What's in a Mane?: Appearance Norms and Racialized Hair Bias*

Tarikua Erda† Jeffrey Shrader‡

December 2024

[Latest version]

Abstract

Hair texture is a phenotypical feature, yet Western norms favor straight hair over the kinky natural textures of Black populations. We conduct experiments in which university students and US adults rate the professionalism, competence, and agreeableness of headshots of hypothetical candidates for STEM professor jobs, with incentive-compatible rewards for truthful revelation. Relative to straightened hair (for women) and clean-cut hair (for men), kinky hair carries large penalties for Black individuals. These penalties are most pronounced for the professionalism and competence ratings of Black women and agreeableness of Black men. Racialized hair penalties also surpass penalties for objective photo elements like poor resolution or not smiling, implying that in mixed-race workplaces, Black individuals face penalties for not tightly adhering to Euro-centric hairstyle norms. Beyond advancing the study of systemic discrimination by using advanced experimental methods to document a pervasive yet underexplored type of bias, this paper also informs current legal discourse to outlaw hair-based racial discrimination in the US.

JEL Codes: J15, Z10
Keywords: bias, norms, STEM innovation and productivity

*We acknowledge generous funding from the Columbia Experimental Lab for the Social Sciences (CELSS) and the Institute for Social and Economic Research and Policy (ISERP). For feedback and comments, we thank Douglas Almond, Elizabeth Ananat, Sandra Black, Mark Dean, Judd Kessler, Corinne Low, Suresh Naidu, Ebonya Washington, Jack Willis, and participants in the CELSS Experimental Lunch (2024). Aryan Ranjan provided excellent research assistance. This study was carried out under Columbia IRB (Protocol: AAAU0417). AEA-Preregistry information: AEARCTR-0010183. All errors are ours.

†Erda: Columbia University, tarikua.erda@columbia.edu.
‡Shrader: Columbia University, jgs2103@columbia.edu.

1 Introduction

Systemic discrimination distorts optimal allocation of talent and economic growth. This is important in every sector of the economy, but particularly so in the Science, Technology, Engineering and Mathematics (STEM) fields which drive innovation, technological advancement, and ultimately economic growth. Recent decades have seen increased representation of racial minorities in STEM fields. Our paper provides novel evidence that, underlying this progress are large and persistent yet invisible and one-sided costs related to hairstyling that have been disproportionately borne by Black individuals.

Since the days of Transatlantic Slave Trade, deep-rooted and systemic biases against the kinky/coily hair phenotype of populations of African origin have significantly impacted how Black people, particularly Black women, navigate everyday life. For women, straight hair—a phenotype of white and other non-Black populations—and, for men, polished/combed hair, are often viewed as the ideal of professional appearance. Black women's hair is more likely to be viewed as unprofessional, and they are also more likely to be sent home from work or school because of their hairstyle. Institutional rules around hair also affect Black men, with several stories of young Black males facing schools suspensions and forced hair cuts before sportsgames due to their dreadlocks—a hairstyle that holds cultural significance to and is typically frequented by Black peoples around the world.[1] As such, Black people often manipulate their hair to maintain employability: 80% of Black women report changing their hair from its natural coily/kinky state to fit in at the office.

While these practices by Black individuals to conform to Euro-centric norms of appearance may seem benign due to their ubiquity, the available evidence and recent legislation concerning hair bias suggest the opposite. They likely impose substantial costs to Black individuals' not only to time and energy, but also to health,[2], cognitive load, psychological well-being, not to mention their finances. While there are no specific statistics on hair care spending across racial groups, the nationally-representative BLS Consumer Expenditure Survey shows that in 2019, White Americans spent 1.09% of income on personal care while Black Americans spent 1.39%—a 30 percent larger income share. In recognition of these multifaceted costs, the CROWN Act (Creating a Respectful and Open World for Natural

[1] According to this account, dreadlocks date back far into human history and are deeply intertwined with many different societies' spiritual and cultural underpinnings, including in ancient Egypt, Ethiopia, India. Today, they are donned by people of many different backgrounds, but the stereotypes of being lazy, dirty, unkempt are most punitive of the less powerful groups, namely racial minorities.

[2] Chemical hair straighteners have been linked to several toxic, and even carcinogenic, outcomes. Thermal straightening can cause heat damage that permanently alters one's curl pattern and may lead to hair loss. Moreover, one in three Black women skips exercise to preserve her hairstyle compared to one in ten white women (Johnson et al., 2017).

hair) was introduced in 2019 to protect Afro-textured hair and Afro-centric hairstyles (e.g. dreadlocks and braids) as an immutable aspect of race and/or cultural identity, thereby making hair-based discrimination punishable by law, and is currently adopted in more than 20 US states. The United Kingdom recognized hair texture and hairstyles associated with race and ethnicity as protected characteristics in 2010 with its nationwide adoption of the Equality Act. Still our understanding of how exactly hair bias operates relative to other elements of appearance that individuals may have more control over, and the intersectionality of race and gender in determining its impacts remains poor. That is the focus of this paper, which studies the topic in STEM academic settings, a crucial input into workforce development and economic growth.

It is unclear how exactly appearance and hairstyle are perceived in academic settings, and whether race plays a factor. On one hand, academia prides itself on meritocracy, seeking to elevate the best ideas and pursue knowledge. Thus, appearance may not matter at all. On the other hand, the human brain takes less than a tenth of a second to decide an individual's attractiveness, trustworthiness, and competence (Willis & Todorov, 2006), with traits like attractiveness having been linked to positive outcomes in seemingly meritocracy-centric settings such as academic careers in economics (Hale et al., 2023) and entrepreneurial pitch competitions with real money at stakes (Brooks et al., 2014). Thus hair bias that has been documented in industry settings[3] may also operate in academic settings. Conversely, looks that may be considered 'less polished' or unruly by wider societal norms (e.g., the popular trope of the mad scientist with crazy hair (Chambers, 1983)) may be positive signals of a scholars' competence, though perhaps differentially by race. To test these hypotheses, we conduct pre-registered experiments to examine how different hairstyles of Black individuals are evaluated according to appearance norms in STEM academic settings.

We survey university students at Columbia University and Barnard College as well as a representative sample of US adults, asking each of them to rate 32 headshots of hypothetical candidates for a STEM professor job on professionalism, competence, and agreeableness. The experimental photos consisted of headshots of four models—a Black female, a white female, a Black male, and a white male—generated from a photoshoot and subsequent Photoshop editing, both of which were both carefully monitored by the authors. To set a common benchmark for what respondents understand to be the pictured models' natural hair texture, we prime them with casual weekend photos of the models at the start of the headshot rating exercise, implying that the white male and female models both have curly hair, and that the Black male and female models have kinky hair. In the rating activity,

[3]See Opie and Phillips (2015); Koval and Rosette (2021)

we include three hairstyle variations: curly/kinky hair for white/Black models, shoulder-length straight hair for female models (and, equivalently, clean-cut hair for male models), and dreadlocks. The curly/kinky natural hairstyles of the models were found to represent comparable levels of effort in prior focus group surveys of white and Black female adults of the respective races who (self-reportedly) had those respective hair textures.

The central challenge with studies of bias like ours is the difficulty of identifying subjects' honest responses, instead of what they deem to be socially-desirable answers, yet our experimental design overcomes this challenge through two key mechanisms. First, to offer subjects plausible deniability and downplay our study's focus on hair, we interweave highly-controlled experimental variation unrelated to hairstyle in the headshots.[4] We vary models' clothing (suit jacket vs t-shirt), facial expression (smiling vs serious expression), and the photo resolution (high resolution vs blurred). To reduce fixation on differences across white and Black models, we exposed each respondent to non-experimental photos of four models of other races for, on average, 14% of her ratings.

Second, we employ a non-deceptive and incentive compatible design to elicit honest responses from subjects. Prior to our main surveys, we had conducted a similar but shorter headshot-rating exercise with a small group of university staff at a New York-area university. We explained this to respondents in our two main samples and promised that we would randomly select one of the headshots rated by the university staff and offer a *bonus* cash prize of $100 each to three respondents who accurately predicted the average rating across the university staff on that photo. This approach of asking about respondents' higher-order beliefs about *other* individuals in STEM academia not only incentivizes them to expend effort in each of their ratings, but also makes it easier/safer to express views they may consider to be socially undesirable, such as negative views of kinky hair or dreadlocks. We then use OLS to investigate the perception of different photo elements, particularly comparing within-model differences in hairstyles (e.g., straight hair versus dreadlocks) across Black and white models of the same gender.

We find pronounced penalties in perceived professionalism and competence for the Black female model when she wears her naturally kinky hair texture instead of straightened hair, including relative to the curly hair penalties faced by the white female model. In the student sample, the Black female's kinky hair penalty is statistically significantly higher than the curly hair penalty of the white female, its magnitude being twice the size of the contribution of smiling for perceived professionalism. While both students and lay adults

[4]The headshots were generated in a highly-controlled photo shoot and subsequent photo-editing, both of which were closely supervised by the authors.

also rate the Black female model as less competent when she wears her naturally kinky hair texture relative to straight hair, students rate her white counterpart's curly hair photos as more competent. In terms of perceived agreeableness, only the lay adult sample penalizes the Black female for wearing her natural hair texture, but we find no statistically significant evidence that this penalty is higher than that faced by the white female.

For the Black male model, wearing a bigger Afro instead of clean-cut/polished hair entails a pronounced differential in his perceived agreeableness relative to his white counterpart.[5] While students' big hair penalty for the Black male model is small and statistically insignificant, they in fact award an economically and statistically significant *premium* on big curly hair for the white male model. Meanwhile, lay adults Black male model as significantly less agreeable when he is pictured in bigger Afro hair but penalize the white male model's comparable hairstyle to a much smaller and statistically insignificant extent. Despite the different underlying mechanisms, there results substantial disparity in the perceived agreeableness of the Black and white male models in both samples, consistent with prior evidence that Black men—who are stereotyped as threatening—are penalized for even generally positive traits like height, while white men are rewarded for them (Hester and Gray (2018)).

Even though both student and lay adult respondents penalized the Black models for wearing dreadlocks relative to straight or clean-cut hairstyles, they penalized the white models far more severely for wearing dreadlocks, especially for the professionalism and competence ratings, expressing that they felt it was a form of cultural appropriation. Despite stated beliefs (in open-ended responses) about dreadlocks as a hairstyle of particular—even exclusive—cultural significance to Black and not white people, respondents' detailed ratings still penalize Black models for the hairstyle, with the penalty sizes being comparable to more objective measures of professionalism and competence like photo resolution. The dreadlock penalties for the white models relative to the Black models largely dissipate when it comes to perceived agreeableness, particularly in the student sample.

We conducted two information intervention treatments on respondents: a vignette highlighting that authenticity to personal style at the workplace increases worker happiness and productivity, and a longer explainer situating hair bias in the broader legacy of racial discrimination in the US.[6] Both treatments had limited efficacy in reducing the penalties Black models face above and beyond the white models for wearing their natural hair texture or locs, particularly the former.

[5] In regards to professionalism and competence, the Black male is penalized wearing a longer Afro instead of polished clean-cut hair but to a lower or comparable extent relative to the white male model.

[6] We also included experimenter demand induction treatments per de Quidt et al. (2018).

This paper contributes to several strands of literature. First, it speaks to the literature on structural discrimination and the costs of assimilation. Previous work shows evidence of racial discrimination on the basis of Black-sounding names (Bertrand & Mullainathan, 2004), darker skin-tone (Darity et al., 2006), African/Black phenotypic facial features (Kahn & Davies, 2011) and other characteristics. Changing these aforementioned characteristics would require long, expensive, and potentially irreversible or medically-invasive processes. In contrast, processes to manipulate Afro-textured coily/kinky hair such as hair straightening processes are commonplace, often reversible, and appear relatively cheap on the surface,[7] though they may have non-negligible and lasting costs to finances and health as discussed at more length in Section 2. The fact that some of the large penalties we document on the perceived professionalism and competence of the Black female model's natural kinky hair texture are higher than the penalties for elements that are actually within individuals' control, such as smiling and setting a high photo resolution, demonstrates novel evidence on the high cognitive load comes with conventional appearance norms to which Black individuals often have to conform. As such, unlike the monetary focus of other identity-specific costs documented in prior literature like gender-based price discrimination (i.e., the pink tax) (de Blasio & Menin, 2015), our paper demonstrates the intersectional, multifaceted, and all-consuming nature of racialized hair discrimination.

This study also adds to the literature on hair discrimination. To our knowledge, this is the first paper to study hair bias in the economics literature. While other studies in adjacent social science disciplines have used experimental approaches and provided evidence on hair bias in industry settings such as advertising and consulting (Koval & Rosette, 2021) and corporate employment (Opie & Phillips, 2015), our study innovates on experimental design and better captures nuances. In particular, neither study includes an incentive-compatible design, examines whether and what kind of hair bias Black men face, nor accounts for social desirability bias, all of which are novel elements included in our pre-registered experiment. Our study is also unique in that it includes information interventions and evaluates their role, if any, on hair bias. Additionally, Opie and Phillips (2015) used largely photo-shopped stock photos, lacking the realism and highly-controlled variation of our experimental photos. In contrast to traditional methods to study bias like the Implicit Association Test (IAT),[8] which reveal the exact bias studied by the test to subjects, our approach of obfuscating the survey's central focus on hair bias through a more holistic and incentive-compatible headshot rating exercise represents a more innovative and reliable approach to study unconscious bias.

[7]Hair is extremely versatile: it can be cut, shaved, grown out, colored, braided, covered (for various personal, religious, or cultural reasons).

[8]This approach was used in prior reports on hair bias (Johnson et al., 2017).

The remainder of the paper is organized as follows. Section 2 provides background on the relationship between hair style and racial biases. Section 3 lays out the experimental design. Section 4 describes the data generated by the experiment. Section 5 presents results. Section 6 discusses the results and concludes.

2 Background

2.1 History of Hair Bias

For centuries, hair was and continues to be used as a marker of various cultural indications such as age, marital status, social status and so on in Africa (Byrd & Tharps, 2002). Since the onset of Slave Trade, however, institutionalized racism has affected, and in part dictated, how Black people in the US navigate hair. Europeans shaved the heads of enslaved Africans upon arrival to the Americas to assert domination, reallocate time toward labor (Chapman 2007, Dash 2006, Morrow 1973). In 1786, the Tignon Law required *all* women of color– including emancipated ones–in New Orleans to wear a scarf over their hair as a visible sign of belonging to the 'slave class.' From late 1800s on, various hair products, mainly chemical and heat-based tools to straighten Black hair were introduced to the market. It was encouraged, even within commonly held wisdom in Black communities, that one had to 'tame' their kinky hair and straighten it to get and maintain jobs. In the 1960s' Civil Rights Movement, more Black people embraced the act of simply wearing their natural hair texture as an Afro as a political act of defiance. This is potentially linked with why wearing Afros may be seen as indicative of increased dominance in Black women in present day psychology studied (Opie & Phillips, 2015). However, it was not too long before hair straightening products gained traction again, and other styles like locs, braids etc on the rise, with greater 'natural hair movement' since early 2000s. Today, the CROWN Act (2019) outlaws hair bias & discrimination (now adopted in more than 20 states).

2.2 Economic Implications of Hair Bias

Deep-rooted biases against Afro-textured hair have blurred the line between racial phenotypes and socio-cultural constructs like professionalism and presentability. Black women's hair is 3.4 times more likely to be perceived as unprofessional (Collective, 2019). Google Image searches for "unprofessional hairstyles" return predominantly pictures of Black women with Afro-textured hair (LaCapria, 2020). Black women are also 1.5 times more likely to be

sent home from work because of their hairstyle (Collective, 2019). Thus, 80 percent of them report changing their hair from its natural curly/coily state using chemicals or heat to fit in at the office (Collective, 2019). One in three Black women also skips exercise to preserve their (straightened) hairstyle compared to one in ten white women (Johnson et al., 2017).[9] Moreover, chemical hair straighteners (called "relaxers") have been linked to several toxic, and even carcinogenic, outcomes[10] while thermal straightening can cause heat damage that permanently alters one's curl pattern and may even result in hair loss.[11] There are financial disparities as well. While there are no specific statistics on hair care spending across racial groups, the nationally-representative BLS Consumer Expenditure Survey finds that in 2019, white Americans spent 1.09 percent of income on personal care while Black Americans spent 1.39 percent—a 30 percent larger income share.[12] Therefore, the Euro-centric norms of presentability and professionalism place an excess grooming-burden on Black people, not only imposing additional time and energy burdens of hairstyling but also costs to finances and physical well-being.

3 Experimental Design

3.1 Experimental Vistual Stimuli

Since our study is centered on appearance, the visual stimuli in our survey are a key element in our survey. We conducted an original photoshoot in May 2022 with four main models: a Black female, a white female, a Black male, and a white male. Models were photographs in different attires and facial expressions and the authors also closely supervised the photoshoot to minimize variation in elements that would not be directly controlled for in our experimental analyses. We subsequently hired a photo editor through the online platform Fiverr to generate pictures of them in different hairstyles. The blurring effect on the originally high resolution photos was added after all photo editing was completed.

A key potential confounder when studying the perception of hair texture, particularly that of women, in large samples of respondents from various gender and racial backgrounds is the level of effort signaled by hairstyle, which we accounted for by running smaller online focus group surveys related to the perception of effort. This is because individuals are more

[9] (Woolford et al., 2016; James-Todd et al., 2011).

[10] See, Eberle et al. (2020); França-Stefoni et al. (2015); De Sá Dias et al. (2007); James-Todd et al. (2011); Woolford et al. (2016).

[11] Callender et al. (2004).

[12] Bureau of Labor Statistics (2019). https://www.bls.gov/cex/

likely to be familiar with their own hair texture and the effort involved with certain hair styles in their gender and racial demographic group, compared to the hair texture and effort of other demographic groups. This is especially relevant to women's hair they face a larger grooming burden in many societies, and hair straightening is common in the US among women of various races. Thus, to account for the role of effort, we surveyed white and Black lay adults who (self-reportedly) had curly and kinky hair rated the level of effort suggested by the models' curly and kinky hairstyles, respectively. We asked subjects to think of effort put into the pictured hairstyle as a combination of time spent on the pictured hair, the amount of hair products applied, and (3) the physical energy the pictured individual spent working with different tools and manipulating their hair to create the pictured hairstyle. The white and Black female models' hair were rated to signal comparable levels of effort by white and Black female focus group respondents.

Since men face a lower grooming burden and are far less likely to manipulate their hair texture (e.g., straighten curly/kinky hair) relative to women, what matters for our study of the perception of men's hair is that there are palpable differences in effort along the lines of the recency of a hair cut.[13] As such, the non-dreadlock hairstyle variations for the male models in our study are based on whether the model has has longer/bigger curly/kinky-Afro hair or shorter hair that looks clean-cut/polished, suggesting a more recent visit to a barber.

The final set of experimental visual stimuli were prepared by combining, for each of the four models, 3 hairstyle variations (kinky/curly, dreadlocks, and straight/clean-cut hair), 2 clothing variations (suit and t-shirt), 2 facial expressions (smiling and non-smiling) and, for all of the above 24 combinations but three of them, 2 photo-resolution levels (high-resolution and blurry). Because of this last step of not including blurry versions for 3 photos per model, we ended up with 84 instead of 96 unique experimental photos, which meaningfully increases each photo's probability of being shown to and rated by respondents while retaining decreasing our ability to recover the role of photo resolution on respondent rating through the combinations of estimates on the rest of the model's photos.

To add racial diversity in the set of photos that subjects rated, we also included 12 stock photos of individuals who presented as neither white nor Black but whose photo characteristics (e.g., photo background, clothing style and angle) generally resembled those of our experimental photos. This brought the total number of visual stimuli from which respondents would rate to 98. Thus, on average, respondents in our experiment encountered the main experimental photos of our four white and Black models about 86% of the time.

[13]Unsurprisingly, the aforementioned definition of 'effort' we gave to focus group participants—which indeed accomplished its purpose of distinguishing effort from hair texture for the female models' photos—resulted in seemingly highly subjective and varied estimates, particularly for the white male.

3.2 Survey Procedure

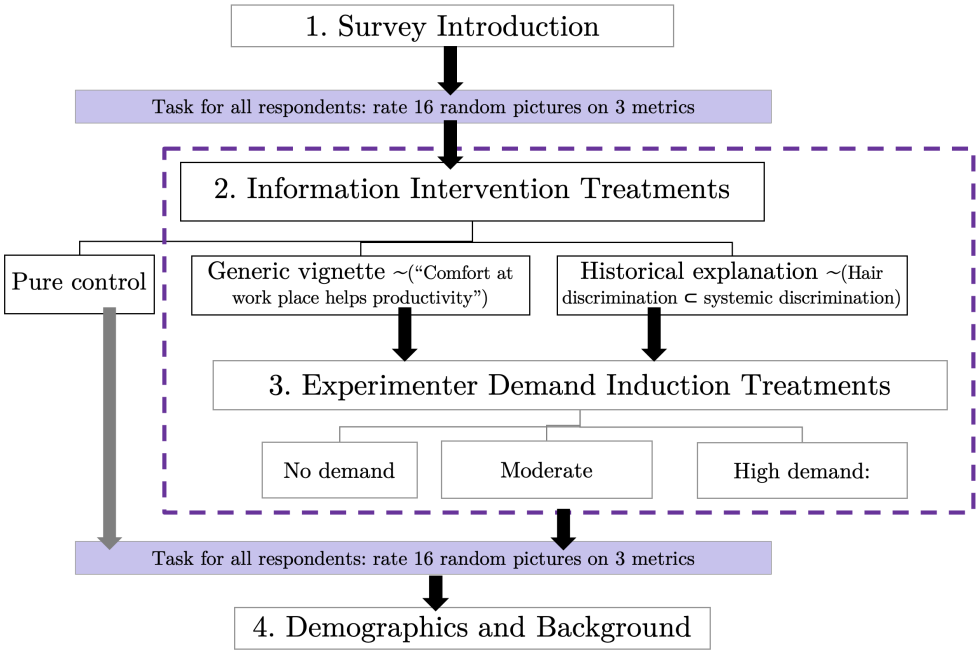
Students at Columbia University and Barnard College were recruited from the Online Recruitment System for Economic Experiments (ORSEE) subject pool while lay adults were recruited through CloudResearch. The main survey activity was rating 32 headshots in four blocks of 8 photos each. Since student participants rated headshots imagining that the pictured models were candidates for a university professor position in the field of their major, they were asked about their major at the start of the survey. Lay adults were all asked to imagine that pictured models were hypothetical candidates for STEM professor jobs. After completing the first two blocks of ratings (after 16 photo ratings), respondents were randomly assigned to one of three categories at an even 1/3 chance: a pure control group, an information treatment involving a vignette about how authenticity at the workplace enhances diversity, and another information intervention that explains the history of hair bias in the US and the CROWN Act. Before proceeding to the last 16 photos, treated subjects were subsequently assigned to one of three groups that received no experimenter demand induction (with 50% chance), a moderate or high experimenter demand (each with 25% chance). The survey flow is presented in Figure 1.

The rating activity involved three metrics: professionalism, competence, and agreeableness. We defined professionalism form them as the state of being in line with the norms of the university as a workplace, competence as being qualified, credible, and capable to rise to the demands of an academic job (teaching and doing research), and agreeableness as the quality of being collegeal and collaborative with other scholars as well as approachable to students seeking answers to their questions or mentorship from faculty. While respondents were asked to rate a Likert scale from 1-10, we also provided a non-response option ("N/A") in case they wanted to skip a rating.[14]

To ensure incentive compatibility, we gave respondents an explanation of the bonus cash prize we were offering for accurate responses right after we introduced the headshot rating activity and rating metrics. We explained that we had conducted a similar headshot rating exercise on university administrative staff who were in a position to train and advise university students in an approach that highly encouraged truthful responses. This was indeed true – we had conducted a headshot-rating exercise with university staff from a New York area higher education institution in May 2023, prior to the launch of the main surveys on students and lay adults in October 2023. As a reward for the participation of the university staff in our survey, we contributed money to a professional headshot session held

[14]For instance, the 'competence' rating, many subjects used this 'N/A' option or at least expressed discomfort with evaluating competence based on a appearance).

Figure 1: Survey Flow



for students of that institution and also shared aggregated insight from the survey responses as input for the headshot session.[15]

Prior to the start of the rating activity, we also presented subjects with a primer to 'acquaint' them with the 8 models (four main and four joker models) whose photos they would subsequently rate using small photos that showed the models on 'a casual weekend morning.' This was intended to implicitly give respondents a benchmark as to the natural state of each model on a casual day lest they make other, potentially heterogeneous, assumptions about the experimental models that interfere with our inference of the photo ratings. For example, without a benchmark, subjects may assume the natural hair texture of the white female model is in fact straight but that she manipulated it to look curly, which would impair our investigation of the perception of similar non-straight natural hair textures in the Black and white female models relative to straightened hair.

[15]This was intended to encourage truthful responses, in a fashion similar to the Incentivized Resume Rating approach pioneered by (Kessler et al., 2019) in which employers rate resumes they know to be hypothetical in return for matches with *real* job candidates.

4 Data and Empirical Strategy

4.1 Sample Descriptions

As Table A.1 shows, there are pronounced differences in the composition of our student and lay adult samples. As it is drawn from a representative sample of the US population, the lay adult sample mirrors the racial break-down of the US population and also has a roughly even split across male and female respondents. In contrast, more than 70% of students identify as female, which partly follows from the fact that a decent share of our respondents attend Barnard College, a women's college. The student sample, which was recruited through the ORSEE which does not allow targeted stratification of the sample, is primarily composed of students with white (46%) and Asian (39%) background, which are both much higher than the Columbia student population distribution. Black students, other multi-racial or minority groups made up a smaller proportion of the student sample, including compared to the relative shares in the Columbia and Barnard student populations. Additionally roughly half of lay adult respondents reported being democrats, and 22% republicans, while 71% of students reported being democrats and less than 3% republican.

About 67% of students had a STEM major. These majors largely include computer science, engineering, chemistry, physics, life and environmental sciences as well as cognitive science and quantitative-driven majors such as economics and data science. All students rated photos imagining that the pictured model was a candidate for a professor position in their major discipline. Our main results presented in Section 5 juxtapose the views of these students and contrast them against those of lay adults, who were all asked to imagine STEM professors in their ratings.

4.2 Estimation Strategy

Using the data from the first half of the ratings, we estimate the below specification.

$$\begin{aligned} V_{ij} = \beta_1 B_j + & \beta_2 G_j + \beta_3 H_{2,j} + \beta_4 H_{3,j} + \beta_5 B_j G_j + \beta_6 B_j H_{1,j} + \beta_7 B_j H_{2,j} \\ & + \beta_8 B_j H_{3,j} + \beta_9 G_j H_{1,j} + \beta_{10} G_j H_{2,j} + \beta_{11} G_j H_{3,j} + \beta_{12} B_j G_j H_{1,j} \\ & + \beta_{13} B_j G_j H_{2,j} + \beta_{14} B_j G_j H_{3,j} + X_{ij}'\theta + \alpha_i + \varepsilon_{ij} \end{aligned}$$

V_{ij} denote respondent i's rating of a photo j on the 1–10 Likert scale. B denotes Black model, G denotes female model, H_1, H_2 and H_3 are hair style indicators for curly/kinky

hair, dreadlocks, and straightened (and for the male models, clean-cut hair), respectively. The other variables in the regression are controls. These include a subject-level fixed effects α_i and, in additional specifications, controls for other photo attributes (represented by X): an indicator for whether the model is wearing a suit jacket (C), and indicator for whether the photo is in high resolution (R), and an indicator for whether the model is smiling (E).

The main hypotheses we will test are if β_{12}=0 and β_{14}=0 to understand whether the Black female model rated the same for curly or loc hairstyles as for straight. Additionally, we will compare the relative penalties faced by Black and white models of of the same gender for curly/kinky hair relative to straight/clean-cut hair.

5 Results

5.1 Baseline ratings

Tables 1, 2, and 3 show regression results from both the university student and lay adult samples where V_{ij} takes values between 1 to 10 for ratings of perceived professionalism, competence, and agreeableness. Within each table, the first and third columns report results related to respondents' ratings for female models, while the second and fourth columns report results related to male models. While the regressions show the role of clothing, facial expression and photo resolution on photo ratings, the main coefficients of interest are the interactions between race and hairstyle, with the photos of the white models with straight hair (with clean-cut hair for the male model) as the base group. We provide t-test results at the bottom of each table for the null hypothesis that the rating difference between naturally curly/kinky and straight hair for the female models and between 'bigger' curly hair and clean-cut hair for the male models is equal Black and white models of the same gender.

In both samples, wearing the natural curly/kinky hair is generally rated as less professional than straight hair for female models, with the Black female model facing a higher penalty. The Black female model's penalty for wearing her natural hair relative to straightening it stands at 0.564 points in the student sample and 0.39 in the lay adult sample. While the white female model also faces a penalty for wearing her naturally curly hair relative to straight hair, the magnitudes are lower at 0.21 and 0.29 points in the student and lay adult samples, respectively. These estimates suggest the presence of large and more pronounced relative penalties for the Black female model relative to the while female model, particularly standing at 0.351 points ($p<0.01$) in the STEM student sample.

Among the male models, these results on rating differentials are somewhat reversed

Table 1: STEM students and lay adults: Professionalism Ratings for Male and Female models

	STEM students		Lay adults	
	Female	Male	Female	Male
Suit	2.287***	2.352***	2.744***	2.798***
	(0.098)	(0.101)	(0.052)	(0.052)
High resolution	0.117	0.211***	0.195***	0.161***
	(0.072)	(0.076)	(0.030)	(0.030)
Smile	0.188***	-0.009	0.379***	0.250***
	(0.059)	(0.065)	(0.030)	(0.029)
Black X Curly	-0.278**	0.382***	-0.352***	-0.067
	(0.110)	(0.111)	(0.053)	(0.055)
Black X Locs	-0.192*	0.084	-0.322***	-0.224***
	(0.109)	(0.115)	(0.053)	(0.060)
Black X Straight	0.286***	0.434***	0.038	0.391***
	(0.094)	(0.115)	(0.051)	(0.052)
White X Curly	-0.212**	-0.412***	-0.291***	-0.580***
	(0.099)	(0.112)	(0.050)	(0.053)
White X Locs	-1.196***	-1.194***	-1.120***	-1.119***
	(0.117)	(0.123)	(0.058)	(0.058)
White X Straight	.	.	.	.
	.	.	.	.
Constant	5.152***	5.039***	4.583***	4.514***
	(0.100)	(0.113)	(0.050)	(0.052)
Rating differentials	**Female**	**Male**	**Female**	**Male**
Black: Curly–straight/clean-cut hair	-0.563***	-0.052	-0.390***	-0.458***
	(0.101)	(0.101)	(0.049)	(0.048)
Black: locs–straight/clean-cut hair	-0.478***	-0.350***	-0.360***	-0.615***
	(0.094)	(0.100)	(0.049)	(0.054)
Black vs white: Curly–straight/clean-cut	-0.351***	0.360**	-0.098	0.123*
	(0.134)	(0.156)	(0.068)	(0.067)
Black vs white: Locs–straight/clean-cut	0.718***	0.844***	0.760***	0.504***
	(0.146)	(0.155)	(0.074)	(0.072)
Observations	2601	2549	12753	12942

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All models contain subject FE

across racial lines. Namely, the Black male model loses about 0.05 points in the student sample and 0.46 in the lay adult sample for wearing a longer Afro compared to a clean-cut look while the white male loses 0.41 and 0.58 points, respectively. These results suggest that clean-cut hair is rated relatively higher in professionalism compared to longer and bigger curly hair, but that the rating penalty for big curly hair is higher for the white male.

We also find that both the white female and male models face large penalties for wearing dreadlocks from both samples. In our free-form response questions as well, several respondents expressed strong opposition against white individuals wearing dreadlocks, calling it a form of cultural appropriation. These severe penalties for white models pictured in dreadlocks largely occur across all three metrics but the magnitudes are much more economically and statistically significant for the professionalism and competence ratings.

Lastly, Table 1 shows that, clothing is the most crucial driver of perceived professionalism: across both the student and lay adult samples, candidates get higher ratings when pictured in a suit jacket as opposed to a t-shirt. Both samples also give higher professionalism ratings for high resolution photos and those in which the candidate is smiling, though the role of smiling appears more pronounced for female candidates than for male candidates.

When it comes to perceived competence, Table 2 suggests, across both samples, the Black female model experiences a large rating differential when she is pictured in her natural kinky/coily hair texture relative to in straightened hair. This differential comes from a large rating premium for straight hair in the student sample, and in the lay adult sample, it comes from a large penalty for wearing her natural texture. On the other hand, the white female model sees no statistically significant rewards or penalties from the student sample for wearing her naturally curly hair texture relative to straightened hair, while she is penalized by the lay adult sample. Ultimately, both of these outcomes lead to a more pronounced relative competence perception penalty for natural kinky hair compared to straight hair for the Black female model. These patterns are not observed in the long-curly hair vs clean-cut hair results of the male models, suggesting the presence of highly gendered patterns in which hairstyle factors into the perceived competence of individuals in STEM.

It is also notable in Table 2 that wearing a suit jacket greatly enhances the perception of competence among both samples. Respondents give higher competence ratings when candidates are pictured in high-resolution photographs and when they are smiling. Here, too, the coefficient sizes suggest a far larger positive benefit of smiling for female models compared to male models.

Table 3 presents regression results related to perceived agreeableness ratings. In the student sample, the Black female model is rated as less agreeable (by about 0.13 points when

Table 2: STEM students and lay adults: Competence Ratings for Male and Female models

	STEM students		Lay adults	
	Female	Male	Female	Male
Suit	1.141***	1.068***	1.517***	1.510***
	(0.080)	(0.079)	(0.040)	(0.040)
High resolution	0.085	0.233***	0.221***	0.234***
	(0.059)	(0.063)	(0.027)	(0.027)
Smile	0.330***	0.099*	0.550***	0.302***
	(0.059)	(0.060)	(0.029)	(0.028)
Black X Curly	0.019	0.444***	-0.336***	0.068
	(0.101)	(0.106)	(0.050)	(0.052)
Black X Locs	-0.035	0.253**	-0.283***	-0.071
	(0.108)	(0.111)	(0.049)	(0.055)
Black X Straight	0.307***	0.531***	-0.031	0.418***
	(0.096)	(0.105)	(0.048)	(0.051)
White X Curly	0.061	-0.160	-0.169***	-0.405***
	(0.095)	(0.100)	(0.048)	(0.048)
White X Locs	-0.936***	-0.861***	-0.907***	-0.938***
	(0.107)	(0.121)	(0.054)	(0.053)
White X Straight	.	.	.	.
	.	.	.	.
Constant	5.836***	5.911***	5.536***	5.543***
	(0.092)	(0.098)	(0.044)	(0.046)
Rating differentials	**Female**	**Male**	**Female**	**Male**
Black: Curly–straight/clean-cut hair	-0.288***	-0.087	-0.305***	-0.350***
	(0.091)	(0.091)	(0.044)	(0.044)
Black: Locs–straight/clean-cut hair	-0.341***	-0.278***	-0.252***	-0.489***
	(0.080)	(0.091)	(0.044)	(0.049)
Black vs white: Curly-straight/clean-cut	-0.349***	0.073	-0.136**	0.056
	(0.132)	(0.141)	(0.062)	(0.064)
Black vs white: Locs-straight/clean-cut	0.594***	0.583***	0.654***	0.449***
	(0.130)	(0.141)	(0.067)	(0.068)
Observations	2583	2525	12744	12929

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All models contain subject FE

she wears her hair in its natural coily/kinky state as opposed to straightening it, yet the white female receives a statistically insigificant yet large 0.08 point *premium* in her agreeableness rating for wearing natural curly hair texture without straightening it. Meanwhile, the lay adult sample rates the white female model as less agreeable (about about 0.127 points) when she wears her natural curly hair, but gives no statistically significant relative penalty to the Black female model for wearing her natural hair texture.

Conversely, the Black male faces large and robust penalties for donning a bigger Afro hair compared to a clean-cut hair. In particular, the student sample rates the white male model as more agreeable when he is pictured in a long, big curly hairstyle as opposed to a clean-cut hairstyle, but significantly penalizes the Black male model for the same action, ultimately resulting in a large relative hair penalty of about 0.39 points ($p<0.01$) for the Black model for wearing a longer Afro as opposed to clean-cut hair. Similarly, the lay adult sample also gives a higher penalty (about 0.163 points, $p<0.05$) to the Black male model relative to the white male model for wearing a large Afro relative to a clean-cut hairstyle.

The lay adult sample rates all four models as lower in agreeableness when they wear their natural hair texture as opposed to straight (for the women) or clean-cut hair (for the men). Here, too, the penalties are higher for the Black models relative to the white models. This is consistent with prior findings that Black women with Afro-centric hairstyles are perceived as aggressive and dominant (Opie & Phillips, 2015).

Unsurprisingly, smiling is the most important driver of high agreeableness ratings among both samples and for all models. As with the other rating metrics, smiling plays a relatively larger role for female models than male ones. Clothing plays a much smaller role compared to its role in professionalism and competence ratings. Notably, however, respondents, particularly students, rate female models pictured in high-resolution photos as lower in agreeableness. One potential explanation for this could be that women who show appear flawed/imperfect may be perceived as more relatable compared to flawless ones.

Treatment Effects

Our treatments had do not seem to have substantially (and statistically significantly) affected respondents' evaluation of hairstyle. The detailed regression results—which we estimate by additionally interacting each hairstyle and race combination in Equation (4.2) with a treatment status and post-treatment indicator (i.e., TREAT X POST)—are presented in Tables A.2-A.4 in the Appendix.

Table 3: STEM students and lay adults: Agreeableness Ratings for Male and Female models

	STEM students		Lay adults	
	Female	Male	Female	Male
Suit	0.348***	0.312***	0.604***	0.606***
	(0.073)	(0.074)	(0.039)	(0.037)
High resolution	-0.140**	0.189***	-0.022	0.229***
	(0.066)	(0.062)	(0.031)	(0.028)
Smile	1.928***	1.673***	2.151***	1.818***
	(0.086)	(0.086)	(0.046)	(0.043)
Black X Curly	0.384***	1.415***	-0.087	0.555***
	(0.111)	(0.110)	(0.058)	(0.056)
Black X Locs	0.308**	1.323***	-0.130**	0.464***
	(0.121)	(0.131)	(0.057)	(0.061)
Black X Straight	0.516***	1.467***	0.077	0.803***
	(0.109)	(0.119)	(0.055)	(0.057)
White X Curly	0.084	0.336***	-0.127**	-0.085
	(0.106)	(0.109)	(0.055)	(0.055)
White X Locs	-0.375***	-0.261**	-0.532***	-0.555***
	(0.116)	(0.114)	(0.057)	(0.057)
White X Straight	.	.	.	.
	.	.	.	.
Constant	5.157***	5.128***	5.058***	5.174***
	(0.097)	(0.113)	(0.050)	(0.052)
Rating differentials	**Female**	**Male**	**Female**	**Male**
Black: Curly–straight/clean-cut hair	-0.132	-0.052	-0.164***	-0.248***
	(0.098)	(0.094)	(0.051)	(0.046)
Black: Locs–straight/clean-cut hair	-0.207**	-0.144	-0.207***	-0.339***
	(0.100)	(0.100)	(0.049)	(0.051)
Black vs white: Curly–straight/clean-cut	-0.216	-0.388***	-0.037	-0.163**
	(0.145)	(0.146)	(0.075)	(0.071)
Black vs white: Locs–straight/clean-cut	0.168	0.117	0.325***	0.216***
	(0.149)	(0.138)	(0.073)	(0.073)
Observations	2593	2537	12751	12941

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All models contain subject FE

6 Discussion and Conclusion

The biases documented in this paper reveal new insight into the disproportionate costs borne by Black individuals in the US STEM workforce which provides the innovation and R&D that drive economic growth. Previous work shows that systemic sexism and racism have hindered the US economy via missing patents, foregone output, and greater inequality.[16] Conversely, diverse perspectives informed by individuals' gender and racial identities, among other attributes, enhance patent novelty (Yang et al., 2022), and medical research attention to diseases (Dossi, 2024). Concerted efforts to grow the representation of Black individuals in STEM fields in recent decades, and have borne fruit. A diverse body of educators in STEM higher education can also be a boon for attracting and retaining a diverse student-body through role-model effects (Redding, 2019; Baker, 2013). Yet the results in this study suggest that there may be more to be done to improve the inclusivity of workplace norms to lower the one-sided costs of assimilation borne by Black individuals.[17]. Lastly, it is important to note that, in realistic situations, there is often more information about a person available than just their picture; e.g., their clothes, voice, mannerisms, accent, or name. It may be possible that our results overstate or understate bias in various ways because the visual setting limits many other contextual factors that people may make judgments on.

[16] Cook and Yang (2018); Hunt et al. (2012).

[17] For example, in a survey about the experiences of underrepresented minorities in economics, a Black associate professor reported that she is "usually the first Black economist [her students] have ever seen. And that in and of itself has its own issues, like are you qualified? Are you competent?" Moreover, she noted that student evaluations about her – and other minority faculties' – courses "veer off [into] personality... They might like my class [but] they'd be like, 'But I don't like.' And it's usually something really personal or how I dress" (Bayer et al., 2020, pp. 12-13).

References

Baker, C. N. (2013, December). Social Support and Success in Higher Education: The Influence of On-Campus Support on African American and Latino College Students. *The Urban Review*, *45*(5), 632–650. doi: 10.1007/s11256-013-0234-9

Bertrand, M., & Mullainathan, S. (2004, September). Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review*, *94*(4), 991–1013. doi: 10.1257/0002828042002561

Brooks, A. W., Huang, L., Kearney, S. W., & Murray, F. E. (2014, March). Investors prefer entrepreneurial ventures pitched by attractive men. *Proceedings of the National Academy of Sciences*, *111*(12), 4427–4431. doi: 10.1073/pnas.1321202111

Bureau of Labor Statistics. (2019). *Consumer Expenditure Survey.* https://www.bls.gov/cex/.

Callender, V. D., McMichael, A. J., & Cohen, G. F. (2004). Medical and surgical therapies for alopecias in black women. *Dermatologic Therapy*, *17*(2), 164–176. doi: 10.1111/j.1396-0296.2004.04017.x

Chambers, D. W. (1983). Stereotypic images of the scientist: The draw-a-scientist test. *Science Education*, *67*(2), 255–265. doi: 10.1002/sce.3730670213

Collective, J. (2019). *The CROWN Research Study.*

Cook, L., & Yang, Y. (2018). Missing Women and African Americans, Innovation, and Economic Growth. *Working paper.*.

Darity, W. A., Mason, P. L., & Stewart, J. B. (2006, July). The economics of identity: The origin and persistence of racial identity norms. *Journal of Economic Behavior & Organization*, *60*(3), 283–305. doi: 10.1016/j.jebo.2004.09.005

de Blasio, B., & Menin, J. (2015, December). From Cradle to Cane: The Cost of Being a Female Consumer.

de Quidt, J., Haushofer, J., & Roth, C. (2018, November). Measuring and Bounding Experimenter Demand. *American Economic Review*, *108*(11), 3266–3302. doi: 10.1257/aer.20171330

De Sá Dias, T. C., Baby, A. R., Kaneko, T. M., & Robles Velasco, M. V. (2007). Relaxing/straightening of Afro-ethnic hair: Historical overview. *Journal of Cosmetic Dermatology*, *6*(1), 2–5. doi: 10.1111/j.1473-2165.2007.00294.x

Dossi, G. (2024). *Race and Science.*

Eberle, C. E., Sandler, D. P., Taylor, K. W., & White, A. J. (2020, July). Hair dye and chemical straightener use and breast cancer risk in a large US population of black and white women. *International Journal of Cancer*, *147*(2), 383–391. doi: 10.1002/

ijc.32738

França-Stefoni, S. A., Dario, M. F., Sá-Dias, T. C., Bedin, V., de Almeida, A. J., Baby, A. R., & Velasco, M. V. R. (2015). Protein loss in human hair from combination straightening and coloring treatments. *Journal of Cosmetic Dermatology*, *14*(3), 204–208. doi: 10.1111/jocd.12151

Hale, G., Regev, T., & Rubinstein, Y. (2023, March). *Do Looks Matter for an Academic Career in Economics?* (SSRN Scholarly Paper No. ID 3805308). Rochester, NY: Social Science Research Network.

Hester, N., & Gray, K. (2018, March). For Black men, being tall increases threat stereotyping and police stops. *Proceedings of the National Academy of Sciences*, *115*(11), 2711–2715. doi: 10.1073/pnas.1714454115

Hunt, J., Garant, J.-P., Herman, H., & Munroe, D. J. (2012, March). *Why Don't Women Patent?* (Tech. Rep. No. w17888). National Bureau of Economic Research. doi: 10.3386/w17888

James-Todd, T., Terry, M. B., Rich-Edwards, J., Deierlein, A., & Senie, R. (2011, June). Childhood Hair Product Use and Earlier Age at Menarche in a Racially Diverse Study Population: A Pilot Study. *Annals of epidemiology*, *21*(6), 461–465. doi: 10.1016/j.annepidem.2011.01.009

Johnson, A. M., Godsil, R. D., MacFarlane, J., Aronson, J., Balcetis, E., Barreto, M., ... Okonofua, J. (2017). PERCEPTION INSTITUTE. , 18.

Kahn, K. B., & Davies, P. G. (2011, July). Differentially dangerous? Phenotypic racial stereotypicality increases implicit bias among ingroup and outgroup members. *Group Processes & Intergroup Relations*, *14*(4), 569–580. doi: 10.1177/1368430210374609

Kessler, J. B., Low, C., & Sullivan, C. D. (2019). Incentivized resume rating: Eliciting employer preferences without deception. , *109*(11), 3713–3744. Retrieved 2021-04-14, from http://www.aeaweb.org/articles?id=10.1257/aer.20181714 doi: 10.1257/aer.20181714

Koval, C. Z., & Rosette, A. S. (2021, July). The Natural Hair Bias in Job Recruitment. *Social Psychological and Personality Science*, *12*(5), 741–750. doi: 10.1177/1948550620937937

LaCapria, K. (2020, June). *'Unprofessional Hairstyles for Men or Women' Versus 'Professional...*

Opie, T. R., & Phillips, K. W. (2015). Hair penalties: The negative influence of Afrocentric hair on ratings of Black women's dominance and professionalism. *Frontiers in Psychology*, *6*. doi: 10.3389/fpsyg.2015.01311

Redding, C. (2019, August). A Teacher Like Me: A Review of the Effect of Student–Teacher

Racial/Ethnic Matching on Teacher Perceptions of Students and Student Academic and Behavioral Outcomes. *Review of Educational Research*, *89*(4), 499–535. doi: 10.3102/0034654319853545

Willis, J., & Todorov, A. (2006). First Impressions: Making Up Your Mind After a 100-Ms Exposure to a Face. *Psychological Science*, *17*(7), 592–598. doi: 10.1111/j.1467-9280.2006.01750.x

Woolford, S. J., Woolford-Hunt, C. J., Sami, A., Blake, N., & Williams, D. R. (2016, July). No sweat: African American adolescent girls' opinions of hairstyle choices and physical activity. *BMC Obesity*, *3*(1), 31. doi: 10.1186/s40608-016-0111-7

Yang, Y., Tian, T. Y., Woodruff, T. K., Jones, B. F., & Uzzi, B. (2022, September). Gender-diverse teams produce more novel and higher-impact scientific ideas. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(36), e2200841119. doi: 10.1073/pnas.2200841119

A Appendix

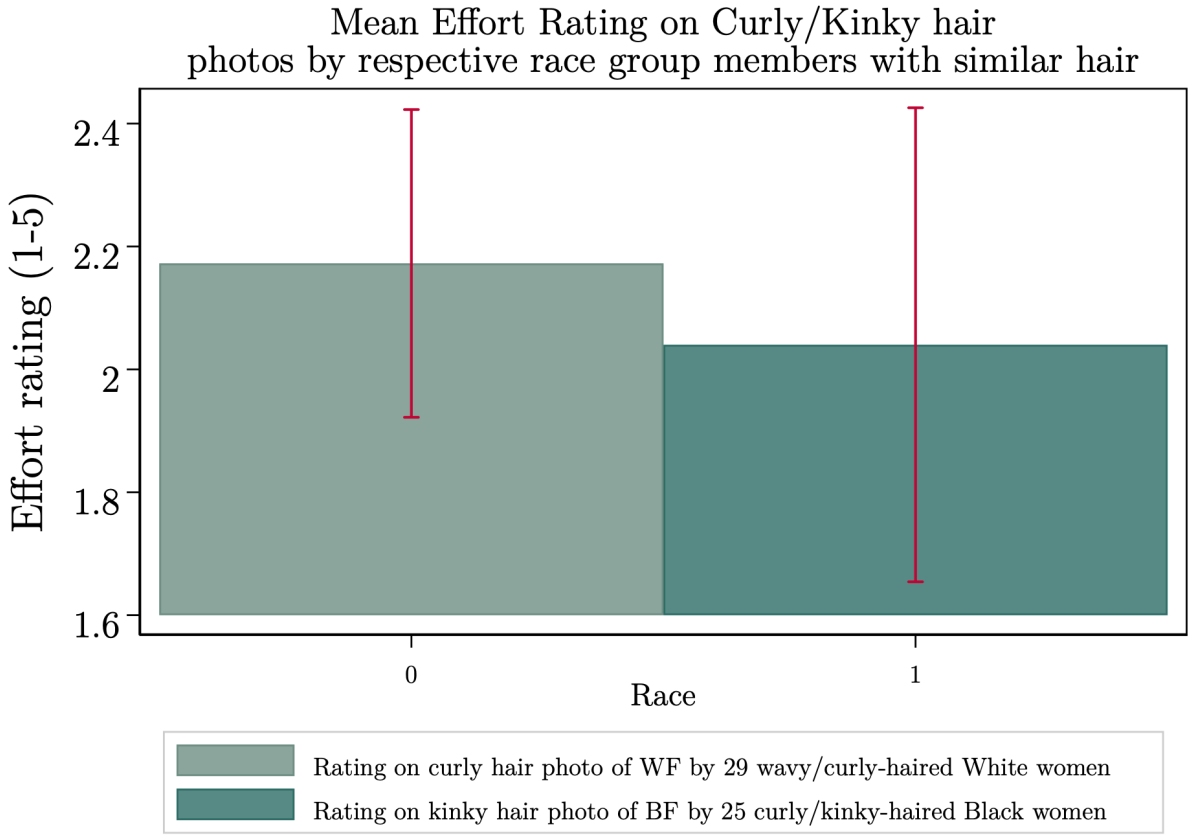
A.1 Summary statistics

Table A.1: Summary statistics

Variable	Mean	Std. Dev.	N
Student sample			
STEM major	0.671	0.47	642
Asian	0.391	0.488	481
Black	0.071	0.257	481
Native_American	0.006	0.079	481
White	0.464	0.499	481
Other	0.069	0.253	481
Female	0.714	0.452	500
Male	0.244	0.43	500
Nonbinary	0.042	0.201	500
Democrat	0.709	0.455	419
Republican	0.024	0.153	419
Independent	0.267	0.443	419
Lay adult sample			
STEM major	-	-	-
Asian	0.037	0.189	1813
Black	0.117	0.321	1813
Native_Am	0.003	0.052	1813
White	0.831	0.375	1813
Other	0.012	0.11	1813
Female	0.501	0.5	1821
Male	0.498	0.5	1821
Nonbinary	0.002	0.041	1821
Democrat	0.492	0.5	1801
Republican	0.224	0.417	1801
Independent	0.284	0.451	1801

A.2 Focus group survey findings

Figure A.1: The curly hair photo of the white female model and the kinky hair photo of the Black female model used in our experiment were rated to represent comparable levels of effort by white and Black women with similar hair textures



A.3 Subjects' self-reported approach to ratings

Figure A.2: Student Respondents' explanations about how they made photo ratings
Note that this wordcloud is made after removing filler words, rating-related metrics, and other common phrases respondents used, in particular ["professionalism professional professionalness competency competence completeness competent agreeableness agreeable agreeability based chose also "]

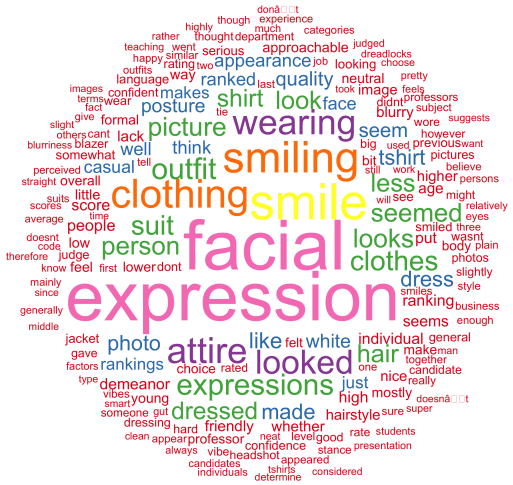


Figure A.3: Lay adult participants' explanations about how they made photo ratings



A.4 Treatment Effects

Table A.2: Treatment effects: Professionalism Ratings

	STEM students						Lay adults					
	No demand		Moderate demand		High demand		No demand		Moderate demand		High demand	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Post-Generic Treatment: Black CURLY penalty	-0.013	-0.001	0.384	0.391	0.286	0.314	0.053	0.055	0.065	0.062	0.353**	0.350*
	(0.203)	(0.202)	(0.257)	(0.253)	(0.305)	(0.300)	(0.104)	(0.104)	(0.159)	(0.159)	(0.179)	(0.179)
Generic TE: Black vs white CURLY penalty	-0.055	0.133	0.331	0.529	0.238	0.451	-0.005	0.077	0.003	0.081	0.295	0.377*
	(0.277)	(0.308)	(0.319)	(0.345)	(0.358)	(0.379)	(0.166)	(0.153)	(0.205)	(0.195)	(0.221)	(0.212)
Post-Historical Treatment: Black CURLY penalty	0.140	0.135	0.306	0.285	0.365	0.331	0.093	0.098	0.206	0.220	0.233	0.239
	(0.186)	(0.186)	(0.323)	(0.322)	(0.276)	(0.275)	(0.108)	(0.108)	(0.171)	(0.171)	(0.171)	(0.171)
Historical TE: Black vs white CURLY penalty	0.098	0.268	0.253	0.424	0.316	0.468	0.036	0.120	0.144	0.239	0.175	0.266
	(0.266)	(0.297)	(0.374)	(0.397)	(0.334)	(0.361)	(0.168)	(0.156)	(0.214)	(0.204)	(0.214)	(0.205)
Post-Generic treatment: Black LOC penalty	-0.013	-0.001	0.384	0.391	0.286	0.314	0.053	0.055	0.065	0.062	0.353**	0.350*
	(0.203)	(0.202)	(0.257)	(0.253)	(0.305)	(0.300)	(0.104)	(0.104)	(0.159)	(0.159)	(0.179)	(0.179)
Generic TE: Black vs white LOC penalty	0.085	0.007	0.126	0.110	0.068	0.057	-0.099	-0.140	0.233	0.190	0.183	0.140
	(0.303)	(0.311)	(0.331)	(0.339)	(0.394)	(0.397)	(0.165)	(0.161)	(0.219)	(0.216)	(0.225)	(0.223)
Post-Historical treatment: Black LOC penalty	-0.015	-0.005	0.509*	0.514*	0.858***	0.849***	0.429***	0.436***	0.272	0.283*	0.060	0.064
	(0.202)	(0.199)	(0.295)	(0.288)	(0.226)	(0.216)	(0.110)	(0.110)	(0.171)	(0.171)	(0.191)	(0.191)
Historical TE: Black vs white LOC penalty	-0.065	-0.176	0.449	0.355	0.801***	0.685**	0.352**	0.317**	0.191	0.162	-0.018	-0.057
	(0.290)	(0.295)	(0.359)	(0.362)	(0.306)	(0.307)	(0.163)	(0.159)	(0.208)	(0.205)	(0.226)	(0.223)
Observations	7273	7299	4832	4858	4933	4959	25191	25189	16864	16862	16854	16852

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All models contain subject FE

Table A.3: Treatment effects: Competence Ratings

	STEM students						Lay adults					
	No demand		Moderate demand		High demand		No demand		Moderate demand		High demand	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Post-Generic Treatment: Black CURLY penalty	0.112	0.156	0.240	0.303	0.215	0.299	0.010	0.013	-0.022	-0.025	0.269	0.270
	(0.170)	(0.174)	(0.234)	(0.240)	(0.205)	(0.215)	(0.108)	(0.108)	(0.135)	(0.135)	(0.169)	(0.169)
Generic TE: Black vs white CURLY penalty	0.141	0.410	0.263	0.557*	0.247	0.553*	0.034	-0.027	-0.002	-0.065	0.292	0.235
	(0.236)	(0.274)	(0.286)	(0.322)	(0.263)	(0.300)	(0.154)	(0.147)	(0.174)	(0.167)	(0.201)	(0.196)
Post-Historical Treatment: Black CURLY penalty	0.056	0.073	-0.004	0.024	0.365	0.370	0.021	0.019	0.237*	0.237*	0.080	0.079
	(0.190)	(0.191)	(0.256)	(0.255)	(0.244)	(0.242)	(0.107)	(0.107)	(0.137)	(0.136)	(0.151)	(0.151)
Historical TE: Black vs white CURLY penalty	0.086	0.327	0.019	0.278	0.398	0.624*	0.045	-0.021	0.258	0.197	0.104	0.044
	(0.252)	(0.285)	(0.304)	(0.330)	(0.292)	(0.322)	(0.153)	(0.146)	(0.176)	(0.168)	(0.186)	(0.181)
Post-Generic treatment: Black LOC penalty	0.112	0.156	0.240	0.303	0.215	0.299	0.010	0.013	-0.022	-0.025	0.269	0.270
	(0.170)	(0.174)	(0.234)	(0.240)	(0.205)	(0.215)	(0.108)	(0.108)	(0.135)	(0.135)	(0.169)	(0.169)
Generic TE: Black vs white LOC penalty	0.358	0.450	0.074	0.181	-0.104	0.035	-0.084	-0.197	0.316	0.202	0.191	0.080
	(0.254)	(0.291)	(0.309)	(0.342)	(0.318)	(0.345)	(0.148)	(0.147)	(0.206)	(0.205)	(0.195)	(0.194)
Post-Historical treatment: Black LOC penalty	-0.095	-0.086	0.046	0.069	0.681***	0.654***	0.258**	0.255**	0.247*	0.247*	-0.022	-0.028
	(0.167)	(0.167)	(0.273)	(0.271)	(0.219)	(0.216)	(0.101)	(0.101)	(0.147)	(0.146)	(0.162)	(0.162)
Historical TE: Black vs white LOC penalty	-0.164	-0.102	-0.033	0.060	0.609**	0.636**	0.287**	0.171	0.273	0.162	0.004	-0.113
	(0.245)	(0.281)	(0.326)	(0.353)	(0.281)	(0.313)	(0.145)	(0.144)	(0.180)	(0.179)	(0.192)	(0.192)
Observations	7152	7185	4717	4750	4844	4877	25170	25169	16846	16845	16832	16831

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All models contain subject FE

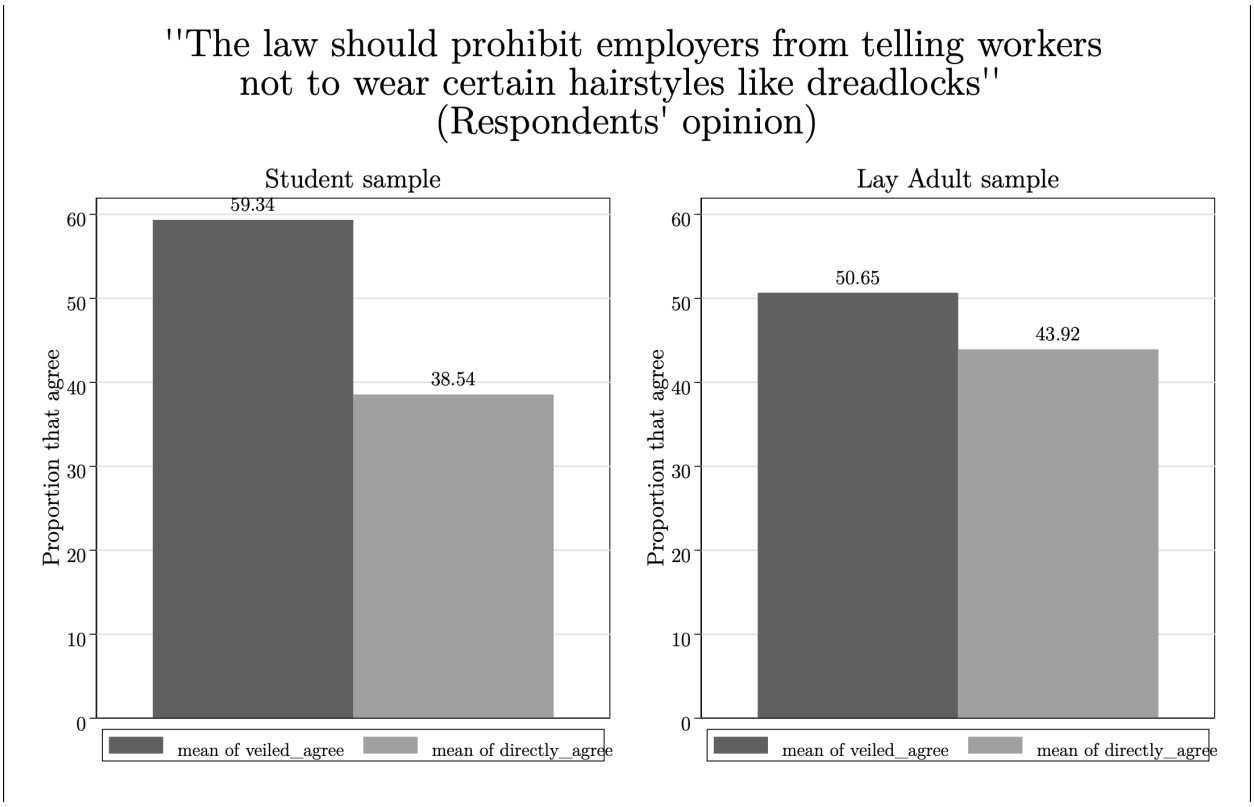
Table A.4: Treatment effects: Agreeableness Ratings

	STEM students						Lay adults					
	No demand		Moderate demand		High demand		No demand		Moderate demand		High demand	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Post-Generic Treatment: Black CURLY penalty	0.002	0.059	0.134	0.273	0.073	0.216	0.049	0.053	-0.017	-0.019	-0.120	-0.112
	(0.183)	(0.182)	(0.290)	(0.288)	(0.266)	(0.263)	(0.127)	(0.126)	(0.186)	(0.185)	(0.191)	(0.189)
Generic TE: Black vs white CURLY penalty	0.019	0.133	0.142	0.343	0.100	0.297	0.077	0.121	0.004	0.046	-0.091	-0.036
	(0.279)	(0.282)	(0.360)	(0.360)	(0.340)	(0.341)	(0.177)	(0.167)	(0.223)	(0.216)	(0.227)	(0.218)
Post-Historical Treatment: Black CURLY penalty	-0.068	-0.055	-0.349	-0.370	0.121	0.093	0.021	0.018	0.226	0.224	0.065	0.071
	(0.188)	(0.191)	(0.288)	(0.294)	(0.273)	(0.281)	(0.123)	(0.122)	(0.176)	(0.175)	(0.164)	(0.163)
Historical TE: Black vs white CURLY penalty	-0.051	0.019	-0.341	-0.300	0.148	0.174	0.048	0.087	0.247	0.288	0.094	0.147
	(0.283)	(0.288)	(0.357)	(0.364)	(0.345)	(0.354)	(0.173)	(0.165)	(0.215)	(0.207)	(0.204)	(0.196)
Post-Generic treatment: Black LOC penalty	0.002	0.059	0.134	0.273	0.073	0.216	0.049	0.053	-0.017	-0.019	-0.120	-0.112
	(0.183)	(0.182)	(0.290)	(0.288)	(0.266)	(0.263)	(0.127)	(0.126)	(0.186)	(0.185)	(0.191)	(0.189)
Generic TE: Black vs white LOC penalty	0.073	0.202	-0.502	-0.301	-0.247	-0.034	0.016	-0.042	0.345	0.288	0.124	0.072
	(0.277)	(0.306)	(0.376)	(0.395)	(0.333)	(0.358)	(0.169)	(0.166)	(0.232)	(0.229)	(0.219)	(0.216)
Post-Historical treatment: Black LOC penalty	-0.303	-0.284	-0.247	-0.218	0.856***	0.828***	0.146	0.138	0.016	0.017	0.017	0.014
	(0.205)	(0.201)	(0.316)	(0.310)	(0.239)	(0.236)	(0.120)	(0.119)	(0.185)	(0.183)	(0.180)	(0.179)
Historical TE: Black vs white LOC penalty	-0.461*	-0.364	-0.408	-0.290	0.688**	0.741**	0.237	0.173	0.104	0.048	0.106	0.050
	(0.277)	(0.306)	(0.366)	(0.385)	(0.305)	(0.333)	(0.168)	(0.164)	(0.220)	(0.216)	(0.215)	(0.212)
Observations	7238	7264	4829	4855	4923	4949	25184	25183	16868	16867	16847	16846

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All models contain subject FE

A.5 List Experiment Results

Figure A.4: Our list experiment suggests that subjects potentially underreport support for legislation like the CROWN Act, which would ban employers from instituting grooming policies that prohibit certain hairstyles



A.6 Robustness

A.6.1 Baseline ratings of ALL students (not just STEM majors)

Table A.5: All students (not just STEM) and Lay adults: Professionalism Ratings

	All students		Lay adults	
	Female	Male	Female	Male
Suit	2.358***	2.427***	2.744***	2.798***
	(0.084)	(0.085)	(0.052)	(0.052)
High resolution	0.138**	0.170***	0.195***	0.161***
	(0.059)	(0.059)	(0.030)	(0.030)
Smile	0.198***	0.011	0.379***	0.250***
	(0.048)	(0.052)	(0.030)	(0.029)
Black X Curly	-0.109	0.501***	-0.352***	-0.067
	(0.093)	(0.087)	(0.053)	(0.055)
Black X Locs	-0.063	0.254***	-0.322***	-0.224***
	(0.091)	(0.091)	(0.053)	(0.060)
Black X Straight	0.404***	0.551***	0.038	0.391***
	(0.083)	(0.090)	(0.051)	(0.052)
White X Curly	-0.143	-0.370***	-0.291***	-0.580***
	(0.094)	(0.091)	(0.050)	(0.053)
White X Locs	-1.100***	-1.044***	-1.120***	-1.119***
	(0.096)	(0.098)	(0.058)	(0.058)
White X Straight	.	.	.	.
	.	.	.	.
Constant	5.036***	4.904***	4.583***	4.514***
	(0.083)	(0.092)	(0.050)	(0.052)
Rating differentials				
Black: Curly–straight/clean-cut hair	-0.512***	-0.050	-0.390***	-0.458***
	(0.084)	(0.084)	(0.049)	(0.048)
Black: Locs–straight/clean-cut hair	-0.467***	-0.297***	-0.360***	-0.615***
	(0.077)	(0.082)	(0.049)	(0.054)
Black vs white: Curly–straight/clean-cut	-0.369***	0.320**	-0.098	0.123*
	(0.122)	(0.124)	(0.068)	(0.067)
Black vs white: Locs–straight/clean-cut	0.633***	0.747***	0.760***	0.504***
	(0.118)	(0.126)	(0.074)	(0.072)
Observations	3841	3845	12753	12942

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All models contain subject FE

Table A.6: All students (not just STEM) and Lay adults: Competence Ratings

	All students		Lay adults	
	Female	Male	Female	Male
Suit	1.205***	1.114***	1.517***	1.510***
	(0.069)	(0.065)	(0.040)	(0.040)
High resolution	0.123**	0.203***	0.221***	0.234***
	(0.051)	(0.054)	(0.027)	(0.027)
Smile	0.354***	0.103**	0.550***	0.302***
	(0.050)	(0.048)	(0.029)	(0.028)
Black X Curly	0.167*	0.668***	-0.336***	0.068
	(0.089)	(0.094)	(0.050)	(0.052)
Black X Locs	0.134	0.516***	-0.283***	-0.071
	(0.090)	(0.102)	(0.049)	(0.055)
Black X Straight	0.449***	0.721***	-0.031	0.418***
	(0.084)	(0.094)	(0.048)	(0.051)
White X Curly	0.109	-0.082	-0.169***	-0.405***
	(0.089)	(0.091)	(0.048)	(0.048)
White X Locs	-0.853***	-0.742***	-0.907***	-0.938***
	(0.090)	(0.101)	(0.054)	(0.053)
White X Straight	.	.	.	.
	.	.	.	.
Constant	5.734***	5.697***	5.536***	5.543***
	(0.083)	(0.090)	(0.044)	(0.046)
Rating differentials				
Black: Curly–straight/clean-cut hair	-0.281***	-0.053	-0.305***	-0.350***
	(0.075)	(0.078)	(0.044)	(0.044)
Black: Locs–straight/clean-cut hair	-0.315***	-0.206***	-0.252***	-0.489***
	(0.067)	(0.079)	(0.044)	(0.049)
Black vs white: Curly–straight/clean-cut	-0.390***	0.029	-0.136**	0.056
	(0.116)	(0.122)	(0.062)	(0.064)
Black vs white: Locs–straight/clean-cut	0.538***	0.536***	0.654***	0.449***
	(0.110)	(0.123)	(0.067)	(0.068)
Observations	3798	3791	12744	12929

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All models contain subject FE

Table A.7: All students (not just STEM) and Lay adults: Agreeableness Ratings

	All students		Lay adults	
	Female	Male	Female	Male
Suit	0.430***	0.276***	0.604***	0.606***
	(0.061)	(0.060)	(0.039)	(0.037)
High resolution	-0.118**	0.152***	-0.022	0.229***
	(0.051)	(0.051)	(0.031)	(0.028)
Smile	1.946***	1.689***	2.151***	1.818***
	(0.073)	(0.071)	(0.046)	(0.043)
Black X Curly	0.482***	1.556***	-0.087	0.555***
	(0.094)	(0.091)	(0.058)	(0.056)
Black X Locs	0.445***	1.452***	-0.130**	0.464***
	(0.097)	(0.105)	(0.057)	(0.061)
Black X Straight	0.603***	1.599***	0.077	0.803***
	(0.095)	(0.095)	(0.055)	(0.057)
White X Curly	0.126	0.360***	-0.127**	-0.085
	(0.094)	(0.089)	(0.055)	(0.055)
White X Locs	-0.363***	-0.159	-0.532***	-0.555***
	(0.096)	(0.101)	(0.057)	(0.057)
White X Straight	.	.	.	.
	.	.	.	.
Constant	5.054***	5.024***	5.058***	5.174***
	(0.084)	(0.092)	(0.050)	(0.052)
Rating differentials				
Black: Curly–straight/clean-cut hair	-0.121	-0.042	-0.164***	-0.248***
	(0.083)	(0.077)	(0.051)	(0.046)
Black: Locs–straight/clean-cut hair	-0.158*	-0.147*	-0.207***	-0.339***
	(0.083)	(0.080)	(0.049)	(0.051)
Black vs white: Curly–straight/clean-cut	-0.246*	-0.402***	-0.037	-0.163**
	(0.127)	(0.120)	(0.075)	(0.071)
Black vs white: Locs–straight/clean-cut	0.205	0.012	0.325***	0.216***
	(0.126)	(0.120)	(0.073)	(0.073)
Observations	3831	3831	12751	12941

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All models contain subject FE