Are within-racial group inequalities by skin color really greater than inequalities between racial groups in the United States?

This article examines the relationship between skin color and educational and labor market outcomes within White, Black, and Hispanic populations in the U.S. By analyzing NLSY97 data, it challenges claims that intra-racial inequalities based on skin color match or surpass inequalities between ethnoracial groups. The findings indicate that, although a darker skin tone correlates with less favorable outcomes across all ethnoracial groups, disparities along the color continuum within the Black population are less pronounced than those between Blacks and Whites as a whole. For Hispanics, the significance of between- and within-race inequality varies depending on the outcome. These insights remain consistent both in descriptive analysis and after adjusting for socioeconomic origins.

Racial stratification in the United States has historically been construed as a system of categorical inequalities among groups whose boundaries are sharply demarcated by common ancestry (e.g., Whites, Blacks, Asians, and others). However, race scholars have long observed that such categorical disparities coexist with a gradational hierarchy rooted in phenotypic traits, with skin color being the most conspicuous. This phenomenon, known as “colorism,” operates through various discriminatory practices that favor individuals with lighter skin and features deemed “European,” over their darker-skinned counterparts.

Consistent with this theory, empirical studies in the U.S. have repeatedly found that, above and beyond well-documented inequalities between races, darker skin is associated with a myriad of adverse social outcomes within racial groups, particularly among Blacks and Hispanics. These include lower educational attainment (Branigan et al. 2013; Monk 2014), reduced earnings, income, and wealth (Hersch 2006; Goldsmith et al. 2006; Monk 2014; Adames 2023), a higher incidence of hypertension (Laidley et al. 2019), an increased likelihood of perceived discrimination, and deteriorated mental health (Monk 2015), among other factors.

Building upon this evidence, some scholars argue that a gradational approach to understanding racial inequalities not only complements but potentially surpasses the traditional categorical perspective in relevance (Monk 2022). For instance, (Monk 2021) (p. 86) contends that “... across nearly every outcome that social scientists study with respect to ethnoracial inequality between Blacks and Whites, African Americans are also significantly stratified by skin tone, so much so that intra-racial inequalities along the color line, across a whole host of outcomes, often rival or exceed ethnoracial inequalities between Blacks and Whites as a whole.”

Using data from the National Longitudinal Survey of Youth 1997 (NLSY97) from 2008 to 2013, this article challenges the aforementioned claim. It does so by examining the association between skin color – measured when respondents were aged 24 to 30 – and four educational and labor market outcomes across White, Black, and Hispanic populations.
in the U.S.: years of schooling, attainment of a college degree, personal earnings, and family income – measured when respondents were aged 26 to 33 years old. The analysis unfolds in two stages: Firstly, using Bayesian regression models, I estimate the expected value of each outcome by skin color within each ethnoracial group. Secondly, recognizing that associations with skin color partly originate from inherited disadvantages due to race and color-based discrimination in previous generations (Flores and Telles 2012), I compute the expected value of each outcome by skin color and ethnoracial group, while controlling for socioeconomic origins (parental education, parental income, and family structure) and basic demographics (age and gender) (Details on measures and modeling strategy in Supplementary Materials). Figure depicts the results of this analysis.
Figure 1: The graphs show the expected value of each outcome, conditional on the interaction of skin color and ethnoracial group. Each row in the panel corresponds to different outcomes, with their respective scales shown on the y-axis. Income variables are graphed on a log10 scale. The x-axis represents the interviewers’ ratings of respondents’ skin tone on a visual palette, ranging from 1 (the lightest) to 11 (the darkest). The first column of plots presents the 'unadjusted' results for each outcome, while the second column displays the 'adjusted' results. These adjusted results represent the expected outcome for individuals of different skin colors and ethnoracial groups if they had comparable socioeconomic origins (i.e., with control variables fixed at their global means).

Solid lines and dots display the outcome gradient across the color spectrum for each ethnoracial group, where the size of each dot is proportional to the share of individuals in each color category within an ethnoracial group. Confidence interval bands correspond to Bayesian 95% credible intervals for the predicted values. Data source: National Lo

The results unveil several important insights. Firstly, the unadjusted analysis shows that darker skin color generally correlates with lower levels of education, a reduced likelihood of attaining a college degree, and lower personal and family income across all ethnoracial groups. However, the extent of this correlation varies significantly among groups. Moreover, variation in the location and amplitude of the color spectrum for each group implies that the “color continuum” carries distinct implications for different ethnoracial groups. For example, while the educational and economic outcomes for Whites display a steep decline as skin color shifts from very light to moderately light, the outcomes for Blacks demonstrate a more gradual rate of decline. However, the color spectrum for Blacks is considerably broader – spanning from moderately light to dark – potentially leading to larger gaps between the extremes.

Secondly, and most importantly, despite the gradient of skin tones within races, these results indicate that – contrary to previous claims – disparities within the Black population along the color spectrum are smaller than the disparities between Blacks and Whites as a whole. In the case of Hispanics, the relative significance of disparities between and within racial groups varies depending on the specific outcome being examined. Furthermore, the adjusted results suggest that if individuals of different skin colors and ethnoracial groups had similar socioeconomic origins, the average income disparities between the White and Black populations would be significantly larger than the inequality by skin color within Blacks. Only for years of schooling does this analysis suggest that the magnitudes of within- and between-race inequalities are comparable.
Overall, these findings cast doubt on the claim that intra-racial inequalities based on skin color gradients equate or exceed categorical inequalities between ethnoracial groups. While such assertions might be valid in certain areas – such as health (Monk 2015) –, this is not the case for educational and labor market outcomes. Importantly, these results are derived from a high quality, nationally representative longitudinal sample of American youth, including one of the most reliable measure of skin color to date within existing surveys (see Abascal et al. 2022, p. 354).


Supplementary Materials

Are within-racial group inequalities by skin color really greater than inequalities between racial groups in the United States?

Data and Measures

This article uses data from the National Longitudinal Survey of Youth 1997, a cohort study of individuals born between 1980 and 1984, who were between 12 and 17 years old at the beginning of the survey in 1997.

The four dependent variables in this study are years of schooling, college degree attainment, personal earnings, and total household income:

- Years of schooling are inferred from the highest degree achieved by the respondent as of 2013.
- College attainment measures whether the respondent had achieved a Bachelor’s degree or higher by 2013.
- Personal earnings are determined by the earnings before taxes as reported by the interviewee, covering wages or salaries, tips, bonuses, overtime pay, and self-employment income. I computed the three-year average incomes for the years 2010, 2011, and 2013.
- Total household income is the combined pre-tax income of all household members contributing to household expenses. Similarly, I computed the three-year average incomes for the years 2010, 2011, and 2013.

The focal independent variable in this study is skin color. NLSY97 relies on interviewer assessments of respondents’ skin tones, using a 10-point visual color palette, where 1 denotes the lightest color and 10 the darkest. This rating occurred during the 12th to 14th rounds of NLSY97, when respondents were aged 24 to 30. Each year, interviewers rated the skin tones of respondents who had not been recorded in earlier assessments.

Other variables in the analyses include:

- Age
- Gender
• Race, measured as self-reported race/ethnicity. Since the survey asks about Hispanicity in a separate question, I have given precedence to Hispanicity over racial categories. Throughout the article, I refer to non-Hispanic Whites and non-Hispanic Blacks as Whites and Blacks, respectively.

• Parental education, measured as the highest level of education completed by any of the biological parent.

• Parental income, defined as the total pre-tax household income from 1997, including earnings from all household members. Note that in the NLSY97, only 5% of the cases involved young respondents working for pay.

• Family structure, measured with an indicator variable for whether the respondent was part of an intact family, meaning both biological parents reside in the household.

The analytic sample is restricted to White, Black, and Hispanic individuals who are not enrolled in an educational institution as of 2010.

Modeling

Each dependent variable was modeled as a function of the interaction between skin tone—treated as a continuous variable — and ethnoracial group. Additionally, the adjusted models accounted for sex, gender, parental education, parental income, and family structure. Parameters were estimated using the Hamiltonian Monte Carlo procedure, as implemented in the Stan software for Bayesian modeling.

1. The general structure of unadjusted models is as follows:

\[ y_i \sim \begin{cases} 
\text{Bernoulli}(\logit^{-1}(\eta_i)) & \text{for binary outcomes} \\
\mathcal{N}(\eta_i, \sigma^2) & \text{for continuous outcomes}
\end{cases} \]

where

\[ \eta_i = \alpha + \beta_{\text{race}_i} + \beta_{\text{color}[\text{race}_i]} \times \text{color}_i \]

\( \alpha \) is the intercept, \( \beta_{\text{race}} \) are the coefficients for each race category, and \( \beta_{\text{color}} \) are the coefficients for color, specific to each race category. The variable \( i \) indexes the observations.

Priors:

• \( \alpha \sim \text{Student’s } t(3, 0, 10) \)
• \( \beta_{\text{race}}, \beta_{\text{color}} \sim \text{Student’s } t(3, 0, 2.5) \)
2. The general structure of **adjusted models** is:

\[
y_i \sim \begin{cases} 
\text{Bernoulli}(\logit^{-1}(\eta_i)) & \text{for binary outcomes} \\
\mathcal{N}(\eta_i, \sigma^2) & \text{for continuous outcomes}
\end{cases}
\]

where

\[
\eta_i = \alpha + \beta_{\text{race}_i} + \beta_{\text{color}[\text{race}_i]} \times \text{color}_i + \mathbf{X}_i \beta_{\text{controls}}
\]

\(\alpha\) represents the intercept, \(\beta_{\text{race}}\) are the coefficients for each race category, \(\beta_{\text{color}}\) are the coefficients for skin color within each race category, and \(\beta_{\text{controls}}\) are the coefficients for the control variables. \(\mathbf{X}_i\) is the vector of control variables for observation \(i\).

**Priors:**

- \(\alpha \sim \text{Student’s } t(3, 0, 10)\)
- \(\beta_{\text{race}}, \beta_{\text{color}}, \beta_{\text{controls}} \sim \text{Student’s } t(3, 0, 2.5)\)

**Priors choice:**

The choice of Student’s t-distribution for the priors, with a mean of 0 and a relatively small scale parameter (10 for \(\alpha\) and 2.5 for the \(\beta\)s), allows for heavier tails than a normal distribution. This accounts for potential outliers or extreme values in the parameters, providing a more robust modeling approach.

**Replication Code**

Code to replicate results and visualizations is available at: [https://osf.io/2zukw/?view_only=77fcc5f48fda4098a97eb4ae9d96c6f3](https://osf.io/2zukw/?view_only=77fcc5f48fda4098a97eb4ae9d96c6f3)