., Gaduh, A., Rothenberg, A. D., & Wong, M. (2019). Unity in diversity? how intergroup contact can foster nation building. *American Economic Review*, *109*(11), 3978–4025. https://doi.org/10.1257/aer.20170947

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., & Topalova, P. (2009). Powerful women: Does exposure reduce bias? *The Quarterly Journal of Economics*, *124*(4), 1497–1540.

Bertrand, M., Hanna, R., & Mullainathan, S. (2010). Affirmative action in education: Evidence from engineering college admissions in india. *Journal of Public Economics*, *94*(1), 16–29.

Bleemer, Z. (2022). Affirmative action, mismatch, and economic mobility after california's proposition 209. *The Quarterly Journal of Economics*, *137*(1), 115–160. https://doi.org/10.1093/qje/qjab027

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.

Breda, T., Grenet, J., Monnet, M., & Effenterre, C. V. (2023). How effective are female role models in steering girls towards stem? evidence from french high schools [Published: 23 February 2023]. *The Economic Journal*, *133*(653), 1773–1809. https://doi.org/10.1093/ej/uead019

Carrell, S. E., Hoekstra, M., & West, J. E. (2019). The impact of college diversity on behavior toward minorities. *American Economic Journal: Economic Policy*, *11*(4), 159–182.

Cassan, G. (2019). Affirmative action, education and gender: Evidence from india. *Journal of Development Economics*, *136*, 51–70.

Chattopadhyay, R., & Duflo, E. (2004). Women as Policy Makers: Evidence from a Randomized Policy Experiment in India. *Econometrica*, *72*(5), 1409–1443. https://ideas.repec.org/a/ecm/emetrp/v72y2004i5p1409-1443.html

Chetty, R., Friedman, J. N., Saez, E., Turner, N., & Yagan, D. (2020). Income segregation and intergenerational mobility across colleges in the united states. *The Quarterly Journal of Economics*, *135*(3), 1567–1633. https://doi.org/10.1093/qje/qjaa005

Corno, L., La Ferrara, E., & Burns, J. (2022). Interaction, stereotypes, and performance: Evidence from south africa. *American Economic Review*, *112*(12), 3848–3875.

Dahl, G. B., Kotsadam, A., & Rooth, D.-O. (2021). Does integration change gender attitudes? the effect of randomly assigning women to traditionally male teams. *The Quarterly Journal of Economics*, *136*(2), 987–1030. https://doi.org/10.1093/qje/qjaa047

Deshpande, A. (2013). Social justice through affirmative action in india: An assessment. In J. Wicks-Lim & R. Pollin (Eds.), *Capitalism on trial*. Edward Elgar Publishing.

Duflo, E., Dupas, P., & Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in kenya. *American Economic Review*, *101*(5), 1739–1774. https://doi.org/10.1257/aer.101.5.1739

Enos, R. D. (2014). Causal effect of intergroup contact on exclusionary attitudes. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(10), 3699–3704. https://doi.org/10.1073/pnas.1317670111

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140. https://doi.org/10.1177/001872675400700202

Finseraas, H., & Kotsadam, A. (2017). Does personal contact with ethnic minorities affect anti-immigrant sentiments? evidence from a field experiment. *European Journal of Political Research*. https://doi.org/10.1111/1475-6765.12199

Ghosh, A. (2025). Religious divisions and production technology: Experimental evidence from india. *Journal of Political Economy*, *133*(10), 3249–3304. https://doi.org/10.1086/736212

Glover, D., Pallais, A., & Pariente, W. (2017). Discrimination as a self-fulfilling prophecy: Evidence from french grocery stores. *The Quarterly Journal of Economics*, *132*(3), 1219–1260.

Gulzar, S., Haas, N., & Pasquale, B. (2020). Does political affirmative action work, and for whom? theory and evidence on india's scheduled areas. *American Political Science Review*, *114*(4), 1230–1246.

Hjort, J. (2014). Ethnic divisions and production in firms. *The Quarterly Journal of Economics*, *129*(4), 1899–1946. https://doi.org/10.1093/qje/qju028

Hoff, K., & Pandey, P. (2006). Discrimination, social identity, and durable inequalities. *American Economic Review*, *96*(2), 206–211.

Krill, Z., Geva, A., & Aloni, T. (2019). The effect of the field of study on the higher education wage premium—evidence from israel. *Labour*, *33*(3), 388–423. https://doi.org/10.1111/labr.12150

Liu, J., Conrad, C., & Blazar, D. (2024). *Computer science for all? the impact of high school computer science courses on college majors and earnings* (Working Paper No. 16758). Institute of Labor Economics (IZA). Bonn, Germany.

Lowe, M. (2021). Types of contact: A field experiment on collaborative and adversarial caste integration. *American Economic Review*, *111*(6), 1807–1844.

Lowe, M. (2025). Has intergroup contact delivered? *Annual Review of Economics*, *17*, 321–344. https://doi.org/10.1146/annurev-economics-081324-091109

Merlino, L. P., Steinhard, M. F., & Wren-Lewis, L. (2024). The long-run impact of childhood interracial contact on residential segregation. *Journal of Public Economics*, *239*, 105242. https://doi.org/10.1016/j.jpubeco.2024.105242

Michelman, V., Price, J., & Zimmerman, S. D. (2022). Old boys' clubs and upward mobility among the educational elite. *The Quarterly Journal of Economics*, *137*(2), 845–909. https://doi.org/10.1093/qje/qjab046

Mousa, S. (2020). Building social cohesion between christians and muslims through soccer in post-isis iraq. *Science*, *369*(6505), 866–870. https://doi.org/10.1126/science.abb3153

Munshi, K. (2019). Caste and the indian economy. *Journal of Economic Literature*, *57*(4), 781–834. https://doi.org/10.1257/jel.20171307

Okunogbe, O. (2024). Does exposure to other ethnic regions promote national integration? evidence from nigeria. *American Economic Journal: Applied Economics*, *16*(1), 157–192. https://doi.org/10.1257/app.20210669

Paluck, E. L., Green, S. A., & Green, D. P. (2019). The contact hypothesis re-evaluated. *Behavioural Public Policy*, *3*(2), 129–158. https://doi.org/10.1017/bpp.2018.25

Pande, R. (2003). Can mandated political representation increase policy influence for disadvantaged minorities? theory and evidence from india. *American Economic Review*, *93*(4), 1132–1151.

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

Rao, G. (2019). Familiarity does not breed contempt: Generosity, discrimination, and diversity in delhi schools. *American Economic Review*, *109*(3), 774–809.

Romppel, M., Herrmann-Lingen, C., Wachter, R., Edelmann, F., Düngen, H.-D., Pieske, B., & Grande, G. (2013). A short form of the general self-efficacy scale (gse-6): Development, psychometric properties and validity in an intercultural non-clinical sample and a sample of patients at risk for heart failure. *Psychosomatic Medicine*, *10*, Doc01. https://doi.org/10.3205/psm000091

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. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (pp. 249–277, Vol. 3). Elsevier. https://doi.org/10.1016/B978-0-444-53429-3.00004-1

Scacco, A., & Warren, S. S. (2018). Can social contact reduce prejudice and discrimination? evidence from a field experiment in nigeria. *American Political Science Review*, *112*(3), 654–677. https://doi.org/10.1017/S0003055418000151

Schäfer, S. J., Kauff, M., Prati, F., Kros, M., Lang, T., & Christ, O. (2021). Does negative contact undermine attempts to improve intergroup relations? deepening the understanding of negative contact and its consequences for intergroup contact research and interventions. *Journal of Social Issues*, *77*(1), 192–219. https://doi.org/10.1111/josi.12422

Sekhri, S. (2020). Prestige matters: Wage premium and value addition in elite colleges. *American Economic Journal: Applied Economics*, *12*(3), 207–225.

Shukla, S. (n.d.). *Making the elite: Class discrimination at multinationals* [Unpublished manuscript].

Tabellini, M. (2020). Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration. *The Review of Economic Studies*, *87*(1), 454–486.

Webb, D. (2025). Silence to solidarity: How communication about a minority affects discrimination [Forthcoming]. *Journal of Political Economy*.

Wren-Lewis, L., Steinhardt, M., Merlino, L., & Rosenbaum, P. (2025). *The impact of childhood interethnic contact on managers' hiring decisions* [Working paper].

Zimmerman, S. D. (2019). Elite colleges and upward mobility to top jobs and top incomes. *American Economic Review*, *109*(1), 1–47. https://doi.org/10.1257/aer.20170522

9 Online Appendix

Table A1: Summary statistics: Mental Health and Attitudes on Affirmative Action at Baseline

	Lower Caste			Upper Caste			Lower Caste - Upper Caste
	N	Mean	Sd	N	Mean	Sd	Difference
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. PHQ4: Anxiety and Depression Symptoms at Baseline							
PHQ4: Nervous or anxious	434	-0.024	0.925	453	-0.002	1.005	-0.022
PHQ4: Can't stop worrying	429	0.058	1.021	450	-0.012	1.009	0.070
PHQ4: Feeling down	431	0.022	0.947	453	0.036	1.019	-0.014
PHQ4: Little interest in activities	423	0.040	0.993	448	-0.080	1.018	0.120
B. Relative unhappiness components at Baseline							
Rel JEE: More happy now (rev)	422	-0.067	0.974	448	0.026	1.028	-0.093
Rel JEE: Find little joy	424	0.102	0.946	447	-0.146	1.062	0.248
Rel JEE: Less control in life	415	-0.013	0.943	443	-0.041	1.035	0.028
Rel JEE: More stressed	423	-0.005	0.932	447	-0.015	1.067	0.010
Rel JEE: Better support system (rev)	420	-0.151	0.999	446	0.186	1.107	-0.337
C. Worry Index Components at Baseline							
Academics	490	-0.058	1.033	502	0.076	1.027	-0.134
Getting an Internship	490	-0.001	1.001	502	0.036	1.023	-0.037
Cultural Adaptation	490	0.383	1.217	501	-0.183	0.864	0.566
Staying Away from Home	490	-0.044	0.957	502	0.029	0.998	-0.073
Discrimination	490	0.230	1.135	501	-0.085	0.860	0.315
Time Management	490	-0.124	1.005	502	0.057	1.010	-0.181
D. Social Justice and Diversity benefits of AA at Baseline							
Promote social justice	433	0.445	0.858	446	-0.414	0.940	0.859
Foster diversity	423	0.302	0.928	445	-0.218	1.112	0.520
Benefit to non-reserved groups	418	0.367	0.809	441	-0.288	1.051	0.655
E. Meritocratic concerns of AA at Baseline							
Hurt reserved groups	429	-0.308	0.973	444	0.250	1.013	-0.558
Unfair to merit	432	-0.509	0.967	447	0.430	0.854	-0.939
Reduce incentive to work	434	-0.380	1.067	451	0.388	0.883	-0.768
Compromise academic quality	429	-0.533	0.972	451	0.318	0.904	-0.851

Notes: The table presents summary statistics for socioeconomic indicators. The sample includes students from entry year of 2024 who had one or two roommates. I report the sample size, mean, and standard deviation separately for upper-caste and lower-caste students. PHQ4 is a standard measure of anxiety and depression symptoms, with higher values indicating worse mental health. Relative Unhappiness Index measures students' happiness relative to their pre-college happiness levels, with higher values indicating worse happiness. Worry Index measures the extent to which students are worried about college life, with higher values indicating more worry. Social Justice and Diversity benefits of Affirmative Action index measures students' attitudes toward the social justice aspects of affirmative action policies, with higher values indicating more support. Meritocratic concerns of Affirmative Action index measures students' attitudes toward whether affirmative action policies undermine meritocracy, with higher values indicating more concern/less supported. Students responded to a series of statements using a five-point Likert scale ranging from strongly disagree to strongly agree. and I standardized the responses to create z-scores for each variable with respect to full sample.

Table A2: Balance table on Attrition, Socio Economic Indicators and College Preparation

	Lower Caste				Upper Caste			
	N	Non-mixed mean	Diff	P-Val	N	Non-mixed mean	Diff	P-Val
Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Survey attrition indicators								
Survey completed	1197	0.872	-0.006	0.804	1253	0.800	-0.005	0.861
Response	1197	0.891	-0.028	0.220	1253	0.838	-0.003	0.894
B. Socio Economic Indicators								
Mother's education (Bachelors or higher)	947	0.434	-0.015	0.677	998	0.669	-0.031	0.361
Father's education (Bachelors or higher)	942	0.528	-0.020	0.567	992	0.639	0.018	0.581
Assets Index	857	-0.171	-0.009	0.856	908	0.166	-0.016	0.727
Taken a loan	890	0.194	0.034	0.241	888	0.148	0.006	0.829
Family income (9L or above)	955	0.425	-0.039	0.286	934	0.699	-0.079	0.047
Speaks English with family	941	0.246	0.048	0.133	926	0.494	-0.027	0.462
Speaks English with friends	941	0.558	0.027	0.451	926	0.705	-0.003	0.934
Knows someone in engg	907	0.490	-0.005	0.888	921	0.682	-0.060	0.088
Knows someone with IIT degree	907	0.168	0.013	0.642	921	0.259	-0.005	0.885
C. JEE Preparation and Academic Background								
Class 12 score (%)	1000	90.175	0.271	0.575	1035	93.089	0.985	0.134
Class 10 score (%)	1001	92.942	-0.374	0.568	1032	94.995	0.424	0.539
Entrance Exam score (out of 360)	1163	139.368	1.192	0.577	1220	204.125	-1.450	0.298
Years: JEE preparation	1036	2.251	-0.020	0.841	1060	2.520	-0.130	0.215
Average daily study hours JEE	1034	8.588	0.467	0.080	1062	8.225	0.167	0.580
Cost of JEE preparation	950	5.058	-0.253	0.202	956	5.051	-0.071	0.729
Took more than 1 attempt	1003	0.317	-0.063	0.066	1044	0.147	-0.004	0.862
Doesn't Know programming	976	0.671	-0.028	0.426	1005	0.455	0.041	0.290
Joint Test p-value				0.540				0.369

Notes: The table presents balance on socioeconomic indicators, JEE preparation, and pre-college academic outcomes. Panel A reports balance on attrition indicators, including whether students responded to the survey and whether they completed the survey. Panels B and C report balance on socioeconomic characteristics and pre-college academic outcomes, respectively. The sample includes students from entry years 2023 and 2024 who had one or two roommates. Columns 2 and 6 report means for same-caste rooms for lower-caste and upper-caste students, respectively. Columns 3–4 and 7–8 report differences in means and corresponding p-values for lower-caste and upper-caste students. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. The last row of each panel reports the p-value from the joint F-test of orthogonality.

Table A3: Balance table on Pre-college Networks

	Lower Caste				Upper Caste			
	N	Non-mixed mean	Diff	P-Val	N	Non-mixed mean	Diff	P-Val
Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Pre-college Networks								
Close friends count	484	9.173	0.595	0.568	488	10.929	-0.057	0.957
Close friends wrote jee	484	9.624	0.103	0.929	488	10.489	-0.550	0.604
Close friends in IITD	484	2.549	0.407	0.455	488	2.979	-0.104	0.865
Close friends in NITs	484	2.692	0.784	0.189	488	1.851	0.819	0.188
Friends wrote JEE but didn't get in	484	5.684	1.080	0.237	488	4.574	0.911	0.290
Prog: Learned from friends	464	0.031	-0.012	0.451	477	0.036	0.018	0.419
Diff caste close friends bf IIT	418	2.956	0.049	0.668	425	1.779	0.260	0.080
B. Consumption Expenditure at Baseline								
Food expenses	503	1500.000	-63.167	0.676	510	1763.514	-34.175	0.856
Movies and partying expenses	503	1021.429	-82.599	0.461	510	1283.784	-36.779	0.837
Education expenses	503	2567.857	-436.217	0.137	510	2354.730	-262.736	0.311
Other miscellaneous expenses	503	2128.571	-353.494	0.203	510	2111.486	-321.836	0.191
Joint Test p-value				0.866				0.239

Notes: The table presents balance on pre-college networks and self-reported skills at baseline. Panels A report balance on friendship networks pre college and panel B represents balance on consumption expenditure at baseline. The sample includes students only from entry year 2024 who had one or two roommates. Columns 2 and 6 report means for same-caste rooms for lower-caste and upper-caste students, respectively. Columns 3–4 and 7–8 report differences in means and corresponding p-values for lower-caste and upper-caste students. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. The last row of each panel reports the p-value from the joint F-test of orthogonality.

Table A4: Early exposure results on Mental Health, Attitudes towards Affirmative Action and Skills

	Lower Caste				Upper Caste			
	N	Non-mixed mean	Diff	P-Val	N	Non-mixed mean	Diff	P-Val
Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mental Health, AA support & Skill indices at Baseline								
PHQ4: Anxiety and Depression	438	0.006	-0.009	0.916	455	-0.013	-0.044	0.639
Relative wellbeing Index	426	-0.000	-0.057	0.366	450	0.009	-0.046	0.540
Worry Index	490	0.015	0.106	0.147	502	-0.006	-0.054	0.408
AA Merit Index (rev)	441	0.392	0.023	0.782	454	-0.354	-0.001	0.994
AA Social Justice	439	0.276	0.084	0.252	454	-0.266	-0.118	0.278
Self rating of Skills	484	-0.303	0.112	0.151	490	0.251	-0.110	0.086
Joint Test p-value				0.473				0.275

Notes: The table presents balance on mental health, attitudes towards Affirmative Action and self-reported skills at baseline. It reports early exposure effects on PHQ4 index, relative happiness index, worry index, academic merit index, social justice index, and skills index. The sample includes students only from entry year 2024 who had one or two roommates. Columns 2 and 6 report means for same-caste rooms for lower-caste and upper-caste students, respectively. Columns 3–4 and 7–8 report differences in means and corresponding p-values for lower-caste and upper-caste students. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. The last row of each panel reports the p-value from the joint F-test of orthogonality.

Table A5: Early exposure results on Mental Health and Attitudes towards Affirmative Action and Skills by Components

	Lower Caste				Upper Caste			
	N	Non-mixed mean	Diff	P-Val	N	Non-mixed mean	Diff	P-Val
Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. PHQ4: Anxiety and Depression Symptoms at Baseline								
PHQ4: Nervous or anxious	434	-0.050	-0.004	0.974	453	0.017	-0.062	0.623
PHQ4: Can't stop worrying	429	0.032	-0.026	0.821	450	-0.047	0.019	0.873
PHQ4: Feeling down	431	0.017	-0.085	0.458	453	-0.026	0.025	0.828
PHQ4: Little interest in activities	423	0.018	0.003	0.979	448	0.003	-0.135	0.210
B. Relative unhappiness components at Baseline								
Rel JEE: More happy now (rev)	422	0.021	-0.132	0.232	448	-0.004	0.038	0.759
Rel JEE: Find little joy	424	0.115	0.047	0.653	447	-0.096	-0.123	0.279
Rel JEE: Less control in life	415	0.087	-0.213	0.084	443	-0.079	-0.018	0.880
Rel JEE: More stressed	423	-0.052	-0.034	0.747	447	0.076	-0.159	0.166
Rel JEE: Better support system (rev)	420	-0.153	0.038	0.735	446	0.148	0.059	0.604
C. Worry Index Components at Baseline								
Academics	490	-0.150	0.155	0.133	502	0.113	-0.178	0.146
Getting an Internship	490	-0.048	0.110	0.335	502	0.042	-0.093	0.394
Cultural Adaptation	490	0.309	0.123	0.342	501	-0.269	0.147	0.107
Staying Away from Home	490	0.034	-0.048	0.666	502	0.004	0.009	0.925
Discrimination	490	0.088	0.261	0.047	501	-0.072	-0.065	0.462
Time Management	490	-0.142	0.036	0.727	502	0.122	-0.123	0.215
D. Attitudes towards AA at Baseline								
Promote social justice	433	0.397	0.030	0.772	446	-0.383	-0.117	0.346
Foster diversity	423	0.180	0.154	0.131	445	-0.170	-0.178	0.219
Benefit to non-reserved groups	418	0.268	0.039	0.705	441	-0.259	-0.047	0.715
Hurt reserved groups	429	-0.236	-0.071	0.531	444	0.233	0.054	0.626
Unfair to merit	432	-0.462	0.033	0.762	447	0.446	0.022	0.824
Reduce incentive to work	434	-0.396	0.037	0.752	451	0.394	0.018	0.857
Compromise academic quality	429	-0.444	-0.093	0.430	451	0.392	-0.128	0.208
E. Self-reported Skills at Baseline and Friendship with Roommate								
Rating of Math & Prog	484	-0.334	0.120	0.142	490	0.286	-0.117	0.117
Rating: English	484	-0.400	0.127	0.305	490	0.336	-0.064	0.493
Rating of Soft Skills	484	-0.295	0.111	0.194	490	0.242	-0.108	0.131
Expected Friendship with roommate	452	2.242	0.004	0.964	473	2.372	-0.197	0.039
Joint Test p-value				0.825				0.031

Notes: The table presents balance on pre-college networks and self-reported skills at baseline. Panel A reports early exposure effects on PHQ4 index components. Panel B, C report effects on components of relative happiness and worry indices. Panel E report effects on attitudes towards Affirmative Action: Social Justice and Meritocratic concerns indices. Panel F reports effects on self-reported skills both soft skills and hard skills and expected frienship with roommate. The sample includes students only from entry year 2024 who had one or two roommates. Columns 2 and 6 report means for same-caste rooms for lower-caste and upper-caste students, respectively. Columns 3–4 and 7–8 report differences in means and corresponding p-values for lower-caste and upper-caste students. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. The last row of each panel reports the p-value from the joint F-test of orthogonality.

Table A6: Effect on PHQ4 index and its components

	PHQ4 Index (1)	Nervous or anxious (2)	Can't stop worrying (3)	Feeling down (4)	Little interest in activities (5)
Panel A: Lower Caste Sample					
Mixed Room	0.080	0.153**	0.076	0.088	0.033
	(0.063)	(0.074)	(0.079)	(0.078)	(0.080)
DV Mean in same caste rm	0.010	-0.048	-0.025	0.054	0.054
R-squared	0.121	0.118	0.134	0.101	0.110
N	934	920	915	916	909
Panel B: Upper Caste Sample					
Mixed Room	0.090	0.096	0.015	0.080	0.143*
	(0.068)	(0.079)	(0.081)	(0.077)	(0.083)
DV Mean in same caste rm	-0.016	0.037	0.009	-0.065	-0.061
R-squared	0.144	0.128	0.113	0.136	0.099
N	971	967	967	966	967
Panel C: Full Sample					
Mixed Room	0.096**	0.131**	0.047	0.097*	0.110*
	(0.047)	(0.056)	(0.059)	(0.055)	(0.058)
DV Mean in same caste rm	-0.003	-0.006	-0.008	-0.005	-0.003
R-squared	0.098	0.083	0.090	0.085	0.069
N	1905	1887	1882	1882	1876

Notes: The table presents results for PHQ-4 index and its components, including anxiety and depression symptoms. Column 1 reports estimates for a composite PHQ-4 Index constructed using standardized scores of the individual PHQ-4 components. Column 2 reports results for feeling nervous, anxious, or on edge. Column 3 reports results for being unable to stop or control worrying. Column 4 reports results for feeling down, depressed, or hopeless. Column 5 reports results for little interest or pleasure in doing things. In all the columns, higher values of these indices indicate worse mental health outcomes. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective dependent variables and a missing indicator for the baseline controls and number of roommates.

Table A7: Effect on Relative well-being index and its components

	Relative well being Index (1)	More happy now (rev) (2)	Find little joy (3)	Less control in life (4)	More stressed (5)	Better support system (rev) (6)
Panel A: Lower Caste Sample						
Mixed Room	0.014	-0.003	-0.006	0.027	-0.012	0.075
	(0.041)	(0.072)	(0.076)	(0.077)	(0.074)	(0.074)
DV Mean in same caste rm	0.036	0.061	0.079	-0.003	0.050	0.005
R-squared	0.116	0.109	0.110	0.079	0.106	0.107
N	850	838	837	830	841	832
Panel B: Upper Caste Sample						
Mixed Room	0.016	0.061	-0.039	-0.020	0.082	-0.026
	(0.051)	(0.081)	(0.076)	(0.080)	(0.086)	(0.080)
DV Mean in same caste rm	-0.055	-0.080	-0.066	-0.010	-0.075	-0.018
R-squared	0.091	0.087	0.086	0.065	0.108	0.093
N	846	834	833	838	835	831
Panel C: Full Sample						
Mixed Room	0.023	0.047	-0.017	0.012	0.027	0.033
	(0.032)	(0.054)	(0.053)	(0.054)	(0.056)	(0.054)
DV Mean in same caste rm	-0.009	-0.008	0.007	-0.006	-0.011	-0.006
R-squared	0.074	0.061	0.062	0.039	0.075	0.053
N	1696	1672	1670	1668	1676	1663

Notes: The table presents results for relative happiness and its components, including anxiety and depression symptoms. Column 1 reports estimates for a composite of relative happiness compared to pre-college period. Column 2 reports results for feeling more unhappy now. Column 3 reports results for finding little joy in doing things. Column 4 reports results for feeling less in control. Column 5 reports results for feeling more stressed. Column 6 reports results for feeling worse about social support. In all the columns, higher values of these indices indicate worse mental health outcomes. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective dependent variables and a missing indicator for the baseline controls and number of roommates.

Table A8: Effect on Worry Index and its components

	Worry Index (1)	Academics (2)	Getting an Internship (3)	Cultural Adaptation (4)	Staying Away from Home (5)	Discrimination (6)	Time Management (7)
Panel A: Lower Caste Sample							
Mixed Room	0.036	0.063	0.026	0.058	-0.023	0.106	0.012
	(0.045)	(0.071)	(0.071)	(0.074)	(0.076)	(0.074)	(0.073)
DV Mean in same caste rm	0.046	0.038	0.003	0.130	0.077	0.016	0.011
R-squared	0.142	0.146	0.124	0.111	0.125	0.070	0.110
N	892	882	873	881	880	871	884
Panel B: Upper Caste Sample							
Mixed Room	0.015	0.026	-0.024	-0.050	0.100	-0.062	0.119
	(0.046)	(0.080)	(0.070)	(0.072)	(0.072)	(0.073)	(0.076)
DV Mean in same caste rm	-0.052	-0.043	-0.014	-0.142	-0.079	-0.026	-0.023
R-squared	0.112	0.102	0.128	0.098	0.114	0.076	0.101
N	896	887	883	890	885	884	889
Panel C: Full Sample							
Mixed Room	0.027	0.033	-0.002	0.016	0.050	0.020	0.054
	(0.033)	(0.053)	(0.050)	(0.052)	(0.052)	(0.052)	(0.053)
DV Mean in same caste rm	-0.003	-0.003	-0.005	-0.005	-0.001	-0.005	-0.006
R-squared	0.095	0.093	0.094	0.088	0.095	0.045	0.062
N	1788	1769	1756	1771	1765	1755	1773

Notes: The table presents results for worry index and its components, including anxiety and depression symptoms. Column 1 reports estimates for a composite of worry index about college aspects. Column 2 reports results for how worried about academic performance. Column 3 reports results for how worried about internship opportunities. Column 4 reports results for how worried about language and cultural adjustment. Column 5 reports results for how worried about being far away from home. Column 6 reports results for how worried about discrimination. In all the columns, higher values of these indices indicate worse mental health outcomes. All the components are measure on scale and 1-5 and standardized with respect to same caste group sample. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective dependent variables and a missing indicator for the baseline controls and number of roommates.

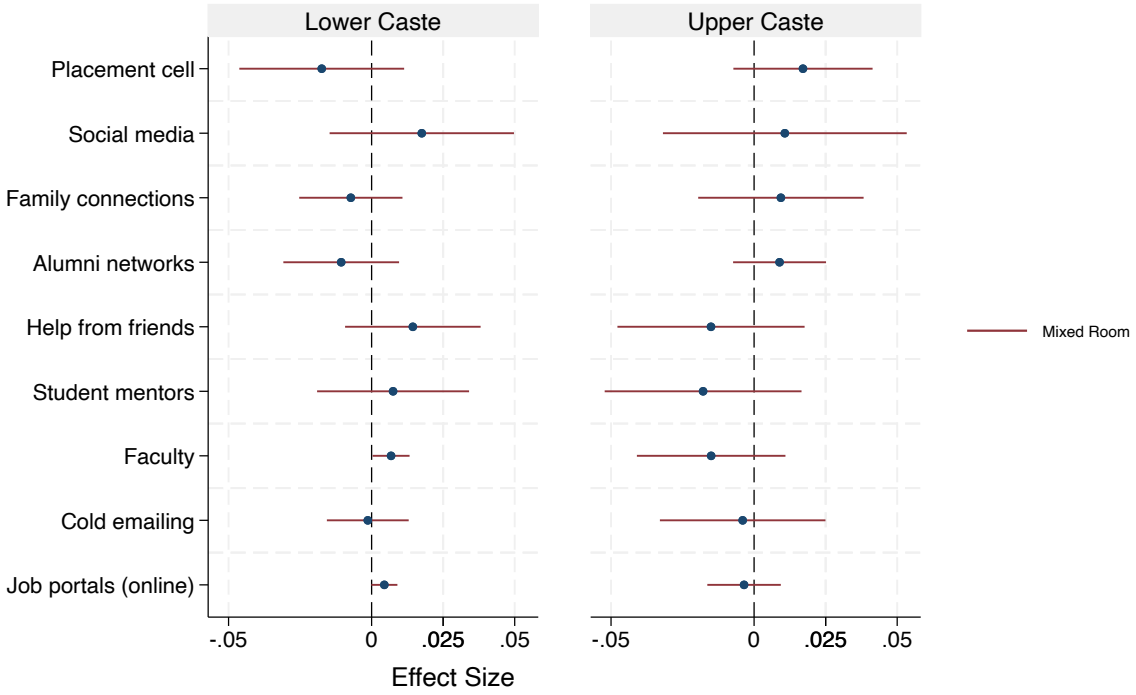


Figure A1: Effect on channels of first year internships or academic projects

Table A9: Effect on Support Affirmative action: Benefits of AA policies to reserved groups

	Benefits of AA to reserved groups (1)	Promote social justice (2)	Foster diversity (3)	Hurt reserved groups (rev) (4)	Reduce incentive to work (rev) (5)
Panel A: Lower Caste Sample					
Mixed Room	-0.022	-0.066	-0.067	-0.080	0.087
	(0.044)	(0.080)	(0.079)	(0.083)	(0.082)
DV Mean in same caste rm	0.239	2.452	2.489	1.662	1.806
R-squared	0.110	0.093	0.071	0.084	0.114
N	883	852	841	865	858
Panel B: Upper Caste Sample					
Mixed Room	0.115**	0.150	0.118	0.110	0.048
	(0.057)	(0.097)	(0.095)	(0.084)	(0.086)
DV Mean in same caste rm	-0.251	1.788	2.168	1.162	1.209
R-squared	0.172	0.127	0.118	0.126	0.125
N	907	891	885	886	890
Panel C: Full Sample					
Mixed Room	0.046	0.037	0.024	0.029	0.056
	(0.037)	(0.064)	(0.063)	(0.061)	(0.060)
DV Mean in same caste rm	-0.002	2.123	2.329	1.415	1.510
R-squared	0.205	0.141	0.082	0.092	0.137
N	1790	1743	1726	1751	1748

Notes: The table presents results for attitudes towards Affirmative Action specific to the reserved groups. Column 1 reports estimates for a index of attitudes that are related to benefits and costs (reverse coded) of AA policies for groups that receive reservation. It is composed of components in columns 2-5. Column 2 reports results for attitudes on if AA is justified to balance historical disadvantage. Column 3 reports results for attitudes on if AA is justified to foster diversity. Column 4 reports results for attitudes on if AA hurt reserved students by placing them in challenging academic environments. Column 5 reports results for if AA policies reduce incentives for reserved students to work hard. Column 4 and 5 are reverse coded and higher estimated value indicate more support on benefits to the reserved groups. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective dependent variables and a missing indicator for the baseline controls and number of roommates.

Table A10: Effect on Support Affirmative action: Costs of AA Policies to non reserved groups

	Costs of AA to non reserved groups (1)	Unfair to merit (2)	Compromise academic quality (3)	Benefit to non-reserved groups (rev) (4)
Panel A: Lower Caste Sample				
Mixed Room	-0.110**	-0.114	-0.047	-0.008
	(0.055)	(0.072)	(0.075)	(0.072)
DV Mean in same caste rm	-0.237	-0.280	-0.304	-0.178
R-squared	0.131	0.146	0.108	0.084
N	874	771	853	753
Panel B: Upper Caste Sample				
Mixed Room	-0.048	0.016	-0.142*	0.019
	(0.058)	(0.078)	(0.077)	(0.083)
DV Mean in same caste rm	0.265	0.295	0.311	0.178
R-squared	0.184	0.137	0.128	0.121
N	901	771	890	757
Panel C: Full Sample				
Mixed Room	-0.097**	-0.046	-0.122**	-0.004
	(0.045)	(0.053)	(0.057)	(0.055)
DV Mean in same caste rm	0.012	0.005	0.003	0.001
R-squared	0.196	0.183	0.132	0.100
N	1775	1542	1743	1510

Notes: The table presents results for attitudes towards Affirmative Action specific to the non reserved groups. Column 1 reports estimates for a index of attitudes that are related to benfits (reverse coded) and cost of AA policies for groups that don't benefit from AA. It is composed of components in columns 2-4. Column 2 reports results for if AA policies are unfair compared to meritocratic policies. Column 3 reports results for attitudes on if AA policies compromise academic standards of the university. Column 4 reports results for attitudes on if AA is justified to foster diversity. Column 4 is reverse coded and higher estimated value indicate more belief on higher costs of AA to the non reserved groups. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective dependent variables and a missing indicator for the baseline controls and number of roommates.

Table A11: Effect on GPA, Internships, Maths and Programming performance: Three Sharing Rooms

	Performance Index (1)	GPA (2)	Maths & Programming (3)	Summer Project or Internship (4)
Panel A: Lower Caste Sample				
Partial Mixed Room	0.218***	0.086	0.124	0.145***
	(0.064)	(0.078)	(0.081)	(0.033)
Full Mixed Room	0.151**	0.058	0.057	0.122***
	(0.073)	(0.090)	(0.095)	(0.038)
P-val (Partial = Full)	0.279	0.745	0.398	0.550
DV Mean in same caste rm	-0.439	-0.452	-0.493	0.045
R-squared	0.341	0.430	0.407	0.132
N	656	584	651	644
Panel B: Upper Caste Sample				
Partial Mixed Room	-0.070	-0.073	-0.032	-0.026
	(0.069)	(0.068)	(0.068)	(0.051)
Full Mixed Room	-0.005	-0.070	-0.009	0.028
	(0.076)	(0.081)	(0.074)	(0.057)
P-val (Partial = Full)	0.250	0.972	0.704	0.229
DV Mean in same caste rm	0.397	0.505	0.472	0.246
R-squared	0.353	0.494	0.389	0.115
N	652	629	652	648

Notes: The table presents results for academic performance outcomes, including GPA, Mathematics and Programming grades, and summer internship or summer academic project completion. Sample is restricted to students with two roommates. Column 1 reports estimates for a composite Performance Index constructed using standardized scores of cumulative GPA, Mathematics and Programming grades, and summer internship or project completion. Column 2 reports results for cumulative GPA, defined as the standardized cumulative GPA up to the end of the first year. Column 3 reports results for Mathematics and Programming grades, measured as an index constructed using standardized scores from Mathematics and Programming courses. Column 4 reports results for summer internship or academic project completion, defined as an indicator for whether a student completed a summer internship or academic project during the summer following the first year. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective college entrance exam scores and number of roommates.

Table A12: Effect on Mental Health & Affirmative Action Support : Three Sharing Rooms

	Mental Health Index	PHQ4 Index	Relative well being Index	Worry Index	Supports AA Index	Benefits of AA to reserved groups	Costs of AA to non reserved groups (rev)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Lower Caste Sample							
Partial Mixed Room	0.105*	0.175**	-0.001	0.035	0.057	-0.010	0.127*
	(0.062)	(0.083)	(0.062)	(0.066)	(0.055)	(0.055)	(0.069)
Full Mixed Room	0.145**	0.164	0.026	0.089	-0.033	-0.072	0.025
	(0.070)	(0.101)	(0.071)	(0.078)	(0.068)	(0.069)	(0.086)
P-val (Partial = Full)	0.526	0.904	0.671	0.449	0.099	0.258	0.148
DV Mean in same caste rm	-0.015	-0.091	0.019	0.056	0.239	0.253	0.223
R-squared	0.201	0.153	0.151	0.181	0.191	0.150	0.229
N	603	568	518	550	542	538	534
Panel B: Upper Caste Sample							
Partial Mixed Room	-0.043	-0.069	-0.048	-0.018	0.145*	0.194**	0.094
	(0.074)	(0.104)	(0.072)	(0.072)	(0.082)	(0.084)	(0.097)
Full Mixed Room	0.021	0.109	-0.073	0.015	0.147	0.198**	0.057
	(0.078)	(0.114)	(0.082)	(0.079)	(0.094)	(0.095)	(0.106)
P-val (Partial = Full)	0.306	0.069	0.720	0.637	0.979	0.960	0.714
DV Mean in same caste rm	0.008	0.024	0.010	-0.004	-0.267	-0.251	-0.301
R-squared	0.171	0.156	0.143	0.107	0.216	0.206	0.198
N	627	593	520	547	568	566	561

Notes: The table presents results for mental health outcomes, including the Mental Health Index, PHQ-4 scores, relative happiness, and worry index scores. Sample is restricted to students with two roommates. Column 1 reports estimates for a composite Mental Health Index constructed using standardized scores of the individual mental health measures. Column 2 reports results for PHQ-4 scores, defined as the standardized PHQ-4 score measuring anxiety and depression symptoms. Column 3 reports results for worry index scores, defined as the standardized worries about various college aspects score. Column 4 reports estimates for a composite Affirmative Action Support Index constructed by mean of dependent variables in columns 5 and 6. Column 5 reports results for metritocratic concerns of affirmative action policies, and are reverse coded for easier interpretation. Column 6 reports results for social justice aspect of affirmative action. In all the columns, higher values of these indices indicate worse mental health outcomes. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective dependent variables and a missing indicator for the baseline controls and number of roommates.

Table A13: Effect on GPA, Internships, Maths and Programming performance: Two Sharing Rooms

	Performance Index (1)	GPA (2)	Maths & Programming (3)	Summer Project or Internship (4)
Panel A: Lower Caste Sample				
Mixed Room	0.036	0.076	0.071	-0.019
	(0.076)	(0.084)	(0.092)	(0.041)
DV Mean in same caste rm	-0.284	-0.477	-0.337	0.143
R-squared	0.390	0.490	0.426	0.124
N	396	377	394	396
Panel B: Upper Caste Sample				
Mixed Room	-0.072	-0.160*	-0.098	0.003
	(0.066)	(0.085)	(0.067)	(0.042)
DV Mean in same caste rm	0.312	0.415	0.448	0.203
R-squared	0.412	0.502	0.529	0.208
N	419	407	418	414
Panel C: Full Sample				
Mixed Room	-0.019	-0.036	-0.005	-0.016
	(0.051)	(0.062)	(0.058)	(0.030)
DV Mean in same caste rm	0.024	-0.010	0.069	0.174
R-squared	0.436	0.536	0.498	0.124
N	827	795	823	822

Notes: The table presents results for academic performance outcomes, including GPA, Mathematics and Programming grades, and summer internship or summer academic project completion. Sample is restricted to students with one roommate. Column 1 reports estimates for a composite Performance Index constructed using standardized scores of cumulative GPA, Mathematics and Programming grades, and summer internship or project completion. Column 2 reports results for cumulative GPA, defined as the standardized cumulative GPA up to the end of the first year. Column 3 reports results for Mathematics and Programming grades, measured as an index constructed using standardized scores from Mathematics and Programming courses. Column 4 reports results for summer internship or academic project completion, defined as an indicator for whether a student completed a summer internship or academic project during the summer following the first year. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective college entrance exam scores and number of roommates.

Table A14: Effect on Mental Health & Affirmative Action Support: Two Sharing Rooms

	Mental Health Index (1)	PHQ4 Index (2)	Relative well being Index (3)	Worry Index (4)	Supports AA Index (5)	Benefits of AA to reserved groups (6)	Costs of AA to non reserved groups (rev) (7)
Panel A: Lower Caste Sample							
Mixed Room	0.077	0.030	0.058	0.028	0.095	0.027	0.116
	(0.066)	(0.102)	(0.060)	(0.069)	(0.077)	(0.076)	(0.092)
DV Mean in same caste rm	0.039	0.097	0.052	0.037	0.241	0.228	0.249
R-squared	0.151	0.155	0.139	0.181	0.120	0.162	0.112
N	372	354	323	332	338	337	333
Panel B: Upper Caste Sample							
Mixed Room	0.172**	0.203*	0.081	0.057	0.011	-0.008	0.005
	(0.072)	(0.103)	(0.082)	(0.071)	(0.076)	(0.076)	(0.086)
DV Mean in same caste rm	-0.086	-0.044	-0.102	-0.086	-0.247	-0.250	-0.236
R-squared	0.212	0.237	0.116	0.221	0.257	0.216	0.266
N	394	371	319	339	336	334	333
Panel C: Full Sample							
Mixed Room	0.113**	0.118	0.064	0.038	0.045	-0.001	0.060
	(0.050)	(0.072)	(0.049)	(0.050)	(0.055)	(0.056)	(0.064)
DV Mean in same caste rm	-0.026	0.025	-0.027	-0.027	-0.004	-0.011	0.006
R-squared	0.149	0.159	0.100	0.160	0.253	0.244	0.214
N	783	739	649	686	681	678	673

Notes: The table presents results for mental health outcomes, including the Mental Health Index, PHQ-4 scores, relative happiness, and worry index scores. Sample is restricted to students with 1 roommate. Column 1 reports estimates for a composite Mental Health Index constructed using standardized scores of the individual mental health measures. Column 2 reports results for PHQ-4 scores, defined as the standardized PHQ-4 score measuring anxiety and depression symptoms. Column 3 reports results for worry index scores, defined as the standardized worries about various college aspects score. Column 4 reports estimates for a composite Affirmative Action Support Index constructed by mean of dependent variables in columns 5 and 6. Column 5 reports results for metritocratic concerns of affirmative action policies, and are reverse coded for easier interpretation. Column 6 reports results for social justice aspect of affirmative action. In all the columns, higher values of these indices indicate worse mental health outcomes. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective dependent variables and a missing indicator for the baseline controls and number of roommates.

Table A15: Effect on GPA, Internships, Maths and Programming performance

	Performance Index (1)	GPA (2)	Maths & Programming (3)	Summer Project or Internship (4)
Panel A: Upper Caste Sample				
OBC Roommate	-0.006	-0.030	-0.014	0.010
	(0.044)	(0.048)	(0.043)	(0.032)
SC ST Roommate	-0.062	-0.103*	-0.072	-0.004
	(0.046)	(0.056)	(0.046)	(0.032)
P-val (Has OBC rm = Has SC/ST rm)	0.276	0.261	0.280	0.710
DV Mean in same caste rm	0.348	0.453	0.458	0.221
R-squared	0.344	0.474	0.422	0.100
N	1078	1043	1076	1069
Panel B: Other Backward Caste Sample				
Upper caste Roommate	0.095	0.031	0.024	0.090**
	(0.067)	(0.080)	(0.078)	(0.039)
SC ST Roommate	-0.088	-0.126	-0.096	-0.010
	(0.064)	(0.081)	(0.075)	(0.039)
P-val (Has UC rm = Has SC/ST rm)	0.006	0.044	0.130	0.013
DV Mean in same caste rm	-0.200	-0.264	-0.183	0.103
R-squared	0.394	0.500	0.450	0.117
N	606	570	604	599
Panel C: Scheduled Caste & Scheduled Tribe Sample				
Upper caste Roommate	0.166**	0.123	0.191*	0.038
	(0.073)	(0.094)	(0.111)	(0.042)
OBC Roommate	0.135*	0.144	0.220**	0.020
	(0.076)	(0.111)	(0.108)	(0.044)
P-val (Has UC rm = Has OBC rm)	0.683	0.829	0.797	0.584
DV Mean in same caste rm	-0.565	-0.744	-0.710	0.090
R-squared	0.296	0.405	0.338	0.120
N	453	398	448	447

Notes: The table presents results for academic performance outcomes, including GPA, Mathematics and Programming grades, and summer internship or summer academic project completion. Column 1 reports estimates for a composite Performance Index constructed using standardized scores of cumulative GPA, Mathematics and Programming grades, and summer internship or project completion. Column 2 reports results for cumulative GPA, defined as the standardized cumulative GPA up to the end of the first year. Column 3 reports results for Mathematics and Programming grades, measured as an index constructed using standardized scores from Mathematics and Programming courses. Column 4 reports results for summer internship or academic project completion, defined as an indicator for whether a student completed a summer internship or academic project during the summer following the first year. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective college entrance exam scores and number of roommates.

Table A16: Effect on Mental Health & Affirmative Action Support

	Mental Health Index (1)	PHQ4 Index (2)	Worry Index (3)	Supports AA Index (4)	Benefits of AA to reserved groups (5)	Costs of AA to non reserved groups (rev) (6)
Panel A: Upper Caste Sample						
OBC Roommate	0.054	0.054	0.004	0.074	0.085	0.046
	(0.048)	(0.070)	(0.047)	(0.054)	(0.056)	(0.060)
SC ST Roommate	0.027	0.056	0.023	0.122**	0.127**	0.098
	(0.050)	(0.070)	(0.049)	(0.054)	(0.058)	(0.064)
P-val (Has OBC rm = Has SC/ST rm)	0.637	0.980	0.747	0.460	0.546	0.519
DV Mean in same caste rm	-0.046	-0.016	-0.052	-0.256	-0.251	-0.265
R-squared	0.135	0.144	0.112	0.200	0.175	0.187
N	1031	971	896	911	907	901
Panel B: Other Backward Caste Sample						
Upper caste Roommate	0.027	0.126	-0.045	0.066	0.001	0.127
	(0.059)	(0.094)	(0.065)	(0.062)	(0.063)	(0.079)
SC ST Roommate	0.041	0.140	-0.052	-0.045	-0.042	-0.040
	(0.063)	(0.096)	(0.066)	(0.064)	(0.067)	(0.080)
P-val (Has UC rm = Has SC/ST rm)	0.826	0.889	0.915	0.089	0.525	0.056
DV Mean in same caste rm	0.028	-0.012	0.033	0.094	0.117	0.064
R-squared	0.175	0.140	0.167	0.193	0.192	0.164
N	564	539	518	524	520	516
Panel C: Scheduled Caste & Scheduled Tribe Sample						
Upper caste Roommate	0.138*	0.075	0.066	-0.111	-0.155**	-0.075
	(0.077)	(0.116)	(0.085)	(0.069)	(0.072)	(0.085)
OBC Roommate	-0.047	-0.038	-0.026	-0.130*	-0.145*	-0.108
	(0.082)	(0.123)	(0.086)	(0.072)	(0.074)	(0.090)
P-val (Has UC rm = Has OBC rm)	0.031	0.408	0.309	0.808	0.903	0.723
DV Mean in same caste rm	-0.004	0.039	0.063	0.456	0.420	0.496
R-squared	0.233	0.214	0.255	0.216	0.210	0.277
N	420	394	373	363	362	357

Notes: The table presents results for mental health outcomes, including the Mental Health Index, PHQ-4 scores, relative happiness, and worry index scores. Column 1 reports estimates for a composite Mental Health Index constructed using standardized scores of the individual mental health measures. Column 2 reports results for PHQ-4 scores, defined as the standardized PHQ-4 score measuring anxiety and depression symptoms. Column 3 reports results for worry index scores, defined as the standardized worries about various college aspects score. Column 4 reports estimates for a composite Affirmative Action Support Index constructed by mean of dependent variables in columns 5 and 6. Column 5 reports results for metritocratic concerns of affirmative action policies, and are reverse coded for easier interpretation. Column 6 reports results for social justice aspect of affirmative action. In all the columns, higher values of these indices indicate worse mental health outcomes. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective dependent variables and a missing indicator for the baseline controls and number of roommates.

Table A17: Effect on Help Provided to roommates: Acads, Intern, Financial, Emotional Support

	Information Index (1)	Academic Information (2)	Job & Internship Information (3)	Administrative Information (4)	Social & Cultural Information (5)
Panel A: Lower Caste Sample					
Mixed Room	0.053	0.073	0.037	0.024	0.078
	(0.060)	(0.070)	(0.072)	(0.072)	(0.073)
DV Mean in same caste rm	-0.030	-0.034	-0.018	-0.003	-0.064
R-squared	0.085	0.073	0.099	0.078	0.063
N	939	939	939	939	939
Panel B: Upper Caste Sample					
Mixed Room	0.054	0.117*	-0.009	0.068	0.039
	(0.062)	(0.069)	(0.076)	(0.081)	(0.073)
DV Mean in same caste rm	0.031	0.035	0.019	0.003	0.067
R-squared	0.097	0.087	0.079	0.100	0.079
N	910	910	910	910	910
Panel C: Full Sample					
Mixed Room	0.062	0.101**	0.025	0.056	0.064
	(0.042)	(0.048)	(0.051)	(0.052)	(0.050)
DV Mean in same caste rm	-0.000	-0.000	0.000	0.000	-0.000
R-squared	0.065	0.058	0.062	0.056	0.049
N	1849	1849	1849	1849	1849

Notes: The table presents results for how easily students get information related to different aspects of college life in mixed rooms, Column 1 reports estimates for access to overall information, an index of different types of information releated to academics, internships, administration, social and cultural activities. Column 2 reports results for information related to academics. Column 3 reports results for information related to internships. Column 4 reports results for information related to administrative tasks. Column 5 reports results for information related to social and actural activities. All the variables are standardized (z-scores) with respect to the full sample. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions number of roommates as control.

Table A18: Effect on Self reported skills: Maths, Programming, English and Soft skills

	Self rating of Skills (1)	Rating: Math & Prog (2)	Rating: English (3)	Rating of Soft Skills (4)
Panel A: Lower Caste Sample				
Mixed Room	-0.044	-0.039	-0.098	-0.046
	(0.058)	(0.064)	(0.079)	(0.064)
DV Mean in same caste rm	-0.128	-0.164	-0.209	-0.114
R-squared	0.063	0.116	0.066	0.062
N	1058	1058	1058	1058
Panel B: Upper Caste Sample				
Mixed Room	-0.038	-0.099*	-0.105	-0.017
	(0.054)	(0.060)	(0.072)	(0.059)
DV Mean in same caste rm	0.138	0.193	0.212	0.116
R-squared	0.072	0.162	0.078	0.071
N	1078	1078	1078	1078
Panel C: Full Sample				
Mixed Room	-0.050	-0.081*	-0.092*	-0.040
	(0.040)	(0.046)	(0.053)	(0.044)
DV Mean in same caste rm	0.003	0.012	-0.001	-0.001
R-squared	0.065	0.135	0.085	0.054
N	2136	2136	2136	2136

Notes: The table presents results for self reported rating on skills, Column 1 reports estimates overall skill index, ranging from maths, programming, english and set of soft skills. Column 2 reports results for math and programming skills. Column 3 reports results for english skills Column 4 reports results for soft skills, measured as skill in communication,leadership,teamwork,time_management,networking All the variables are standardized (z-scores) with respect to the full sample. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective dependent variables and a missing indicator for the baseline controls and number of roommates.

Table A19: Effect on study time, attendance in classes, and time spent in jobs and student clubs

	Study hours : normal week (1)	Study hours: exam week (2)	Attendance: Last Sem (3)	Hours in internships/jobs (4)	Hours in student clubs (5)
Panel A: Lower Caste Sample					
Mixed Room	-0.031	0.050	-0.103*	-0.540	0.285
	(0.085)	(0.089)	(0.062)	(0.696)	(0.818)
DV Mean in same caste rm	0.101	-0.061	0.095	6.216	8.714
R-squared	0.095	0.116	0.079	0.108	0.069
N	1055	1055	978	1048	1049
Panel B: Upper Caste Sample					
Mixed Room	0.182**	0.068	-0.159*	0.979	-0.064
	(0.076)	(0.091)	(0.091)	(0.823)	(0.905)
DV Mean in same caste rm	-0.111	0.050	-0.099	5.111	9.970
R-squared	0.095	0.088	0.124	0.103	0.079
N	1077	1077	989	1071	1075
Panel C: Full Sample					
Mixed Room	0.080	0.059	-0.145**	0.304	0.107
	(0.058)	(0.066)	(0.057)	(0.532)	(0.602)
DV Mean in same caste rm	-0.003	-0.006	-0.000	5.672	9.334
R-squared	0.046	0.061	0.088	0.086	0.051
N	2132	2132	1967	2119	2124

Notes: The table presents results for academic effort and time spent in jobs and student clubs, Column 1 reports estimates studying during normal week in college. Column 2 reports results for time spent studying in exam week in college Column 3 reports results for standardized attendance in last semester. Column 4 reports results for time spent in internships Column 5 reports results for time spent in student clubs. All the variables are standardized (z-scores) with respect to the full sample. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective dependent variables and a missing indicator for the baseline controls and number of roommates.

Table A20: Effect on GSE, Leadership and Career Aspirations

	GSE Index (1)	Expected salary (2)	Clubs Leadership (3)	Preferred job sector: Finance/ Consulting (4)
Panel A: Lower Caste Sample				
Mixed Room	-0.069	5.323*	0.001	-0.019
	(0.081)	(3.104)	(0.056)	(0.049)
DV Mean in same caste rm	0.023	37.164	-0.161	0.308
R-squared	0.086	0.080	0.073	0.083
N	460	527	475	483
Panel B: Upper Caste Sample				
Mixed Room	-0.033	2.125	-0.127**	-0.106**
	(0.087)	(3.178)	(0.060)	(0.049)
DV Mean in same caste rm	-0.025	45.356	0.137	0.520
R-squared	0.050	0.212	0.097	0.128
N	475	533	505	496
Panel C: Full Sample				
Mixed Room	-0.051	3.076	-0.055	-0.065*
	(0.061)	(2.326)	(0.040)	(0.034)
DV Mean in same caste rm	-0.002	41.485	0.000	0.421
R-squared	0.029	0.124	0.110	0.115
N	935	1060	980	979

Notes: The table presents results for GSE index, expectes salary, if they held any student leadership positions, and preference for job in finance and consulting sector repectively in columns 1-4. These questions were asked only for 2024 cohort module. All the variables are standardized (z-scores) with respect to the full sample. All regressions include hostel and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level.

Table A21: Effect on GPA, Internships, Maths and Programming performance: Hetrogeneity by absolute difference of entrance scores with Roommate

	Performance Index (1)	GPA (2)	Maths & Programming (3)	Internship (4)
Panel A: Lower Caste Sample				
Mixed Room	0.129**	0.078	0.080	0.058***
	(0.053)	(0.060)	(0.068)	(0.022)
Absolute Score diff with RM (z-score)	-0.046	-0.003	-0.025	-0.005
	(0.032)	(0.038)	(0.046)	(0.011)
Absolute Score diff with RM x Mixed Room	0.035	0.016	0.044	-0.004
	(0.038)	(0.047)	(0.055)	(0.014)
DV Mean in same caste rm	-0.357	-0.466	-0.411	0.032
R-squared	0.338	0.435	0.374	0.083
N	1059	968	1052	1047
Panel B: Upper Caste Sample				
Mixed Room	-0.025	-0.068	-0.032	0.028
	(0.049)	(0.054)	(0.050)	(0.021)
Absolute Score diff with RM (z-score)	-0.041	-0.021	-0.017	-0.026**
	(0.033)	(0.040)	(0.038)	(0.010)
Absolute Score diff with RM x Mixed Room	0.018	-0.010	-0.002	0.029**
	(0.037)	(0.043)	(0.042)	(0.014)
DV Mean in same caste rm	0.348	0.453	0.458	0.064
R-squared	0.346	0.475	0.422	0.075
N	1078	1043	1076	1069
Panel C: Full Sample				
Mixed Room	0.053	-0.014	0.044	0.039**
	(0.036)	(0.041)	(0.042)	(0.015)
Absolute Score diff with RM (z-score)	-0.040*	-0.014	-0.026	-0.008
	(0.023)	(0.027)	(0.031)	(0.007)
Absolute Score diff with RM x Mixed Room	0.017	0.005	0.007	0.005
	(0.026)	(0.032)	(0.036)	(0.009)
DV Mean in same caste rm	-0.009	0.000	0.018	0.048
R-squared	0.411	0.515	0.457	0.048
N	2137	2011	2128	2116

Notes: The table presents results for academic performance outcomes, including GPA, Mathematics and Programming grades, and summer internship or summer academic project completion. I look at heterogeneity by including an interaction of mixed room with absolute difference with roommates entrance exam score. I took average absolute difference of entrance exam score with roommates and then standardized it. Column 1 reports estimates for a composite Performance Index constructed using standardized scores of cumulative GPA, Mathematics and Programming grades, and summer internship or project completion. Column 2 reports results for cumulative GPA, defined as the standardized cumulative GPA at the end of the first year. Column 3 reports results for Mathematics and Programming grades, measured as an index constructed using standardized scores from Mathematics and Programming courses. Column 4 reports results for summer internship or academic project completion, defined as an indicator for whether a student completed a summer internship or academic project during the summer following the first year. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective college entrance exam scores and number of roommates.

Table A22: Effect on Mental Health and Affirmative action: Hetrogeneity by absolute difference of entrance scores with Roommate

	Mental Health Index (1)	PHQ4 Index (2)	Relative well being Index (3)	Worry Index (4)	Supports AA Index (5)	Benefits of AA to reserved groups (6)	Costs of AA to non reserved groups (rev) (7)
Panel A: Lower Caste Sample							
Mixed Room	0.074	0.080	-0.022	0.013	0.017	-0.036	0.063
	(0.047)	(0.069)	(0.047)	(0.053)	(0.051)	(0.051)	(0.063)
Absolute Score diff with RM (z-score)	0.044	0.063	-0.029	-0.003	0.013	0.018	0.003
	(0.038)	(0.051)	(0.029)	(0.036)	(0.032)	(0.038)	(0.043)
Absolute Score diff with RM x Mixed Room	-0.018	-0.038	0.057*	0.034	0.026	0.010	0.052
	(0.040)	(0.057)	(0.034)	(0.043)	(0.037)	(0.042)	(0.048)
DV Mean in same caste rm	0.014	0.010	0.036	0.046	0.240	0.239	0.237
R-squared	0.117	0.102	0.090	0.115	0.076	0.063	0.107
N	985	934	850	892	888	883	874
Panel B: Upper Caste Sample							
Mixed Room	0.041	0.057	0.027	-0.057	0.103*	0.136**	0.064
	(0.052)	(0.075)	(0.054)	(0.053)	(0.062)	(0.065)	(0.067)
Absolute Score diff with RM (z-score)	0.064*	0.087	0.097**	0.017	0.037	0.018	0.046
	(0.039)	(0.053)	(0.043)	(0.040)	(0.055)	(0.055)	(0.059)
Absolute Score diff with RM x Mixed Room	-0.065	-0.089	-0.131**	0.039	-0.072	-0.062	-0.077
	(0.044)	(0.062)	(0.052)	(0.047)	(0.062)	(0.062)	(0.067)
DV Mean in same caste rm	-0.046	-0.016	-0.055	-0.052	-0.256	-0.251	-0.265
R-squared	0.101	0.098	0.060	0.108	0.117	0.128	0.111
N	1031	971	846	896	911	907	901
Panel C: Full Sample							
Mixed Room	0.075**	0.094*	0.030	-0.005	0.072*	0.054	0.085*
	(0.035)	(0.049)	(0.034)	(0.037)	(0.041)	(0.041)	(0.049)
Absolute Score diff with RM (z-score)	0.038	0.058	0.029	-0.001	0.003	0.013	-0.011
	(0.027)	(0.036)	(0.025)	(0.026)	(0.034)	(0.029)	(0.049)
Absolute Score diff with RM x Mixed Room	-0.035	-0.057	-0.041	0.034	0.001	-0.022	0.026
	(0.030)	(0.041)	(0.029)	(0.030)	(0.037)	(0.034)	(0.052)
DV Mean in same caste rm	-0.016	-0.003	-0.009	-0.003	-0.005	-0.002	-0.012
R-squared	0.076	0.064	0.037	0.079	0.177	0.166	0.154
N	2016	1905	1696	1788	1799	1790	1775

Notes: The table presents results for mental health outcomes, including the Mental Health Index, PHQ-4 scores, relative happiness, and worry index scores. Column 1 reports estimates for a composite Mental Health Index constructed using standardized scores of the individual mental health measures. Column 2 reports results for PHQ-4 scores, defined as the standardized PHQ-4 score measuring anxiety and depression symptoms. Column 3 reports results for worry index scores, defined as the standardized worries about various college aspects score. Column 4 reports estimates for a composite Affirmative Action Support Index constructed by mean of dependent variables in columns 5 and 6. Column 5 reports results for metritocratic concerns of affirmative action policies, and are reverse coded for easier interpretation. Column 6 reports results for social justice aspect of affirmative action. In all the columns, higher values of these indices indicate worse mental health outcomes. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective dependent variables and a missing indicator for the baseline controls and number of roommates.

Table A23: Effect on GPA, Internships, Maths and Programming performance: Hetrogeneity by difference of entrance scores with Roommate

	Performance Index (1)	GPA (2)	Maths & Programming (3)	Internship (4)
Panel A: Lower Caste Sample				
Mixed Room	0.029	0.041	-0.023	0.038
	(0.075)	(0.090)	(0.102)	(0.025)
Indicator for positive score diff with RM	0.154**	0.023	0.195*	0.051**
	(0.076)	(0.092)	(0.106)	(0.025)
Entrance Exam Score diff with RM (z-score)	0.092**	0.104	0.060	-0.004
	(0.046)	(0.064)	(0.074)	(0.021)
Std Score diff with RM*Ind for positive diff	-0.105	0.014	-0.059	-0.008
	(0.084)	(0.099)	(0.126)	(0.029)
Entrance Exam Score diff with RM x Mixed Room	-0.175**	-0.132	-0.131	-0.032
	(0.079)	(0.094)	(0.095)	(0.037)
Std Score diff with RM x Mixed Room x Ind for positive diff	0.180	0.046	0.146	0.034
	(0.119)	(0.139)	(0.159)	(0.045)
DV Mean in same caste rm	-0.357	-0.466	-0.411	0.032
R-squared	0.348	0.441	0.380	0.086
N	1059	968	1052	1047
Panel B: Upper Caste Sample				
Mixed Room	-0.115	-0.145*	-0.109	-0.005
	(0.074)	(0.081)	(0.079)	(0.030)
Indicator for positive score diff with RM	-0.140**	-0.179**	-0.153*	-0.011
	(0.068)	(0.082)	(0.079)	(0.029)
Entrance Exam Score diff with RM (z-score)	0.103	0.121	0.094	0.049**
	(0.066)	(0.077)	(0.071)	(0.020)
Std Score diff with RM*Ind for positive diff	-0.125	-0.063	-0.075	-0.086***
	(0.096)	(0.120)	(0.107)	(0.030)
Entrance Exam Score diff with RM x Mixed Room	-0.077	-0.082	-0.069	-0.052**
	(0.067)	(0.077)	(0.072)	(0.022)
Std Score diff with RM x Mixed Room x Ind for positive diff	0.028	-0.053	0.048	0.080
	(0.127)	(0.158)	(0.171)	(0.053)
DV Mean in same caste rm	0.348	0.453	0.458	0.064
R-squared	0.350	0.479	0.425	0.075
N	1078	1043	1076	1069
Panel C: Full Sample				
Mixed Room	0.038	-0.019	0.050	0.032
	(0.049)	(0.057)	(0.058)	(0.019)
Indicator for positive score diff with RM	0.003	-0.084	0.018	0.017
	(0.046)	(0.054)	(0.056)	(0.020)
Entrance Exam Score diff with RM (z-score)	0.095**	0.128**	0.067	0.012
	(0.038)	(0.050)	(0.053)	(0.015)
Std Score diff with RM*Ind for positive diff	-0.085	-0.011	-0.045	-0.023
	(0.062)	(0.075)	(0.087)	(0.020)
Entrance Exam Score diff with RM x Mixed Room	-0.050	-0.075	-0.013	-0.015
	(0.041)	(0.053)	(0.055)	(0.016)
Std Score diff with RM x Mixed Room x Ind for positive diff	0.034	-0.000	-0.013	0.017
	(0.071)	(0.088)	(0.097)	(0.025)
DV Mean in same caste rm	-0.009	0.000	0.018	0.048
R-squared	0.412	0.518	0.459	0.049
N	2137	2011	2128	2116

Notes: The table presents results for academic performance outcomes, including GPA, Mathematics and Programming grades, and summer internship or summer academic project completion. I look at heterogeneity by including an interaction of mixed room with standardized difference with roommates entrance exam score. I took average difference difference of extrance exam score with roommates and then standardized it. I also include an indicator for positive difference (roommates having higher entrance exam score than student) and its interaction with mixed room to look at slope changes. Column 1 reports estimates for a composite Performance Index constructed using standardized scores of cumulative GPA, Mathematics and Programming grades, and summer internship or project completion. Column 2 reports results for cumulative GPA, defined as the standardized cumulative GPA at the end of the first year. Column 3 reports results for Mathematics and Programming grades, measured as an index constructed using standardized scores from Mathematics and Programming courses. Column 4 reports results for summer internship or academic project completion, defined as an indicator for whether a student completed a summer internship or academic project during the summer following the first year. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective college entrance exam scores and number of roommates.

Table A24: Effect on Mental Health and Affirmative action: Hetrogeneity by difference of entrance scores with Roommate

	Mental Health Index (1)	PHQ4 Index (2)	Relative well being Index (3)	Worry Index (4)	Supports AA Index (5)	Benefits of AA to reserved groups (6)	Costs of AA to non reserved groups (rev) (7)
Panel A: Lower Caste Sample							
Mixed Room	0.107	0.145	-0.049	-0.042	0.023	-0.030	0.057
	(0.074)	(0.105)	(0.070)	(0.078)	(0.074)	(0.074)	(0.094)
Indicator for positive score diff with RM	-0.082	-0.167	-0.007	0.073	-0.124	-0.080	-0.155
	(0.083)	(0.121)	(0.083)	(0.079)	(0.078)	(0.074)	(0.104)
Entrance Exam Score diff with RM (z-score)	-0.053	0.001	0.019	-0.036	0.148**	0.131**	0.168**
	(0.060)	(0.093)	(0.057)	(0.066)	(0.059)	(0.060)	(0.072)
Std Score diff with RM*Ind for positive diff	0.070	0.119	-0.124	-0.053	0.016	0.032	-0.011
	(0.099)	(0.138)	(0.082)	(0.103)	(0.094)	(0.095)	(0.132)
Entrance Exam Score diff with RM x Mixed Room	0.024	-0.041	-0.087	0.028	-0.054	-0.023	-0.102
	(0.091)	(0.136)	(0.091)	(0.091)	(0.090)	(0.086)	(0.115)
Std Score diff with RM x Mixed Room x Ind for positive diff	-0.011	-0.038	0.197	0.092	-0.042	-0.090	0.034
	(0.132)	(0.185)	(0.124)	(0.134)	(0.125)	(0.119)	(0.178)
DV Mean in same caste rm	0.014	0.010	0.036	0.046	0.240	0.239	0.237
R-squared	0.120	0.104	0.092	0.118	0.105	0.091	0.125
N	985	934	850	892	888	883	874
Panel B: Upper Caste Sample							
Mixed Room	0.062	0.077	0.144*	-0.055	0.252***	0.311***	0.191**
	(0.074)	(0.111)	(0.083)	(0.085)	(0.083)	(0.087)	(0.096)
Indicator for positive score diff with RM	-0.108	-0.163	-0.018	0.056	0.133*	0.244***	0.038
	(0.078)	(0.120)	(0.076)	(0.088)	(0.080)	(0.081)	(0.100)
Entrance Exam Score diff with RM (z-score)	0.075	0.110	-0.041	0.002	-0.061	-0.098	-0.016
	(0.065)	(0.092)	(0.073)	(0.068)	(0.088)	(0.089)	(0.100)
Std Score diff with RM*Ind for positive diff	0.120	0.186	0.228*	0.050	0.250*	0.214	0.265*
	(0.103)	(0.147)	(0.121)	(0.109)	(0.139)	(0.138)	(0.153)
Entrance Exam Score diff with RM x Mixed Room	-0.061	-0.094	0.090	-0.076	0.094	0.126	0.058
	(0.066)	(0.095)	(0.079)	(0.072)	(0.092)	(0.092)	(0.105)
Std Score diff with RM x Mixed Room x Ind for positive diff	-0.206	-0.280	-0.433***	0.009	-0.442**	-0.483**	-0.388*
	(0.134)	(0.202)	(0.158)	(0.154)	(0.188)	(0.187)	(0.212)
DV Mean in same caste rm	-0.046	-0.016	-0.055	-0.052	-0.256	-0.251	-0.265
R-squared	0.109	0.107	0.063	0.109	0.130	0.145	0.120
N	1031	971	846	896	911	907	901
Panel C: Full Sample							
Mixed Room	0.095**	0.126*	0.062	-0.029	0.097*	0.094*	0.094
	(0.046)	(0.066)	(0.048)	(0.053)	(0.057)	(0.056)	(0.070)
Indicator for positive score diff with RM	-0.072	-0.134*	-0.061	0.085	-0.024	0.027	-0.065
	(0.049)	(0.074)	(0.050)	(0.052)	(0.054)	(0.051)	(0.068)
Entrance Exam Score diff with RM (z-score)	-0.014	0.027	-0.002	-0.035	0.075	0.047	0.102*
	(0.047)	(0.067)	(0.042)	(0.047)	(0.051)	(0.052)	(0.059)
Std Score diff with RM*Ind for positive diff	0.062	0.109	0.039	-0.013	0.064	0.098	0.028
	(0.074)	(0.101)	(0.069)	(0.075)	(0.102)	(0.079)	(0.148)
Entrance Exam Score diff with RM x Mixed Room	0.031	-0.000	0.039	-0.025	-0.041	-0.021	-0.055
	(0.050)	(0.071)	(0.048)	(0.051)	(0.058)	(0.057)	(0.068)
Std Score diff with RM x Mixed Room x Ind for positive diff	-0.059	-0.107	-0.088	0.093	-0.063	-0.119	-0.006
	(0.082)	(0.115)	(0.079)	(0.086)	(0.110)	(0.089)	(0.156)
DV Mean in same caste rm	-0.016	-0.003	-0.009	-0.003	-0.005	-0.002	-0.012
R-squared	0.076	0.066	0.038	0.080	0.188	0.177	0.163
N	2016	1905	1696	1788	1799	1790	1775

Notes: The table presents results for mental health outcomes, including the Mental Health Index, PHQ-4 scores, relative happiness, and worry index scores. I look at heterogeneity by including an interaction of mixed room with score difference with roommates entrance exam score (roommate - student). I also include an indicator for positive difference (roommates having higher entrance exam score than student) and its interaction with mixed room to look at slope changes. I took average absolute difference of entrance exam score with roommates and then standardized it. Column 1 reports estimates for a composite Mental Health Index constructed using standardized scores of the individual mental health measures. Column 2 reports results for PHQ-4 scores, defined as the standardized PHQ-4 score measuring anxiety and depression symptoms. Column 3 reports results for worry index scores, defined as the standardized worries about various college aspects score. Column 4 reports estimates for a composite Affirmative Action Support Index constructed by mean of dependent variables in columns 5 and 6. Column 5 reports results for metritocratic concerns of affirmative action policies, and are reverse coded for easier interpretation. Column 6 reports results for social justice aspect of affirmative action. In all the columns, higher values of these indices indicate worse mental health outcomes. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective dependent variables and a missing indicator for the baseline controls and number of roommates.

Table A25: Effect on GPA, Internships, Maths and Programming performance: Hetrogeneity by difference of entrance scores with Roommate

	Performance Index (1)	GPA (2)	Maths & Programming (3)	Internship (4)
Panel A: Lower Caste Sample				
Mixed Room	0.029	0.041	-0.023	0.038
	(0.075)	(0.090)	(0.102)	(0.025)
Indicator for negative score diff with RM	-0.154**	-0.023	-0.195*	-0.051**
	(0.076)	(0.092)	(0.106)	(0.025)
Entrance Exam Score diff with RM (z-score)	-0.013	0.118*	0.001	-0.013
	(0.063)	(0.071)	(0.093)	(0.019)
Std Score diff with RM*Ind for negative diff	0.105	-0.014	0.059	0.008
	(0.084)	(0.099)	(0.126)	(0.029)
Entrance Exam Score diff with RM x Mixed Room	0.005	-0.086	0.015	0.002
	(0.063)	(0.073)	(0.093)	(0.019)
Std Score diff with RM x Mixed Room x Ind for negative diff	-0.180	-0.046	-0.146	-0.034
	(0.119)	(0.139)	(0.159)	(0.045)
DV Mean in same caste rm	-0.357	-0.466	-0.411	0.032
R-squared	0.348	0.441	0.380	0.086
N	1059	968	1052	1047
Panel B: Upper Caste Sample				
Mixed Room	-0.115	-0.145*	-0.109	-0.005
	(0.074)	(0.081)	(0.079)	(0.030)
Indicator for negative score diff with RM	0.140**	0.179**	0.153*	0.011
	(0.068)	(0.082)	(0.079)	(0.029)
Entrance Exam Score diff with RM (z-score)	-0.022	0.058	0.019	-0.036*
	(0.068)	(0.088)	(0.080)	(0.020)
Std Score diff with RM*Ind for negative diff	0.125	0.063	0.075	0.086***
	(0.096)	(0.120)	(0.107)	(0.030)
Entrance Exam Score diff with RM x Mixed Room	-0.049	-0.135	-0.020	0.028
	(0.094)	(0.118)	(0.135)	(0.040)
Std Score diff with RM x Mixed Room x Ind for negative diff	-0.028	0.053	-0.048	-0.080
	(0.127)	(0.158)	(0.171)	(0.053)
DV Mean in same caste rm	0.348	0.453	0.458	0.064
R-squared	0.350	0.479	0.425	0.075
N	1078	1043	1076	1069
Panel C: Full Sample				
Mixed Room	0.038	-0.019	0.050	0.032
	(0.049)	(0.057)	(0.058)	(0.019)
Indicator for negative score diff with RM	-0.003	0.084	-0.018	-0.017
	(0.046)	(0.054)	(0.056)	(0.020)
Entrance Exam Score diff with RM (z-score)	0.010	0.117**	0.023	-0.011
	(0.045)	(0.054)	(0.066)	(0.013)
Std Score diff with RM*Ind for negative diff	0.085	0.011	0.045	0.023
	(0.062)	(0.075)	(0.087)	(0.020)
Entrance Exam Score diff with RM x Mixed Room	-0.017	-0.075	-0.026	0.003
	(0.047)	(0.057)	(0.068)	(0.014)
Std Score diff with RM x Mixed Room x Ind for negative diff	-0.034	0.000	0.013	-0.017
	(0.071)	(0.088)	(0.097)	(0.025)
DV Mean in same caste rm	-0.009	0.000	0.018	0.048
R-squared	0.412	0.518	0.459	0.049
N	2137	2011	2128	2116

Notes: The table presents results for academic performance outcomes, including GPA, Mathematics and Programming grades, and summer internship or summer academic project completion. I look at heterogeneity by including an interaction of mixed room with standardized difference with roommates entrance exam score. I took average difference difference of extrance exam score with roommates and then standardized it. I also include an indicator for negative difference (roommates having higher entrance exam score than student) and its interaction with mixed room to look at slope changes. Column 1 reports estimates for a composite Performance Index constructed using standardized scores of cumulative GPA, Mathematics and Programming grades, and summer internship or project completion. Column 2 reports results for cumulative GPA, defined as the standardized cumulative GPA at the end of the first year. Column 3 reports results for Mathematics and Programming grades, measured as an index constructed using standardized scores from Mathematics and Programming courses. Column 4 reports results for summer internship or academic project completion, defined as an indicator for whether a student completed a summer internship or academic project during the summer following the first year. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective college entrance exam scores and number of roommates.

Table A26: Effect on Mental Health and Affirmative action: Hetrogeneity by difference of entrance scores with Roommate

	Mental Health Index (1)	PHQ4 Index (2)	Relative well being Index (3)	Worry Index (4)	Supports AA Index (5)	Benefits of AA to reserved groups (6)	Costs of AA to non reserved groups (rev) (7)
Panel A: Lower Caste Sample							
Mixed Room	0.107	0.145	-0.049	-0.042	0.023	-0.030	0.057
	(0.074)	(0.105)	(0.070)	(0.078)	(0.074)	(0.074)	(0.094)
Indicator for negative score diff with RM	0.082	0.167	0.007	-0.073	0.124	0.080	0.155
	(0.083)	(0.121)	(0.083)	(0.079)	(0.078)	(0.074)	(0.104)
Entrance Exam Score diff with RM (z-score)	0.017	0.119	-0.105	-0.090	0.164***	0.163***	0.157
	(0.071)	(0.092)	(0.064)	(0.071)	(0.063)	(0.058)	(0.104)
Std Score diff with RM*Ind for negative diff	-0.070	-0.119	0.124	0.053	-0.016	-0.032	0.011
	(0.099)	(0.138)	(0.082)	(0.103)	(0.094)	(0.095)	(0.132)
Entrance Exam Score diff with RM x Mixed Room	0.013	-0.079	0.109*	0.120*	-0.096	-0.112*	-0.068
	(0.066)	(0.089)	(0.062)	(0.070)	(0.064)	(0.060)	(0.104)
Std Score diff with RM x Mixed Room x Ind for negative diff	0.011	0.038	-0.197	-0.092	0.042	0.090	-0.034
	(0.132)	(0.185)	(0.124)	(0.134)	(0.125)	(0.119)	(0.178)
DV Mean in same caste rm	0.014	0.010	0.036	0.046	0.240	0.239	0.237
R-squared	0.120	0.104	0.092	0.118	0.105	0.091	0.125
N	985	934	850	892	888	883	874
Panel B: Upper Caste Sample							
Mixed Room	0.062	0.077	0.144*	-0.055	0.252***	0.311***	0.191**
	(0.074)	(0.111)	(0.083)	(0.085)	(0.083)	(0.087)	(0.096)
Indicator for negative score diff with RM	0.108	0.163	0.018	-0.056	-0.133*	-0.244***	-0.038
	(0.078)	(0.120)	(0.076)	(0.088)	(0.080)	(0.081)	(0.100)
Entrance Exam Score diff with RM (z-score)	0.195**	0.296***	0.186**	0.052	0.189**	0.116	0.249***
	(0.077)	(0.107)	(0.090)	(0.081)	(0.085)	(0.083)	(0.096)
Std Score diff with RM*Ind for negative diff	-0.120	-0.186	-0.228*	-0.050	-0.250*	-0.214	-0.265*
	(0.103)	(0.147)	(0.121)	(0.109)	(0.139)	(0.138)	(0.153)
Entrance Exam Score diff with RM x Mixed Room	-0.267***	-0.374**	-0.343***	-0.067	-0.348***	-0.357***	-0.329**
	(0.101)	(0.155)	(0.120)	(0.113)	(0.133)	(0.133)	(0.151)
Std Score diff with RM x Mixed Room x Ind for negative diff	0.206	0.280	0.433***	-0.009	0.442**	0.483**	0.388*
	(0.134)	(0.202)	(0.158)	(0.154)	(0.188)	(0.187)	(0.212)
DV Mean in same caste rm	-0.046	-0.016	-0.055	-0.052	-0.256	-0.251	-0.265
R-squared	0.109	0.107	0.063	0.109	0.130	0.145	0.120
N	1031	971	846	896	911	907	901
Panel C: Full Sample							
Mixed Room	0.095**	0.126*	0.062	-0.029	0.097*	0.094*	0.094
	(0.046)	(0.066)	(0.048)	(0.053)	(0.057)	(0.056)	(0.070)
Indicator for negative score diff with RM	0.072	0.134*	0.061	-0.085	0.024	-0.027	0.065
	(0.049)	(0.074)	(0.050)	(0.052)	(0.054)	(0.051)	(0.068)
Entrance Exam Score diff with RM (z-score)	0.048	0.136**	0.037	-0.047	0.139*	0.145***	0.130
	(0.052)	(0.068)	(0.054)	(0.053)	(0.073)	(0.044)	(0.120)
Std Score diff with RM*Ind for negative diff	-0.062	-0.109	-0.039	0.013	-0.064	-0.098	-0.028
	(0.074)	(0.101)	(0.069)	(0.075)	(0.102)	(0.079)	(0.148)
Entrance Exam Score diff with RM x Mixed Room	-0.027	-0.107	-0.049	0.068	-0.104	-0.140***	-0.061
	(0.051)	(0.069)	(0.052)	(0.053)	(0.073)	(0.048)	(0.116)
Std Score diff with RM x Mixed Room x Ind for negative diff	0.059	0.107	0.088	-0.093	0.063	0.119	0.006
	(0.082)	(0.115)	(0.079)	(0.086)	(0.110)	(0.089)	(0.156)
DV Mean in same caste rm	-0.016	-0.003	-0.009	-0.003	-0.005	-0.002	-0.012
R-squared	0.076	0.066	0.038	0.080	0.188	0.177	0.163
N	2016	1905	1696	1788	1799	1790	1775

Notes: The table presents results for mental health outcomes, including the Mental Health Index, PHQ-4 scores, relative happiness, and worry index scores. I look at heterogeneity by including an interaction of mixed room with score difference with roommates entrance exam score (roommate - student). I also include an indicator for negative difference (roommates having higher entrance exam score than student) and its interaction with mixed room to look at slope changes. I took average absolute difference of entrance exam score with roommates and then standardized it. Column 1 reports estimates for a composite Mental Health Index constructed using standardized scores of the individual mental health measures. Column 2 reports results for PHQ-4 scores, defined as the standardized PHQ-4 score measuring anxiety and depression symptoms. Column 3 reports results for worry index scores, defined as the standardized worries about various college aspects score. Column 4 reports estimates for a composite Affirmative Action Support Index constructed by mean of dependent variables in columns 5 and 6. Column 5 reports results for metritocratic concerns of affirmative action policies, and are reverse coded for easier interpretation. Column 6 reports results for social justice aspect of affirmative action. In all the columns, higher values of these indices indicate worse mental health outcomes. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective dependent variables and a missing indicator for the baseline controls and number of roommates.

Table A27: Effect on GPA, Internships, Maths and Programming performance: Effects by caste distance

	Performance Index (1)	GPA (2)	Maths & Programming (3)	Internship (4)
Panel A: Lower Caste Sample				
Caste distance with roommates	0.015	0.019	0.023	0.009
	(0.036)	(0.048)	(0.053)	(0.014)
DV Mean in same caste rm	-0.284	-0.474	-0.307	0.025
R-squared	0.330	0.433	0.371	0.073
N	1059	968	1052	1047
Panel B: Upper Caste Sample				
Caste distance with roommates	-0.050	-0.090**	-0.057*	0.013
	(0.031)	(0.040)	(0.032)	(0.013)
DV Mean in same caste rm	0.353	0.459	0.464	0.064
R-squared	0.344	0.475	0.422	0.072
N	1076	1041	1074	1067
Panel C: Full Sample				
Caste distance with roommates	-0.032	-0.050	-0.037	0.009
	(0.024)	(0.033)	(0.029)	(0.009)
DV Mean in same caste rm	-0.009	0.000	0.018	0.048
R-squared	0.409	0.516	0.457	0.044
N	2135	2009	2126	2114

Notes: The table presents results for academic performance outcomes, including GPA, Mathematics and Programming grades, and summer internship or summer academic project completion. Instead of looking at the effect of mixed room, I look at the effect of caste distance with roommates on these outcomes. I took average distance of caste (absolute) with roommates and then standardized it. I assigned upper caste as 1, OBC as 2, SC and ST as 3 for this purpose and took absolute distance with roommates and took average with roommates. Column 1 reports estimates for a composite Performance Index constructed using standardized scores of cumulative GPA, Mathematics and Programming grades, and summer internship or project completion. Column 2 reports results for cumulative GPA, defined as the standardized cumulative GPA at the end of the first year. Column 3 reports results for Mathematics and Programming grades, measured as an index constructed using standardized scores from Mathematics and Programming courses. Column 4 reports results for summer internship or academic project completion, defined as an indicator for whether a student completed a summer internship or academic project during the summer following the first year. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective college entrance exam scores and number of roommates.

Table A28: Effect on Mental Health and Affirmative action: Effects by caste distance

	Mental Health Index (1)	PHQ4 Index (2)	Relative well being Index (3)	Worry Index (4)	Supports AA Index (5)	Benefits of AA to reserved groups (6)	Costs of AA to non reserved groups (rev) (7)
Panel A: Lower Caste Sample							
Caste distance with roommates	0.103***	0.109**	0.036	0.059	0.091***	0.045	0.136***
	(0.038)	(0.055)	(0.036)	(0.043)	(0.034)	(0.036)	(0.040)
DV Mean in same caste rm	0.010	0.008	0.049	0.059	0.250	0.251	0.246
R-squared	0.117	0.102	0.089	0.114	0.078	0.062	0.108
N	985	934	850	892	888	883	874
Panel B: Upper Caste Sample							
Caste distance with roommates	0.046	0.075	0.015	0.019	0.048	0.059	0.023
	(0.031)	(0.046)	(0.036)	(0.032)	(0.036)	(0.037)	(0.042)
DV Mean in same caste rm	-0.049	-0.021	-0.057	-0.055	-0.262	-0.256	-0.273
R-squared	0.099	0.098	0.053	0.103	0.112	0.122	0.106
N	1029	969	844	894	909	905	899
Panel C: Full Sample							
Caste distance with roommates	0.059**	0.082**	0.019	0.027	0.071***	0.052**	0.082***
	(0.025)	(0.035)	(0.025)	(0.026)	(0.026)	(0.026)	(0.032)
DV Mean in same caste rm	-0.016	-0.003	-0.009	-0.003	-0.005	-0.002	-0.012
R-squared	0.074	0.063	0.036	0.077	0.179	0.167	0.155
N	2014	1903	1694	1786	1797	1788	1773

Notes: The table presents results for mental health outcomes, including the Mental Health Index, PHQ-4 scores, relative happiness, and worry index scores. Instead of looking at the effect of mixed room, I look at the effect of caste distance with roommates on these outcomes. I took average distance of caste (absolute) with roommates and then standardized it. I assigned upper caste as 1, OBC as 2, SC and ST as 3 for this purpose and took absolute distance with roommates and took average with roommates. Column 1 reports estimates for a composite Mental Health Index constructed using standardized scores of the individual mental health measures. Column 2 reports results for PHQ-4 scores, defined as the standardized PHQ-4 score measuring anxiety and depression symptoms. Column 3 reports results for worry index scores, defined as the standardized worries about various college aspects score. Column 4 reports estimates for a composite Affirmative Action Support Index constructed by mean of dependent variables in columns 5 and 6. Column 5 reports results for metritocratic concerns of affirmative action policies, and are reverse coded for easier interpretation. Column 6 reports results for social justice aspect of affirmative action. In all the columns, higher values of these indices indicate worse mental health outcomes. All regressions include hostel-by-cohort and specialization fixed effects and control for the number of roommates. Standard errors are clustered at the room level. All the regressions include baseline controls for the respective dependent variables and a missing indicator for the baseline controls and number of roommates.

A Proofs for Section 6

Throughout, $\Phi_{ij}(B) = v(B) + \pi_{ij}(1-B) - c(1-B;\,\Delta_i)$, $B(s) = \iota_c + \tau_c\, s$, $\tau_c = \lambda_M - \iota_c$. Primes on v, c denote derivatives in B, R respectively.

Proof of Proposition 1. Take expectations of (5) conditional on room type; the noise term integrates to zero.

Same-caste baseline. In a same-caste room, $\lambda_{ij} = 1$ and $d_{ij} = 0$. Two students drawn from the same caste distribution satisfy $\Pr(\theta_j > \theta_i \mid \text{both caste } c) = 1/2$ by symmetry, so

$$\mathbb{E}\big[(\theta_j - \theta_i)_+ \mid \text{both } c\big] = \tfrac{1}{2}\,\delta_c,$$

and $\mathbb{E}[g_i \mid \text{same-caste}] = \theta_i + \beta \cdot \frac{1}{2} \cdot \delta_c$.

Mixed room. For a lower-caste student i with an upper-caste roommate j, the event $\{\theta_j > \theta_i\}$ has probability p (defined in (3)) and conditional gap δ:

$$\mathbb{E}\big[(\theta_j - \theta_i)_+ \mid i \text{ lower-caste, mixed}\big] = p\,\delta.$$

Hence $\mathbb{E}[g_i \mid \text{mixed}] = \theta_i + \beta\,\lambda_{ij}\,p\,\delta - \mu_L\,d_{ij}$. The contrast against the same-caste baseline is

$$\mathbb{E}[g_i \mid \text{mixed}] - \mathbb{E}[g_i \mid \text{same-caste}] = \beta\big[\lambda_{ij}\,p\,\delta - \delta_L/2\big] - \mu_L\,d_{ij},$$

giving (i). For an upper-caste student i, $(\theta_j - \theta_i)_+ > 0$ only on the reversal event, probability $1-p$, conditional gap δ':

$$\mathbb{E}[g_i \mid \text{mixed}] - \mathbb{E}[g_i \mid \text{same-caste}] = \beta\big[\lambda_{ij}\,(1-p)\,\delta' - \delta_U/2\big] - \mu_U\,d_{ij},$$

giving (ii). The upper-caste bracket is unambiguously negative: the new cross-caste transmission $(1-p)\delta'$ is small (weighted by the reversal probability) and discounted by $\lambda_{ij} < 1$, while the forgone within-caste transmission $\delta_U/2$ is subtracted in full.

Empirical magnitudes. The between-caste gap $\Delta\theta$ is slightly greater than one within-type standard deviation of the entrance-score distribution, so $p \in [0.85, 0.90]$ (see Table [*]). Reversals occur in the overlap region where scores are close, implying $\delta' \ll \delta$ and, a fortiori, $(1-p)\delta' \ll p\,\delta$. Within-caste upward gaps δ_c are bounded above by the within-type standard deviation, which is substantially smaller than the cross-caste gap δ; accordingly $\delta_L/2$ is small relative to $\lambda_{ij}\,p\,\delta$, and the lower-caste bracket is positive well before the distance cost is subtracted. The composites $\beta[\lambda_{ij}\,p\,\delta - \delta_L/2]$ and $\beta[\lambda_{ij}(1-p)\delta' - \delta_U/2]$ are what the reduced-form grade regressions estimate; the individual objects δ, δ', δ_c are not

separately identified from β and λ_{ij}. □

Proof of Proposition 2. (i) With $r_i = \theta_i$, standing is fixed by the entrance score and does not depend on room type. Differencing (7) between $d_i > 0$ and $d_i = 0$, the ego wedge $-\varphi\kappa\, v(r_i)$ cancels. By (8) and $\ell(0) = 0$, $\psi_i^B\big|_{d_i=0} = \psi(b_i)$; the difference of posteriors is $\ell(d_i)(z_i - \psi(b_i))$, giving (9). Since $\psi'(b_i) < 0$, larger b_i raises $z_i - \psi(b_i)$; since $\ell' > 0$, the contrast rises in d_i.

(ii) Now let r_i depend on room type through g_i. From (7), $\alpha_i = \alpha_0 + \varphi\,\psi_i^B - \varphi\kappa\, v(r_i) + \zeta_i$. Differencing across room types:

$$\Delta\alpha_i = \varphi\big[\psi_i^B\big|_{\text{mixed}} - \psi_i^B\big|_{\text{same}}\big] - \varphi\kappa\big[v(r_i\big|_{\text{mixed}}) - v(r_i\big|_{\text{same}})\big].$$

The first term is $\varphi\,\ell(d_i)(z_i - \psi(b_i))$ as in part (i). For the second, $v' > 0$, so $\operatorname{sign}[v(r_i|_{\text{mixed}}) - v(r_i|_{\text{same}})] = \operatorname{sign}[r_i|_{\text{mixed}} - r_i|_{\text{same}}]$. If r_i includes g_i, then by Proposition 1(i) $r_i|_{\text{mixed}} > r_i|_{\text{same}}$ for lower-caste students (grades rise), making the residual $-\varphi\kappa[\cdot] < 0$: the correction is attenuated. By Proposition 1(ii) the inequality reverses for upper-caste students (grades fall), making $-\varphi\kappa[\cdot] > 0$: the correction is amplified. □

Proof of Proposition 3. Fix the correction size $z_i - \psi(b_i)$ and caste distance d_i. From (12), $\Delta\alpha_i^k = \varphi^k\,\ell^k(d_i, c_i)(z_i - \psi(b_i))$, where $\ell^k(d_i, c_i) = \tau(d_i)/(\tau(d_i) + \tau_0^k(c_i))$ is strictly decreasing in $\tau_0^k(c_i)$: $\partial\ell^k/\partial\tau_0^k = -\tau(d_i)/(\tau(d_i) + \tau_0^k)^2 < 0$.

(i) $\tau_0^{ben}(L) > \tau_0^{ben}(U)$ implies $\ell^{ben}(d_i, L) < \ell^{ben}(d_i, U)$. Since $\varphi^{ben} > 0$ and the correction size is common, $\Delta\alpha_U^{ben} > \Delta\alpha_L^{ben}$.

(ii) $\tau_0^{fair}(U) > \tau_0^{fair}(L)$ implies $\ell^{fair}(d_i, U) < \ell^{fair}(d_i, L)$. Hence $\Delta\alpha_L^{fair} > \Delta\alpha_U^{fair}$.

(iii) Both corrections are positive when $z_i - \psi(b_i) > 0$ (both groups revise ψ upward). Define the dimension-specific between-caste gap $G^k \equiv \alpha_U^k - \alpha_L^k$. The change induced by mixing is

$$\Delta G^k = \big(\Delta\alpha_U^k - \Delta\alpha_L^k\big) = \varphi^k\big[\ell^k(d, U) - \ell^k(d, L)\big]\big(z_i - \psi(b_i)\big).$$

For benefits: $\ell^{ben}(d, U) > \ell^{ben}(d, L)$ by (i), so the bracket is positive and $\Delta G^{ben} > 0$—the benefits gap closes.

For fairness: $\ell^{fair}(d, L) > \ell^{fair}(d, U)$ by (ii), so the bracket is negative and $\Delta G^{fair} < 0$—the fairness gap widens.

The opposite signs $\Delta G^{ben} > 0 > \Delta G^{fair}$ are a joint implication of the prior-precision ordering. An exposure-only specification, which delivers a common shift $\Delta\alpha$ across dimensions, predicts ΔG^{ben} and ΔG^{fair} of the same sign; the opposite-sign pattern therefore distinguishes the correction model from a pure dose-response.

The net effect on the overall gap, $\Delta G = \Delta G^{ben} + \Delta G^{fair}$, is unsigned: it depends on $\varphi^{ben}[\ell^{ben}(d,U) - \ell^{ben}(d,L)]$ versus $\varphi^{fair}[\ell^{fair}(d,L) - \ell^{fair}(d,U)]$, which the model does not rank.

For the attenuation prediction: a student with prior cross-caste contact has observed A from the other side, raising the precision of the dimension that was otherwise low. This narrows the gap $\tau_0^k(U) - \tau_0^k(L)$, compressing both $\ell^{ben}(d,U) - \ell^{ben}(d,L)$ and $\ell^{fair}(d,L) - \ell^{fair}(d,U)$ toward zero, which attenuates the dimensional split. □

Proof of Corollary 1. The numerator precision $\tau(d_i)$ in (11) is common to both dimensions and increasing in d by assumption. Since $\mu_c d$ is also increasing in d, students with larger $\mu_c d$ have larger $\tau(d_i)$ and hence larger $\ell^k(d_i, c_i)$ for each k, so $|\Delta\alpha_i^k|$ is increasing in $\mu_c d$. The argument uses only the monotonicity of $\tau(d)$ and of $g(\tau) = \tau/(\tau + \tau_0^k)$ in τ, independent of the dimensional split. □

Proof of Lemma 1. $B(s)$ is affine and monotone in s, so choosing s is equivalent to choosing B on $[\min\{\lambda_{ij}, \iota_c\}, \bar{B}_c]$. Substituting $R = 1 - B$ into (13) and dropping the (choice-irrelevant) grade term gives (14). $\Phi''_{ij}(B) = v''(B) - c''(1 - B;\, \Delta_i) < 0$, so Φ_{ij} is strictly concave; the linear friction term preserves concavity, so the maximizer is the interior $B^\dagger$ where feasible, else the nearer boundary, with $B^* = \bar{B}_c$ exactly when $\Phi'_{ij}(\bar{B}_c) > 0$. □

Proof of Proposition 4. (i) At $B = 1$ ($R = 0$), $\Phi'_{ij}(1) = v'(1) - \pi_{ij}$; concavity implies $B^\dagger < 1$ iff $\pi_{ij} > v'(1)$. Implicit differentiation of $\Phi'_{ij}(B^\dagger;\, \Delta, \pi_{ij}) = 0$ gives $\partial B^\dagger/\partial\Delta = -c_{R\Delta}/\Phi''_{ij} \geq 0$ and $\partial B^\dagger/\partial\pi_{ij} = 1/\Phi''_{ij} < 0$.

(ii) $\bar{B}_c \geq B^\dagger$ iff $\Phi'_{ij}(\bar{B}_c) \leq 0$ iff $\pi_{ij} \geq \underline{\pi}_c(\Delta)$. Since $\bar{B}_L < \bar{B}_U$, strict concavity of v gives $v'(\bar{B}_L) > v'(\bar{B}_U)$, and $c'(1 - \bar{B}_L;\, \Delta) > 0 > c'(1 - \bar{B}_U;\, 0) = 0$, so $\underline{\pi}_L(\Delta) > \underline{\pi}_U$, widening in Δ.

(iii) Follows from (i)–(ii). □

Interpretation of Proposition 4

The unconstrained optimum $B^\dagger$ is pinned by one equation: the student wants belonging up to the point where the marginal value of one more unit,

$$\underbrace{v'(B)}_{\substack{\text{direct utility}\\ \text{of belonging}}} + \underbrace{c'(1 - B;\, \Delta_i)}_{\substack{\text{bridging cost}\\ \text{saved by retreating}}} = \underbrace{\pi_{ij}}_{\substack{\text{placement value}\\ \text{of bridging forgone}}},$$

balances the placement return to bridging. The left side is the pull toward belonging: its direct utility plus the psychological cost of bridging one would avoid. The right side is the pull toward bridging: what the roommate's network is worth on the labor market. Both pulls vary with the assignment.

Part (i) traces how $B^\dagger$ responds. A wider preparation gap Δ_i raises the marginal cost of bridging ($c_{R\Delta} \geq 0$), pulling the student toward more belonging: $\partial B^\dagger / \partial \Delta_i \geq 0$. A more advantaged roommate raises the marginal return to bridging, pulling the student toward less belonging: $\partial B^\dagger / \partial \pi_{ij} < 0$. For a lower-caste student in a mixed room, these two forces work against each other—the roommate is simultaneously high-Δ (bridging is painful) and high-π (bridging pays)—and $B^\dagger$ settles wherever the two pulls balance.

The student is bonding-constrained when $\bar{B}_c < B^\dagger$: no feasible allocation delivers the belonging they want. What determines the ceiling $\bar{B}_c$ is the quality of the outside option. For an arena-capable upper-caste student, $\bar{B}_U = \max\{\lambda_M, \iota_U\} = 1$: the arena is perfectly aligned ($\iota_U = 1$), so full belonging is always attainable by spending all social time there. They are never bonding-constrained in a mixed room. For an arena-capable lower-caste student, $\bar{B}_L = \max\{\lambda_M, \iota_L\} < 1$: neither the cross-caste roommate nor the imperfectly aligned arena delivers full belonging, so the ceiling binds whenever $B^\dagger$ exceeds it. For an arena-excluded student of either caste, the ceiling collapses to $\bar{B}_i = \lambda_M < 1$ regardless of ι_{c_i}, because the arena is simply unavailable.

Part (ii) defines the threshold $\underline{\pi}_c(\Delta) = v'(\bar{B}_c) + c'(1 - \bar{B}_c; \Delta)$: the marginal value of belonging evaluated at the ceiling. A student is unconstrained iff the placement return exceeds this threshold—the return to bridging is high enough that they voluntarily choose an interior allocation below the ceiling. The ordering $\underline{\pi}_U < \underline{\pi}_L(\Delta)$ follows from two compounding forces: the upper-caste ceiling is higher ($\bar{B}_U > \bar{B}_L$), so the marginal utility of belonging is lower there by concavity ($v'(\bar{B}_U) < v'(\bar{B}_L)$); and the upper-caste student holds zero bridging at their ceiling ($R = 0$, so $c'(0; 0) = 0$), while the lower-caste student holds positive bridging even at their ceiling ($c'(1 - \bar{B}_L; \Delta) > 0$). Both forces raise the lower-caste threshold: it takes a much more valuable roommate to pull a lower-caste student into voluntary bridging.

Part (iii) follows immediately. Three regimes, indexed by how valuable the roommate's network is:

- *Low π_{ij} ($\leq v'(1)$):* the roommate's network is not worth sacrificing belonging for. Even the upper-caste student, with the best outside option, wants full belonging. Everyone is constrained.

- *High π_{ij} ($\geq \underline{\pi}_L(\Delta)$):* the network is so valuable that even the lower-caste student, with the worst outside option and the highest bridging cost, voluntarily bridges. No one is constrained.

- *Intermediate π_{ij}:* the return clears the low upper-caste threshold but not the high lower-caste one. The upper-caste student voluntarily bridges; the lower-caste stu-

dent is stuck at the ceiling wanting more belonging than any allocation can deliver. This is the regime where the asymmetry bites, and it is the empirically relevant one.

Proof of Corollary 2. $\bar{B}_i$ depends on caste only through ι_{c_i} and on the student only through A_i; the target $B^\dagger$ depends on Δ_i, π_{ij}. An arena-excluded student has $a_i = 0$ (Assumption 4), so $B_i = \lambda_M < 1$ regardless of ι_{c_i}, constrained whenever $B^\dagger > \lambda_M$. By continuity and monotonicity of $B^\dagger(\Delta)$ (part (i) above), $\{\Delta : B^\dagger(\Delta) \leq \bar{B}_L\}$ is an interval $[0, \Delta_{\max}]$. Bunching follows since $B(s)$ is monotone, so the ceiling binds at $s = 1$ ($\tau_c > 0$) or $s = 0$ ($\tau_c < 0$). □

Proof of Proposition 5. Wellbeing. In a same-caste room $\lambda_{ij} = 1$ and $R = 0$ is feasible, so both bridging cost and friction can be zero; in a mixed room $\bar{B}_c < 1$ forces $R > 0$, hence $c(R;\, \Delta) > 0$ and any $s^* > 0$ carries friction: $S^*(\lambda_M) < S^*(1)$. The shortfall $\Phi_{ij}(B^\dagger) - \Phi_{ij}(\bar{B}_c)$ is positive only when constrained and, by strict concavity of Φ_{ij} around a common $B^\dagger$, weakly larger the lower the ceiling—hence larger for the lower-caste student ($\bar{B}_L < \bar{B}_U$). It is decreasing in $\Phi'_{ij}(\iota_c)$, which rises in arena alignment and access.

Divergence. By Proposition 1 the mixed-room grade change for a lower-caste student is positive under the stated condition. In a same-caste room $\bar{B} = 1$ and $\chi_c(0) = 0$, so the student achieves $B^* = \min\{B^\dagger, 1\}$ with $\Delta = 0$. In a mixed room with $\bar{B}_L < B^\dagger$, the student is constrained at $B^* = \bar{B}_L < \min\{B^\dagger, 1\}$: $v(B)$ falls, and the bridging cost rises from $c(1 - \min\{B^\dagger, 1\};\, 0)$ to $c(1 - \bar{B}_L;\, \Delta)$ with $\Delta > 0$, a positive change since c is increasing in both arguments. Wellbeing falls by the sum plus any friction; the grade and wellbeing changes therefore have opposite signs. □

Proof of Corollary 3. $\partial S_i / \partial \theta_j = -(\partial c / \partial \Delta_i)(\partial \Delta_i / \partial \theta_j)$. For a lower-caste student, $\Delta_i = \theta_j - \theta_i$ so $\partial \Delta_i / \partial \theta_j = 1$ and the derivative is $-\partial c / \partial \Delta_i < 0$, larger in magnitude at low θ_i by convexity of c in Δ. For an upper-caste student $\Delta_i \approx 0$ and $\partial \Delta_i / \partial \theta_j = 0$, so the derivative vanishes. □

Proof of Corollary 4. The grade effect is Proposition 1(i). From (4), $R_i = (1 - \lambda_{ij}) s_i + (1 - \iota_{c_i}) a_i$; for a bonding-constrained student s_i is pinned at the boundary and does not respond to θ_j, so $\partial R_i / \partial \theta_j = 0$ and placement moves only through the grade channel: $\partial Y_i / \partial \theta_j = \pi_g \, \partial g_i / \partial \theta_j$. The worry effect is Corollary 3. □

Proof of Corollary 5. $\pi_{ij} = \pi(A_j)$ is increasing in A_j; $\underline{\pi}_c(\Delta_i)$ does not depend on A_j. The unconstrained set $\{A_j : \pi_{ij} \geq \underline{\pi}_c(\Delta_i)\}$ is therefore an upper interval, and by Proposition 4(i) the shortfall $\Phi_{ij}(B^\dagger) - \Phi_{ij}(\bar{B}_c)$ is decreasing in A_j wherever positive, so the wellbeing loss attenuates; placement rises directly since $Y_i = \pi(A_j)\, R_i + \pi_g \, g_i$ with $\pi' > 0$. □

Proof of Proposition 7. (i) Immediate from the functional form of P_i: increasing in ι_{c_i} (larger for upper caste, Assumption 2) and in θ through attainment; A_i enters only via the arena-access gate, not the production function.

(ii) Log-differentiating gives the stated expression; by Lemma 1 an upper-caste student facing friction shifts from roommate time toward the arena ($\Delta \log s < 0$, $\Delta \log a > 0$), so the expression is linear in η with a unique zero $\bar{\eta} \in (0, 1)$.

(iii) Lower-caste students are gated by $\iota_L < \iota_U$ and, when constrained, hold $R^* = 1 - \bar{B}_L$ above the unconstrained level (raising placement, not positions); flatness in θ_j follows from $\partial R_i / \partial \theta_j = 0$ (Corollary 4). □

Proof of Proposition 8. Write $V_c(\lambda) = \Phi_{ij}(B(s^*(\lambda);\, \lambda)) - \chi_c(d)\, s^*(\lambda)$, $B(s;\, \lambda) = \iota_c + (\lambda - \iota_c)\, s$. Differentiating and using the first-order condition (which zeroes the term multiplying $ds^*/d\lambda$ at an interior optimum, and $ds^*/d\lambda = 0$ at the $s^* = 0$ corner) leaves $dV_c/d\lambda = \Phi'_{ij}(B^*)\, s^*_c$. Substituting $\Phi'_{ij}(B^*) = \chi_c(d)/\tau_c$ (valid at an interior optimum) gives (15). At $s^* = 1$ the first-order condition holds only as $\tau_c\, \Phi'_{ij}(B) \geq \chi_c(d)$ ($\tau_c > 0$), so the true derivative $\Phi'_{ij}(B)\, s^*$ weakly exceeds the formula's $\chi_c(d)/\tau_c$ pointwise, and the integrated welfare change is weakly more negative than (15) reports. □

Proof of Proposition 9. From (13), $\Delta U_i = \Delta S_i + \pi_{ij}\, \Delta R_i + \pi_g\, \Delta g_i$ is additively separable. A student leaving a mixed room reveals $\Delta U_i < 0$ (rearranging to the stated inequality); a stayer reveals $\Delta U_i \geq 0$; an indifferent marginal student has $\Delta U_i = 0$. Since π_g, π_{ij} are estimable from the assignment and ΔR_i, Δg_i observed, the wellbeing term is recovered without further structure. □

Proof of Theorem 1. $\Delta W = \Delta U + \omega_W\, \Delta \alpha$ per student under the stated W. Setting $\Delta W \geq 0$ and dividing by $\Delta \alpha > 0$ gives $\omega_W \geq -\Delta U / \Delta \alpha$. For the two-versus-three-person comparison, preferring configuration (2) requires $\Delta U^{(2)} + \omega_W\, \Delta \alpha^{(2)} \geq \Delta U^{(3)} + \omega_W\, \Delta \alpha^{(3)}$, rearranging to the stated threshold (dividing by the positive quantity $\Delta \alpha^{(2)} - \Delta \alpha^{(3)}$). □