



Complicating Colorism: Race, Skin Color, and the Likelihood of Arrest

Amelia R. Branigan¹, Christopher Wildeman², Jeremy Freese³, and Catarina I. Kiefe⁴

Socius: Sociological Research for a Dynamic World
 Volume 3: 1–17
 © The Author(s) 2017
 Reprints and permissions:
sagepub.com/journalsPermissions.nav
 DOI: 10.1177/2378023117725611
srd.sagepub.com



Abstract

Both conventional public beliefs and existing academic research on colorism presuppose that variation in skin color predicts social outcomes among minorities but is inconsequential among whites. The authors draw on social psychological research on stereotyping to suggest that in quick, low-information decisions such as an arrest, the opposite may be true. Contrary to findings for longer-term socioeconomic outcomes, the authors find that black men's probability of arrest remains constant across the spectrum of skin color, while white men's probability of arrest decreases continuously with lighter skin. Beyond posing an exception to the modern conception of colorism, these results have implications for efforts to ameliorate the epidemic of incarceration among black men, as well as for understanding how elements of visible phenotype may serve as a unique category of predictors in models of social inequality.

Keywords

skin color, race, arrests, incarceration

The descriptive “person of color” makes explicit the common understanding that racial and ethnic minorities have skin color, while white people functionally do not. From a literal standpoint, this distinction is of course false, as individuals identifiable as white do have a skin color. But the popular conception that the color of white individuals' skin is irrelevant for social outcomes has important implications for stratification research, as it begs the question of whether color predicts social outcomes only as a categorical indicator of race or as a separate continuous quantity operating also within racial designations.

Both constructions of color, the categorical and the continuous, present plausible hypotheses regarding how skin color may operate as a determinant of social outcomes. Prior research on the socioeconomic consequences of skin color have found that color does indeed operate as a continuous predictor within race in a variety of contexts, ranging from wages (Goldsmith, Hamilton, and Darity 2007; Keith and Herring 1991) to education (Branigan et al. 2013; Hersch 2006; Hunter 2002; Keith and Herring 1991) to the marriage market (Goering 1972; Hughes and Hertel 1990; Udry, Bauman, and Chase 1971). In contrast, a categorical construction of skin color would imply that color is relevant only in its association with a racial group identity that is associated with differences in a given outcome, while variation in

skin color within race is not associated with differences in that outcome. This would be the case if, for example, dark-skinned black men are not arrested more frequently than light-skinned black men, even though it is well established that black men are arrested at a far higher frequency than whites (Federal Bureau of Investigation 2014). The vast majority of inequality research implicitly assumes such categorical construction by controlling for race without additional measures of skin color.

That color is so rarely considered separately from race in models of social stratification leaves gaps in our understanding of when color does function only as an indicator of race, as opposed to when it is associated with variation in social outcomes even within a single racial group. This omission becomes practically important when attempts are made to

¹University of Illinois at Chicago, IL, USA

²Cornell University, Ithaca, NY, USA

³Northwestern University, Evanston, IL, USA

⁴University of Massachusetts Medical School, Worcester, MA, USA

Corresponding Author:

Amelia R. Branigan, University of Illinois at Chicago, Department of Sociology, 4112 Behavioral Sciences Building, 1007 West Harrison Street (MC 312), Chicago, IL 60607, USA.
 Email: arbi94@uic.edu



ameliorate a racial disparity, such as in the probability of arrest, without interrogating the extent to which the observed inequity is one of race versus also of color. Indeed, despite the extensive public and academic discourse surrounding the disproportionate arrest and incarceration rates of black Americans (Dumont et al. 2012; Federal Bureau of Investigation 2014), research on the relationship between skin color and interactions with the criminal justice system remains limited. The small body of research on this topic has typically found evidence for the continuous relevance of skin color: lighter-skinned minorities are sentenced less severely than darker-skinned minorities (Burch 2015; Viglione, Hannon, and DeFina 2011). Color may affect likelihood of incarceration, although King and Johnson (2016) found evidence of a consistent relationship only among the darkest-skinned black arrestees.¹ The sole study of the relationship between skin color and arrests was equally inconclusive (White 2015).²

The traditional understanding of colorism would suggest that skin color should operate continuously within race in predicting a given social outcome among minorities, while variation in skin color within race should not be associated with differences in that outcome among whites (e.g., Darity, Dietrich, and Hamilton 2005; Hochschild and Weaver 2007; Pearce-Doughlin, Goldsmith, and Hamilton 2013). Here we draw on social psychological research on stereotyping to suggest that for quick low-information decisions, such as an arrest (Smith 1986), the opposite may be true. Individuals have been repeatedly found to perceive more physical variation among social categories of which they are members than among categories to which they do not belong (Linville, Fischer, and Salovey 1989; Quattrone and Jones 1980); termed the “out-group homogeneity effect,” this phenomenon describes the bias expressed in phrases such as “they all look alike, but we don’t” (Quattrone and Jones 1980:142). Arrest likelihood stands as a unique outcome relative to all

others thus far considered in the colorism literature in that an arrest is a binary decision, made by single (or few) individuals during short interactions. By comparison, long-term socioeconomic outcomes—such as educational attainment, income, occupation, and marital choices—all reflect an accumulation of complicated interactions and decisions, both by individuals and an array of relevant gatekeepers, such as teachers, employers, and potential mates. Even sentencing outcomes involve a range of gatekeepers, such as judges and juries, carefully considering a detailed battery of background information on the individual being sentenced. The comparative lack of opportunity to exchange individuating information on a potential arrestee can be expected to exacerbate the extent to which group stereotypes become relevant for decision making (Ostrom and Sedikides 1992).

In the United States today, 75 percent of law enforcement officers are white (Reaves 2010); in our sample from the mid-1980s, three quarters of respondents lived in cities where more than 95 percent of law enforcement officers were white.³ The decision to make an arrest will thus be overwhelmingly made by white officers, who may, as per the out-group homogeneity effect, simply perceive less physical variation among individuals who are not also white. As such, continuous variation in skin color may well predict white men’s probability of arrest, as white officers may be more likely to perceive physical differences between same-race others even in a short interaction. Relative lightness may serve as a quick shorthand indicator for certainty of in-group membership among white men in this context, with darkness connoting a lower likelihood of immediate visual classification as in-group by a white arresting officer. Within-race variation in skin color is expected to be less relevant among black men, whom white officers may be more likely to perceive as physically homogenous.

Here we test for that pattern of associations, asking whether skin color is associated with arrest among black and white men separately. Using data from the Coronary Artery Risk Development in Young Adults (CARDIA) study, we find that the probability of arrest is indeed constant across skin color among black men, while darker skin is associated with higher arrest likelihood among white men. This finding poses an exception to the standard expectations regarding how colorism functions, and suggests a need to interrogate basic assumptions about how, for whom, and in what contexts skin color becomes socially relevant. From a practical standpoint, our results have potential implications for efforts to address the incarceration epidemic among black men, as they may suggest a disparity in how white versus black individuals are perceived by gatekeepers in the critical decision-making moments that make up a criminal record. Finally, as measurable aspects of

¹King and Johnson also found a continuous relationship between an interviewer-coded measure of skin color and likelihood of incarceration among whites, wherein darker skin color was associated with higher likelihood of incarceration. However, in a sample of 264 white offenders who were reportedly non-Hispanic, this relationship was driven entirely by 13 individuals with identifiably Hispanic surnames. The relationship between “Afrocentric features” and incarceration among white respondents in their sample was robust to controls for Hispanic surname.

²White (2015) found a significant bivariate association between skin color and arrests among black respondents in the National Longitudinal Study of Adolescent to Adult Health data, but this effect disappeared after controlling for sex. Because women are both lighter than men on average in every population thus far assessed for skin color (Jablonski and Chaplin 2000) and are also far less likely to be arrested than men (Snyder 2011), it is unclear whether White’s finding reflects colorism versus differences by sex in both skin color and arrest probability.

³More detail on the relative percentage of white police in the cities in our sample at the time of data collection is presented in the “Data and Methods” section.

the body are becoming increasingly common in models of social stratification, this finding demonstrates a need to consider how visible phenotypic characteristics may operate differently from more traditional predictors of socioeconomic outcomes, as the physical body need only be perceived by others to be socially meaningful.⁴

Background

Although the material advantages of skin lightness during slavery undoubtedly had lasting effects on the legacy of American colorism (Franklin 2000; Hill 2000; Myrdal 1944), preference for particular skin coloration is far from unique to the U.S. context (van den Berghe and Frost 1986). Anthropologists have observed skin color valuation in societies ranging across all inhabited continents, with great variation in cultural practices, level of development, and colonial history (van den Berghe and Frost 1986). In the vast majority of cases, the social preference is for lightness (van den Berghe and Frost 1986).

Yet despite the breadth of societies in which skin color functions as a social sorting mechanism (van den Berghe and Frost 1986), the assumption that skin color variation is irrelevant for white individuals has been pervasive not only as conventional lay knowledge, but also within the sociological literature on skin color (Branigan et al. 2013). A small number of analyses have used white respondents as a homogeneous comparison group against which to interpret the associations between color and socioeconomic outcomes among minorities (e.g., Goldsmith et al. 2007), but the majority of studies that consider skin color differences simply omit white respondents altogether.⁵ The rationale for this omission is rooted in key assumptions regarding how colorism operates. As per Hochschild and Weaver (2007), colorism is when people

attribute higher status and grant more power and wealth to one group, typically those designated as white . . . [and then] attribute higher status and grant more power and wealth to people of one complexion, typically light skin, within the groups designated as non-white. (p. 646)

Goldsmith et al. (2007) similarly described a “preference for whiteness,” in which minorities are differentially advantaged by the extent to which they visibly resemble the white in-group. In sum, existing research suggests that first we stratify skin color categorically by race; then we sort continuously, but only among nonwhites (Hochschild and Weaver 2007).

⁴Here we differentiate measurable phenotypic characteristics, such as skin color, from social categories such as race and gender that are frequently associated with phenotypic differences but are not themselves defined by visible aspects of the body.

⁵For two recent exceptions, see Branigan et al. (2013) and King and Johnson (2016).

This understanding of colorism is self-reinforcing in the academic literature, as the few social surveys that have collected information on respondent skin color have almost exclusively relied on interviewer coding scales that functionally preclude within-race analyses of white respondents.⁶ White Americans have only about half the variance in skin color (at least in terms of percent reflectance) as do black Americans (Branigan et al. 2013), and all color-coding instruments thus far used have required interviewers to code all respondents on a single scale. The comparatively limited variance among whites, combined with the small number of coding categories, results in little variation being actually captured: across all available data sources using categorical interviewer coding, the vast majority of white respondents consistently fall into the lightest one or two color categories. Although other critiques have been leveled at interviewer-coded color ratings—for example, concern that subjective factors other than a respondent’s skin may also be captured in the color measurement (Caruso, Mead, and Balcetis 2009; Hill 2002a)—a critical constraint of interviewer-coded color scales lies in their construction around a presupposition that skin color is socially relevant only for minorities. To avoid that presupposition, here we use a data source with a mechanical reading of skin color (skin reflectance), which captures enough variation among white respondents for separate analysis by self-reported race to be feasible.

Limiting research on the relationship between skin color and social outcomes to minorities is problematic not only in that it renders colorism invisible among whites, but also because comparing patterns of association between race may help better contextualize findings within-race. Branigan et al. (2013), for example, found an association between skin lightness and educational attainment among black respondents of both sexes as well as among white women, but not among white men, suggesting that in the educational sphere, colorism may apply to anyone who does not occupy the dominant position at the intersection of race and sex. Conceptualizing associations between skin color and social outcomes as comparative relationships between social groups also allows increased engagement with related lines of social psychological research theorizing how individuals may be expected to perceive others who are more or less like themselves (“in-group” vs. “out-group”).

Sociological research on how skin color is associated with socioeconomic outcomes is generally framed within more

⁶Surveys in which interviewer-coded color data have been used to investigate socioeconomic disparities by color include the National Survey of Black Americans (Gullickson 2005; Hersch 2006; Hill 2002b; Hughes and Hertel 1990; Keith and Herring 1991; Thompson and Keith 2001), the Multi-City Study of Urban Inequality (Hersch 2006; Hill 2002a), multiple waves of the General Social Survey (Gullickson 2005), the Detroit Area Study (Hersch 2006), and the New Immigrant Survey (Hersch 2008). The color scales contain between 3 and 11 categories.

macro-level social theory—the social construction of race, how institutional dynamics produce social gradients in skin tone (e.g., Gullickson 2005; Keith and Herring 1991)—while rarely engaging more micro-level theories of when and why visible physical characteristics such as skin color might be perceived differently in one-on-one interactions (Hill 2002a). Hill stands as a unique example in this respect, testing the out-group homogeneity effect in a large social survey and finding that as predicted, both black and white interviewers do indeed report greater physical variation among same-race respondents than other-race respondents. Mean skin color among other-race respondents was also exaggerated, such that black interviewers perceived white individuals as much lighter than did white interviewers, while white interviewers perceived black individuals as much darker than did black interviewers (Hill 2002a). As Hill concluded, these results suggest a problematically limited ability to distinguish physical difference among other-race persons.

Far from being a new or marginal area of research, studies of the out-group homogeneity effect and the related “cross-race effect”—the phenomenon of individuals being able to more accurately recognize same-race faces than cross-race faces (Malpass and Kravitz 1969; Young et al. 2012)—spans more than a century (Feingold 1914). This body of work offers suggestions into the characteristics of a given interaction that may affect how stereotyping is triggered: in particular, a lack of opportunity for the exchange of individuating information, such as in the time-constrained context of an arrest decision, may be expected to exacerbate the extent to which group stereotypes will be used for decision making (Ostrom and Sedikides 1992). It is on this basis that we pose arrests as a unique outcome relative to the longer-term measures typically considered in population research on colorism, and suggest the potential for color to matter quite differently in this case than in prior literature.

Data and Methods

The CARDIA study is a widely-used health-related cohort study collected by the National Heart, Lung, and Blood Institute (NHLBI). Data collection has been carried out in eight waves 2 to 5 years apart, beginning in 1985 with 5,115 community-dwelling non-Hispanic blacks and whites ages 18 to 30 years and continuing to the present (data are currently available through 2010–2011). Respondents were randomly selected after stratification by race, sex, age, and education in four U.S. cities: Birmingham, Alabama; Chicago; Minneapolis; and Oakland, California (Hughes et al. 1987). Notably, the percentage of white police drastically exceeded the percentage of white population in all four data collection cities during the time period of our arrest data: police forces in the Birmingham, Chicago, and Minneapolis metropolitan areas were all more than 95 percent white, while in the Oakland metropolitan area the police

force was more than 75 percent white (U.S. Department of Justice 2012).⁷

A basic sociodemographic questionnaire has been administered in each wave of data collection. A “life events” questionnaire administered in the first survey wave included a question on whether respondents had been arrested in the year prior to the first survey (1985–1986), making CARDIA one of the few medical cohort studies with data on criminal justice contact (Wang et al. 2014). The same question was asked again in the second wave (1987–1988), but with an option to respond that an arrest had occurred but not in the previous year. From these, we generate a binary measure of whether a respondent reported having been arrested on either of the available life events questionnaires. Given the age range of CARDIA respondents, these measures will include arrest events that occurred when respondents were at maximum 32 years old. Consistent with the vast gender gap in arrest and incarceration (Snyder 2011), we limit our sample to men only, because there were too few reported arrests among women for separate analysis to be feasible.

The skin color measure in CARDIA was taken in the fourth wave of data collection (1992–1993) as the percentage of light reflected off the skin, assessed using a Photovolt 577 spectrophotometer at the upper volar arm (the underside of the upper arm). This serves as a measure of “constitutive” skin color—baseline skin color at regions not exposed to light—which stays relatively constant in the same person over time compared with other locations on the body (Pershing et al. 2008). In contrast, “facultative” color—skin color at photo-exposed sites such as the forehead—might be a better indicator of how an individual appears to others at the time of measurement, but would also be far more sensitive to variables such as season of the year or cosmetic tanning. Although a social survey would ideally collect both constitutive and facultative skin color, in a single-time point assessment, the stability of constitutive skin color holds a distinct advantage over the risk of capturing seasonal or other extremes of facultative skin color.

Spectrophotometer readings were taken with three filters (amber, green, and blue), but as correlations among the three sets of readings ranged from .96 to .98, we follow previous literature (e.g., Sweet et al. 2007) in using only the reading taken with the amber filter. Higher reflectance scores denote lighter skin, because lighter colors reflect more light. Of male respondents interviewed in the fourth wave of data collection, 97 percent (1,777) had color data recorded, and

⁷Percentages were calculated using the 1987 Law Enforcement Management and Statistics Survey (U.S. Department of Justice 2012). Minneapolis had the highest percentage of white police (98 percent), with Chicago and Birmingham close behind (95 percent). Oakland had the lowest percentage of white police (76 percent) but also has not had more than 40 percent white residents since the 1970s (Bay Area Census 1980).

respondents without recorded skin color data were excluded from the analysis. Although as in other comparable medical studies (Wang et al. 2014), attrition is higher among black men than any other race-by-sex subgroup (88 percent of initially empaneled white men were interviewed at wave 4, compared with 78 percent of black men), attrition by wave 4 among men who reported arrest records in the first two survey waves does not differ by race.⁸ Among respondents for whom skin color data were recorded, 20 percent of black men and 8 percent of white men reported having been arrested in one of the first two survey waves.

Although our measure of skin color was taken chronologically after collection of our arrest data, base skin color has been found to remain relatively constant with age (Mayes et al. 2010) and should thus be thought of as a generally stable physical quantity, similar to adult height. Of greater concern is that an association between skin color and any outcome of interest associated with socioeconomic status might result partly from sociodemographic differences in tanning, either cosmetic tanning or tanning during outdoor labor, particularly among white respondents. This concern is minimized by the use of the skin color measurements collected in a physical location generally hidden from the sun; indeed, Branigan et al. (2013) reported no significant differences in CARDIA skin reflectance scores by season of measurement for black respondents and negligible differences by season among white respondents. Furthermore, although we have no information regarding frequency of cosmetic tanning in our sample, indoor tanning is known to be far more common among women than men (Heckman, Coups, and Manne 2008).⁹

The distribution of skin reflectance readings are similar between black and white men who do and do not report having been arrested (Figure 1). Consistent with findings from data sources collecting interviewer-coded skin color measures, we observe minimal overlap in reflectance between self-reported white and black men.¹⁰ The overlap is only 3 percent in the full sample, and the overlap among respondents

with a reported arrest history is approximately identical: only 2 percent of white men who report arrest records are below a skin reflectance of 36 percent, while for black men with arrest records, only 2 percent are above a skin reflectance of 36 percent. The variance in skin reflectance among black men is approximately double the variance among white men, both in the full sample and among those with arrest records (Figure 1). Given the bimodal distribution of skin color, all models are run separately by race.

Race is self-reported as white or black. Respondents were screened for non-Hispanic ethnicity, and the 14 individuals who nonetheless self-reported as Hispanic in the initial telephone interview were dropped, along with 93 respondents who were foreign born. Unfortunately, no further detail on ethnicity is available. Respondents reported the race of their mother and father, and to capture any differences in skin color associated with being multiracial, we include a control variable indicating whether either parent was reported as belonging to a racial group other than the respondent's own. A count of biological siblings is coded as per respondent report in the first wave of data collection, as higher sibship size is associated with lower individual receipt of parental resources (Downey 1995; Jaeger 2008). We include fixed effects on birth year, as well as for the four data collection sites from which the respondent pool was drawn.

Because skin color is known to be associated with socioeconomic disadvantage more broadly, one might expect to observe a bivariate relationship between skin color and arrest solely because arrest is also associated with socioeconomic disadvantage. As such, here we control for the occupation and educational attainment of the respondent himself and both of his parents, a unique advantage over studies using police records or other administrative data, which typically lack any measures of socioeconomic background (King and Johnson 2016; Viglione et al. 2011). Our control for respondents' educational attainment is taken from the first wave of data collection, assessed as a scale ranging from 0 to 20 years of education, with 20 or more years coded as 20. Occupation is coded as a socioeconomic index (SEI) score on the basis of the three-digit 1980 census occupational code.

Educational attainment for both parents was also reported by the respondent, again as the number of years completed from 0 to 20. Additional categories were provided for respondents who reported that they did not know their parents' education, and we coded these responses to zero and included a dummy variable indicating replacement. Occupation for both parents was again coded as an SEI score on the basis of the three-digit 1980 census occupational code.¹¹ Parents who were reported as being

⁸Both black and white men who report arrest records are 41 percent more likely to leave the sample by wave 4 than are same-race men with no reported arrest record. Although we cannot directly test the possibility of differential rates of attrition in wave 4 by color within race, models estimating the relationship between skin reflectance at wave 4 and likelihood of attrition in later waves yielded no significant or substantively meaningful association.

⁹To ensure that estimates are not upwardly biased due to individuals who work in outdoor labor being both more tanned and more likely to have been arrested, supplemental models excluded the 48 white men working in outdoor occupations at the survey wave in which skin reflectance was measured. Results were substantively similar and the statistical significance of coefficients was the same.

¹⁰Interviewer-coded skin color measures in both the National Longitudinal Study of Youth 1997 and the General Social Survey 2012 and 2014 yielded a comparable lack of overlap of skin color between white and black respondents (see Supplementary Figure 1).

¹¹SEI scores for occupations were those used by the General Social Survey 1980 (Davis, Smith, and Marsden 2009), with a scale running from 0 through 100. In the few cases in which the occupation of a "responsible adult" was reported in lieu of a parent, we coded the data provided.

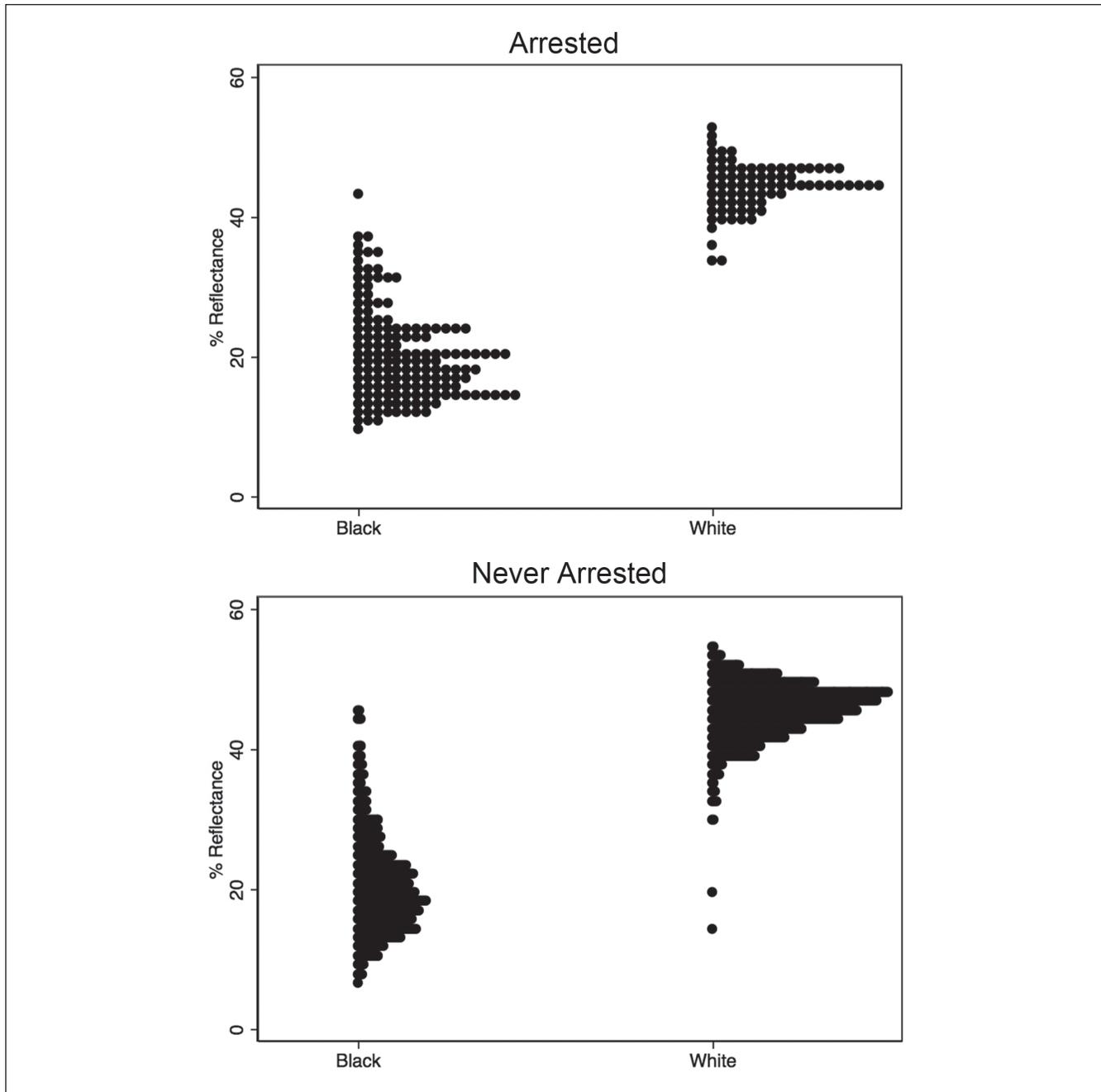


Figure 1. Spectrophotometer Readings for Black and White Respondents, by Arrest Record.

Note: Higher reflectance indicates lighter skin color.

unemployed or with occupations that do not correspond to a prestige score (such as homemakers) were coded to zero, with a dummy variable indicating replacement. Respondents providing no codable information on parental occupation were also coded to zero.

Although an alternative approach to handling missingness on parental occupation and education data would be multiple imputation (Rubin 1987), the decision not to impute was based on concern that values are quite plausibly not missing at random (Allison 2000). The vast

majority of “missing” data, particularly in the case of parental education, resulted from respondents reporting that they “did not know” the educational attainment of their parents. If the mechanisms through which one is unable to report basic socioeconomic status information about a parent, such as parental absenteeism, are associated with lower true values of parental education and SEI scores, then basic assumptions necessary for imputation are violated. Unfortunately, parental absenteeism is not directly queried in the CARDIA data.

Because our first round of arrest data is fielded at the baseline survey wave, we are unable to determine whether any survey measures that would be expected to predict subsequent delinquency, such as drug use or particular personality traits, preceded or resulted from an experience of arrest. As such, these measures are not viable controls in our models predicting arrest likelihood. However, as a key assumption here is that arrests are subjective decisions and do not simply reflect objective differences by skin color in crime perpetration, we separately affirm that skin color is not associated with an array of established correlates of arrest: abuse of alcohol or illicit drugs (marijuana, crack/cocaine, or amphetamines) or psychometric measures of trait hostility and trait anxiety.

Using self-reports of substance use from the first two survey waves, we code an indicator for heavy drinking as consuming 14 or more drinks per week (National Institute on Alcohol Abuse and Alcoholism 2016), and a set of indicators for having used any of the three categories of illicit drugs listed above more than 10 times ever. Supplemental models using indicators for having ever used any of the three categories of illicit drugs yielded no meaningful differences in the coefficients of interest. Although rates of drug use appear high in this sample, they are generally consistent with rates of reported drug use among urban black and white men of the same birth years in the 1985 National Household Survey of Drug Abuse (National Institute on Drug Abuse 2015).¹² Furthermore, among CARDIA respondents coded as drug users, the majority reported no use in the past month.

Trait hostility is assessed as the score on the Cook and Medley Hostility (“Ho”) Scale derived from the Minnesota Multiphasic Personality Inventory, a psychometric test of personality and psychopathology commonly used in criminal justice settings (Contrada and Jussim 1992; Han et al. 1995; Pope, Smith, and Rhodewalt 1990). Trait anxiety, a measure of general negative affect (Balsamo et al. 2013) that has been found to moderate the relationship between violence exposure and subsequent delinquency (Jencks and Burton 2013), is assessed as the score on the Spielberger Trait Anxiety Inventory (Spielberger 2010). Although hostility and anxiety can be temporary states triggered by specific stressors, the “trait” assessment here is intended to capture hostility and anxiety as enduring dispositions.

¹²Among urban respondents in the 1985 National Household Survey on Drug Abuse (NHSDA) who were born between 1954 and 1967, approximately 78 percent of white men reported ever using marijuana and 43 percent reported ever using cocaine ($n = 237$), compared with 82 percent and 48 percent, respectively, in CARDIA. Approximately 60 percent of black men reported ever using marijuana, and 23 percent reported ever using cocaine in the NHSDA, compared with 78 percent and 35 percent, respectively, among black men in CARDIA. Differences between the estimates for black men in the two surveys may reflect the small sample size ($n = 135$) for this subgroup in the NHSDA.

Of the 1,065 white men and 957 black men who were not foreign born and responded to the arrest questions in wave 1 or 2 of CARDIA, 915 white men and 729 black men have skin reflectance readings. Our analytical sample includes the 888 white men and 703 black men who also have codable parental socioeconomic status data, as per the exclusion criteria detailed above. Summary statistics on all variables described are presented in Table 1.

Analytic Strategy

To address the question of whether skin color is associated with likelihood of arrest, we use the logistic regression written

$$\ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_r r_i + \beta_P P_i + \beta_C C_i + \beta_B B_i + \beta_S S_i + \varepsilon_i, \quad (1)$$

in which p_i is the probability that individual i will be arrested. The spectrophotometer reflectance score for each individual i is denoted by r , with higher reflectance scores signifying lighter skin.¹³ P is a vector of family background measures, including education at the first survey wave, occupational prestige (SEI score) at the first survey wave, parental education, parental occupational prestige (SEI score), sibship size, and whether either parent is of a different race than the respondent.¹⁴ C denotes fixed effects on the survey center, and B denotes fixed effects on respondent age. To account for any variation by sun exposure due to seasonal fluctuation and time spent in outdoor labor, S is a set of indicators for the season in which a respondent’s skin reflectance was assessed.¹⁵ We fit the models presented separately by race.¹⁶ For the results presented in Table 2,

¹³Additional specifications to test for nonlinearity in r included using indicators for quartiles of the skin color range within race and including r as a continuous measure with higher order terms. Although significance fluctuated for white respondents in models using the quartile specification because of the relatively small number of individuals who had been arrested or incarcerated within each color quartile, effects were consistent in direction with those presented using the continuous specification of r . For black respondents, models using the quartile specification were consistent with those presented using the continuous specification. Higher order terms on r were nonsignificant and yielded negligible improvement to model fit.

¹⁴Tests revealed no evidence of multicollinearity between any of the control variables and either our outcome or our independent variable of interest in any of the models presented (Allison 2012).

¹⁵Supplemental models using month of skin color measurement in place of season of skin color measurement yielded no meaningful differences.

¹⁶With our full models run as a pooled sample with an indicator for race, a continuous measure of percentage reflectance, and a race-by-reflectance interaction term, both main effects and the interaction are significant at standard levels, affirming that race is independently associated with probability of arrest even net of reflectance.

Table 1. Summary Statistics for Selected Variables by Race: the CARDIA Study, 1985–1988.

Variable	All Men (<i>n</i> = 1,591)	White Men (<i>n</i> = 888)	Black Men (<i>n</i> = 703)
Arrested (%)	13.498 (34.181)	8.343 (27.668)	20.116 (40.116)
Percentage reflectance	34.553 (13.470)	45.418 (3.939)	20.592 (6.915)
SEI (std)	0.054 (0.958)	0.338 (0.960)	-0.303 (0.828)
Education (years)	14.237 (2.455)	14.992 (2.504)	13.273 (2.015)
Age (years)	25.034 (3.515)	25.517 (3.319)	24.419 (3.664)
Father's education (years)	11.491 (5.685)	13.427 (4.931)	9.013 (5.626)
Mother's education (years)	12.349 (4.186)	13.101 (3.982)	11.392 (4.250)
Father's SEI (std)	-0.007 (0.951)	0.286 (0.969)	-0.369 (0.782)
Mother's SEI (std)	-0.001 (0.839)	0.147 (0.774)	-0.181 (0.874)
Number of siblings	3.070 (2.391)	2.683 (1.907)	3.566 (2.819)
Non-same-race parent	0.038 (0.190)	0.028 (0.166)	0.049 (0.217)
Birth year	1959.802 (3.549)	1959.331 (3.373)	1960.399 (3.681)
Substance use (%)			
Alcohol (≥14 drinks/week)	0.154 (0.362)	0.172 (0.378)	0.132 (0.338)
Marijuana (≥10 times ever)	0.745 (0.436)	0.793 (0.405)	0.684 (0.465)
Crack/cocaine (≥10 times ever)	0.176 (0.381)	0.190 (0.393)	0.158 (0.365)
Amphetamines (≥10 times ever)	0.118 (0.323)	0.165 (0.371)	0.058 (0.235)
Trait hostility (Ho scale)	0.066 (0.996)	-0.228 (0.919)	0.449 (0.962)
Trait anxiety (Trait Anxiety Inventory)	-0.089 (0.947)	-0.182 (0.927)	0.035 (0.959)
Missing father's education data	0.139 (0.346)	0.070 (0.256)	0.228 (0.420)
Missing mother's education data	0.067 (0.251)	0.052 (0.222)	0.087 (0.282)
Missing father's SEI	0.094 (0.291)	0.043 (0.203)	0.158 (0.365)
Missing mother's SEI	0.307 (0.461)	0.352 (0.478)	0.248 (0.432)
Black (%)	43.791 (49.629)		

Note: Skin reflectance was measured in 1992–1993 (wave 4). Higher skin reflectance denotes lighter skin. Educational attainment and socioeconomic index (SEI) score were measured at wave 1. Standard errors are in parentheses.

we first present the bivariate association; we then introduce fixed effects on age and survey center; finally, we introduce all remaining covariates.

We then affirm that skin color is not itself associated with known correlates of arrest including substance abuse, trait hostility, and trait anxiety. To assess the relationship between skin color and alcohol or illicit drug use, we run the logistic regression model above (equation 1) substituting indicators of heavy drinking and use each of the three categories of illicit drugs as the outcome. Because of the low frequency of amphetamine use in particular, many birth years have no respondents who were users, so birth year is included in these models as a continuous measure rather than as a battery of indicators; supplemental models using birth-year indicators yielded no meaningful differences in the coefficients of interest. To assess whether skin color is associated with trait hostility and trait anxiety, we use the ordinary least squares regression equation

$$y_i = \alpha + \beta_r r_i + \beta_P P_i + \beta_C C_i + \beta_S S_i + \varepsilon_i, \quad (2)$$

wherein i denotes an individual respondent, y is respondent i 's score on the Ho scale or the Trait Anxiety Inventory, and all remaining variables are defined as in equation 1.

Although we are unable to determine whether the outcomes here preceded or followed an experience of arrest in the CARDIA sample, we do expect that arrest should be significantly associated with established correlates of delinquency; to affirm this assumption, the models presented control for arrest record, although supplemental models excluding the control for arrests or excluding all respondents with an arrest record yielded no substantive differences in the associations between skin color and the outcomes of interest. Coefficients on skin color and arrest record from these models are presented in Table 3.

Replication

Because we hypothesize a pattern of associations between skin color and arrests by race that diverges from findings in studies of skin color and other social outcomes, caution is warranted in affirming that findings are not exclusive to our sample. Although the CARDIA data offer a precise measure of skin color, the sample is relatively small and was drawn from four specific urban centers, and our arrest data were collected three decades ago, a mere 20 years after the passage of the Civil Rights Act. Such limitations pose potential threats to both validity and generalizability, particularly when interpreting a null effect such as that hypothesized among black men.

Although we know of no data source with sufficient variation in the skin color measure among white respondents for a feasible replication of our models for white men, we run a simplified version of our logistic regression model (equation 1) on black men in the National Longitudinal Study of

Table 2. Logistic Regression Models: Respondent Arrested on Percentage Skin Reflectance.

	White Men			Black Men		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Percentage skin reflectance	-0.072** (0.027)	-0.080** (0.028)	-0.072* (0.032)	-0.007 (0.014)	-0.018 (0.015)	0.004 (0.017)
Education			-0.266*** (0.073)			-0.289*** (0.069)
SEI (std)			-0.018 (0.012)			-0.012 (0.010)
Mother's education			0.020 (0.069)			-0.032 (0.052)
Father's education			0.044 (0.061)			-0.004 (0.046)
Mother's SEI (std)			0.317 (0.334)			-0.243 (0.226)
Father's SEI (std)			0.044 (0.020)			0.132 (0.217)
Number of siblings			-0.009 (0.070)			0.010 (0.038)
Non-same-race parent			-0.961 (0.847)			-0.082 (0.472)
Missing father's education			0.280 (0.910)			0.344 (0.582)
Missing mother's education			0.602 (1.009)			-0.973 (0.731)
Missing father's SEI			1.150 (0.782)			0.201 (0.526)
Missing mother's SEI			0.764 (0.747)			-0.745 (0.450)
Constant	0.857 (1.188)	2.206 (1.408)	5.470** (2.063)	-1.240*** (0.299)	-1.279 (0.731)	2.330 (1.586)
<i>n</i>		888			703	

Note: Model 1 is the bivariate association, model 2 introduces fixed effects on age and survey center, and model 3 introduces all remaining covariates. Standard errors in parentheses.

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

Youth 1997 (NLSY97; Horrigan and Walker 2001). The NLSY97 is a nationally representative sample of approximately 9,000 respondents ages 12 to 16 years on December 31, 1996, conducted by the Bureau of Labor Statistics (Horrigan and Walker 2001). Respondents have been surveyed annually since 1997. As in past research using the CARDIA measure of skin reflectance (Branigan et al. 2013), an association has been found between the interviewer-coded measure of skin color in the NLSY97 and both employment (Kreisman and Rangel 2014) and educational outcomes (Hannon, DeFina, and Bruch 2013). The NLSY97 is thus a useful complement to the CARDIA data, offering a less precise measure of skin color but a larger, more recent, and nationally representative sample.

The outcome of interest is an indicator of whether a respondent reported having been arrested by the 2005 wave of data collection, at which point respondents were a minimum of 20

years old, and thus roughly comparable in age with our CARDIA sample. The independent variable of interest is an interviewer-coded skin color rating (Massey and Martin 2003), which we treat as a continuous measure. Values on the skin color scale range from 0 to 10 and are coded to match the direction of the CARDIA reflectance measure, such that lower numbers denote lighter skin. In Table 4, we first present the bivariate association, and then introduce fixed effects on birth year and census region at the time respondents were impaneled (Northeast, Midwest, South, and West).

Results

Results of the logistic regression models estimating the relationship between skin color and likelihood of arrest are presented in Table 2. As can be seen in the bivariate models (model 1), skin reflectance does predict likelihood of arrest

Table 3. Substance Abuse, Trait Hostility, and Trait Anxiety on Percentage Skin Reflectance (Coefficients on Percentage Skin Reflectance and Arrests Only).

	Substance Abuse				Traits	
	Alcohol	Crack/Cocaine	Marijuana	Amphetamines	Hostility	Anxiety
Black men						
Arrest	1.152*** (0.255)	1.082*** (0.259)	0.569* (0.242)	1.191** (0.410)	0.208* (0.090)	0.200* (0.096)
Percentage skin reflectance	-0.009 (0.018)	0.009 (0.017)	-0.022 (0.014)	-0.013 (0.027)	-0.003 (0.005)	0.000 (0.006)
<i>n</i>	698	679	701	678	673	650
White men						
Arrest	1.130*** (0.282)	1.390*** (0.282)	1.376* (0.534)	0.422 (0.299)	0.239* (0.114)	0.247* (0.119)
Percentage skin reflectance	-0.027 (0.023)	-0.014 (0.024)	-0.016 (0.023)	-0.006 (0.026)	-0.003 (0.008)	0.001 (0.008)
<i>n</i>	888	888	888	885	876	869

Note: Coefficients on arrest and skin color in models predicting substance abuse are calculated using the logistic regression defined in equation 1. Coefficients on arrest and skin color in models predicting hostility and trait anxiety are calculated using the ordinary least squares regression defined in equation 2. All models include controls for parental education and occupation with indicators for missingness; number of siblings; and whether either parent is of a different race than the respondent, birth year, and survey center fixed effects. Standard errors in parentheses.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 4. Logistic Regression Models: Ever Arrested on Interviewer-coded Skin Color among Black Men in the National Longitudinal Study of Youth 1997.

	Model 1	Model 2
Skin color	-0.002 (0.035)	-0.005 (0.037)
Constant	0.183 (0.143)	0.470 (0.294)
<i>n</i>	855	

Note: Model 1 is the bivariate association; model 2 introduces fixed effects on birth year and on census region at the time a respondent was impaneled. The interviewer-coded skin color scale is run as a continuous measure ranging from 0 to 10, coded such that as in the Coronary Artery Risk Development in Young Adults data, lower values denote darker skin. Standard errors in parentheses.

among white men ($p < .01$), whereas for black men there is no significant association. The magnitude of the coefficient for white men varies slightly with the addition of fixed effects on birth year and survey center in model 2, and again with the addition of the full battery of controls in model 3. Among black men, the nonsignificant coefficient fluctuates negligibly across models, and by model 3 is still nonsignificant and substantively small.¹⁷

Although the significant coefficient on skin color for white men contrasts with the nonsignificant and near-zero coefficient on skin color for black men, the magnitudes of

these coefficients are not directly comparable across race given the differences by race in the distribution of skin color. As such, in Figure 2 we present the probability of arrest by percentile of skin reflectance, holding all other variables in model 3 in Table 2 at the mean within race.¹⁸ For a black man in the darkest 10 percent of skin reflectance, the probability of arrest is 19.99 percent, while for a black man in the lightest 10 percent of reflectance, the probability of arrest is 20.25 percent. As per Table 2, this negligible change in arrest probability is not statistically significant. For white men, on the other hand, moving from the bottom 10 percent to the top 10 percent of skin reflectance is associated with a significant 5.4-percentage-point decrease in probability of arrest. White men in the bottom 5 percent of skin reflectance, whose skin tone falls in the small region of overlap between the range of reflectance for white and black respondents, have a probability of arrest closer to that of black respondents than to the lightest-skinned white respondents. That said, there are exceedingly few individuals in the region of common support, so estimates at these extremes should be interpreted cautiously.

This reversal of the standard expectations regarding how skin color functions for black and white men is not explained by an association between skin color and a battery of known predictors of arrest, including substance abuse, trait hostility, or trait anxiety (Table 3). As expected, arrest is indeed significantly associated with substance

¹⁷This finding is not driven by city-specific differences in mean skin color among white men who were arrested versus those who were not ($p = .139$).

¹⁸Figure 2 was generated using the `-spost-` package of commands written for Stata 13 (Long and Freese 2014).

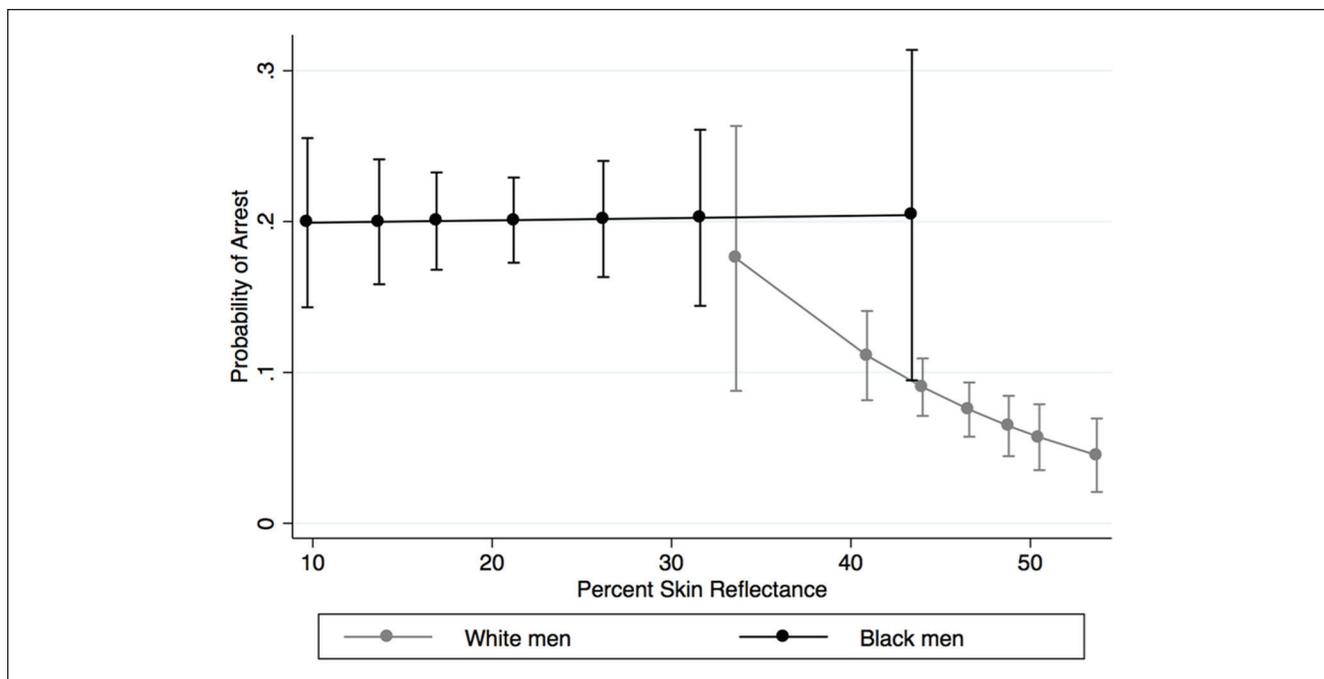


Figure 2. Probability of Arrest by Percentile of Skin Reflectance.

Note: Probabilities are estimated at the 1st, 10th, 25th, 50th, 75th, 90th, and 99th percentiles of skin reflectance by race. Probabilities are based on predictions from model 3 in Table 2, holding all other covariates at the mean.

use, with the sole exception of amphetamines for white men. Arrest is also significantly associated with both trait hostility and trait anxiety. Skin color, on the other hand, is not significantly associated with any of the outcomes tested, and all coefficients on percent reflectance are near zero in magnitude. As noted, although we are unable to determine whether the outcomes of interest preceded or followed an experience of arrest in the CARDIA sample, the null and near-zero associations between skin color and all outcomes in Table 3 were robust to alternative model specifications excluding the control for arrests and excluding all respondents with an arrest record. As an additional robustness check, we ran Table 2 with the delinquency correlates as covariates, despite the problem with time ordering; including these controls yielded no meaningful effect on the coefficients on skin color for either race (Supplementary Table 1).

Our null finding for the relationship between skin color and arrest likelihood among black men is indeed replicated in the NLSY97 sample (Table 4). Although an association has been found between the NLSY97 skin color measure and longer term socioeconomic measures of interest (Hannon et al. 2013; Kreisman and Rangel 2014), as in the CARDIA sample, the difference in arrest probability between black men in the categories corresponding to the darkest 10 percent and lightest 10 percent of skin reflectance is less than 1 percentage point. As discussed above, we know of no data source with sufficient variation in the skin color measure

among white respondents for a feasible replication of our models for white men.

Discussion

The notion that minorities are “people of color” while whites are people without color is pervasive not only as common lay knowledge, but also within the academic research community. Studies of the relevance of skin color for social stratification have generally taken for granted that lightness is a blanket characteristic of whites, who experience no meaningful within-race differentiation by skin tone (Hochschild and Weaver 2007). Among minorities, on the other hand, color is expected to matter continuously, with privilege attached to lightness. Colorism has therefore been implicitly assumed to be a problem only among minorities, as first we stratify skin color categorically by race, and then we sort continuously among nonwhites only.

The findings here do affirm part of that hypothesis: we find evidence that skin color can function categorically for individuals of one race, while functioning continuously for individuals of another race. Beyond that, however, the results presented pose an exception to the common understanding of how colorism operates. The standard construction of colorism would predict that white men’s probability of arrest should remain constant across the spectrum of skin color, while black men’s probability of arrest decreases continuously with lighter skin. We find precisely the opposite: black men’s probability

of arrest remains constant across the spectrum of skin color, while white men's probability of arrest decreases continuously with lighter skin. Rather than white respondents' being categorically advantaged, while minorities are differentially advantaged on the basis of their proximity to aesthetic lightness, we find black respondents to be categorically disadvantaged, while white respondents are disadvantaged differentially on the basis of their proximity to aesthetic darkness.

Where Branigan et al. (2013) demonstrated the potential for skin color to affect social outcomes among white women, to our knowledge, the findings presented here stand as the first in which skin color predicts a social outcome among white men. As all human skin has a color, and skin color is a characteristic of visible phenotype with a long and diverse history of being differentially socially valued (van den Berghe and Frost 1986), this finding should not come entirely as a surprise. Acknowledging skin color as a relevant stratifying quantity among white individuals should not be interpreted as minimizing the legacy of discrimination against minorities, either by race or by color; to the contrary, the findings here suggest that recognizing how social outcomes are stratified by color among individuals of all races, white included, may itself emphasize the pervasiveness of blanket discrimination by race against minorities. Even the darkest-skinned white respondents in our sample remain less likely to be arrested than the lightest-skinned black respondents.

Despite increasing interest in the social consequences of skin color, sociological research on colorism still rarely engages related social psychological theory on stereotyping and cross-race perception of physical appearance. Rather than attributing our observed pattern of results to a "preference for whiteness" (or, conversely, a dispreference for darkness), we draw from research on stereotyping to propose an alternative explanation (Goldsmith et al. 2007). The "out-group homogeneity effect" describes the tendency to perceive out-group members as "all looking alike," while in-group members are perceived as more physically variable (Linville et al. 1989; Ostrom and Sedikides 1992; Quattrone and Jones 1980). To that end, this study builds on Hill's (2002a) test of the out-group homogeneity effect in a population of interviewers for a large social survey, in which he found that same-race interviewers did indeed report less variation in the skin color of other-race respondents than in same-race respondents. As Hill (2002a) noted, this finding was particularly concerning because it suggests a limited ability to perceive differentiating physical characteristics in other-race individuals. Rather, "perception of other-race individuals is filtered through a powerful social prism, which provides fertile ground for the perpetuation of ethnocentric stereotypes and race-related conflict" (Hill 2002a:106).

As in the present study, Hill (2002a) included only black and white individuals in his analysis, despite lived experience in the U.S. entailing interactions with a far more continuous array of skin colors. As the ranges of interviewer-coded color observed among Asians, Hispanics, and American

Indians indeed fall between the color ranges occupied by black and white individuals (Supplementary Figure 1), skin color will not be an accurate identifier of persons who definitively are of one's same race. The continuous association between skin color and arrest among white men may reflect this ambiguity, wherein the lightest respondents are most unambiguously visually identifiable as white, whereas darker white men occupy a more racially ambiguous position on the color continuum. Findings of higher incarceration probability and harsher sentencing (Blair, Judd, and Chapleau 2004; King and Johnson 2016) among whites with more "Afrocentric" facial features lends support to this interpretation. Although lay knowledge regarding white colorlessness does tend to include quiet exception clauses for "not-all-the-way white" ethnic whites (Raffo 1998), we unfortunately lack data on ethnicity in our sample other than that respondents are non-Hispanic, leaving the relationship between ethnicity and arrests as an area for future research.¹⁹

In the case of black and white Americans, their positioning at the relative extremes of the continuum of skin color means that color should serve as a reasonably accurate identifier of individuals who are clearly *not* of their same race. As noted, at the time of our data collection, all CARDIA survey centers were in cities with meaningfully higher percentages of white police than residents: in all cities except for Oakland, the police force was more than 95 percent white (U.S. Department of Justice 2012). Because black arrestees would have been almost exclusively arrested by white officers, our results follow what would be predicted by the out-group homogeneity effect if arresting officers simply perceive more variation in skin color among same-race versus other-race potential arrestees. That white men in the lowest percentiles of skin reflectance have a probability of arrest more similar to the lightest-skinned black men than to the lightest-skinned white men is also consistent with this explanation, as that ambiguous range of skin color would afford the highest probability of white individuals being misclassified as out-group. This does not suggest that darkest-skinned white men would be necessarily perceived as black in particular (as opposed to Hispanic, for example), but simply that when skin tone no longer allows immediate identification as same-race by a white arresting officer, a white man's arrest probability approaches that of non-same-race individuals. The overlap between the distributions of reflectance for black and

¹⁹To the extent that "not-all-the-way white" ethnicity correlates with religion, we are able to determine that religious denomination in CARDIA is not associated with meaningful differences in likelihood of arrest. Separately by race, Protestants, Catholics, and nonaffiliates (atheist, agnostic, or "none") are represented at approximately equal proportions among men who had and had not been arrested or incarcerated; our sample included too few religious minorities for separate comparison (of the 5 self-reported Muslims, none reported criminal records; of the 53 self-reported Jews, 3 reported criminal records).

white respondents is minimal, however, and estimates at these extremes should be interpreted cautiously.

Future research on this subject would ideally include Americans of racial and ethnic designations other than non-Hispanic black and white, as the implications of this work for those populations remains at this point entirely speculative. It is particularly unclear how this process might work among ethnic groups such as Asians or Native Americans, for example, whose range of skin color meaningfully overlaps with that of white Americans (Supplementary Figure 1). Studies considering how skin color matters relative to visible racial cues other than skin color, such as King and Johnson's (2016) work on "Afrocentric" features, could be particularly useful here, asking whether color functions as a primary screening mechanism relative to other physical characteristics in quick low-information situations such as an arrest. Having a skin color that allows one to be perceived as potentially same-race relative to an arresting officer could convey advantage regardless of any other visible physical signs of race, although perhaps more plausibly, color might be an interactive factor with other racialized physical attributes such as hair color and texture or facial features (Blair et al. 2004).

An alternative explanation for the results presented would be that our controls for socioeconomic background fail to truly capture the relationship between color and socioeconomic disadvantage more broadly, which is in turn associated with higher likelihood of encountering the criminal justice system. Although we cannot rule out this causal pathway, our models account for a far richer battery of background measures than are available in court or police records, the most common source of data for population research on color and criminal justice outcomes (e.g., King and Johnson 2016; Viglione et al. 2011). Furthermore, if color is merely a correlate of socioeconomic disadvantage, that darker skin tone is associated with lower educational attainment and occupational prestige among black men and not among white men in the CARDIA sample (Branigan et al. 2013) would suggest that color should be also associated with arrest among blacks and not whites: the opposite of what we find here. As a supplemental test of whether darker skin color is associated with life stress more broadly for white men, we ran our logistic regression model (equation 1) on our sample of white men using as our outcome a range of additional "life events" from the same survey questionnaire from which our arrest data were drawn. Outcomes included whether respondents reported troubles at work, having moved to a worse neighborhood, or going on or off of welfare. Mirroring findings from our models predicting drug use, trait hostility, and trait anxiety, skin reflectance among white men was not associated with any of the additional outcomes considered.

Omitted variable bias nonetheless remains the most consistent threat to the validity of findings in this and any of the many studies interpreting a remaining association between skin color and social outcomes net of controls as indicative of discrimination (e.g., Goldsmith et al. 2006; Hughes and

Hertel 1990; Keith and Herring 1991), a strategy that Fryer (2010) dubbed "a competition of 'name that residual'" (p. 2). Although future studies on this topic should endeavor to better control for background measures collected prior to the time of arrest, broadening the range of background controls will never nullify the possibility that skin color is simply functioning as a proxy for key unobserved variables. Although we know of no data source that currently contains both skin reflectance and sibling data, sibling fixed effects would be one useful strategy for better addressing unobserved family-level heterogeneity in future studies relating skin color to social outcomes such as arrest. In addition, we again emphasize the utility of modeling the association between skin color and social outcomes as a comparative relationship between social groups, as although group differences do not invalidate concerns over omitted variable bias, they may complicate the argument that the results observed are likely to result strictly from unobservables. For example, if the results of the present study are entirely explicable via key omitted variables, it would suggest that skin color serves as a proxy for unobserved correlates of arrests only among white men and not among black men, an interesting finding in of itself.

Although we present our results as a complication to the standard conceptualization of colorism, generalizing these findings to the current American population should be done with caution. The CARDIA sample is drawn from four specific urban centers, two midwestern, one southern, and one western. Furthermore, our arrest data are from 30 years ago, a mere two decades after the passage of the Civil Rights Act. Shifts in the American racial landscape since the mid-1980s may have altered the probability of arrest from the pattern of results observed here, or altered the association between color and any of the other socioeconomic outcomes tested.

That said, our results from the NLSY97 replication suggest that at least the null association between skin color and arrest probability among young black men may well persist to the present. Although our analysis remains correlational and not causal, the portrait of arrest probability painted here aligns with contemporary concern that excessive police brutality against black men and boys may be due partially to the perception of black individuals as homogenous across critical demographic differences such as physical size or age (Patton 2014). The police commentary following the fatal shooting of 12-year-old Tamir Rice by a white officer in November 2014 poses a relevant example: as per the president of the Cleveland Police Patrolman's Association, the officer simply "had no clue he was a 12-year-old" (Bever 2014). Although police departments in the CARDIA collection cities are more integrated now than they were in the 1980s, the percentage of white officers remains disproportionate to the racial composition of the resident populations (Ashkenas and Park 2015).

There are several limitations to the available data beyond those previously discussed. First, knowing the month of the

year in which respondents were arrested would be useful to affirm that an association between skin color and arrests is not capturing differences in arrest probability by season.²⁰

Second, although the police forces in the cities from which our population was drawn were almost exclusively white at the time of data collection, information on both racial identification and skin color of the arresting officers would be useful to directly test our proposed explanation for the pattern of results observed. This is particularly true given that even with only a handful of minority officers, those officers may well be disproportionately assigned to minority neighborhoods; furthermore, even with an overwhelmingly white police force, there may be differences in responsiveness to skin color in others as a function of the skin color of the officer making the arrest.

Finally, collection of our arrest data in the same wave as our measures of drug use and personality traits poses a problem for determining whether these measures were a cause or a consequence of arrest. That said, even were we to know the specific charges for each arrest, an argument of discrimination in this case is predicated in part on an assumption that arrests may not be purely objective measures of crimes committed, but are also partly subjective decisions that depend on characteristics of the arresting officer interacting with characteristics of the individual being arrested. Although our results align with this interpretation, we remain unable to definitively separate such objective versus subjective components of officer decision making.

Our findings emphasize the need to consider color as a separate quantity from race in models of social outcomes. This point has practical implications, as the tools for recording skin color data in social surveys have been developed with the implicit understanding that collecting color data among white Americans is not a priority. As such, mechanical measurement like that used in this analysis remains the sole option for assessing the relationship between skin color and social outcomes without first imposing categorical assumptions about the relevance of skin color by race. The increasing number of social data sources collecting interviewer-coded skin color ratings—including the General Social Survey, the Fragile Families and Child Wellbeing Study, and the NLSY97, as used here—suggests that color data are of sufficient interest to the research population to merit serious consideration of how they should best be quantified. That the coding tool constructed for the New Immigrant Survey (Massey and Martin 2003) was subsequently used in all three surveys noted above marks a meaningful improvement over past efforts to collect interviewer-coded skin color, as it provides a basis for cross-cohort and cross-survey analysis. However, as colorimeter readings

can now be taken using a smartphone (Chang 2012), eliminating the cost and burden of lab-standard equipment, interviewer coding may no longer have the benefit of efficiency to outweigh the comparative loss of precision.

Findings additionally speak to the potential utility of any policy intervention that alters in-group versus out-group identities. The designation “in-group” is not synonymous with “same race,” but rather extends to any social identity with which a gatekeeper in question self-affiliates. For example, in many racially diverse communities where the police remain disproportionately white, a large percentage of officers live outside the cities in which they are employed (Silver 2014). This holds at present in the cities from which the CARDIA sample was drawn: Minneapolis has only 5 percent of white officers living within the city limits, while in Oakland that figure is a mere 3 percent (Silver 2014). Working in disproportionately white departments and living in neighboring towns, officers may be most frequently encountering minorities in the context of crimes committed, reifying the classification of “nonwhite” as inherently out-group. Requiring police officers to live in the communities where they are employed could well have the effect of defining a new in-group of “my neighbors,” providing officers a basis on which to affiliate with non-same-race community members instead of viewing them as strictly “other.”

Finally, although it is increasingly accepted that appearance matters for social outcomes, the results presented emphasize the importance of considering how measurable quantities of the visible body function similarly to and differently from more traditional quantities of interest in population research. Whereas education, income, and other measures of socioeconomic status may be visibly displayed through dress or carriage, aspects of the visible body such as skin color or body fatness have the unique property that they may be rendered socially relevant even in the absence of individual action, when one’s body is observed by others. As such, quantitative research engaging the visible physical body should consider aspects of social interaction that may affect how bodies are perceived by relevant gatekeepers, such as context and timing, as well as theories of how the visible body may be understood differently across social categories such as gender and race.

Funding

The CARDIA study is supported by contracts HHSN268201300025C, HHSN268201300026C, HHSN268201300027C, HHSN268201300028C, HHSN268201300029C, and HHSN268200900041C from the NHLBI, the Intramural Research Program of the National Institute on Aging, and an intra-agency agreement between the National Institute on Aging and the NHLBI (AG0005).

References

- Allison, Paul. 2000. “Multiple Imputation for Missing Data: A Cautionary Tale.” *Sociological Methods and Research* 28(8):301–309.

²⁰Although seasonality in both skin color and arrests could be a concern among respondents drawn from Chicago and Minneapolis, it should be of less concern in Oakland and Birmingham. The direction of our association persists when models are run on white men from each city separately, albeit with some fluctuation in statistical significance as per the reduced sample size.

- Allison, Paul. 2012. "When Can You Safely Ignore Multicollinearity?" *Statistical Horizons*, September 10. Retrieved August 8, 2017 (<http://statisticalhorizons.com/multicollinearity>).
- Ashkenas, Jeremy, and Haeyoun Park. 2015. "The Race Gap in America's Police Departments." *The New York Times*, April 8.
- Balsamo, Michela, Roberta Romanelli, Marco Innamorati, Gabriele Ciccarese, Leonardo Carlucci, and Aristide Saggino. 2013. "The State-Trait Anxiety Inventory: Shadows and Lights on Its Construct Validity." *Journal of Psychopathology and Behavioral Assessment* 35(4):475–86.
- Bay Area Census. 1980. "City of Oakland Census." Retrieved August 8, 2017 (<http://www.bayareacensus.ca.gov/cities/Oakland70.htm>).
- Bever, Lindsey. 2014. "Cleveland Police Kill 12-year-old Boy Wielding BB Gun That Looked Like a Semiautomatic Pistol." *The Washington Post*, November 24.
- Blair, Irene V., Charles M. Judd, and Kristine M. Chapleau. 2004. "The Influence of Afrocentric Facial Features in Criminal Sentencing." *Psychological Science* 15(10):674–79.
- Branigan, Amelia R., Jeremy Freese, Assaf Patir, Thomas W. McDade, Kiang Liu, and Catarina I. Kiefe. 2013. "Skin Color, Sex, and Educational Attainment in the Post-Civil Rights Era." *Social Science Research* 42(6):1659–74.
- Burch, Traci. 2015. "Skin Color and the Criminal Justice System: Beyond Black-white Disparities in Sentencing." *Journal of Empirical Legal Studies* 12(3):395–420.
- Caruso, Eugene M., Nicole L. Mead, and Emily Balcetis. 2009. "Political Partisanship Influences Perception of Biracial Candidates' Skin Tone." *Proceedings of the National Academy of Sciences* 106(48):20168–73.
- Chang, Byoung-Yong. 2012. "Smartphone-based Chemistry Instrumentation: Digitization of Colorimetric Measurements." *Bulletin of the Korean Chemical Society* 33(2):549–52.
- Conrada, Richard J., and Lee Jussim. 1992. "What Does the Cook-Medley Hostility Scale Measure? In Search of an Adequate Measurement Model." *Journal of Applied Social Psychology* 22(8):615–27.
- Darity, William A., Jr., Jason Dietrich, and Darrick Hamilton. 2005. "Bleach in the Rainbow: Latin Ethnicity and Preference for Whiteness." *Transforming Anthropology* 13(2):103–9.
- Davis, James A., Tom W. Smith, and Peter V. Marsden. 2009. "General Social Surveys, 1972–2006" [Cumulative File]. Inter-University Consortium for Political and Social Research. Retrieved (<http://dx.doi.org/10.3886/ICPSR04697>).
- Downey, Douglas B. 1995. "When Bigger Is Not Better: Family Size, Parental Resources, and Children's Educational Performance." *American Sociological Review* 60(5):746–61.
- Dumont, Dora M., Brad Brockmann, Samuel Dickman, Nicole Alexander, and Josiah D. Rich. 2012. "Public Health and the Epidemic of Incarceration." *Annual Review of Public Health* 33(April):325–39.
- Federal Bureau of Investigation. 2014. "Crime in the United States 2013." Washington, DC: U.S. Department of Justice.
- Feingold, Gustave A. 1914. "The Influence of Environment on Identification of Persons and Things." *Journal of the American Institute of Criminal Law and Criminology* 5(1):39–51.
- Franklin, John Hope. 2000. *From Slavery to Freedom: A History of African Americans*. 8th ed. New York: Knopf.
- Fryer, Roland G. 2010. "Racial Inequality in the 21st Century: The Declining Significance of Discrimination." NBER Working Paper No. 16256. Retrieved August 8, 2017 (<http://www.nber.org/papers/w16256.pdf>).
- Goering, John M. 1972. "Changing Perceptions and Evaluations of Physical Characteristics among Blacks: 1950–1970." *Phylon* 33(3):231–41.
- Goldsmith, Arthur H., Darrick Hamilton, and William Darity, Jr. 2006. "Shades of Discrimination: Skin Tone and Wages." *American Economic Review* 96(2):242–45.
- Goldsmith, Arthur H., Darrick Hamilton, and William Darity, Jr. 2007. "From Dark to Light: Skin Color and Wages among African-Americans." *Journal of Human Resources* 42(4):701–38.
- Gullickson, Aaron. 2005. "The Significance of Color Declines: A Re-analysis of Skin Tone Differentials in Post-Civil Rights America." *Social Forces* 84(1):157–80.
- Han, Kyunghye, Nathan C. Weed, Robert F. Calhoun, and James N. Butcher. 1995. "Psychometric Characteristics of the MMPI-2 Cook-Medley Hostility Scale." *Journal of Personality Assessment* 65(3):567–85.
- Hannon, Lance, Robert DeFina, and Sarah Bruch. 2013. "The Relationship between Skin Tone and School Suspension for African Americans." *Race and Social Problems* 5(4): 281–95.
- Heckman, Carolyn J., Elliot J. Coups, and Sharon L. Manne. 2008. "Prevalence and Correlates of Indoor Tanning among US Adults." *Journal of the American Academy of Dermatology* 58(5):769–80.
- Hersch, Joni. 2006. "Skin-tone Effects among African Americans: Perceptions and Reality." *American Economic Review* 96(2):251–55.
- Hersch, Joni. 2008. "Profiling the New Immigrant Worker: The Effects of Skin Color and Height." *Journal of Labor Economics* 26(2):345–86.
- Hill, Mark E. 2000. "Color Differences in the Socioeconomic Status of African American Men: Results of a Longitudinal Study." *Social Forces* 78(4):1437–60.
- Hill, Mark E. 2002a. "Race of the Interviewer and Perception of Skin Color: Evidence from the Multi-city Study of Urban Inequality." *American Sociological Review* 67(1):99–108.
- Hill, Mark E. 2002b. "Skin Color and the Perception of Attractiveness among African Americans: Does Gender Make a Difference?" *Social Psychology Quarterly* 65 (1): 77–91.
- Hochschild, Jennifer L., and Vesla Weaver. 2007. "The Skin Color Paradox and the American Racial Order." *Social Forces* 86(2):643–70.
- Horrigan, Michael, and James Walker. 2001. "The NLSY97: An Introduction." *Monthly Labor Review*, August 3.
- Hughes, Glenn H., Gary Cutter, Richard Donahue, Gary D. Friedman, Steve Hulley, Enid Hunkeler, David R. Jacobs, et al. 1987. "Recruitment in the Coronary Artery Disease Risk Development in Young Adults (CARDIA) Study." *Controlled Clinical Trials* 8(4 Suppl. 1): 68–73.
- Hughes, Michael, and Bradley R. Hertel. 1990. "The Significance of Color Remains: A Study of Life Chances, Mate Selection, and Ethnic Consciousness among Black Americans." *Social Forces* 68(4):1105–20.
- Hunter, Margaret L. 2002. "'If You're Light You're Alright': Light Skin Color as Social Capital for Women of Color." *Gender & Society* 16(2):175–93.

- Jablonski, Nina G., and George Chaplin. 2000. "The Evolution of Human Skin Coloration." *Journal of Human Evolution* 39(1): 57–106.
- Jaeger, Mads Meier. 2008. "Do Large Sibships Really Lead to Lower Educational Attainment?" *Acta Sociologica* 51(3): 217–35.
- Jencks, Jennifer W., and David L. Burton. 2013. "The Role of Trait Anxiety in Reducing the Relationship between Childhood Exposure to Violence/Victimization and Subsequent Violent Behavior among Male Delinquent Youth." *International Journal of Offender Therapy and Comparative Criminology* 57(8):985–95.
- Keith, Verna M., and Cedric Herring. 1991. "Skin Tone and Stratification in the Black Community." *American Journal of Sociology* 97(3):760–78.
- King, Ryan, and Brian Johnson. 2016. "A Punishing Look: Skin Tone and Afrocentric Features in the Halls of Justice." *American Journal of Sociology* 122(1):90–124.
- Kreisman, Daniel, and Marcos A. Rangel. 2014. "On the Blurring of the Color Line: Wages and Employment for Black Males of Different Skin Tones." *Review of Economics and Statistics* 97(1):1–13.
- Linville, Patricia W., Gregory W. Fischer, and Peter Salovey. 1989. "Perceived Distributions of the Characteristics of In-group and Out-group Members: Empirical Evidence and a Computer Simulation." *Journal of Personality and Social Psychology* 57(2):165–88.
- Long, J. Scott, and Jeremy Freese. 2014. *Regression Models for Categorical Dependent Variables Using Stata*. 3rd ed. College Station, TX: Stata Press.
- Malpass, Roy S., and Jerome Kravitz. 1969. "Recognition for Faces of Own and Other Race." *Journal of Personality and Social Psychology* 13(4):330–34.
- Massey, Douglas S., and Jennifer A. Martin. 2003. "The NIS Skin Color Scale."
- Mayes, A. E., P. G. Murray, D. A. Gunn, C. C. Tomlin, S. D. Catt, Y. B. Wen, L. P. Zhou, H. Q. Wang, M. Catt, and S. P. Granger. 2010. "Ageing Appearance in China: Biophysical Profile of Facial Skin and Its Relationship to Perceived Age." *Journal of the European Academy of Dermatology & Venereology* 24(3):341–48.
- Myrdal, Gunnar. 1944. *An American Dilemma; the Negro Problem and Modern Democracy*. New York: Harper.
- National Institute on Alcohol Abuse and Alcoholism. 2016. "Drinking Levels Defined." Retrieved August 8, 2017 (<https://www.niaaa.nih.gov/alcohol-health/overview-alcohol-consumption/moderate-binge-drinking>).
- National Institute on Drug Abuse. 2015. "National Household Survey on Drug Abuse, 1985." Inter-University Consortium for Political and Social Research. Retrieved (<http://doi.org/10.3886/ICPSR06844.v4>).
- Ostrom, Thomas M., and Constantine Sedikides. 1992. "Out-group Homogeneity Effects in Natural and Minimal Groups." *Psychological Bulletin* 112(3):536–52.
- Patton, Stacey. 2014. "In America, Black Children Don't Get to Be Children." *The Washington Post*, November 26. Retrieved (http://www.washingtonpost.com/opinions/in-america-black-children-dont-get-to-be-children/2014/11/26/a9e24756-74ee-11e4-a755-e32227229e7b_story.html).
- Pearce-Doughlin, Shavonn, Arthur Goldsmith, and Darrick Hamilton. 2013. "Colorism." In *International Encyclopedia of Race and Racism*, edited by Patrick L. Mason II. New York: Macmillan Reference.
- Pershing, Lynn K., Vidyant P. Tirumala, Joel L. Nelson, Judy L. Corlett, Anne G. Lin, Laurence J. Meyer, and Sancy A. Leachman. 2008. "Reflectance Spectrophotometer: The Dermatologists' Sphygmomanometer for Skin Phototyping?" *Journal of Investigative Dermatology* 128(7):1633–40.
- Pope, Mary Katherine, Timothy W. Smith, and Frederick Rhodewalt. 1990. "Cognitive, Behavioral, and Affective Correlates of the Cook and Medley Hostility Scale." *Journal of Personality Assessment* 54(3/4):501–14.
- Quattrone, George A., and Edward E. Jones. 1980. "The Perception of Variability within In-groups and Out-groups: Implications for the Law of Small Numbers." *Journal of Personality and Social Psychology* 38(1):141–52.
- Raffo, Susan. 1998. "Not Always as Easy as Breathing." In *Duraggia: Writing by Women of Italian Descent*, edited by Domenica DiLeo, Gabriella Micallef, and Patrizia Tavourmina. Toronto, Canada: Women's Press.
- Reaves, Brian A. 2010. "Local Police Departments, 2007." NCJ 231174. Retrieved August 8, 2017 (<http://www.bjs.gov/content/pub/pdf/lpd07.pdf>).
- Rubin, D. B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley.
- Silver, Nate. 2014. "Most Police Don't Live in the Cities They Serve." FiveThirtyEight (DataLab). August 20. Retrieved (<http://fivethirtyeight.com/datalab/most-police-dont-live-in-the-cities-they-serve/>).
- Smith, Douglas A. 1986. "The Neighborhood Context of Police Behavior." Pp. 313–41 in *Communities and Crime*, edited by A. J. Reiss and M. Tonry. Chicago: University of Chicago Press.
- Snyder, Howard N. 2011. "Arrest in the United States, 1980–2009." NCJ 234319. Patterns and Trends. Washington, DC: U.S. Department of Justice.
- Spielberger, Charles D. 2010. *State-Trait Anxiety Inventory*. Hoboken, NJ: Wiley Online Library.
- Sweet, Elizabeth, Thomas W. McDade, Catarina I. Kiefe, and Kiang Liu. 2007. "Relationships between Skin Color, Income, and Blood Pressure among African Americans in the CARDIA Study." *American Journal of Public Health* 97(12):2253–59.
- Thompson, Maxine S., and Verna M. Keith. 2001. "The Blacker the Berry: Gender, Skin Tone, Self-esteem, and Self-efficacy." *Gender & Society* 15(3):336–57.
- Udry, J. Richard, Karl E. Bauman, and Charles Chase. 1971. "Skin Color, Status, and Mate Selection." *American Journal of Sociology* 76(4):722–33.
- U.S. Department of Justice. 2012. "Law Enforcement Management and Administrative Statistics (LEMAS), 1987." Inter-University Consortium for Political and Social Research [distributor]. Retrieved (<http://doi.org/10.3886/ICPSR09222.v3>).
- van den Berghe, Pierre L., and Peter Frost. 1986. "Skin Color Preference, Sexual Dimorphism and Sexual Selection: A Case of Gene Culture Co-evolution?" *Ethnic and Racial Studies* 9(1):87–113.
- Viglione, Jill, Lance Hannon, and Robert DeFina. 2011. "The Impact of Light Skin on Prison Time for Black Female Offenders." *Social Science Journal* 48(1):250–58.
- Wang, Emily A., Jenerius A. Aminawung, Christopher Wildeman, Joseph S. Ross, and Harlan M. Krumholz. 2014. "High

Incarceration Rates among Black Men Enrolled in Clinical Studies May Compromise Ability to Identify Disparities.” *Health Affairs* 33(5):848–55.

White, Karletta M. 2015. “The Salience of Skin Tone: Effects on the Exercise of Police Enforcement Authority.” *Ethnic and Racial Studies* 38(6):993–1010.

Young, Steven G., Kurt Hugenberg, Michael J. Bernstein, and Donald F. Sacco. 2012. “Perception and Motivation in Face Recognition: A Critical Review of Theories of the Cross-race Effect.” *Personality and Social Psychology Review* 16(2):116–42.

Author Biographies

Amelia R. Branigan is a visiting assistant professor in the Department of Sociology at the University of Illinois at Chicago. She is a social demographer with central interests in inequality, health, and the criminal justice system.

Christopher Wildeman is an associate professor of policy analysis and management and sociology (by courtesy) at Cornell University, as well as a senior researcher at the Rockwool Foundation Research Unit. His research focuses on the prevalence, causes, and consequences of criminal justice and child protective services contact for individuals, families, and communities.

Jeremy Freese is a professor of sociology at Stanford University. He conducts research on various topics related to the relationship between social advantage and embodied advantage. This includes work differences in physical health, cognitive functioning, health behaviors, and the role of differential use of knowledge and innovations toward producing differences. He is co-leader of the Health Disparities Working Group for the Stanford Center for Population Health Sciences, and he is co-principal investigator of the General Social Survey and the Time-sharing Experiments in the Social Sciences.

Catarina I. Kiefe is, by training, an internist (MD) and a mathematician (PhD). She is currently a professor and the inaugural chair of the Department of Quantitative Health Sciences at the University of Massachusetts Medical School (UMMS) and holds the inaugural Cutler Chair for Biomedical Research. She has founded and lead multiple research centers at UMMS and the University of Alabama at Birmingham. Dr. Kiefe’s major research interests are in the development and application of innovative methods to measuring quality in health care, and in cardiovascular epidemiology and outcomes and effectiveness research. She has been the principal investigator on federally funded research grants continuously for more than two decades, has more than 250 peer-reviewed publications, has served on or led multiple national and international scientific advisory boards; and is co-editor-in-chief of *Medical Care*, a premier journal in health services research.