Racial demographics explain the link between racial disparities in traffic stops and county-level racial attitudes

Pierce D. Ekstrom[a], Joel M. Le Forestier[b], Calvin K. Lai[c]

Last Modified: 8/15/2021

Citation: Ekstrom, P. D., Le Forestier, J. M., & Lai, C. K. (In Press). Racial demographics explain the link between racial disparities in traffic stops and county-level racial attitudes. *Psychological Science*. https://doi.org/10.31234/osf.io/znhuf

Forthcoming in *Psychological Science*

Acknowledgments: We thank Christopher Federico, George Wood, and Rafael Aguilera for their comments on an earlier draft of this manuscript.

[a] Please direct correspondence concerning this article to Pierce Ekstrom.
Email: pierce.ekstrom@gmail.com
University of Nebraska – Lincoln, Department of Political Science
Address: 511 Oldfather Hall, Lincoln, NE 68588

[b] University of Toronto, Department of Psychology
Address: 4th Floor, Sidney Smith Hall, 100 St. George Street, Toronto, ON M5S 3G3

[c] Washington University in St. Louis, Department of Psychological and Brain Sciences
Address: Somers Family Hall, One Brookings Drive, St. Louis, MO 63130

Abstract

Disparities in the treatment of Black and White Americans in police stops are pernicious and widespread. We examine racial disparities in police traffic stops by leveraging data on traffic stops from hundreds of U.S. counties from the Stanford Open Policing Project and corresponding county-level data on implicit and explicit racial attitudes from the Project Implicit research website. We find that Black-White traffic stop disparities are associated with county-level implicit and explicit racial attitudes and that this association is attributable to racial demographics: counties with a higher proportion of White residents had larger racial disparities in police traffic stops. We also examined racial disparities in several post-stop outcomes (e.g., arrest rates) and found that they were not systematically related to racial attitudes, despite evidence of disparities. These findings indicate that racial disparities in counties' traffic stops are reliably linked to counties' racial attitudes and demographic compositions.

Keywords: Race, Attitudes, Policing, Geographic Analysis

Statement of Relevance

Millions of Americans are stopped by police every year. Prior research indicates that stops are inequitably concentrated among Black civilians, who are also more likely than White civilians to suffer serious legal, physical, or fatal consequences following the stop. We investigate regional variation in racial attitudes as one potential correlate of racial disparities in traffic stops, reasoning that region-level attitudes may reflect the region's culture and its institutions in ways that manifest in police behavior. We find that racial disparities in a county's traffic stops are pervasive and related to the racial attitudes and the racial demographics that prevail in a county. Our findings indicate a need to consider the psychological and structural correlates of racial disparities in police traffic stops.

Racial demographics explain the link between racial disparities in traffic stops and county-level racial attitudes

Americans' experiences with police traffic stops are distressing, common, and disproportionately concentrated among Black civilians. In a survey of North Carolina drivers, 22% reported having been stopped by police at least once in the past year, with 18% of White drivers and 26% of Black drivers reporting at least one stop (Warren et al., 2006). More recent nationwide analyses indicate that this racial disparity in stop rates is the rule rather than the exception (Davis et al., 2018), and is unlikely to be attributable to racial differences in traffic offense rates (Pierson et al., 2020). This experience can be particularly harrowing for Black drivers, who may fear that even "routine" stops could spiral out of control with life-changing or fatal consequences (Epp et al., 2014).

Recent psychological research has used macro-level variables—including *geographic* variation in explicit and implicit racial attitudes—to help explain racial disparities. County-level variation in racial attitudes has been found to predict Black-White racial disparities in infant health (Orchard & Price, 2017), school disciplinary practices (Riddle & Sinclair, 2019), mortality rates due to circulatory disease (Leitner et al., 2016a) and mortality rates in general (Leitner et al., 2016b). This work has even documented an association between regional variance in racial attitudes and the disproportionate use of lethal force by police against Black people (Hehman et al., 2018). In the current study, we examine whether county-level racial attitudes also predict the disproportionate rates at which Black drivers are stopped by police.

Racial attitudes at the county level include components that are absent at the individual level. To understand an individual's racial attitude, one only needs to assess the direction and intensity of that one individual's attitude. Meanwhile, counties are comprised of many

individuals who vary in race, gender, age, and other characteristics. These individuals aggregate into subgroups that vary in their racial attitudes. To understand a county's racial attitudes, then, one must account for the *direction and intensity* of various subgroups' racial attitudes and for the *prevalence* of each subgroup within the county. We have two goals in the current study. First, we examine the relation between stop disparities and county-level attitudes. Second, we explore which specific components of county-level attitudes best predict stop disparities: the direction and intensity of subgroups' racial attitudes, subgroups' prevalence, or other systematic regional variance in racial attitudes.

We theorize that county-level racial attitudes reflect the attitudes of governing officials or of community members, the culture that prevails within law enforcement agencies, the racial attitudes of individual police officers, or some combination of all of these. In turn, cultural norms and the attitudes of community members, officers, and public officials may manifest in local laws and law enforcement priorities, the areas that officers are instructed to patrol, and how willing officers are to stop Black vs. White drivers (Alpert et al., 2005; Beckett et al., 2006; Roh & Robinson, 2009). The current study seeks to examine the plausibility of these mechanisms but does not have the resolution to causally isolate them from each other. From our perspective, county-level racial attitudes may be uniquely useful for understanding racial disparities in police traffic stops *because* they can capture the diverse range of individual, institutional, and cultural biases that prevail within that region.

To examine the relation between regional racial attitudes and disparities in police traffic stops, we combine traffic-stop data from the Stanford Open Policing Project with racial-attitude data from Project Implicit. We predicted that Black-White stop disparities would be most pronounced in counties with the highest anti-Black/pro-White bias.

Method

Stanford Open Policing Project (SOPP)

All variables pertaining to traffic stops were derived from a version of the Stanford Open Policing Project Dataset (SOPPD; Pierson et al., 2020) retrieved in October 2019, available at our project's OSF page: https://osf.io/6r4zy/. The SOPP collected records of tens of millions of traffic stops made by law enforcement agencies across the United States. Because data were sparse prior to 2011 (Pierson et al., 2020), we focus specifically on traffic stops by state patrol officers that occurred in the five-year period from 2011 through 2015.

The time period we examined includes 65,114,995 total traffic stops from 1,448 counties in 24 states. We analyzed the subset of those stops made in counties for which Project Implicit data were also available, limiting our analysis to 64,695,961 total traffic stops from 1,382 counties across 24 states. States included an average of 58 analyzed counties (ranging from 1 in Illinois to 254 in Texas). The specific information available for any given traffic stop (e.g., whether an arrest or search was made) varied considerably across states. See our extended supplement (https://osf.io/b7m84/) for lists and maps of which data were available in which regions.

Stop Rates. All 64,695,961 analyzed stops from all 1,382 analyzed counties included the driver's race and the county in which the driver was stopped, which allowed us to compute the proportion of Black and White driving-age residents who were stopped in each county by summing the number of stops within each racial category and dividing that aggregated total by the White or Black population share within the county. This procedure yielded race-specific stop rates (i.e., a White stop rate and a Black stop rate) that account for differences in population size across racial groups and across counties. We also aggregated separate county stop totals for Black and White men and women grouped into seven 10-year age brackets, but missing gender

and age data across states made these totals less useful for our present analyses. Still, these finer-grained stop rates are available in the dataset posted on OSF.

Arrests, searches, hit rates, and search thresholds. The extended supplement presents analyses of racial disparities in arrest rates, search rates, "hit" rates (i.e., the proportion of searches that uncover contraband), and search thresholds. Black (and White) arrest rates were computed as the proportion of Black (or White) stopped drivers who were subsequently arrested. Black (and White) search rates were computed as the proportion of Black (or White) stopped drivers who were subsequently searched. Black (and White) hit rates were computed as the proportion of Black (or White) searched drivers who were discovered to have contraband. All of these computations are possible using the SOPPD alone. Search thresholds were estimated by Pierson and colleagues, using a procedure they describe at length in their paper on the SOPPD (Pierson et al., 2020). They provided these estimates at our request, and we have appended them to the dataset on OSF.

Project Implicit

We estimated counties' implicit and explicit racial attitudes using data from Project Implicit (PI). PI has recorded responses from millions of volunteer visitors to their website, but to estimate county-level attitudes within the U.S. during the time period covered by the SOPPD, we focused specifically on those recorded from 2011 through 2015 from participants in the 50 U.S. States and Washington D.C. (N = 818,398; 39.8% men, 60.2% women; 70.4% White, 12.3% Black, 17.4% other races; Mean age = 26.9 years, SD = 11.7).

***Implicit Attitudes*.** PI measured participants' implicit attitudes using the Implicit Association Test (IAT; Greenwald et al., 1998). Positive IAT D scores indicate more anti-Black/pro-White

implicit attitudes (i.e., faster responses when categorizing Black (White) faces with negative (positive) stimuli than when the pairing is reversed).

We estimated county-level implicit attitudes using multi-level regression with post-stratification (MrP)—a technique often used to estimate public opinion at sub-national levels of analysis using broader national and state-level trends (see for example Hoover & Dehghani, 2019; Lax & Phillips, 2009; Riddle & Sinclair, 2019). This estimation procedure follows from our conceptual understanding of county-level racial attitudes. It first estimates the *direction and intensity* of residents' racial attitudes within demographic subgroups in each county. It then uses the *prevalence* of each subgroup within the county population to weight its contribution to a county-level attitude estimate.

MrP begins with a *multi-level regression* predicting the quantity of interest. In this case, we fit a model to predict implicit attitudes at the participant level from participant race (White, Black, or neither), gender (male or female), and age (using 8 categorical indicator variables: 15-17, 18-19, 25-29, 30-34, 35-44, 45-54, 55-64, 65 and older; 20-24 was the reference category). We estimated random intercepts at both the state and county level and allowed the effects of each of our predictors to vary randomly across states. If we instead attempt to estimate random effects at the county level, the model does not converge.

This first stage of MrP offers our best estimate for the implicit bias of each demographic subgroup (defined by race, gender and age) in each county. It captures *the direction and intensity* of county subgroups' racial attitudes. These estimates include three components: nationwide demographic effects (of age, gender, and race on individual-level bias), state-specific demographic effects (of age, gender, and race on individual-level bias), and any systematic, county-specific variance in implicit bias over and above these demographic effects. For example,

race had a strong nationwide effect. All else equal, Black respondents showed less anti-Black/pro-White implicit bias than White respondents did (b = -0.42; 95% CI: [-0.43, -0.42]). In Mississippi, the effect of race was slightly stronger than their nationwide counterpart. In Kansas, the effect of race was slightly weaker. At the same time, respondents from some counties in Mississippi and in Kansas were systematically more or less biased than respondents from other counties within the same state, independent of their race, gender, and age.

The second stage of MrP *post-stratifies* these scores using population prevalence estimates to generate a weighted average of the predicted scores from that county to capture the *prevalence of biased subgroups* in our county-level attitude estimate. For example, if Black men of age 25-34 comprise 5% of the population of a hypothetical county according to the 2014 5-year ACS estimates, the average IAT *D*-score for that group would comprise 5% of the county's overall *D*-score. We used the same race, gender, and age categories described above to post-stratify county-level implicit attitude estimates. In short, the first stage of MrP estimates the racial attitudes of demographic subgroups within each county, and the second uses the prevalence of these subgroups to estimate county-level attitudes.

The resulting estimate characterizes the racial attitudes of the county as a whole. To determine whether traffic stop disparities were specifically associated with variation in the racial attitudes of specific racial groups, we also estimated county-level attitudes among White residents, Black residents, and Non-Black Non-White residents separately. To do this, we changed only the second stage of the MrP estimation procedure. To estimate county-level attitudes among White residents, for example, we took the predicted attitude scores for White men and women of every age bracket and averaged them together, weighting the scores of each subgroup according to its share of the *White* population within that county.

Although MrP is technically capable of estimating implicit attitudes in counties with no PI participants using national and state-level trends, we only estimated attitude scores for counties for which at least one PI participant was available. In practice, county *n*s ranged from 1 to 16,295, with an average of 242 participants per county and 14,327 participants per state. With this procedure, we generated implicit attitude estimates for 3,018 counties in total, 1,383 of which could be matched to stop data from the SOPPD.

The illustrative models we use to present our key results yield the same substantive conclusions if we set a stricter minimum county sample size (n = 50).

***Explicit Attitudes*.** PI participants indicated "how warm or cold [they] feel" toward "African Americans/Black people" and "European Americans/White people" on separate 11-point scales. Item wording (e.g., "Black people" vs. "African Americans") varied randomly across respondents. We took the difference between these two feeling thermometer measures to create a measure of participants' explicit racial attitudes, with higher scores indicating more anti-Black/pro-White attitudes.

To estimate explicit attitudes at the county level, we used exactly the same MrP procedure described above, substituting our explicit attitude measure for IAT *D*-scores. These estimates are based on county-level *n*s ranging from 1 to 16,627, with an average of 245 participants per county and 14,516 participants per state. We generated explicit attitude estimates for 3,021 counties in total, 1,387 of which could be matched to stop data from the SOPPD.

In total, 1,343 counties had implicit attitude scores, explicit attitude scores, and traffic stop data.

Additional County-Level Variables

The population estimates we use for the MrP procedure (e.g., the number of Black driving-age residents within each county) are based on the 2010-2014 5-year estimates from the American Community Survey (ACS). We also used the ACS estimates to compute the proportion of driving-age county residents who are Black (or White), to compute Black and White stop rates within counties, and to post-stratify our county-level attitude estimates. We set 15 as the minimum driving age, given the availability of learners' permits and hardship licenses to 15-year-olds. Our estimate of Black-White racial segregation—a dissimilarity index (Duncan & Duncan, 1955)—is also based on the 2010-2014 ACS estimates (University of Wisconsin Population Health Institute, 2021). In some analyses, we controlled for additional county demographic characteristics included in the original Project Implicit dataset (originally derived from the 2009-2013 ACS). These included: county population, county population density (population per square mile as of 2010), median household income, and percent of residents age 25 or older with a bachelor's degree or higher.

Ethics

We affirm that the analysis of secondary data presented here conforms to ethical standards required by law in the United States.

Results

We pre-registered a set of analyses of these secondary data at https://osf.io/nz7sb. However, we deviated from our pre-registered analysis plan to account for unanticipated complexity in the dataset (e.g., extreme outlying observations, lack of racial diversity in many U.S. counties). To maximize transparency in light of this decision, we instead rely on a multiverse analysis (Steegen et al., 2016), which reports results from all of the data exclusion criteria and model specifications that we tested. For the sake of presentation, our results in the

main text focus on a single, illustrative set of models, along with a brief summary of the multiverse results that are described in full detail in both the abbreviated and extended supplements. See Figure 1 for an illustration of how available outcome data varied across model specifications.

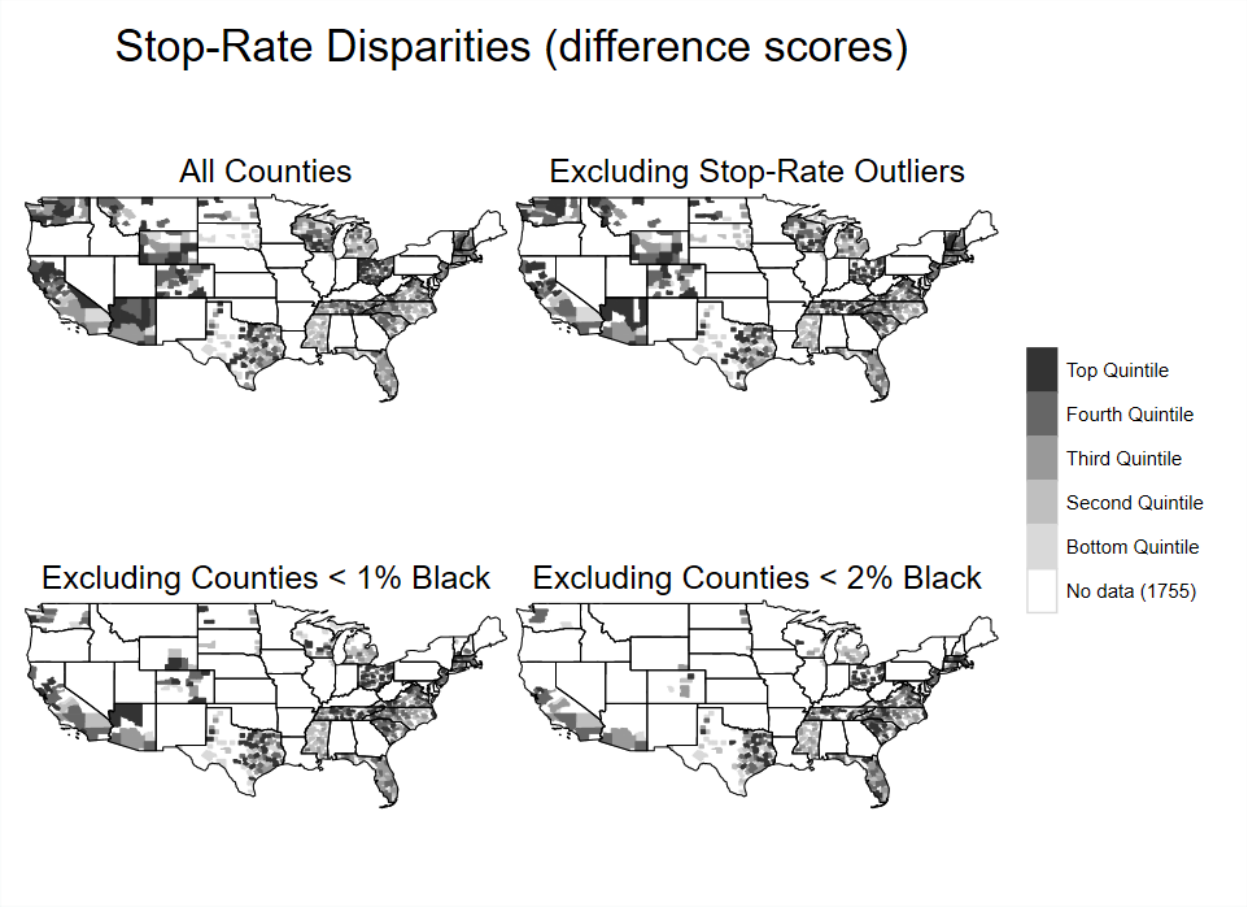

Figure 1. Figure depicts stop-rate disparities (Black stop rate – White stop rate) in the counties we used in our analyses. The top left panel shows stop-rate disparities for all counties in which these difference scores could be computed. Other panels exclude counties with outlying stop-rate difference scores (> 3 IQRs away from the median) or counties with populations < 1% or < 2% Black.

Racial Disparities in Traffic Stops

Consistent with prior findings using the SOPPD (Pierson et al., 2020), Black drivers were stopped at higher rates than White drivers from 2011 to 2015. Because we were concerned that

stop rates might be skewed by a small number of outlying scores, we compared the rates at which Black and White drivers were stopped in 912 counties from 24 states where (1) at least 1% of the driving-age residents were Black and (2) racial stop disparities fell within 3 interquartile ranges of the median score. We found sizeable racial disparities in stop rates ($t(911) = 13.05$, $p < 0.001$, Cohen's $d = 0.31$), with the average county stopping 101 Black drivers for every 100 driving-age Black residents (*SD* = 106) and 71 White drivers for every 100 driving-age White residents (*SD* = 84) during this 5-year period. See Figure 2.

When we instead include *all* counties with stop data (*N* = 1,347), the average rate at which Black drivers are stopped exceeds the average stop rate for White drivers by a staggering margin: on average, police stop 555 Black drivers for every 100 Black driving-age residents (SD = 5393) compared to 93 White drivers for every 100 White driving-age residents (SD = 181). However, this difference is distorted by counties with outlying scores. For example, two counties (where less than 0.5% driving-age residents were Black) stopped over 1,000 Black drivers for *each* driving-age Black resident.

To prevent these outlier counties from distorting our inferential tests of the predictors of stop disparities, we apply the exclusion criteria described above to all subsequent analyses described in our main text. Alternate analyses without these exclusion criteria are described in both the abbreviated and extended supplements.

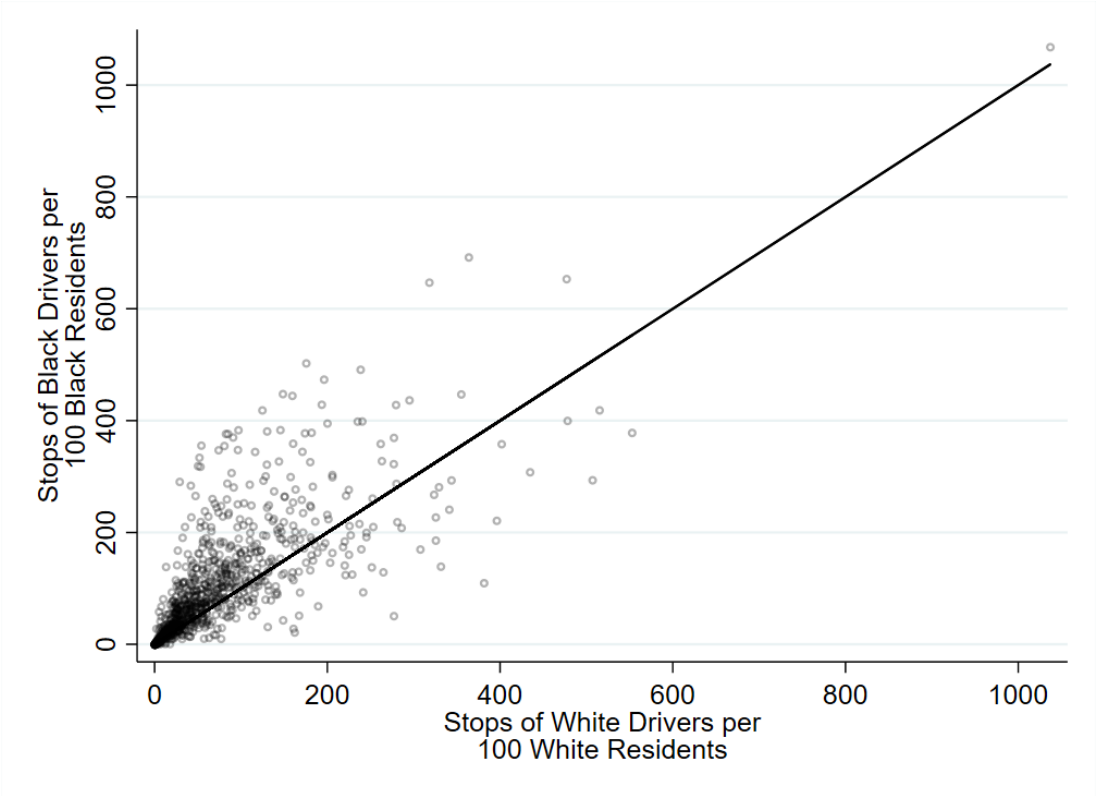

Figure 2. Racial Disparities in County-Level Traffic Stop Rates. Figure depicts only counties where the driving-age population is at least 1% Black and where the stop-rate disparity difference score (Black Stop Rate – White Stop Rate) is within 3 interquartile ranges of the median (N = 912). The line indicates stop-rate parity (Black Stop Rate = White Stop Rate).

County-Level Attitudes Predict Traffic Stop Disparities

We next examined whether county-level racial attitudes predicted disparities. In our illustrative model in which we quantified disparities with a difference score (i.e., Black Stop Rate – White Stop Rate), we found that stop disparities were correlated with our county-level estimates of implicit attitudes ($r(830) = 0.31$, $p < 0.001$) and explicit attitudes ($r(857) = 0.29$, $p < 0.001$), such that counties with relatively anti-Black/pro-White attitudes showed relatively anti-Black/pro-White stop disparities. These correlations exclude counties with outlying variable scores and counties with < 1% Black driving-age residents. However, the observations used to compute these correlations are not strictly independent. Counties are geographically, legally, and culturally nested within states. A more rigorous test of the relation between stop disparities and

county-level attitudes would account for this interdependence across observations. We therefore regressed traffic stop disparities on county-level attitude estimates in a series of multi-level linear models in which we allowed all coefficients (slope and intercept) to vary randomly across states.

For presentation purposes, figures below focus on illustrative models that exclude counties where less than 1% of the driving-age residents are Black and counties with outlying scores (i.e., scores more than 3 interquartile ranges above or below the median score) on the outcome or attitude variables. Because our estimates of county-level implicit and explicit racial attitudes were highly correlated ($r(3013) = 0.95$, $p < 0.001$), these models also only include one of these predictors at a time.[1] As many other defensible model specifications were possible, we report a larger "multiverse" (Steegen et al., 2016) of models in our abbreviated and extended supplements. These models are distinguished by whether they include or exclude counties with outlying scores, whether they include counties where at least 2%, 1%, or 0% of the driving-age population is Black, and whether they include implicit and explicit attitudes as simultaneous or separate predictors of stop disparities. We also estimated a version of all of these models using an alternative dependent variable: the *ratio* of the Black stop rate to the White stop rate (rather than the difference score). In total, this strategy yielded 36 models predicting traffic-stop disparities from racial attitudes, with 24 possible coefficient estimates for implicit attitudes and another 24 possible coefficient estimates for explicit attitudes. The "illustrative models" that we highlight below are generally representative of this larger multiverse.

Across most model specifications, counties with relatively anti-Black/pro-White attitudes had more pronounced racial disparities in traffic stops. In our illustrative models, both implicit attitudes ($b = 512$, $p < 0.001$, 95% CI: 401, 623; $N = 832$) and explicit attitudes ($b = 78$, $p <$

[1] This high collinearity is a consequence of aggregation and our MrP estimation procedure (Riddle & Sinclair, 2019). See the supplement for a detailed discussion of this issue.

0.001, 95% CI: 54, 103; N = 859) significantly predicted greater disparities. The coefficient for implicit attitudes was positive and significant in 18 of the 24 models in which it was estimated. Most of the non-significant coefficients (5/6) were observed in models that controlled for explicit attitudes. The coefficient for explicit attitudes was positive and significant in 11 of the 24 models in which it was estimated, with significance in 0 of the 12 models that controlled for implicit attitudes and significance in 11 of the 12 models that did not. The lack of significance for either implicit and explicit attitudes when both predictors were included in the model suggests that the strong relation between these variables (r = 0.95, $p < 0.001$ using all available counties) created multicollinearity in these models. Neither predictor was significant in models that included counties that were less than 1% Black and counties with outlying variable scores.

The size of these effects varied across model specifications, but Figure 3 depicts one example using our illustrative models (overlaid on a scatter plot of observed values). In counties with the most anti-Black/pro-White implicit attitudes (IAT D-score ≈ 0.42), the illustrative model predicts pronounced anti-Black stop disparity; if such a county had equal numbers of Black and White driving-age residents, 79% more Black drivers would be stopped (i.e., the predicted stop-rate difference score is 79). In contrast, the model predicts that counties with the least anti-Black/pro-White implicit attitudes (IAT D-score ≈ 0.21) would have stopped 29 *fewer* Black drivers per 100 Black driving-age residents than White drivers per 100 White driving-age residents. In short, moving across the full range of county-level implicit attitudes (from 0.21 to 0.42) produces a 108-point increase in the stop-rate difference score, from a 29-point anti-White bias to a 79-point anti-Black bias. In other models predicting stop-rate difference scores, the corresponding difference ranged from 52 to 971 (*Median Difference* = 100). See supplemental

Figure S1 for all unstandardized regression coefficients, including those predicting stop-rate ratios.

Explicit bias predicted stop disparities less robustly but with effect sizes of similar magnitude. In counties with the most anti-Black/pro-White explicit attitudes (White Warmth – Black Warmth ≈ 1.10), the illustrative model predicts a 90-point difference in stop rates, to Black drivers' disadvantage. In counties with the least anti-Black/pro-White explicit attitudes (White Warmth – Black Warmth ≈ -0.44), the model predicts a 31-point difference in stop rates to White drivers' disadvantage. This implies a 121-point increase in the stop-rate difference score across the full range of county-level explicit attitudes (from -0.44 to 1.10). In other models predicting stop-rate difference scores, the corresponding difference ranged from -247 to 1063 (*Median Difference* = 93). See supplemental Figure S1 for all unstandardized regression coefficients, including those predicting stop-rate ratios.

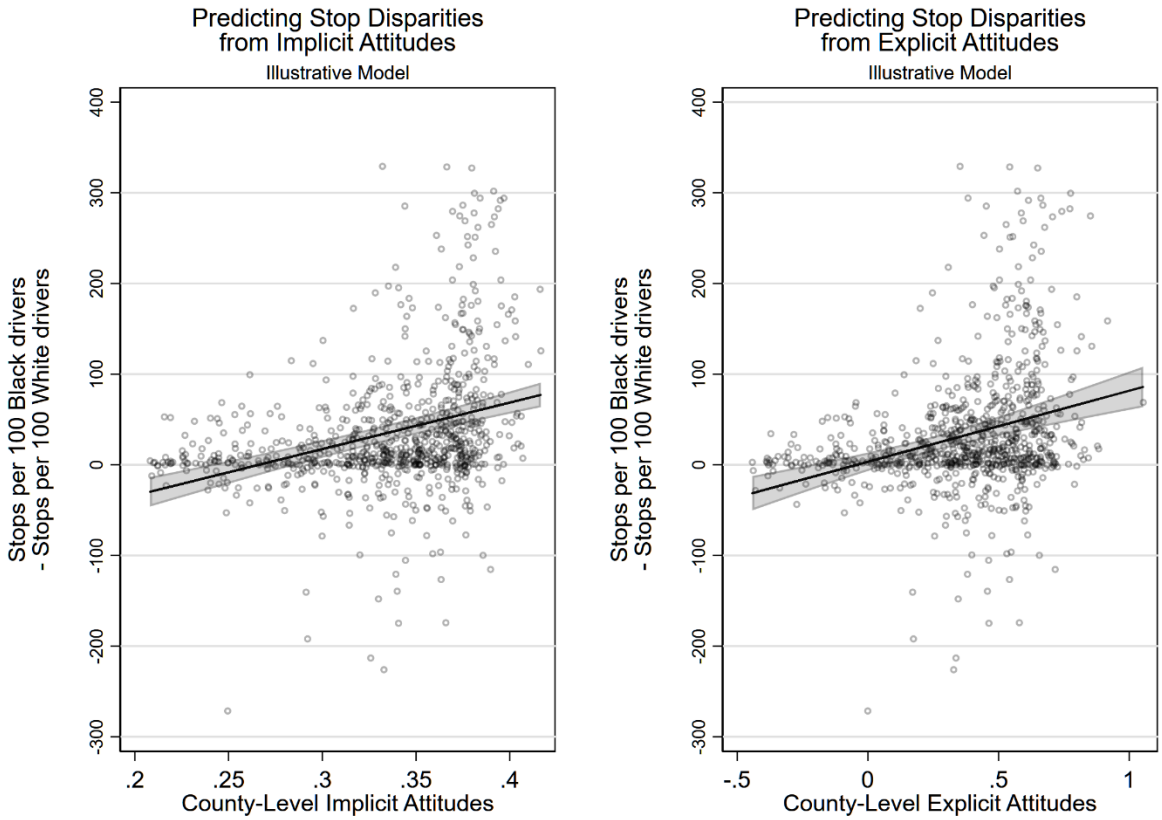

Figure 3. County-level attitudes predict stop-rate disparities. Figure depicts observed and predicted stop-rate disparities as a function of county-level implicit and explicit attitudes. More positive scores indicate greater anti-Black/pro-White attitudes or disparities. Circles represent observed values. Lines represent predicted values based on the fixed portion of a multi-level linear model predicting stop rate difference scores (Black stop rate – White stop rate) from either implicit or explicit attitudes; the slope and intercept varied randomly across states. Models exclude counties with outlying scores on the predictor or outcome and counties where < 1% driving-age residents were black. Error bands indicate 95% CIs.

Explaining the Link between County-Level Attitudes and Disparities

We next examine which components of county-level attitudes explain the relationship between county-level racial attitudes and racial disparities in police traffic stops. We have argued that counties' racial attitudes reflect a combination of *how biased* residents are and *how many* residents are biased, and our method of estimating county-level bias intentionally combines these into a single summary attitude score for each county. Our next analyses distinguish each of these from the other.

We first consider the direction and intensity of residents' anti-Black/pro-White attitudes. Perhaps stop disparities reflect the attitudes of all racial groups within a county. Or perhaps traffic stops are only related to certain residents' attitudes. For example, officers (or law enforcement agencies) may feel more accountable to White than to non-White civilians, lending White residents' attitudes more (direct or indirect) weight in whom officers stop. Or, if Black residents' racial attitudes predict their willingness to engage in political action, their attitudes may have a stronger impact on police behavior. We therefore computed separate county-level attitude estimates (both implicit and explicit) for three racial groups: White residents, Black residents, and Non-Black Non-White residents. If traffic stop disparities are more pronounced in counties where people are *more intensely biased*, then some or all of these race-specific attitude estimates should predict stop disparities. This analysis sets aside *how many* residents of each county were White, Black, or neither—an important component of county-level racial attitudes.

We found that *no* specific racial group's estimated racial attitudes was related traffic stop disparities. First, we estimated the bivariate correlations between our stop-disparity difference score and each attitude estimate (once again, excluding counties with outlying variable scores and counties with < 1% Black driving-age residents). These correlations were uniformly weak and non-significant (rs = -0.02 to 0.06; ps = 0.07 to 0.58). Second, we regressed traffic-stop disparities on the county-level racial attitudes of White residents (and then, in separate models, Black residents, then non-Black non-White residents), using the same multi-level models we used to analyze overall county-level attitudes.[2]. Of the 144 coefficients estimated across these analyses, only 1 was statistically significant. These analyses are presented in both the abbreviated and extended supplements.

[2] We also estimated county-level bias among the male and White male residents of counties and tested these subgroup biases as predictors of disparities. These analyses are summarized in the extended supplement.

These consistently null results indicate that traffic-stop disparities *are not related* to the variations in county-level racial attitudes within racial groups (i.e., White residents, Black residents, or non-White/non-Black residents). For example, counties where White residents harbored especially anti-Black/pro-White attitudes did not show more pronounced stop-rate disparities than counties where White residents harbored less anti-Black/pro-White attitudes. Why, then, would traffic-stop disparities be related to county-level racial attitudes after aggregating across these groups?

Perhaps traffic stop disparities are primarily related to *how many* White, Black, and other-race residents live in a county. Previous research (Nosek et al., 2007) and our own multi-level models indicate that White Americans tend to have the strongest anti-Black/pro-White racial attitudes and Black Americans tend to have the strongest pro-Black/anti-White racial attitudes of any racial group. Given that trend, counties' racial demographics are an important component of county-level racial attitudes, both as we conceptualize them and as we estimate them using MrP. In short, counties' racial demographics tell us a great deal about *how many* residents harbor (relatively) anti-Black/pro-White racial attitudes.

Consequently, we next examined the link between traffic-stop disparities and counties' racial demographics. Stop-disparity difference scores were significantly correlated with the percentage of driving-age county residents who were White ($r(910) = 0.36$, $p < 0.001$), who were Black ($r(910) = -0.33$), and who were neither Black nor White ($r(910) = -0.10$, $p = 0.003$; analyses exclude counties with outlying stop-disparity difference scores or with < 1% Black residents). Black drivers were stopped at higher rates than White drivers in counties where more of the population of White, where less of the population was Black, and (to a weaker degree), where less of the population was neither Black nor White.

Multi-level models permit a more rigorous test of the relation between racial attitudes, county demographics, and stop disparities. We re-estimated the multi-level linear models we used to predict stop disparities from county-level racial attitudes (weighted and aggregated across all racial subgroups, as in Figure 3), adding two predictors to each model: the proportion of driving-age residents who are White (*P*-White) and the proportion who are Black (*P*-Black). By including all of these predictors simultaneously, we can distinguish counties' racial composition from any systematic variation in county-level racial attitudes that is independent of race and compare the predictive utility of each to the other. To avoid multicollinearity, we did not include the percentage of non-Black, non-White drivers as a predictor in these models. We allowed all slopes and intercepts to vary randomly across states. In total, we estimated 36 models predicting stop disparities from county-level attitudes and demographics, which yielded 36 distinct coefficient estimates for each demographic variable (*P*-White and *P*-Black). The coefficient for *P*-White was positive and significant in 29 of the 36 models. The coefficient for *P*-Black was negative and significant in 15 of the 36 models. The coefficients for implicit and explicit attitudes were nearly uniformly non-significant across all 36 models (with the exception of one significant coefficient for implicit attitudes). See the abbreviated or extended supplement for details.

Figure 4 depicts predictions from our two illustrative models overlaid on scatter plots of the raw data. In these models, both county demographic variables significantly predicted racial disparities, but implicit and explicit attitudes did not. Stop disparities were more pronounced in counties where a higher proportion of driving-age residents were White ($b_{\text{P-White-implicit}} = 101$, $p < 0.001$, 95% CI: 64, 137; $N = 832$; $b_{\text{P-White-explicit}} = 104$, $p < 0.001$, 95% CI: 68, 141; $N = 859$). Using the illustrative model controlling for implicit attitudes, counties with the highest

proportion of White residents (0.98) stopped 56 more Black drivers per 100 Black driving-age residents than White drivers per 100 White driving-age residents. These stop rates only reach predicted parity when the White population share drops to about 0.50. Stop disparities were also more pronounced in counties where a lower proportion of driving-age residents were Black ($b_{\text{P-Black-implicit}}$ = -258, p = 0.003, 95% CI: -430, -86; N = 832; $b_{\text{P-Black-explicit}}$ = -257, p < 0.001, 95% CI: -397, -117; N = 859). Using the illustrative model controlling for implicit attitudes, we estimate that counties with the lowest proportion of Black residents (0.01) stopped 50 more Black drivers per 100 Black driving-age residents than White drivers per 100 White driving-age residents, and that these stop rates only reach predicted parity where the Black population share is at least 0.20. Because most counties in the United States (and in our sample) are more than 75% White and less than 10% Black, our model would predict anti-Black/pro-White stop disparities in most U.S. counties.

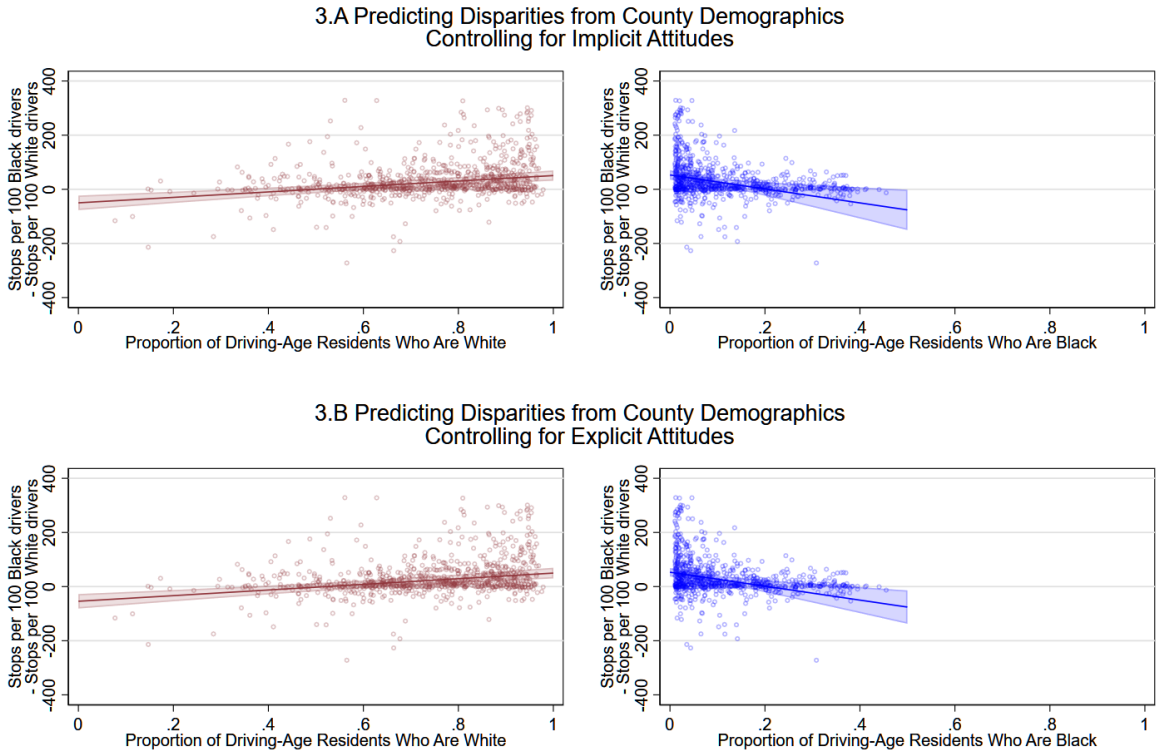

Figure 4. County demographics predict stop-rate disparities. Figure depicts observed and predicted stop-rate disparities as a function of proportions of White and Black driving-age residents. More positive scores indicate greater anti-Black/pro-White stop disparities. Circles represent observed values. Lines represent predicted values based on the fixed portion of multi-level linear models predicting stop rate difference scores (Black stop rate – White stop rate) from either implicit or explicit attitudes and the two demographic variables; the intercept varied randomly across states. Models exclude counties with outlying scores on the predictor or outcome and counties where < 1% driving-age residents were black. Error bands indicate 95% CIs.

In sum, county racial demographics were much more robust predictors of racial disparities in traffic stops than other sources of systematic variation in county racial attitudes were. Racial disparities were robustly more pronounced in counties that were relatively uniformly White and (less robustly) in counties that had relatively few Black residents.

To ensure that our *P-White* and *P-Black* variables were not merely proxies for some other demographic confound in our analysis, we re-estimated these 36 models with the following additional covariates: county population, population density, median household income, Black-White racial segregation, and percent of residents age 25 or older with a college education. These

additional models do not affect our conclusions. The coefficient for *P-White* is, if anything, more robust in models that include additional controls (positive and significant in all 36 models), whereas the coefficient for *P-Black* became slightly less robust (negative and significant in 13 of 36 rather than 15 of 36 models). The effects of county-level implicit and explicit attitudes, meanwhile, remained relatively weak and infrequently significant after taking demographics into account (6/24 coefficients for implicit attitudes were significant; 0/24 for explicit attitudes). Some of the covariates we included – especially median income – may also be related to racial disparities in traffic stops. These are discussed in the extended supplement.

Post-Stop Racial Disparities

Our analysis focuses on racial disparities in stop rates. The SOPPD also includes information about arrest rates, search rates, and the proportion of searches that yield contraband. Once stopped, Black drivers were more often arrested or searched than White drivers, and searched Black drivers were less likely than searched White drivers to be found with contraband. However, county-level racial attitudes did not consistently predict any of these disparities. Analyses of arrest, search, and hit-rate disparities are described at length in the extended supplement.

Discussion

The present study yields three key findings that are robust across a multiverse of model specifications. First, consistent with other work in this domain, we find that Black drivers (compared to White drivers) are stopped by police at a rate disproportionate to their share of the driving-age population in counties across the United States. Second, we find that these disparities in traffic stops are related to county-level racial attitudes. Finally, we find evidence that the relation between stop disparities and racial attitudes is more closely related to the *prevalence* of

anti-Black/pro-White racial attitudes than to the *direction and intensity* of racial attitudes. Whether one county's White residents (or Black residents, or Non-Black non-White residents) were more or less racially biased than another's did not uniquely predict its traffic stop disparities. What mattered was simply the proportion of a county's population that was White (and less robustly, the proportion that was Black). As White Americans show the strongest anti-Black/pro-White attitudes on average, this proportion is closely tied to the prevalence of racial bias in a county. Where White Americans were more prevalent, anti-Black/pro-White attitudes were more prevalent, and Black stop rates tended to exceed White stop rates.

We did not find any relation between county-level attitudes and racial disparities in searches or arrests. Traffic stops differ from searches and arrests in that stop decisions are often made prior to any direct interaction and may be based on more ambiguous information. Decisions that involve more ambiguous information tend to be more strongly guided by racial attitudes (Dovidio & Gaertner, 2000; Swencionis & Goff, 2017), so decisions to stop may be more consistently tied to racial attitudes than decisions to search or arrest. Stops may also be more directly subject to regional biases in how police resources are deployed than the arrests or searches that officers make after stops. Finally, because arrests and searches are *contingent* on stops, our analyses may underestimate the total effect of attitudes on these post-stop outcomes (Knox et al., 2020).

Implications

The present study combines the largest existing datasets on racial attitudes and policing in the U.S. to highlight a previously undocumented relation between county-level attitudes and racial disparities in traffic stops—a relation that is attributable to county racial demographics in particular. Although we cannot draw strong causal inferences, one possibility consistent with our

results is that between-groups differences in Black and White racial attitudes that prevail throughout the U.S. have a more racializing influence on police practices than do within-group differences in attitudes or county-to-county variation in the intensity of those attitudes. In fact, net of racial demographics, we observed very little systematic county-level variation in racial attitudes. After taking into account Project Implicit respondents' race, gender, and age, the county in which participants resided accounted for little variance in their racial attitudes (for implicit bias, ICC_{county} = 0.004, 95% CI: 0.003, 0.006; for explicit bias, ICC_{county} = 0.012, 95% CI: 0.010, 0.017). This means that, empirically, there was not a great deal of systematic, county-to-county variance in the direction and intensity of racial attitudes independent of race, gender, age, or state-level trends. The difference between Black and White Americans' racial attitudes seems to dwarf geographic variance in racial attitudes within these groups.

That said, we are not saying that "race matters and racial attitudes don't." Race may be so important in our analyses *because* Black and White Americans differ so consistently in their racial attitudes. By our estimates, the "average" White resident of every single county within our dataset harbored some degree of anti-Black/pro-White attitudes. This consistent anti-Black bias among White Americans may provide the political will to enact laws and policies that advantage White Americans and disadvantage Black Americans in counties with large proportions of White residents, regardless of the *particular* extent of Whites' anti-Black/pro-White bias in that county. Legislators may pass laws with racially disparate impacts (e.g., Henricks & Harvey, 2017), law enforcement agencies may deploy officers to over-police predominantly Black areas (Chohlas-Wood et al., 2018; Mummolo, 2018), and police officers who racially discriminate may avoid accountability (Luh, 2020). These sources of racial disparities may loom larger in counties with relatively large proportions of White residents, if, on average, White residents with even a

modest preference for their own race are willing to tolerate them. Our analysis cannot rule out the possibility that large-scale changes in Whites' racial attitudes would reduce stop disparities. But by our estimation, the naturally occurring county-to-county variance in Whites' attitudes that existed from 2010-2015 was not sufficient.

It is also possible that other phenomena related to counties' racial demographics matter more than between-race differences in attitudes. First, racial demographics may be linked to other region-level psychological characteristics that could contribute to stop disparities, such as race-danger stereotypes or authoritarianism (Altemeyer, 1988; Hehman et al., 2018). Second, White residents might "use" the police (intentionally or not) to preserve their advantaged position in society (e.g., by asserting their control over spaces that they perceive to be theirs; Laniyonu, 2017). As dominant group members sometimes defend their power and status even when their attitudes toward other groups are ostensibly positive, White residents in predominantly White counties might be empowered to implement anti-Black/pro-White policing laws and practices even if their attitudes *were* unbiased (Bobo, 1999; Dixon et al., 2012; Jackman, 1994; Sidanius & Pratto, 1994). Third, reduced racial diversity in a county may cause officers to see Black drivers as "out of place" and police them closely (Gelman et al., 2007; Meehan & Ponder, 2002). Finally, predominantly White counties may have a history of racially discriminatory laws, policies, and institutions in decades past that have enduring impacts on racial stop disparities. These impacts may persist independent of current residents' attitudes.

Limitations

Our study is limited in at least three important ways. First, our data are correlational. Racial disparities in traffic-stops could be a cause rather than a consequence of county-level attitudes or counties' racial demographics. For example, our results are consistent with the

possibility that Black Americans move away from counties where police treat them unfairly. Moreover, unmeasured correlates of racial attitudes, county racial demographics, and stop disparities could bias our estimates. Controlling for population, population density, median income, segregation, and college education did not change our results, but other confounds may matter.

Second, our inferences may not generalize to counties or individuals outside our sample. Because the PI data come from an online convenience sample, our estimates of racial attitudes are imperfect, even when weighted using MrP. Our stop-disparity data come from 24 states, which may systematically differ from states in which data were unavailable. And although racial disparities in policing are not uniquely American, the trends we observe here may differ in other countries with different racial politics.

Finally, our analysis focuses on a particular outcome—traffic stop rates per 100 driving-age residents of a given race. This measure captures racial disparities in stops imperfectly because the racial composition of a county's residents may not perfectly reflect that of its drivers. Some counties with few Black residents, for example, may include a large number of Black drivers who regularly pass through them. Another issue with this measure is that we only know which people were stopped by police. We do not know who actually violated traffic laws and whom police opted not to stop. If there are real group differences in traffic violations, our measure of racial disparities in stop rates could underestimate the racial inequity of traffic stops in an area (if White drivers offend more often than Black drivers) or overestimate this inequity (if Black drivers offend more often than White drivers). The ideal denominator would be the number of drivers of a given race in a county who actually violate traffic laws, but records of traffic law violations are subject to the same potential biases as stops themselves (Hehman et al.,

2018). A related problem may influence our inferences about the role of attitudes in post-stop outcomes. Because the post-stop outcomes were only observed for drivers who were pulled over, the effects of region-level attitudes on arrests and searches may only *appear* to be null if initial stops are strongly related to regional attitudes. If officers stopped drivers at perfectly equal or equitable rates, attitudes might be more closely related to post-stop outcomes (Knox et al., 2020).

Conclusion

We find that more homogeneously White counties show more pronounced racial disparities in traffic stops, suggesting that Black Americans in these counties are systematically subjected to a greater degree of scrutiny and suspicion than White Americans. Multiple processes could account for this pattern. The relation we observe may arise from differences across counties in how police are trained, where police are deployed, which neighborhoods are rich, poor, or dangerous, which crimes people commit, or which crimes people report. But regardless of the source of these disparities, they have a meaningful geographical component. American drivers' experiences with the police depend both on who they are and where they travel.

References

Alpert, G. P., MacDonald, J. M., & Dunham, R. G. (2005). Police suspicion and discretionary decision making during citizen stops. *Criminology, 43,* 407-434. https://doi.org/10.1111/j.0011-1348.2005.00012.x

Altemeyer, B. (1988). *Enemies of freedom: Understanding right-wing authoritarianism*. Jossey-Bass.

Beckett, K., Nyrop, K., & Pfingst, L. (2006). Race, drugs, and policing: Understanding disparities in drug delivery arrests. *Criminology, 44,* 105-137. https://doi.org/10.1111/j.1745-9125.2006.00044.x

Bobo, L. D. (1999). Prejudice as group position: Microfoundations of a sociological approach to racism and race relations. *Journal of Social Issues, 55,* 445-472. https://doi.org/10.1111/0022-4537.00127

Chohlas-Wood, A., Goel, S., Shoemaker, A. & Shrof, R. *An Analysis of the Metropolitan Nashville Police Department's Traffic Stop Practices.* Retrieved from https://policylab.stanford.edu/media/nashville-trafc-stops.pdf (Stanford Computational Policy Lab, 2018).

Davis, E., Whyde, A., & Langton, L. (2018). Contacts between police and the public, 2015. *US Department of Justice Office of Justice Programs Bureau of Justice Statistics Special Report*, 1-33. http://www.bjs.gov/index.cfm?ty=pbdetail&iid=6406

Dixon, J., Levine, M., Reicher, S., & Durrheim, K. (2012). Beyond prejudice: Are negative evaluations the problem and is getting us to like one another more the solution?. *Behavioral and Brain Sciences, 35,* 411-425. https://doi.org/10.1017/S0140525X11002214

Dovidio, J. F., & Gaertner, S. L. (2000). Aversive racism and selection decisions: 1989 and 1999. *Psychological Science*, *11*, 315–319. https://doi.org/10.1111/1467-9280.00262.

Duncan, O. D., & Duncan, B. (1955). A methodological analysis of segregation indexes. *American Sociological Review*, *20*, 210-217. https://doi.org/10.2307/2088328

Epp, C. R., Maynard-Moody, S., & Haider-Markel, D. P. (2014). *Pulled over: How police stops define race and citizenship*. University of Chicago Press.

Gelman, A., Fagan, J., & Kiss, A. (2007). An analysis of the New York City Police Department's "stop-and-frisk" policy in the context of claims of racial bias. *Journal of the American Statistical Association*, *102*, 813-823. https://doi.org/10.1198/016214506000001040

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464-1480. https://doi.org/10.1037//0022-3514.74.6.1464

Hehman, E., Flake, J. K., & Calanchini, J. (2018). Disproportionate use of lethal force in policing is associated with regional racial biases of residents. *Social Psychological and Personality Science*, *9*, 393-401. https://doi.org/10.1177/1948550617711229

Henricks, K., & Harvey, D. C. (2017). Not one but many: Monetary punishment and the Fergusons of America. *Sociological Forum*, *32*(S1), 930–951. https://doi.org/10.1111/socf.12360

Hoover, J. & Dehghani, M. (2019). The big, the bad, and the ugly: Geographic estimation with flawed psychological data. *Psychological Methods*, *25*, 412-429. https://doi.org/10.1037/met0000240

Jackman, M. R. (1994). *The velvet glove: Paternalism and conflict in gender, class, and race relations.* University of California Press.

Knox, D., Lowe, W., & Mummolo, J. (2020). Administrative records mask racially biased policing. *American Political Science Review*, *114*, 619-637. https://doi.org/10.1017/S0003055420000039

Laniyonu, A. (2018). Coffee shops and street stops: Policing practices in gentrifying neighborhoods. *Urban Affairs Review*, *54*(5), 898–930. https://doi.org/10.1177/1078087416689728

Lax, J. R. & Phillips, J. H. (2009). How should we estimate public opinion in the states? *American Journal of Political Science, 53,* 107-121. https://doi.org/10.1111/j.1540-5907.2008.00360.x

Leitner, J. B., Hehman, E., Ayduk, O., & Mendoza-Denton, R. (2016a). 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

Leitner, J. B., Hehman, E., Ayduk, O., & Mendoza-Denton, R. (2016b). Racial bias is associated with ingroup death rate for Blacks and Whites: Insights from Project Implicit. *Social Science & Medicine, 170,* 220-227. https://doi.org/10.1016/j.socscimed.2016.10.007

Luh, E. (2020). *Not so black and white: Uncovering racial bias from systematically misreported trooper reports*. Unpublished manuscript. Retrieved from https://elizluh.github.io/files/highwaybias.pdf.

Meehan, A. J. & Ponder, M. C. (2002). Race and place: The ecology of racial profiling African American motorists. *Justice Quarterly, 19,* 399-430. https://doi.org/10.1080/07418820200095291

Mummolo, J. (2018). Militarization fails to enhance police safety or reduce crime but may harm police reputation. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(37), 9181–9186. https://doi.org/10.1073/pnas.1805161115

Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., Smith, C. T., Olson, K. R., Chugh, D., Greenwald, A. G., & Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology, 18,* 36-88. https://doi.org/10.1080/10463280701489053

Orchard, J. & Price, J. (2017). County-level racial prejudice and the black-white gap in infant health outcomes. *Social Science & Medicine, 181,* 191-198. https://doi.org/10.1016/j.socscimed.2017.03.036

Pierson, E., Simoiu, C., Overgoor, J., Corbett-Davies, S., Jenson, D. Shoemaker, A., Ramachandran, V., Barghouty, P., Phillips, C., Shroff, R., & Goel, S. (2020). A large-scale analysis of racial disparities in police stops across the United States. *Nature Human Behaviour, 4,* 736-745. https://doi.org/10.1038/s41562-020-0858-1

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, 116,* 8255-8260. https://doi.org/10.1073/pnas.1808307116

Roh, S. & Robinson, M. (2009). A geographic approach to racial profiling: The microanalysis and macroanalysis of racial disparity in traffic stops. *Police Quarterly, 12,* 137-169. https://doi.org/10.1177/1098611109332422

Sidanius, J., & Pratto, F. (2001). *Social dominance: An intergroup theory of social hierarchy and oppression*. Cambridge University Press.

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science, 11*, 702-712. https://doi.org/10.1177/1745691616658637

Swencionis, J. K., & Goff, P. A. (2017). The psychological science of racial bias and policing. *Psychology, Public Policy, and Law*, *23*, 398–409. https://doi.org/10.1037/law0000130.

University of Wisconsin Population Health Institute. (2021). *Residential segregation - black/white.* https://www.countyhealthrankings.org/explore-health-rankings/measures-data-sources/county-health-rankings-model/health-factors/social-and-economic-factors/family-social-support/residential-segregation-blackwhite

Warren, P., Tomaskovic-Devey, D., Smith, W., Zingraff, M., & Mason, M. (2006). Driving while Black: Bias processes and racial disparity in police stops. *Criminology*, *44*, 709-738. https://doi.org/10.1111/j.1745-9125.2006.00061.x

Xu, K., Nosek, B., & Greenwald, A. G. (2014). Data from the race implicit association test on the Project Implicit demo website. Journal of Open Psychology Data, 2, 1-3. https://doi.org/10.5334/jopd.ac