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

Tarikua Erda† Jeffrey Shrader‡

August 24, 2026

Abstract

We use an incentive-compatible survey experiment to investigate a malleable identity marker that individuals alter to conform to dominant norms at recurring private cost: hair. Subjects rate headshots of hypothetical job candidates on professionalism, competence, and agreeableness. Relative to straightened hair, wearing one's naturally curly or kinky hair texture carries larger penalties for Black women than for White women across all metrics. Smiling partially attenuates these penalties. For Black men, racialized penalties are concentrated on agreeableness. Our findings inform ongoing legal discourse around recognizing hair texture as a protected racial characteristic and outlawing hair-based discrimination in the U.S.

JEL Codes: J15, Z10
Keywords: Norms; Implicit Bias; Discrimination; Assimilation; Identity; STEM; Innovation and R&D

*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, Alex Eble, Judd Kessler, Corinne Low, Suresh Naidu, Ebonya Washington, Jack Willis, and participants in the 2026 NBER Economics of Race and Stratification Workshop, CELSS Experimental Lunch, and the Workshop on Entrepreneurial Finance and Innovation (WEFI) group. This study was carried out under Columbia IRB (Protocol: AAAU0417). AEA-Preregistration: AEARCTR-0010183. All errors and omissions are ours.

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

"As someone with naturally curly hair, I spent decades damaging it with flat irons to conform to a mainstream, 'professional' appearance, feeling that I had little choice if I wanted to advance up the corporate ladder."

— Survey Respondent

1 Introduction

The existing literature on discrimination predominantly identifies bias through markers that evaluators treat as largely fixed, such as names or facial features (Bertrand & Mullainathan, 2004; Kahn & Davies, 2011). But many traits that shape individuals' day-to-day experiences—how they speak, dress, or wear their hair—are mutable to varying degrees. An emerging 'constructivist' perspective in economics also emphasizes that race is not simply observed, but is interpreted through such traits (Rose, 2023). Where conforming to dominant group norms carries economic returns (Akerlof & Kranton, 2000), the existing literature's focus on fixed markers misses the recurring private costs that managing such malleable traits generates. We provide direct evidence on the bias underlying these costs, focusing on hair, a trait where the perceptions and the costs of assimilation are both observable and substantial.

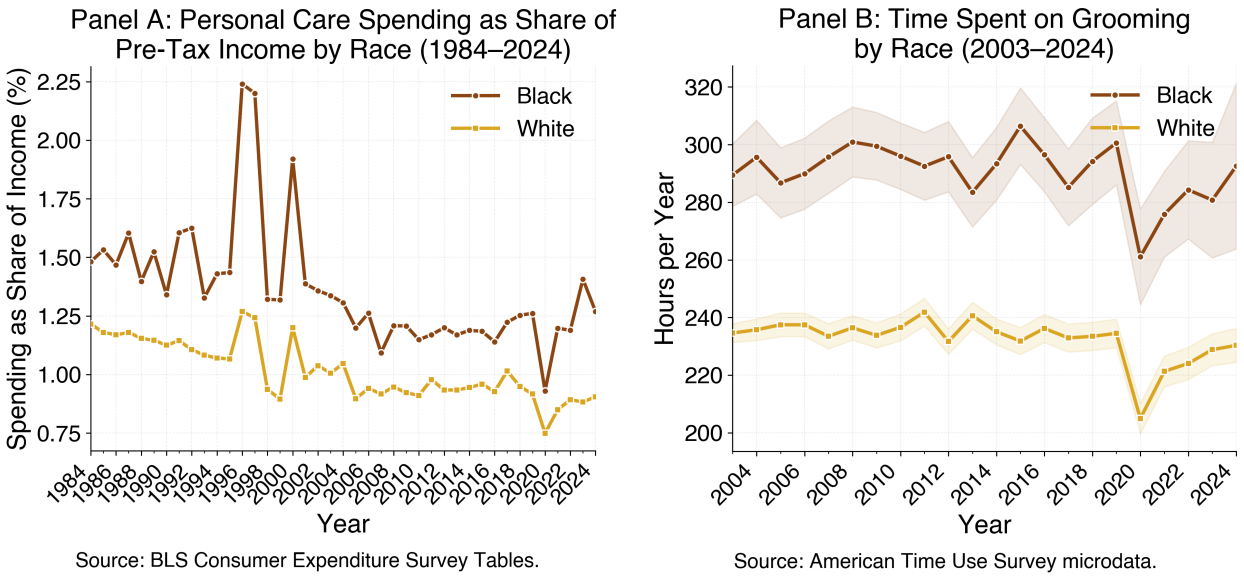
Hair *texture*—whether strands are straight, wavy, curly, or kinky—is an immutable feature, yet hair *style* can be altered using chemicals, heat, and various accessories. Bias against the hair texture of populations of African origin has roots in the Transatlantic Slave Trade era, and is one of many cases across societies and centuries where hair has been used to enforce cultural conformity or dominance.[1] Mainstream professionalism norms that favor straightened hair impose private costs along several margins: Black Americans spend a 34% larger share of their income, and dedicate roughly 57 more hours per year on personal grooming relative to their White counterparts (see Figure 1). About 80% of Black women report changing their hair from its natural kinky state to fit in at the office (JOY Collective, 2019). Hair straightening also carries health risks, including cancer-related ones. These costs primarily fall on women, but young Black men also face hair-related penalties, including suspensions from school or sports for wearing locs (also known as dreadlocks).[2]

We conduct incentive-compatible, pre-registered experiments to assess appearance norms in academia, where the role of appearance is ex-ante ambiguous. In principle, given that academia prides itself on meritocracy, appearance may not matter. Yet, attractiveness pre-

[1] European slave traders shaved the heads of enslaved Africans upon arrival in the Americas (Byrd & Tharps, 2002). US boarding schools forcibly cut Native American children's hair as part of assimilation campaigns (Adams, 1995). Nazi concentration camps shaved Jewish prisoners' heads (Levi, 1958).

[2] See APNews (2023) and Reuters (2024).

Figure 1: Large, persistent Black-White gaps in money and time spent on personal care



In Panel (A), "White" denotes the BLS category "White and Other" prior to 2003 and "White" thereafter. Panel (B) plots annual hours spent on washing, dressing, and grooming oneself (ATUS activity code 010201) by race, using ATUS microdata for $N \approx 203{,}000$ non-Hispanic White and Black respondents. Shaded bands in Panel (B) represent 95% confidence intervals around survey-weighted means.

dicts success even in ostensibly merit-based contexts like economics faculty careers (Hale et al., 2023) and entrepreneurial pitch competitions (Brooks et al., 2014). Hair-based bias documented in industry[3] may thus extend to academia, but it is still unclear whether 'unconventional' hair would be penalized or rewarded as a signal of competence (for instance, it might invoke the mad scientist trope (Chambers, 1983)). We give special attention to STEM fields, which carry high stakes for both efficiency and equity: STEM academics train the innovation and R&D workforce that drives productivity growth and also command large wage premia, so discrimination that deters talented Black individuals from entering or advancing can distort aggregate growth and worsen racial income gaps.[4]

We survey university students and a representative sample of US adults, asking each respondent to rate 32 headshots of hypothetical professor candidates on professionalism, competence, and agreeableness. Students are embedded in the academic norms they are evaluating, while lay adults provide a benchmark for whether our findings are specific to academia or reflect broader societal attitudes about norms in STEM academia. The headshots feature four models—a Black female, a White female, a Black male, and a White

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

[4]Hsieh et al. (2019) estimate that declining barriers faced by women and Black Americans account for a significant share of U.S. productivity growth since 1960, while Cook (2014) and Cook and Yang (2018) document innovation losses from the historical exclusion of Black Americans.

male—generated through an original photoshoot and photo-editing, and depict three hair looks: natural curly/kinky hair, locs, and straightened hair (clean-cut for men), with additional variation in attire, facial expression, and photo resolution to obscure our central focus on hair.[5]

Our design incorporates several features to ensure both clean identification and honest responses. First, we employ higher-order belief elicitation: respondents in each sample were offered one of three bonus $100 cash prizes if they could most accurately predict how the university staff we had previously surveyed under incentives rated the same photos, making it safer to express potentially stigmatized views. Second, before they start rating, we show respondents casual, weekend photos of the models to establish that all models have naturally curly/kinky hair, so that subsequent hairstyle variation is understood as resulting from styling choices relative to a shared baseline texture. Third, we include non-experimental photos of other-race models to reduce fixation on Black-White comparisons. Fourth, our main estimates come from regressions that exploit within-model variation in hairstyle, comparing how the same model is rated when wearing their natural curly texture or locs relative to straightened hair for women and clean-cut hair for men.

We find pronounced penalties in perceived professionalism, competence, and agreeableness for the Black female model when she wears her naturally kinky hair instead of straightened hair. The White female model receives smaller, if any, penalties for wearing her natural curly hair relative to straightening it, suggesting that there is a racialized nature to the bias against curly/kinky textures. In the student sample, these racialized penalties reduce the Black female model's ratings across all elicited outcomes by about 5% of the baseline average rating.[6] These relative effects are about twice as large as the penalty for a blurry, low-resolution photograph. Lay adults show a similar pattern, though the racialized gap is attenuated. That students show larger penalties suggests that proximity to the ivory tower, if anything, sharpens appearance-based bias.

For the Black male model, the excess penalty is concentrated on agreeableness: students award the White male a significant *premium* for curly hair while penalizing the Black male, consistent with natural Black hair on men activating priors associated with perceived threat (Hester & Gray, 2018). On professionalism and competence, the pattern reverses, with the

[5]The natural hair texture looks of the models were rated to represent comparable levels of effort in prior focus group surveys of lay adults with matching (self-reported) race and hair texture.

[6]Given that we elicited subject views on three different outcomes related to professionalism, competence, and agreeableness, any of which could be indicative of appearance-based penalties, we conducted multiple hypothesis test corrections. In the student sample, the professionalism and competence results survive these corrections, while the agreeableness result becomes marginal ($p = 0.074$). See the main results for further discussion.

White male penalized more for deviating from clean-cut hair. Evaluators appear to map the same deviation from straight/clean-cut hair norms onto different penalties depending on race and gender—less polished and competent for Black women, less agreeable for Black men—highlighting the intersectional nature of the bias.

We also find a statistical discrimination channel behind the agreeableness penalties. The racialized agreeableness penalty is large when Black models wear a neutral expression but, for the Black female model, vanishes when she smiles, consistent with evaluators updating unfavorable priors in response to a warmth signal (Phelps, 1972). One source of these priors may be historical: during the Civil Rights era of the 1960s, the "Black is Beautiful" movement encouraged Black Americans to embrace their natural hair as an act of self-acceptance, but wearing an Afro was widely perceived by mainstream society as a political statement and a sign of militancy (Byrd & Tharps, 2002), an association that experimental work in psychology finds persists today in the form of higher perceived dominance attributed to Afrocentric hairstyles (Opie & Phillips, 2015).

When it comes to locs, the racial gap in ratings is reversed as both student and lay adult respondents severely penalize White models, especially on professionalism and competence, with many citing 'cultural appropriation.' Despite these stated beliefs about locs as a hairstyle of particular cultural significance to Black people, respondents still penalize Black models in absolute terms.

Finally, a primary concern that we wanted to guard against relates to social desirability bias. We designed the incentives in the experiment to mitigate this concern, and we took two additional measures to assess whether such bias remained. We randomized exposure to induced experimenter demand per de Quidt et al. (2018) and to two information interventions: a vignette on workplace authenticity and productivity, and an explainer on the history of hair-based racial discrimination. Neither treatment significantly affected our primary results on racialized penalties for natural hair styles, suggesting that the experiment did isolate honest estimates. The fact that information on historical hair bias did not change responses also indicates that these penalties are not solely driven by ignorance.

Our findings are relevant to a growing set of legislative efforts in the US to outlaw hair discrimination: the CROWN ("Creating a Respectful and Open World for Natural Hair") Act has been adopted in more than 25 US states since 2019.[7]

While our study uses hair as an unusually clean, tractable setting, our findings gener-

[7] In the UK, the Equality Act of 2010 prohibits race discrimination, under which hair texture and hairstyles associated with racial or ethnic identity may be protected in practice. In France, the National Assembly approved at first reading a bill to prohibit hair-based discrimination in workplaces in March 2024.

alize to a broader class of malleable traits that individuals routinely alter to access economic opportunity. Millions in South Asia and Africa bleach their skin because lighter skin carries a premium in labor and marriage markets (Glenn, 2008); racial and regional speech patterns carry wage penalties whereas code-switching by minority workers can carry rewards (Grogger, 2011; Grogger et al., 2020); and in South Korea, where the vast majority of employers require a photograph on job applications and half of hiring managers report rejecting candidates based on appearance, millions undergo cosmetic surgery—called *chwieop seonghyeong* ("employment surgery")—with many procedures requiring periodic maintenance over subsequent years.[8] Our study thus sheds new light on the unique double-bind associated with such malleable traits: where individuals must either deviate from mainstream norms and face penalties or conform and bear ongoing, private costs.

Related Literature: This paper makes several key contributions. First, we add to the study of discrimination by shifting attention to traits that are malleable enough that nonconformity is treated as a choice and penalized accordingly. Prior work documents discrimination on the basis of names (Bertrand & Mullainathan, 2004; Kline et al., 2022) and phenotypic facial features (Kahn & Davies, 2011)—characteristics that are costly, complicated, or irreversible to alter. A related literature documents returns to appearance attributes such as beauty (Hamermesh & Biddle, 1994), height (Persico et al., 2004), weight (Cawley, 2004), and dental health (Glied & Neidell, 2008; Gallego et al., 2024), as well as the costs of adjusting identity signals through surname changes (Arai & Skogman Thoursie, 2009) or resume whitening (Kang et al., 2016). Hair falls between these two literatures: it is a phenotypic marker of race, but one that can be (and often is) altered routinely and reversibly. The relevant cost is therefore not a one-time barrier at the point of hiring but a continuous tax on daily life which can compound over a career along financial, temporal, health—and as our smile attenuation results suggest—emotional margins. Our paper thus draws attention to a dimension of discrimination that the existing literature does not capture—the private costs that minorities bear to participate in institutions governed by dominant group norms—thus potentially understating the welfare toll of discrimination.

As such, our focus on a malleable racial cue connects to an emerging constructivist perspective that views race not as fixed in data but as interpreted through observable char-

[8]South Korea's beauty standards trace partly to a post-colonial project of reconstructing national identity after Japanese occupation (Holliday & Elfving-Hwang, 2012). In a 2017 survey by a job portal, 93% of firms required a photo on applications and nearly half of HR managers reported they had rejected candidates based on appearance (Saramin, 2017); a 2019 amendment to Korea's Fair Hiring Procedure Act restricted appearance-related information requests but exempts identification photographs, and resume photographs remain widespread. A 2024 survey found that 66% of Gen Z job seekers had undergone beauty treatments for employment purposes (Catch, 2024).

acteristics like skin tone, speech, and hair (Rose, 2023). While we do not study whether hairstyle alters racial categorization per se, our results show that varying a single cue within a recognized racial category is sufficient to generate large shifts in evaluation. This suggests that the markers through which race is perceived can carry independent evaluative weight, including when it comes to bias. Additionally, individuals may alter observable behaviors or appearance, incurring substantial costs in doing so, to manage how they are perceived by others (Bursztyn & Jensen, 2017). We demonstrate evidence on a setting in which an identity-relevant cue is not only interpreted by evaluators but can itself be manipulated in anticipation of that evaluation at a cost, part of which we quantify.

Second, we broaden the scope of identity-specific costs documented in prior work. Existing estimates of such costs—most notably, gender-based price discrimination (i.e., the pink tax) (de Blasio & Menin, 2015)—have centered on pecuniary margins. That smiling, a costly signal of agreeableness, attenuates the racialized hair penalty for Black models is suggestive of non-pecuniary costs, a sort of warmth-tax. Additionally, consistent with recent empirical work showing that labor market penalties associated with multiple marginalized identities interact in non-additive ways (Paul et al., 2022), our use of an explicitly intersectional approach (Crenshaw, 1989) reveals that hair penalties fall disproportionately on Black women, and cannot be recovered by examining race or gender in isolation.

Third, this study advances the methodology of studying unconscious bias and contributes new evidence on the costs and mechanisms of hair bias, and if/how it can be mitigated. Prior work in adjacent social science disciplines has used experimental approaches to document hair bias in settings such as advertising and consulting (Koval & Rosette, 2021) and corporate employment (Opie & Phillips, 2015). We build on this foundation by introducing several innovations: our experiment is the first to use pre-registered, incentive-compatible designs, to examine hair bias as it applies to both women and men, and to explicitly account for social desirability bias and potential interventions to mitigate bias. Our experimental photos are designed to achieve high realism and precisely controlled variation as well. Rather than using methods that reveal the bias under study to subjects—such as the Implicit Association Test (IAT)—our approach obfuscates the survey's central focus on hair through a holistic headshot rating exercise, offering a more reliable measure of unconscious bias.

Finally, our closer focus on STEM academia contributes to the literature on the underrepresentation of racial and gender minorities. Self-selection is commonly posited as one explanation for representation gaps, yet growing evidence is showing the potential role of discriminatory or unwelcoming environments, including stereotype threat and discriminatory evaluation (Steele et al., 2002; Francis & Darity, 2020), discrimination in professional

networks (Ajzenman et al., 2025), and the absence of senior faculty mentors who could serve as role models for minority students (Bettinger & Long, 2005; Redding, 2019). That natural curly hair earns the White female model a competence premium while penalizing the Black female model illustrates the kind of double standard that, compounded over a career, may erode the sense of belonging this literature identifies as critical to retention.

The rest of the paper proceeds as follows: Section 2 provides background on hair bias. Section 3 describes the experiment. Section 4 describes the data and estimation strategy. Section 5 shows the main results and robustness checks. Section 6 discusses implications for current debates around discrimination protections. Section 7 concludes.

2 Background

2.1 Historical Overview of Hair Bias

For centuries, hair has served as a marker of various aspects of cultural and social identity such as age, marital status, social status and so on across societies, including those in Africa (Byrd & Tharps, 2002). Since the onset of the Transatlantic Slave Trade, however, systemic racism has affected, and in part dictated, how Black people in the US wear and manage their hair. Europeans shaved the heads of enslaved Africans upon arrival to the Americas to assert domination and reallocate time toward labor (Chapman, 2007; Dash, 2006; Morrow, 1973). In 1786, the Tignon Law required *all* women of color—including free and emancipated ones—in New Orleans to wear a scarf over their hair as a visible sign of belonging to the 'slave class.' The targeting of hair for purposes of cultural subordination is not unique to the Black American experience, but has been practiced in several other contexts against minority and persecuted groups including Native American and Chinese minorities in the U.S. and Jewish prisoners in Nazi concentration camps.[9] Hair is visible, personal, and alterable, thus often becoming a lever for institutions seeking to suppress group identity or assert control.

[9] As part of broader assimilation campaigns, US federal boarding schools banned native languages and replaced cultural clothing with uniforms from the 1870s onward (Adams, 1995). The Chinese queue–a single braid worn with a shaved forehead–was itself a product of forced assimilation: the Qing dynasty's 1645 *Tifayifu* edict required all Han Chinese men to adopt the Manchu hairstyle on pain of death (Godley, 1994). By the time Chinese immigrants arrived in the US, the queue had become a marker of cultural identity. San Francisco's Queue Ordinance of 1876 exploited this by mandating that prisoners' hair be cut to one inch—ostensibly for hygiene, but in practice targeting Chinese men for whom losing the queue meant both disgrace and potential punishment upon return to China. Nazi concentration camps also shaved prisoners' heads upon entry (Levi, 1958).

Figure 2: A 2000 advertisement for a hair relaxer asks whether the pictured Black woman got her job because of her resume or her straightened hair

Source: *Ebony Magazine*, 56, no. 1 (November 2000), p. 161.

Figure 3: A 1928 advertisement for chemical permanent waves promises a "natural" wave, not "unbecoming" kinky results

Source: *Princeton Herald*, Volume 5, Number 17, 24 February 1928.

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. Advertisers also played into this link between hair straightening and professional and labor market success, as demonstrated by the Raveen hair relaxer ad in Figure 2.

This is not to say that mainstream norms have *only* excluded Black individuals, or that they have *always* disfavored curls. Historical and contemporary accounts document similar pressures felt by other ethnic-minority women, such as Jewish women, Latina women, and South Asians, to straighten hair as a means of assimilating to dominant appearance norms.[10] White individuals have also widely adopted chemically-altered waves or loose curls through a process called permanent waving (essentially the opposite of hair relaxing) that was popular in the early 1900s, with a specific focus on the 'right' types of curls. At the time, advertisers and fashion writers took great care to distinguish for their White consumer base the desired outcomes from permanent waving as 'natural' waves that would stay truer to Caucasian phenotypes rather than tightly coiled "kinky" hair that they labeled as "unbecoming" (see Figure 3).[11] Thus, while trends in hairstyling have come and gone—with substantial grooming burden implications for women of various racial backgrounds—these examples suggest a racialized hierarchy of hair texture in which tightly coiled or "kinky" hair was particularly stigmatized, even in beauty advertising aimed at women seeking to make naturally straight hair curlier. This historical texture hierarchy motivates our focus on Black–White relative differences in the penalties associated with curly hair today.

Later, during the Civil Rights and Black Power movements of the 1960s and 1970s, more Black people embraced the act of simply wearing their natural hair texture as an Afro as an act of self-acceptance and racial pride (Byrd & Tharps, 2002; Kuumba & Ajanaku, 1998). Locs, which were associated in the modern era with Rastafari, and later, with broader Black liberation movements, similarly came to signify Black identity and collective resistance (Kuumba & Ajanaku, 1998; Lemi & Brown, 2019). Yet these styles continued to face institutional penalties, as employers and schools have at times explicitly banned locs and other natural or protective Black hairstyles, and courts historically provided limited protection

[10] See for example (Meyers, 2021; Pérez-Rosario, 2018; Goel et al., 2021)

[11] In a strikingly similar sentiment, when discussing the importance of doing a strand test to preview how permanent waves would turn out before full application, a staff writer for a popular 1920s fashion magazine writes: "A little girl I know emerged from a $5.00 permanent with a head that resembled that of a native of the Congo. Fortunately, she was supremely satisfied and just child enough for its tousled outline not to prove unbecoming. But the tragedy it might have been if it had chanced to alight on an older and more serious head as it might very easily have done! And that is why I strongly urge a preliminary test." (Source: *Fashion Service Magazine*, January 1928.)

against such grooming rules (e.g., see Robinson and Robinson (2020)). In part in response to this history, the CROWN Act was introduced in 2019, explicitly recognizing that restrictions against Afros, braids, twists, and locs can disproportionately burden Black workers and applicants, and extending race-discrimination protections to hair texture and protective hairstyles.[12] Today, the CROWN Act has been enacted in over 25 U.S. states.

2.2 Economic Implications of Hair Bias

Hair bias carries substantial costs along several margins. Using the Consumer Expenditure Survey—which tracks total personal care products and services—we show in Figure 1 (Panel A) that, over the past four decades (1984–2024), Black households have devoted roughly 1.34 times as much of their income to personal care as White households. This ratio has remained stable across business cycles, recessions, and shifting consumer trends. In the most recent year (2024), the gap stands at about 0.53 percentage points—Black households spend 1.41 percent of income on personal care versus 0.88 percent for White households—one of the widest gaps in the post-2000 period. This gap is doubly regressive because Black households both spend more in absolute dollars while earning less than Whites.

The time burden is also notable. Panel (B) of Figure 1 plots annual hours spent on grooming by race using ATUS microdata from 2003 to 2024. Black respondents spend roughly 290 hours per year on washing, dressing, and grooming, compared to about 230 for White respondents—a gap of 57.2 hours, or more than seven full working days per year. This gap has persisted across two decades, narrowing briefly during the COVID-19 pandemic when both groups reduced grooming time.

Both the CES and ATUS measures capture personal care broadly rather than hair care alone. Part of the racial gap may reflect a wider pattern: stereotypes associating Blacks with dirtiness and poor hygiene have deep roots in American history, propagated through advertising, medical discourse, and segregation-era sanitation policies (McClintock, 1995; Zimring, 2015), and have put pressure on Black Americans to invest more in personal presentation across categories. But regression-adjusted analysis of CES Interview Survey microdata from over 17,130 consumer units confirms that hair is a central driver of the spending gap: controlling for income, age, household size, education, region, and urban versus rural status, Black households spend $146–$173 more per year on personal care than observably similar White households, with the difference driven almost entirely by salon and barbershop services (see Figure C.10).

[12] "Protective" hairstyles refer to styles such as braids, twists, and locs that reduce frequent manipulation of tightly curled hair and can help limit friction, moisture loss, and breakage.

Beyond the time and monetary costs, chemical hair straighteners (called "relaxers") have been linked to several toxic, and even carcinogenic, outcomes (Eberle et al., 2020; França-Stefoni et al., 2015; De Sá Dias et al., 2007; James-Todd et al., 2011; Woolford et al., 2016). Thermal straightening can also cause heat damage that permanently alters one's curl pattern and results in hair loss (Callender et al., 2004). Additionally, as straightened hairstyles do not last long under moisture, one in three Black women also skips exercise to preserve their hairstyle compared to one in ten White women (see Johnson et al. (2017)). Therefore, mainstream norms of professionalism impose a multifaceted burden on Black Americans, with ongoing costs that compound over a lifetime.

3 Experimental Design

3.1 Experimental Visual Stimuli

Since our study is centered on appearance, the visual stimuli in our survey are a key element. We conducted an original photoshoot with four main models: a Black female, a White female, a Black male, and a White male. Models were photographed 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 the models in different hairstyles. Blurred versions of the photos were also created 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 a hairstyle, which we accounted for by running smaller online focus group surveys related to the perception of effort. This is because individuals are more likely to be familiar with their own hair texture and the effort involved with certain hair styles in their own gender and racial demographic group, compared to the hair texture and effort of other demographic groups. Thus, to account for the role of effort, we had White and Black lay adults who (self-reportedly) have curly and kinky hair rate 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 (1) time spent on the pictured hair, (2) 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 survey responses indicated that the White and Black female

models' hair signalled comparable levels of effort (see Appendix Figure C.11).

Since men typically face a lower grooming burden and are far less likely to manipulate their hair texture (e.g., straighten their 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. As such, the non-dreadlock hairstyle variations for the male models in our study are based on whether the model has big and long curly/kinky-Afro hair or shorter hair that looks clean-cut/polished, suggesting a recent visit to a professional barber.

In addition to hair style (3 categories), model race (2 categories), and model gender (2 categories), we also had 2 clothing variations (suit and t-shirt), 2 facial expressions (smiling and non-smiling) and 2 photo-resolution levels (high-resolution and blurry). To increase power to identify our primary race-gender-hair interactions, we used 84 out of the 96 total photos by excluding one of the blurry versions of each race-gender-hair by other attribute combinations.[13] We used these 84 photos to form choice sets through a bounded incomplete block design that allows us to use 16 choices per experimental subject to identify preferences over attribute interactions up to 3rd order.

As an example of the visual stimuli shown to respondents, Figure 4 presents the head-shots of the Black (top) and White (bottom) female models wearing a suit and smiling in all photos, with the variations showing their high-resolution curly hair, locs, straight hair, and low-resolution (intentionally-blurred) straight hair photos from left to right. Respondents were able to see the full face of the models they were rating, but we have redacted the eyes of models in this figure to protect their anonymity.

To help mitigate potential experimenter demand effects, 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) were similar to our experimental photos. These photos appeared to the experimental subjects just like all of the main visual stimuli, and every subject saw the same set of these additional photos. The ratings for the photos are not used in the empirical analysis.

3.2 Survey Procedure

University students were recruited from the Online Recruitment System for Economic Experiments (ORSEE) subject pool while lay adults were recruited through CloudResearch.

[13]The dropped blurry photos were chosen to have non-overlapping attributes so that we can still identify the effect of blurriness, just not the effect of blurriness alongside all interactions.

Figure 4: Examples of visual stimuli in our survey

Note: The eyes of models have been 'redacted' in this rendition to protect their privacy, but survey respondents were able to see the face of the models in their rating activity.

To avoid bias that might arise from self-selection into participation, our survey recruitment note was intentionally kept highly general as a study about "individual perceptions." The recruitment text for university students is presented in Figure A.1 in the Appendix.

The main survey activity was rating 32 headshots in four blocks of 8 photos each. The full instructions for the rating activity, and an example rating question are presented in Figures A.2 and A.3 in the Appendix. Student participants rated headshots imagining that the pictured models were candidates for a university professor position in the field of their major, and 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. The survey flow is presented in Figure 5 and is explained in chronological order below.

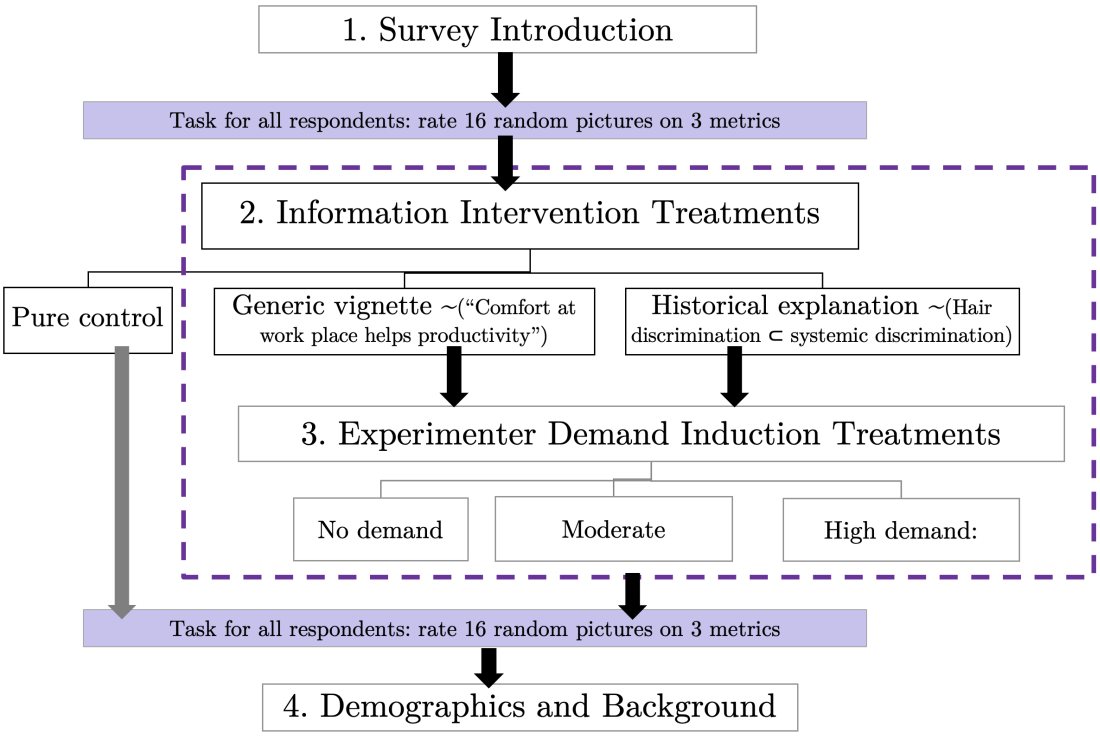
Prior to the start of the rating activity, we presented subjects with initial photos to acquaint them with the 8 models (four main Black and White models and four additional models) whose photos they would subsequently rate. We used photos that showed the models on what we described to subjects as "a casual weekend morning." This was intended to give subjects 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 affect their photo rankings. 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.

To ensure incentive compatibility, we then introduced the headshot rating activity and rating metrics and gave respondents an explanation of the bonus cash prize we were offering for accurate responses. The bonus description to student respondents is presented in Figure A.4 as an example. In it, we explained that we had already conducted a similar headshot rating exercise—also with an incentivized approach that encouraged truthful responses—on university staff who were in a position to train and advise graduate students preparing to become educators. This was indeed true: we had conducted a headshot-rating exercise with staff from a New York-area higher education institution prior to the launch of the main surveys on students and lay adults. As a reward for the participation of the university staff in our survey, we contributed money to a professional headshot session held for students of that institution and also shared aggregated insight from the survey responses as input for the session.[14] The fact that we layered multiple incentive-compatible designs *and* clarified this to students/lay adults—our main survey subjects of interest—helps allay concerns those

[14]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.

subjects might report biased higher-order beliefs because they expect university staff might have *themselves* self-censored due to social desirability bias.

Figure 5: Survey Flow



The subjects were then asked to rate photos. The rating activity involved three metrics: professionalism, competence, and agreeableness. We defined professionalism for them as the state of being in line with the norms of the university as a workplace; competence as being qualified, credible, and capable of rising to the demands of an academic job (teaching and doing research); and agreeableness as the quality of being collegial 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 on a 1–10 Likert scale, we also provided a non-response option ("N/A") in case they wanted to skip a rating.[15] Figure C.12, which presents a word cloud of the open-ended responses that respondents gave for 2 randomly chosen photos where we asked them why they rated the photo the way they did, shows that they were primarily focused on facial expression and clothing, which reassures us that respondents were not fixated on hair and gives initial evidence against experimenter demand effects.

We included interventions in the midpoint of the survey to further test for experi-

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

menter demand or social desirability biases. In particular, 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 employee productivity, and second information treatment that explains the history of hair bias in the US and the CROWN Act. The first treatment was a generic vignette reading, "Recent research shows that when employees are allowed to be fully authentic—including in their appearance and personal style—at the workplace, they experience positive outcomes in their well-being and productivity, thus, being better performers for their employer." The second one was more specifically focused on systemic discrimination and explained history and costs of hair bias tracing back to the Trans-Atlantic Slave Trade, and ended with a brief mention about the CROWN Act. For further details on how these treatments appeared, see Figures A.5 and A.6.

Subjects who were not in the pure control group were further assigned to one of three groups for an experimenter demand treatment following de Quidt et al. (2018). Half (50%) of subjects received no further experimenter demand induction; the remaining half were equally split into a moderate or high experimenter demand treatment. The moderate experimenter demand treatment read "We expect that participants who saw the previous message will rate curly/coily hairstyles or dreadlocks less negatively than they normally would" while the strong demand prompt read "Now that you have seen this message, you would be doing us a favor if you rated curly/coily hairstyles or dreadlocks as more professional, competent, agreeable than previously."

The final section of the survey asked respondents basic questions about their demographics, socioeconomic background, experience with dress-code norms in their own work experience, and views about higher education.

4 Data and Empirical Strategy

4.1 Sample Descriptions

As Table 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 sizable share of our respondents

attend a women's college. The student sample, which was recruited through 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 substantially higher than the US population's racial distribution. Roughly half of lay adult respondents reported being Democrats, and 24% 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.

Table 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			
Asian	0.049	0.216	2,412
Black	0.127	0.333	2,412
Native American	0.013	0.114	2,412
White	0.830	0.376	2,412
Other	0.012	0.111	2,412
Female	0.501	0.500	2,418
Male	0.497	0.500	2,418
Nonbinary	0.002	0.050	2,418
Democrat	0.490	0.500	2,390
Republican	0.235	0.424	2,390
Independent	0.275	0.447	2,390

Note: Observation counts change due to item non-response.

4.2 Estimation Strategy

Using the data from the first half of the ratings (before experimenter demand treatments), we estimate the following specification.

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

V_{ij} denotes 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, locs, and straightened (and for the male models, clean-cut hair), respectively. The other variables in the regression are controls: we include subject-level fixed effects α_i, which helps us absorb persistent respondent-level heterogeneity in how subjects use the rating scale across their multiple ratings, increasing precision relative to a purely between-subject design (List, 2025). We also include controls for other photo attributes (represented by X): namely, an indicator for whether the model is wearing a suit jacket, an indicator for whether the photo is in high resolution, and an indicator for whether the model is smiling.

Because the regression equation is saturated, the main results of interest—the within-race and gender ratings for curly/kinky or loc hair versus straight hair—are estimated based on within-model variation. The main hypotheses we will test are if $\beta_{12}-\beta_{14}=0$ and $\beta_{13}-\beta_{14}=0$ to understand whether the Black female model is rated the same for curly or loc hairstyles as for straight. We also test similar hypotheses for the male models, though prior literature suggests that penalties will be stronger or more pronounced for Black women. If the penalty a Black model faces for a given haircut (relative to straightened hair for women, and to clean-cut hair for men) is higher than what the White model faces, this difference-in-differences across models of the same gender but different races would suggest a 'racialized' hair penalty.

5 Results

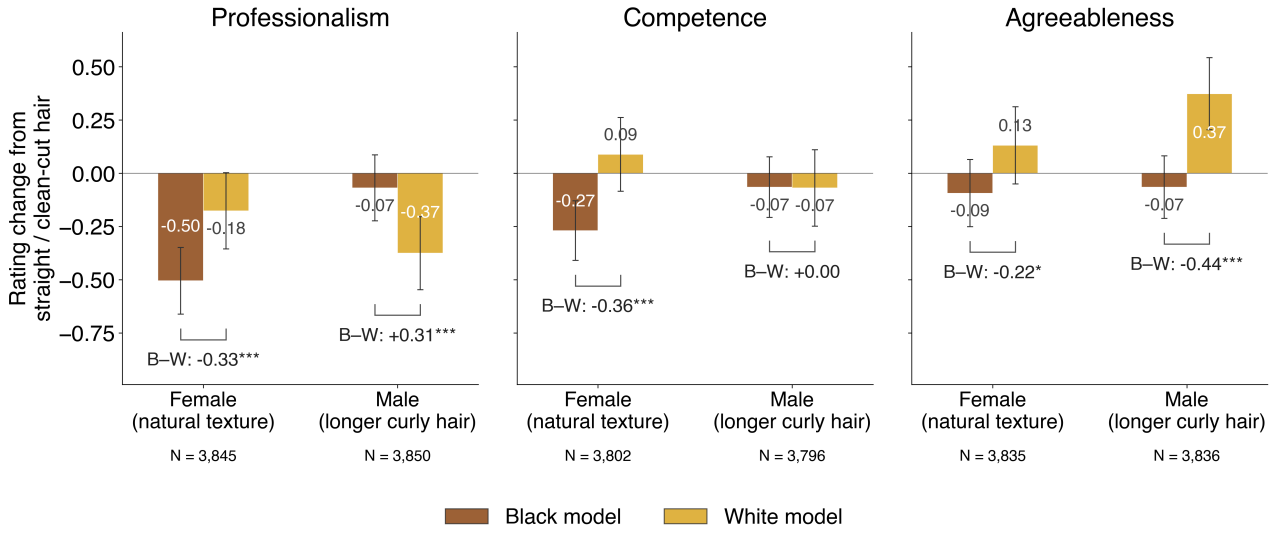
In this section, we report the findings from the primary sample, university students. The findings from the lay adult sample show broadly similar patterns with attenuated racialized gaps and are reported in Appendix Section B. For readability, the main text presents results graphically as within-model rating differentials across races; full regression tables with

detailed coefficients for all photo attributes are in Appendix Tables C.1–C.3.[16]

5.1 Baseline Rating differentials for curly/kinky hair texture

Figure 6 presents the racialized hair penalties for natural curly/kinky hair relative to straightened (or clean-cut) hair in the student sample, for both female and male models based on the first 16 ratings (i.e., the first half of the survey). Each bar shows the deviation from the model's own straight-hair predicted rating, derived from Equation (1) evaluated at control-variable means. The B–W gap reported beneath each pair is the difference-in-differences: the Black model's curly-versus-straight penalty minus the White model's.

Figure 6: Racialized penalties for natural curly/kinky hair: Student sample



Note: Bars show deviations from each model's own straight-hair (or for males, clean-cut hair) predicted rating (set to zero), based on Equation (1) evaluated at control-variable means. Error bars are 95% CIs. Brackets report the "B–W gap" (difference-in-differences: Black model's hair penalty minus White model's), from the difference between the Black and White models' hairstyle coefficients. All regressions include subject fixed effects, model fixed effects, and controls for clothing, facial expression, and photo resolution. Observation counts differ across outcomes because of item non-response. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The asymmetry in how natural hair affects ratings of the Black versus White female models is stark. Relative to straightened hair, wearing naturally kinky hair reduces the Black female model's professionalism rating by 0.5 points and her competence rating by 0.27 points on a 10-point scale. The White female model's corresponding penalties are much smaller—0.18 points for professionalism, and only marginally significant—and on competence, she

[16]Since the regression equation is saturated, we estimate and report those detailed results in the numerically equivalent and digestible parameterization with model fixed effects and model-specific hairstyle indicators. Thus, in Tables C.1–C.3, every hairstyle coefficient compares the ratings the same model gets when wearing curly/kinky hair or locs to their own straightened (clean-cut for men) photo baseline.

actually receives a statistically insignificant premium of 0.11 points when wearing curly hair. This suggests, students rate the Black female as less competent for wearing her natural hair texture, while viewing the White female model, if anything, as *more* competent for wearing hers. The racialized curly hair gaps are −0.33 for professionalism and −0.36 for competence. To a limited extent, a similar pattern occurs with the agreeableness ratings, where the White female model receives an imprecisely estimated but positive premium when wearing her curly hair, while the Black model receives an imprecisely estimated penalty. The racialized gap for agreeableness ratings is −0.22, significant at the 10% level.

The pattern is different for the male models. On professionalism and competence, the White male is penalized *more* than the Black male for deviating from clean-cut hair: the racialized gap is a positive 0.31 for professionalism and an imprecisely estimated 0 for competence. This pattern is reversed on agreeableness. The White male receives a large and highly significant premium of 0.37 points for wearing curly hair while the Black male receives a small penalty of 0.07 points, yielding a racialized gap of −0.44 points. Thus, our results suggest that respondents do not see the Black male as less professional in an Afro, but they do see him as less approachable.[17]

The curly hair results, taken together, suggest that evaluators map the same deviation from straight-hair norms onto different perceived deficits depending on the model's race and gender. For Black women, the penalties concentrate on professionalism and competence, dimensions tied to conformity with workplace appearance standards. For Black men, the penalty concentrates on agreeableness, a dimension tied to interpersonal warmth.

Benchmarking

To aid interpretation of the racialized hair penalty's magnitude, we benchmark it against the rating patterns associated with two other attributes that enter the same specification: wearing a t-shirt (relative to a suit) and submitting a blurry photograph (relative to a high-resolution one). Both attributes are within the applicant's control, and photo resolution in particular is effectively costless to change, unlike the natural hair texture they are born with. The results of this comparison are presented in Table C.4 in the Appendix.

For the Black female model, the racialized curly hair penalty is roughly twice the size of the blurry-photo penalty across the professionalism and competence dimensions,[18] and

[17] Section 5.5 presents multiple hypothesis test corrected results. The student sample's greater penalties for curly hair generally remain significantly higher for the Black models relative to the White ones after applying the (Romano & Wolf, 2005) stepdown procedure.

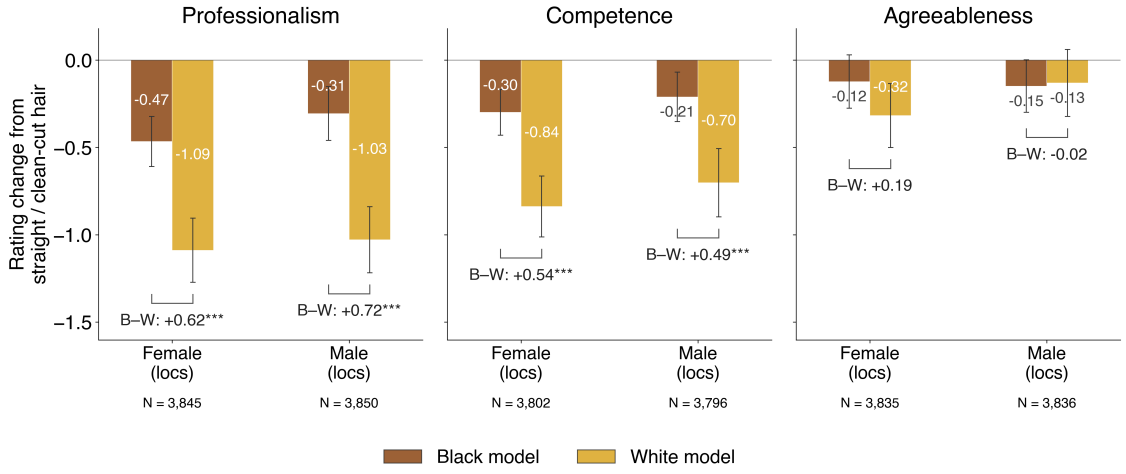
[18] The blur penalty on agreeableness is close to zero, making that ratio uninformative.

erodes between 14 and 64 percent of the benefit of wearing a suit rather than a t-shirt. The racialized excess penalty that evaluators impose on a Black woman for wearing her natural hair texture thus meaningfully exceeds the penalty she would incur from something that is more objectively within her control, such as submitting a low-resolution image.

5.2 Rating differentials for locs

Figure 7 presents the corresponding results for locs. Here the racial asymmetry reverses: White models bear a consistently heavier penalty, and Black models, though still penalized in absolute terms, are penalized considerably less.

Figure 7: Racialized penalties for locs: Student sample



Note: Bars show deviations from each model's own straight-hair (or, for males, clean-cut hair) predicted rating (set to zero), based on Equation (1) evaluated at control-variable means. Error bars are 95% CIs. Brackets report the "B–W gap" (difference-in-differences: Black model's hair penalty minus White model's), from the difference between the Black and White models' hairstyle coefficients. All regressions include subject fixed effects, model fixed effects, and controls for clothing, facial expression, and photo resolution. Observation counts differ across outcomes because of item non-response. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Among both the male and female models, White models are penalized far more severely than Black models for wearing locs, particularly on professionalism and competence. For the female models, the White female's professionalism drops by 1.09 points in locs relative to straight hair, compared to 0.47 for the Black female. On competence, the penalties are 0.84 versus 0.30. The racialized locs gaps are large and positive—0.62 for professionalism and 0.54 for competence—indicating that locs are penalized far more when worn by the White model. On agreeableness, the gap is smaller and not statistically distinguishable from zero. Similarly, among the male models, the White male's professionalism falls by 1.03 points in locs versus clean-cut hair, compared to 0.31 for the Black male. On competence, 0.70 versus

0.21. The racialized locs gaps are 0.72 for professionalism and 0.49 for competence. On agreeableness, however, the racialized locs gap is near zero (-0.02), which is a departure from the large negative curly hair gap on this dimension reported earlier.

One possible explanation for this pattern comes from responses to our open-ended questions, where several respondents expressed strong opposition to White individuals wearing locs, characterizing it as cultural appropriation. Yet Black models are still penalized for wearing locs in absolute terms on every metric. Respondents appear to hold two evaluative standards simultaneously. Locs on a White model read as a norm violation and are penalized accordingly, especially where respondents invoke cultural appropriation. In contrast, locs on a Black model are recognized as culturally congruent, yet are still judged as insufficiently professional. Even when a hairstyle is acknowledged as authentic to Black identity, the underlying appearance norm against non-straight hair persists.

5.3 Heterogeneity by respondent characteristics

Our pre-analysis plan specified heterogeneity checks along five respondent dimensions: race, gender, political leaning, confidence in higher education, and, for students, STEM major. Appendix Tables C.5–C.8 present the findings from those analyses, with each row reporting the difference in the B-W model racialized hair penalty between the indicated respondent group and a reference (omitted) group.

The main takeaway from these analyses is that there is limited heterogeneity across groups in the relative rating differentials they give to Black versus White models of the same gender. With the exception of respondent race and confidence in higher education, and far more sporadically by gender, the majority of the group differences on other dimensions of heterogeneity are statistically indistinguishable from zero: we find no evidence of systematic differences in relative Black-White rating gaps across democrats and republicans or across STEM vs non-STEM major students.

We find more severe relative penalties in agreeableness ratings (a difference of 1.36 points) for the Black female model's kinky hairstyle among Black students than among White students. We document similar marginally significant differences on the male models' professionalism and competence ratings by Black students as well. Our student sample had a very small (about 5%) share of Black students, so these comparisons rest on a small group and are imprecisely estimated. Yet, the direction of heterogeneity is similar in the lay adult sample as well, where the share of Black respondents mirrors the U.S. population share—here, the Black female model sees a marginally higher agreeableness penalty when wearing

her natural kinky hair texture.[19] These results are consistent with 'in-group' evaluators (i.e., evaluators who share a similar racial background as the pictured model) having more information about these hairstyles and their social reception, which may enable them to display a more informed form of bias. Respondents who have experienced hair-based judgment or biases themselves may also be anticipating how evaluators would respond to such candidates. In either scenario, our findings suggest that in-group membership does not soften the racialized penalty—a finding that is consistent with prior evidence that Black women are rated as more 'dominant' when wearing Afro-centric hair (Opie & Phillips, 2015), including by Black evaluators.

Confidence in the higher education system also appears as an important dimension of heterogeneity in the lay adult sample in particular. Almost uniformly across the different metrics, respondents with moderate or high confidence in the higher education system extend systematically smaller favorable locs differentials to Black models than respondents with low confidence. This is more because they are more lenient on White models pictured in locs (which some respondents characterize as cultural appropriation), rather than because they penalize the Black models more harshly for wearing locs. A similar pattern emerges, though less consistently, when looking at the rating differentials for curly/kinky hair relative to straight hair among the female models. Overall, our findings suggest higher Black-White relative penalties associated with evaluators that have greater confidence in higher education systems rather than those more skeptical of the ivory tower, primarily because of the former are more lenient on hairstyle deviations of the White models.

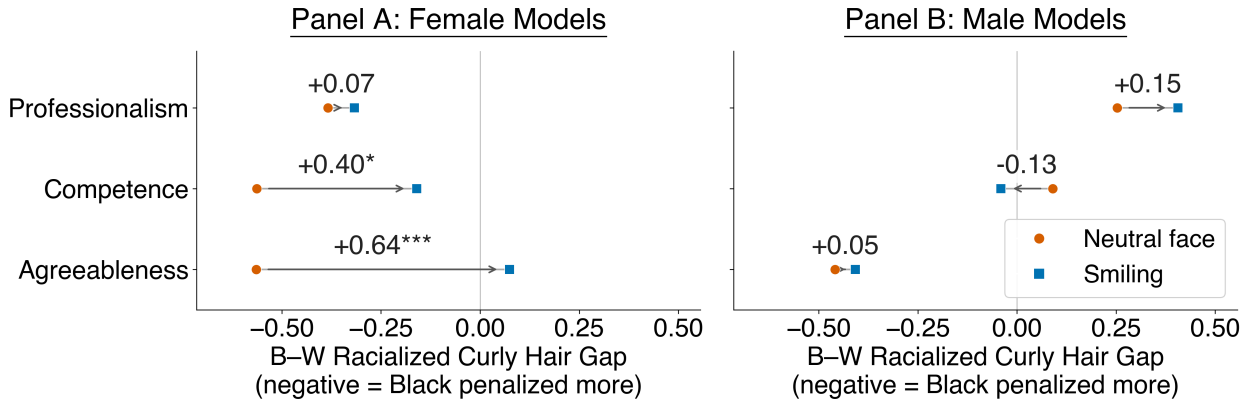
5.4 Facial Expression as a Moderator of Racialized Penalties

Standard theories of statistical discrimination (see Phelps (1972)) would predict that, if evaluators hold prior beliefs that non-mainstream hair signals poorer qualities about Black subjects—such as low agreeableness or low polish—then smiling may provide a countervailing positive signal that can neutralize penalties as evaluators update their beliefs. To test this hypothesis, we study whether and how the Black-White gap in curly/kinky hair ratings changes when models are shown with a smile versus a neutral (or serious) facial expression.

We find that smiling indeed substantially attenuates the racialized curly hair penalty for the Black female model, particularly on dimensions related to agreeableness and competence. Among university students, the Black-White gap in the curly hair penalty on agreeableness falls from -0.56 points when the model is pictured with a neutral (i.e., serious) expression

[19] When it comes to locs as well, Black lay adults extend less of the favorable locs differential to the Black female model than White lay adults do, by 0.40 to 0.43 points across the three rating dimensions.

Figure 8: Smiling attenuates racialized penalties



Note: Each row shows the Black–White racialized curly hair penalty (curly minus straight, Black minus White) under serious (circle) and smiling (square) conditions. Values above arrows report the smile moderation effect (change in B–W gap). $*p < 0.10, **p < 0.05, ***p < 0.01$.

to 0.07 when smiling—an attenuation of 0.64 points. On competence, the gap narrows by 0.4 points. Lay adults exhibit a similar pattern: smiling closes the racialized curly hair gap on competence (a moderation of 0.23 points) and agreeableness (0.30 points), though the professionalism moderation (0.17) is not statistically significant (see Figure B.8). The smile moderation effects point in the same direction for the Black male model as well but are muted and statistically indistinguishable from zero in both samples. Overall, our findings suggest that smiling attenuates racialized hair penalties for the Black female model in particular, consistent with costly signaling of warmth to offset unfavorable priors.

5.5 Robustness Checks

5.5.1 Respondent effort

Our pre-analysis plan specified checks for respondent effort based on survey completion time. First, we flag respondents below the 3rd or above the 97th percentile of completion time and re-estimate our main specifications with these observations dropped. Second, we exclude respondents whose completion time fell below one-fifth of the sample median, a threshold designed to identify those who may have clicked through the survey without engaging with the headshot ratings. Our main results are robust to both restrictions. Point estimates remain stable in magnitude and direction, and statistical significance is preserved across the primary outcomes. We also re-estimate our main specifications restricting to respondents who passed every attention-check question they were shown. The female-model racialized curly penalties retain their sign and significance in both samples and in fact strengthen in

magnitude, though the smaller subset of respondents passing all attention checks widens the confidence intervals, while the male-model patterns attenuate. These results are reported in Table C.12 in the Appendix.

5.5.2 Multiple Hypothesis Tests

Our design tests for racialized hair penalties across 2 races, 2 genders, 2 hairstyle contrasts (curly vs. straight, locs vs. straight), and 3 outcome measures — yielding a large family of simultaneous tests. To guard against false positives from multiple comparisons, we apply the Romano-Wolf step-down procedure (500 bootstrap replications) within each sample, defining families of 3 outcomes (professionalism, competence, agreeableness) for each gender $\times$ hair-type cell. This approach controls the familywise error rate while preserving power by exploiting the joint dependence structure across outcomes. The results are shown in Appendix Figure C.15.

The racialized curly hair penalty for Black women relative to Whites—our central finding—survives correction on professionalism and competence in the student sample ($p =$ 0.010 and 0.002, respectively) and is marginal on agreeableness (p = 0.074). Among lay adults, the Black-White curly hair rating gaps for the female models' perceived competence and the male models' perceived agreeableness survive correction, at $p = 0.04$ and $p = 0.002$, respectively. The Black-White ratings gaps for locs are substantially robust: the racialized gap favoring Black models survives correction across both samples, both genders, and nearly every metric.

We apply the same correction to the smile moderation results, where the coefficient of interest is the triple interaction of expression, hair type, and photo race. For curly hair, the female agreeableness moderation survives among students (p = 0.016), consistent with a warmth signal strong enough to offset the penalty on that dimension but not on professionalism or competence. The moderation is more broadly robust for locs, with results surviving for all metrics in both samples. The overall pattern of these multiple hypothesis testing exercises reinforces our key finding that the racialized curly hair penalty, and the partial ability of smiling to attenuate it, are most clearly detected by evaluators who are closest to academic institutions.

5.5.3 Sensitivity to controls

As pre-specified in our analysis plan, we perform machine learning-based robustness checks to the inclusion of controls. Following Ludwig et al. (2019), we residualize each rating

on subject fixed effects and use a ten-fold cross-validated LASSO to select among a pool of candidate covariates—respondent fatigue and effort proxies and higher-order photo-attribute interactions—before re-estimating the difference-in-differences. As Table C.13 shows, the LASSO-augmented estimates are nearly identical to our main specification in both magnitude and significance, indicating that the penalty is not an artifact of our control choices.

5.5.4 Experimenter demand and social desirability bias

To estimate the effect of our experimenter demand interventions, we interact each hairstyle and race combination in Equation (1) with a treatment status and post-treatment indicator (i.e., TREAT × POST). We calculate these as $\left(\bar{y}^{POST}_{\text{TREAT}} - \bar{y}^{PRE}_{\text{TREAT}}\right) - \left(\bar{y}^{\text{POST}}_{\text{CONTROL}} - \bar{y}^{PRE}_{\text{CONTROL}}\right)$ with Y being the rating differential between curly (or locs) vs. straight/clean-cut hair for the Black model relative to the White model of the same gender.

Overall, we find no strong evidence for experimenter demand effects. The treatments do not appear to have substantially or statistically significantly shifted respondents' racialized evaluation of hairstyle. The few individually significant estimates that do emerge are concentrated among locs, where the historical treatment is associated with modest reductions in the locs penalty for professionalism and competence for the male Black model—for example, a 0.42-point reduction on the professionalism penalty among lay adults—though the White male model's penalty also falls to a similar extent, leaving the relative Black-White gap unchanged. We also observe suggestive narrowing of the Black female model's own curly hair penalties on professionalism, and to a limited extent, competence: point estimates are positive under both treatments, though statistically significant only in the experimenter-demand arms, while the *relative* penalty compared to the White female model is largely unresponsive. If anything, the two treatment messages differ in whether they move evaluations at all: the generic vignette, which frames authenticity in race-neutral terms by claiming that authenticity at work helps productivity, produces no systematic movement; in contrast, the historical treatment yields modest improvements in locs ratings, but again for models of both races, leaving the racial gap itself intact. However, these few statistically significant findings are sporadic, do not survive multiple hypothesis testing and are not consistently echoed across genders or samples. These patterns are consistent with the interpretation that racialized hair perceptions are deeply ingrained and resistant to informational nudges. Detailed results are presented in Tables C.9–C.11 in the Appendix.

6 Discussion

The preceding sections document a racialized penalty on natural Black hair that evaluators impose while rating photos in a setting where racial composition, hairstyle, and all other attributes are experimentally controlled. If the penalty is real, (at least) two sets of instruments exist for reducing it. The first operates on evaluator preferences or beliefs directly, through information that reframes either the trait being judged (authenticity at work) or the history that shaped the norm being applied (the legacy of hair-based discrimination in the United States), which the discussion in Section 5.5.4 above touches on. The second operates on evaluator behavior regardless of preferences, by prohibiting appearance-based decisions that are correlated with race. Our survey engages with this latter channel through a list experiment aimed at identifying respondents' extent of support for legislation against employer hair discrimination.

Introduced in 2019, the CROWN Act prohibits discrimination in employment and public schools on the basis of hair texture and protective hairstyles historically associated with race. In practice it bars grooming codes that, while written in race-neutral language, disallow afros, cornrows, braids, locs, twists, or Bantu knots. California adopted the first version in 2019 and roughly half of U.S. states have since passed analogous statutes, though a federal bill that cleared the House in 2022 has not been enacted. State-level CROWN Acts do not, however, always bind: in *Darryl George v. Barbers Hill Independent School District* (2024), a Texas court held that a school-district policy limiting hair length did not violate the state's CROWN Act, allowing the continued discipline of a Black student for wearing locs. Whether such policies are race-neutral in practice is the empirical question our rating experiment addresses.

To measure the prevalence of privately-held support for laws like the CROWN Act, the survey included a list experiment (e.g., following Aksoy et al. (2025)) at the end of the survey. Respondents were randomly assigned to one of two arms. The control arm saw four innocuous statements and reported only how many they agreed with. The treated arm saw the same four statements plus a fifth sensitive statement: "The law should prohibit employers from telling workers not to wear certain hairstyles like dreadlocks." The difference in the mean count across arms identifies the share of respondents who privately agree with the sensitive statement, without requiring any respondent to reveal her own view. All respondents were separately asked whether they agreed with the same statement in a direct, non-veiled question posed later in the survey.

Figure C.14 reports both measures. In the student sample, 59.3 percent of respondents

agree with the sensitive statement under the veiled measure and 52 percent under the direct question, a gap of approximately 7.4 percentage points. In the lay adult sample the gap goes in the same direction, with roughly 52% and 47% of respondents showing veiled and direct support, respectively—a difference of 5.1 percentage points. While the veiled estimates are imprecise, and the relative gaps statistically insignificant, the gaps suggest that respondents underreport their support for legislation like the CROWN Act. Taken together with the results in Section 5.5 above, this pattern implies that the light-touch information treatments may not be sufficient to shift respondents' racialized evaluations of hair, and that around half of respondents in both samples privately support the legal route that the CROWN Act takes, even when they do not state this support openly.

7 Conclusion

Despite affecting the daily lives of millions of Black Americans by imposing well-documented monetary, time, and health costs, the racialized hair bias we document has long remained outside the scope of empirical economics. This paper introduces the first pre-registered, incentive-compatible experimental design for measuring these penalties in perceptions of presentability using academia, specifically STEM fields, as a setting. We document that these penalties are large, intersectional, and operate through distinct channels for Black women and Black men, and also uncover a statistical discrimination mechanism through which smiling attenuates warmth-related penalties while the excess penalty for professionalism persists. Light-touch information interventions do not appear to reduce bias. This suggests that the penalties do not solely arise from ignorance, with implications for ongoing legal efforts to recognize hair texture and hairstyles of cultural significance (such as locs, or braids) as a protected characteristic of race via the CROWN Act.

Our focus on hair offers unusually clean identification, but the logic extends to any trait that marks identity yet remains (at least) partially malleable. This includes, for instance, accent code-switching among immigrants, skin bleaching in South Asia and Sub-Saharan Africa, or cosmetic procedures in South Korea known colloquially as "employment surgery," each rewarded in labor and marriage markets. Where conforming to dominant group norms comes with economic benefits (Akerlof & Kranton, 2000), minority groups can face a double bind where they can only either deviate and face penalties, or assimilate by paying the ongoing private cost of manipulating those malleable traits.

Our work also adds to a growing literature on representation in academia, and specifically in STEM fields, showing that systemic barriers have hindered the US economy via

missing patents, foregone output, and greater inequality (Cook & Yang, 2018; Hunt et al., 2012), while diverse perspectives enhance patent novelty (Yang et al., 2022) and broaden research attention to underserved populations (Dossi, 2024). A diverse body of educators in STEM higher education can also attract and retain a diverse student body through role-model effects (Redding, 2019; Baker, 2013). Our results suggest that even as representation grows, more remains to be done on the inclusivity of workplace norms that govern who is perceived as professional, competent, and approachable.[20]

We note two caveats. First, in realistic settings there is often more information available about a person than a photograph: voice, mannerisms, accent, credentials. Our estimates may therefore overstate or understate bias depending on how these additional signals interact with appearance. Second, each race-by-gender cell is represented by a single model. Within-model variation is used to estimate relative penalties, but comparison across models could still be affected. We take this concern seriously but highlight several results that are hard to reconcile with model-specific confounds. In particular, the racialized agreeableness penalty and, to a lesser extent, the competence penalty attenuate when the model smiles, but the professionalism penalty does not. This selective attenuation tracks what statistical discrimination theory predicts (Phelps, 1972) and is not easily generated by any idiosyncrasy of one model's appearance. The locs results also show a consistent cross-gender pattern in which White models are penalized more, consistent with cultural-appropriation priors operating at the group level. Our focus group calibration confirmed that the hairstyle stimuli were rated as reflecting comparable effort across race-matched respondents. We see these results as establishing the existence and structure of racialized hair penalties with enough precision to warrant further work with multiple models per cell.

Overall, our findings broaden the welfare analysis of discrimination by showing that its costs extend well beyond the discrete decision points the existing literature has studied, into the ongoing expenditure of money, time, health, and, as we now show, emotional labor required to conform.

[20]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 faculty "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).

References

Adams, D. W. (1995). *Education for extinction: American indians and the boarding school experience, 1875–1928*. Lawrence, KS: University Press of Kansas.

Ajzenman, N., Ferman, B., & Sant'Anna, P. C. (2025). Discrimination in the formation of academic networks: A field experiment on #EconTwitter. *American Economic Review: Insights*, *7*(3), 357–375.

Akerlof, G. A., & Kranton, R. E. (2000). Economics and identity. *Quarterly Journal of Economics*, *115*(3), 715–753. doi: 10.1162/003355300554881

Aksoy, B., Carpenter, C. S., & Sansone, D. (2025, January). Understanding Labor Market Discrimination Against Transgender People: Evidence from a Double List Experiment and a Survey. *Management Science*, *71*(1), 659–677. Retrieved 2026-04-21, from https://pubsonline.informs.org/doi/10.1287/mnsc.2023.02567 doi: 10.1287/mnsc.2023.02567

APNews. (2023, September). *A Black student was suspended for his hairstyle. The school says it wasn't discrimination.* Retrieved 2026-03-10, from https://apnews.com/article/hairstyles-dreadlocks-racial-discrimination-crown-act-034a59b9f2652881470dc606b39e5243 (Section: Education)

Arai, M., & Skogman Thoursie, P. (2009). Renouncing personal names: An empirical examination of surname change and earnings. *Journal of Labor Economics*, *27*(1), 127–147.

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

Bayer, A., Hoover, G. A., & Washington, E. (2020, August). How You Can Work to Increase the Presence and Improve the Experience of Black, Latinx, and Native American People in the Economics Profession. *Journal of Economic Perspectives*, *34*(3), 193–219. doi: 10.1257/jep.34.3.193

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

Bettinger, E. P., & Long, B. T. (2005). Do faculty serve as role models? The impact of instructor gender on female students. *American Economic Review*, *95*(2), 152–157.

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

Bursztyn, L., & Jensen, R. (2017). Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure. *Annual Review of Economics*, *9*(1), 131–153. doi: 10.1146/annurev-economics-063016-103625

Byrd, A. D., & Tharps, L. L. (2002). *Hair story : untangling the roots of Black hair in America*. New York : St. Martin's Griffin. Retrieved 2025-04-17, from http://archive.org/details/isbn_9780312283223

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

Catch. (2024). *Survey of Gen Z job seekers on appearance and employment.* (Reported in The Korea Herald. Available at: https://www.koreaherald.com/view.php?ud=20240730050575)

Cawley, J. (2004). The impact of obesity on wages. *Journal of Human Resources*, *39*(2), 451–474.

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

Chapman, Y. (2007). *"I am not my hair! Or am I?": Black women's transformative experience in their self perceptions of abroad and at home.* (Thesis).

Cook, L. (2014). Violence and Economic Growth: Evidence from African American Patents, 1870- 1940. *Journal of Economic Growth*, *19*(2), 221–257.

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

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. In *University of chicago legal forum* (Vol. 1989, pp. 139–167).

Dash, P. (2006, January). Black hair culture, politics and change. *International Journal of Inclusive Education*. Retrieved 2025-04-17, from https://www.tandfonline.com/doi/full/10.1080/13603110500173183 doi: 10.1080/13603110500173183

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

Francis, D. V., & Darity, W. A., Jr. (2020). Isolation: An alternative to the "acting white" hypothesis in explaining black under-enrollment in advanced courses. *Journal of Economics, Race, and Policy*, *3*(2), 117–122.

Gallego, F. A., Larroulet Philippi, C., & Repetto, A. (2024). What's behind her smile? Health, looks, and self-esteem. *American Economic Journal: Applied Economics*, *16*(2), 359–388.

Glenn, E. N. (2008). Yearning for lightness: Transnational circuits in the marketing and consumption of skin lighteners. *Gender & Society*, *22*(3), 281–302. doi: 10.1177/0891243208316089

Glied, S., & Neidell, M. (2008). *The economic value of teeth* (Working Paper No. 13879). National Bureau of Economic Research.

Godley, M. R. (1994). The end of the queue: Hair as symbol in Chinese history. *East Asian History*(8), 53–72.

Goel, N. J., Thomas, B., Boutté, R. L., Kaur, B., & Mazzeo, S. E. (2021). Body image and eating disorders among south asian american women: What are we missing? *Qualitative Health Research*, *31*(13), 2512–2527. doi: 10.1177/10497323211036896

Grogger, J. (2011). Speech patterns and racial wage inequality. *Journal of Human Resources*, *46*(1), 1–25.

Grogger, J., Steinmayr, A., & Winter, J. (2020). *The wage penalty of regional accents* (Working Paper No. 26719). National Bureau of Economic Research.

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.

Hamermesh, D. S., & Biddle, J. E. (1994). Beauty and the labor market. *American Economic Review*, *84*(5), 1174–1194.

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

Holliday, R., & Elfving-Hwang, J. (2012). Gender, globalization and aesthetic surgery in South Korea. *Body & Society*, *18*(2), 58–81. doi: 10.1177/1357034X12440828

Hsieh, C.-T., Hurst, E., Jones, C. I., & Klenow, P. J. (2019). The allocation of talent and U.S. economic growth. *Econometrica*, *87*(5), 1439–1474.

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.

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

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

Kang, S. K., DeCelles, K. A., Tilcsik, A., & Jun, S. (2016). Whitened résumés: Race and self-presentation in the labor market. *Administrative Science Quarterly*, *61*(3), 469–502.

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

Kline, P., Rose, E. K., & Walters, C. R. (2022). Systemic discrimination among large U.S. employers. *The Quarterly Journal of Economics*, *137*(4), 1963–2036. doi: 10.1093/qje/qjac024

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

Kuumba, M. B., & Ajanaku, F. (1998). Dreadlocks: The hair aesthetics of cultural resistance and collective identity formation. *Mobilization: An International Quarterly*, *3*(2), 227–243. doi: 10.17813/maiq.3.2.nn180v12hu74j318

Lemi, D. C., & Brown, N. E. (2019). Melanin and curls: Evaluation of black women candidates. *The Journal of Race, Ethnicity, and Politics*, *4*(2), 259–296. doi: 10.1017/rep.2019.18

Levi, P. (1958). *Survival in Auschwitz*. New York: Orion Press. (Originally published as *Se questo è un uomo*, De Silva, 1947)

List, J. A. (2025, February). *The experimentalist looks within: Toward an understanding of within-subject experimental designs* (Working Paper No. 33456). National Bureau of Economic Research. doi: 10.3386/w33456

Ludwig, J., Mullainathan, S., & Spiess, J. (2019, May). Augmenting Pre-Analysis Plans with Machine Learning. *AEA Papers and Proceedings*, *109*, 71–76. Retrieved 2026-06-19, from https://www.aeaweb.org/articles?id=10.1257%2Fpandp.20191070 doi: 10.1257/pandp.20191070

McClintock, A. (1995). *Imperial leather: Race, gender, and sexuality in the colonial contest*. New York: Routledge.

Meyers, H. (2021). *Movie-made jews: An american tradition*. New Brunswick, NJ: Rutgers University Press.

Morrow, W. L. (1973). *400 Years Without a Comb*. Black Publishers of San Diego.

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

Paul, M., Zaw, K., Hamilton, D., & Darity, W. A., Jr. (2022). Returns in the labor market: A nuanced view of penalties at the intersection of race and gender in the US. *Feminist Economics*, *28*(2), 1–31.

Pérez-Rosario, V. (2018). Latina feminist theory and writing. In J. M. González & L. Lomas (Eds.), *The cambridge history of latina/o american literature* (pp. 488–509). Cambridge University Press. doi: 10.1017/9781316869468.025

Persico, N., Postlewaite, A., & Silverman, D. (2004). The effect of adolescent experience on labor market outcomes: The case of height. *Journal of Political Economy*, *112*(5), 1019–1053.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *American Economic Review*, *62*(4), 659–661.

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

Reuters. (2024, February). Darryl George: Texas judge rules school district can discipline Black teen over hairstyle. *Reuters*. Retrieved 2026-03-10, from https://www.reuters.com/legal/texas-judge-rules-school-district-can-discipline-black-teen-over-hairstyle-2024-02-22/

Robinson, D. E., & Robinson, T. (2020). Between a loc and a hard place: A socio-historical, legal, and intersectional analysis of hair discrimination and title vii. *University of Maryland Law Journal of Race, Religion, Gender and Class*, *20*(2), 263–288.

Romano, J. P., & Wolf, M. (2005). Stepwise multiple testing as formalized data snooping. *Econometrica*, *73*(4), 1237–1282.

Rose, E. K. (2023). A constructivist perspective on empirical discrimination research. (Working Paper, University of Chicago)

Saramin. (2017). *Survey of HR managers on appearance-based hiring practices.* (Reported in Quartz. Available at: https://qz.com/1016682)

Steele, C. M., Spencer, S. J., & Aronson, J. (2002, January). Contending with group image: The psychology of stereotype and social identity threat. In *Advances in Experimental Social Psychology* (Vol. 34, pp. 379–440). Academic Press. doi: 10.1016/S0065-2601(02)80009-0

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

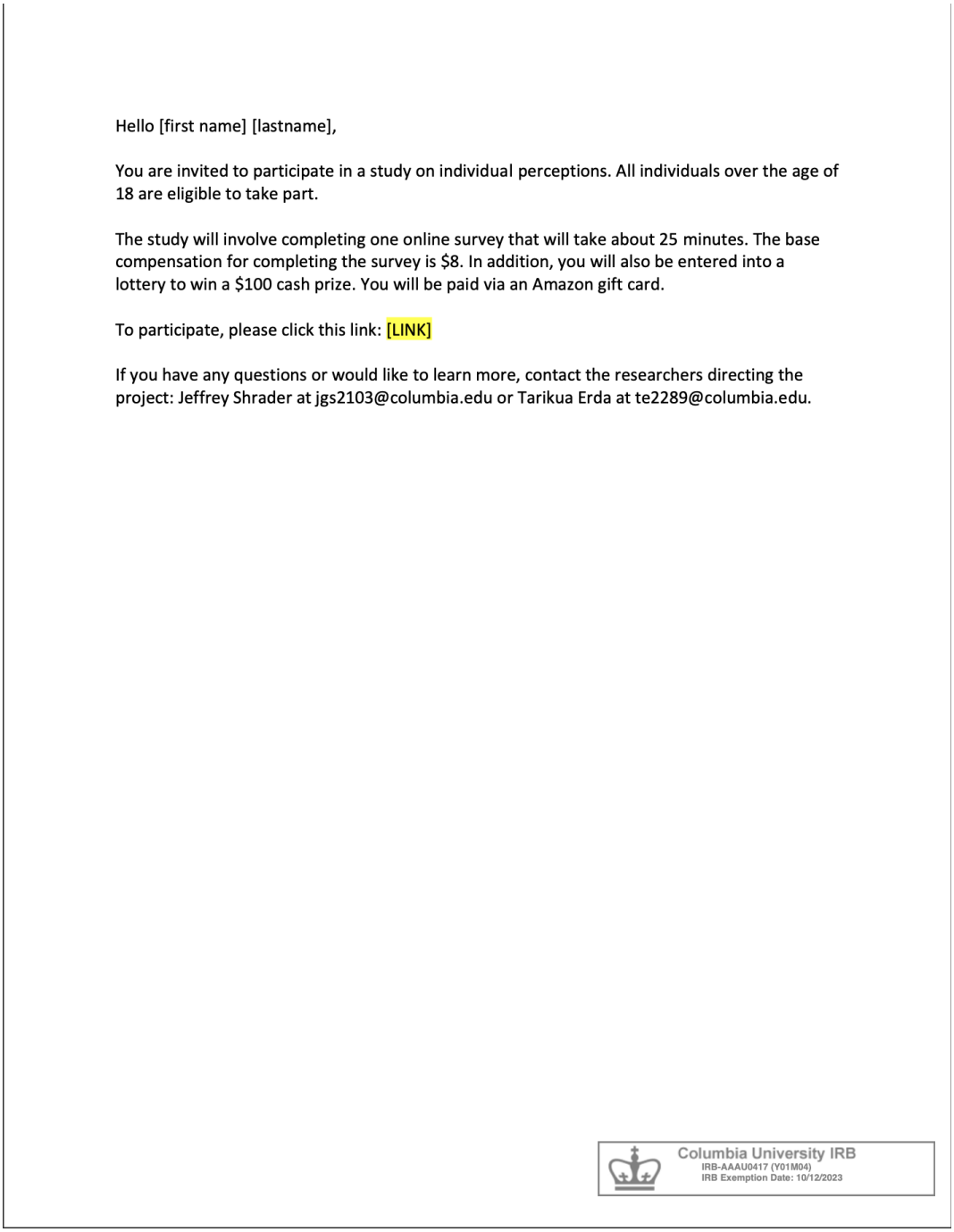
Yang, Y., Tian, T. Y., Woodruff, T. K., Jones, B. F., & Uzzi, B. (2022, September). Gender-diverse teams produce more novel and higher-impact scientific ideas. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(36), e2200841119. doi: 10.1073/pnas.2200841119

Zimring, C. A. (2015). *Clean and white: A history of environmental racism in the United States.* New York: New York University Press.

Appendix

A Details on Survey Design and Content

Figure A.1: Personalized recruitment note sent to university students



Hello [first name] [lastname],

You are invited to participate in a study on individual perceptions. All individuals over the age of 18 are eligible to take part.

The study will involve completing one online survey that will take about 25 minutes. The base compensation for completing the survey is $8. In addition, you will also be entered into a lottery to win a $100 cash prize. You will be paid via an Amazon gift card.

To participate, please click this link: [LINK]

If you have any questions or would like to learn more, contact the researchers directing the project: Jeffrey Shrader at jgs2103@columbia.edu or Tarikua Erda at te2289@columbia.edu.

Columbia University IRB
IRB-AAAU0417 (Y01M04)
IRB Exemption Date: 10/12/2023

Figure A.2: Instructions for rating activity

Over the next few minutes, you will be asked your opinion of the appearance of a set of hypothetical applicants for the job of professor in Chemistry. If it helps, you can also imagine them as someone who may teach you a course in this subject or give you academic advising.

Please read all instructions carefully and answer the questions truthfully.

You will be presented with 32 headshots below and asked to rate them from 1 (lowest) to 10 (highest) on the three metrics below:

1. **Professionalism**: does the picture suggest that the individual is professional? Do they appear to be in line with the norms, rules and expectations of their workplace (i.e. a university)?
2. **Competence**: does the picture suggest that the individual is qualified, capable, and credible enough to rise to the demands of an academic job (e.g. teaching undergraduate and graduate students, producing high quality research)?
3. **Agreeableness**: does this picture suggest that the individual is collegial and able to maintain cordial, lasting collaborations with other scholars? University professors also teach undergraduate and graduate students. Does the picture suggest that they would be approachable to their students and engage them (e.g. answering their questions, mentoring them)?

Figure A.3: Example rating question with a stock photo

Imagine this individual is applying to be a professor who teaches and does research in a STEM field.

How professional, competent, and agreeable does this individual appear to you?

	1 (least)	2	3	4	5	6	7	8	9	10 (extremely)	N/A
Professional	○	○	○	○	○	○	○	○	○	○	○
Competent	○	○	○	○	○	○	○	○	○	○	○
Agreeable	○	○	○	○	○	○	○	○	○	○	○

Figure A.4: Bonus description shown to student survey respondents

Bonus - Lottery for a $100 prize

In addition to your base compensation, your completion of this survey will make you eligible to enter a lottery for a $100 prize. At the end of this survey, you will be asked to enter an email address to receive your compensation as an Amazon gift card.

Many of the photographs you will see in this survey have already been rated by university administrative staff, who were highly incentivized to share their honest views with us about appropriate appearance for educators and university staff. We are now assessing whether their appearance norms for the academic setting are accurately understood by university students like yourself. We will randomly select one photograph and award $100 *each* to 3 respondents whose ratings on the various attributes for that photo are closest to the average rating given by university staff. You would be competing against approximately 500 individuals in this lottery.

Note that you will rate each photograph on three metrics--professionalism, competence, and agreeability--and so *every single* rating you give will be considered for a $100 lottery prize. Thus, it is in your best interest to answer every single question truthfully and thoughtfully to increase your chances of winning an additional $100 prize!

Figure A.5: Information Treatment: Generic vignette

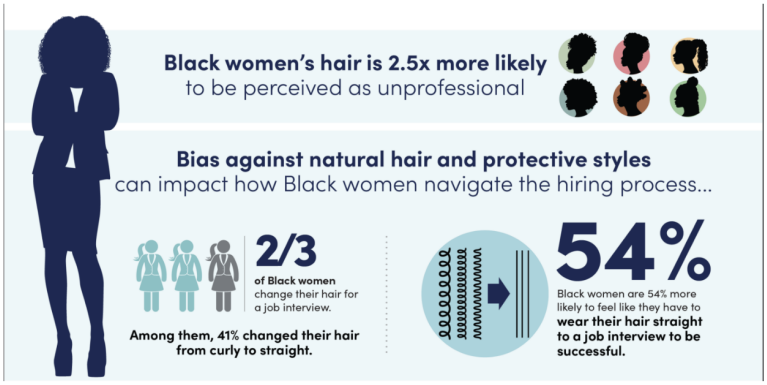
Recent research shows that when employees are allowed to be fully authentic--including in their appearance and personal style--at the workplace, they experience positive outcomes in their well-being and productivity, thus, being better performers for their employer.

Figure A.6: Information Treatment: Historical background on hair bias

Historians document that, in Africa, **hair** was used to denote age, religion, social rank, and marital status as well as other status symbols for centuries.

Yet, with the onset of the Trans-Atlantic Slave Trade, Europeans shaved the hair of enslaved Africans upon arrival to the Americas, a practice that historians claim represented a removal of any trace of cultural identity and a forceful push toward assimilation.

These deep-rooted biases are still alive and well today. Recent studies find persistent biases and discrimination against the curly/coily natural hair texture of Black individuals and their hairstyles (e.g., dreadlocks, braids) at work, school and other settings.



In 2019, The CROWN Act was created to ensure legal protection against discrimination based on hair texture and race-based hairstyles, including protective styles such as braids, dreadlocks, twists, and knots.

It prohibits discrimination on the basis of hairstyle or hair texture that is commonly associated with a particular race or national origin. The CROWN Act has now been adopted in more than 20 US states.

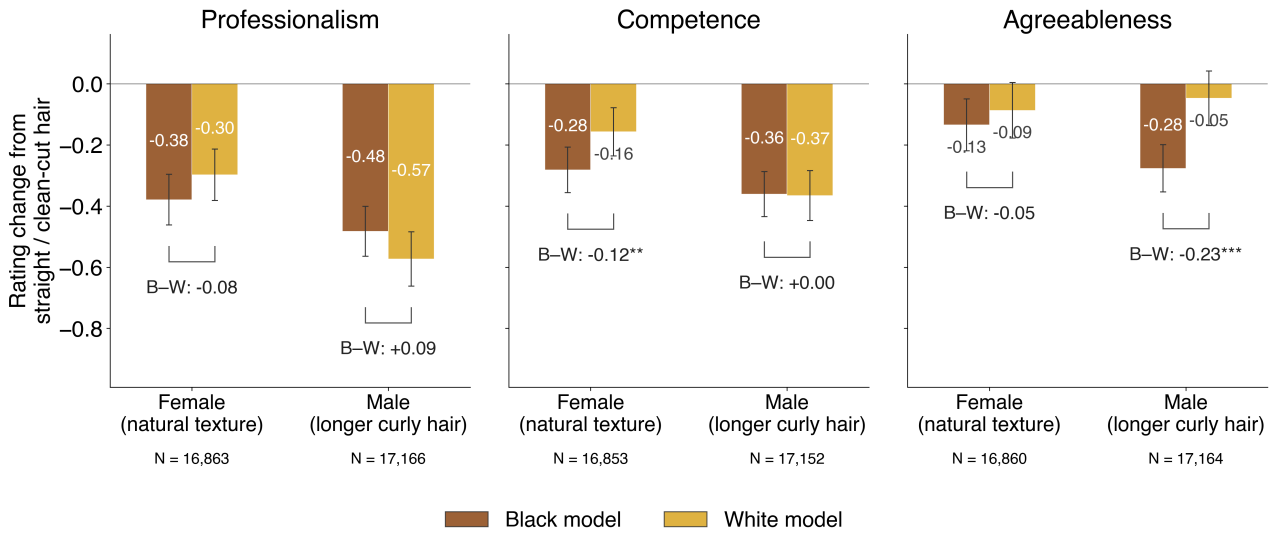
B Lay Adult Results

Figures B.7 and B.9 replicate the main analysis for the representative sample of US lay adults. The broad patterns align with the student sample, though the racialized gaps for curly hair are attenuated.

B.0.1 Rating differentials for natural curly/kinky hair

We find that lay adults penalize both the Black and White female models for wearing natural hair relative to straightened hair: the Black female loses 0.38 points on professionalism and 0.28 on competence, while the White female loses 0.30 and 0.16. Both models are penalized, but the racialized gaps are smaller than in the student sample—−0.08 for professionalism (not statistically significant) and −0.12 for competence, which is significant and survives correction for multiple comparisons. On agreeableness, the racialized gap is an imprecisely estimated −0.05 points.

Figure B.7: Racialized penalties for natural curly/kinky hair: Lay adult sample

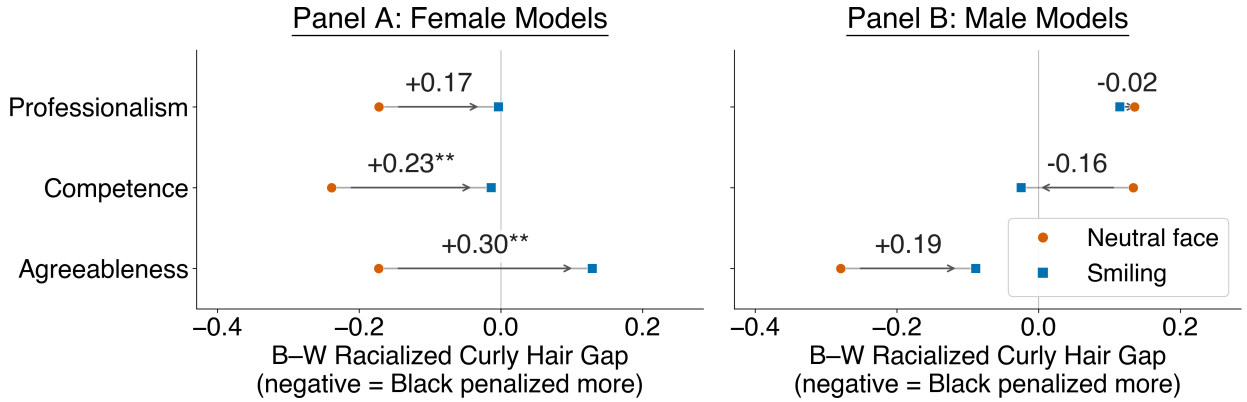


Bars show deviations from each model's own straight-hair (or, for males, clean-cut) predicted rating (set to zero), based on Equation (1) evaluated at control-variable means. Error bars are 95% CIs. Brackets report the "B–W gap" (difference-in-differences: Black model's hair penalty minus White model's), from the difference between the Black and White models' hairstyle coefficients. All regressions include subject fixed effects, model fixed effects, and controls for clothing, facial expression, and photo resolution. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

For the male models, lay adults follow a similar pattern to students. The racialized curly hair gap on agreeableness is −0.23 and highly statistically significant, with the Black

male penalized more for wearing a bigger Afro. While the professionalism and competence relative penalties are estimated imprecisely, their magnitudes suggest that the penalty of the White male model is slightly higher in terms of professionalism (relative gap being +0.09), and the competence penalties are equivalent (0.005). As in the student sample, the Black male faces excess penalties on warmth (agreeableness) rather than on professionalism or competence.

Figure B.8: Smiling attenuates racialized penalties in lay adult sample



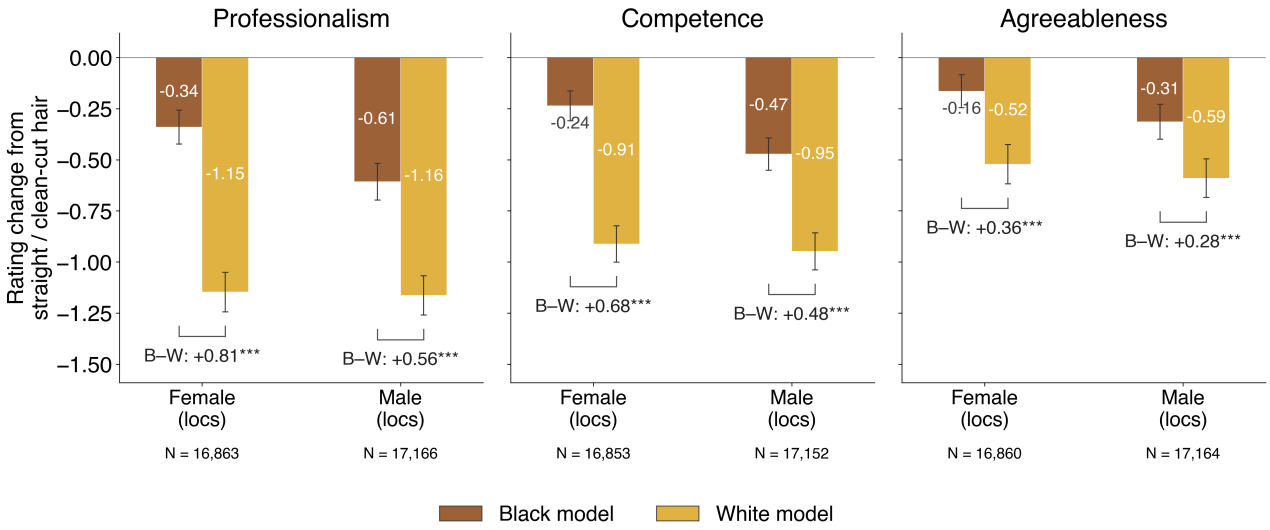
Note: Each row shows the Black–White racialized curly hair penalty (curly minus straight, Black minus White) under serious (circle) and smiling (square) facial expressions. Values above arrows report the smile moderation effect (change in B–W gap). $*p < 0.10, **p < 0.05, ***p < 0.01$.

Figure B.8 shows that smiling attenuates racialized penalties for curly hair in the lay adult sample as well, particularly for the Black female model's perceived competence (by 0.23 points) and agreeableness (0.30 points), both statistically significant and surviving Romano-Wolf correction for multiple hypothesis testing, at least marginally. This further confirms that costly signaling of warmth through smiling may be a non-pecuniary, emotional cost of norms disfavoring Black individuals' natural hair texture.

B.0.2 Rating differentials for locs

The locs results among lay adults closely mirror the student sample. White models are penalized far more severely than Black models for wearing locs. The White female loses 1.15 points on professionalism versus 0.36 for the Black female; on competence, the loc penalties are 0.91 versus 0.24 for the White versus Black female models; and on agreeableness, 0.59 versus 0.31. Therefore, the racialized Black-White gaps for locs are +0.81 for professionalism, +0.68 for competence, and +0.36 for agreeableness, all highly significant. Relative to the Black male model, the White male's ratings drop by larger magnitudes when he is pictured

Figure B.9: Racialized penalties for locs: Lay adult sample



Bars show deviations from each model's own straight-hair (or clean-cut) predicted rating (set to zero), based on Equation (1) evaluated at control-variable means. Error bars are 95% CIs. Brackets report the "B–W gap" (difference-in-differences: Black model's hair penalty minus White model's), from the difference between the Black and White models' hairstyle coefficients. All regressions include subject fixed effects, model fixed effects, and controls for clothing, facial expression, and photo resolution. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

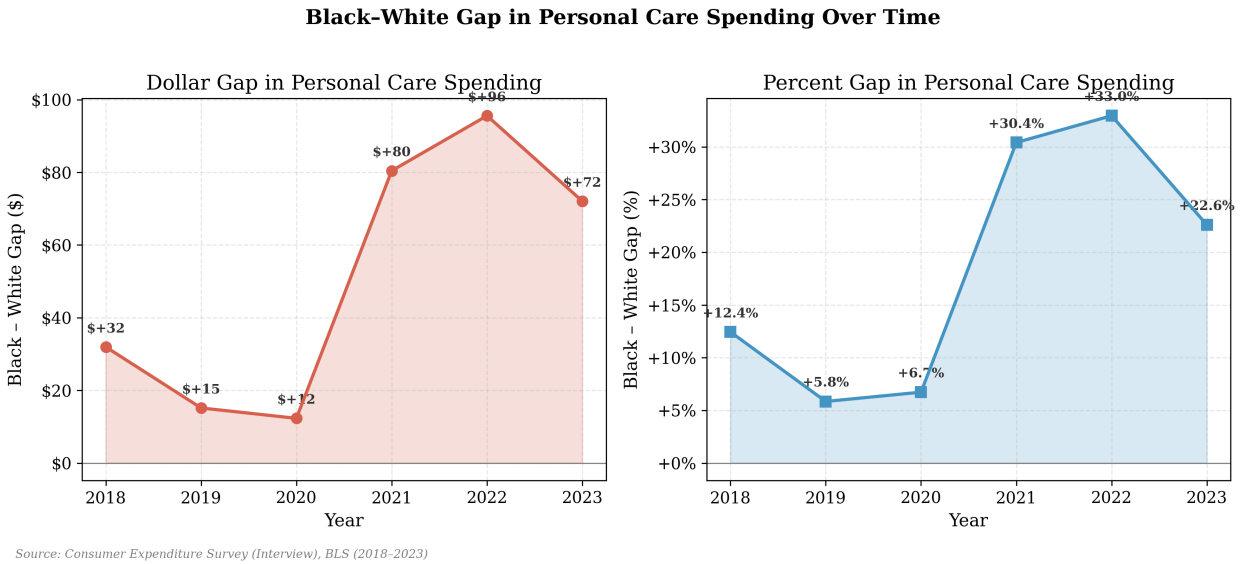
in locs versus clean-cut hair. In particular, the Black-White locs rating gaps are 0.56 for professionalism, +0.48 for competence, and 0.28 for agreeableness, all highly significant.

Like in the student sample, lay adults also point to cultural appropriation as an explanation to why they penalize the White models more than Black ones for wearing locs. However, it is notable that Black models are also penalized for wearing this hairstyle that is purported to have specific cultural meaning for them.

The robustness of the locs results across both samples contrasts with the attenuation of the curly hair Black-White relative penalty findings among lay adults. It is consistent with the view that the two hairstyles activate different evaluative channels: curly hair penalties appear tied to institutional appearance norms that students have absorbed through direct exposure to academia, while the locs penalties reflect broader cultural associations that are legible to any evaluator regardless of institutional context.

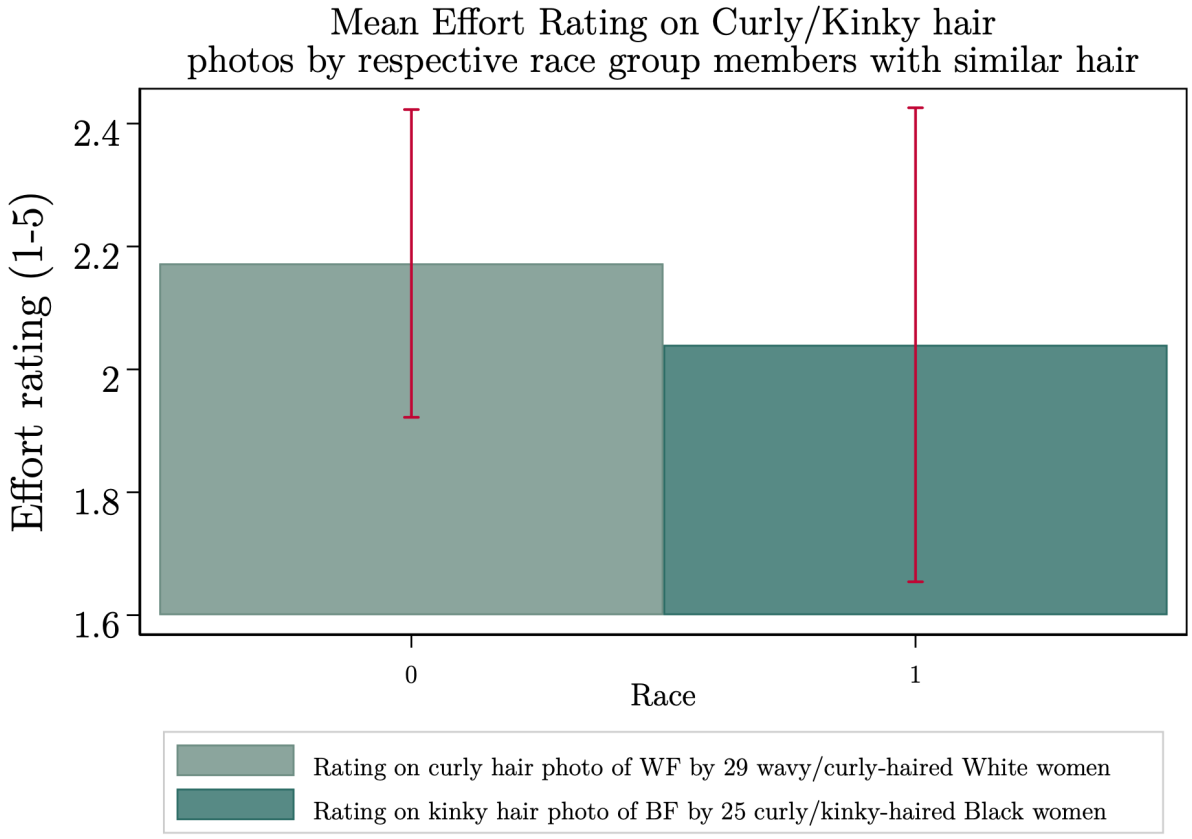
C Additional Tables and Figures

Figure C.10: Gap in spending on personal care



C.1 Focus group survey findings

Figure C.11: 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



C.2 Subjects' self-reported approach to ratings

Figure C.12: Student respondents' explanations about how they made photo ratings

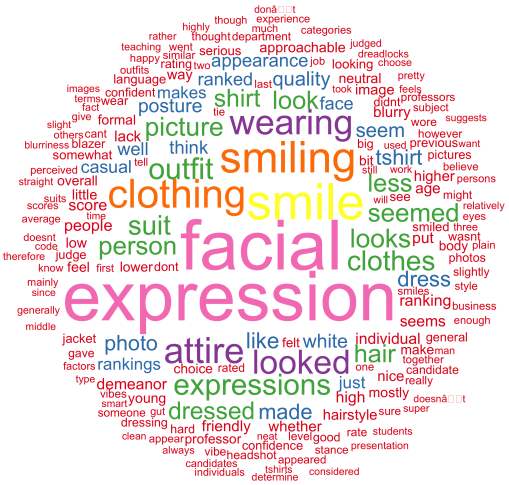


Figure C.13: Lay adult participants' explanations about how they made photo ratings

[20]Note that these wordclouds were 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 "]

Table C.1: Students and Lay adults: Professionalism Ratings

	All students	Lay adults
Suit	2.372***	2.765***
	(0.075)	(0.041)
High resolution	0.163***	0.157***
	(0.046)	(0.019)
Smile	0.094***	0.295***
	(0.036)	(0.019)
Black female X Curly	-0.505***	-0.379***
	(0.080)	(0.042)
Black female X Locs	-0.467***	-0.340***
	(0.073)	(0.042)
Black male X Curly	-0.069	-0.483***
	(0.079)	(0.042)
Black male X Locs	-0.306***	-0.607***
	(0.078)	(0.046)
White female X Curly	-0.176*	-0.298***
	(0.091)	(0.043)
White female X Locs	-1.089***	-1.147***
	(0.094)	(0.049)
White male X Curly	-0.375***	-0.573***
	(0.087)	(0.045)
White male X Locs	-1.029***	-1.163***
	(0.097)	(0.049)
Constant	5.224***	4.697***
	(0.058)	(0.029)
Rating differentials (B–W gaps)		
F Black vs white: Curly–straight/clean-cut	-0.329***	-0.081
	(0.116)	(0.058)
F Black vs white: Locs–straight/clean-cut	0.623***	0.807***
	(0.115)	(0.063)
M Black vs white: Curly–straight/clean-cut	0.306***	0.091
	(0.118)	(0.058)
M Black vs white: Locs–straight/clean-cut	0.723***	0.556***
	(0.123)	(0.062)
Observations	7695	34029

Notes: All models contain subject FE and photo-model FE. Each hair coefficient (e.g., Black female X Curly) is that model's penalty for wearing the hairstyle relative to their own mainstream style photos (straightened hair for women; clean-cut hair for men). Rating differentials are the Black–White gaps in these within-model penalties, for models of the same gender. Standard errors, clustered by respondent, in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table C.2: Students and Lay adults: Competence Ratings

	All students	Lay adults
Suit	1.149***	1.508***
	(0.057)	(0.031)
High resolution	0.178***	0.204***
	(0.040)	(0.018)
Smile	0.210***	0.412***
	(0.036)	(0.019)
Black female X Curly	-0.269***	-0.282***
	(0.072)	(0.038)
Black female X Locs	-0.299***	-0.237***
	(0.067)	(0.037)
Black male X Curly	-0.065	-0.361***
	(0.073)	(0.038)
Black male X Locs	-0.211***	-0.472***
	(0.072)	(0.040)
White female X Curly	0.089	-0.157***
	(0.088)	(0.040)
White female X Locs	-0.838***	-0.912***
	(0.089)	(0.045)
White male X Curly	-0.069	-0.366***
	(0.091)	(0.042)
White male X Locs	-0.703***	-0.948***
	(0.100)	(0.046)
Constant	6.015***	5.665***
	(0.050)	(0.024)
Rating differentials (B–W gaps)		
F Black vs white: Curly–straight/clean-cut	-0.358***	-0.124**
	(0.113)	(0.053)
F Black vs white: Locs–straight/clean-cut	0.539***	0.675***
	(0.110)	(0.057)
M Black vs white: Curly–straight/clean-cut	0.003	0.005
	(0.118)	(0.055)
M Black vs white: Locs–straight/clean-cut	0.492***	0.476***
	(0.120)	(0.059)
Observations	7598	34005

Notes: All models contain subject FE and photo-model FE. Each hair coefficient (e.g., Black female X Curly) is that model's penalty for wearing the hairstyle relative to their own mainstream style photos (straightened hair for women; clean-cut hair for men). Rating differentials are the Black–White gaps in these within-model penalties, for models of the same gender. Standard errors, clustered by respondent, in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table C.3: Students and Lay adults: Agreeableness Ratings

	All students	Lay adults
Suit	0.348***	0.658***
	(0.046)	(0.027)
High resolution	0.017	0.087***
	(0.038)	(0.018)
Smile	1.805***	1.899***
	(0.060)	(0.034)
Black female X Curly	-0.093	-0.135***
	(0.080)	(0.043)
Black female X Locs	-0.123	-0.165***
	(0.078)	(0.041)
Black male X Curly	-0.065	-0.277***
	(0.075)	(0.040)
Black male X Locs	-0.149*	-0.314***
	(0.077)	(0.043)
White female X Curly	0.131	-0.087*
	(0.092)	(0.046)
White female X Locs	-0.317***	-0.522***
	(0.093)	(0.049)
White male X Curly	0.373***	-0.048
	(0.086)	(0.046)
White male X Locs	-0.131	-0.590***
	(0.098)	(0.048)
Constant	5.595***	5.371***
	(0.053)	(0.026)
Rating differentials (B–W gaps)		
F Black vs white: Curly–straight/clean-cut	-0.224*	-0.047
	(0.122)	(0.063)
F Black vs white: Locs–straight/clean-cut	0.195	0.357***
	(0.122)	(0.062)
M Black vs white: Curly–straight/clean-cut	-0.438***	-0.229***
	(0.116)	(0.060)
M Black vs white: Locs–straight/clean-cut	-0.018	0.276***
	(0.117)	(0.063)
Observations	7671	34024

Notes: All models contain subject FE and photo-model FE. Each hair coefficient (e.g., Black female X Curly) is that model's penalty for wearing the hairstyle relative to their own mainstream style photos (straightened hair for women; clean-cut hair for men). Rating differentials are the Black–White gaps in these within-model penalties, for models of the same gender. Standard errors, clustered by respondent, in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table C.4: Benchmarking the racialized hair penalty against penalties for attributes within the applicant's control (student sample)

	Penalty (rating points, 10-point scale)			Ratio: |Hair| / | · |	
Outcome	Hair (B–W DiD)	t-shirt	Blur	t-shirt	Blur
Panel A. Black female model					
Professionalism	−0.33	−2.37	−0.16	0.14	2.0*
Competence	−0.36	−1.15	−0.18	0.31	2.0*
Agreeableness	−0.22	−0.35	−0.02	0.64	13.0*
Panel B. Black male model					
Professionalism	+0.31†	−2.37	−0.16	0.13	1.9
Competence	+0.00†	−1.15	−0.18	0.00	0.0
Agreeableness	−0.44	−0.35	−0.02	1.26	25.4***

Notes. "Hair"is the racialized curly-hair penalty: the difference between the Black and White model's within-model curly-vs-straight (clean-cut for men) rating change, from one specification per outcome with subject and photo-model fixed effects and all photo attributes. "t-shirt"and "Blur"are the within-subject effects of wearing a t-shirt (relative to a suit) and of submitting a blurry photo (relative to a high-resolution one) from the same regression; because all four models enter a single regression, these comparison penalties are identical in Panels A and B by construction. A negative value denotes a penalty. The right two columns report |Hair| as a fraction of the absolute magnitude of each comparison effect. Stars on the |Hair|/|Blur| ratio indicate the one-sided test that |Hair| exceeds |Blur|, run only where the hair DiD is itself a penalty: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

† Hair DiD is positive because the white male is penalized *more* than the Black male for deviating from clean-cut hair. The ratio columns report |DiD|; interpret directionally.

Table C.5: Heterogeneity in the racialized curly hair penalty (student sample)

	Female models (Black–White)			Male models (Black–White)		
	Prof.	Comp.	Agree.	Prof.	Comp.	Agree.
(1) Respondent race (reference: White)						
Black − White	-0.103	-0.536	-1.355**	-0.807*	-0.824*	0.246
	(0.539)	(0.544)	(0.595)	(0.459)	(0.495)	(0.630)
Asian − White	-0.117	-0.010	-0.051	0.150	-0.233	0.508*
	(0.285)	(0.279)	(0.308)	(0.297)	(0.287)	(0.275)
Mixed (2+) − White	-0.214	-0.272	-0.045	-0.577	-0.408	0.131
	(0.464)	(0.501)	(0.464)	(0.411)	(0.428)	(0.386)
(2) Respondent gender (reference: Male)						
Female − Male	0.315	0.125	0.127	-0.071	0.030	0.283
	(0.273)	(0.254)	(0.276)	(0.268)	(0.270)	(0.279)
(3) Political leaning (reference: Independent)						
Democrat − Independent	-0.176	-0.150	0.476	0.236	0.141	0.511*
	(0.301)	(0.282)	(0.328)	(0.301)	(0.303)	(0.310)
Republican − Independent	1.119	1.049	0.815	0.265	-0.319	0.476
	(1.011)	(1.072)	(1.213)	(0.860)	(0.769)	(0.548)
(4) Confidence in higher education (reference: Low)						
Moderate − Low	0.401	0.444	0.282	-0.458	0.174	0.040
	(0.357)	(0.341)	(0.351)	(0.355)	(0.350)	(0.342)
High − Low	-0.026	0.102	0.085	-0.421	-0.054	-0.046
	(0.353)	(0.325)	(0.341)	(0.341)	(0.333)	(0.318)
(5) STEM major (reference: Non-STEM)						
STEM − Non-STEM	0.043	0.003	-0.055	0.120	0.162	0.150
	(0.274)	(0.271)	(0.272)	(0.252)	(0.257)	(0.247)
Observations	5,128	5,063	5,113	5,128	5,063	5,113

Notes: Each panel reports one regression per outcome in which respondent-group indicators are interacted with the model-specific hair indicators: all four photo models enter one regression per dimension, with subject fixed effects and photo-model-by-respondent-group fixed effects; clothing, smile, and blur controls are common across groups. Each row reports the difference in the racialized hair penalty (the Black model's hair–straight rating change minus the White model's, within photo gender) between the indicated respondent group and the panel's reference group, so each standard error provides a test of equality of the two groups' penalties; the reference groups' own penalties are not displayed. The female- and male-model columns are linear combinations from the same regression. Small groups not displayed (respondent race: Other / American Indian / Native Hawaiian; respondent gender: non-binary) are included in the regression with their own interactions. Standard errors clustered at the respondent level in parentheses. Respondents who selected Prefer not to say on a given dimension are dropped from that dimension's regression. * p<0.10, ** p<0.05, *** p<0.01.

Table C.6: Heterogeneity in the racialized curly hair penalty (lay adult sample)

	Female models (Black–White)			Male models (Black–White)		
	Prof.	Comp.	Agree.	Prof.	Comp.	Agree.
(1) Respondent race (reference: White)						
Black − White	-0.049	-0.157	-0.366*	0.076	0.088	-0.311
	(0.225)	(0.192)	(0.221)	(0.208)	(0.201)	(0.206)
Asian − White	-0.075	-0.290	-0.145	0.105	0.060	-0.178
	(0.299)	(0.269)	(0.325)	(0.272)	(0.270)	(0.310)
Mixed (2+) − White	-0.453*	-0.716***	-0.517	0.460*	0.327	0.419
	(0.273)	(0.247)	(0.367)	(0.280)	(0.263)	(0.321)
(2) Respondent gender (reference: Male)						
Female − Male	-0.194	-0.214**	-0.180	0.139	0.026	-0.014
	(0.119)	(0.108)	(0.129)	(0.118)	(0.112)	(0.123)
(3) Political leaning (reference: Independent)						
Democrat − Independent	-0.066	0.038	-0.147	-0.068	-0.021	0.033
	(0.142)	(0.129)	(0.152)	(0.135)	(0.130)	(0.141)
Republican − Independent	0.207	0.203	-0.055	-0.194	-0.169	0.088
	(0.173)	(0.154)	(0.183)	(0.171)	(0.162)	(0.176)
(4) Confidence in higher education (reference: Low)						
Moderate − Low	-0.279*	-0.085	-0.293*	0.157	0.016	0.035
	(0.159)	(0.149)	(0.176)	(0.148)	(0.145)	(0.164)
High − Low	-0.114	0.091	-0.275*	0.075	-0.073	-0.164
	(0.147)	(0.134)	(0.163)	(0.145)	(0.139)	(0.152)
Observations	22,729	22,714	22,731	22,729	22,714	22,731

Notes: Each panel reports one regression per outcome in which respondent-group indicators are interacted with the model-specific hair indicators: all four photo models enter one regression per dimension, with subject fixed effects and photo-model-by-respondent-group fixed effects; clothing, smile, and blur controls are common across groups. Each row reports the difference in the racialized hair penalty (the Black model's hair–straight rating change minus the White model's, within photo gender) between the indicated respondent group and the panel's reference group, so each standard error provides a test of equality of the two groups' penalties; the reference groups' own penalties are not displayed. The female- and male-model columns are linear combinations from the same regression. Small groups not displayed (respondent race: Other / American Indian / Native Hawaiian; respondent gender: non-binary) are included in the regression with their own interactions. Standard errors clustered at the respondent level in parentheses. Respondents who selected Prefer not to say on a given dimension are dropped from that dimension's regression. * p<0.10, ** p<0.05, *** p<0.01.

Table C.7: Heterogeneity in the racialized locs penalty (student sample)

	Female models (Black–White)			Male models (Black–White)		
	Prof.	Comp.	Agree.	Prof.	Comp.	Agree.
(1) Respondent race (reference: White)						
Black − White	0.737	0.632	0.255	-1.014*	-0.065	0.364
	(0.588)	(0.678)	(0.665)	(0.590)	(0.573)	(0.685)
Asian − White	-0.337	-0.115	-0.095	-0.286	-0.365	0.179
	(0.270)	(0.260)	(0.301)	(0.302)	(0.297)	(0.272)
Mixed (2+) − White	-0.509	-0.135	0.106	-0.489	-0.043	0.115
	(0.451)	(0.393)	(0.425)	(0.441)	(0.452)	(0.531)
(2) Respondent gender (reference: Male)						
Female − Male	0.638**	0.416	0.421	0.049	-0.115	-0.229
	(0.290)	(0.279)	(0.291)	(0.287)	(0.292)	(0.273)
(3) Political leaning (reference: Independent)						
Democrat − Independent	0.319	0.144	0.392	-0.105	0.218	0.159
	(0.327)	(0.312)	(0.319)	(0.332)	(0.291)	(0.292)
Republican − Independent	-0.399	-0.356	0.571	0.069	0.362	0.207
	(0.693)	(0.586)	(0.930)	(0.832)	(1.252)	(1.184)
(4) Confidence in higher education (reference: Low)						
Moderate − Low	0.070	-0.268	0.029	-0.371	-0.175	0.174
	(0.356)	(0.354)	(0.374)	(0.351)	(0.310)	(0.332)
High − Low	-0.261	-0.449	-0.277	-0.079	0.153	0.225
	(0.340)	(0.329)	(0.352)	(0.348)	(0.310)	(0.324)
(5) STEM major (reference: Non-STEM)						
STEM − Non-STEM	0.218	0.120	-0.253	0.253	0.189	0.383
	(0.244)	(0.245)	(0.275)	(0.262)	(0.268)	(0.263)
Observations	5,132	5,069	5,117	5,132	5,069	5,117

Notes: Each panel reports one regression per outcome in which respondent-group indicators are interacted with the model-specific hair indicators: all four photo models enter one regression per dimension, with subject fixed effects and photo-model-by-respondent-group fixed effects; clothing, smile, and blur controls are common across groups. Each row reports the difference in the racialized hair penalty (the Black model's hair–straight rating change minus the White model's, within photo gender) between the indicated respondent group and the panel's reference group, so each standard error provides a test of equality of the two groups' penalties; the reference groups' own penalties are not displayed. The female- and male-model columns are linear combinations from the same regression. Small groups not displayed (respondent race: Other / American Indian / Native Hawaiian; respondent gender: non-binary) are included in the regression with their own interactions. Standard errors clustered at the respondent level in parentheses. Respondents who selected Prefer not to say on a given dimension are dropped from that dimension's regression. * p<0.10, ** p<0.05, *** p<0.01.

Table C.8: Heterogeneity in the racialized locs penalty (lay adult sample)

	Female models (Black–White)			Male models (Black–White)		
	Prof.	Comp.	Agree.	Prof.	Comp.	Agree.
(1) Respondent race (reference: White)						
Black − White	-0.407*	-0.403**	-0.425**	-0.037	0.322	0.149
	(0.247)	(0.204)	(0.204)	(0.233)	(0.210)	(0.236)
Asian − White	-0.036	-0.335	-0.583*	0.260	0.087	0.216
	(0.309)	(0.298)	(0.352)	(0.352)	(0.341)	(0.354)
Mixed (2+) − White	-0.042	-0.544**	-0.215	0.612*	0.389	0.551*
	(0.313)	(0.251)	(0.369)	(0.352)	(0.311)	(0.323)
(2) Respondent gender (reference: Male)						
Female − Male	0.080	-0.033	-0.068	-0.021	-0.120	-0.138
	(0.129)	(0.117)	(0.126)	(0.128)	(0.120)	(0.129)
(3) Political leaning (reference: Independent)						
Democrat − Independent	-0.041	-0.029	-0.063	-0.151	-0.277**	-0.134
	(0.155)	(0.137)	(0.146)	(0.147)	(0.141)	(0.155)
Republican − Independent	0.092	0.152	0.075	0.022	-0.090	0.022
	(0.188)	(0.172)	(0.184)	(0.184)	(0.172)	(0.186)
(4) Confidence in higher education (reference: Low)						
Moderate − Low	-0.340**	-0.218	-0.314*	-0.098	-0.314*	0.098
	(0.173)	(0.160)	(0.171)	(0.170)	(0.161)	(0.175)
High − Low	-0.278*	-0.146	-0.278*	-0.085	-0.369**	-0.204
	(0.162)	(0.146)	(0.159)	(0.155)	(0.145)	(0.158)
Observations	22,660	22,645	22,659	22,660	22,645	22,659

Notes: Each panel reports one regression per outcome in which respondent-group indicators are interacted with the model-specific hair indicators: all four photo models enter one regression per dimension, with subject fixed effects and photo-model-by-respondent-group fixed effects; clothing, smile, and blur controls are common across groups. Each row reports the difference in the racialized hair penalty (the Black model's hair–straight rating change minus the White model's, within photo gender) between the indicated respondent group and the panel's reference group, so each standard error provides a test of equality of the two groups' penalties; the reference groups' own penalties are not displayed. The female- and male-model columns are linear combinations from the same regression. Small groups not displayed (respondent race: Other / American Indian / Native Hawaiian; respondent gender: non-binary) are included in the regression with their own interactions. Standard errors clustered at the respondent level in parentheses. Respondents who selected Prefer not to say on a given dimension are dropped from that dimension's regression. * p<0.10, ** p<0.05, *** p<0.01.

C.3 Treatment Effects

Table C.9: Treatment effects: Professionalism Ratings

	Students			Lay adults		
	No demand	Moderate demand	High demand	No demand	Moderate demand	High demand
Post-Generic: Black CURLY penalty change (F)	0.404	1.063***	0.506	-0.104	-0.191	0.334
	(0.278)	(0.333)	(0.364)	(0.147)	(0.213)	(0.206)
Generic TE: Black vs white CURLY penalty (F)	0.099	0.100	0.851	-0.327	-0.417	0.012
	(0.521)	(0.548)	(0.741)	(0.261)	(0.342)	(0.317)
Post-Historical: Black CURLY penalty change (F)	0.367	1.073***	1.106**	0.107	-0.088	0.041
	(0.240)	(0.413)	(0.436)	(0.152)	(0.223)	(0.201)
Historical TE: Black vs white CURLY penalty (F)	-0.127	1.006*	0.143	-0.069	-0.014	-0.133
	(0.428)	(0.609)	(0.594)	(0.256)	(0.305)	(0.296)
Post-Generic: Black LOC penalty change (F)	0.257	0.552*	0.338	-0.037	0.110	0.339*
	(0.294)	(0.318)	(0.310)	(0.149)	(0.230)	(0.189)
Generic TE: Black vs white LOC penalty (F)	-0.347	-0.130	-0.125	-0.188	0.297	-0.341
	(0.584)	(0.632)	(0.581)	(0.255)	(0.357)	(0.318)
Post-Historical: Black LOC penalty change (F)	0.334	0.473	0.757*	0.231	0.226	0.026
	(0.262)	(0.321)	(0.387)	(0.169)	(0.201)	(0.210)
Historical TE: Black vs white LOC penalty (F)	-0.716	-0.440	-0.801	0.129	0.247	-0.126
	(0.475)	(0.591)	(0.648)	(0.282)	(0.322)	(0.335)
Post-Generic: Black CURLY penalty change (M)	-0.282	-0.136	0.167	0.159	0.123	0.340*
	(0.285)	(0.364)	(0.474)	(0.129)	(0.209)	(0.194)
Generic TE: Black vs white CURLY penalty (M)	-0.133	-0.183	-0.775	0.199	0.153	0.076
	(0.523)	(0.671)	(0.751)	(0.243)	(0.339)	(0.303)
Post-Historical: Black CURLY penalty change (M)	-0.108	-0.307	-0.230	-0.097	0.614***	0.258
	(0.259)	(0.395)	(0.407)	(0.132)	(0.224)	(0.205)
Historical TE: Black vs white CURLY penalty (M)	-0.549	-0.697	-1.058	0.159	0.679*	0.153
	(0.496)	(0.545)	(0.657)	(0.247)	(0.347)	(0.326)
Post-Generic: Black LOC penalty change (M)	0.102	-0.072	-0.002	-0.067	0.391	0.315
	(0.278)	(0.355)	(0.509)	(0.155)	(0.246)	(0.220)
Generic TE: Black vs white LOC penalty (M)	-0.527	-0.712	-1.563**	-0.339	0.268	-0.528
	(0.567)	(0.621)	(0.779)	(0.267)	(0.367)	(0.337)
Post-Historical: Black LOC penalty change (M)	-0.335	0.479	0.904***	0.416***	0.380*	0.151
	(0.279)	(0.387)	(0.312)	(0.140)	(0.209)	(0.244)
Historical TE: Black vs white LOC penalty (M)	-1.201**	0.049	-0.532	0.057	-0.071	-0.424
	(0.548)	(0.673)	(0.570)	(0.249)	(0.339)	(0.332)
Observations	9817	7376	7477	44576	33563	33528

Notes: All models contain subject FE and model x post x treatment cell FE; SEs clustered by subject in parentheses. Post-T rows: post-pre change in the Black model's within-model hair penalty in arm T (positive = penalty shrinks). TE rows: that change minus the control-arm change, minus the white model's DDD (Black vs white). Demand columns include respondents with no demand induction. * p<0.10, ** p<0.05, *** p<0.01.

Table C.10: Treatment effects: Competence Ratings

	Students			Lay adults		
	No demand	Moderate demand	High demand	No demand	Moderate demand	High demand
Post-Generic: Black CURLY penalty change (F)	0.255	0.538*	0.446*	0.058	-0.053	0.334*
	(0.226)	(0.320)	(0.247)	(0.127)	(0.176)	(0.183)
Generic TE: Black vs white CURLY penalty (F)	0.074	0.029	0.985	0.006	-0.119	0.150
	(0.464)	(0.556)	(0.680)	(0.234)	(0.308)	(0.288)
Post-Historical: Black CURLY penalty change (F)	0.312	0.448	0.593	-0.038	-0.141	-0.046
	(0.248)	(0.397)	(0.399)	(0.134)	(0.182)	(0.172)
Historical TE: Black vs white CURLY penalty (F)	-0.008	0.220	0.211	-0.076	-0.072	-0.239
	(0.435)	(0.528)	(0.620)	(0.225)	(0.273)	(0.272)
Post-Generic: Black LOC penalty change (F)	0.308	0.122	0.150	0.012	0.274	0.232
	(0.248)	(0.350)	(0.270)	(0.130)	(0.197)	(0.164)
Generic TE: Black vs white LOC penalty (F)	-0.131	-0.018	-0.213	-0.002	0.319	-0.151
	(0.516)	(0.631)	(0.536)	(0.232)	(0.312)	(0.285)
Post-Historical: Black LOC penalty change (F)	0.111	0.117	0.324	0.037	0.032	-0.074
	(0.225)	(0.312)	(0.355)	(0.147)	(0.174)	(0.172)
Historical TE: Black vs white LOC penalty (F)	-0.612	-0.394	-0.893	-0.098	0.134	-0.011
	(0.470)	(0.517)	(0.621)	(0.248)	(0.291)	(0.288)
Post-Generic: Black CURLY penalty change (M)	0.172	0.188	0.221	0.136	0.035	0.302*
	(0.256)	(0.316)	(0.359)	(0.135)	(0.179)	(0.182)
Generic TE: Black vs white CURLY penalty (M)	0.478	0.055	-1.004	0.098	0.274	0.091
	(0.487)	(0.576)	(0.628)	(0.238)	(0.294)	(0.292)
Post-Historical: Black CURLY penalty change (M)	-0.200	-0.285	0.239	-0.005	0.466**	0.129
	(0.277)	(0.275)	(0.362)	(0.121)	(0.181)	(0.172)
Historical TE: Black vs white CURLY penalty (M)	0.162	0.031	-0.110	0.349	0.418	-0.288
	(0.468)	(0.534)	(0.627)	(0.235)	(0.311)	(0.275)
Post-Generic: Black LOC penalty change (M)	0.637***	0.284	-0.057	-0.041	0.226	0.111
	(0.240)	(0.374)	(0.418)	(0.141)	(0.229)	(0.190)
Generic TE: Black vs white LOC penalty (M)	0.341	-0.226	-0.639	-0.289	-0.118	-0.676**
	(0.563)	(0.640)	(0.685)	(0.246)	(0.326)	(0.291)
Post-Historical: Black LOC penalty change (M)	-0.233	-0.045	0.959***	0.321**	0.331*	0.109
	(0.243)	(0.311)	(0.308)	(0.126)	(0.177)	(0.201)
Historical TE: Black vs white LOC penalty (M)	-0.583	-0.092	-0.108	0.119	-0.208	-0.583**
	(0.513)	(0.599)	(0.627)	(0.228)	(0.286)	(0.281)
Observations	9615	7180	7307	44537	33527	33488

Notes: All models contain subject FE and model x post x treatment cell FE; SEs clustered by subject in parentheses. Post-T rows: post-pre change in the Black model's within-model hair penalty in arm T (positive = penalty shrinks). TE rows: that change minus the control-arm change, minus the white model's DDD (Black vs white). Demand columns include respondents with no demand induction. * p<0.10, ** p<0.05, *** p<0.01.

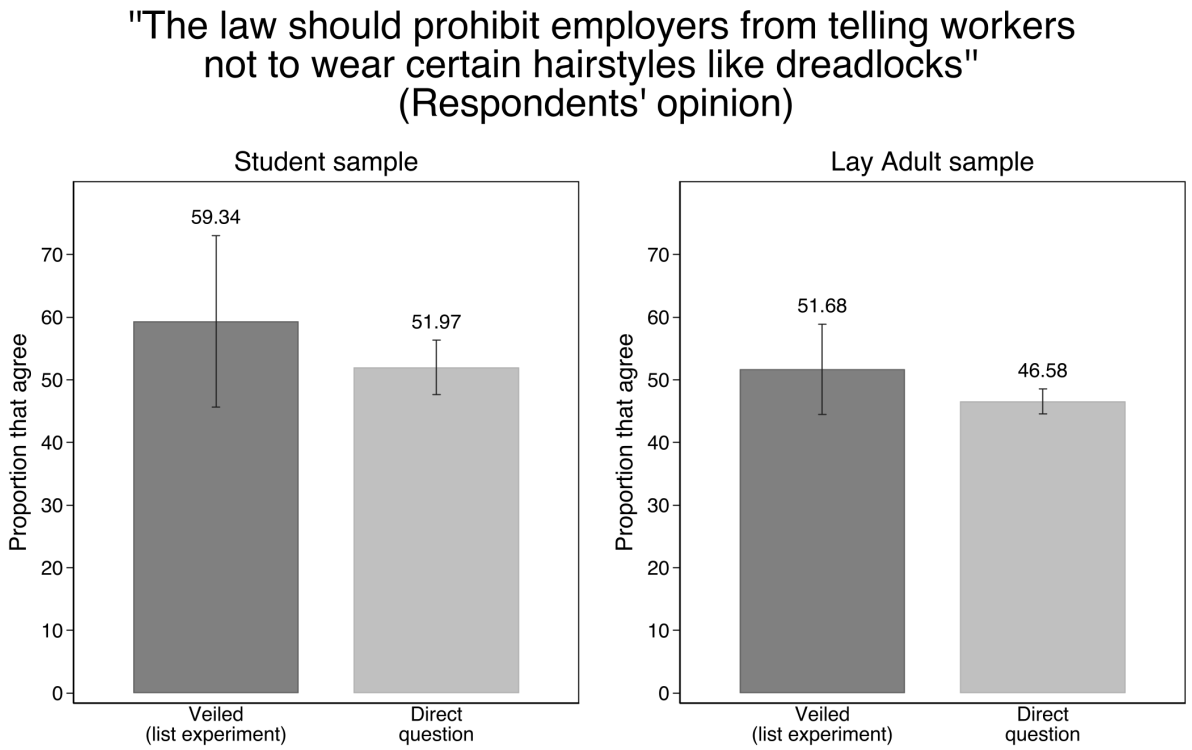
Table C.11: Treatment effects: Agreeableness Ratings

	Students			Lay adults		
	No demand	Moderate demand	High demand	No demand	Moderate demand	High demand
Post-Generic: Black CURLY penalty change (F)	0.186	0.658	0.217	0.029	0.094	-0.024
	(0.259)	(0.464)	(0.350)	(0.152)	(0.217)	(0.198)
Generic TE: Black vs white CURLY penalty (F)	-0.029	0.285	0.242	-0.293	-0.020	-0.710**
	(0.478)	(0.659)	(0.759)	(0.272)	(0.366)	(0.324)
Post-Historical: Black CURLY penalty change (F)	-0.023	0.023	0.261	0.094	-0.082	0.076
	(0.223)	(0.416)	(0.392)	(0.148)	(0.233)	(0.203)
Historical TE: Black vs white CURLY penalty (F)	-0.349	-0.183	-0.114	-0.111	0.027	-0.476
	(0.473)	(0.590)	(0.707)	(0.271)	(0.368)	(0.335)
Post-Generic: Black LOC penalty change (F)	0.161	-0.482	0.095	0.084	0.126	0.142
	(0.275)	(0.371)	(0.306)	(0.133)	(0.204)	(0.174)
Generic TE: Black vs white LOC penalty (F)	-0.314	-0.998	-0.949	-0.150	0.348	-0.408
	(0.509)	(0.765)	(0.693)	(0.256)	(0.376)	(0.323)
Post-Historical: Black LOC penalty change (F)	-0.209	-0.264	0.441	-0.039	0.037	-0.156
	(0.258)	(0.414)	(0.302)	(0.151)	(0.196)	(0.191)
Historical TE: Black vs white LOC penalty (F)	-1.124**	-0.700	-1.060*	-0.272	0.278	-0.483
	(0.472)	(0.659)	(0.539)	(0.286)	(0.333)	(0.345)
Post-Generic: Black CURLY penalty change (M)	0.025	-0.003	0.361	0.154	-0.066	-0.003
	(0.276)	(0.339)	(0.333)	(0.144)	(0.199)	(0.201)
Generic TE: Black vs white CURLY penalty (M)	-0.023	-0.425	0.096	0.041	-0.255	-0.422
	(0.553)	(0.719)	(0.643)	(0.262)	(0.343)	(0.326)
Post-Historical: Black CURLY penalty change (M)	-0.268	-0.451	0.158	-0.134	0.606***	-0.012
	(0.316)	(0.420)	(0.423)	(0.131)	(0.194)	(0.186)
Historical TE: Black vs white CURLY penalty (M)	0.013	-0.495	0.119	-0.074	0.744**	-0.075
	(0.518)	(0.693)	(0.636)	(0.251)	(0.342)	(0.315)
Post-Generic: Black LOC penalty change (M)	0.278	-0.067	0.026	-0.098	0.433*	0.039
	(0.249)	(0.390)	(0.397)	(0.164)	(0.242)	(0.207)
Generic TE: Black vs white LOC penalty (M)	-0.168	-0.715	-0.680	-0.270	0.181	-0.518
	(0.521)	(0.683)	(0.650)	(0.279)	(0.386)	(0.338)
Post-Historical: Black LOC penalty change (M)	-0.335	-0.362	1.100***	0.231*	0.150	0.070
	(0.263)	(0.347)	(0.344)	(0.137)	(0.222)	(0.202)
Historical TE: Black vs white LOC penalty (M)	-0.333	-0.830	0.825	0.146	0.198	-0.259
	(0.481)	(0.565)	(0.651)	(0.257)	(0.348)	(0.301)
Observations	9779	7370	7464	44561	33562	33514

Notes: All models contain subject FE and model x post x treatment cell FE; SEs clustered by subject in parentheses. Post-T rows: post-pre change in the Black model's within-model hair penalty in arm T (positive = penalty shrinks). TE rows: that change minus the control-arm change, minus the white model's DDD (Black vs white). Demand columns include respondents with no demand induction. * p<0.10, ** p<0.05, *** p<0.01.

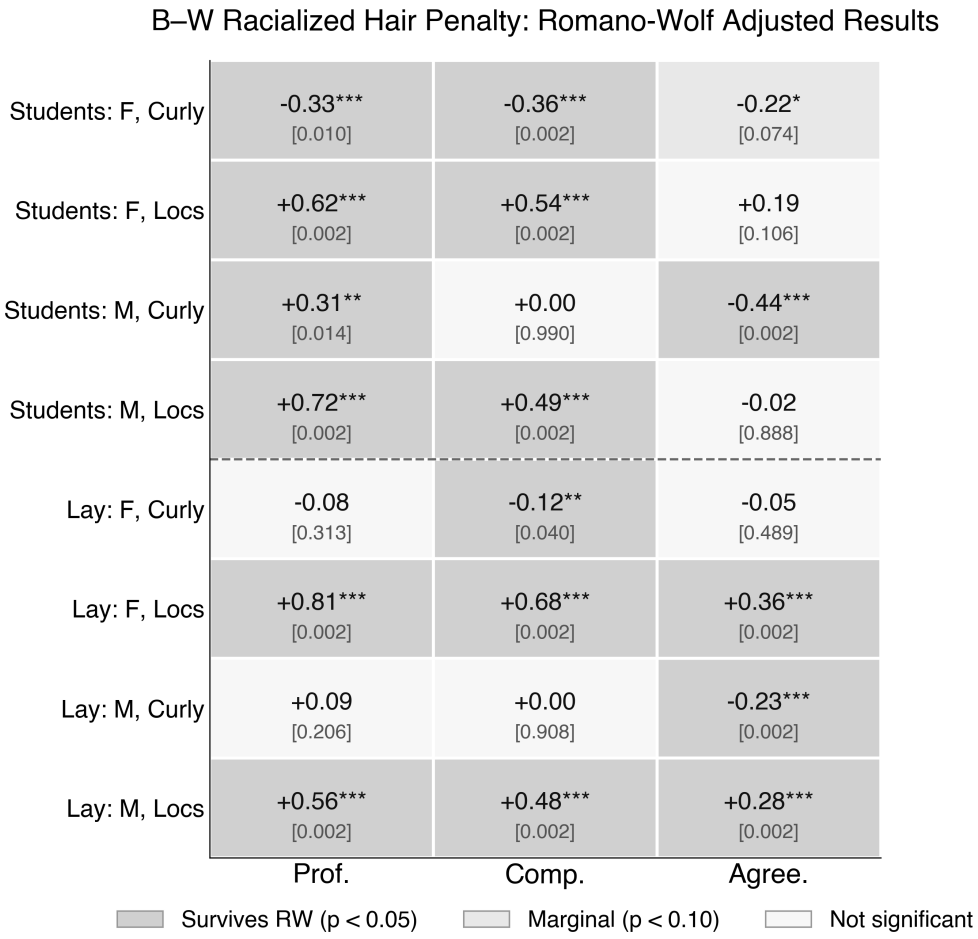
C.4 List Experiment Results

Figure C.14: Veiled vs direct support for legislation on employers banning hairstyles



C.5 Multiple Hypothesis Testing Robustness Checks

Figure C.15: Racialized Hair Penalty estimates with Romano-Wolf Correction



C.6 Participant Effort Robustness Checks

Table C.12: Respondent-effort robustness: racialized hair penalty under completion-time restrictions

	Full sample		Time $\in [P3, P97]$		Time $\geq$ med/5		Passed attn.	
	F	M	F	M	F	M	F	M
Panel A. Student sample								
Curly hair (vs. straight/clean-cut)								
Professionalism	-0.329***	0.306***	-0.326***	0.324***	-0.329***	0.299**	-0.362*	0.032
	(0.116)	(0.118)	(0.119)	(0.119)	(0.116)	(0.118)	(0.192)	(0.198)
Competence	-0.358***	0.003	-0.369***	0.031	-0.358***	-0.002	-0.520**	-0.319
	(0.113)	(0.118)	(0.115)	(0.118)	(0.113)	(0.118)	(0.204)	(0.212)
Agreeableness	-0.224*	-0.438***	-0.224*	-0.443***	-0.223*	-0.440***	-0.354*	-0.708***
	(0.122)	(0.116)	(0.125)	(0.117)	(0.122)	(0.116)	(0.215)	(0.193)
Locs (vs. straight/clean-cut)								
Professionalism	0.623***	0.723***	0.595***	0.747***	0.629***	0.714***	1.014***	0.140
	(0.115)	(0.123)	(0.116)	(0.125)	(0.115)	(0.123)	(0.175)	(0.210)
Competence	0.539***	0.492***	0.513***	0.520***	0.542***	0.486***	0.844***	0.240
	(0.110)	(0.120)	(0.112)	(0.122)	(0.110)	(0.120)	(0.189)	(0.210)
Agreeableness	0.195	-0.018	0.160	0.007	0.194	-0.020	0.302	-0.331*
	(0.122)	(0.117)	(0.124)	(0.119)	(0.122)	(0.117)	(0.199)	(0.196)
Observations (N)	3845	3850	3717	3741	3838	3847	1494	1456
Panel B. Lay adult sample								
Curly hair (vs. straight/clean-cut)								
Professionalism	-0.081	0.091	-0.081	0.092	-0.079	0.090	-0.015	0.089
	(0.058)	(0.058)	(0.059)	(0.060)	(0.058)	(0.058)	(0.069)	(0.069)
Competence	-0.124**	0.005	-0.132**	0.003	-0.122**	0.005	-0.113*	0.012
	(0.053)	(0.055)	(0.054)	(0.057)	(0.053)	(0.055)	(0.062)	(0.066)
Agreeableness	-0.047	-0.229***	-0.051	-0.231***	-0.048	-0.228***	-0.034	-0.190***
	(0.063)	(0.060)	(0.065)	(0.061)	(0.063)	(0.060)	(0.075)	(0.071)
Locs (vs. straight/clean-cut)								
Professionalism	0.807***	0.556***	0.793***	0.557***	0.807***	0.555***	0.899***	0.600***
	(0.063)	(0.062)	(0.065)	(0.063)	(0.063)	(0.062)	(0.076)	(0.074)
Competence	0.675***	0.476***	0.675***	0.471***	0.674***	0.475***	0.718***	0.529***
	(0.057)	(0.059)	(0.058)	(0.060)	(0.057)	(0.059)	(0.068)	(0.071)
Agreeableness	0.357***	0.276***	0.360***	0.263***	0.357***	0.277***	0.356***	0.348***
	(0.062)	(0.063)	(0.063)	(0.064)	(0.062)	(0.063)	(0.075)	(0.076)
Observations (N)	16863	17166	16126	16422	16852	17159	12153	12405

Notes: Standard errors in parentheses, clustered at the respondent level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each cell reports the racialized hair penalty (Black curly–straight or Black locs–straight, minus the same contrast for the White model) from the baseline specification: a specification with subject and photo-model fixed effects and model-specific hairstyle indicators; the F and M columns are linear combinations from the same regression. The first three restrictions operate on the respondent's total survey completion time: the P3 and P97 cutoffs are the 3rd and 97th percentiles; respondents below the sample median divided by 5 are dropped in the third column group. The fourth column group restricts to respondents who passed all attention-check questions.

Table C.13: Racialized hair penalty: baseline vs. LASSO-augmented specification

	Student sample				Lay adult sample			
	Baseline		LASSO		Baseline		LASSO	
	F	M	F	M	F	M	F	M
Curly (vs. straight/clean-cut)								
Professionalism	-0.329***	0.306***	-0.322***	0.290**	-0.081	0.091	-0.109*	0.082
	(0.116)	(0.118)	(0.116)	(0.119)	(0.058)	(0.058)	(0.058)	(0.059)
Competence	-0.358***	0.003	-0.354***	-0.003	-0.124**	0.005	-0.142***	-0.001
	(0.113)	(0.118)	(0.112)	(0.118)	(0.053)	(0.055)	(0.053)	(0.056)
Agreeableness	-0.224*	-0.438***	-0.242**	-0.455***	-0.047	-0.229***	-0.073	-0.231***
	(0.122)	(0.116)	(0.121)	(0.116)	(0.063)	(0.060)	(0.063)	(0.060)
Locs (vs. straight/clean-cut)								
Professionalism	0.623***	0.723***	0.603***	0.714***	0.807***	0.556***	0.801***	0.579***
	(0.115)	(0.123)	(0.117)	(0.123)	(0.063)	(0.062)	(0.063)	(0.062)
Competence	0.539***	0.492***	0.528***	0.486***	0.675***	0.476***	0.672***	0.490***
	(0.110)	(0.120)	(0.110)	(0.120)	(0.057)	(0.059)	(0.057)	(0.059)
Agreeableness	0.195	-0.018	0.185	0.003	0.357***	0.276***	0.361***	0.295***
	(0.122)	(0.117)	(0.121)	(0.117)	(0.062)	(0.063)	(0.062)	(0.063)
Observations	7695	7695	7695	7695	34029	34029	34029	34029

Notes: Each cell reports the racialized hair penalty (the Black model's within-model curly–straight or locs–straight rating change, minus the White model's) from the *baseline* specification: one regression per sample and outcome including all four photo models, subject and photo-model fixed effects, and clothing, smile, and resolution controls; the F and M columns are linear combinations from the same regression. The *LASSO* columns re-estimate that specification after adding controls chosen by a 10-fold cross-validated LASSO from the candidate pool: the pre-rating count and its square and cube, log survey duration, and the two- and three-way photo-attribute interactions. The LASSO runs on the outcome residualized on subject and photo-model fixed effects; treatment terms are never penalized (Ludwig et al., 2019). Standard errors in parentheses, clustered at the respondent level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.