

CrimRxiv

Are Young Black Men Traveling Together Targeted for Traffic Stops?

Lance Hannon¹ Molly Biddle

¹Villanova University

Published on: Mar 05, 2026

DOI: <https://doi.org/10.21428/cb6ab371.f1d81a4b>

License: [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License \(CC-BY-NC-ND 4.0\)](#)

ABSTRACT

This study explores whether police officers are especially likely to stop young Black males traveling in parties of two or more. Our analysis is guided by social psychological theory and research suggesting that racial stereotypes about criminality are applied with greater intensity to Black people encountered in groups. Investigatory stop data for Philadelphia indicate that relative to young White male drivers, young Black male drivers are about five times more likely to be pulled over in the company of other young men of the same race. Using the veil-of-darkness test, we show that this disparity cannot be easily explained by racial differences in motorist behavior (such as carpooling). Our analyses reveal that young Black men in groups experience a significantly elevated probability of being stopped for traffic violations when daylight visibility facilitates officer discernment of vehicle passengers. Previous studies have overlooked this dynamic, implicitly assuming that multi-occupant vehicle stops are equally distributed across racial groups and/or that having a lone young Black male driver would produce a ceiling effect in the level of disparate scrutiny and selective traffic enforcement. We find that while young Black males traveling alone are disadvantaged in inter-racial comparisons, intra-racially they are less likely to be targeted than young Black males in groups. Overall, our results suggest that beyond a present-versus-absent dichotomy, the greater the presence of young Black men in a vehicle, the greater the likelihood of being selected for a pretextual traffic stop.

Introduction and Legal Background

Numerous legal scholars have argued that *Whren v. United States* (1996) greatly enabled police misconduct (Carbado and Feingold 2022; Epp et al. 2014; Stuart VI 2022). In this momentous decision, the U.S. Supreme Court ruled that pretextual traffic stops were not in violation of 4th Amendment protections against unreasonable search and seizure. The case essentially constitutionalized a police officer's ability to exploit trivial traffic infractions as a way to circumvent required evidentiary thresholds for criminal investigations. Still, the Court left open the possibility that pretextual vehicle stops could be challenged on 14th Amendment grounds if it could be shown that the tactic was being unequally applied based on protected characteristics like race, age, and sex.

The facts in the (in)famous *Whren* case help inform our research on demographic disparities in urban traffic stops, specifically racial disparities linked to vehicle occupancy rates. On June 10th, 1993, Michael Whren was a passenger in a SUV that paused at a Washington, DC intersection. This pause drew the attention of plainclothes vice officers in a nearby unmarked car who testified that they could see Whren looking down at his lap. It was 8:25 pm, but the sun had not yet set in DC and so daylight visibility was reasonably good. Of course, if the officers had a line of sight where it was possible to perceive Whren's gaze, they could also almost certainly discern that Whren was a young Black man, as was the vehicle's driver, James Brown, who was also said to be looking at Whren's lap. While pausing to look down is not a stop-worthy offense, the vehicle was

eventually pulled over on the basis of Brown's failure to signal before turning right as well as unreasonable speed in making that right turn. One of the lead officers admitted that despite detaining Brown and Whren for traffic violations, he had no intention of issuing citations: "I'm a vice investigator; I'm not out there to give tickets." A close examination of the vehicle revealed drug contraband in Whren's hands, leading to the arrest of both vehicle occupants. Whren's lawyers were unsuccessful in their attempts to suppress the seized contraband on the grounds that the vice officers lacked reasonable suspicion or probable cause to initiate an investigatory stop (see Whren's [appellate brief](#)).

Since the *Whren* decision (1996), pedestrian stop-and-frisk practices have been successfully challenged with 4th Amendment unreasonable search and seizure claims, but challenges related to vehicle stops have focused on 14th Amendment concerns—especially racial bias in the enforcement of motor vehicle code violations (Whitney 2008). Demonstrating in court proceedings that racial disparities in traffic stops are a product of racial bias is a notoriously difficult endeavor (Butler 2017). A key issue is the absence of a reliable statistical benchmark. For example, if police officers were to act without any unlawful bias, what is the expected percentage of stops involving Black motorists? Courts have often rejected allegations of bias based on questionable benchmarks that assume that driver demographics must match those of neighborhood residents or that all demographic groups violate traffic laws to the same extent (Hannon 2019).

One of the most respected methods for demonstrating bias in traffic enforcement, the veil-of-darkness test, was explicitly engineered to address this challenging benchmarking issue (Knode et al. 2024, Smith et al. 2021). Indeed, the first words in Grogger and Ridgeway's (2006: 878) abstract introducing the technique are "the key problem in testing racial profiling in traffic stops is estimating the risk set, or 'benchmark,'" while the final words are "the data yield little evidence of racial profiling in traffic stops" [in their empirical example of Oakland, California]. Grogger and Ridgeway (2006) won the 2007 Outstanding Statistical Application Award from the American Statistical Association for the development of this methodology. Their influential innovation exploits variation in daylight throughout the year and is based on the notion that officers who are engaged in bias-based policing will be somewhat less capable of discerning a driver's race when it is dark outside. They argue that if, when holding constant the time of day and location, stops made in darkness have a significantly smaller proportion of Black drivers than stops made in direct sunlight, then this suggests racial bias associated with varying awareness of who is in the car (Grogger and Ridgeway 2006).

The veil-of-darkness test produces a notoriously conservative estimate of bias-based policing; artificial lighting, car profiling, and urban segregation can all reduce the extent of practical difference between daylight and nighttime enforcement. A defining feature of the veil-of-darkness test is that it focuses on likely variation, however slight, in the amount of information officers have regarding a vehicle's occupants and whether this is at all correlated with discretionary actions to pull the car over. Relative to previously employed statistics, it makes a much more cautious assumption that any tendencies for certain populations to disproportionately drive at specific hours or commit traffic infractions will be equally present in daylight and darkness. Adding to the

measure's conservativeness, existing research suggests that, if anything, Black drivers appear to drive more carefully in high visibility conditions relative to other racial groups (Kalinowski et al. 2017).

In relation to the specifics of the *Whren* case, the veil-of-darkness methodology implies that Whren and Brown would have stood a better chance of avoiding being profiled had they ventured out just one hour later under the cover of darkness. Along with numerous other academic and policy studies in a variety of jurisdictions (e.g., Knode et al. 2024, Pierson et al. 2020, Ritter and Bael 2009, Taniguchi et al. 2017, and Vito et al. 2026), recent research utilizing the veil-of-darkness test in Philadelphia has shown that the proportion of single-occupant vehicle stops involving young Black men significantly declines with nightfall (Hannon et al. 2021, Hannon and Biddle 2025). However, to our knowledge, there is no existing research on whether the number of people in the vehicle makes a difference for the likelihood of visibility-based racial profiling.

The present study asks whether young Black male motorists in groups are especially likely to be stopped for traffic infractions when daylight enhances the visibility of vehicle passengers. Framed in the context of the *Whren* case, the key issue is whether Whren might have had a lower chance of being judged as criminally suspicious had he been alone rather than in the company of another young Black man.

Related Research in Social Psychology

While we are unaware of any previous research on whether the number of vehicle occupants affects the race-specific probability of being pulled over, an insightful study in social psychology examines how pedestrian stop-and-frisks in New York City were more intense for Black people encountered in groups of 2 or more when compared to those Black people stopped alone (Cooley et al. 2019). Cooley et al.'s focus on pedestrian stops meant that the veil-of-darkness method could not be applied to their particular data (since vehicle glass is a necessary part of the 'veil'). While Cooley et al. were unable to directly ascertain whether Black people were more likely to be stopped when walking in groups, the researchers could compare the extent of intrusiveness in solo versus group stops after holding constant potential confounders. Cooley et al. reported that frisks, extensive searches, and uses of force were all significantly more common in stops of Black people among groups (consistent with data from other locations, only 7% of the suspects in their stop-question-and-frisk sample were female). The relative intensity of these interrogations suggests that groups of Black (male) pedestrians may have been particularly targeted for investigation.

Cooley et al. (2019) noted that their stop-and-frisk results were consistent with a large body of experiment-based studies demonstrating that misperceptions about dangerousness are heightened in group contexts. For example, Insko and Schopler's (1998) multifaceted review of the 'discontinuity effect' evidence concluded that distrust of groups is significantly greater than that of individuals in otherwise identical transactions. This psychological phenomenon has been shown for groups as small as two (McGlynn et al. 2009). Other research in social psychology has pointed to an 'anger bias' in the evaluation of groups where observers were more likely to characterize subtle expressions as angry in the context of a crowd (Mihalache et al. 2021). This bias is

thought to reflect an intrinsic human tendency for people to lean toward negative evaluations in situations of potential threat, possibly serving an evolutionary protective function.

While there is considerable research suggesting that observers are universally more inclined to perceive people in groups with greater suspicion than people met individually, there is also evidence indicating that cultural stereotypes can amplify broadly human defensive tendencies (Cooley and Payne 2017, 2018). Importantly, observers appear more prone to assign cultural labels to a group of representatives than a single representative of that group. Cooley et al. (2019: 762) argue that theoretically, “Such findings suggest that stereotypes about entire racial categories are applied more strongly to homogenous racial groups than individuals from those groups—perhaps because racial categories are themselves large groups.” Ultimately, people might individuate the solitary strangers they encounter and process the situation based on its specific merits. However, for strangers encountered in groups, the same people might respond to cognitive bandwidth strain by ignoring the multitude of specifics and instead relying on learned generalizations (in this case, cultural schemas concerning the racial typification of crime).

The Racial Typification of Crime in the U.S.

Criminological research in the United States has long been interested in investigating tendencies to associate blackness with criminal offending, violence, and danger (Welch 2007). Russell-Brown (1998) famously coined the term “criminalblackman” to encapsulate how young Black men have become practically synonymous with criminality in the American psyche. Early empirical findings in this literature chronicle the overestimation of the amount of crime committed by Black youth as well as the overestimation of the dangerousness of majority Black neighborhoods (Quillian and Pager 2001). Scholarship in the conflict tradition argues that this racial typification of crime has been instrumental in various subordination and exploitation regimes throughout the history of the United States (Alexander 2010, Cobb 2025, Wacquant 2001). At a less macro level, a central result in survey work in this area is that the racial typification of crime predicts support for punitive criminal justice policies even after adjusting for ideology, crime salience, and other covariates. For example, Chiricos et al. (2004) show that respondents who more strongly characterize crime as a Black phenomenon are more supportive of harsh punishment measures, and that this relationship is especially evident among Whites—including those not expressing any overt prejudice.

Contemporary research on Black criminal stereotypes emphasizes that racial typification is not merely an aberrant psychological bias, but a socially produced schema reinforced through longstanding structural imbalances and slanted media narratives. In tandem with the continuation of residential segregation and institutionalized inequity, a central explanation for the persistence of racial typification is that crime information is widely obtained through distorted media depictions rather than direct experience, and this is particularly true for White people’s understanding of crime. Using a rich dataset with respondent geographic identifiers, Baranauskas and Drakulich (2018) show that local television news and crime dramas are linked to perceptions of crime and support for tough policy frames (as others have documented), but these relationships

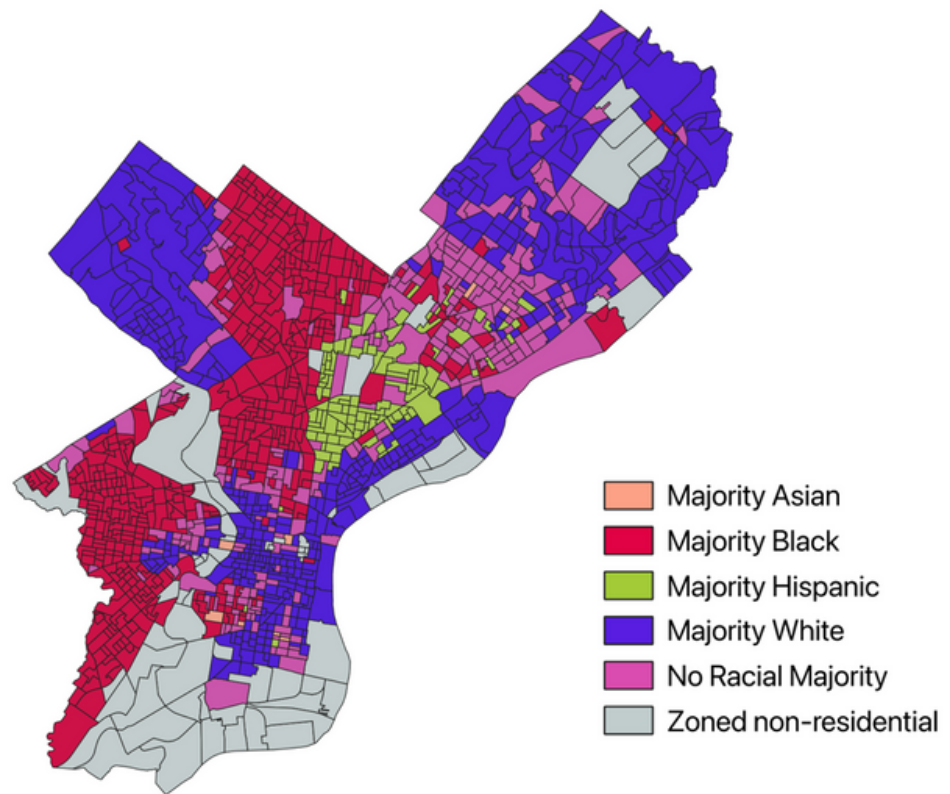
are significantly conditioned by the racial context in which media are consumed. In particular, the impact of local television news and crime dramas was stronger for Whites living in proximity to large numbers of Black people.

Extending beyond television, Intravia and Pickett (2019) show that social media news consumption (but not social media posting, commenting, or sharing) is positively associated with the racial typification of Black people as offenders. Focusing on the social media presence of local law enforcement agencies, Grunwald et al. (2022) analyze nearly 100,000 posts from almost 14,000 Facebook pages and find that, on average, viewers are exposed to posts that overrepresent Black suspects relative to local arrest rates by a substantial margin—a pattern that Grunwald et al. conclude could reinforce stereotypes and intensify punitive preferences. Grunwald et al. also reported that the degree of overrepresentation of Black suspects in local police department Facebook posts increases with the percent of non-Black residents in an area.

Data and Methods

The 2021-2024 vehicular investigatory stop data for the present analyses were originally collected by the Philadelphia Police Department (PPD) and made publicly available on [OpenDataPhilly.org](https://opendata.philly.org/). In Philadelphia, highway stops (e.g., Interstate 95) are mostly handled by the Pennsylvania State Police who do not participate in this data publishing practice. Almost all of the PPD stops utilized in the current study took place in Philadelphia's residential neighborhoods. While Philadelphia is a diverse city where no single ethnoracial category constitutes the overall majority, it is also highly residentially segregated (Anderson 2014, Massey and Denton 1998). For example, a single racial group predominates in 81% of all residential census block groups in Philadelphia, and these small areas tend to be clustered together by racial predominance (see Figure 1). Moreover, while Black and White Philadelphians constitute about 39% and 34% of the city's population respectively, most Black residents live in census block groups that are less than 5% White (U.S. Census Bureau 2025).

Figure 1. Residential Segregation in Philadelphia Census Block Groups



Note: Derived from 5-year 2022 American Community Survey estimates

Philadelphia's intense segregation complicates cross-race comparisons. Most obviously, there is limited racial overlap in the areas where people are commonly stopped, and therefore even basic statistical comparisons of 'similarly situated' individuals from different racial groups are strained (for example, only 5% of the stops of young Black men occur in majority White districts). More important, unlike the locations of most highway stops, stops in Philadelphia's neighborhoods are not race neutral. As Sampson and Raudenbush (2004: 320) have convincingly demonstrated, "dark skin is an easily observable trait that has become a statistical marker in American society, one imbued with meanings about crime and disorder that *stigmatize not only people but also the places in which they are concentrated*" [emphasis added]. Even when sample sizes are large enough to conduct interracial analyses with geographic controls, defining what constitutes a similar situation for Whites and racial minorities in the context of segregation is conceptually fraught. For example, stops of White motorists traveling through predominately Black areas are arguably qualitatively different from stops of Black motorists in those exact same spaces; while the former might be due to an out-of-place scrutinization effect, the latter could be influenced by the interaction of individual racial identity and racialized neighborhood stigma (an intensification effect due to fitting in)—and these two separate effects need not be symmetrical in

magnitude nor necessarily occur in reverse in majority White neighborhoods (Hannon et al. 2021; Hannon and Siegel 2019).

Ultimately, although we still see great utility in the traditional cross-race veil-of-darkness test (Grogger and Ridgeway 2006), we believe that intra-racial versions of the tool should also be employed as a complement, particularly when using data from segregated urban jurisdictions (Hannon and Biddle 2025). Methodologically, one benefit of intra-racial analyses is that, like panel fixed-effects regressions, unmeasured factors associated with racial categorization can be controlled by concentrating on within-race variation. In this case, for instance, using an all-Black sample provides a robust n for subgroup statistical comparisons and mitigates concerns about difficult-to-measure confounding influences that are race-specific in nature.

While at first glance it might seem paradoxical to test for racial profiling in an all-Black sample, such thinking reflects our deeply ingrained uncritical acceptance of bureaucratic ‘state categories’ as perfectly capturing complex social constructions like race (and racism). As Monk (2022: 7) has argued, census categories used by the state are intrinsically reductionist: “The state reduces what are fundamentally continuous social differences to crisp, nominal differences that are more easily handled in bureaucratic systems...as if processes of social categorization are not continuous, graded, and subcategorical, as the contemporary interdisciplinary consensus among philosophers, social psychologists, and cognitive scientists strongly suggests.” In essence, Monk (2002) argues that researchers should supplement the use of broad state classifications of group membership by incorporating the relevance of categorical typicality on a continuum. In the context of employment discrimination, for example, the question expands from just ‘does being Black negatively impact performance evaluations’ to ‘does being more stereotypically Black negatively impact performance evaluations.’ The former question implies a non-Black reference group while the latter allows for comparisons among Black employees, perhaps using skin darkness as the clearest indicator of racial typicality.

Unfortunately, the public stop data for Philadelphia do not include assessments of driver or passenger skin complexion. Still, the public data can be configured to produce ordinal measures of the extent of racial presence in the stopped vehicle, such as the number or proportion of investigated occupants that are classified as Black. Given existing theory and research in social psychology suggesting that racial stereotypes about criminality are amplified in group contexts (Cooley et al. 2019), we operationalize the extent of racial typicality for the intra-racial component of our analyses with an indicator of racially homogenous multiple occupancy (versus single occupancy).

In addition to our focus on vehicle stops that are racially homogenous in occupant classifications, we also limit the sample to (1) investigatory stops of Black or White motorists (the two largest racial categories in Philadelphia), (2) stops exclusively involving young men between the ages of 18 and 29 (for consistency with previous research, e.g., Rosenfeld et al. 2011), (3) stops based on motor vehicle code (MVC) violations (as opposed to ‘vehicle involved in crime’ or ‘vehicle matches flash information’), (4) stops that are not associated with an arrest (only about 2.5% of the total involved an arrest), (5) records that are not exact duplicates in the

data (only about 1.7% were duplicates) or missing information on driver or investigated passenger characteristics (<1% missing), and (6) stops occurring during the inter-twilight period.

Consistent with Knode et al.'s (2024) suggested refinements to the veil-of-darkness test, we define the inter-twilight period as the interval between the earliest civil dusk time and the latest sunset time for a given location. During this portion of the day, some dates fall in full daylight while others fall in darkness. For our Philadelphia sample, this window spans from 5:08 pm to 8:35 pm. Like most veil-of-darkness studies, we utilize the `{suncalc}` package in R. In line with the traditional approach taken by Grogger and Ridgeway (2006), we exclude stops that occur in the roughly 30-minute span between sunset and dusk on any given day, since these events cannot be clearly classified as taking place in either definite daylight or definite darkness. The remaining inter-twilight stops in our focal sample are split approximately evenly between those occurring in sunlight and those occurring after dark, enabling appropriate counterfactual comparisons.

We analyze stops that are entirely homogenous across race, age group, and sex classifications for drivers and passengers to limit the complexity of the analyses (and because for multi-occupant vehicle stops, we do not have a reliable indicator of who is in the driver position—a position that may be uniquely scrutinized). We focus on stop rationales based on MVC violations as these are most likely to be 'could stop' offenses, as opposed to 'must stop' situations like spotting a specific vehicle linked to a series of robberies (Epp et al. 2014). Discretion is, of course, a necessary component of discrimination. Including 'must stop' offenses in the analyses would thus bias our results toward the null hypothesis of no difference. We also exclude the relatively small number of stops associated with an arrest to help focus the sample on cases where officers choose to act following their discretionary judgment rather than feeling compelled to act based on the intrinsic seriousness of the infraction. In our remaining sample of MVC violations, only 12% were deemed worthy of a traffic ticket.

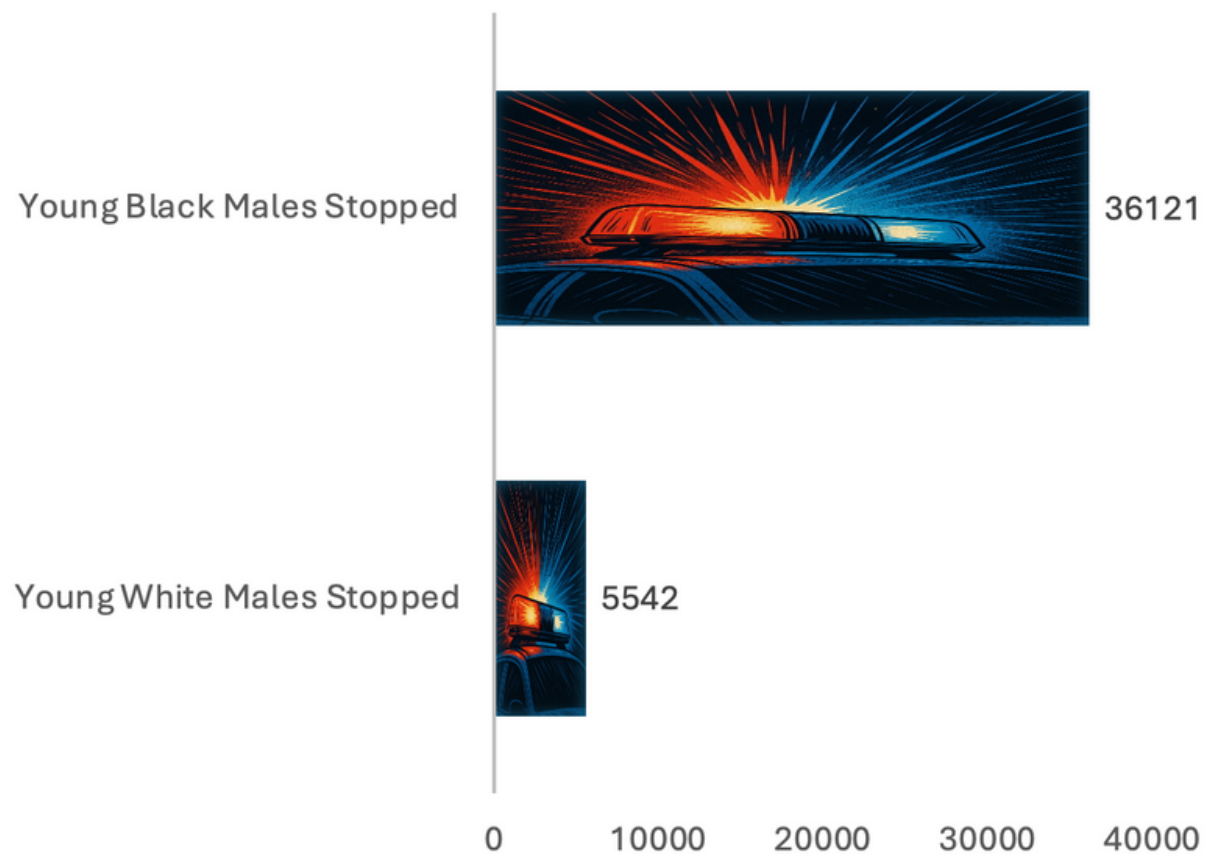
Following Grogger and Ridgeway's (2006) original formulation of the veil-of-darkness test, our primary explanatory variable is whether the stop took place in darkness, and our key control variable is clock time (minute) operationalized with a natural spline with six degrees of freedom. Our other selected controls are indicator variables for day of the week, year, police service area (PSA) where the stop occurred (65 PSAs/local beats), officer assigned unit (54 work assignments tied to police districts or cross-district initiatives), and summer month (June, July, or August).

Knode et al. (2024) recently proposed a useful weighting adjustment for the veil-of-darkness test that addresses concerns about seasonality—that is, the risk that findings may be influenced by seasonal patterns in behavior, particularly those tied to summer or winter. Their method uses a quadratic kernel estimator designed to give greater weight to stops occurring on dates when the inter-twilight interval contains a more even mix of daylight and darkness, which typically happens during spring or fall. We incorporate this weight into our logistic regression models predicting (1) whether a young male driver is Black (versus White) and (2) whether a young Black male driver is traveling with at least one young Black male passenger (versus alone).

Results

Figure 2 displays the stark disparity in the number of traffic stops for young Black males relative to young White males in Philadelphia. Despite very similar populations sizes of young Black and White men in the city (and despite young Black men being much more likely to use public transportation), young Black men were represented in the traffic stop data at a rate 6.5 times that of young White men. We suspect that a significant portion of this huge disparity reflects police deployment decisions driven by perceived neighborhood crime control needs rather than traffic safety priorities. This is what Grogger and Ridgeway (2006: 879) refer to as “neighborhood profiling.” Unfortunately, the PPD does not publicly provide deployment data, or an explicit rationale for [possibly] concentrating traffic enforcement based on crime rather than traffic patterns. It is also conceivable that some portion of the wide gap in Figure 2 can be explained by greater rates of motor vehicle code violation among young Black male drivers.

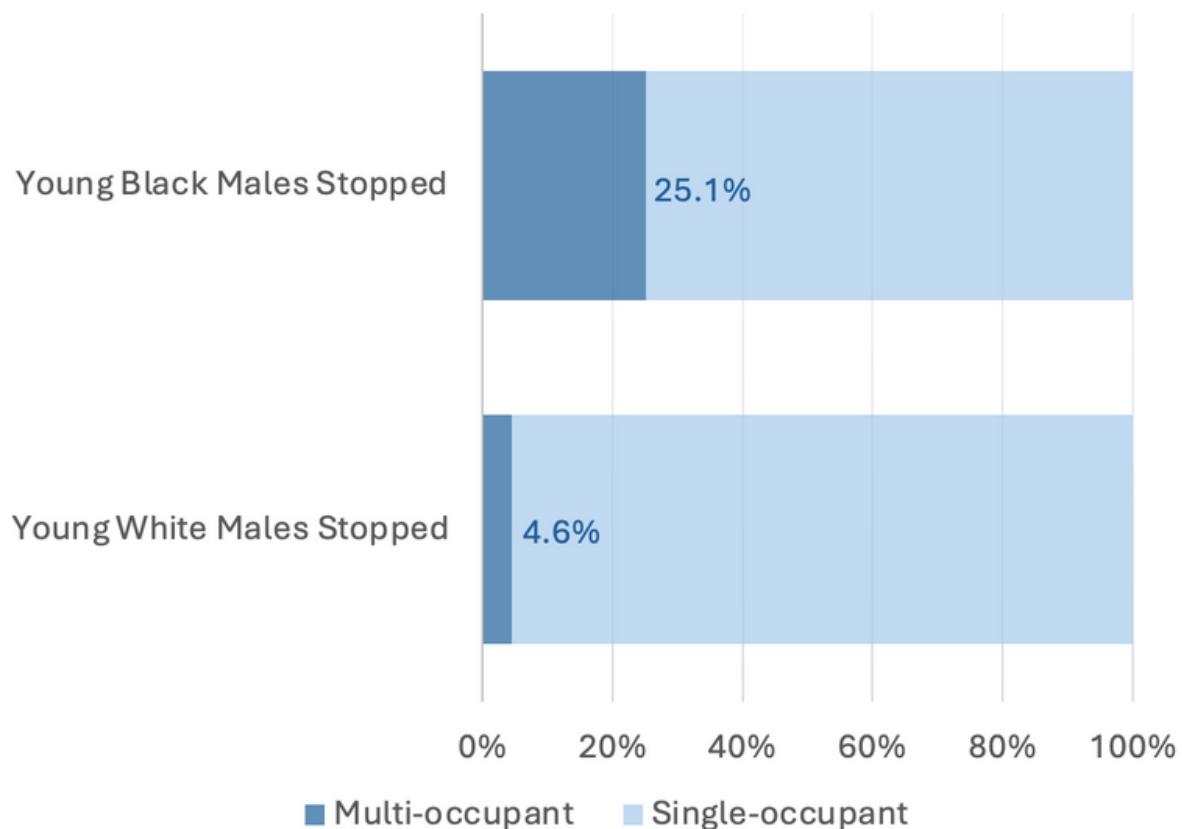
Figure 2. Frequency of Young Males Pulled Over for Philadelphia's Two Largest Racial Groups



Note: These samples focus on motor vehicle code stops occurring between 2021 and 2024 in Philadelphia during the inter-twilight period (between 5:08pm and 8:35pm). Only multi-person stops involving all young (18-29) males of the same race as the young male driver are included in each tally. None of the stops sampled involved an arrest.

Potentially greater MVC violation rates cannot, however, reasonably explain why young Black male drivers are several times more likely to be found in the company of other young Black men than young White male drivers are investigated with other young White men (see Figure 3). A more plausible explanation for this particular vast difference might be race-specific carpooling behaviors, especially considering our examined time range of 5:08 pm to 8:35 pm. Yet data from the American Community Survey's commuting component suggest that carpooling rates for those taking a private vehicle to work are roughly equivalent at around 14% for Black and White Philadelphians (U.S. Census Bureau 2025). Of course, carpooling happens outside of work commutes too, and many of those traveling through Philadelphia are not residents. Thus, as discussed earlier, the absence of a consensus benchmark for the expected rate at which a particular group 'should' be pulled over emerges [again] as a significant obstacle.

Figure 3. Percent Multi-occupant Vehicle Stops for Young Males Pulled Over by Race



Note: $n=36,121$ young Black male motorists and $n=5,542$ young White male motorists. These samples focus on motor vehicle code stops occurring between 2021 and 2024 in Philadelphia during the inter-twilight period (between 5:08pm and 8:35pm). Only those multi-person stops involving all young (18-29) males of the same race as the young male driver are included in each sample. None of the stops sampled involved an arrest.

Fortunately, however, as previously discussed, the award-winning veil-of-darkness test was constructed for the exact purpose of circumventing the benchmarking problem. We utilize this powerful statistical tool to help isolate the officer's discretionary judgment to conduct a traffic stop from the driver's decision making, whether the driver's decision making is related to venturing out to a particular destination at a particular time, committing an MVC violation, bringing along a passenger, or anything else.

Table 1 provides the traditional veil-of-darkness estimates for the log odds that a young male driver is Black. The first model includes controls for clock time, day of the week, and year, while the second model includes controls for police service area, police assignment, and summer month and is weighted to adjust for seasonality. In both cases, the veil-of-darkness coefficient is negative and statistically significant, suggesting

that an obscured view reduces the likelihood that the young male driver pulled over will be Black. For the full model with the seasonality weight, the veil-of-darkness coefficient indicates that darkness reduces the odds that the driver is Black by about 17% ($\exp -.188 = .829$). This estimate is generally in line with what others have reported for Philadelphia (Hannon et al. 2021; Pierson et al. 2020). Among the controls, one set of robust effects worth mentioning are those for the year dichotomies. The year coefficients show that 2021 had a greater proportion of young Black male drivers stopped than the years that followed—suggesting recent progress in the reduction of racial disparities in traffic stops.

Table 1: Traditional Inter-racial Veil-of-Darkness Logistic Regressions

	Predicting Whether Stopped Driver is Black	
	Time-Day-Year Controls Only	Full Controls + Seasonality Weight
	1	2
Obscured View	-.116***	-.188***
	(.032)	(.046)
Spline 1 (minute)	.114	.021
	(.092)	(.106)
Spline 2 (minute)	.169	.058
	(.114)	(.131)
Spline 3 (minute)	.217*	.116
	(.101)	(.119)
Spline 4 (minute)	.144	.215*
	(.092)	(.110)

Spline 5 (minute)	.210	.400
	(.184)	(.210)
Spline 6 (minute)	.057	.228*
	(.084)	(.100)
MONDAY	.222***	-.014
	(.054)	(.063)
TUESDAY	.221***	.041
	(.052)	(.061)
WEDNESDAY	.229***	.004
	(.053)	(.062)
THURSDAY	.210***	.008
	(.053)	(.062)
SATURDAY	-.017	-.059
	(.055)	(.064)
SUNDAY	-.113*	-.096
	(.055)	(.065)
Year 2022	-.036	-.132*

	(.048)	(.056)
Year 2023	-.459***	-.426***
	(.045)	(.053)
Year 2024	-.649***	-.614***
	(.042)	(.050)
Stop Location (65 PSAs)	---	YES
Assignment (54 Units)	---	YES
Summer (Jun-Aug)	---	YES
Seasonality Weight	---	YES
Observations	37,168	37,168
Significance level:	* P < .05	
	** P < .01	
	*** P < .001	

Note: The sample focuses on motor vehicle code stops of young males occurring between 2021 and 2024 in Philadelphia during the inter-twilight period (between 5:08pm and 8:35pm). n=31,738 young Black male drivers stopped and n=5,430 young White male drivers stopped. Only those multi-person stops involving all young men of the same race as the young male driver are included. None of the stops sampled involved an

arrest. Quasi-binomial models are employed to adjust for under/overdispersion, particularly with the seasonality weight.

Table 2 provides the veil-of-darkness estimates for a non-traditional dependent variable (whether the vehicle has passengers) in a non-traditional homogenous sample (exclusively young Black males). As noted earlier, intra-racial analysis can serve as an important complement to the classic inter-racial implementation of the veil-of-darkness test and arguably may be better suited for data associated with a segregated urban environment. For the intra-racial analysis in Table 2, the test gauges whether an obscured view reduces the log odds of young Black males being stopped in groups versus solo. Substantively, the question is whether more visibility of who is in the car is associated with more stops of vehicles carrying multiple Black occupants deemed worthy of investigation.

Table 2: Novel Intra-racial Veil-of-Darkness Logistic Regressions		
	Predicting Whether Stopped Black Driver Has Black Passenger(s)	
	Time-Day-Year Controls Only	Full Controls + Seasonality Weight
	1	2
Obscured View	-.242***	-.269***
	(.035)	(.044)
Spline 1 (minute)	.104	.210*
	(.102)	(.103)
Spline 2 (minute)	.067	.061
	(.123)	(.125)
Spline 3 (minute)	.319**	.380***
	(.109)	(.112)

Spline 4 (minute)	.131	.198
	(.098)	(.104)
Spline 5 (minute)	.087	.175
	(.207)	(.207)
Spline 6 (minute)	.318 ^{***}	.397 ^{***}
	(.092)	(.095)
MONDAY	-.131 [*]	-.107
	(.060)	(.061)
TUESDAY	-.018	-.066
	(.056)	(.057)
WEDNESDAY	.089	.042
	(.056)	(.057)
THURSDAY	-.025	-.032
	(.057)	(.058)
SATURDAY	.015	.071
	(.062)	(.062)
SUNDAY	-.257 ^{***}	-.237 ^{***}
	(.067)	(.069)

Year 2022	-.038	-.067
	(.044)	(.046)
Year 2023	-.156***	-.160***
	(.046)	(.047)
Year 2024	-.220***	-.113*
	(.044)	(.045)
Stop Location (65 PSAs)	---	YES
Assignment (54 Units)	---	YES
Summer (Jun-Aug)	---	YES
Seasonality Weight	---	YES
Observations	31,738	31,738
Significance level:	* P < .05	
	** P < .01	
	*** P < .001	

Note: The sample focuses on motor vehicle code stops of young Black males occurring between 2021 and 2024 in Philadelphia during the inter-twilight period (between 5:08pm and 8:35pm). n=27,056 young Black male drivers stopped alone and n=4,682 young Black male drivers stopped with at least one other young Black male

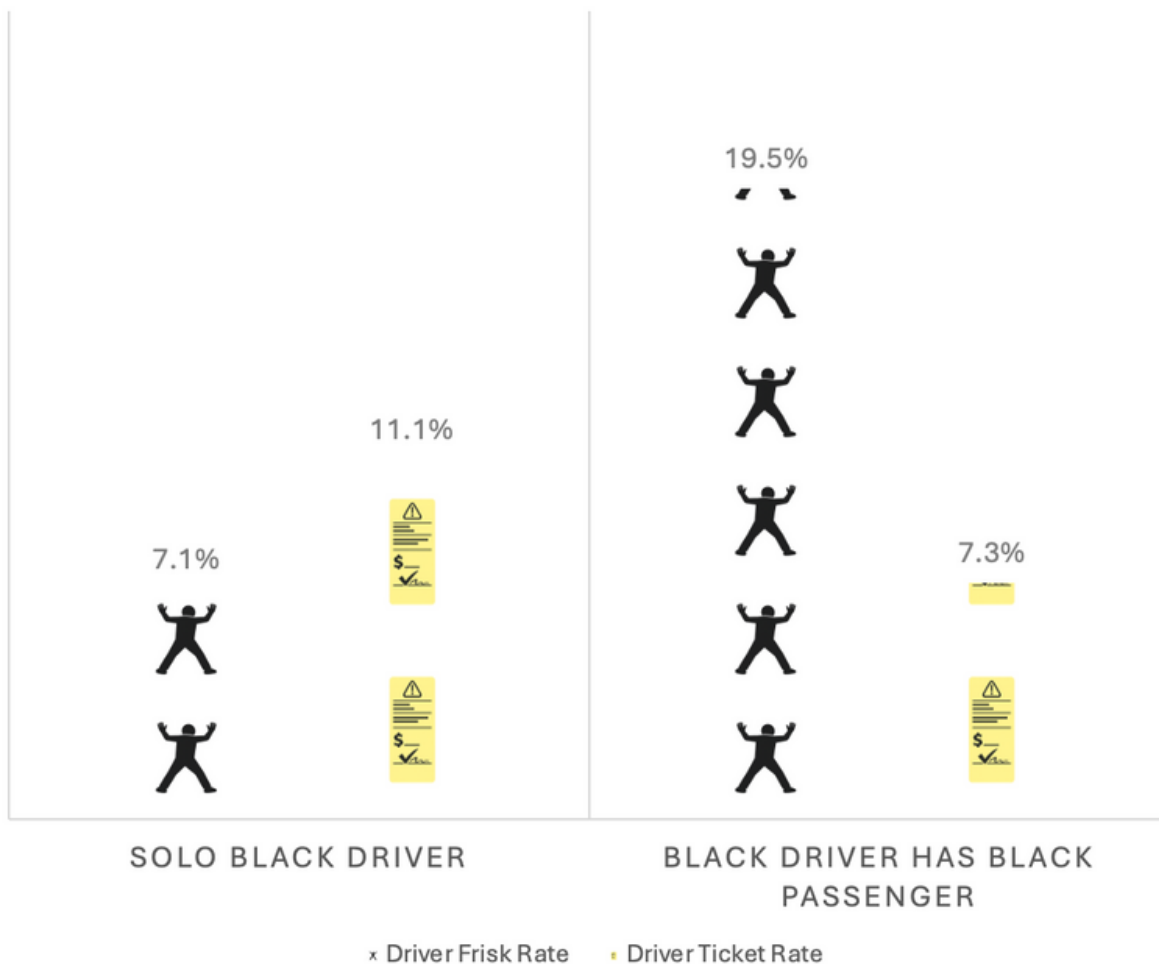
in the vehicle. Only those multi-person stops involving all young Black men are included. None of the stops sampled involved an arrest. Quasi-binomial models are employed to adjust for under/overdispersion, particularly with the seasonality weight.

The results from our intra-racial analysis align well with our inter-racial findings. As in Table 1, both the reduced and full model with sampling weight produced negative and statistically significant veil-of-darkness coefficients (Table 2). The full model estimate in Table 2 suggests that darkness reduces the odds that a young Black male driver will have a young Black male passenger by approximately 24% ($\exp -.269 = .764$). Overall, these results imply that young Black men in groups are significantly more likely to be targeted for pretextual traffic stops than young Black men driving alone, who previous analyses show are significantly more likely to be targeted than others.

Despite the relatively small sample size, we also tried estimating intra-racial models predicting the presence of young White male passengers for vehicles driven by young White men. The veil-of-darkness estimates were always statistically insignificant at the .05 level. We tried more parsimonious models, expanding the sample to include other years, relaxing sampling restrictions, and using robust regression techniques. In the end, we found no meaningful evidence of any bias, positive or negative, associated with the concentration of multiple young White men in a vehicle.

Also supplemental to our main investigation, we explored some of the potential consequences of an elevated rate of multi-person vehicle stops for young Black men [using the data available to us]. Consistent with Cooley et al.'s (2019) results for the policing of pedestrians, we found that frisk rates of young Black male drivers were nearly three times higher for stops involving at least one young Black male passenger (see Figure 4). Thus, Cooley et al.'s theory concerning the amplification of racial threat stereotypes in group contexts appears supported in this exploratory analysis. The substantially higher frisk rate observed in our data was associated with a reduced ticket rate for drivers, although the consolation of a lower ticket rate may actually be indicative of how stops of young Black men in groups are more likely to be pretextual in the first place (see our earlier discussion of the details underlying the infamous *Whren* case). Arguably, the fact that young Black men traveling together are 2-3 times more likely to be fruitlessly frisked than to receive a traffic ticket speaks to the centrality of pretextual policing practices in Philadelphia.

Figure 4. Frisk and Ticket Rates for Young Black Male Drivers by Whether There is a Young Black Male Passenger in the Car



Note: These samples focus on motor vehicle code stops of young Black males occurring between 2021 and 2024 in Philadelphia during the inter-twilight period (between 5:08pm and 8:35pm). $n=27,056$ young Black male drivers stopped alone, $n=4,682$ young Black male drivers stopped with at least one other young Black male in the vehicle. Only those multi-person stops involving all young Black men are included. None of the stops sampled involved an arrest.

Conclusion

In practical terms, the *Whren v. United States* decision (1996) means that as long as police officers have good reason to suspect that a traffic infraction has occurred, they can stop and investigate motorists for serious crimes as an ulterior motive. The Supreme Court rejected the claim that pretextual vehicle stops are a violation of 4th Amendment protections against unreasonable search and seizure. Critics of the decision have argued that

the Court's judgement renders officer intentions irrelevant and opens the door to racial profiling, as police officers can always find a technical justification for pulling someone over (Epp et al. 2014; Harris 2002). Still, the Court was clear that vehicle stop practices could be challenged on 14th Amendment equal protection grounds if it could be shown that selective enforcement was based on race or some other protected status.

Our findings indicate that pretextual traffic stops in Philadelphia do not live up to the promise of the 14th Amendment; a significant number are based on the visibility of driver and passenger race, age, and sex. In particular, young Black men traveling together are investigated much more intensely than others. This is true both in terms of the frequency of stops and the level of stop intrusiveness. For instance, young Black male drivers stopped with young Black male passengers are almost three times more likely to be removed from the vehicle and frisked (without arrest) than they are to receive a ticket for the offense allegedly motivating the encounter in the first place.

Our empirical findings align with previous research in criminology on the historical and ongoing typification of crime as a Black phenomenon (Grunwald et al. 2022; Welch 2007) as well as work in social psychology suggesting that the application of racial stereotypes is amplified in police encounters with people walking in groups of two or more (Cooley et al. 2019). Our results reinforce earlier analyses of data for Philadelphia indicating that daylight visibility increases the ratio of stopped young Black male drivers to stopped young White male drivers (e.g., Hannon et al. 2021) but further specify that racial disparities are likely greater when the sample and analytic approach are broadened to incorporate race-of-passenger effects.

More generally, in terms of methodology, our findings underscore concerns about aggregation bias in the traditional application of the veil-of-darkness test where weak or null results for inter-racial comparisons may be due to lack of attention to differences in theoretically relevant intra-racial variation (Hannon and Biddle 2025; Taniguchi et al. 2017). Put differently, it is important to consider the possibility that the overall ratio of Black to non-Black drivers might change very little while the 'type' of motorists stopped could vary in race-specific ways with greater ease of visual identification by officers. This possibility increases in the context of urban racial segregation. Future research should continue to apply the veil-of-darkness test both inter-racially and intra-racially to assess both equity in the relative quantity of stops and in the distributional quality or character of those stops. Where the data allow it, the single versus multi-occupant nature of vehicular investigations should be examined with both intra- and inter-racial veil-of-darkness tests. While previous inter-racial analyses have already documented how officers particularly scrutinize young Black male drivers relative to young White male drivers, our innovative intra-racial models suggest that even in a sample of all young Black male drivers, a greater presence of young Black men in the vehicle means a greater chance of being pulled over for selective traffic enforcement. (There was little evidence in our data that young White men in groups were detained any differently than young White men driving alone.)

Besides further distinction between inter and intra-racial dynamics in the context of segregation, future research can build on our analyses in several other ways. First, for the sake of parsimony, we did not examine race-specific contraband recovery rates for single versus multi-occupant vehicle searches. Future analyses in this vein would benefit from recent advances in statistical techniques and available code for handling issues of inframarginality (Simoiu et al. 2017). Second, although our data included both vehicle driver and passenger information that could easily be linked to specific stops, we did not have a definitive indicator variable for who was driving in multi-occupant situations. For simplicity in presentation and modeling, we assumed the driver was the first listed record and focused the sample on stops where driver and passenger demographics were equivalent. Likewise, we did not have any data about where investigated passengers were located in the vehicle. Additional research with richer data could explore whether physical position in the vehicle matters. Third, while the veil-of-darkness test is a strongly conservative gauge of officer discrimination, false positives are still theoretically possible (in this case for example, while we think it highly implausible, daylight might motivate get-togethers much more for young Black men than young White men). Future ethnographic and survey work should ask police officers about how the constitutionality of pretextual stops as a crime fighting tool affects discretionary decisions to enforce traffic laws (see, for example, Stuart VI 2022). Our conclusions would be bolstered by direct evidence indicating that officers perceive young Black men in groups as especially criminally suspect relative to young Black men driving alone.

In addition, it is important to consider the obvious limits of generalizability of a city-level veil-of-darkness test: it only examines a narrow timeframe around rush hour(s), it says nothing about how diffuse profiling behavior is within the relevant police agency, and it is limited to just one municipality's fallible data recording efforts. Nevertheless, the veil-of-darkness is widely regarded as one of the best statistical tests available for uncovering unequal and unlawful policing of vehicle stops (Knode et al 2024; Smith et al. 2021). Moreover, while repeated test results indicating bias cannot tell you how many officers are significantly engaged in bias-based policing, such results can tell you something about whether such behavior is normatively accepted (or perhaps expected) by the organization. Likewise, as one of the oldest and largest police forces in the United States, Philadelphia's police department is certainly not average—but that is also arguably why it might set trends elsewhere.

In sum, our findings add to a considerable body of evidence suggesting that, across the United States, Black people are policed in a fundamentally different manner than others, and stereotypes of criminality likely play an important role in this differential treatment. Taken for granted freedoms, such as the right to travel on public streets without excessive scrutiny, appear more tenuous for Black citizens. As Cooley (2020) has noted, even commonsense mottos like 'there is safety in numbers' might be less sensible for young Black men fearful of being targeted by the police. Along the same lines, while timing a road trip with friends to maximize daylight visibility might seem like the universally safest choice, the enhanced threat of racial profiling through the glass can complicate the calculation for Black motorists.

References

- Alexander, M. (2010). *The new Jim Crow: Mass incarceration in the age of colorblindness*. The New Press.
- Anderson, E. (2015). The white space. *Sociology of Race and Ethnicity*, 1(1), 10–21.
<https://doi.org/10.1177/2332649214561306>
- Baranauskas, A. J., & Drakulich, K. M. (2018). Media construction of crime revisited: Media types, consumer contexts, and frames of crime and justice. *Criminology*, 56(4), 679–714. <https://doi.org/10.1111/1745-9125.12189>
- Butler, P. (2017). Race and adjudication. In E. Luna (Ed.), *Reforming criminal justice* (Vol. 3, pp. 211–226). Academy for Justice.
- Carbado, D., & Feingold, J. (2022). Rewriting *Whren v. United States*. *UCLA Law Review*, 68, 1678–1701.
- Chiricos, T., Welch, K., & Gertz, M. (2004). Racial typification of crime and support for punitive measures. *Criminology*, 42(2), 358–390. <https://doi.org/10.1111/j.1745-9125.2004.tb00523.x>
- Cobb, J. (2025). *Three or more is a riot: Notes on how we got here 2012–2025*. Penguin Random House.
- Cooley, E., & Payne, B. K. (2017). Using groups to measure intergroup prejudice. *Personality and Social Psychology Bulletin*, 43, 46–59. <https://doi.org/10.1177/0146167216675331>
- Cooley, E., & Payne, B. K. (2018). A group is more than the average of its parts: Why existing stereotypes are applied more to the same individuals when viewed in groups than when viewed alone. *Group Processes and Intergroup Relations*, 1–15. <https://doi.org/10.1177/1368430218756494>
- Cooley, E., Hester, N., Cipolli, W., Rivera, L. I., Abrams, K., Pagan, J., Sommers, S. R., & Payne, K. (2019). Racial biases in officers' decisions to frisk are amplified for Black people stopped among groups. *Social Psychological and Personality Science*, 11(6), 761–769. <https://doi.org/10.1177/1948550619876638>
- Cooley, E. (2020). When safety in numbers becomes racism in numbers. *Society for Personality and Social Psychology Blog*. <https://spsp.org/news-center/character-context-blog/when-safety-numbers-becomes-racism-numbers>
- Epp, C., Maynard-Moody, S., & Haider-Markel, D. (2014). *Pulled over: How police stops define race and citizenship*. University of Chicago Press.
- Grogger, J., & Ridgeway, G. (2006). Testing for racial profiling in traffic stops from behind a veil of darkness. *Journal of the American Statistical Association*, 101(475), 878–887.
<https://doi.org/10.1198/016214506000000168>

Grunwald, B., Nyarko, J., & Rappaport, J. (2022). Police agencies on Facebook overreport on Black suspects. *Proceedings of the National Academy of Sciences*, 119(45). <https://doi.org/10.1073/pnas.2203089119>

Hannon, L. (2019). Neighborhood residence and assessments of racial profiling using census data. *Socius: Sociological Research for a Dynamic World*, 5. <https://doi.org/10.1177/2378023118818746>

Hannon, L., & Siegel, A. (2019). Racially profiling people and places. *Contexts*, 18(2), 66–67. <https://doi.org/10.1177/1536504219854723>

Hannon, L., Neal, M., & Gustafson, A. R. (2021). Out-of-place and in-place policing: An examination of traffic stops in racially segregated Philadelphia. *Crime & Delinquency*, 67(6–7), 868–890. <https://doi.org/10.1177/0011128720926122>

Hannon, L., & Biddle, M. (2025). Unequal policing of Black motorists in Black communities by age and gender: A short report on intraracial applications of the veil-of-darkness test. *American Journal of Criminal Justice*, 50, 1081–1090. <https://doi.org/10.1007/s12103-025-09879-8>

Harris, D. (2002). *Profiles in injustice: Why racial profiling cannot work*. The New Press.

Horrace, W. C., & Rohlin, S. M. (2016). How dark is dark? Bright lights, big city, racial profiling. *The Review of Economics and Statistics*, 98(2), 226–232. https://doi.org/10.1162/REST_a_00543

Insko, C. A., & Schopler, J. (1998). Differential distrust of groups and of individuals. In C. Sedikides, J. Schopler, & C. A. Insko (Eds.), *Intergroup cognition and intergroup behavior* (pp. 75–107). Erlbaum.

Intravia, J., & Pickett, J. T. (2019). Stereotyping online? Internet news, social media, and the racial typification of crime. *Sociological Forum*, 34(3), 616–642. <https://doi.org/10.1111/socf.12515>

Kalinowski, J., Ross, S. L., & Ross, M. B. (2021). *Endogenous driving behavior in tests of racial profiling in police traffic stops* (NBER Working Paper No. 28789). <https://doi.org/10.3386/w28789>

Knode, J. L., Wolfe, S. E., & Carter, T. M. (2024). Pulling back the veil of darkness: A proposed road map to disentangle racial disparities in traffic stops. *Criminology*, 62, 364–375. <https://doi.org/10.1111/1745-9125.12366>

Massey, D. S., & Denton, N. A. (1998). *American apartheid: Segregation and the making of the underclass* (2nd ed.). Harvard University Press.

McGlynn, R. P., Harding, D. J., & Cottle, J. L. (2009). Individual–group discontinuity in group–individual interactions: Does size matter? *Group Processes & Intergroup Relations*, 12(1), 129–143. <https://doi.org/10.1177/1368430208098781>

Mihalache, D., Lamer, S. A., Allen, J., Maher, M., & Sweeny, T. D. (2021). Anger bias in the evaluation of crowds. *Journal of Experimental Psychology: General*, 150(9), 1870–1889.

<https://doi.org/10.1037/xge0001025>

Monk, E. P. (2022). Inequality without groups: Contemporary theories of categories, intersectional typicality, and the disaggregation of difference. *Sociological Theory*, 40(1), 3–27.

<https://doi.org/10.1177/07352751221076863>

Philadelphia Police Department. (2024). *Vehicle & pedestrian investigations* [Data set].

<https://www.opendataphilly.org/dataset/vehicle-pedestrian-investigations>

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>

Quillian, L., & Pager, D. (2001). Black neighbors, higher crime? The role of racial stereotypes. *American Journal of Sociology*, 107, 717–767. <https://doi.org/10.1086/338938>

Ritter, J. A., & Bael, D. (2009). Detecting racial profiling in Minneapolis traffic stops: A new approach. *CURA Reporter*, 11–17.

Rosenfeld, R., Rojek, J., & Decker, S. (2011). Age matters: Race differences in police searches of young and older male drivers. *Journal of Research in Crime and Delinquency*, 49(1), 31–55.

Russell-Brown, K. (1998). *The color of crime: Racial hoaxes, white fear, Black protectionism, police harassment, and other macroaggressions*. New York University Press.

Sampson, R. J., & Raudenbush, S. W. (2004). Seeing disorder: Neighborhood stigma and the social construction of broken windows. *Social Psychology Quarterly*, 67, 319–342.

Simoiu, C., Corbett-Davies, S., & Goel, S. (2017). The problem of inframarginality in outcome tests for discrimination. *The Annals of Applied Statistics*, 11(3), 1193–1216. <https://doi.org/10.1214/17-AOAS1058>

Smith, M., Tillyer, R., Lloyd, C., & Petrocelli, M. (2021). Benchmarking disparities in police stops: A comparative application of 2nd- and 3rd-generation techniques. *Justice Quarterly*, 38(3), 513–536.

<https://doi.org/10.1080/07418825.2019.1660395>

Stuart, J. E. B. VI. (2022). Pretextual stops: The rest of the story. *University of Richmond Law Review*, 56, 1–18.

Taniguchi, T. A., Hendrix, J. A., Levin-Rector, A., Aagaard, B. P., Strom, K. J., & Zimmer, S. A. (2017). Extending the veil of darkness approach: An examination of racial disproportionality in traffic stops in Durham, NC. *Police Quarterly*, 20(4), 420–448. <https://doi.org/10.1177/1098611117721665>

United States Census Bureau. (2025). *American Community Survey 2022* [Data set]. <https://www.census.gov/acs/www/data/data-tables-and-tools/data-profiles/2022/>

Vito, A. G., Rocheleau, G., Teed, E., & Dent, L. (2026). Daylight, race, gender, and age: Testing the veil of darkness and luxury brand car to examine racial profiling. *International Journal of Police Science & Management*. <https://doi.org/10.1177/14613557251409947>

Wacquant, L. (2001). Deadly symbiosis: When ghetto and prison meet and mesh. *Punishment & Society*, 3(1), 95–133. <https://doi.org/10.1177/14624740122228276>

Welch, K. (2007). Black criminal stereotypes and racial profiling. *Journal of Contemporary Criminal Justice*, 23(3), 276–288. <https://doi.org/10.1177/1043986207306870>

Whitney, M. (2008). The statistical evidence of racial profiling in traffic stops and searches: Rethinking the use of statistics to prove discriminatory intent. *Boston College Law Review*, 49, 263–299.

Whren v. United States, 517 U.S. 806 (1996).

Worden, R. E., McLean, S. J., & Wheeler, A. P. (2012). Testing for racial profiling with the veil-of-darkness method. *Police Quarterly*, 15, 92–111.