Title

Interracial contact at work: Does workplace diversity reduce bias?

Journal
Group Processes & Intergroup Relations, 24(7)

ISSN
1368-4302

Authors
Darling-Hammond, Sean
Lee, Randy T
Mendoza-Denton, Rodolfo

Publication Date
2021-10-01

DOI
10.1177/1368430220932636

Article

Group Processes & Intergroup Relations

GPIR

Group Processes & Intergroup Relations
1–18

DOI: 10.1177/1368430220932636
journals.sagepub.com/home/gpi
SAGE

Interracial contact at work: Does workplace diversity reduce bias?

Sean Darling-Hammond,[1] Randy T. Lee[2] and Rodolfo Mendoza-Denton[1]

Abstract

Research suggests that anti-Black bias among White Americans is persistent, pervasive, and has powerful negative effects on the lives of both Black and White Americans. Research also suggests that intergroup contact in workplaces can reduce bias. We seek to address two limitations in prior research. First, the workplaces reviewed in prior studies may not be typical. Second, previously observed relationships between workplace contact and bias may stem from selection bias—namely, that White individuals who tend to work with Black individuals are systematically different from those who do not, and those systematic differences explain lower bias levels. To address these issues, we review records (N = 3,359) of White, non-Hispanic, working adults in a nationally representative survey to examine the relationship between workplace contact and racial closeness bias after adjusting for an exhaustive set of potential confounders. Using propensity score matching, we compare individuals who work with Black individuals with their "virtual twins"—individuals who have the same propensity of working with Black individuals but do not. We estimate that having a Black coworker causes a statistically significant reduction in racial closeness bias for White, non-Hispanic adults.

Keywords

affective bias, explicit bias, intergroup contact, racial bias, racial closeness bias, workplace diversity

Paper received 23 April 2019; revised version accepted 17 May 2020.

> "The single most promising arena of racial integration—at least for adults—is the workplace."
>
> Cynthia Estlund (2003, p. 9)

A broad and growing body of research suggests that intergroup contact can reduce bias (e.g., Pettigrew, Tropp, Wagner, & Chris, 2011). At the same time, polling suggests that many White adults do not experience intergroup contact in their friend groups—as many as 40% of White Americans do not have a single non-White friend (Dunsmuir, 2013). Notably, that figure drops by

[1]University of California, Berkeley, USA
[2]Cornell University, Ithaca, USA

Corresponding author:
Sean Darling-Hammond, Goldman School of Public Policy, University of California, 2607 Hearst Ave, Berkeley, CA 94720, USA.
Email: sean.darling.hammond@berkeley.edu

10 percentage points when "workplace acquaintances" are included. Against this backdrop, Estlund (2003) has argued that workplaces represent the most critical source of intergroup contact for many adults, and that, because workplaces encourage common goals and cooperation, contact in workplaces may reduce bias.

However, while some studies have found evidence that intergroup contact in workplaces reduces bias (Pettigrew & Tropp, 2006), prior literature is dated, and the workplaces assessed in prior literature may not be typical. Moreover, research shows that intergroup contact in contexts where individuals feel threatened can *increase* bias (Pettigrew et al., 2011). Workplaces may elicit feelings of threat, and workplace contact could increase bias. Given the gap in available literature, we aim to ascertain whether intergroup contact in a typical workplace decreases bias.

We focus here on "racial closeness bias," a form of explicit bias whereby a person indicates feeling "closer" to members of their own race than to members of other races. We focus on racial closeness bias because, like the biases evaluated in research showing explicit bias is related to circulatory disease death rates (Leitner, Hehman, Ayduk, & Mendoza-Denton, 2016) and racial disparities in school disciplinary actions (Riddle & Sinclair, 2019), it is affective (or measures generalized emotions about a group). Thus, evaluating the causal impact of having a Black coworker on this form of bias will allow more natural bridging to other research. We focus on the biases of White, non-Hispanic individuals because, collectively, they represent a majority of individuals in the United States (U.S. Census Bureau, 2020) and because they are sufficiently represented in available data to ensure rigorous analyses.

Prior Research on Intergroup Contact in Workplaces

Psychologists have long studied mechanisms for reducing bias. Allport (1954) suggested that intergroup contact can, under appropriate conditions, reduce bias. In his seminal work, Allport found that contact best reduces bias when four features are present: (a) equal status of the groups in the situation, (b) common goals, (c) intergroup cooperation, and (d) the support of authorities, laws, or customs. Pettigrew et al. (2011) conducted a meta-analysis to test this intergroup group contact theory, reviewing over 500 studies with over a quarter million subjects. They found that contact was associated with lower bias in 94% of studies, and that contact was associated with larger declines in bias when most of Allport's conditions were present. The meta-analysis also found that contact in situations where the participants felt threatened led to increases in bias. Building on this research, Paluck, Green, and Green (2018) reviewed all 27 intergroup contact studies published at the time of their review that involved random assignment and delayed outcome measures. Consistent with Pettigrew et al. (2011), they found that contact generally appeared to cause lower levels of later bias, but that the strength of contact effects varied substantially.

Estlund (2003) has argued that workplaces often exhibit Allport's contact-effect-enhancing conditions, and that increasing workplace diversity could therefore be an effective strategy for reducing interracial bias. Four studies support Estlund's position. Brophy (1945) found that White merchant seamen who went on more voyages with Black seamen developed genuine bonds and had less bias. Kephart (1957) similarly found that White police officers in Philadelphia who worked with Black colleagues expressed fewer objections to Black individuals joining their previously all-White police districts, to teaming with a Black partner, and to taking orders from qualified Black officers.

In a widely cited study on the topic, Cook (1984) reported results from a controlled experiment testing the effects of contact on bias. He recruited White, "highly prejudiced" women and had them work alongside two confederate coworkers—one Black and one White—as railway operators. At multiple points, the Black confederate described a personal experience with racial discrimination and the White confederate expressed disapproval of the discrimination.

After the experiment, participants in the experimental condition reported more positive attitudes about Black people than they had expressed in baseline surveys, and these gains were larger than those experienced by women in the control condition (who did not work as railway operators). Effects lasted at least several months. Finally, Pettigrew and Tropp (2006) found that contact in workplaces was consistently associated with reductions in bias—and that it was associated with larger reductions in bias than contact in educational, residential, and travel settings (although contact in laboratory and recreational settings was associated with even larger reductions). While the meta-analysis does not provide detailed information about the nature of the workplaces evaluated, a review of identifiable studies suggests that workplaces included those described before (marine ships, railway stations, and police departments) and also included a military base (Butler & Wilson, 1978).

Limitations of Prior Research

While these results are encouraging, the workplaces analyzed in prior studies may systematically differ from the "typical" workplace in ways that enhance the effect of contact on bias. Mariners, police officers, and military personnel may have stronger incentives to get along with their peers than typical workers. Effective cooperation may help protect them from the dangers posed by turbulent seas, armed suspects, and enemy soldiers. Such threats may provide an atypically strong incentive to overcome biases to work productively with colleagues. It seems worth investigating whether contact may not shift bias in the same manner in other workplace contexts that may lack similar incentives for cooperation.

Moreover, while the railway study involved random selection into treatment (exposure to a Black coworker), the other studies did not account for the possibility of treatment selection effects. Generally, selection effects refer to systematic differences between treated and control groups which may drive differences in outcomes. For example, perhaps it is not that certain White Philadelphia police officers became less biased by working with Black partners; instead, some third factor (such as political attitudes or age) may predict both selecting a Black partner and being less biased. One class of selection effects is "reverse directionality"—that those who seek the treatment are predisposed to have different levels on the outcome. For example, it may be that less biased officers tend to seek out Black partners, and more biased officers tend to avoid them. Research on college student friend groups by Sidanius, Levin, van Laar, and Sears (2008) and Binder et al. (2009) lends credence to the possibility of reverse directionality. They found that while a prior measure of college students' intergroup friendships (contact) predicted later bias, a prior measure of bias also predicted a later measure of contact. It is thus unclear whether findings from observational studies on workplace contact reflect selection effects or the impact of contact.

Finally, in the railway study, the "treatment" was not natural intergroup contact with a typical coworker, but exposure to two scripted confederate coworkers where one was Black and one was White. While it is encouraging that these scripted, cross-group interactions appeared to cause stable reductions in bias, it remains unclear whether unscripted, typical intergroup interactions in typical workplaces would yield the same effects.

In summary, prior research has not evaluated contact effects in typical workplaces; all but one of the prior workplace contact studies appear vulnerable to selection bias; and the sole randomized controlled trial on workplace contact involved scripted, rather than natural, intergroup encounters. Extant research thus leaves a hole in our understanding of the relationship between workplace contact and bias. The present research attempts to provide new clarity.

The Present Research

To ascertain if contact in a typical workplace causes reductions in racial bias, we leveraged 12 years of geocoded data from the General Social Survey (GSS). The GSS is a nationally representative survey created and regularly collected by the University of Chicago since 1972. In each survey year, GSS interviewers ask thousands of adults (18 or older) around the country a battery of questions regarding life experiences, workplace characteristics, and attitudes. By investigating a large subsample of working, White, non-Hispanic adults drawn from this nationally representative dataset, we can assess (a) whether contact with Black individuals in "typical" workplace environments is associated with lower levels of bias and (b) whether any such association persists after adjusting for relevant confounders. Next, so long as we are persuaded that we have a sufficiently robust set of covariates on which to match individuals in our sample, we can attempt to ascertain (c) whether contact *causes* reductions in bias.

Estimating causal effects using cross-sectional data is a complex endeavor and the legitimacy of estimates invariably rests on researcher assumptions. As we discuss more thoroughly in what follows, we assume here that the wide array of variables in the GSS provide sufficient clarity regarding whether and why individuals had a Black coworker to ascertain the causal effect of having one. However, even if one does not agree entirely with this assumption, this research still holds value. Compared to prior observational studies, it provides a cleaner estimate of the relationship between real-world workplace contact and bias.

Relevant Confounders

In order to isolate the relationship between contact and bias, we must first identify potential confounders. By definition, a confounder is a variable that is correlated with both the treatment and the outcome of interest (Agresti, 2018). We adjust for confounders because failure to do so confounds, or confuses, our understanding of the relationship between treatment and outcome. As explained by Brookhart et al. (2006), when building models to predict the relationship between a treatment and an outcome, there are many approaches one might take for determining which potential confounders to include. One approach is to only include confounders if they are statistically significantly related to both the treatment and the outcome. By setting a high bar for including variables in the model (and adjusting for them), this approach risks underadjustment of bias and can return overstated causal estimates.

A more conservative approach is to include as a confounder any variable that is a significant predictor of either the treatment or the outcome, even if it is only marginally related to the other of the two. In this case, that would mean including any variable that is a significant predictor of *either* having a Black coworker or of racial bias. The benefit of this approach is that it adjusts away more bias. The drawback is that it can result in larger standard errors, increasing the risk of a Type II error.

Our goal, and we believe the goal of any researcher using observational, cross-sectional data to estimate contact effects, is to overcome as much potential confoundedness as possible and approximate a causal estimate, even if doing so risks Type II errors. We therefore take, and recommend, the more conservative approach. Based on our analyses of relationships between variables in the GSS, we argue that assessments of the relationship between workplace contact and bias should adjust for each of the 11 confounders depicted in Table 1.

Next, we provide additional justification for the general type of confounder that each variable belongs to. We provide more general justifications in the hopes that doing so will increase the utility of this analysis to other types of contact research.

Other types of contact. Intergroup contact theory suggests that various forms of contact predict lower levels of bias (Pettigrew et al., 2011). Certain forms of contact are also correlated with one another. For example, it is not hard to imagine that workplaces located in diverse neighborhoods are more able to hire diverse workforces. Individuals in those neighborhoods may thus be more likely to experience intergroup contact both in their communities *and* workplaces. As depicted in Table 1, we found evidence for this proposition, $r(3739) = .18$, $p < .001$. Thus, we argue that to

Table 1. Relationships between each confounder and having a Black coworker/racial closeness bias.

	Black coworker		Racial closeness bias	
	r / F	p	r / F	p
Black neighbor**	$r(3739) = .18$	$< .001$	$r(6049) = -.12$	$< .001$
Conservatism**	$r(3831) = -.04$	.008	$r(6192) = .08$	$< .001$
Age*	$r(3907) = -.05$	$< .001$	$r(6326) = .01$	.275
Education**	$r(3912) = .12$	$< .001$	$r(6335) = -.03$	.022
Female*	$r(3913) = -.02$	.179	$r(6339) = .03$	.009
Family income*	$r(3595) = .05$	.001	$r(5681) = -.02$	.162
Marital status**	$F(4, 3909) = 3.54$	.007	$F(4, 6333) = 4.82$	$< .001$
Social class*	$F(3, 3898) = 6.65$	$< .001$	$F(3, 6312) = 1.51$	.210
Commuting zone**	$F(190, 3724) = 2.73$	$< .001$	$F(192, 6147) = 2.09$	$< .001$
Occupation classification**	$F(24, 3890) = 7.90$	$< .001$	$F(24, 6316) = 1.60$	.033
Year**	$F(6, 3908) = 2.92$	.008	$F(6, 6334) = 4.68$	$< .001$

Note. The r / F column depicts the Pearson's correlation (r) or Fischer test statistic (F). In parentheses, the column also depicts the number of degrees of freedom and, in the case of F statistics, the number of categories, minus 1. In general, the correlation coefficient (r) ranges from −1 to 1, and F ranges from 0 to infinity. In both cases, values further from 0 indicate a greater degree of relatedness.
p column depicts the probability of obtaining an r or F statistic of a given magnitude under the null hypothesis of no relatedness.
*p values for either Black coworker or racial closeness bias are less than .05. **p values for both Black coworker and racial closeness bias are less than .05.

the extent that one hopes to ascertain the relationship between a single form of contact and bias, adjustment for other forms of contact is essential.

Social attitudes. Certain social attitudes, such as political conservatism, are correlated with racial bias. Neal (2017) found that more conservative individuals are less likely to believe racism is a major problem in America. To the extent that views about racism are related to views about Black individuals, these individuals may also be more biased, and we indeed found that conservatism and bias are correlated in the GSS, $r(6192) = .08, p < .001$. We offer here an important cautionary note about "bad controls." When trying to predict the relationship between a "treatment" and an "outcome," a "bad control" is a variable that is *impacted by* (rather than simply predictive of) the "treatment" (Angrist & Pischke, 2015). Such controls should not be included in models, as including them biases assessments of the relationship between treatment and outcome. Here, our "treatment" is contact, and our "outcome" is bias. Many social attitudes are likely correlated with contact and bias. However, many of these attitudes are arguably *measures* of bias, which themselves would be impacted by contact and which, therefore, would be "bad controls." For example, the most relevant social attitude for predicting workplace contact might be an individual's response to the question "How much do want to work with Black individuals?" However, this could arguably be considered a measure of bias, or at least a reflection of it. If our hypothesis is that contact shifts bias, and we believe this measure is also a reflection of bias, then we would expect that contact will also shift responses to this measure, rendering it a "bad control." We strongly recommend against including this and similar "bad controls" in models predicting the relationship between forms of contact and bias. Our inclusion of conservatism here represents our belief, supported by recent research (see e.g., Hatemi & Verhulst, 2015), that political attitudes are largely stable and thus unlikely to change due to workplace contact with Black individuals.

Sociodemographic characteristics. Modern research suggests that certain demographics can predict certain forms of contact. Reviewing data from users of its dating application, OKCupid found that White women were much more likely than White men to have romantic relationships with Black individuals (OKCupid, 2014). In the GSS, meanwhile, we see that contact is statistically significantly related to each of the following: age, $r(3{,}907) = -0.05$, $p < .001$; education, $r(3{,}912) = 0.12$, $p < .001$; family income, $r(3{,}595) = 0.05$, $p < .001$; marital status, $F(4, 3909) = 3.54$, $p = .007$; and social class, $F(3, 3898) = 6.65$, $p < .001$. Sociodemographic characteristics may also predict bias levels. We found, for example, that racial closeness bias is statistically significantly related to education, $r(6{,}335) = -0.03$, $p = .022$; sex, $r(6{,}339) = 0.03$, $p = .009$; and marital status, $F(4, 6333) = 4.82$, $p < .001$. We suggest inclusion of sociodemographic characteristics in models predicting the relationship between contact and bias for two reasons. First, as discussed before, they are frequently predictors of contact and bias and thus are often strong candidates for confounders. Second, and perhaps more importantly, they are stable characteristics which should not be impacted by contact, and thus are unlikely to be "bad controls." Thus, there is minimal, if any, risk that including these variables will engender bias in a model predicting the relationship between contact and bias.

Spatial and temporal characteristics. A great deal of both contact and bias is driven by the characteristics of the places we call home. Underscoring the power of spatial characteristics, Leitner et al. (2016) and Riddle and Sinclair (2019) found substantial variation in county-level anti-Black explicit bias (which they linked to deleterious social phenomena). In the GSS, we find that where a person lives (as defined by their commuting zone, or the multicounty job market they belong to) is a statistically significant predictor of both having a Black coworker, $F(190, 3724) = 2.73$, $p < .001$, and racial bias, $F(192, 6147) = 2.09$, $p < .001$. We thus suggest adjustment for spatial characteristics in models ascertaining the relationship between contact and bias. Just like the place we call home, the *time* we live in exerts a large effect on contact and on bias. We find evidence of this effect in the GSS, with survey year being a statistically significant predictor of having a Black coworker, $F(6, 3908) = 2.92$, $p = .008$, and racial closeness bias, $F(6, 6334) = 4.68$, $p < .001$.

Predictors of the specific form of contact. As noted before, we include as a confounder any variable that significantly predicts either treatment or outcome. Thus, we recommend that any variables that predict a specific form of contact be included and adjusted for in models designed to ascertain the impact of a given form of contact on bias. A prime example from the GSS is occupational classification, which describes the sector an individual works in, and which likely predicts whether or not a person has a Black coworker. For example, the military is known to be remarkably diverse (Barroso, 2019), and White individuals in military-specific occupations would be expected to be more likely to have Black coworkers than individuals in many other sectors. Workplace classification is a statistically significant predictor of having a Black coworker in the GSS, $F(24, 3890) = 7.90$, $p < .001$.

Confounders Included in Models

For the aforementioned reasons, in our model predicting the relationship between workplace contact and bias, we include the following potential confounders, subdivided by confounder type:

- **Other forms of contact**: has a Black neighbor.
- **Social attitudes:** political conservativism.
- **Sociodemographic characteristics:** respondent age, educational attainment, sex, marital status (married, widowed, divorced, separated, never married), social class (lower class, working class, middle class, upper class), and family income.

- **Spatial and temporal characteristics:** residential commuting zone and year of GSS interview.
- **Predictors of contact:** occupation classification.

While these variables do not represent an exhaustive list of every possible variable of each type, together, we believe they account for the core of the confoundedness we hope to adjust for.

Analysis Plan

Our analysis proceeds in three steps. First, we will conduct a bivariate regression of racial closeness bias on having a Black coworker. Next, we will conduct a multivariate regression of racial closeness bias on having a Black coworker and each of the 11 confounders discussed before. We hypothesize that, consistent with contact theory, having a Black coworker will be associated with statistically significantly less pro-White bias, and that the relationship will endure even after adjusting for the potential confounders discussed before.

In the third step, we will use a method known as propensity score matching (PSM) to attempt to ascertain the causal effect of having a Black coworker on racial closeness bias. In order to assess whether contact *causes* lower levels of bias, we need to develop a better understanding of what distinguishes each individual who has a Black coworker from those that do not. Using the GSS's rich data about life experiences, attitudes, and workplace characteristics, we can develop a model that predicts how likely each individual in our data was to have a Black coworker. We can then create pairs of virtual "twins" comprised of one individual who had a Black coworker and another individual who had the same propensity of having a Black coworker but, by dint of luck, did not have one. By seeing the average difference in bias between individuals with a Black coworker and their virtual "twin," we can estimate the impact of having a Black coworker on bias.

Relative to multivariate regression analysis, PSM has two distinct advantages. First, by matching treated individuals with control individuals who are similar on covariates, you create "apples to apples" comparisons that are less sensitive to imbalance issues between treated and control groups. Second, matching-based estimations are less sensitive to functional form assumptions. In other words, in the case of regression, our estimate would be biased if we were to omit even a single relevant interaction term or fail to account for any nonlinear relationships between variables. In contrast, a matching-based estimate is less reliant on getting these functional form aspects of the model right.

Using this "apples to apples" comparative approach, we hypothesize that matched individuals with a Black coworker will have statistically significantly lower levels of bias than similarly situated, matched individuals without a Black coworker.

Mechanics of Propensity Score Matching

PSM, as operationalized here, proceeds in three steps: propensity score generation, matching, and comparison. First, we use logistic regression (logit) to predict the likelihood of receiving treatment (having a Black coworker) based on relevant neighborhood as well as personal and workplace characteristics. Specifically, our logit model includes the confounders described before. We plug each individual's actual scores on covariates into the logit model to predict their unique p score, or probability of having a Black coworker. Second, we match each treated individual to the one control individual (one to one) with the closest p score (their nearest neighbor) to create a matched dataset for "apples to apples" comparison. We do so without replacement, meaning we do not allow a given control case to be matched as a control for multiple treated cases. It is worth noting that matching with replacement may be appropriate in instances where matching without replacement fails to address potential sources of bias from confounders. However, because matching with replacement involves utilizing a single control case as a match for multiple treated cases, it can engender situations where a small number of repeatedly utilized control cases overwhelmingly skew the data. Fortunately, we find evidence (discussed in the pages that follow) that, in our data, matching without replacement reduces observable sources of bias to negligible levels and is therefore adequate. This type of matching (one to one, nearest neighbor, without replacement) is quite commonly utilized and finds support in a range of published research on PSM (see e.g., Austin, 2011a).

Finally, to glean the effect of treatment, we compare mean bias levels of the treated and control individuals in our matched dataset. And to conduct hypothesis tests, we conduct a two independent samples *t* tests to ascertain whether there is a statistically significant difference between our matched treated and matched control groups in their mean levels of bias.

Assumptions Underlying Propensity Score Matching

The difference-in-means statistic derived from PSM is considered an unbiased estimate of the causal effect of treatment on outcome so long as two conditions are satisfied. First, and most importantly, the logit model predicting treatment must include all confounders. Given the rich set of confounders available in the GSS and described before, we are persuaded that our model approaches this condition. However, as noted before, we are (as all researchers using PSM should be) mindful that some confoundedness lingers. Via these methods and data, we can only at best *approach* a causal estimate. We believe, however, that the confoundedness that remains is marginal enough that our failure to account for it will not meaningfully bias our estimate. The second condition is that the PSM model must have overlap, which is to say that if a treated individual has a given propensity of treatment, there must be at least one control individual with a very similar propensity of treatment who can serve as a comparator, or virtual twin. We demonstrate that this condition is satisfied in Appendix B.

As discussed before, the goal of matching is to achieve balance in covariates. As Table 2 depicts, prior to matching, there is a great deal of imbalance in two numerical variables (has a Black neighbor and education) and moderate imbalance in three others (age, conservatism, and family income). A multivariate regression would fail to account for these imbalances. However, via one-to-one, nearest neighbor, matching without replacement, standardized differences in average values between treated and control cases diminish markedly to negligible levels.

Method

Treatment Measure

The GSS asks respondents to describe their workplaces as "all White," "mostly White," "half White, half Black," "mostly Black," or "all Black." Using responses from this question, we constructed a binary treatment measure indicating whether a given person has Black coworkers (1) or does not have Black coworkers (0).

Outcome Measure

The GSS also asks, on 9-point scales, "In general, how close do you feel to Whites?" and "In general, how close do you feel to Blacks?" Our measure of "pro-White bias" is the difference between scores on these two questions (-8 = *extreme pro-Black bias*, 8 = *extreme pro-White bias*).

Participants

Participants include 3,359 White, non-Hispanic, working adults (age 18 or higher) who participated in General Social Survey (GSS) interviews between 2002 and 2014 and who had scores on each relevant variable. Importantly, there were 12,651 White, non-Hispanic, working adults in the geocoded GSS from 2002 to 2014. However, among these, only 3,887 provided answers to the questions that comprise our treatment and outcome measures. Moreover, because our goal was to isolate the impact of treatment on outcome by creating "apples to apples" comparisons, we also needed to ensure participants had provided

Table 2. Covariate means and standardized differences in treated and control groups for unmatched sample (U) and matched sample (M).

Variable	Sample	Treated	Control	Standardized difference
Has a Black neighbor	U	0.75	0.56	0.40
	M	0.59	0.56	0.05
Age	U	43.74	45.32	0.12
	M	45.45	45.32	0.01
Years of education	U	14.62	13.88	0.28
	M	13.88	13.88	0.00
Conservatism (1–7)	U	4.08	4.26	0.13
	M	4.35	4.26	0.06
Female	U	0.47	0.50	0.05
	M	0.51	0.50	0.02
Family income	U	$45,510	$40,534	0.13
	M	$42,978	$40,534	0.06

Note. Two samples are perfectly balanced if the standardized difference between mean values on all numeric covariates is zero. Standardized differences range from zero to infinity. Standardized differences below .1 are generally considered to be indicators of very good balance.

answers to questions related to the relevant confounder measures described before. Our analysis sample (N = 3,359) represents the subset of White, non-Hispanic, working adults that provided answers to questions related to our treatment, outcome, and confounder measures.

When one starts with a representative sample and experiences this kind of sample attrition, it presents three potential problems: (a) power—too small of a sample size, and thus too large of standard errors to detect any statistically significant relationships between treatment and control that might exist in the population; (b) representativeness—an analysis sample that is systematically different from the larger sample; and (c) bias—an analysis sample that exhibits an uncharacteristically, and therefore deceptively, large relationship between treatment and outcome. Overall, our analysis sample is not underpowered (as we will show in the Results section), and, as explained more fully in Appendix A, our analysis sample appears relatively representative of the larger GSS sample (which, itself, is designed to be representative of the United States). Notably, however, our analysis sample is, on net, meaningfully younger, more educated, less likely to be female, and wealthier than the full sample. In terms of potential bias occasioned by attrition, the relationship between treatment and outcome appears nearly identical in the analysis and full samples, suggesting that the attrition required to conduct covariate adjustment will not introduce unanticipated bias.

Results

Descriptive Findings

As depicted in Figure 1, a large percentage of White individuals in our analysis sample (about 50%) exhibited some degree of pro-White bias, or a score above zero; 48% exhibited "no bias," or a score of zero; and less than 3% exhibited what might be termed "pro-Black bias," or scores lower than zero.

Regression Analyses

As depicted in Table 3, having a Black coworker statistically significantly ($p < .001$) predicted lower levels of pro-White bias in both a bivariate regression (without confounders) and a multivariate regression (including confounders). In the bivariate model, having a Black coworker predicted a bias score 0.57 points lower ($N = 3{,}359$, $p < .001$,

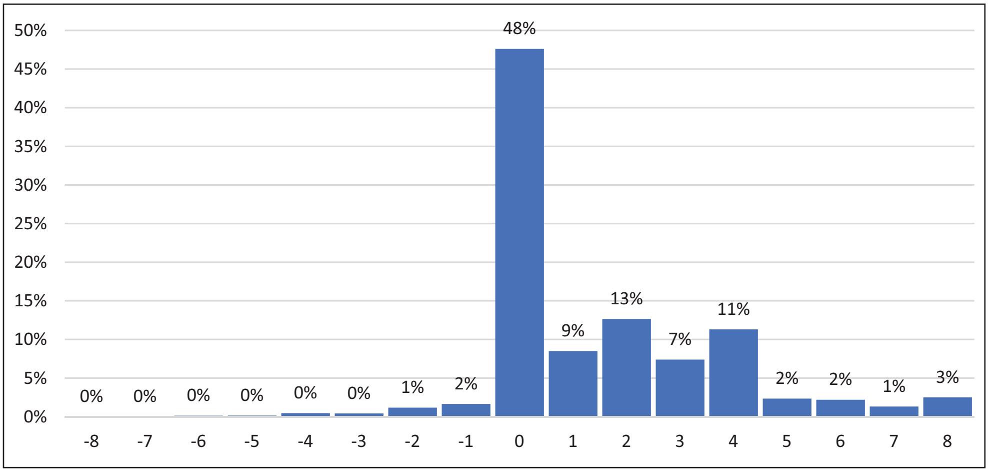

Figure 1. Distribution of pro-White bias scores among individuals in the analysis sample.

Table 3. OLS regressions predicting pro-White bias.

Model	One	Two†
Black coworker	−.57*** (0.08)	−0.34*** (0.09)
Black neighbor		−0.28** (0.09)
Conservatism		0.18*** (0.03)
Age		−0.01 (0.00)
Educational attainment		−0.02 (0.02)
Female		0.10 (0.09)
Marital status: Widowed		0.34 (0.23)
Marital status: Divorced		−0.11 (0.11)
Marital status: Separated		−0.50* (0.25)
Marital status: Never married		−0.02 (0.11)
Social class: Working class		0.22 (0.24)
Social class: Middle class		0.29 (0.24)
Social class: Upper class		0.12 (0.33)
Family income		−0.00 (0.00)
Constant	1.85*** (0.07)	1.36 (0.89)

Note. Coefficients and related standard errors are shown, with the latter within parentheses.
†Fixed effects were included for commuting zones, employment sector category, and year of survey.
***$p < .001$. **$p < .01$. *$p < .05$.

$d = 0.27$), and in the multivariate model, having a Black coworker predicted a bias score 0.34 points lower ($N = 3{,}359$, $p < .001$, $d = 0.16$).

Consistent with Pettigrew and Tropp (2006), these results indicate a negative relationship between workplace contact and bias. But is there a negative causal relationship? The next section discusses an initial attempt at answering this question.

Propensity Score Matching Analysis to Ascertain the Causal Relationship Between Workplace Contact and Bias

The aforementioned multivariate regression is sensitive to two major limitations. First, a multivariate regression of the kind employed cannot ensure "apples to apples" comparisons, which is to say the sample of "treated" individuals (those with Black coworkers) may differ in material ways from the sample of "control" individuals to which they are compared (those without Black coworkers). In addition, the multivariate regression may be biased to the extent that it fails to account for nonlinear relationships between variables or for interaction effects. Thus, in extreme cases, even when a multivariate regression includes all relevant confounders, it can yield a remarkably biased estimate of the causal effect of treatment (Austin, 2011a).

Propensity Score Matching-Based Causal Estimation

PSM is a curative measure for the issues described before. As noted previously, propensity scores are unbiased so long as they include all confounders and have overlap in propensity score values between treated and control individuals. Among our 2,443 treated individuals (those who had Black coworkers), propensity scores ranged from 0.068 to 0.996. Among the 911 control individuals (those who did not have Black coworkers), propensity scores ranged from 0.026 to 0.976. Thus, even in the most extreme case, treated individuals with propensity scores at the higher end were able to match with control cases with meaningfully similar propensity scores. For example, the treated case with the highest propensity score matched with a control case whose propensity score was less than 0.02 p-score units away.

Via PSM, we end up with a sample of 1,822 individuals comprised of 911 pairs of treated individuals with meaningful counterfactual controls ("virtual twins"). As explained before (see Table 2), our matched sample is extremely balanced across numerical predictors, and markedly more balanced than our unmatched sample. This suggests that conducting PSM substantially diminished potential sources of bias. In this new sample, comparing mean outcomes between treated and control individuals, we glean a causal estimate of -0.45 ($p < .001$, $d = 0.21$). As depicted in Appendix C, we also ran a myriad of robustness checks demonstrating that our point estimates are consistently negative when using a range of less common, but also valid, methods for determining matches.

Discussion

The present research attempted to leverage 12 years of geocoded GSS data to ascertain whether working with Black individuals is associated with or causes lower levels of bias among White individuals. We first found that about half of the White, working adult Americans in our sample exhibited some degree of pro-White bias. Using OLS regressions, we found that, controlling for a myriad of confounders, White individuals who had with Black coworkers had statistically significantly ($p < .001$) lower average pro-White bias scores than their counterparts who did not have Black coworkers. Specifically, after adjusting for a range of confounders, those who had Black coworkers had, on average, 0.34 points lower bias scores than those who did not.

We next recruited PSM to estimate the causal effect of working with a Black individual on bias. In our model, having a Black coworker appeared to cause a reduction in bias of about 0.45 points, relative to not having one, and the effect was statistically significant ($p < .001$). As noted repeatedly, while these results are certainly exciting, they should be taken with a grain of salt. Any PSM model is only unbiased so long as it includes all confounders. While we believe we have come meaningfully close to this goal such that we can glean a meaningfully unbiased estimate, we do not believe we have included every possible confounder in our model.

Even given this caveat, prior research had not established that intergroup contact in a "typical" workplace was associated with lower levels of bias. This research suggests it is. Prior research also had not established whether intergroup contact in workplaces causes reductions of bias. This research provides initial support for the notion that interracial contact in workplaces causes statistically significant reductions in bias.

However, even assuming this research has correctly estimated the causal effect of workplace contact on bias as a reduction of 0.45 points (or thereabouts), what remains unclear is what this reduction in bias portends in terms of other outcomes of interest. For example, could a reduction of 0.45 points, repeated across a sufficient number of individuals, stem the negative association between bias and circulatory death rates for Black individuals and White individuals observed in Leitner et al.'s (2016) research? Future research could attempt to answer this and related questions and help tease out the social meaning of the causal estimate we present here.

In addition, it remains unclear to what extent participants in this study worked in settings that exhibited Allport's contact-effect-enhancing conditions of equal status, common goals, cooperation, and support from authorities. Future research could evaluate whether, as Estlund (2003) claims, workplaces exhibit these factors and whether, as Allport (1954) would predict, workplaces that exhibit more of these factors demonstrate larger contact effects. More generally, future research could ascertain whether contact effects differ by workplace characteristics.

Finally, like any attempt at causal estimation, this research would benefit from attempts to corroborate, or refute, its findings. For example, researchers might use an instrumental variable approach to assess whether phenomena that increase workplace diversity (such as receiving a government-funded workplace diversity grant) catalyze reductions in bias. Or they might look to panel data and use difference-in-difference or event study designs to ascertain whether shifts in workplace contact lead to concomitant or subsequent shifts in bias.

Conclusion: Segregation at Work, and the Work Ahead

As discussed in our introduction, many White American adults do not have Black individuals in their social networks. Thus, barring some other mechanism for intergroup contact, they will not experience this critical debiasing phenomenon. Polling data show that workplaces can be an important source of intergroup contact (Dunsmuir, 2013), and this research suggests that intergroup contact in workplaces can, indeed, reduce bias. These results come at an important moment. Sophisticated spatial research by Ferguson and Koning (2018) suggests that between-workplace segregation (e.g., the number of workplaces that are largely homogenous) has actually increased so much that it is higher today than it was in the 1970s. It is possible that fewer White Americans are experiencing the debiasing effects of workplace contact.

We thus believe that investments in further research regarding workplace contact should be paired with meaningful, strategic efforts to increase workplace contact. For example, governmental, private sector, and philanthropic organizations could fund efforts to widen pipelines for people of color to work in largely White workplaces, or could create incentives for White individuals to work with people of color. These efforts could yield long- and short-term benefits. In the long term, these efforts might provide a clearer, more actionable lens into the power of workplace contact to reduce bias. In the short term, they may help create a more connected society.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Sean Darling-Hammond https://orcid.org/0000-0002-6353-4670

References

Agresti, A. (2018). *Statistical methods for the social sciences*. Harlow, UK: Pearson Education.

Allport, G. (1954). *The nature of prejudice*. Cambridge, MA: Addison-Wesley.

Angrist, J. D., & Pischke, J. (2015). *Mastering metrics: The pathway from cause to effect*. Princeton, NY: Princeton University Press.

Austin, P. C. (2011a). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, *46*, 399–424. https://doi.org/10.1080/00273171.2011.568786

Austin, P. C. (2011b). Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical Statistics*, *10*, 150–161. https://doi.org/10.1002/pst.433

Austin, P. C., Grootendorst, P., & Anderson, G. M. (2007). A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: A Monte Carlo study. *Statistics in Medicine*, *26*, 734–753. https://doi.org/10.1002/sim.2580

Barroso, A. (2019). *The changing profile of the U.S. military: Smaller in size, more diverse, more women in leadership*. Retrieved from Pew Research Center website: https://www.pewresearch.org/fact-tank/2019/09/10/the-changing-profile-of-the-u-s-military/

Binder, J., Zagefka, H., Brown, R., Funke, F., Kessler, T., Mummendey, A., . . .Leyens, J. P. (2009). Does contact reduce prejudice or does prejudice reduce contact? A longitudinal test of the contact hypothesis among majority and minority groups in three European countries. *Journal of Personality and Social Psychology*, *96*, 843–856. https://doi.org/10.1037/a0013470

Brookhart, M., Schneeweiss, S., Rothman, K., Glynn, R., Avorn, J., & Stürmer, T. (2006). Variable section for propensity score models. *American Journal of Epidemiology*, *163*, 1149–1156. https://doi.org/10.1093/aje/kwj149

Brophy, I. (1945). The luxury of anti-Negro prejudice. *The Public Opinion Quarterly*, *9*, 456–466. https://doi.org/10.1086/265762

Butler, J., & Wilson, K. (1978). "The American soldier" revisited: Race relations and the military. *Social Science Quarterly*, *59*, 451–467. Retrieved from https://www.jstor.org/stable/42860376

Cook, S. (1984). Cooperative interaction in multiethnic contexts. In N. Miller & M. B. Brewer (Eds.), *Groups in contact: The psychology of desegregation* (pp. 155–185). Cambridge, MA: Academic Press.

Dunsmuir, L. (2013, August 8). *Many Americans have no friends of another race: Poll.* Retrieved from https://www.reuters.com/article/us-usa-poll-race/many-americans-have-no-friends-of-another-race-poll-idUSBRE97704320130808

Estlund, C. (2003). *Working together: How workplace bonds strengthen a diverse democracy*. New York, NY: Oxford University Press.

Ferguson, J. P., & Koning, K. (2018). Firm turnover and the return of racial establishment segregation. *American Sociological Review*, *83*, 445–474. https://doi.org/10.1177/0003122418767438

Hatemi, P. K., & Verhulst, B. (2015). Political attitudes develop independently of personality traits. *PLoS ONE*, *10*. https://doi.org/10.1371/journal.pone.0118106

Kephart, W. (1957). *Racial factors and urban law enforcement*. Philadelphia: University of Pennsylvania Press.

Leitner, J. B., Hehman, E., Ayduk, O., & Mendoza-Denton, R. (2016). Blacks' death rate due to circulatory diseases is positively related to Whites' explicit racial bias: A nationwide investigation using Project Implicit. *Psychological Science*, *27*, 1299–1311. https://doi.org/10.1177/0956797616658450

Neal, S. (2017). *Views of racism as a major problem increase sharply, especially among Democrats*. Retrieved from Pew Research Center website: http://www.pewresearch.org/fact-tank/2017/08/29/views-of-racism-as-a-major-problem-increase-sharply-especially-among-democrats/

OKCupid. (2014). *Race and attraction, 2009–2014: What changed in five years?* Retrieved from https://theblog.okcupid.com/race-and-attraction-2009-2014-107dcbb4f060

Paluck, E. L., Green, S. A., & Green, D. (2018). The contact hypothesis re-evaluated. *Behavioral Public Policy*, *3*, 129–158. https://doi.org/10.1017/bpp.2018.25

Pettigrew, T., & Tropp, L. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*, 751–783. https://doi.org/10.1037/0022-3514.90.5.751

Pettigrew, T., Tropp, L. R., Wagner, U., & Chris, O. (2011). Recent advances in intergroup contact theory. *International Journal of Intercultural Relations*, *35*, 271–280. https://doi.org/10.1016/j.ijintrel.2011.03.001

Riddle, T., & Sinclair, S. (2019). Racial disparities in school-based disciplinary actions are associated with county-level rates of racial bias. *Proceedings of the National Academy of Sciences of the USA*, *116*, 8255–8260. https://doi.org/10.1073/pnas.1808307116

Sidanius, J., Levin, S., van Laar, C., & Sears, D. O. (2008). *The diversity challenge: Social identity and intergroup relations on the college campus*. New York, NY: Russel Sage Foundation.

United States Department of Agriculture (USDA). (2012). *Commuting zones and labor market areas*. Retrieved from https://www.ers.usda.gov/data-products/commuting-zones-and-labor-market-areas/

U.S. Census Bureau. (2020). United States of America. Retrieved from https://data.census.gov/cedsci/profile?q=United%20States&g=0100000US&tid=ACSDP1Y2018.DP05

Appendix A

Here we explain analyses conducted to determine that the sample loss occasioned by adding confounders to our model did not result in representativeness issues or bias.

Representativeness issues. To ascertain the representativeness of our analysis sample, we utilize a common method for comparing mean characteristics of samples (see e.g., Austin, Grootendorst, & Anderson, 2007). Table 1A summarizes the results of this analysis, showing, for each numeric variable and for both the full sample and the analysis sample, the number of individuals who provided information on a given measure and the mean values on that measure. It also shows the standardized difference between means on each measure. Standardized differences are generally below 0.1, suggesting that the analysis sample is relatively representative of the overall sample. However, in four cases, standardized differences are above 0.1, indicating that the analysis sample is, on net, meaningfully younger, more educated, less likely to be female, and wealthier than the full sample.

While representativeness concerns are certainly worthy of consideration, it is important to put them in perspective. Most analyses of the relationship between contact and bias involve contact by specific individuals in specific contexts, which, by its very nature, generates unrepresentative results. For example, the widely cited "railway study" by Cook (1984) was limited to 84 White, college-age, female individuals. Our results certainly suffer from representativeness challenges, but ones that are nowhere near as pronounced as one would expect in a typical causal test of the intergroup contact theory.

Bias. We might glean a biased estimate of the relationship between treatment and outcome if the relationship between treatment and outcome were stronger in the analysis sample than in the full sample. In our case, we see very similar relationships between treatment and outcome in our analysis and total samples. In the larger sample, the correlation between treatment and

Table 1A. Mean values on outcome, treatment, and numerical covariates for total GSS sample of White, non-Hispanic, working adults as compared to analysis sample.

Variable	Total responding (total sample)	Mean (total sample)	Mean (analysis sample)	Standardized difference
Feel closer to Whites	6,341	1.45	1.44	0.00
Has a Black coworker	3,915	0.72	0.73	0.01
Has a Black neighbor	9,160	0.67	0.70	0.07
Age	12,618	49.68	44.10	0.36
Years of education	12,638	13.89	14.39	0.18
Conservatism (1–7)	10,353	4.20	4.14	0.04
Female	12,651	0.54	0.48	0.12
Family income	11,301	$38,464	$43,772	0.14

Note. While individuals in the analysis sample are less likely to have a Black coworker than individuals in the total sample, the groups are otherwise largely indistinguishable from one another. Moreover, individuals in both samples are remarkably likely to have a Black coworker, making the two relatively comparable even on this measure.
GSS = General Social Survey.

outcome is $r(3887, p < .001) = -.12$, while in our analysis sample, the correlation is $r(3359, p < .001) = -.12$. To ascertain if these correlation coefficients are meaningfully distinct, we conduct a Fischer Z-transformation on both correlation coefficients, take the difference of the two, and conduct a hypothesis test on the null hypothesis that the difference between the two is zero. Using this process, we found that our z-transformed correlation coefficients are, respectively, $z = 0.12$ and $z = 0.12$, the difference between the two is 0.00203, and the two-tailed p value associated with the null hypothesis that this difference is truly zero (i.e., that the two correlations are identical) is 0.99. At any level of statistical significance testing, we cannot reject the null hypothesis that these two correlation coefficients are identical. This suggests that by limiting ourselves to this smaller sample we have not biased our estimates.

Appendix B

A necessary condition for conducting propensity score matching (PSM) is what is known as "common support." In essence, for each treated individual (with their given propensity score), we must be sure we can find at least one control individual with a similar propensity score who can serve as a counterfactual comparator. The most common means to visually demonstrate that this threshold is met is to graph two density functions side by side: one for propensity scores for treated individuals and one for propensity scores among control individuals. This provides a mechanism for identifying propensity score ranges in which one might have treated individuals but no similar comparators. As Figure 1A demonstrates, at each range point at which we have treated individuals, we also have control individuals who can serve as comparators.

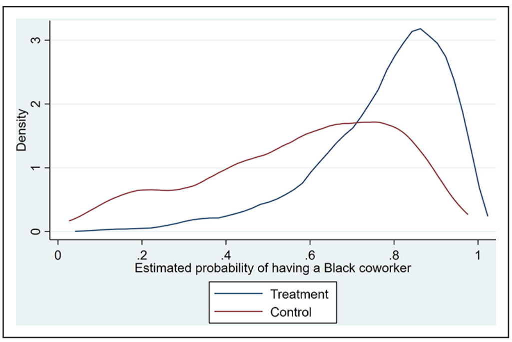

Figure 1A. Kernel density functions for propensity scores of treated and untreated individuals.
Note. Technically, there is an infinitesimal range between 0.9864 and 1 at which there are treated individuals but not control individuals. There are still, however, control individuals with remarkably similar propensity scores who can serve as a comparator.

Appendix C

Matching is a general term referring to a number of approaches for constructing a counterfactual control sample for a given treated (or control) sample. The most common matching approach is "one-to-one, nearest neighbor, no replacement, propensity score matching," which is the method employed in this research. However, other methods exist, and one robustness check is to ascertain if causal estimates vary depending on the matching method used to construct the counterfactual sample. As depicted in Table 2A, the overall result (that interracial contact in workplaces causes reductions in bias) is generally robust to common matching specifications, with 11 out of 11 returning negative point estimates. Moreover, seven returned a statistically significant ($p < .05$) result, and two more yielded a marginally statistically significant ($p < .1$) result.

In addition, we ascertain whether results are sensitive to our choice to execute matching without setting a caliper. A caliper is a propensity score range around which the PSM will allow a match. In our model, we allow each control case to be matched to the nearest treated case regardless of how far away that treated case is. As depicted in Table 3A, we found that our estimate is not sensitive to whether we use a caliper and, if so, what size we utilize. In all cases, the causal estimate is negative and statistically significant ($p < .001$).

Table 2A. Comparison of results from 11 separate matching estimations.

	N	Difference	*t*-statistic	*p* value
Radius: Default Caliper***	3,252	−0.55	−11.85	$< .001$
Nearest neighbor, no replacement***	1,822	−0.45	−4.26	$< .001$
Doubly robust regression***	1,822	−0.44	−4.03	$< .001$
Kernel: Normal**	3,252	−0.36	−3.05	.002
Mahalanobis**	3,359	−0.43	−2.67	.008
Kernel: Uniform*	3,252	−0.31	−2.48	.013
Kernel: Epanechnikov*	3,252	−0.28	−2.14	.032
Kernel: Tricube♦	3,252	−0.26	−1.93	.054
Local linear regression: Normal♦	3,252	−0.24	−1.67	.095
Radius: Caliper = 2 *SD*	3,252	−0.23	−1.56	.119
Nearest neighbor, with replacement	3,252	−0.05	−0.28	.779

Note. Sorted by *t*-statistic value.
♦$p < .1$. *$p < .05$. **$p < .01$. ***$p < .001$.

Table 3A. Comparison of results from one-to-one, nearest neighbor, no replacement matching with various caliper options.

	N	Difference	*t*-statistic	*p*
No caliper***	1,822	−0.45	−4.26	$< .001$
Caliper = 0.1***	1,677	−0.42	−3.66	$< .001$
Caliper = 0.01***	1,677	−0.42	−3.75	$< .001$
Caliper = 0.001***	1,562	−0.41	−3.37	$< .001$

*$p < .05$. **$p < .01$. ***$p < .001$.

Appendix D

In this section, we provide the STATA code utilized to conduct the anlyses discussed above.

*Preparing STATA for big datasets and analyses, loading data, and creating bulk of variables needed for analysis:

set maxvar 30000
set matsize 10000
use "C:\Users\Administrator\Desktop\Articles, Presentations, Applications, Consulting\Contact and Racial Attitudes, GSS, Project Implicit\Data and Analysis\Full GSS 2002-2014 Geocoded.dta"
gen white_non_hispanic = 1 if race == 1 & hispanic == 1
replace white_non_hispanic = 0 if race > 1 | hispanic > 1
gen closer_white = closewht - closeblk
gen black_coworker = 0 if racwork == 1
replace black_coworker = 1 if racwork > 1 & racwork < 6
gen black_neighbor = 1 if raclive == 1
replace black_neighbor = 0 if raclive == 2
gen age_recode = age - 0
gen education_years = educ - 0
gen conservatism = polviews - 0
gen female = sex - 1
gen marital_status = marital - 0
gen family_income = realinc - 0
gen social_class = class - 0

*Here, we wrap Federal Information Processing Standards (FIPS) county codes into commuting zones based on United States Department of Agriculture's 2000 commuting zone and labor market crosswalk (USDA, 2012). We do not include related code as it spans dozens of pages. It is available upon request as a STATA do file.

*Next, we wrap 2010 Census occupation codes (census.gov) up into occupation classifications based on "Industry and Occupation Code: Lists and Crosswalks" (available at https://www.census.gov/topics/employment/industry-occupation/guidance/code-lists.html). We do not include related code as it also spans many pages. It is available upon request as a STATA do file.
*Now we restrict our sample to White, non-Hispanic, working adults who provided responses to all relevant measures:
keep if closer_white != . & black_coworker != . & black_neighbor != . & age_recode != . & education_years != . & conservatism != . & female != . & marital_status != . & social_class != . & family_income != . & year != . & commuting_zone != . & occupation_classification != . & white_n == 1 & age > 17 & occupation_classification < 9999
*Bivariate and multivariate regressions:
reg closer_white black_coworker
reg closer_white black_coworker black_neighbor conservatism age_recode education_years female i.marital_status i.social_class family_income i.year i.commuting_zone i.occupation_classification
*Running Logit to predict probability of having a Black coworker for White, non-Hispanic working adults:
quietly logit black_coworker black_neighbor age_recode education_years conservatism female i.marital_status i.social_class family_income i.year i.commuting_zone i.occupation_classification
predict prop_treatment
*Checking region of common support:
kdensity prop_treatment if black_coworker==1, addplot(kdensity prop_treatment if black_coworker==0) legend(label(1 "treatment") label (2 "control"))
*Ascertaining region of common support numerically:
summ prop_t if black_c == 0
summ prop_t if black_c == 1
*Estimating average effect of treatment on the treated (ATT):
psmatch2 black_coworker, outcome(closer_white) pscore(prop_treatment) neighbor(1) noreplacement
*Running balance test, and determining improvement on balance occasioned by PSM:
pstest black_neighbor age_recode education_years conservatism female family_income, both
*Robustness checks using a number of other matching methods. First, doubly robust:
reg closer_w black_c black_n age_re education_y conservatism female family_income i.marital_sta i.social_c i.commuting_z i.occupation_c i.year if _weight == 1
*Now matching with replacement:
psmatch2 black_coworker, outcome(closer_white) pscore(prop_treatment) neighbor(1)
*Now, radius with default caliper size:
psmatch2 black_coworker, radius outcome(closer_white) pscore(prop_treatment)
*Radius with caliper = .2 * *SD*(*p* score), consistent with Austin (2011b) "Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies":
psmatch2 black_coworker, radius caliper(.02940574) outcome(closer_white) pscore(prop_treatment)
*Kernel matching via different kinds of kernels:
psmatch2 black_coworker, kernel outcome(closer_white) kerneltype(normal) pscore(prop_treatment)
psmatch2 black_coworker, kernel outcome(closer_white) kerneltype(epan) pscore(prop_treatment)
psmatch2 black_coworker, kernel outcome (closer_white) kerneltype(uniform) pscore(prop _treatment)
psmatch2 black_coworker, kernel outcome(closer_white) kerneltype(tricube) pscore(prop_treatment)
*Local linear regression:
psmatch2 black_coworker, llr outcome(closer_white) kerneltype(normal) pscore(prop_treatment)
*Mahalanobis matching:
psmatch2 black_coworker, mahalanobis(education _years i.social_class black_neighbor female i.marital_status conservatism family_income age_recode i.commuting_zone i.occupation_cla i.year) outcome(closer_white)
*Going back to the original model, determing how influenced the model is by the decision to match *every* control case to a treated case, regardless of *p*-score distance. Seeing how results may be affected by caliper decisions:
psmatch2 black_coworker, outcome(closer_white) pscore(prop_treatment) neighbor(1) noreplacement caliper(.1)

psmatch2 black_coworker, outcome(closer_white) pscore(prop_treatment) neighbor(1) noreplacement caliper(.01)
psmatch2 black_coworker, outcome(closer_white) pscore(prop_treatment) neighbor(1) noreplacement caliper(.001)
*In all cases, the ATT hovers around −.4 and is statistically sgnificant $p < .001$. Adding a caliper is unnecessary:
*A final note: one could conduct a very similar analysis using the nongeocoded GSS from the same time period, but would have to swap the "commuting zone" variable out for the GSS variable "region," which indicates the region in which the interview took place. This would certainly yield an underadjusted estimate but could serve as a proxy or proof of the methods described herein. When we do so, we find that:
*The bivariate regression, of course, yields the same estimate ($b = -0.57, p < .001$).
*The multivariate regression now yields a slightly larger estimate than before, still statistically significant ($b = -0.43, p < .001$).
*The PSM estimate is ever so slightly smaller, still statistically significant ($b = -0.42, p < .001$).