Interaction, Stereotypes and Performance. Evidence from South Africa

Lucia Corno, Eliana La Ferrara, Justine Burns*

This version: November 2018

Abstract

We exploit a policy designed to randomly allocate roommates in a large South African university to investigate whether inter-racial interaction affects stereotypes, attitudes and performance. Using Implicit Association Tests, we find that living with a roommate of a different race reduces white students' negative stereotypes towards blacks and increases inter-racial friendships. Interaction also affects academic outcomes: blacks in mixed rooms improve their GPA, pass more exams and have lower dropout rates. This positive effect is not driven by the ability of the roommate and is stronger the lower the roommate's prejudice, suggesting a complementarity between stereotype reduction and performance gains.

*We thank Nicola Gennaioli, Caroline Hoxby, Michael Kremer, Monica Martinez-Bravo, Rohini Pande, Enrico Rettore, Piero Veronese and seminar participants at NBER Political Economy, CIFAR at LSE, IEA World Congress, CEPR Public Policy, EUDN, Boston University, Tufts, University of Chicago, University of Turin, SciencesPo, INSEAD, Oxford, Universidade Nova de Lisboa, Pompeu Fabra, EUI, University of Essex, IRVAPP, Stockholm University, University of Cape Town, CERGE-EI and Hebrew University for helpful comments. We are grateful to Selene Ghisolfi for invaluable help with data collection and analysis. Nicola Fontana and Martina Zanella provided excellent research assistance. La Ferrara gratefully acknowledges financial support from the ERC Advanced Grant ASNODEV. The usual disclaimer applies. Corno: Cattolica University, LEAP and IFS, lucia.corno@unicatt.it; La Ferrara: Bocconi University, LEAP and IGIER, eliana.laferrara@unibocconi.it; Burns: University of Cape Town, justine.burns@uct.ac.za

1 Introduction

Contemporary societies are becoming increasingly diverse along ethnic lines and this trend has brought to the forefront of the public debate the significance of inter-group prejudice and stereotypes. While stereotypes serve a cognitive role in allowing for faster processing of information, they often induce distortions in social behavior, e.g., leading to over-simplified judgement and discrimination. The consequences for groups subject to negative stereotypes are unequal access to economic, social and political opportunities, as well as self-confirming 'aspiration traps'. Understanding if, and how, such stereotypes can be changed is therefore crucial to improve our ability to formulate effective policies in diverse societies.

This paper tests whether a policy designed to exogenously generate exposure to members of different racial groups induces changes in attitudes and stereotypes, and whether this translates into sizeable performance gains for the individuals involved. We study this in the context of South Africa, a country where the experience of apartheid made people relatively prone to stereotyping and led to the economic marginalization of black South Africans. We take advantage of a policy implemented by the University of Cape Town (UCT) with the aim of promoting racial integration. This policy randomly allocates students across university residences and –in some of the residences– to roommates, thus providing a unique opportunity to estimate the causal effect of a roommate's race on individual attitudes and behavior. We address two sets of questions: (i) Does interaction with someone of a different race change individual stereotypes towards that race? Does this occur through belief updating (that is, in response to the characteristics of the roommate) or through mere exposure (e.g., because of increased empathy or anxiety reduction)? (ii) What are the effects of inter-group contact on academic outcomes? Are they mediated by prejudice, e.g., do individuals learn more when they are paired with someone of another race who is not prejudiced against their group? To the best of our knowledge, ours is the first study to assess this potential complementarity and to contextually estimate the effect of inter-group contact on stereotypes and performance.

Whether increased interaction with members of other groups would increase or decrease stereotypes is theoretically unclear. Allport's (1954) 'contact hypothesis' maintained that, under certain conditions, exposure to outgroup members would allow individuals to better understand their char-

acteristics and points of view, thus diminishing reliance on stereotypes and eventually ameliorating inter-group relations.[1] On the other hand, a recent literature on the 'negative contact hypothesis' contends that negative contact makes categories more salient than positive contact, potentially leading to an increase –as opposed to a reduction– in negative outgroup stereotyping (Paolini et al., 2010; Barlow et al., 2012).[2]

To address the above questions, we collected data on students living in UCT residences who were exposed to the policy of random roommate allocation. Our sample includes 495 freshmen living in double rooms, whom we interviewed at the beginning and at the end of the 2012 academic year. Our first outcome of interest is prejudice or stereotype held against members of different racial groups. To gain a possibly objective measure, we administered a series of Implicit Association Tests (IATs). The IAT was first introduced by Greenwald and Banaji (1995) and is a tool used by social psychologists. It exploits variation in the time that individuals take to complete a rapid categorization task that involves associating concepts with visual cues about race. The underlying idea is that subjects who are systematically slower in associating certain pairs implicitly reveal mental processes that tend to perceive those pairs as less common. In addition to the typical "Population IAT", which elicits associations between generally "positive" concepts and race, we designed an IAT to elicit associations between academic ability and race –we refer to this as the "Academic IAT". The advantage of IATs over self-reported measures of prejudice is significant, especially in contexts where subjects may be reluctant to disclose prejudice or may not be fully aware of it. To our knowledge, ours is the first paper that uses IATs to estimate the impact of a policy to promote integration on racial stereotypes.[3]

Our first main result shows that exposure to members of a different race led to significant changes in stereotypes: whites became relatively less prejudiced against blacks. The effect is sizeable, corresponding to .58 of a standard deviation of the Population IAT. The magnitude

[1] The conditions that need to be satisfied in Allport's view include the fact that groups enjoy equal status in the relationship, share common goals, benefit from cooperation and that interaction is supported by a commonly recognized authority. All of these conditions can be said to apply in the university setting we study.

[2] For a theoretical model of stereotype formation, see Bordalo et al. (2016).

[3] Recent work by Lowes et al. (2015) uses the IAT to provide descriptive evidence on implicit attitudes towards different ethnic groups in the Democratic Republic of Congo. Barnhardt (2009) exploits a natural experiment with public housing in India to study the effects of geographic proximity on religious prejudice using an IAT for the categories of Hindu and Muslim.

of the estimated effect suggests that the treatment closes 84 percent of the gap in stereotypes between whites and blacks found in the Population IAT. We do not find corresponding changes in the Academic IAT, on average, suggesting that interaction *per se* does not necessarily induce an update in beliefs on ability in a direction that is favorable to black South Africans.

To uncover the role played by information in stereotype reduction, we test if the effect is stronger for individuals whose roommate has more positive characteristics (e.g., academic ability, altruism) than the individual's own prior regarding the other group. We do not find differential effects on the Population IAT score, suggesting that belief updating may not be the driving force and that empathy may play a significant role (Pettigrew and Tropp, 2008). On the other hand, the Academic IAT strongly responds to the realization of roommate's type: white students who were assigned non-white roommates whose ability is a 'positive surprise' significantly revise their priors upward. This is not surprising, as we expect ability stereotypes to be more responsive to performance than stereotypes on generic good or bad characteristics.

Our second set of results concerns academic performance. We find significant effects of inter-group contact on academic achievement, heterogeneous across groups. Grade point average (GPA) scores improve by .26 standard deviations for black students sharing a room with non-blacks. This closes about 1/3 of the gap in GPA between whites and blacks. Black students in mixed rooms also pass a higher number of exams and are more likely to be eligible to continue to the following year (as opposed to dropping out). The positive effect on academic performance is long lasting: it persists in the second year, when most of the students are no longer in residences. No significant impact on academic performance is found for white students in mixed rooms.

What are the channels through which this performance effect operates? First of all, the effect is not due to the fact that a black exposed to a white is on average exposed to a higher ability peer: in all regressions we control for the roommate's admission score into UCT as a proxy for ability, and the latter is typically insignificant. Interestingly, we find that the positive effect on performance for black students is stronger if they are paired with less-prejudiced roommates: black students in mixed rooms improved their GPA by more when their roommate was (at baseline) less prejudiced against blacks. This could be due to a more productive academic interaction, a reduction in inter-

group anxiety experienced by blacks, and possibly a lower perceived risk of confirming negative stereotypes, as highlighted in the 'stereotype threat' literature (Steele and Aronson, 1995). While previous work has examined the effect of random roommate allocation on academic performance, to our knowledge ours is the first attempt to study how this effect varies with the stereotypes held by the subjects involved. If academic gains from interaction are to be realized, it seems plausible that stereotypes would play a significant role in making the interaction viable and beneficial for the parties involved.

To further explore the mechanisms of interaction, we examine the effects of a roommate's race on a variety of attitudinal and behavioral measures. We find that exposure to a roommate of a different race increases inter-racial interactions outside the room: students in mixed rooms report that they hang out more often with people of a different race, and their desired and actual number of friends and study-mates of a different race is higher than that of students in same-race rooms. Self-reported attitudes also improve, as students in mixed rooms report talking more frequently about race and feeling more comfortable about it. They also feel less self-conscious dancing with or dating a person of a different race. We find weaker evidence that students paired with roommates from other groups exhibit more generalized pro-social behavior, as measured by participation in voluntary organizations and social services, and cooperative behavior in a prisoner dilemma game.

Research in economics has widely studied the effects of ethnic diversity and racial segregation. In this literature, integration policies are often proposed as a means of reducing racial gaps in outcomes, reducing the costs of ethnic divisions and leveraging the potential benefits of diversity.[4] Social psychologists, on the other hand, have studied diversity and underlined the importance of identity and stereotype formation. In this literature, the main role served by integration is not that of improving (economic) performance, but of changing individual attitudes and stereotypes, possibly reducing prejudice and inter-group conflict.[5] Our paper is an attempt to bring together these two sides of the problem to show that attitudinal change and performance gains may go

[4]For a review of the literature on ethnic diversity, see Alesina and La Ferrara (2005). A recent study of the economic costs of ethnic divisions is Hjort (2014), while a challenge to the assumption of coethnic bias is presented by Berge et al. (2015). On diversity-enhancing policies, see, among others, Fryer and Loury (2013) for a theoretical analysis and Kling et al. (2007) for an evaluation of the Moving to Opportunity program.

[5]For a meta-analysis of the relation between inter-group contact and prejudice, see Pettigrew and Tropp (2006) and Paluck, Green and Green (2018).

hand-in-hand.

Our work relates to three strands of literature. The first is the literature on the effects of integration policies on inter-group attitudes. Boisjoly et al. (2006) find that randomly matching roommates of different races in a US university increases support for affirmative action and empathy towards other groups. Van Laar et al. (2005) use housing assignments of first-year college students at University of California and find that having a roommate from another ethnic group decreases prejudice. Carrell et al. (2016) exploit data from the United States Air Force Academy to show that white males randomly assigned to both a higher number, as well as higher-aptitude, black peers in their freshman year are more likely to match with a black roommate in their sophomore year. Scacco and Warren (2018) randomly assign Christian and Muslim young men to computer training groups in Nigeria and find that inter-group contact did not affect prejudice but reduced discriminatory behavior. These papers use self-reported attitudes, different from our work that relies on Implicit Association Tests. Glover, Pallais and Pariente (2015) examine the performance of cashiers in a French grocery store chain and find that minorities perfom more poorly when they work with managers biased against them (as determined by Implicit Association Tests). Our setting is different from the latter as it is one where interaction occurs among peers, as opposed to people at different levels of the hierarchy, and according to Allport's (1954) hypothesis this should be key for realizing the gains from contact. Also, we look at a different set of outcomes than Glover et al. (2015), including stereotypes, social interaction and academic performance.

Recent papers by Barnhardt (2009) and Rao (forthcoming) study different forms of integration in India. Barnhardt (2009) examines the effects of neighborhood religious composition on inter-religious attitudes, while Rao (2015) studies the impact of changes in the wealth composition of children's classmates on pro-social behavior and test scores in Delhi's private schools. With respect to these authors, we focus on a different dimension (i.e., race as opposed to religion or social class) and we also study the interplay between prejudice reduction and educational outcomes.

A second body of literature brings the notion of identity to the forefront of economic analysis, embedding concepts developed in the social psychology literature, including the seminal work by Tajfel et al. (1971). Contributions include, among others, Akerlof and Kranton (2000), Hoff and

Pandey (2006) and Shayo (2009). While we do not directly elicit notions of self-identification, some of our results on the salience of race and on revised beliefs regarding out-group members speak to the issue of identity formation.

Finally, our paper contributes to the vast literature on the effect of peers on human capital formation. Most of this literature studies the effect of peers' ability and academic performance (Sacerdote, 2001; Lyle, 2009; Garlick, 2013) and derives implications for policies such as tracking (Duflo et al., 2011; Carrell, Sacerdote and West, 2013). While we do not estimate 'endogenous' peer effects, our paper aims at shedding light on the interplay between peers' ability and racial stereotypes. Our results suggest that stereotypes may act as barriers to interaction and communication among peers, so that the effect of exposure to another student with a certain ability will differ depending on the stereotypes held by that student.

The remainder of the paper is organized as follows. Section 2 provides some institutional background and describes the data we collected. Section 3 shows some descriptive statistics and discusses the identifying assumption underlying our work. In section 4 we present our empirical strategy. Section 5 contains the econometric results and section 6 concludes.

2 Background and data collection

2.1 Institutional setting

The University of Cape Town (UCT) is a public research university located in Cape Town. It is the oldest and most prestigious university in South Africa and it enrols approximately 5000 incoming freshmen every year, more than half of whom live in university residences. Incoming students were historically allocated into dormitories to live with students whose academic performance in standardized high school graduation tests was similar to their own. This tracking regime was replaced in 2006 with a policy of randomly assigning incoming students to dormitories.[6]

Students submit applications to the university between July and October to start studying in January of the following year. UCT's admission policy is mainly based on a measure called Admission Points Score, computed from the high school grades in the final school year, but it is

[6] Garlick (2013) compares the two policies in terms of GPA distributions using a difference-in-differences strategy.

also designed in order to build a student body that reflects the demographics of South African society. In 2012 the incoming freshmen class was composed of 36 percent blacks, 16 percent Coloureds, 40 percent whites, and 8 percent Indians, Asians or other race.[7]

In the application form, students may request to live in university residences. Only students living outside the Cape Town area can apply for accommodation. Exceptions are made for disadvantaged students or for those with great academic merit. The policy and criteria for admission to UCT student housing assume that a first year student will enter a first-tier (catering) residence and in subsequent years move to second- or third-tier (senior catering or self-catering) accommodation. While second year students may express a preference for the residence they wish to be assigned to, freshmen assignment to residences relies on a random allocation system.

Once first year students are randomly allocated to one of the fifteen residences, they are assigned a room which can be either single or double occupancy. All rooms are single gender. The criteria for allocation to specific rooms within the residence are decided by the warden and may vary by residence.[8] Our analysis focuses exclusively on eight residences with double rooms in which during the year of our study the wardens stated that they implemented a random allocation mechanism, conditional on gender, and where white and black students co-exist.[9]

Approximately one week before the beginning of the academic year, each residence organizes an open day with first year students to explain the rules and features of the residence. During the open day, each student is assigned to a room. In some residences, the random assignment takes the form of extracting the room number from an urn: if the room extracted is a single room, the number is removed from the urn; if it is a double room, it is placed back in the urn so that a roommate may extract it again. In other residences the wardens randomly select students and their roommates from the list of students enrolled in the residence.[10]

[7] In South Africa, the term 'Coloured' denotes people of mixed ethnic origin who may have diverse ancestries, including European, Asian and local Khosa and Bantu ethnic groups.

[8] Therefore the random allocation of students across dormitories is a policy at the university level, while the random allocation of roommates is a discretionary policy at the dormitory level.

[9] We originally collected data for ten residences that told us they were implementing a random room allocation scheme, but two of the residences had a very small number of double rooms and did not have any white student in double rooms. We drop these residences, which accounted for a total of 9 observations in the original sample. Our results are unchanged if we keep them in the analysis.

[10] It is possible that wardens may 'adjust' the composition of some rooms, e.g., to ensure that each floor or wing has a certain composition. While this was not described to us as a standard procedure, we cannot rule it out. We will however provide evidence that such exceptions, if they occurred, did not lead to significant deviations from a

Approximately 50 percent of undergraduate students living on campus are in shared rooms in their first year in residence. The Residence Management Services is in charge of residence applications; it records, for each student, his/her room number and the dates in which he/she moved in or out of the residence. Rooms are never reserved irrevocably and may be switched. First year students may also decide to swap residences. In our sample 20 percent of the students interviewed at follow-up declare that they changed roommate since the beginning of the year. In all our analysis we will use the initial room assignment, thus reporting "intention to treat" estimates.[11]

2.2 Sample

Our working sample includes one cohort of students: 495 freshmen who joined UCT in 2012 and who were assigned to double rooms in any of the eight residences that reported randomly allocating rooms. We conducted two rounds of data collection: a baseline in February 2012, during the first week of the academic year, and a follow-up survey in September 2012, near the end of the academic year and just before students took their final exams. As part of the data collection, we conducted a series of Implicit Association Tests (IATs) both at baseline and at follow up. During the follow up survey, besides collecting data through questionnaires and IATs, we also conducted lab experiments with the same individuals who took the baseline survey.

Students were recruited to participate in the project through a variety of channels. First, the project was advertised during a residence meeting among wardens and students. Second, posters advertising the project were hung in visible places (e.g., residence halls) about one week before the kick-off. Third, we sent an e-mail to all the students in the participating residences to schedule an appointment for the survey at their most convenient time. In all cases we described the goal of our research as being about "student life at UCT", without mentioning race.

The survey questionnaire, the IATs and the experimental game were conducted in each residence on laptops and under the supervision of two enumerators per residence. We did our best to ensure

random allocation in terms of most observable characteristics.

[11]Note that, although 20 percent of the students change room or roommate, only 2 go from a mixed race room to a non-mixed room, and no one goes from non-mixed to mixed. For this reason, we do not report local average treatment effect (LATE) estimates.

no communication among students during the survey. To try not to contaminate the IATs scores, we conducted them first, followed by the survey questionnaire and then by the experimental game. For their participation in the survey, each student received a monetary incentive of 30 South African Rands, i.e., approximately 3.5 US dollars (corresponding to 1/2 hour of an enumerator's wage).

Our initial sample size for the baseline was 625 freshmen.[12] Of these, 508 were traced successfully for the follow-up and for 495 of them we have non-missing values for IAT tests, with a tracking rate of 79 percent. Appendix Table A1 summarizes the study sample and attrition. The p-value reported in the last column indicates that there is no statistically significant difference in the attrition rate across students allocated to a roommate of a different race (*Mixed Room*) and students allocated to a roommate of the same race (*Non Mixed Room*). In Appendix Table A2 we examine the correlates of the decision to participate in the follow-up round. Again, we find no differential attrition between respondents in mixed and non mixed rooms (column 1). Importantly, attrition does not depend on our measures of stereotypes, the Population IAT (column 2) and the Academic IAT score (column 3), nor on the interaction between these scores and *Mixed Room* (columns 4 and 5) and between IAT score and race dummies (columns 6 and 7). Looking at other controls, it emerges that white, Coloured and richer students (as measured by their consumption) are less likely to participate in the follow-up survey compared to blacks. This is possibly due to the fact that monetary incentives for participation were relatively low, and these groups, on average, come from wealthier families.

2.3 Implicit association tests

Our key outcome of interest is racial stereotypes, as measured by the Implicit Association Test (IAT). The IAT is an experimental method introduced by Greenwald and Banaji (1995) and Greenwald et al. (1998), based on the idea that respondents who more rapidly pair two concepts in a rapid categorization task more strongly associate those concepts (e.g., how fast people pair images of black versus white people with descriptions of leadership roles). Slower speed in associating

[12] The sample size of 504 students mentioned above refers to our final working sample, i.e., students for whom we have both baseline and follow-up survey data with non-missing values for the Academic and Population IAT.

certain pairs denotes mental processes that tend to perceive those pairs as less common. This tool has been widely employed in social psychology to understand implicit cognition, that is, cognitive processes of which an individual may not be aware and that include, among others, perception, stereotyping, and memory. IATs have been correlated to explicit attitudes and behaviors (e.g., Nosek, Banaji and Greewald, 2002) and their use is growing in other disciplines such as neuroscience, marketing research and economics (e.g., Bertrand et al., 2005; Beaman et al., 2009). For our purposes, a particularly useful feature of IATs is that they implicitly reveal attitudes that individuals may be uncomfortable disclosing, such as racial prejudice. Thus, we use IATs to complement subjective and self-reported perceptions of inter-ethnic attitudes with more 'objective' measures of racial bias.

We conducted two types of IATs.[13] The first was a standard test in which tasks involved pairing positive and negative attributes (e.g., "happy", "good", "terrible", "failure") with the racial categories of White South African and Black South African. We refer to this as the Population IAT. The second IAT was less standard and was designed to elicit associations between academic ability and race. We asked people to match pictures of different gender and race with different percentiles of the grade distribution. We denote this as the Academic IAT. The goal of conducting this second IAT was to test whether differential interaction or cooperation with members of the opposite race may reflect priors on how much one can benefit in terms of learning and academic success, based on the beliefs that one holds about the academic performance of the other race. In this sense, the Academic IAT should be seen as a measure of beliefs regarding academic performance rather than a measure of prejudice.

Figure 1 shows the density of the Population IAT score (panel A) and the Academic IAT score (panel B) at baseline, separately for whites and blacks.[14] Lower values of the IAT score indicate more negative stereotypes towards blacks relative to whites.

Two interesting patterns emerge from this figure. First, on average, all groups hold negative stereotypes towards blacks, as the mean of the Population and the Academic IAT is always negative. Second, while white students are more prejudiced against blacks when we consider the Population

[13] Experimental instructions are available from the authors.

[14] The pattern when we pool blacks, Coloureds and other races in a single category is very similar to that of blacks alone.

IAT (Panel A), no significant difference across races emerges when looking at distributions of the Academic IAT (Panel B). This is interesting because it suggests that the Academic IAT is more likely to reflect statistical information on the relative academic performance of the two groups, while the Population IAT may embed more elements of 'taste'.

2.4 Attitudinal and behavioral measures

Through our survey questionnaire, we collected information on students' socioeconomic backgrounds, beliefs, friendships and attitudes towards other ethnic groups. More specifically, we elicited information about the following.

Friendship patterns. We asked: how often the respondent socialized with people of a different race in the past month; the self-reported preferred number of people of a different race in a hypothetical study group or a leisure group formed by 7 people; the share of (actual) best friends who are black or white; and the share of (actual) study-mates who are black or white.

Inter-racial attitudes. We asked about the frequency and comfort in discussing issues of race; agreement with abolishing affirmative action in university admission; feelings of self-consciousness in dancing with a person of another race and in dating a person of another race.

Pro-social behavior. We asked if the respondent was a member of any community service or volunteer organizations and how much money he/she gave to charities (excluding churches) in the last year. For the follow-up survey, we also collected an experimental measure of pro-social behavior running a prisoner dilemma game. We employed the strategy method and revealed the racial identity of participants using photographs.[15] The payoffs were as follows: R50 each if both players chose 'Cooperate'; R40 each if both players chose 'Block'; and if one player chose 'Block' while the other chose 'Cooperate', the former earned R75 while the latter earned R15.

2.5 Academic Performance

To measure students' academic performance we rely on administrative data. First of all, we know the Admission Point Score that the student received based on his/her performance in high school. This is the sum of high school final grades, with weights depending on the specific department the

[15]Note that the photos used in the game do not include any of the subjects in our sample. Experimental instructions are available from the authors.

student enrolls in. We denote this variable as '*UCT admission score*' and we use it as a proxy for students' ability at the beginning of their career at UCT.

We then have several measures of performance at the end of the first and second academic year, collected by the university's registry. These include the total number of exams passed and failed and their grade point average (GPA). In our analysis we employ the average GPA the student obtained, standardized to have mean 0 and standard deviation 1 in each year and we denote this variable as '*GPA*'. In addition to the GPA and the number of exams passed, we use a third indicator based on students' academic evaluation by the Faculty Examination Committees.[16] This indicator, denoted as '*Eligible to continue*', takes value one if the student is in good standing and eligible to continue studying in the next academic year, possibly subject to passing some makeup exams.

3 Descriptive statistics and randomization

Our working sample consists of those students successfully interviewed at baseline and follow-up with non missing values for both IAT tests, which is 495 out of the 508 students who took both rounds of the survey. Out of these 495 students, 154 were allocated to a roommate of a different race and 341 were sharing the room with a student of their own race. The racial composition of this sample is as follows: 329 respondents are black, 116 are white, 18 are Coloured and 32 are Indian, Asian or other race. Notice that this composition mechanically generates differences in the probability of being in a mixed room for different races, with the more numerous group (blacks) having lower probability of being in a mixed room.[17] If we take the shares of the various groups in each residence (including freshmen who did not take our survey but were in double rooms) and simulate random pairings based on 10,000 replications, we obtain an average probability of being in a mixed room equal to 0.51. In our sample the empirical probability is 0.3, significantly lower. This may raise concerns that the wardens deviated from the random allocation policy, which may invalidate our inference. To assess this point we bring a number of pieces of evidence showing that,

[16] The students' evaluation is conducted at the end of the academic year and takes into account the number of exams passed and their grades.

[17] In our sample the fraction living with a roommate of a different race is .22 for blacks, .34 for whites, .72 for Indians, Asians and others and 1 for Coloureds, who all happen to be allocated to non-Coloured roommates.

despite this discrepancy, we believe the allocation can be considered as good as random.

First, Table 1 reports summary statistics at baseline for the main outcomes of interest and the controls for the full sample (columns 1-2) and separately by treatment status, i.e., students in mixed versus non-mixed rooms (columns 3 to 6). The last two columns show the difference in means between non-mixed and mixed rooms and the associated p-value. In Panel A we include the full sample, while in Panels B and C we separately consider white and black students, respectively, since we will be running regressions on these sub-samples.[18] Overall, none of the variables considered are statistically different between the two groups. Importantly, our two main outcomes of interest, the Population and the Academic IAT, are well balanced between mixed and non-mixed rooms, both on aggregate and within racial groups. In the full sample, the mean of the Population and Academic IAT scores is -0.21 (recall that negative values of the IAT indicate prejudice against blacks relative to whites). When we split by race, white students exhibit higher prejudice levels than blacks, though even the latter group is found to hold negative stereotypes against itself. Turning to other controls, it is noteworthy that the UCT admission score, a proxy for academic ability at baseline, is on average identical for students in mixed and non-mixed rooms in all three panels. Socioeconomic controls are also well balanced across treatment arms.

Second, as an additional check, in Table 2a we report the coefficients of a regression of the likelihood of being assigned to a mixed room on individual pre-treatment characteristics for the full sample and separately by race.[19] For our key outcome variables, the Population and Academic IAT scores, no evidence of sorting appears at baseline. This is also true for other controls when we look at whites and blacks separately. The only variables that are significant at the 5 percent level are the race dummies in the full sample (column 1). However, this is to be expected: the fact that whites, Coloureds and other races have lower population shares in the sample of students living in dorms compared to blacks (the omitted category) mechanically increases their probability of having a roommate of a different race. In Table 2b we look at the correlation between the

[18] Summary statistics at follow-up for all variables, including academic outcomes, are reported in Appendix Table A3.

[19] Note that the number of observations in columns 2 and 3 do not sum up to 504 because Coloureds and other races are also included in the full sample. While we do not have a large enough sample to run separate estimations for these groups, when we pool them with blacks the results are very similar to those reported in column 3 (results available from the authors).

probability of being in a mixed room and explicit inter-racial attitudes at baseline. Columns 1, 3 and 5 employ a principal component index that we will use in our analysis later on, while columns 2, 4 and 6 show the individual variables entering this index. Once again, the allocation to a mixed room is not significantly correlated with pre-treatment attitudes towards inter-group interaction.

Finally, we test for sorting by forming all $N * (N-1)$ possible dyads and estimating the probability that two individuals i and j are in the same room as a function of the distance in observables between i and j. In particular, we follow Fachamps and Gubert (2007) and estimate the undirectional dyadic regression:

$$SameRoom_{ij} = \alpha + \beta |X_i - X_j| + \gamma(X_i + X_j) + \varepsilon_{ij} \quad (1)$$

where $SameRoom_{ij}$ takes value one if i and j are assigned to the same room, and X includes baseline values of the IAT scores, attitudes and socio-demographic characteristics. The coefficient of interest is β : a negative value of this coefficient would indicate sorting, i.e., that individuals who differ more in observables are less likely to be roommates.[20] Column 1 of Table 3 shows that the only variable that is marginally significant (at the 10 percent level) is whether two students have both attended a private high school. The difference in their UCT admission score does not predict rooming with each other. Importantly, in column 2 we show that differences in IAT scores do not predict being roommates, neither for the Population nor for the Academic IAT. Finally, column 3 shows that there is no sorting on explicit attitudes either.

Overall, the results in tables 1, 2 and 3 increase our confidence that the roommate assignment mechanism can be considered as good as random for the purpose of our analysis.

4 Empirical strategy

To estimate the effects of exposure to a roommate of a different race on stereotypes, we estimate the following benchmark specification on the full sample and on the subsamples of blacks and whites:

[20] The term $(X_i + X_j)$ is included to capture level effects and, as discussed in Fafchamps and Gubert (2007), its coefficient is not interpretable. For this reason we do not display it in the table.

$$Y_{ijkt} = \alpha Y_{ijk0} + \beta MixRoom_{ik0} + \boldsymbol{\gamma} Race_i + \lambda X_{ik0} + \mu X_{jk0} + \delta_k + \varepsilon_{ijkt} \quad (2)$$

where Y_{ijkt} is the IAT score of student i paired with student j, in residence k, in the follow-up survey (time t) and Y_{ijk0} is the baseline (time 0) value of the same variable; $MixRoom$ is a dummy equal to 1 if at baseline the student was assigned a roommate of a different race; $Race_i$ is a vector of race dummies (White, Coloured, Indian or Asian or Other, with Black as omitted category); X_{ik0} is a set of individual controls measured at baseline that include: gender, UCT admission score, household wealth, the student's monthly consumption expenditure, and a dummy equal to one if the respondent is not from South Africa.[21] The vector X_{jk0} is the same set of controls for the roommate; δ_k is a set of residence dummies, and ε_{ijkt} is the error term.[22] Our coefficient of interest is β. A positive value of this coefficient indicates a reduction in prejudice against blacks (recall that negative values of IATs indicate negative stereotypes against blacks relative to whites and positive value stereotypes in favor of blacks).

For our other outcomes of interest, academic performance, attitudes and prosocial behavior, we estimate a specification analogous to (2) but without including the lagged dependent variable when not available (e.g., because academic grades or lab experiments are only measured at follow-up). Also, in the specification where the dependent variable is academic performance, we add a set of dummies indicating the program in which the student is enrolled in.[23] We estimate (2) using OLS with robust standard errors. For those attitudinal variables that are categorical (and ordered), we employ an ordered logit model.

[21] *Household wealth* is an index that measures per capita ownership of durable goods in the respondent's household. It is calculated by applying principal component analysis to the following categories of goods: computer, fridge, TV, landline and mobile phones, bicycles, motorbikes, bakkies, cars, electric stove, gas stove, kettle, and geyser. *Consumption* includes expenditure on lunches, dinners, food, alcohol, cigarettes, cell phone minutes and entertainment.

[22] To avoid restricting the sample due to missing values of some control variables, for control variables other than the baseline IAT score we replace missing values with the means and include in the regression dummies taking value one for those observations in which the missing value has been replaced.

[23] In the specifications where the dependent variable comes from the prisoner dilemma game, we also include a dummy for whether the respondent knows the game player he/she has been matched with.

5 Results

5.1 Implicit Association Tests

Table 4 contains our first main result. It reports the estimated coefficients for equation (2) for the full sample and for whites and blacks separately. The coefficient on *Mixed Room* captures the change in IAT for students who have been allocated a roommate of a different race, compared to students who have a roommate of their own race. In the full sample (column 1), results show that being exposed to a roommate of a different race has no significant effect on prejudice on average, as measured by the Population IAT. This is not surprising because, due to the way in which the IAT is defined (it captures relative prejudice towards one group versus the other), if each group became less prejudiced towards the other this would imply movements of the IAT in opposite directions. Indeed, when we split the sample between whites and blacks (columns 2 and 3), we find that exposure to a roommate of different race significantly reduces prejudice against blacks for white students and has no significant effect for blacks. The magnitude of the effect for whites (0.282) corresponds to about half of a standard deviation of the IAT score, and indicates that exposure to a different race brings the average white student to the same level of prejudice as the average black student, which is quite noteworthy.[24]

The coefficient on Mixed Room in the black subsample has the opposite sign (indicating a relative movement in favor of the other group) but is not significant. The lack of a significant effect on blacks is not surprising in the context of psychological theories of intergroup contact. Tropp and Pettigrew's (2005) meta-analysis found that the effect of interaction on the reduction of intergroup bias was significantly stronger for members of 'high-status' groups than for members of 'low-status' groups. While the literature reviewed by these authors mostly includes lab experiments and correlational studies, our findings confirm that this result holds in a field setting when we exploit exogenous variation in contact.

In columns 4-6 the dependent variable is stereotypes regarding academic performance, as measured by the Academic IAT. This variable captures the belief on the relative academic ability of

[24] As reported at the bottom of Table 4, the mean of the dependent variable at follow-up in the subsample of students with same-race roommate is -0.41 for whites and -0.08 for blacks.

blacks versus whites at the end of the first year, i.e., after the respondent has seen the performance of the roommate but also of other students at UCT. Results show that, on average, the simple exposure to members of another race does not change stereotypes on academic ability. Presumably, the 'academic quality' of the roommate should matter when determining whether beliefs on relative academic performance are affected, as we show below.

The results in Table 4 condition on roommate characteristics in order to isolate the effect of roommates' race. In Appendix Table A4 we replicate Table 4 without including roommate controls. The estimates are very similar to those reported in Table 4, although the coefficient on *Mixed Room* in the white subsample drops by about 1/3 and becomes significant at 10 percent level in the black subsample. This suggests that part of the overall effect of exposure to a different group comes from the characteristics of that group, but an improvement in implicit attitudes towards outgroup members is observed with or without conditioning on such characteristics.

In Appendix Table A5 we conduct a sensitivity analysis to show that our results are not driven by outliers, given the relatively small sample size. Once we exclude influential observations using the DFBETA method, the results become even stronger: the coefficient on *Mixed Room* for the sample of white students increases to 0.345 and becomes significant at the 1 percent level.

5.1.1 Do stereotypes respond to new information?

To understand what mechanisms underlie stereotype change, we next focus on the role played by information and belief-updating. We are interested in understanding to what extent the change in stereotypes held by students in mixed rooms responds to new information on the characteristics of the other type, or to mere exposure which could, for example, affect the taste for interaction. To this end, we estimate the effect of having a roommate of a different race separately for three categories of individuals: the first category is that of people who held certain beliefs at the beginning of the academic year and saw them 'confirmed' by the particular realization of the roommate that they got. The second category includes people whose roommate generated a 'positive surprise' by conveying more positive information on the characteristics of blacks relative to the ex ante beliefs of the respondent; and the third category includes people who received a 'negative surprise' in the sense that the characteristics of their roommate conveyed negative information on blacks compared

to the individual's ex ante beliefs.

To operationalize the concept of a positive and negative surprise, we exploit some variables that we collected at baseline. First, we define a measure of surprise based on *academic ability*. This is the measure that we expect should be most relevant for updating stereotypes on academic performance, i.e., the variable that should have the most explanatory power when the outcome is the Academic IAT. We start by classifying individuals with a higher than average Academic IAT score at baseline as people with 'high (ex ante) beliefs' on black students' achievements. We next consider the UCT admission score of the roommate, which measures the academic skills of the roommate when he/she entered UCT, and we classify roommates with an above-average UCT score as 'high ability'. We then construct two 'surprise' variables based on various combinations of these indicators. *Positive surprise* takes value 1 when a respondent with 'low beliefs' had a 'high ability' black roommate or a 'low ability' white roommate. We use the term 'positive' to indicate that in both cases the respondent should update upwards his/her prior on the academic prospects of blacks versus whites.[25] *Negative surprise* takes value 1 when a respondent with 'high beliefs' had a 'low ability' black roommate or a 'high ability' white roommate: in this case the respondent should update downwards his/her prior on the academic achievement of blacks versus whites. The residual category is that of people whose ex ante beliefs were confirmed. Appendix Figure A1 summarizes our definitions of 'surprise', with the sign "+" indicating positive surprise, "−" negative surprise, and "0" no surprise.[26]

Our second measure of surprise aims at capturing dimensions that should be more relevant for updating *general* racial stereotypes as proxied by the Population IAT. Here it is less obvious what types of information would lead someone to change their mental associations of good and bad concepts with racial categories. We consider three statements that we included in our baseline survey asking people if they agreed or disagreed: (i) "It is ok to develop friendships just because you know they can be of use to you"; (ii) "Those in need have to learn to take care of themselves and not depend on others"; and (iii) "The first value you will teach your children is to help others who

[25] Recall that IATs are always defined in *relative* terms, comparing one group of individuals to another, and not in absolute terms.

[26] In our sample, the share of individuals who received a positive surprise about academic ability is 0.18 and those receiving a negative surprise 0.21.

are in difficulty". With a slight abuse of terminology, we define 'altruistic' the respondents who, at baseline, disagreed with (i) and (ii) while agreeing with (iii). Our conjecture is that interaction with a roommate who does not view friendships in a purely opportunistic way and who places high value on helping others should convey favorable information on the 'positive' characteristics of a group. As before, we classify individuals with a higher than average Population IAT score at baseline as people with 'high (ex ante) beliefs', i.e., low prejudice against blacks. We define *Positive surprise* as a dummy taking value 1 when a respondent with 'low beliefs' has an 'altruistic' black roommate or a 'non-altruistic' white roommate. Conversely, *Negative surprise* takes value 1 when a respondent with 'high beliefs' had a 'non-altruistic' black roommate or an 'altruistic' white roommate. Again, the residual category is that of people whose ex ante beliefs were confirmed.[27]

Table 5 reports the results of our analysis when we augment specification (2) introducing interactions of *Mixed Room* with *Positive* and *Negative surprise*. In Panel A the dependent variable is the Population IAT, while in Panel B it is the Academic IAT. In columns 1 to 3 surprise is defined in terms of academic ability, while in columns 4 to 6 it is defined in terms of altruistic attitudes, as detailed above.

Panel A shows results that are generally consistent with our earlier analysis. Racial stereotypes do not change on average (columns 1 and 4) but improve significantly for white respondents in mixed rooms (columns 2 and 5). Notice that in these columns the interactions with our surprise variables are not significant, suggesting that the mere exposure to the other group led to a reduction in prejudice, independent of the new information conveyed by the roommate regarding the academic ability or altruistic attitudes of blacks.

When we take the Academic IAT as dependent variable (Panel B), however, a different result emerges. In column 2, while the coefficient on the standalone *Mixed Room* variable is not significant for whites, the interaction with *Positive surprise* is positive and significant at the 1 percent level. This suggests that academic stereotypes do respond to new information regarding ability and are not changed by mere exposure. Furthermore, columns 4-6 show that new information on roommates' altruistic attitudes does not affect stereotypes about school performance – a dimension

[27] In this case, the share of individuals who received a positive surprise is 0.2 and the share who received a negative surprise is 0.28.

that is possibly orthogonal to altruism.

It is interesting to discuss the different results on the Population and Academic IAT in the light of existing evidence from social psychology. In a meta-analysis involving both correlational studies and laboratory experiments, Pettigrew and Tropp (2008) tested the relative importance of the three most commonly invoked 'mediators' through which contact may reduce prejudice: increased knowledge about outgroup members, reduction in anxiety about contact, and increased empathy (e.g., greater ability to take the other's perspective). They find that knowledge plays a weaker role than anxiety reduction and empathy. This is consistent with what we find in Panel A of Table 5, since the measure of stereotypes we employ here is very similar to measures of racial prejudice in the psychology literature. On the other hand, stereotypes on academic ability are not commonly studied in the literature on intergroup contact, and it is reasonable that those types of stereotypes may respond more to hard evidence on academic performance (i.e., knowledge) than to empathy or anxiety.

5.2 Academic Performance

A central motivation underlying the policies of many universities that apply random assignment of roommates is to mix students who have different academic achievement. It is therefore important to investigate the effect of having a roommate of a different race on a student's academic performance.

Appendix Figure A2 shows the distribution of the average (standardized) GPA for blacks and whites at the end of the first academic year.[28] White students have a higher academic performance compared to black students, and the gap is substantial: the mean GPA for whites is about .6, while that for blacks is $-.21$ (see Appendix Table A3).

In Table 6 we test whether having a roommate of a different race significantly affects this gap. We use three different outcomes for academic performance and we report robust standard errors in parentheses and p-values adjusted for multiple hypothesis testing in square brackets. Specifically, we adjust p-values using Westfall and Young's (1993) free step-down resampling method for the family-wise error rate (FWER), defined as the probability of rejecting at least one true null

[28] The pattern when we pool blacks, Coloureds and other races in a single category is very similar to that of blacks.

hypothesis.[29]

In columns 1 to 3 the dependent variable is the average GPA at the end of the first year. Regressors (not shown) include all the controls of the benchmark specification (Table 3) for the respondent and for the roommate, plus academic program fixed effects.[30] On average, students in mixed rooms have a higher GPA than students in same-race rooms (column 1), but this effect masks an important heterogeneity. Black students who share the room with someone of a different race improve their GPA by .26 of a standard deviation (column 3), while whites who room with non-whites see a, albeit insignificant, deterioration in their GPA of .12 standard deviations (column 2). The magnitude of the positive effect for blacks amounts to closing about one third of the gap between white and black students (the average GPA for whites and blacks in same-race rooms are, respectively, 0.67 and -0.24).

In the remaining specifications of Table 6 we consider two additional measures of academic performance. The dependent variable in columns 4-6 is the number of exams passed during the first year (ranging from 0 to 9). Being allocated to a mixed race room increases the number of exams passed in the first year by .46 on average. This effect is particularly strong among black students –an increase of .72 exams, significant at the 1 percent level after adjusting for multiple hypothesis testing – while the (negative) impact on white students is smaller and insignificant.

In columns 7-9, the outcome we consider is *Eligibility to continue,* an assessment made by the Faculty Examination Committee based on the number of exams passed and the grades of each student. In particular, this variable takes value one if the student is declared eligible to continue studying, possibly subject to passing makeup exams, and zero otherwise. Once again, students in mixed rooms are more likely to be eligible to continue and this effect is driven by black students. For this subsample the estimated effect of 17 percentage points implies a 19 percent increase in the probability to be eligible to continue university over the control group mean, significant at the 1 percent level. In this case the effect on whites is a precise zero.

In the last three columns of Table 6 we show the effect of having a roommate of a different race on a summary index of performace, computed as the first principal component of *GPA*, *Number*

[29] The p-values in brackets are based on 10,000 replications.

[30] Academic programs are a subdivision of college tracks within each faculty. Within each academic program, students in the first year typically take the same (compulsory) courses.

of exams passed and *Eligibility to continue*. Being in mixed room is strongly positively correlated with a general index of performance for black students.

Estimates without controls for roommate characteristics and sensitivity analysis excluding influential observations are reported in Appendix Tables A6 and A7, respectively. These estimates are virtually identical to those in Table 6.

Appendix Table A8 reports analogous estimates for academic performance in the second year. Given that students are free to change room –and to leave the residence altogether– in their second year, it is interesting to test whether the effects of exposure during the first year are persistent. The estimates in Table A8 show that the effect on GPA fades away, while the effects on the number of exams passed and on eligibility to continue remain highly significant in the full sample (marginally significant for blacks) and of similar magnitude. This is remarkable because the outcomes in Table A8 do not cumulate performance in years 1 and 2, but only refer to year 2, i.e.: GPA in second year exams, number of exams passed in the second year, and eligibility to continue to the third year. The lack of persistence of the gains in GPA may suggest that in their second year students, who had been exposed to a different race roommate, focus on staying on track more than on improving their grades in every single exam.

We next try to uncover the mechanism through which the improvement in academic performance of black students in mixed rooms occurs. A first hypothesis is that blacks who have a non-black roommate are on average matched with someone who has a stronger academic background. As noted above, UCT applies a certain degree of affirmative action in admissions, and as a consequence black students have on average lower UCT admission scores compared to other groups, reflecting the legacy of unequal apartheid schooling systems by race. Thus, one may conjecture that it is not exposure to a different race but exposure to a higher-achieving peer that generates the gains in academic performance for blacks. While plausible, this explanation is not supported by the data. The regressions in Table 6 directly control for roommate's ability (proxied by UCT admission score) and this variable is never significant.[31] A further indication in this respect comes from Appendix Table A9. Here we control for whether an individual and his/her roommate are in

[31] We also tried including an interaction between *Mixed Room* and roommate's UCT admission score, but did not find significant effects.

the same faculty (31 percent of the students in our sample are). For freshmen in the same faculty it should be easier to study together and help each other with coursework. When we interact *Mixed Room* with an indicator for respondents in the same faculty, we find that the coefficient on the interaction is typically positive for blacks, but never significant. This suggests that purely academic interaction is not responsible for our results.

In Table 7 we start exploring to what extent dimensions other than mere academic characteristics of one's roommate matter for academic performance. Specifically, we test whether having a roommate of a different race has a heterogeneous impact depending on the roommate's stereotypes. We do so by including an interaction between *Mixed Room* and *Roommate's Population IAT* at baseline. The underlying idea is that a black student paired with a roommate of a different race may be able to improve his/her academic performance by more, the less prejudiced the roommate is. Note also that, given that we are considering the sample of individuals whose roommates have taken the IAT test, the sample used in Table 7 is smaller compared to the ones used in tables 5 and 6.

The results are quite interesting. In column 1, the coefficient on the interaction term is positive and significant, suggesting that in the full sample students with roommates of a different race, who hold less negative stereotypes against blacks, improve their GPA score by more. When we disaggregate by race, this coefficient is positive both for whites and for blacks but statistically significant only for blacks. To gauge the economic significance of the effect, consider the following thought experiment. Suppose we placed a black student in a room with an average white student and we changed the white student's IAT score from the (baseline) sample average for whites ($-.36$) to absence of prejudice ($Population\ IAT = 0$). Based on our estimates, other things equal, the GPA of the black student would increase by 0.12 standard deviations compared to a black student with a white roommate who has average prejudice. For the number of exams passed (columns 4-6) and for eligibility to continue (columns 7-9) the direction of the impact is similar but not significant at conventional levels.

The findings in Table 7 are consistent with at least two (non mutually-exclusive) explanations. The first is that students of different races have a more positive social interaction when they

are paired with someone who is not prejudiced against them. This may improve their ability to communicate (Lang, 1986), lead to greater well-being and possibly reduce intergroup anxiety (Stephan and Stephan, 1985), which allows them to perform better academically.[32] The second explanation is related to the stereotype threat literature (e.g., Steele and Aronson, 1995). To the extent that black students paired with white roommates realize that the prejudice of the latter is reduced, they may perceive a lower risk of confirming negative stereotypes about their own group, which could improve their performance in the exams.

Overall the results in Table 7 suggest that there is significant complementarity between the reduction of negative stereotypes against lower status groups and the performance gains from intergroup contact. To the best of our knowledge, our analysis is the first to document this phenomenon, which seems very relevant for the debate on racial integration in universities.

5.3 Attitudinal and behavioral measures

Our final set of results test for the presence of behavioral changes induced by intergroup interaction that could be consistent with the stereotype reduction and academic gains documented above. We start from a set of 15 outcomes that include behavioral and attitudinal measures that we collected through our survey and through lab experiments. To avoid over-rejecting the null hypothesis due to multiple inference (Anderson, 2008), we first group the 15 outcomes into three main indices that aggregate information across related variables: we label them 'friendship', 'attitudes' and 'pro-social behavior'. Since the original outcomes are measured at different levels (some are binary, other ordinal, other continuous), we use as a summary index the first principal component computed using the polychoric correlation matrix of the outcomes in each group, which is theoretically invariant to changes in the number of levels.[33] We also construct a global index of 'social behavior' that extracts the first principal component from all 15 outcomes at once. Since even with the summary indices we are still testing multiple hypotheses, we also report FWER-adjusted p-values

[32] In the case of white students, two opposite effects coexist. On the one hand, the reduction in intergroup anxiety should allow them to perform better; on the other, whites in mixed rooms have a roommate who is on average lower performing. The two effects go in opposite directions and may account for the lack of a positive effect on whites' academic performance.

[33] Results are robust to different index choices, such as the mean effects approach used in Kling, Liebman and Katz (2007) or the one in Anderson (2008).

in square brackets.[34] Table 8 reports these estimates; results for the disaggregated individual outcomes are reported in Appendix Tables A10, A11 and A12.

Columns 1-3 of Table 8 report the effect of having a roommate of a different race on inter-racial friendships. The index of friendships includes the following variables: (i) the number of times the respondent socialized with someone of a different race in the last month (excluding the roommate); (ii) the last time the respondent socialized with people of a different race;[35] (iii) the fraction of actual friends and study mates of a different race (excluding the roommate);[36] (iv) the number of desired friends of a different race in a hypothetical leisure group and in a hypothetical study group (excluding the roommate).[37] Column 1 shows that exposure to a roommate from a different racial group induces a sizeable increase in inter-racial friendships in the full sample, and this effect is significant at the 1 percent level. When we disaggregate by race of the respondent, the effect is particularly strong for white students, in line with our previous results for the Population IAT score. Note that these effects are not mechanical given that we exclude the roommate among the friends or studymates. These findings suggest that contact with a roommate of a different race leads to changes in the pattern of social interactions, which likely affects a range of individual outcomes, including stereotypes and academic performance.

In columns 4 to 6 we measure the effect of the treatment on an index of attitudinal measures. This index includes the following five variables: (i) the frequency with which the respondent talked about discrimination and racial bias with friends in the last month;[38] (ii) an indicator for whether the respondent feels comfortable talking to people about race and discrimination; (iii) a dummy taking value one if the respondent disagrees that affirmative action in university admissions should be abolished; (iv) an indicator for whether the individual disagrees with the statement: "I would probably feel a little self-conscious dancing with a person of another race in a public place"; and

[34] We apply Westfall and Young's (1993) resampling method with $10,000$ interactions.

[35] Possible responses were: yesterday, last week, last month, last year, never.

[36] Friends were defined as "those you can turn to for help if needed" and we asked respondents to list the first name, gender, age and race of up to five friends.

[37] We asked our respondents how many people of different race they would want in a group of 7 people (including themselves), and they could choose a number between 0 and 6. We use a dummy equal to one if the respondent states that she would like to have more than half of the members from a difference race from her own in a leisure or study group. This captures a situation where the respondents wants to retain a majority of group members of different race from it own race.

[38] Possible answers were: never, rarely, sometimes, most of the time, always.

(v) an indicator for whether the individual disagrees with the statement "I would probably feel a little self-conscious having a girlfriend or boyfriend of another race". The coefficient on *Mixed Room* is positive and significant in the full sample (column 4), indicating that on average exposure to a roommate of a different race improved inter-racial attitudes. When we disaggregate by race of the respondent, again we find that the effect is driven by whites, consistent with the reduction in negative stereotypes by this group that emerged from our IAT results (although we lose statistical significance after correcting for multiple hypothesis testing).

Columns 7 to 9 examine impacts on prosocial behavior, measured through an index that aggregates behavioral and experimental measures, in particular: (i) membership in community service or volunteer organizations; (ii) the amount of money given to a charity in the past year; and two experimental measures elicited through a prisoner dilemma game: (iii) an indicator for whether the respondent chose to cooperate, and (iv) an indicator for whether the respondent believed that the partner would cooperate.[39] Our estimates suggest an improvement in pro-social behavior for all groups, but the effect is only (marginally) significant for blacks after correcting for multiple hypothesis testing.

Finally, in the last three columns of Table 8 we aggregate all individual outcomes into a 'global' index of social behavior. We find positive effects across the board, significant at the 1 percent level in the full sample and in the white sub-sample and at the 10 percent in the black sub-sample. The pattern for explicit attitudes and behaviors thus mirror our earlier findings on implicit attitudes, where the stronger effects of inter-racial contact on stereotypes were found among whites.

As a last test of behavioral change, we adopt a revealed preference approach and study residential choices for the students in our sample after the first year. If the students have opted to remain in the residences, through the administrative data we know in which residence and room they are and, if it is a double room, we know the identity of their roommate. If the students have left the university residence system we do not have any information on their location. On average, only 15.3 percent of the students still live in a residence during the second year. This limits our ability to estimate rich regression models. From the top panel of Table 9, however, we can see

[39] In our sample, 57 percent of the subjects chose to cooperate and 61 percent stated that they believed the partner would cooperate.

that having been assigned a roommate of a different race in the first year is uncorrelated with the choice of staying in a university residence during the second year. This is encouraging for our results because, if inter-racial contact had generated significant 'unhappiness', we would expect students in mixed rooms to be more likely to opt out of the residence system, and we find they are not.

For the small number of students from our original sample who remain in residence and who are in a double room in the second year, we can perform an additional test. Given that students have the freedom to pick their roommate after the first year, we can test whether the likelihood of changing roommates in year two differs for students who were exposed to a different race in year 1. We do this in Panel B of Table 9, testing the difference in means of the probability of having the same roommate, both for the full sample and for blacks and whites separately. We find no significant differences between students assigned to mixed and non-mixed rooms at baseline, which is again reassuring because it suggests that roommates of a different race were not differentially disliked or preferred.

6 Conclusions

Social diversity does not always translate into equal opportunities for different groups in society: concerns exist about the possibility that today's schools, workplaces and social networks display significant segregation along income or racial lines. Integration policies have been proposed as a way of improving equality of opportunities but also of reducing inter-group prejudice and conflict. In this paper, we take advantage of a policy designed to randomly allocate roommates in some residences at University of Cape Town to investigate the effects of exposure to individuals of a different race on implicit and explicit attitudes, behavior and academic performance. To the best of our knowledge, ours is the first attempt to link implicit bias, as measured by the Implicit Association Test (IAT), with academic performance in the context of a real world policy setting.

Our results point to a number of positive effects from inter-racial contact generated through this policy. First, we find that living with a roommate of a different race during the first year reduces white students' negative stereotypes against blacks, as measured by the IAT. This effect is quite

remarkable because a number of transformation initiatives have happened in post-apartheid South Africa that have made inroads in reducing the salience of race. Yet, the interaction generated by the random policy allocation is able to (further) reduce prejudice. We also find significant positive effects on explicit attitudes towards the other race and on inter-racial friendships, again most pronounced for white students.

Finally, we show positive effects of inter-group contact on the academic performance of the negatively stereotyped group: black students in mixed rooms significantly improved their GPA, passed more exams and were more likely to be eligible to continue university. The latter two effects (but not the first) persisted into the second year of university, when most of the students were in a different residential setting. Interestingly, the prejudice level of one's roommate turns out to be a key ingredient in explaining academic gains: blacks paired with whites do better, the less prejudiced their roommate is. This is a potentially important result for the literature on academic peer effects, which has focused on the potential benefits of integration in terms of exposure to a different set of skills. Our findings point to the importance of assessing the impact of integration policies on both attitudinal measures (e.g., stereotype reduction) and performance, given that the two positively reinforce each other.

It is worth discussing to what extent our findings generalize beyond the context we study. Our experiment takes place in South Africa, a place that is certainly at the higher end of the distribution if one considers its history of inter-racial conflict and the significance of racial stereotypes. If anything, we think that this should have reduced the ex ante probability of success of a policy like the one we study. First, many would assume that such stereotypes were so deeply ingrained that it would have been difficult to affect them; second, because a large number of 'transformation' initiatives have been put in place in South Africa since the end of apartheid, the marginal effect of the roommate allocation policy we study may have been assumed to be small. The fact that in-depth interaction with the other group can reduce negative stereotypes even in historically charged contexts is a very encouraging message in this respect.

A second issue to consider in assessing external validity is that our sample is made of university students who are certainly not representative of the entire South African population. On the one

hand, these students are likely to be the political and economic elites of tomorrow's South African society, so that a change in their attitudes and behavior can bring a certain significance for the country as a whole. On the other hand, if we consider how this may have affected our estimates, the direction of a potential bias is not obvious. If these students are more 'malleable' than the general population, then our estimates would be an upper bound. But if they were already more open-minded and less reliant on preconceptions when they entered university, then there would have been a lower margin for the intervention to have an effect and the bias could be in the opposite direction.

Overall, we believe the specificity of the South African context should not be seen as a limitation of our results but, if anything, a strong proof of concept. More work to investigate the generalizability of our findings to different social and economic environments will help improve our ability to design integration policies in our increasingly diverse societies.

References

[1] Akerlof, G. and R. Kranton (2000), "Economics and Identity", *Quarterly Journal of Economics*, 115, 715–53.

[2] Alesina, A., E. La Ferrara (2005), "Ethnic Diversity and Economic Performance", *Journal of Economic Literature*, 43(3), 762-800.

[3] Allport, G.W. (1954), *The Nature of Prejudice*. Cambridge, MA: Perseus Books.

[4] Anderson, M.L. (2008), "Multiple Inference and Gender Differences in the Effects of Early Interventions: A re-evaluation of the Abecedarian, Perry Preschool, and Early Training Projects", *Journal of the American Statistical Association*, 103 (484), 1481-95.

[5] Barlow, F. K., Paolini, S., Pedersen, A., Hornsey, M. J., Radke, H. R. M., Harwood, J., Rubin, M., & Sibley, C. G. (2012). "The contact caveat: Negative contact predicts increased prejudice more than positive contact predicts reduced prejudice". *Personality and Social Psychology Bulletin*, 38, 1629-1643.

[6] Barnhardt, S. (2009), "Near and Dear? Evaluating the Impact of Neighbor Religion on Inter-Religious Attitudes: Experimental Evidence from India", mimeo, IFMR.

[7] Beaman, L., R. Chattopadhyay, E. Duflo, R. Pande and P. Topalova (2009), "Powerful Women: Does Exposure Reduce Bias?", *Quarterly Journal of Economics*, 124 (4).

[8] Berge, L.I., K. Bjorvatn, S. Galle, E. Miguel, D. Posner, B. Tungodden and K. Zhang (2015), "How Strong Are Ethnic Preferences?", NBER Working Paper 21715.

[9] Bertrand, M., D. Chugh and S. Mullainathan (2005), "Implicit Discrimination", *American Economic Review*, 95(2), 94-98.

[10] Boisjoly, J., G. J. Duncan, M. Kremer, D. Levy, and J. Eccles (2006), "Empathy or Antipathy? The Impact of Diversity", *American Economic Review*, 96(5): 1890–1905.

[11] Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2016), "Stereotypes", *Quarterly Journal of Economics*, 131(4), 1753-94.

[12] Carrell, S., B. Sacerdote and J. West (2013), "From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation," *Econometrica*. 81(3): 855-882.

[13] Carrell, S., M. Hoekstra and J. West (2016), "The Impact of College Diversity on Behavior Toward Minorities," *NBER working paper 20940*.

[14] Duflo, E., P. Dupas and M. Kremer (2011), "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya", *American Economic Review*, 101(5), pp. 1739-74.

[15] Fafchamps, M. and F. Gubert (2007), "The Formation of Risk-Sharing Networks", *Journal of Development Economics*, 83(2), 326-350.

[16] Fryer, R. and G. Loury (2013), "Valuing Diversity", *Journal of Political Economy,* 121(4), 747-774.

[17] Garlick, R. (2013), "Academic Peer Effects with Different Group Assignment Rules: Residential Tracking versus Random Assignment", mimeo, University of Michigan.

[18] Glover D., A. Pallais and W. Pariente (2015) "Discrimination as a Self-Fulfilling Prophecy: Evidence from French Grocy Stores", working paper.

[19] Greenwald, A. G., & Banaji, M. R. (1995). "Implicit social cognition: Attitudes, self-esteem, and stereotypes." *Psychological Review*, 102, 4-27.

[20] Greenwald, A. G., McGhee, D. E., & Schwartz, J. K. L. (1998). "Measuring individual differences in implicit cognition: The Implicit Association Test." *Journal of Personality and Social Psychology*, 74, 1464-1480.

[21] Hoff, K., Pandey P. (2006), "Discrimination, Social Identity and Durable Inequalities", *American Economic Review*, 96(2): 206-211..

[22] Hjort, J. (2014), "Ethnic divisions and production in firms", *Quarterly Journal of Economics*, 1899-1946.

[23] Kling, J.R., J.B. Liebman and L.F. Katz (2007), "Experimental Analysis of Neighborhood Effects." *Econometrica*, 75(1), 83-119.

[24] Lang, K. (1986), "A Language Theory of Discrimination," *Quarterly Journal of Economics*, 101, 363-382.

[25] Lowes, S., N. Nunn, J. A. Robinson and J. Weigel (2015), "Understanding Ethnic Identity in Africa: Evidence from the Implicit Association Test (IAT)." *American Economic Review Papers and Proceedings,* 105 (5), 340-45.

[26] Lyle, D. (2009), "The Effects of Peer Group Heterogeneity on the Production of Human Capital at West Point", *American Economic Journal: Applied Economics*, 1(4), 69-84.

[27] Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002), "Harvesting Implicit Group Attitudes and Beliefs from a Demonstration Web Site", *Group Dynamics: Theory, Research, and Practice*, 6, 101–115.

[28] Paluck, E., S. Green and D. Green (2018), "The contact hypothesis re-evaluated", *Behavioural Public Policy*, 1-30.

[29] Paolini, S., Harwood, J. and Rubin, M. (2010), "Negative intergroup contact makes group memberships salient: Explaining why intergroup conflict endures." *Personality and Social Psychology Bulletin.* 36, 1723-1738.

[30] Pettigrew, T. and L. Tropp (2006), "A Meta-analytic Test of Intergroup Contact Theory". *Journal of Personality and Social Psychology*, 90(5), 751-783.

[31] Pettigrew, T. and L. Tropp (2008), "How Does Intergroup Contact Reduce Prejudice? Meta-analytic Tests of Three Mediators". *European Journal of Social Psychology*, 38, 922-34.

[32] Rao, G. (forthcoming), "Familiarity Does Not Breed Contempt: Diversity, Discrimination and Generosity in Delhi Schools," *American Economic Review*.

[33] Sacerdote, B. (2001): "Peer effects with random assignment: Results for Dartmouth roommates," *The Quarterly Journal of Economics*, 116(2), 681-704.

[34] Scacco, A. and S. Warren (2018), "Can Social Contact Reduce Prejudice and Discrimination? Evidence from a Field Experiment in Nigeria", *American Political Science Review*, 112(3), 654-677.

[35] Shayo, M. (2009), "A Model of Social Identity with an Application to Political Economy: Nation, Class, and Redistribution", *American Political Science Review*, 103(2), 147-174.

[36] Steele, C. and J. Aronson (1995), "Stereotype Threat and the Intellectual Test Performance of African Americans", *Journal of Personality and Social Psychology*, 69(5), 797-811.

[37] Stephan, W. G. and C. Stephan (1985), "Intergroup Anxiety", *Journal of Social Issues*, 41(3), 157-175.

[38] Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971), "Social categorization and intergroup behaviour," *European Journal of Social Psychology*, 1, 149-178.

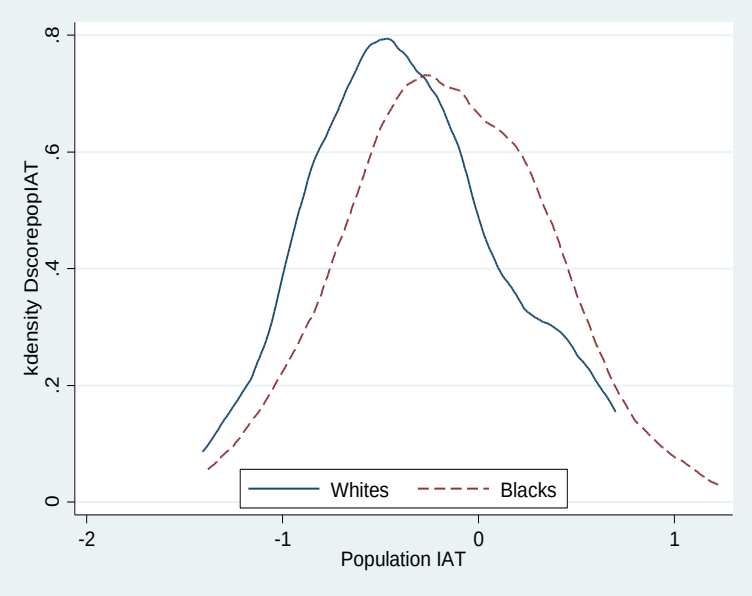
[39] Tropp, L. and Pettigrew, T. (2005), "Relationships between intergroup contact and prejudice among minority and majority status groups." *Psychological Science*, 16, 651-53.

[40] Van Laar, C., S. Levin, S. Sinclair and J. Sidanius (2005), "The effect of university roommate contact on ethnic attitudes and behavior", *Journal of Experimental Social Psychology*, 41, 329–345.

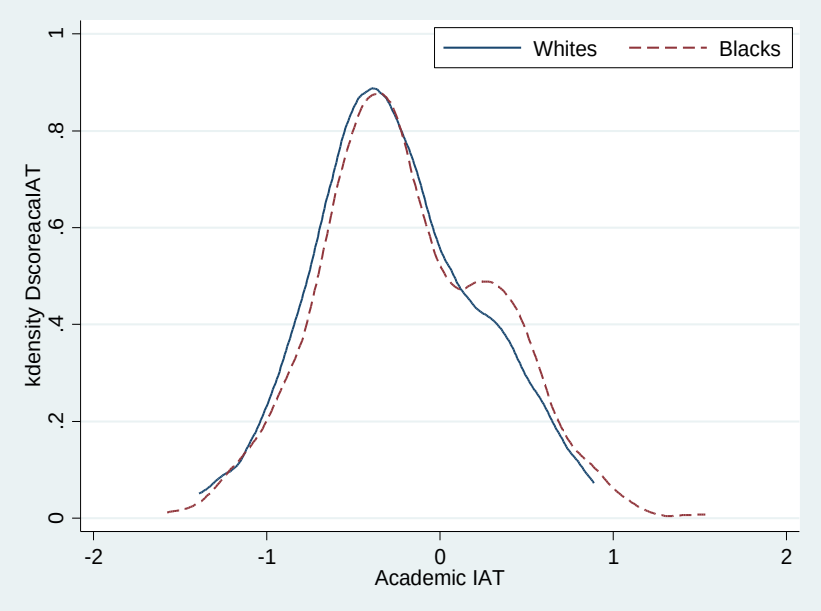
[41] Westfall, P. and S. Young (1993), *Resampling-based Multiple Testing: Examples and Methods for P-value Adjustment*, vol. 279, John Wiley & Sons.

Figures and Tables

Figure 1: Stereotypes as measured by IAT at baseline



Panel A: Population IAT



Panel B: Academic IAT

Table 1: Descriptive statistics

	Full sample		*Mixed rooms*		*Non mixed rooms*		*Non mixed - Mixed*	
	Mean	*Std. Dev.*	*Mean*	*Std. Dev.*	*Mean*	*Std. Dev.*	*Difference*	*P-value*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Full Sample								
Population IAT	-0.21	0.51	-0.25	0.50	-0.20	0.51	0.05	0.29
Academic IAT	-0.21	0.49	-0.21	0.52	-0.21	0.48	-0.01	0.89
Female	0.67	0.47	0.66	0.47	0.67	0.47	0.01	0.89
UCT admission score	0.46	0.05	0.47	0.05	0.46	0.05	0.00	0.36
Wealth Index	-0.04	2.10	-0.03	1.86	-0.04	2.20	-0.02	0.93
Consumption	0.93	0.85	0.99	0.92	0.89	0.82	-0.11	0.20
Foreign	0.11	0.32	0.14	0.34	0.10	0.30	-0.04	0.23
Private high school	0.59	0.49	0.61	0.48	0.59	0.49	-0.02	0.65
No. Obs.	495		154		341		495	
Panel B: Whites								
Population IAT	-0.36	0.49	-0.34	0.46	-0.37	0.51	-0.03	0.77
Academic IAT	-0.25	0.46	-0.23	0.49	-0.26	0.45	-0.03	0.71
Female	0.67	0.47	0.64	0.49	0.69	0.47	0.05	0.61
UCT admission score	0.49	0.04	0.49	0.04	0.49	0.04	0.00	0.91
Wealth Index	0.75	1.790	0.40	1.27	0.93	1.98	0.53	0.13
Consumption	1.18	0.92	1.18	1.01	1.18	0.87	0.00	0.99
Foreign	0.07	0.25	0.10	0.31	0.05	0.22	-0.05	0.31
Private high school	0.73	0.44	0.69	0.47	0.75	0.43	0.06	0.49
No. Obs.	116		39		77		116	
Panel C: Blacks								
Population IAT	-0.15	0.51	-0.18	0.54	-0.15	0.50	0.04	0.57
Academic IAT	-0.19	0.50	-0.15	0.51	-0.20	0.49	-0.05	0.44
Female	0.68	0.47	0.73	0.45	0.67	0.47	-0.06	0.34
UCT admission score	0.45	0.05	0.45	0.05	0.45	0.05	0.00	0.95
Wealth Index	-0.44	2.03	-0.57	1.57	-0.39	2.15	0.17	0.51
Consumption	0.81	0.80	0.89	0.91	0.79	0.77	-0.10	0.35
Foreign	0.12	0.32	0.15	0.36	0.11	0.31	-0.04	0.36
Private high school	0.53	0.50	0.54	0.50	0.53	0.50	-0.01	0.92
No. Obs.	329		74		225		329	

Notes: The p-value in col. 8 is for the two-sided test that the difference in means between cols. 5 and 3 is zero. "UCT admission score" is the sum of high school final grades, with weights depending on the specific department the student enrolls in; the "Wealth index" measures per capita ownership of durable goods in the respondent's household and is constructed applying principal component analysis to the following categories of goods: computer, fridges, TV, landline and mobile phones, bicycles, motorbikes, bakkies, electricity, gas, kettles, geysers and cars; "Consumption" is the monthly consumption in Rands on lunches, dinners, food, alcohol, cigarettes, cell phone minutes, entertainment; "Foreign" is a dummy equal to one if the respondent is not from South Africa, “Private high school” is equal to one if the respondent was enrolled in a private high school before joining UCT.

Table 2 (a): Probability of being in a mixed room at baseline

Dependent variable = 1 if roommate of a different race at baseline			
Sample:	*Full Sample*	*Whites*	*Blacks*
	(1)	(2)	(3)
Population IAT	-0.023	0.053	-0.044
	(0.041)	(0.092)	(0.050)
Academic IAT	0.033	0.006	0.047
	(0.041)	(0.117)	(0.049)
White	0.138**		
	(0.056)		
Coloured	0.792***		
	(0.029)		
Indian/Other	0.538***		
	(0.087)		
Female	0.035	-0.059	0.066
	(0.041)	(0.100)	(0.050)
UCT admission score	-0.169	-0.207	-0.016
	(0.433)	(1.226)	(0.501)
Foreign	0.087	0.207	0.081
	(0.077)	(0.201)	(0.089)
Private high school	-0.020	-0.112	0.005
	(0.042)	(0.114)	(0.050)
Wealth index	-0.016*	-0.033	-0.015
	(0.008)	(0.021)	(0.010)
Consumption	0.012	0.026	0.021
	(0.025)	(0.054)	(0.032)
Constant	0.248	0.527	0.145
	(0.197)	(0.618)	(0.225)
Mean of the dep. variable	0.311	0.336	0.225
Residence fixed effects	X	X	X
R-squared	0.171	0.066	0.023
No. Obs.	495	116	329

Notes: OLS estimates with robust standard errors in parentheses. *** p<.01, ** p<.05, * p<.10. All regressions include residence fixed effects.

Table 2 (b): Probability of being in a mixed room at baseline

Dependent variable = 1 if roommate of a different race at baseline

	Full Sample		*Whites*		*Blacks*	
	(1)	(2)	(3)	(4)	(5)	(6)
Attitudes index, PCA	-0.001		0.006		-0.013	
	(0.025)		(0.049)		(0.032)	
Talk about race		-0.027		-0.046		-0.033
		(0.019)		(0.049)		(0.022)
Comfortable talking about race		-0.040		-0.104		-0.006
		(0.051)		(0.137)		(0.059)
Disagree to abolish affirmative action		-0.056		-0.077		-0.109*
		(0.046)		(0.099)		(0.061)
Not conscious dancing with a person of another race		0.001		-0.032		-0.002
		(0.055)		(0.123)		(0.066)
Not conscious having boyfriend/girlfriend of another race		0.041		0.090		0.024
		(0.045)		(0.093)		(0.056)
Controls	X	X	X	X	X	X
R-squared	0.172	--	0.076	--	0.022	--
No. of Obs.	452	--	111	--	293	--

Notes: Coefficients reported in cols. 2, 4, 6 are from different regression. OLS estimates with robust standard errors in parentheses. *** p<.01, ** p<.05, * p<.10. Controls include a dummy equal to one if the respondent is female, UCT admission score, wealth index, consumption, foreign, private high school, as defined in the footnote of Table 1. All regressions include residence fixed effects.

Table 3: Probability of being in the same room: dyadic analysis

Dependent variable = 1 if i and j are in the same room			
	(1)	(2)	(3)
Difference in population IAT		0.00095	
		(0.0013)	
Difference in academic IAT		-0.000546	
		(0.0012)	
Difference in Attitudes Index, PCA			-1.14E-05
			(0.0006)
Difference in UCT admission score	-0.000306	-0.000292	-0.000309
	(0.0005)	(0.0005)	(0.0005)
Difference in Foreign	0.00213	0.00222	0.00233
	(0.0027)	(0.0027)	(0.0028)
Difference in Private high school	-0.00204*	-0.00203*	-0.00217*
	(0.0011)	(0.0011)	(0.0012)
Difference in Wealth index	-0.000272	-0.000281	-0.00027
	(0.0002)	(0.0002)	(0.0002)
Difference in Consumption	-1.40E-07	-1.30E-07	-1.12E-09
	(0.000)	(0.000)	(0.000)
Constant	0.00813***	0.00789***	0.00832***
	(0.0011)	(0.0015)	(0.0013)
Mean of dep. var	0.00615	0.00615	0.00637
No. Obs.	34,474	34,474	31,716

Notes: OLS estimates with standard errors corrected for dyadic correlation of errors in parentheses. *** p<.01, ** p<.05, * p<.10. Variables are defined in the footnote of Table 1.

Table 4: Stereotypes and exposure to a roommate of different race

Dependent variable:	*Population IAT*			*Academic IAT*		
Sample:	*Full Sample*	*Whites*	*Blacks*	*Full Sample*	*Whites*	*Blacks*
	(1)	(2)	(3)	(4)	(5)	(6)
Mixed Room	-0.034	0.282**	-0.079	-0.063	0.042	-0.036
	(0.055)	(0.135)	(0.073)	(0.046)	(0.099)	(0.060)
Controls	X	X	X	X	X	X
Roommate controls	X	X	X	X	X	X
Mean of dep.var. in same race room	-0.166	-0.414	-0.084	-0.208	-0.298	-0.185
R-squared	0.122	0.209	0.095	0.092	0.280	0.096
No. Obs.	495	116	329	495	116	329

Notes: OLS estimates with robust standard errors in parentheses. * p<.10, ** p<.05, *** p<.01. Higher values of the dependent variable (IAT) indicate less prejudice against blacks. All control variables are measured at baseline. All regressions include university residence fixed effects and the dependent variable at baseline. Controls include IAT at baseline, a dummy equal to one if the respondent is female, UCT admission score, wealth index, consumption, foreign, private high school, as defined in the footnote of Table 1. In cols. 1 and 4, controls also include the race of the respondent (white, colored and Indian/others) with black as the omitted category. Controls for roommate include: UCT admission score, wealth index, consumption, foreign and private high school (female not included as all rooms are single sex).

Table 5: Information and stereotypes

Sample:	*Full sample* (1)	*Whites* (2)	*Blacks* (3)	*Full sample* (4)	*Whites* (5)	*Blacks* (6)
Panel A: Dependent variable: Population IAT						
	Surprise defined on the basis of					
	Academic Ability			*Altruistic Attitudes*		
Mixed Room	0.001	0.259**	-0.022	-0.015	0.293**	-0.077
	(0.061)	(0.130)	(0.085)	(0.062)	(0.150)	(0.084)
Mixed Room*Positive Surprise	-0.038	0.133	-0.109	-0.114	-0.015	-0.074
	(0.160)	(0.432)	(0.201)	(0.115)	(0.299)	(0.163)
Mixed Room*Negative Surprise	-0.144	-0.067	-0.226*	0.056	-0.085	0.354***
	(0.093)	(0.306)	(0.134)	(0.092)	(0.202)	(0.127)
Mean of dep. var. in same race room	-0.166	-0.414	-0.083	-0.166	-0.414	-0.083
R-squared	0.126	0.214	0.099	0.126	0.212	0.103
No. Obs.	489	115	324	489	115	324
Panel B: Dependent variable: Academic IAT						
	Surprise defined on the basis of					
	Academic Ability			*Altruistic Attitudes*		
Mixed Room	-0.093*	-0.053	-0.050	-0.093*	-0.019	-0.044
	(0.051)	(0.101)	(0.069)	(0.052)	(0.104)	(0.071)
Mixed Room*Positive Surprise	0.221**	0.586***	0.076	0.069	0.198	0.012
	(0.111)	(0.201)	(0.155)	(0.089)	(0.150)	(0.112)
Mixed Room*Negative Surprise	0.038	0.213	0.001	0.137	0.234	-0.018
	(0.096)	(0.220)	(0.141)	(0.099)	(0.181)	(0.202)
Mean of dep. var. in same race room	-0.204	-0.298	-0.180	-0.204	-0.298	-0.180
R-squared	0.103	0.340	0.098	0.100	0.308	0.097
No. Obs.	489	115	324	489	115	324

Notes: OLS estimates with robust standard errors in parentheses. * p<.10, ** p<.05, *** p<.01. Higher values of the dependent variable (IAT) indicate less prejudice against blacks. All regressions include university residence fixed effects, the dependent variable at baseline, controls and roommate controls as in table 4. In the full sample, we also control for respondent's race. In Cols.1 -3, "Positive surprise" is a dummy equal to one if a respondent's Academic IAT at baseline is lower than the mean and her black (white) roommate's UCT admission score is higher (lower) than the mean. "Negative surprise" takes value one if a respondent's Academic IAT at baseline is higher than the mean and her black (white) roommate's UCT admission score is lower (higher) than the mean. In Cols. 4-6, "Positive surprise" is a dummy equal to one if a respondent's Population IAT at baseline is lower than the mean and her black (white) roommate reported altruistic (non-altruistic) attitudes at baseline. "Negative surprise" takes value one if a respondent's Population IAT at baseline is higher than the mean and her black (white) roommate reported non-altruistic (altruistic) attitudes at baseline.

Table 6: Impact on academic performance

Dependent variable:	*GPA*			*Number of exams passed*			*Eligible to continue*			*Index of Performance*		
	Full sample	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Mixed Room	0.170*	-0.119	0.258**	0.463**	-0.364	0.715***	0.126***	0.033	0.169***	0.333***	-0.099	0.481***
	(0.096)	(0.244)	(0.122)	(0.198)	(0.561)	(0.252)	(0.031)	(0.067)	(0.042)	(0.110)	(0.270)	(0.145)
	[0.078]*	[0.848]	[0.029]**	[0.031]**	[0.848]	[0.007]***	[0.000]***	[0.848]	[0.000]***			
UCT entry score	7.826***	10.001***	5.616***	13.239***	9.943	11.595***	0.611*	-0.187	0.710	7.846***	7.653***	6.321***
	(1.085)	(2.566)	(1.321)	(2.270)	(6.763)	(2.940)	(0.362)	(0.773)	(0.483)	(1.218)	(2.425)	(1.595)
Roommate's UCT entry score	-0.149	1.882	-0.088	-0.434	0.938	-0.784	0.226	0.672	0.167	0.094	2.094	-0.001
	(0.679)	(1.621)	(0.875)	(1.297)	(3.605)	(1.759)	(0.184)	(0.443)	(0.250)	(0.712)	(1.719)	(0.976)
Controls	X	X	X	X	X	X	X	X	X	X	X	X
Roommate controls	X	X	X	X	X	X	X	X	X	X	X	X
Academic program FE	X	X	X	X	X	X	X	X	X	X	X	X
Mean of dep. var. in same race room	-0.018	0.673	-0.242	4.944	6.481	4.444	0.869	0.922	0.849	-0.031	0.727	-0.280
R-squared	0.421	0.562	0.388	0.704	0.702	0.714	0.319	0.408	0.416	0.442	0.405	0.452
No. Obs.	487	116	325	487	116	325	487	116	325	487	116	325

Notes: OLS Estimates with robust standard errors in parentheses. P-values adjusted for multiple inference are shown in squared brackets. Adjusted p-values are constructed using the resampling method of Westfall and Young (1993) with 10.000 interactions. * p<.10. ** p<.05. *** p<.01. Asterisks in the top row denote significance according to the robust standard errors; asterisks near the p-value denote significance after the FWER correction. All regressions include university residence. Controls and roommate controls are measured at baseline and are the same as in table 4. In cols.1, 4, 7 and 10 controls also include the race of the respondent (white, coloured and Indian/others) with black as the omitted category.

Table 7: Academic performance, heterogeneous effects by roommate's prejudice

Dependent variable:	*GPA*			*Number of exams passed in the first year*			*Eligible to continue*			*Index of Performance*		
Sample:	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full Sample*	*Whites*	*Blacks*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Mixed Room	0.306**	-0.124	0.507***	0.699***	-0.576	1.031***	0.115***	-0.028	0.186***	0.445***	-0.212	0.710***
	(0.127)	(0.214)	(0.191)	(0.250)	(0.671)	(0.337)	(0.040)	(0.083)	(0.059)	(0.138)	(0.247)	(0.205)
	[0.017]**	[0.769]	[0.009]***	[0.011]**	[0.717]	[0.006]***	[0.011]**	[0.769]	[0.006]***			
Mixed Room * Roommate pop IAT baseline	0.466**	0.635	0.660**	0.706*	0.200	0.835	0.055	0.040	0.092	0.474**	0.469	0.658*
	(0.201)	(0.442)	(0.319)	(0.381)	(1.108)	(0.567)	(0.058)	(0.121)	(0.106)	(0.214)	(0.485)	(0.355)
	[0.053]*	[0.254]	[0.097]*	[0.109]	[0.925]	[0.241]	[0.332]	[0.925]	[0.378]			
Roommate pop IAT baseline	-0.262**	-0.234	-0.329**	-0.408*	-0.329	-0.483	-0.085**	-0.025	-0.127**	-0.331**	-0.23	-0.434**
	(0.116)	(0.274)	(0.162)	(0.239)	(0.657)	(0.348)	(0.040)	(0.103)	(0.052)	(0.132)	(0.304)	(0.191)
Controls	X	X	X	X	X	X	X	X	X	X	X	X
Roommate controls	X	X	X	X	X	X	X	X	X	X	X	X
Academic program FE	X	X	X	X	X	X	X	X	X	X	X	X
Mean of dep. var. in same race room	-0.008	0.701	-0.233	4.955	6.644	4.407	0.874	0.932	0.853	-0.017	0.785	-0.277
R-squared	0.441	0.634	0.394	0.714	0.755	0.719	0.357	0.479	0.465	0.475	0.492	0.479
No. Obs.	364	85	248	364	85	248	364	85	248	364	85	248

Notes: OLS Estimates with robust standard errors in parentheses. Adjusted p-values are constructed using the resampling method of Westfall and Young (1993) with 10.000 interactions. * p<.10. ** p<.05. *** p<.01. Asterisks in the top row denote significance according to the robust standard errors; asterisks near the p-value denote significance after the FWER correction. All the regressions include university residence. Controls and roommate controls are measured at baseline and are the same as in Table 4. In cols.1, 4, 7 and 10 controls also include the race of the respondent (white, coloured and Indian/others) with black as the omitted category.

Table 8: Impact on friendships, attitudinal measures and pro-social behaviour

Dependent variable:	*Index of friendship*			*Index of attitudinal measures*			*Index of pro-social behavior*			*Global Index of social behavior*		
Sample:	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Mixed Room	0.575***	0.767***	0.437**	0.250**	0.433**	0.107	0.119	0.566*	0.297*	0.703***	1.314***	0.453*
	(0.159)	(0.227)	(0.214)	(0.106)	(0.216)	(0.142)	(0.137)	(0.290)	(0.171)	(0.199)	(0.405)	(0.268)
	[0.001]***	[0.004]***	[0.090]*	[0.054]*	[0.140]	[0.839]	[0.377]	[0.140]	[0.090]*			
Controls	X	X	X	X	X	X	X	X	X	X	X	X
Roommate controls	X	X	X	X	X	X	X	X	X	X	X	X
Mean of dep.var. in same race room	-1.469	-1.377	-1.496	-0.150	-0.806	0.054	-0.735	-0.907	-0.672	-1.941	-2.018	-1.915
R-squared	0.147	0.476	0.128	0.162	0.356	0.077	0.098	0.414	0.098	0.130	0.444	0.115
No. Obs.	366	93	273	449	105	296	384	93	250	279	78	201

Notes: OLS Estimates with robust standard errors in parentheses. P-values adjusted for multiple inference are shown in squared brackets. Adjusted p-values are constructed using the resampling method of Westfall and Young (1993) with 10.000 interactions. * p<.10. ** p<.05. *** p<.01. Asterisks in the top row denote significance according to the robust standard errors; asterisks near the p-value denote significance after the FWER correction. The indices used as dependent variables are constructed using a polychoric principal component analysis. The index of friendship (cols.1-3) includes the following variables: (i) no. of times respondent hung out with people of different race in the last month: =0 if never. =1 if once. =2 if 2-5 times. =3 if 5-10 times. =4 if more than 10 times; (ii) last time respondent hung out with people of different race: =0 if never. =1 if last year. =2 if last month. =3 if last week. =4 if yesterday; (iii) fraction of friends and study mates of a different race (excl. roommate); (iv) dummy for whether respondent desires >50% of members of different race in leisure group and academic group. The index of attitudes (cols.4-6) includes the following variables: (i) "In the last month. how often did you talk with any friends of yours about topics of discrimination. prejudice and racial bias?": =0 if never. =1 if rarely. =2 if sometimes. =3 if most of the time. =4 if always; (ii) a dummy for whether respondent is comfortable talking about race; (iii) a dummy for whether respondent does not agree that affirmative action in University admission should be abolished; (iv) a dummy for whether respondent do not feel conscious dancing with a person of another race; (v) dummy for whether a respondent does not feel conscious having a boyfriend/girlfriend of another race. The index of pro-social behavior (cols.7-9) includes the following variables: (i) member of community service or volunteer organization; (ii) amount of money given to charity in the past year; (iii) dummy for whether respondent cooperated in the prisoner dilemma game; (iv) dummy for whether respondent believed their partner would cooperate in prisoner's dilemma. The global index of social behavior (cols.10-12) includes all the variables listed in the previous three indexes. All regressions include residence fixed effects, individual controls and roommate controls as in Table 4.

Table 9: Residential choices at the end of the first year

Panel A: Dependent Variable = Still in Residence in year 2

Sample:	*Full sample*	*Whites*	*Blacks*
Mixed Room	0.047	0.044	0.038
	(0.039)	(0.064)	(0.055)
R-squared	0.188	0.151	0.219
No. Obs.	495	116	329

Panel B: T test - Same Roommate in year 2

	1 if in Mixed Room at baseline	*1 if not in Mixed Room at baseline*	*p-value*
Full sample	0.19	0.18	0.89
Whites	0.67	0.4	0.54
Blacks	0.16	0.21	0.64

Notes: Panel A reports OLS estimates for the full sample with robust standard errors in parentheses. * p<.10. ** p<.05. *** p<.01. All regressions include residence fixed effects, individual controls and roommate controls at baseline as in table 4. Panel B reports a t-test for the difference in means in the probability of having the same roommate in year 2. for the subsample of respondents who are still in a residence and in a double room in year 2.

Online Appendix – Not for publication

Figure A1: Definition of Positive and Negative Surprise

		Roommate's Race			
		Black		White	
		Roommate's type		**Roommate's type**	
		Low	*High*	*Low*	*High*
Individual ex-ante beliefs	*Low*	0	+	+	0
	High	–	0	0	–

Notes: Roommate's type are related to academic ability and altruistic attitudes. See more detailed description in the text.

Figure A2: GPA distribution, by race

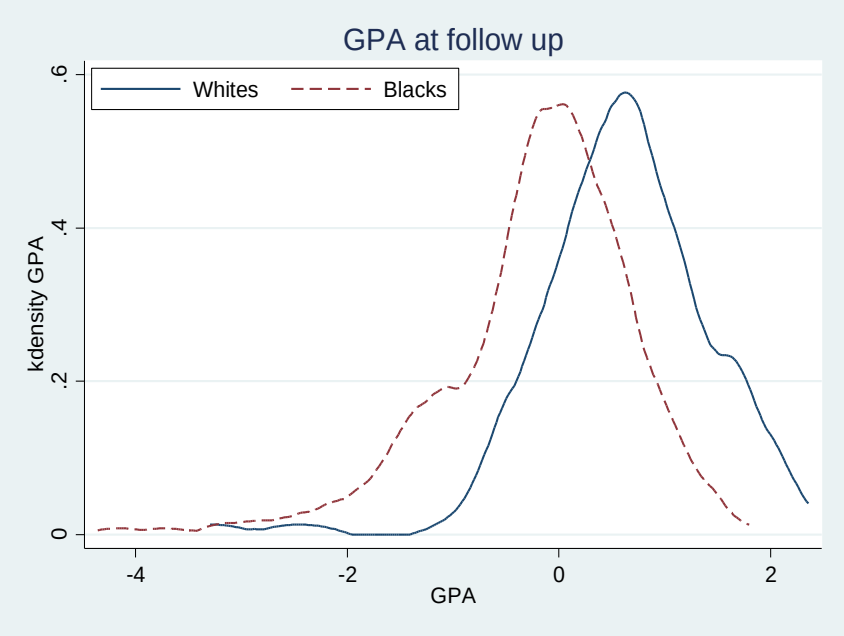


Table A1: Sample size and attrition

	All		*Mixed Room*		*Non Mixed Room*		*p-value*
	Obs	*%*	*Obs*	*%*	*Obs*	*%*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Baseline	625		203		422		
Follow-up	495	79.2	116	75.9	329	77.9	0.169

Notes: In column 7 we report the p-value for the difference in attrition between mixed and non-mixed rooms.

Table A2: Correlates of attrition

Dependent variable = 1 if respondent participated in follow-up survey

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mixed Room	-0.006			-0.021	-0.001	-0.006	-0.007
	(0.038)			(0.042)	(0.040)	(0.038)	(0.038)
Population IAT		0.000		0.022		-0.011	
		(0.032)		(0.037)		(0.036)	
Academic IAT			0.026		0.017		0.032
			(0.031)		(0.038)		(0.036)
Mixed Room*Population IAT				-0.066			
				(0.070)			
Mixed Room*Academic IAT					0.027		
					(0.063)		
White*Population IAT						0.092	
						(0.077)	
Coloured*Population IAT						-0.163	
						(0.168)	
Indian/Other*Population IAT						-0.064	
						(0.169)	
White*Academic IAT							-0.007
							(0.084)
Coloured*Academic IAT							0.099
							(0.162)
Indian/Other*Academic IAT							-0.136
							(0.105)
White	-0.149***	-0.150***	-0.148***	-0.147***	-0.147***	-0.114**	-0.147***
	(0.047)	(0.046)	(0.046)	(0.047)	(0.047)	(0.052)	(0.051)
Coloured	-0.173*	-0.177**	-0.172*	-0.169*	-0.165*	-0.204**	-0.129
	(0.093)	(0.088)	(0.088)	(0.092)	(0.093)	(0.098)	(0.108)
Indian/Other	0.013	0.010	0.012	0.013	0.016	-0.003	-0.016
	(0.075)	(0.070)	(0.071)	(0.075)	(0.076)	(0.087)	(0.083)
Female	-0.081	-0.082	-0.079	-0.077	-0.077	-0.086	-0.080
	(0.070)	(0.068)	(0.068)	(0.069)	(0.070)	(0.071)	(0.070)
UCT admission score	0.214	0.214	0.237	0.200	0.234	0.182	0.215
	(0.387)	(0.388)	(0.388)	(0.389)	(0.389)	(0.391)	(0.390)
Foreign	0.098*	0.098*	0.098*	0.099*	0.100*	0.104*	0.103*
	(0.057)	(0.057)	(0.058)	(0.058)	(0.058)	(0.057)	(0.057)
Private high school	-0.034	-0.034	-0.034	-0.034	-0.034	-0.036	-0.034
	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Durables p/c (PCA)	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.001
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)
Consumption	-0.043*	-0.043*	-0.043*	-0.042*	-0.043*	-0.044*	-0.045*
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
Constant	0.978***	0.977***	0.969***	0.981***	0.969***	1.001***	0.980***
	(0.196)	(0.197)	(0.196)	(0.197)	(0.197)	(0.197)	(0.198)
R-squared	0.092	0.092	0.093	0.094	0.094	0.097	0.096
No. Obs.	623	623	623	623	623	623	623

Table A3: Descriptive statistics at follow-up

	Full sample		*Mixed rooms*		*Non mixed rooms*	
	Mean	*Std. Dev.*	*Mean*	*Std. Dev.*	*Mean*	*Std. Dev.*
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Full Sample						
Population IAT	-0.19	0.50	-0.26	0.49	-0.17	0.50
Academic IAT	-0.22	0.44	-0.24	0.44	-0.21	0.45
Female	0.67	0.47	0.66	0.47	0.67	0.47
UCT admission score	0.46	0.05	0.47	0.05	0.46	0.05
Wealth Index	-0.04	2.10	-0.03	1.86	-0.05	2.20
Consumption	0.93	0.85	1.00	0.92	0.89	0.82
Foreign	0.11	0.31	0.14	0.34	0.10	0.30
Private high school	0.59	0.49	0.61	0.48	0.58	0.49
GPA	0.02	0.96	0.10	0.93	-0.02	0.97
No. Obs.	495		154		341	
Panel B: Whites						
Population IAT	-0.38	0.51	-0.30	0.48	-0.41	0.52
Academic IAT	-0.31	0.46	-0.35	0.40	-0.30	0.49
Female	0.67	0.47	0.64	0.49	0.69	0.47
UCT admission score	0.49	0.04	0.49	0.04	0.49	0.04
Wealth Index	0.75	1.78	0.40	1.27	0.93	1.98
Consumption	1.18	0.92	1.18	1.01	1.18	0.87
Foreign	0.07	0.25	0.10	0.31	0.05	0.22
Private high school	0.73	0.44	0.69	0.47	0.75	0.43
GPA	0.60	0.83	0.45	0.97	0.67	0.74
No. Obs.	116		39		77	
Panel C: Blacks						
Population IAT	-0.11	0.48	-0.20	0.49	-0.08	0.47
Academic IAT	-0.19	0.43	-0.22	0.43	-0.18	0.43
Female	0.68	0.47	0.73	0.45	0.67	0.47
UCT admission score	0.45	0.05	0.45	0.05	0.45	0.05
Wealth Index	-0.44	2.03	-0.57	1.57	-0.40	2.15
Consumption	0.81	0.80	0.89	0.91	0.79	0.77
Foreign	0.12	0.32	0.15	0.36	0.11	0.31
Private high school	0.53	0.50	0.54	0.50	0.53	0.50
GPA	-0.21	0.90	-0.08	0.79	-0.24	0.93
No. Obs.	329		74		225	

Notes: "UCT admission score" is the sum of high school final grades, with weights depending on the specific department the student enrolls in; the "Wealth index" measures per capita ownership of durable goods in the respondent's household and is constructed applying principal component analysis to the following categories of goods: computer, fridges, TV, landline and mobile phones, bicycles, motorbikes, bakkies, electricity, gas, kettles, geysers and cars; "Consumption" is the monthly consumption in Rands on lunches, dinners, food, alcohol, cigarettes, cell phone minutes, entertainment; "Foreign" is a dummy equal to one if the respondent is not from South Africa, “Private high school” is equal to one if the respondent was enrolled in a private high school before joining UCT.

Table A4: Stereotypes and exposure to a roommate of different race, no roommate controls

Dependent variable:	*Population IAT*			*Academic IAT*		
Sample:	*Full Sample*	*Whites*	*Blacks*	*Full Sample*	*Whites*	*Blacks*
	(1)	(2)	(3)	(4)	(5)	(6)
Mixed Room	-0.053	0.205*	-0.119*	-0.065	0.017	-0.056
	(0.053)	(0.124)	(0.066)	(0.045)	(0.098)	(0.056)
Controls	X	X	X	X	X	X
Mean of dep. var. in same race room	-0.166	-0.414	-0.084	-0.208	-0.298	-0.185
R-squared	0.109	0.151	0.067	0.071	0.179	0.056
No. Obs.	495	116	329	495	116	329

Notes: OLS estimates with robust standard errors in parentheses. * p<.10, ** p<.05, *** p<.01. Higher values of the dependent variable (IAT) indicate less prejudice against blacks. All control variables are measured at baseline. All regressions include university residence fixed effects and the dependent variable at baseline. (a) Controls include IAT at baseline, a dummy equal to one if the respondent is female, UCT admission score, wealth index, consumption, foreign, private high school, as defined in the footnote of Table 1. In cols. 1 and 4, controls also include the race of the respondent (white, colored and Indian/others) with black as the omitted category. (b) Controls for roommate include: UCT admission score, wealth index, consumption, foreign and private high school (female not included as all rooms are single sex).

Table A5: Stereotypes and exposure to a roommate of different race, sensitivity analysis

Dependent variable:	*Population IAT*			*Academic IAT*		
Sample:	*Full Sample*	*Whites*	*Blacks*	*Full Sample*	*Whites*	*Blacks*
	(1)	(2)	(3)	(4)	(5)	(6)
Mixed Room	-0.068	0.345***	-0.060	-0.036	0.079	-0.029
	(0.051)	(0.125)	(0.076)	(0.045)	(0.105)	(0.066)
Controls	X	X	X	X	X	X
Roommate controls	X	X	X	X	X	X
Mean of dep. var. in same race room	-0.172	-0.435	-0.074	-0.213	-0.296	-0.185
R-squared	0.110	0.384	0.131	0.084	0.397	0.109
No. Obs.	461	104	304	469	109	312

Notes: OLS estimates excluding influential observations with the DFBETA method. Robust standard errors in parentheses. * p<.10, ** p<.05, *** p<.01. Higher values of the dependent variable (IAT) indicate less prejudice against blacks. All control variables are measured at baseline. All regressions include university residence fixed effects and the dependent variable at baseline. (a) Controls include IAT at baseline, a dummy equal to one if the respondent is female, UCT admission score, wealth index, consumption, foreign, private high school, as defined in the footnote of Table 1. In cols. 1 and 4, controls also include the race of the respondent (white, colored and Indian/others) with black as the omitted category. (b) Controls for roommate include: UCT admission score, wealth index, consumption, foreign and private high school (female not included as all rooms are single sex).

Table A6: Impact on academic performance, no roommate controls

Dependent variable:	*GPA*			*Number of exams passed*			*Eligible to continue*			*Index of Performance*		
	Full sample	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Mixed Room	0.157*	-0.190	0.238**	0.438**	-0.161	0.676***	0.115***	0.031	0.145***	0.308***	-0.109	0.434***
	(0.090)	(0.232)	(0.109)	(0.188)	(0.485)	(0.227)	(0.030)	(0.068)	(0.039)	(0.103)	(0.269)	(0.130)
	[0.077]*	[0.395]	[0.025]**	[0.033]**	[0.863]	[0.004]***	[0.000]***	[0.863]	[0.000]***			
Controls	X	X	X	X	X	X	X	X	X	X	X	X
Academic program FE	X	X	X	X	X	X	X	X	X	X	X	X
Mean of dep. var. in same race room	-0.018	0.673	-0.242	4.944	6.481	4.444	0.869	0.922	0.849	-0.031	0.727	-0.28
R-squared	0.411	0.540	0.372	0.695	0.682	0.704	0.293	0.329	0.373	0.426	0.364	0.426
No. Obs.	487	116	325	487	116	325	487	116	325	487	116	325

Notes: OLS Estimates with robust standard errors in parentheses. P-values adjusted for multiple inference are shown in squared brackets. Adjusted p-values are constructed using the resampling method of Westfall and Young (1993) with 10.000 interactions. * p<.10. ** p<.05. *** p<.01. Asterisks in the top row denote significance according to the robust standard errors; asterisks near the p-value denote significance after the FWER correction. All regressions include university residence. Controls are measured at baseline and are the same as in table 4. In cols.1, 4, 7 and 10 controls also include the race of the respondent (white. coloured and Indian/others) with black as the omitted category.

Table A7: Impact on academic performance, sensitivity analysis

Dependent variable:	*GPA*			*Number of exams passed*			*Eligible to continue*			*Index of Performance*		
	Full sample	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Mixed Room	0.107	-0.116	0.262**	0.337*	-0.327	0.634**	0.044*	0.018	0.076**	0.231**	0.022	0.464***
	(0.087)	(0.180)	(0.120)	(0.173)	(0.314)	(0.268)	(0.025)	(0.041)	(0.038)	(0.096)	(0.182)	(0.135)
Controls	X	X	X	X	X	X	X	X	X	X	X	X
Rommate Controls	X	X	X	X	X	X	X	X	X	X	X	X
Academic program FE	X	X	X	X	X	X	X	X	X	X	X	X
Mean of dep. var. in same race room	0.045	0.687	-0.171	5.060	6.621	4.502	0.920	0.971	0.900	0.079	0.758	-0.171
R-squared	0.373	0.712	0.411	0.774	0.909	0.784	0.402	0.646	0.543	0.468	0.614	0.570
No. Obs.	453	96	292	449	95	296	451	100	293	452	100	292

Notes: OLS estimates excluding influential observations using the DFBETA method. Robust standard errors in parentheses. * p<.10, ** p<.05, *** p<.01. All regressions include university residence and academic program fixed effects. Controls are measured at baseline and are the same as in table 4. In cols.1, 4, 7 and 10 controls also include the race of the respondent (white. coloured and Indian/others) with black as the omitted category.

Table A8: Academic performance in the second year

Dependent variable:	*GPA*			*Number of exams passed*			*Eligible to continue*			*Index of Performance*		
Sample:	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Mixed Room	-0.003	-0.091	0.041	0.771***	0.557	0.875**	0.099**	-0.025	0.106*	0.278**	-0.045	0.365*
	(0.052)	(0.094)	(0.076)	(0.283)	(0.823)	(0.401)	(0.041)	(0.0769)	(0.062)	(0.137)	(0.308)	(0.213)
	[0.955]	[0.665]	[0.586]	[0.020]**	[0.723]	[0.078]*	[0.031]**	[0.729]	[0.162]			
UCT entry score	3.491***	7.348***	1.630**	3.675	-6.199	6.199	0.775*	0.636	0.762	6.395***	9.415***	4.392**
	(0.601)	(1.137)	(0.747)	(3.102)	(9.477)	(4.128)	(0.461)	(0.795)	(0.742)	(1.434)	(2.582)	(2.084)
Roommate's UCT entry score	0.456	0.663	0.756**	1.734	2.311	1.699	0.623*	0.678	0.658	1.762*	2.223	2.201**
	(0.312)	(0.634)	(0.318)	(1.800)	(4.396)	(2.430)	(0.352)	(0.460)	(0.486)	(0.984)	(1.613)	(1.104)
Controls	X	X	X	X	X	X	X	X	X	X	X	X
Roommate controls	X	X	X	X	X	X	X	X	X	X	X	X
Academic program FE	X	X	X	X	X	X	X	X	X	X	X	X
Mean of dep. var. in same race room	0.382	0.685	0.241	5.45	6.1	5.112	0.89	0.957	0.874	-0.073	0.545	-0.349
R-squared	0.505	0.788	0.448	0.534	0.658	0.59	0.264	0.519	0.378	0.448	0.657	0.475
No. Obs.	343	104	200	343	104	200	342	104	199	342	104	199

Notes: OLS Estimates with robust standard errors in parentheses. * p<.10. ** p<.05. *** p<.01.All the regressions include university residence. Controls and roommate controls are measured at baseline and are the same as in table 4. In cols. 1, 4, 7 and 10 controls also include the race of the respondent (white, coloured and Indian/others) with black as the omitted category. Adjusted p-values are constructed using the resampling method of Westfall and Young (1993) with 10.000 interactions. Asterisks in the top row denote significance according to the robust standard errors; asterisks near the p-value denote significance after the FWER correction.

Table A9: Academic performance, heterogeneous effect by same faculty

Dependent variable:	*GPA*			*Exams passed*			*Eligible to continue*			*Index of Performance*		
Sample:	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Mixed Room	0.186*	-0.04	0.214	0.371*	0.209	0.401	0.177***	0.028	0.209***	0.385	0.458	0.446
	(0.107)	(0.203)	(0.147)	(0.212)	(0.558)	(0.291)	(0.033)	(0.058)	(0.049)	(0.117)	(0.204)	(0.173)
	[0.133]	[0.930]	[0.229]	[0.133]	[0.930]	[0.229]	[0.000]***	[0.930]	[0.000]***			
MixRoom*Samefaculty	-0.107	-0.474	0.071	0.124	-2.586*	0.895*	-0.113	-0.036	-0.081	-0.174	-0.794	0.11
	(0.211)	(0.693)	(0.24)	(0.422)	(1.472)	(0.503)	(0.069)	(0.227)	(0.081)	(0.246)	(0.81)	(0.277)
	[0.807]	[0.702]	[0.766]	[0.807]	[0.192]	[0.178]	[0.231]	[0.876]	[0.485]			
Same Faculty	0.004	-0.063	-0.007	-0.088	0.584	-0.339	0.026	-0.091	0.022	0.017	-0.038	-0.039
	(0.111)	(0.289)	(0.14)	(0.24)	(0.481)	(0.324)	(0.041)	(0.093)	(0.048)	(0.136)	(0.315)	(0.169)
Controls	X	X	X	X	X	X	X	X	X	X	X	X
Roommate controls	X	X	X	X	X	X	X	X	X	X	X	X
Residence fixed effects	X	X	X	X	X	X	X	X	X	X	X	X
R-squared	0.43	0.572	0.388	0.7	0.72	0.716	0.34	0.432	0.441	0.45	0.433	0.463
No. Obs.	461	111	305	461	111	305	461	111	305	461	111	305

Notes: OLS Estimates with robust standard errors in parentheses. * p<.10. ** p<.05. *** p<.01. Adjusted p-values are constructed using the resampling method of Westfall and Young (1993) with 10.000 interactions. Asterisks in the top row denote significance according to the robust standard errors; asterisks near the p-value denote significance after the FWER correction All regressions include university residence and academic program fixed effects. Controls are measured at baseline and are the same as in Table 4.

Table A10: Impact on friendship

Dependent variable:	*# Times hang out with people of different race over past month*			*Last time hang out with people of different race*			*% of friends of a different race (excl. roommate)*		
Sample:	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mixed Room	0.788***	1.645**	0.710**	0.913***	1.602	0.752*	0.104***	0.114**	0.047
	(0.235)	(0.685)	(0.294)	(0.275)	(1.035)	(0.340)	(0.031)	(0.052)	(0.041)
	[0.006]	[0.111]	[0.102]	[0.006]	[0.497]	[0.141]	[0.002]	[0.111]	[0.536]
Controls	X	X	X	X	X	X	X	X	X
Roommate controls	X	X	X	X	X	X	X	X	X
R-squared							0.215	0.320	0.122
No. Obs.	477	110	317	476	110	316	458	109	303

Dependent Variable	*% of study-mates of a different race (excl. roommate)*			*=1 if desires >50% of members of different race in:*					
				Leisure group			*Academic group*		
Sample:	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Mixed Room	0.069**	0.136*	0.032	0.116**	0.255***	0.092	0.087*	0.115	0.107*
	(0.035)	(0.070)	(0.042)	(0.049)	(0.096)	(0.068)	(0.047)	(0.108)	(0.056)
	[0.080]	[0.111]	[0.536]	[0.080]	[0.036]	[0.536]	[0.183]	[0.646]	[0.178]
Controls	X	X	X	X	X	X	X	X	X
Roommate controls	X	X	X	X	X	X	X	X	X
R-squared	0.082	0.409	0.059	0.074	0.257	0.077	0.075	0.269	0.098
No. Obs.	390	96	294	479	111	319	479	111	318

Notes: Cols. 1-6 report ordered logit estimates; cols. 7-18 OLS estimates. Robust standard errors in parentheses. P-values adjusted for multiple inference are shown in squared brackets. Adjusted p-values are constructed using the resampling method of Westfall and Young (1993) with 10.000 iterations. * p<.10. ** p<.05. *** p<.01. Asterisks in the top row denote significance according to the robust standard errors; asterisks near the p-value denote significance after the FWER correction. All regressions include university residence fixed effects. Controls and roommate controls are measured at baseline and are the same as in table 4. "# times hang out more with people of different race in the last month": =0 if never. =1 if once. =2 if 2-5 times. =3 if 5-10 times. =4 if more than 10 times. "Last time hang out with people of different race": =0 if never. =1 if last year. =2 if last month. =3 if last week. =4 if yesterday.

Table A11: Impact on attitudinal measures

Dependent variable:	*Talked about race*			*Comfortable talking about race*			*Disagree to abolish affirmative action*			*Not conscious dancing with a person of another race*			*Not conscious having boyfriend/girlfriend of another race*		
Sample:	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Mixed Room	0.406*	-0.401	0.681**	0.088*	0.021	0.085	0.010	-0.106	0.034	0.095**	0.179**	0.026	0.091*	0.174	0.066
	(0.222)	(0.555)	(0.274)	(0.045)	(0.090)	(0.055)	(0.048)	(0.127)	(0.054)	(0.043)	(0.085)	(0.058)	(0.048)	(0.123)	(0.060)
	[0.291]	[0.957]	[0.022]	[0.190]	[0.999]	[0.417]	[0.994]	[0.777]	[0.873]	[0.057]	[0.077]	[0.873]	[0.190]	[0.159]	[0.417]
Controls	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Roommate controls	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
R-squared				0.081	0.216	0.130	0.231	0.161	0.092	0.118	0.391	0.100	0.237	0.427	0.134
Observations	453	107	320	442	105	290	448	106	294	446	105	294	450	107	295

Notes: Cols. 1-3 report ordered logit estimates; cols. 4-15 OLS estimates. Robust standard errors in parentheses. P-values adjusted for multiple inference are shown in squared brackets. Adjusted p-values are constructed using the resampling method of Westfall and Young (1993) with 10.000 iterations. * p<.10. ** p<.05. *** p<.01. All regressions include University residence fixed effects. All control and roommate controls are measured at baseline and are the same as in table 4. Talked about race (ordinal): "In the last month. how often did you talk with any friends of yours about topics of discrimination and racial bias?" 1 Never. 2 Rarely. 3 Sometimes. 4 Most of the time. 5 Always. Comfortable talking about race: =1 if comfortable talking to people about race. Disagree to abolish Affirmative Action: = 1 if does not agree that that affirmative action in university admission should be abolished. Not conscious dancing with a person of another race: = 1 if respondent does not feel conscious dancing with a person of another race. Not conscious having a boyfriend/girlfriend of another race = 1 if respondent does not feel conscious having a boyfriend/girlfriend of another race.

Table A12: Impact on pro-social behaviour

Dependent variable:	*Member of Volunteer Organization*			*Money given to a charity*			*Cooperate in Prisoner dilemma*			*Belief partner will cooperate in prisoner dilemma*		
Sample:	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Mixed Room	0.097*	0.249**	0.075	91.312	121.989	22.700	0.046	0.230*	0.090	0.037	0.078	0.103
	(0.057)	(0.110)	(0.075)	(73.120)	(344.743)	(41.702)	(0.055)	(0.120)	(0.070)	(0.053)	(0.128)	(0.068)
	[0.320]	[0.084]	[0.792]	[0.513]	[0.976]	[0.931]	[0.639]	[0.105]	[0.356]	[0.639]	[0.536]	[0.130]
Controls	X	X	X	X	X	X	X	X	X	X	X	X
Roommate controls	X	X	X	X	X	X	X	X	X	X	X	X
R-squared	0.057	0.334	0.066	0.110	0.249	0.125	0.080	0.377	0.074	0.044	0.236	0.036
No. Obs.	463	109	309	401	98	257	489	113	327	48	113	327

Notes: OLS Estimates with robust standard errors in parentheses. P-values adjusted for multiple inference are shown in squared brackets. Adjusted p-values are constructed using the resampling method of Westfall and Young (1993) with 10.000 iterations. * p<.10. ** p<.05. *** p<.01. All the estimates include university residence fixed effects. Controls and roommate controls are measured at baseline and are the same as in table 4. In cols 7-12. controls also include a dummy indicating whether the respondent knows the partner in the game.