Beholding Inequality: Race, Gender, and Returns to Physical Attractiveness in the United States

Ellis P. Monk Jr.
*Harvard University*

Michael H. Esposito and Hedwig Lee
*University of Washington, St. Louis*

Physical attractiveness is an important axis of social stratification associated with educational attainment, marital patterns, earnings, and more. Still, relative to ethnoracial and gender stratification, physical attractiveness is relatively understudied. In particular, little is known about whether returns to physical attractiveness vary by race or significantly vary by race and gender combined. In this study, we use nationally representative data to examine whether (1) socially perceived physical attractiveness is unequally distributed across race/ethnicity and gender subgroups and (2) returns to physical attractiveness vary significantly across race/ethnicity and gender subgroups. Notably, the magnitude of the earnings disparities along the perceived attractiveness continuum, net of controls, rivals and/or exceeds in magnitude the black-white race gap and, among African-Americans, the black-white race gap and the gender gap in earnings. The implications of these findings for current and future research on the labor market and social inequality are discussed.

INTRODUCTION

Running mostly in parallel to studies of ethnoracial and gender inequality, a burgeoning body of research strongly suggests that perceived physical

---

1 An earlier version of this article was presented at the 2017 Annual Meeting of the Population Association of America (PAA). Direct correspondence to Ellis Monk, Department

© 2021 by The University of Chicago. All rights reserved.
0002-9602/2021/12701-0005$10.00

194  *AJS* Volume 127 Number 1 (July 2021): 194–241
attractiveness is a powerful source of social inequality and stratification (see Holla and Kuipers 2015, p. 290; see also Kuipers 2015). Researchers have found, for example, that perceived physical attractiveness is significantly associated with wealth, relationship length and quality (Yela 2001), being perceived as credible (Madera et al. 2007), being judged to be fit and healthy (Weeden and Sabini 2005), being socially desired by others (Anderson et al. 2010), as well as with educational attainment, social capital, and social network structures, occupational status and wages (Hamermesh and Biddle 1994; Mobius and Rosenblat 2006; Jæger 2011; Mears 2014, 2015; Scholz and Sicinski 2015), and even differences in some of these outcomes among siblings (Conley 2004; Gordon, Crosnoe, and Wang 2013; Bauldry et al. 2016). Given all of this, it is quite surprising how little attention this dimension of inequality receives in sociology relative to other status characteristics such as race/ethnicity and gender.

This research builds upon and substantiates the notion that “what is beautiful is good” (and the related notion that “what is ugly is bad”; see Eagly et al. 1991), which goes back to empirical studies following the lead of Dion (1972) and even earlier work by Gowin (1915). These studies reveal myriad psychological, socioeconomic, and social benefits of physical attractiveness starting as early as childhood (see, e.g., Jæger 2011, p. 983; Langlois et al. 2000). While individuals perceived to be physically attractive may be ascribed positive traits such as intelligent, competent, friendly, likeable, and more, those judged as physically unattractive tend to be ascribed negative traits (i.e., the inverse of the aforementioned positive traits). Beauty, then, operates as a powerful form of social status or status characteristic (see Webster and Driskell 1983).

While it is true that research on perceived physical attractiveness across the social sciences has identified a number of important relationships between perceived physical attractiveness and a wide array of outcomes, the main focus of this research has been the labor market (Hamermesh and Biddle 1994; Conley 2004; Mobius and Rosenblatt 2006; Fletcher 2009; Hamermesh 2011; Jæger 2011; Conley and McCabe 2011; Pfeifer 2012; Kuwabara and Thebaud 2017). One study, for example, finds that wage returns to perceived physical attractiveness among high school graduates is actually larger than the returns for actual ability (Fletcher 2009). While a one-standard-deviation increase in

---

McClintock (2014) finds that beauty may not, however, be exchanged in the marital market. Instead partners tend to match on physical attractiveness. This is very similar to the finding that while there may be some evidence of status exchange with respect to skin tone and marital patterns among African-Americans (lighter-skin women marrying higher-status male spouses more often than darker-skinned black women), most couples are of similar or approximate skin tones (Monk 2014).
ability is associated with 3%–6% higher wages, attractive or very attractive individuals earn 5%–10% more than average-looking individuals (Fletcher 2009, p. 322). Another study even finds that returns to perceived attractiveness unfold over the life course and are robust to a wide array of potentially relevant controls, such as educational attainment, parental background, personality traits, IQ, and so on (Scholz and Sicinski 2015).

These studies, while perhaps atypical in their analytic foci, have certain advantages over conventional studies of ethnoracial and gender wage gaps. Indeed, intense debates persist over the causes of ethnoracial and gender gaps, how to properly estimate their magnitude, and how possible shifts in their size over time have continued across generations of researchers and across multiple academic disciplines. Still, these bodies of research are mostly united in their overall conceptual and methodological approaches (Altonji and Blank 1999; Morris and Western 1999; Grodsky and Pager 2001; Browne and Askew 2005; Greenman and Xie 2008; Leicht 2008; Pager and Shepard 2008). In the main, researchers look to compare earnings between ethnoracial and gender categories, which are taken to be groups (see Brubaker 2002), using nominal measures based largely on self-identification. This is what Leicht (2008) refers to as the “group gaps” approach. While these studies are undoubtedly important and influential, the similarities in their approaches bracket out at least two important truths: (1) within-category inequalities on the labor market are steadily growing and are often larger than what obtains between categories (Morris and Western 1999; Leicht 2008) and crucially, (2) many social categories, especially race/ethnicity and gender, are fundamentally embodied (Kawakami et al. 2017).

In other words, these studies not only miss the increasing importance of intra-categorical inequality, but also how the body “gives off” signals in interaction, which cue social categories and signify social status, all of which makes the body a potentially critical locus of inequality. In so doing, common approaches sidestep the insights of Goffman’s (1963) Stigma, which is a pioneering work on the importance of the body in shaping social interactions and producing inequality, though it tends to focus on solely on negatively viewed traits. Similarly, research on status characteristics tends to highlight the importance of bodily cues (see Webster and Driskell 1983). Put bluntly, many existing studies on ethnoracial and gender wage gaps, then, render the physical body invisible and bracket out the often fundamentally interactional and relational nature of inequality (see Tomaskovic-Devey and Avent-Holt 2019). This is not a trivial issue for an area of research that often interprets its findings as being the result of interactional biases in myriad face-to-face encounters (e.g., discrimination), particularly in the context of the labor market (e.g., recruitment through social networks, job interviews, initial wage negotiations, performance evaluations for raises and/or promotions, etc.).
Still, while studies on returns to physical attractiveness on the labor market do take the important step of highlighting the role of the physical body as a fount of social status that is profoundly implicated in the production of inequality (see Webster and Driskell 1983; Ridgeway 2011; Monk 2015, 2019) and even tend to take within-category inequalities seriously, existing studies in this area tend to neglect asking whether and to what extent returns to physical attractiveness significantly vary by race, by gender, or within and between race-by-gender combinations. This study uses nationally representative data to address this important question and, consequently, bridge the gap between research on race/ethnicity and gender on the labor market and returns to physical attractiveness on the labor market.

Race/Ethnicity, Gender, and the Consequences of Perceived Attractiveness

Specifically, this study builds upon prior research on returns to perceived physical attractiveness on the labor market by taking an intersectional approach to examine whether and to what extent (1) perceived physical attractiveness is unevenly distributed by race and gender, (2) perceived physical attractiveness is associated with earnings, net of relevant controls, overall and within and across race-by-gender combinations, (3) returns to perceived physical attractiveness on the labor market significantly vary at the intersection of race and gender, and to provide context for these findings (4) compares estimates of the returns to physical attractiveness within and across race-by-gender pairings to what obtains following more conventional approaches to wage disparities (e.g., ethnoracial inequality between Blacks and Whites and gender inequality between men and women).

To be sure, some studies have attempted to adjudicate whether men or women have larger returns to physical attractiveness (see, e.g., Hamermesh and Biddle 1994; Andreoni and Petrie 2008; Hamermesh 2011; Jæger 2011; Wong and Penner 2016), but these results have gone in multiple directions, ultimately resulting in a lack of consensus on the matter. Furthermore, in many analyses, race is either not considered at all or used as a control to “account for possible confounding effects” (Fletcher 2009; Wong and Penner 2016, p. 116) given that earnings, in general, vary by race (Monnat, Raffa- lovich, and Tsao 2012). Even Hamermesh (2011), whose seminal work on beauty is perhaps the best known, only dedicates a small fraction of a page to whether returns to beauty vary by race in Beauty Pays. He claims that a lack of data allows him only to speculate that “the effects of beauty within the African-American population might be smaller [than among whites]” (Hamermesh 2011, p. 58). His remark demonstrates that little is known about how race may affect the significance of physical attractiveness on the labor
market. It also remains unclear how gender may affect the significance of physical attractiveness on the labor market.

Consequently, not much is known about whether returns to physical attractiveness on the labor market vary at the intersection of race and gender. This study seeks to shed light on this heretofore unexamined issue by using longitudinal measures of attractiveness. These data allow us to move beyond previous quantitative and qualitative research on attractiveness, which has tended to rely on “single snapshot” ratings of attractiveness.

A central contention of our study is that to the extent that social perceptions of physical attractiveness are racialized and gendered (see hooks 1996; Hunter 2002; Craig 2002; Cottom 2018), then beauty should be unequally distributed along ethnoracial and gender lines. Indeed, this is the crux of the matter when scholars decry the hegemony of Eurocentric beauty standards (see Jha 2015). After all, physical markers of race, such as relatively dark skin tone, which tends to co-vary with coarser hair types and more Afro-centric facial features, are widely perceived as aesthetically inferior to Eurocentric physical traits (see Anderson and Cromwell 1977; Keith and Herring 1991; Hill 2002; Craig 2002; Tate 2008; Cottom 2018).

These attitudes around skin color and physical aesthetics, however, we must remember, are never neutral. They are a function of power relations and ethnoracial domination. Beauty is, after all, in the eye of the beholder – and the beholder learns what to deem beautiful and what to deem ugly. The standards themselves, of course, shift over time and space. As Tate (2008) points out, Kant rigorously argues in Critique of Judgment that there is no objective basis for deeming any object beautiful, even though a degree of intersubjective agreement is sought and may be necessary for such judgments. “A judgement of beauty cannot rightfully just belong to an individual but is based in sociality. It is based in the discourses on beauty and ugliness and the embodied practices of beauty which sediment in our structures of feeling over centuries of transnational political debate and stylization. It is these dominant discourses and practices which demand assent from us as they declare this ‘beautiful’ and that ‘ugly’ through norms which delineate who will qualify as a subject of recognition within regimes of beauty truth” (Tate 2008, p. 4). Put more simply, beauty is fundamentally relational, intersubjective, and as such, supraindividual. Indeed, notions of what is or is not beautiful hinge upon “culture” (see Lizardo 2017) understood not only as discourse, but practice, manifold processes of valuation and evaluation (see Lamont, Beljean, and Clair 2014), and relatively consensual intersubjective “agreement” at the cognitive level (see DiMaggio 1997; Patterson 2014), which acts as a crucially important factor of inequality and stratification.

To be sure, this is not to suggest that the use of physical attractiveness as an important aspect of evaluation is purely subjective. Nor is it to gainsay research that has identified physical characteristics such as facial symmetry
and facial averageness, which seem to have cross-cultural relevance as markers of perceived attractiveness among adults (see Langlois et al. 2000; Todorov et al. 2015). As one influential study found, by two months of age, infants hold their gazes longer at faces that adults judge to be attractive than at faces judged to be unattractive. (Langlois et al. 2000). For many researchers this has been interpreted as evidence that some aspects of perceived attractiveness have deep evolutionary and biological roots.\(^3\) Recent studies, however, temper this claim by pointing out that infant preferences for faces may be a matter of novelty and not an aesthetic preference per se (see Rhodes et al. 2002). Similarly, evidence suggests that emotional expressions (e.g., happiness, sadness, anger, etc.), and the degree to which one’s facial attributes are prototypical or atypical for their sex or gender, may influence judgments of physical attractiveness across different cultures (Zebrowitz and Montepare 2008). Thus, it is important to recognize that some physical attributes may have cross-cultural relevance for judgments of beauty.

Nevertheless, it would be a grievous mistake not to recognize the substantial role of culture, understood as intersubjectively shared cognitive processes (e.g., categorization, biases, stereotyping, evaluative and valuative tendencies, etc.), in shaping judgments of beauty (especially the content of claims of what is or is not beautiful). That is, the content of attractiveness is indelibly shaped by culture. Consider, for example, evidence that Dominicans have a different somatic norm image than most whites, which considers a racially mixed aesthetic more attractive than a prototypically European appearance (Candelario 2007). Thus, one aspect of the role of culture on somatic norms (i.e., judgments of beauty) is that to the extent that one lives in a society with ethnoracial and gender hierarchies, notions of attractiveness are likely to be racialized and gendered. And given the racialization of beauty, one may expect that “looking the part” may be profoundly more difficult for some compared to others (Warhurst and Nickson 2001; Hunter 2002; Witz, Warhurst, and Nickson 2003; Mears 2011; Otis 2011; Hoang 2015; Cottom 2018; Stepanova and Strube 2018; Walters 2018).

After all, on average, ethnoracial minorities, especially African-Americans, are less likely to have the facial features, hair types, or skin tones that match thoroughly Eurocentric standards of beauty. This means that ethnoracial minorities, from the vantage point of members of dominant ethnoracial categories (who are much more likely to be gate-keepers in the labor market), may be significantly more likely to be perceived as unattractive than members of dominant ethnoracial categories. Consequently, certain ethnoracial minorities

---

\(^3\) Similarly, some research claims that waist-to-hip ratio (WHR) is also a key predictor of perceptions of female attractiveness that is culturally invariant (Singh et al. 2010). Still, this is a tendentious position because some researchers claim that the role WHR may play in judgments of attractiveness are a function of its confounding with weight and BMI (see Swami et al. 2006).
may face discrimination on the basis of their ethnoracial category and stark penalties for not being perceived as possessing this valued form of capital (i.e., beauty as bodily capital) on the labor market, to the extent that they do not conform to hegemonic aesthetic norms.

These dimensions of social difference may interact multiplicatively, yielding complex patterns of inequality. Given the racialization and gendering of perceived beauty, we should expect such interactions. In short, while Black men may face *double jeopardy* on the labor market (race and beauty), Black women may face *triple jeopardy* (race, gender, and beauty). Furthermore, to the extent that Latinx (or other ethnoracial minorities) are darker skinned and have other physical traits that do not adhere to hegemonic standards of beauty (see Hunter 2002), they too should face penalties on the labor market. This study directly considers this possibility by adopting an intersectional approach (see Crenshaw 1989, 1991; Collins 2002; Browne and Misra 2003; McCall 2005; Hancock 2007; Glenn 2009) to returns to physical attractiveness on the labor market by running fully interactive statistical models and stratified race-by-gender models (see the appendix) that formally examine the extent to which returns to perceived physical attractiveness are contingent upon the combination of race and gender, net of relevant controls (for more on quantitative approaches to intersectionality, see Hancock [2013], Bauer [2014], and Bowleg and Bauer [2016]).

In concert with existing research, we find that perceived physical attractiveness is a significant factor of inequality regardless of race and gender—a fact that further emphasizes its importance as a dimension of inequality that is worthy of further scholarly attention. In fact, we find that the magnitude of the earnings disparity among white men along the perceived attractiveness continuum rivals or perhaps, exceeds the canonical Black-White wage gap in magnitude; and the earnings disparity among White women along this same continuum is larger than the Black-White wage gap using these same data. These findings add yet more evidence to our contention that perceived

---

4 Some of the earliest uses of the notion of double or multiple jeopardy with respect to the consequences of intersecting dimensions of social difference (e.g., intersectionality) have been traced to Frances Beale’s ([1970] 2008) “Double Jeopardy: To Be Black and Female” and Deborah K. King’s (1988) “Multiple Jeopardy, Multiple Consciousness: The Context of a Black Feminist Ideology.” While the language of multiple or double jeopardy may sound additive, in the context of statistical analysis, many researchers agree that capturing double or multiple jeopardy requires methods that are *not* additive (e.g., interaction models, multilevel models, hierarchical models, ecological models, or fuzzy set methods; see McCall 2005, p. 1788; Hancock 2007). As one prominent researcher in this area puts it, “the very language used in intersectionality can create confusion for quantitative researchers” (Bauer 2014, p. 12). After all, theories of intersectionality were not necessarily crafted with quantitative research in mind (for more on quantitative approaches to intersectionality, see Hancock [2013] and Bowleg and Bauer [2016]).
physical attractiveness is a powerful, yet relatively neglected, dimension of inequality and stratification worthy of much more scholarly attention.

Still, with respect to the novel questions we address, which turn around whether perceived attractiveness is intersectional, not only do we find that perceived physical attractiveness is unequally distributed by race and gender but returns to perceived physical attractiveness also significantly vary at the intersection of race and gender. Black men and women face especially stark penalties on the basis of their perceived physical attractiveness (e.g., double and triple jeopardy)—so much so that the magnitude of the wage gaps between those rated the least and most attractive within these race-by-gender groups exceeds both the Black-White wage gap and the gender gap in earnings. These penalties persist after controlling for family background, their own educational attainment, occupational status, BMI, and even respondents’ skin tone. Thus, given that “lookism” (i.e., discrimination based on physical appearance), unlike ethnoracial and gender discrimination, is legal (see Rhode 2010), perceived physical attractiveness may be a technically legal channel through which ethnoracial inequality is maintained on the labor market, in addition to lookism’s consequences for everyone regardless of their race and/or gender.

For the relatively small number of Black men and women who are perceived to be highly physically attractive, however, there are great rewards. In fact, Black women judged to be highly physically attractive find that their earnings converge and, perhaps, cross-over with that of white women of the same level of perceived physical attractiveness. A finding quite similar to a study, which reported that the wages of light-skinned Blacks were virtually indistinguishable from whites (Goldsmith, Hamilton, and Darity 2007). As we explain in the following section, this, perhaps, counterintuitive dynamic is mostly anticipated by Bourdieu’s (1984) formulation of capital, a promising theoretical approach to conceptualizing perceived physical attractiveness on the labor market and as a dimension of inequality and stratification in general. Next, we turn our attention to explicating this theoretical approach.

BEAUTY AS BODILY CAPITAL

Currently, one of the most prominent approaches to capturing the significance of perceived physical attractiveness is conceptualizing it as a form of capital (e.g., Mulford et al. 1998, Martin and George 2006, Mears 2011, and Bauldry et al. 2016). The most compelling of these approaches turn to and develop Bourdieu’s (1984) notion bodily capital and sidestep the anti-relational approach of Hakim’s (2010, 2011) erotic capital, which “completely glosses over how the structures of race, class, age, and context—the structures of a field—systematically organize what constitutes “desirability,” that is, the political
The Social Psychology of Beauty as Bodily Capital

Though this remains underdeveloped in Bourdieu’s conception, there are important social psychological insights that can help us make sense of how bodily capital operates. First and foremost, social psychological research convincingly shows that we draw inferences from bodily cues about people’s intelligence, competence, warmth, “moral uprightness,” and much more (see Kawakami et al. 2017; Zebrowitz 1996; Todorov et al. 2015). One of the key social psychological phenomena at the heart of the efficacy of bodily capital is the “halo” effect, where we tend to impute myriad positive traits to those perceived to be physically attractive. Concomitantly, negative traits are imputed to those who are perceived to be physically unattractive. Put differently, from the vantage point of “expectation states” theory, we have certain expectations of others’ competence and future behaviors linked to their physical appearance (again, see Webster and Driskell 1983).

Ultimately, these inferences (e.g., stereotypes and heuristics), drawn from social perceptions of bodies, guide our interactions with others in myriad social contexts, which indelibly shape processes of evaluation (Lamont 2012) within and across key organizations and institutions (e.g., the labor market, the criminal justice system, education systems, marital market, electoral...
politics, etc.). Crucially, this often takes place beneath the level of explicit consciousness and is hard to suppress even when one tries to, which results in cognitive biases that advantage some while disadvantaging others (Zebrowitz 1996).

As people who are perceived as physically attractive tend to be advantaged across a plethora of interpersonal exchanges (see Mulford et al. 1998), we should expect that insofar as the labor market is concerned, this form of bodily capital should act as a vital resource shaping social networks that are important in initial recruitments into various occupations and sorting into hierarchical positions within occupations (Smith 2007), affecting the success of job interviews (see Rivera 2015), influencing performance evaluations linked to raises and promotions (Elvira and Town 2001; Tomaskovic-Devey and Avent-Holt 2019), and enhancing (or detracting from) wage negotiations.

In other words, those who enjoy “larger” volumes of this form of capital may enjoy increased earnings through multiple pathways on the labor market and throughout their careers (Scholz and Sicinski 2015). Overall, bearers of high volumes of bodily capital are likely to be advantaged in relational claims-making: “the discursive articulation of why one actor is more deserving of organizational resources than others” (Tomaskovic-Devey and Avent-Holt 2019: 188). This not only includes their own claims-making (e.g., job applications, requests for raises, etc.), but also the probability of having powerful and influential actors within and across organizations seek to employ them, promote them, and channel them resources. Indeed, bearers of large volumes of bodily capital, a crucial symbolic resource (see above), are more likely to have claims-making on their behalf be considered legitimate (Tomaskovic-Devey and Avent-Holt 2019, pp. 191–96).

Still, as prior research has convincingly shown, race and gender also play important roles on the labor market. Unfortunately, however, we know very little about the conjoint consequences of race, gender, and attractiveness on the labor market. Namely, it is unclear whether the significance of perceived physical attractiveness as a form of bodily capital is itself contingent upon these other important factors (e.g., race and gender). In other words: is the significance of beauty as a form of bodily capital on the labor market, intersectional? As we explain in the following section, given that beauty itself is intersectional (see hooks 1996; Craig 2002; Hunter 2002; Cottom 2018; Jha 2015), and [bodily] capital is a relational good, we should indeed expect this to be the case.

RELATIONALITY, INTERSECTIONALITY, AND RETURNS TO BODILY CAPITAL ON THE LABOR MARKET

The inherent relationality of capital is central to both Marx’s and Bourdieu’s take on the concept and the latter’s oeuvre in general. Following Marx’s
classic formulation (Marx 1894, p. 590), however, “capital is not a thing, but rather a definite set of social relations which belong to a definite historical period in human development, and which give the things enmeshed within these relations their specific content as social objects.” As Bourdieu (1990, p. 126) explains, “One has to break away from the mode of thought that Cassirer calls substantialist, and which leads people to recognize no realities except those that are available to direct intuition in ordinary experience, individuals and groups.” The major contribution of what one has to call the structuralist revolution consisted in applying to the social world a relational way of thinking, which is that of modern physics and mathematics, and which identifies the real not with substances but relations (Bourdieu and Wacquant 1992).

In other words, capital locates individuals in certain positions or points in the broader social space and/or within fields—in relation to others (Bourdieu 1984). The “value” of a particular form of capital fluctuates and shifts depending on the context (e.g., characteristics of the perceiver, social settings, etc.) much like Goffman’s (1963) rendering of stigma. Moreover, and this is especially critical for this study, the value of capital is often contingent upon its bearer (Bourdieu 1986). Indeed, “Educational qualifications [institutionalized cultural capital] never function perfectly as currency. They are never entirely separable from their holders: their value rises in proportion to the value of their bearer” (Bourdieu 1986, p. 29). Likewise, bodily capital is unlikely to function perfectly as currency; instead its value is likely to be inextricably linked to the value of its bearer. Here “value” can be thought of as referring to whether these particular agents belong to stigmatized or dominant social categories. Members of stigmatized social categories, by definition, are likely to be relatively deprived of symbolic capital regardless of their overall portfolio of capital—this, of course, is one of the hallmarks of stigmatization’s insidiousness (see Lamont 2016).

The Racialization and Gendering of Beauty as Bodily Capital

The generalized dishonor faced by members of a stigmatized category canonically described by Max Weber ([1918] 1978; see also Wacquant 2005; Brubaker 2015, p. 46) is what so much research is essentially referring to when it highlights the role of gender or ethnoracial discrimination as a cause of social inequality and stratification via social closure and the operation of social boundaries (Lamont and Molnar 2002; Wimmer 2008). By virtue of belonging to a stigmatized social category, members of the category are already much more likely to be prejudged or stereotyped as lacking competence, intelligence, trustworthiness, and so on (Fiske et al. 2002; Massey 2007). Given that, conceptually, bodily capital is especially suited to act as naturalized and
legitimate indicia of honor (i.e., symbolic capital), it stands to reason that the value of beauty (as a form of bodily capital) is quite likely to interact with properties that already signify honor and/or dishonor.

This means that the distribution of attractiveness as a form of bodily capital is likely to be unequal across social categories—especially ethnoracial categories given the dominance of Eurocentric beauty ideals (Craig 2002; Cottom 2018). Indeed, as Mears (2014, p. 1337) compellingly explains, “assessments of good looks are grounded in power relations, which [leads to] assigning differential value to workers according to interlocking social positions”—that is, race/ethnicity and gender. Social judgments of physical attractiveness are indelibly shaped by notions of masculinity and femininity, which are themselves indelibly shaped by the complex imbrication of ethnoracial and gender stereotypes (Hill 2002; García and Abascal 2015; Kuwabara and Thebaud 2017). After all, evidence suggests that the social perception of race and gender and are inextricably linked down to the bedrock of human cognition—as some researchers have put it, “race is gendered”: “race and sex categories are psychologically and phenotypically confounded” (Johnson, Freeman, and Pauker 2012).

A result of this profound confounding is that Black women may be especially unlikely to be perceived as physically attractive by Whites given pervasive stereotypes that view all African-Americans, regardless of their gender, as relatively more masculine (and dangerous) than members of other ethnoracial categories (see Eberhardt et al. 2004; Eberhardt 2005; Galinsky, Hall, and Cuddy 2013). In other words, to the extent that stereotypical traits associated with blackness (e.g., dark skin tone, Afrocentric facial features, etc.) are read as inherently masculine, then Black women, especially darker-skinned Black women, may be penalized for being perceived as not feminine enough. We must keep in mind the role of sex/gender prototypicality in judgments of attractiveness—appearing “feminine” and appearing “masculine” is associated with attractiveness for women and men, respectively.

Thus, on one hand, adherence to hegemonic, Eurocentric aesthetic norms about facial features, skin tone, and hair may cause Black men to be seen as relatively unattractive. On the other hand, however, the link between dark skin and perceived masculinity may also mean that relatively dark-skinned Black men, in particular, and Black men, in general, may not be as harshly penalized with respect to perceptions of perceived attractiveness compared to Black women (for compelling evidence of this possibility, see Hill [2002]). In fact, one study finds that British women rated Black male faces as more attractive than white male faces (Lewis 2011).

Ultimately, then, Hamermesh’s (2011) hypothesis that skin tone may be treated similarly to beauty (see also Conley 2004), even though it is not reducible or equivalent to beauty (see Stepanova and Strube 2018), is especially relevant and compelling here. A dominant interpretation of skin tone

Beholding Inequality

legitimate indicia of honor (i.e., symbolic capital), it stands to reason that the value of beauty (as a form of bodily capital) is quite likely to interact with properties that already signify honor and/or dishonor.

This means that the distribution of attractiveness as a form of bodily capital is likely to be unequal across social categories—especially ethnoracial categories given the dominance of Eurocentric beauty ideals (Craig 2002; Cottom 2018). Indeed, as Mears (2014, p. 1337) compellingly explains, “assessments of good looks are grounded in power relations, which [leads to] assigning differential value to workers according to interlocking social positions”—that is, race/ethnicity and gender. Social judgments of physical attractiveness are indelibly shaped by notions of masculinity and femininity, which are themselves indelibly shaped by the complex imbrication of ethnoracial and gender stereotypes (Hill 2002; García and Abascal 2015; Kuwabara and Thebaud 2017). After all, evidence suggests that the social perception of race and gender and are inextricably linked down to the bedrock of human cognition—as some researchers have put it, “race is gendered”: “race and sex categories are psychologically and phenotypically confounded” (Johnson, Freeman, and Pauker 2012).

A result of this profound confounding is that Black women may be especially unlikely to be perceived as physically attractive by Whites given pervasive stereotypes that view all African-Americans, regardless of their gender, as relatively more masculine (and dangerous) than members of other ethnoracial categories (see Eberhardt et al. 2004; Eberhardt 2005; Galinsky, Hall, and Cuddy 2013). In other words, to the extent that stereotypical traits associated with blackness (e.g., dark skin tone, Afrocentric facial features, etc.) are read as inherently masculine, then Black women, especially darker-skinned Black women, may be penalized for being perceived as not feminine enough. We must keep in mind the role of sex/gender prototypicality in judgments of attractiveness—appearing “feminine” and appearing “masculine” is associated with attractiveness for women and men, respectively.

Thus, on one hand, adherence to hegemonic, Eurocentric aesthetic norms about facial features, skin tone, and hair may cause Black men to be seen as relatively unattractive. On the other hand, however, the link between dark skin and perceived masculinity may also mean that relatively dark-skinned Black men, in particular, and Black men, in general, may not be as harshly penalized with respect to perceptions of perceived attractiveness compared to Black women (for compelling evidence of this possibility, see Hill [2002]). In fact, one study finds that British women rated Black male faces as more attractive than white male faces (Lewis 2011).

Ultimately, then, Hamermesh’s (2011) hypothesis that skin tone may be treated similarly to beauty (see also Conley 2004), even though it is not reducible or equivalent to beauty (see Stepanova and Strube 2018), is especially relevant and compelling here. A dominant interpretation of skin tone
stratification (or colorism) among African-Americans is that light skin is a relatively rare physical attribute among African-Americans, which signifies their proximity to whiteness within a context of ethnoracial domination (Goldsmith et al. 2007; Glenn 2009; Cottom 2018). This trait is highly valued by Whites who are the most likely to serve as gatekeepers on the labor market) and many Blacks as well (again, there is often widespread agreement on these aesthetic judgments, which may itself be a form of symbolic violence; see Monk 2015). Still, as we explain next, there is compelling reason to believe that Hamermesh may not be correct that beauty may matter less to Blacks on the labor market.

Unequal Distribution, Unequal Returns: The Relationality of Bodily Capital

Why might Hamermesh (2011) be wrong? Precisely because the distribution of attractiveness may be unequal across social categories, returns to physical attractiveness may actually be greatest among the most stigmatized given the rarity of them being perceived as highly physical attractive. One way of thinking about this is through the notion typicality (see Wittgenstein 1953; Schutz [1962] 1982; Rosch and Mervis 1975; Rosch 1978; Maddox 2004; Monk 2015). Members of stigmatized categories who manage to be perceived as physically attractive by members of dominant social categories (i.e., hold large volumes of bodily capital) are atypical to their stigmatized category and may be greatly rewarded for “standing out” or serving as an “exception” to the myriad stereotypes associated with their category (see, e.g., Maddox 2004; Ridgeway and Kricheli-Katz 2013; Pedulla 2014).

Walters (2018), for instance, details how white women are generally privileged by retail managers, but when retail managers do seek to “diversify” their brands they do so by seeking out the lightest-skinned, most racially ambiguous, and thus, highly atypical ethnoracial minorities. The general dynamic at work here, again, is similar to colorism, a literature that emphasizes the advantages of atypical minorities (e.g., lighter-skinned African-Americans) across several outcomes, from educational attainment to the criminal justice system (see Monk 2014, 2015, 2019). As this literature demonstrates, African-Americans who possess phenotypical cues of “whiteness,” are often advantaged relative to other African-Americans who are far more prototypical of this stigmatized category (Maddox 2004). Visibly atypical African-Americans possess an important “countersignal” against the stereotypes and inferred negative traits that membership in the stigmatized category may usually “signal.” These traits, in their racialized and gendered evaluation by alters, distance them in terms of their perceived similarity from other members of the stigmatized category and modulate the frequency and/or harshness of the stereotypes they face.
Beholding Inequality

With respect to the labor market, then, atypical minorities may be relatively more likely than other members of their category to integrate majority-white spaces and workplaces. Again, given that skin tone is deeply linked to perceptions of attractiveness (see Hill 2002) and the notion that attractiveness is racialized, the general mechanisms behind the consequences of color and attractiveness are much the same. Both skin tone and perceived physical attractiveness are forms of bodily capital linked to profits for those relatively atypical and rare ethnoracial minorities who possess it in high volumes. In short colorism and lookism are intimately linked; consequently, one way of thinking about this study is that it considers the conjoint consequences of lookism and colorism on the labor market.

The potentially heightened profits for atypical minorities are anticipated by Bourdieu’s (1986) theory of capital(s). As he (Bourdieu 1984, p. 19) explains, the profit produced by capital is “mediated by the relationship of (objective and/or subjective) competition between himself and the other possessors of capital competing for the same goods, in which scarcity—and through it, social value—is generated.” Thus, while the vast majority of members of the stigmatized category will face steep penalties for not “living up to” racialized and gendered aesthetic norms, those few members of that same category who are deemed to be beautiful—especially from the vantage point of dominant members of society who set and maintain hegemonic aesthetic norms—may see substantial profits for possessing high volumes of bodily capital relative to their compatriots.

Thus, we should expect that returns to perceived physical attractiveness should be strongest among African-Americans, contra Hamermesh’s (2011) conjecture. That being said, most category members should see significant wage penalties even after adjusting for their educational credentials and family backgrounds, to the extent that bodily capital is an important factor of inequality and stratification on the labor market. To be sure, beauty as bodily capital should, as previous literature suggests, be significantly associated with wages regardless of race and gender. This is true even though beauty premia and penalties are likely to be more muted among members of dominant ethnoracial categories. Nevertheless, even inequalities among members of dominant ethnoracial categories are likely to be considerable given that within-category inequalities are increasing and often rival what obtains between categories (Leicht 2008).

By foregrounding how race and gender combine with physical attractiveness to affect wages on the labor market we adopt a thoroughly intersectional approach (Crenshaw 1989, 1991; Collins 2002; Glenn 2009; Browne and Misra 2003; McCall 2005; Hancock 2007; Greenman and Xie 2008; Choo and Ferree 2010). Further, by examining the significance of race, gender, and beauty within and across social categories as opposed to only between them (i.e., the “group gaps” approach) we adopt what McCall (2005,
The categorical approach focuses on the complexity of relationships among multiple social groups within and across analytical categories and not on complexities within single social groups, single categories, or both. The subject is multigroup, and the method is systematically comparative. Unlike single-group studies, which analyze the intersection of a subset of dimensions of multiple categories, however, multigroup studies analyze the intersection of the full set of dimensions of multiple categories and thus examine both advantage and disadvantage explicitly and simultaneously. It is not the intersection of race, class, and gender in a single social group that is of interest but the relationships among the social groups defined by the entire set of groups constituting each category. (McCall 2005, p. 1786; emphasis added)

The analyses below follow this approach by estimating stratified race-by-gender models (see the appendix; for a similar approach to quantitative intersectional analysis, please see Hancock [2013]), in addition to fully pooled models with interactions for attractiveness, race, and gender. In so doing, we seek to render comprehensively how perceived attractiveness, comparatively, shapes labor market outcomes within and across ethnoracial and gender categories. Before turning to the analyses, however, next we will describe our data, measures, and analytic plan.

DATA

Data come from the National Longitudinal Study of Adolescent to Adult Health (Add Health). Add Health is an ongoing, nationally representative survey of a group of individuals who were in grades 7–12 in 1994 (Harris et al. 2009). In the original sample, 20,745 individuals from 132 U.S. schools were selected for in-depth home interviews. Add Health currently includes four waves of data (wave 1/1994–95; wave 2/1996; wave 3, 2001–2; wave 4, 2007–8). In wave 4, all respondents were between 24 and 32 years old. The ethnoracial diversity of Add Health’s sample, in addition to its rich information on SES and physical appearance, make these data useful for our analysis. Our initial analytical sample is composed of respondents who appeared in all four waves of the data and had valid sample weights (n = 9,345).

Measurement

Our response, earnings, encodes a respondent’s annual income. At wave 4 individuals were asked, “Over the past year, how much income did you receive from personal earnings before taxes, that is, wages or salaries, including tips, bonuses, overtime pay, and earnings from self-employment?”
Responses were taken as a measure of earnings, in dollars. Note that we winsorize earnings at $200,000 and $1,000 per year to lessen the sensitivity of results to extreme values. This results in approximately 7% of individuals with top or bottom coded incomes. Substantive conclusions are robust to varying these cut-points.

Our focal predictor, *socially perceived physical attractiveness*, is constructed from information provided by Add Health interviewers. At each wave of the data, interviewers were asked to evaluate the physical attractiveness of the respondent. Possible ratings included very unattractive, unattractive, average, attractive, and very attractive. Interviewers were given little to no explicit training about *how* to judge attractiveness, as the measure is designed to capture subjective, social perceptions of respondents’ beauty. This follows decades of prior research, including that of labor economists (see Mobius and Rosenblat 2006; Hamermesh 2011; Pfeifer 2012), which have used similar methods to measure perceived attractiveness. In other words, interviewer ratings help proxy processes of bias and discrimination that survey respondents may face on the labor market and other realms of life.

Toward that end, it is important to note that the vast majority of interviewers (with measured demographic characteristics) are White (just under 70%), and female. Furthermore, the sample of interviewers was highly educated, with 20% having a postgraduate degree, 28% having a college degree, and 31% having some college experience. Again, this interviewer pool represents actors that respondents may typically encounter as gatekeepers in the labor market. This is helpful for the purposes of our study.

Still, even though most interviewers are White, given the potential relationality of perceptions of attractiveness some may wonder about potential differences in ratings across race categories and/or own-race biases in attractiveness ratings. To assess potential differences in ratings across race categories and/or own-race biases in attractiveness ratings we ran a Bayesian multilevel cumulative logit regression model (see Bürkner and Vuorre 2019) of W4 attractiveness ratings (results available by request). An interviewer-specific (random) intercept was included to account for the fact that ratings are clustered within interviewers. We find overwhelming consistency in the attractiveness scores given by White women and White men. The only difference being that thesmen rated respondents slightly lower overall, with even heavier penalties against Black women. This lines up with our main findings—Black women, in particular, face harsh penalties on the labor market with respect to perceived attractiveness in a labor market dominated by White gatekeepers.

Most of the Black interviewers were women. While Black female interviewers’ ratings were mostly consistent with White female interviewers, Black female interviewers appeared to give slightly lower ratings overall than
White women (except for when evaluating Hispanic respondents). The small number of Black male, Hispanic male, and Hispanic female interviewers makes the uncertainty estimates for their predicted ratings much wider. Still, one thing is clear: their ratings are mostly consistent with white interviewers with the only exception being that Black male interviewers tended to give lower scores to male respondents regardless of their race/ethnicity.

Overall, then, there seems to be overwhelming consensus on ratings of attractiveness across interviewers regardless of race/ethnicity and gender. Furthermore, our analyses reveal virtually no evidence of same-race bias in their attractiveness ratings. Thus, taken together, even though these measures are subjective and there is some variation, interviewers exhibit consensus, which is in line with the findings of research on judgments of attractiveness (see Zebrowitz and Montepare 2008; Hamermesh 2011). The parallels between lookism and colorism appear even more striking—after all, one reason why colorism persists is due to relative agreement on the value of light and dark skin across ethnoracial categories (see Monk 2015).

While perceived attractiveness ratings do correlate within respondents in Add Health—such that individuals coded as attractive in one wave are likely to be coded as attractive in other waves—there is enough fluctuation in scores across waves to dissuade us from representing this construct with a single interviewer’s assessment. Following the lead of prior studies that have used Add Health’s attractiveness measures, we use each physical attractiveness indicator in conjunction with one another to reduce measurement error and extract a more stable measure of each interviewee’s overall socially perceived physical attractiveness trait (Bauldry et al. 2016).

The underlying assumption made here is that individuals assessed as attractive by four separate interviewers are significantly more likely to possess traits that highly correspond with notions of socially perceived physical beauty, while individuals consistently assessed as unattractive across waves are more likely to possess traits that are highly discordant with hegemonic notions of perceived beauty. To be clear, however, the idea here is not to estimate “objective” attractiveness, but rather to have a relatively stable estimate of a respondent’s socially perceived, and necessarily subjective, physical attractiveness.

To combine attractiveness scores, we—similar to Bauldry et al. (2016)—fit a structural equation

---

5 From wave 1 to wave 2, e.g., approximately 50% of respondents were perceived as the same level of attractiveness on this 5-point scale; 40% of respondents’ ratings deviated by only one category, while the remaining 10% were assessed as 2 or more points away from their original score. In other words, some 90% of the respondents were assessed as having roughly the same level of perceived attractiveness (+/- 1 point). See the appendix for more on how perceived attractiveness scores overlapped across waves.
Beholding Inequality

model where each wave’s attractiveness indicator loads onto a single, latent perceived attractiveness construct. (This measurement model fits the data well, with a root mean square error of approximation of 0.054 and a comparative fit index of 0.980). Scores estimated from this model range from approximately −1.3 (i.e., least likely to be assessed as attractive by interviewers) to +1.1 (i.e., most likely to be assessed as attractive by others). For more information, please see the appendix.

Given the use of multiple waves of data one may wonder about differences in the interviewer characteristics across waves. We find that despite the fact that Add Health recruited an almost entirely new set of reviewers between waves (e.g., 72% of wave 4 interviewers did not participate in wave 3), demographic characteristics of the raters at wave 3 were largely the same as those at wave 4. While we cannot be sure that the characteristics presented above represent interviewers at waves 1 and 2, the fact that Add Health recruited a largely similar pool of individuals seven years apart (wave 3 in 2001; wave 4 in 2008) gives us some indication of the type of interviewers that Add Health targeted across waves. Still, we acknowledge that not having concrete information on wave 1 and 2 interviewers is a limitation.

Finally, it is also important to acknowledge that a limitation of the interviewer ratings of perceived physical attractiveness in Add Health is that they are also potentially influenced, at least in part, by nonverbal cues, vocal cues, prior answers to questions on the survey, and home environment cues. While our extensive control variables, robustness checks, and use of multiple perceived attractiveness scores over time should help mitigate this issue to some extent, the physical attractiveness ratings we use are arguably less “pure” than what may obtain in an experimental setting. The trade-off here, however, is that our core analyses on the potentially intersectional consequences of perceived attractiveness are possible because of the rich set of variables and covariates available in Add Health that are typically unavailable in experiments. In the next section we discuss some of these important sociodemographic variables.

Sociodemographic Measures

In addition to socially perceived attractiveness, sex category and race/ethnicity are key features in our analysis. Sex category, a self-reported measure of one’s sex, is coded as either male or female. Race, a measure of one’s self-reported race/ethnicity, is coded as either non-Hispanic White; non-Hispanic Black; Hispanic; Asian; Native American/American Indian; or Other. Note that self-reported race/ethnicity is taken from the wave 1 in-home interview file. Results that operationalize race using other self-reports available in Add Health (e.g., wave 3 self-reported race) yield similar substantive conclusions.
American Journal of Sociology

to what is presented below (results available by request). Given their small sample sizes, Asian; Native American/American Indian; and Other individuals are excluded from the present analysis.

To assess the association between socially perceived physical attractiveness and income net of human capital, family background, and other relevant factors that may influence earnings, we include a number of additional features as controls. Covariates include wage 4 age (in years); U.S.-born status (born in the United States; or born elsewhere), education level of the highest educated parent/guardian in a respondent's household at wave 1 (either less than a high school diploma; high school diploma or some college; or a college degree), wage 4 educational attainment (either less than a high school diploma; high school diploma or some college; or a college degree), wage 4 occupational status (1 if working in a managerial/professional occupation; 0 otherwise), wage 4 marital status (1 if married; 0 otherwise), body mass index (calculated from respondents’ measured heights and weights; see Conley and McCabe 2011), and interviewer assessed skin tone (light, medium, or dark).

Because we are primarily interested in understanding how earnings are generated among individuals in the workforce, we restrict the sample according to wage 4 employment status (1 if working at least 10 hours per week; 0 otherwise). Note that respondents with missing information on covariate values (approximately 5% of the data) were excluded from the analysis. Our final analytical sample consist of 6,090 working, Black, White or Hispanic individuals.

Analytic Plan

If, net of controls, perceived physical attractiveness is predictive of earnings, a model of earnings that includes a measure of physical attractiveness should fit the data better than a model that does not. To assess whether this is the case for our data, we compare (1) the Akaike’s information criterion (AIC; Long and Freese 2001) of a model that regresses earnings on the set of controls listed above with (2) the AIC of a model that regresses earnings on the same controls, as well as physical attractiveness. If the latter model fits the data better, as indicated by a lower AIC score, we will have evidence that physical attractiveness is relevant in predicting earnings.

Similarly, if the strength of the association among attractiveness and earnings varies at the intersection of race and gender, models of earnings that allow for race-by-sex specific attractiveness slopes should fit the data better than a model that constrains attractiveness to have the same effect for all groups. To assess whether this holds for our sample, we fit (1) a model that regresses income on attractiveness, race-by-gender social categories,
and controls and (2) the same model with additional interaction terms that allow for each race-by-gender group to have discrete attractiveness slopes. Again, we use AICs to discern between these specifications.

To estimate the models needed for this analysis, we use a gamma generalized linear model (GLM) with a log link-function. We choose this model given that our response is continuous, nonnormal—in that it is always positive—and right skewed (Faraway 2016). Note that choosing a log-link function means that the exponentiated model can be interpreted as multiplicative effects. All analyses are weighted using Add Health’s sampling weights and R’s survey package (Lumley 2004).

RESULTS

Tables 1 and 2 display weighted descriptive statistics for our analytical samples. Note that personal earnings in Add Health generally correspond to national income patterns: White males have the highest average levels of earnings, while Black females have the lowest. Earnings among Hispanic individuals are, relative to Whites, somewhat higher than national estimates—likely because of the limited age and perhaps because of the relatively light skin tone of most Hispanics individuals sampled in Add Health. Consider that over 80% of Hispanic respondents have skin tones that are

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>DESCRIPTIVE STATISTICS, MALE SAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>VARIABLE</td>
<td>WHITE</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>Income ($)</td>
<td>42,746</td>
</tr>
<tr>
<td>Perceived attractiveness score</td>
<td>-0.06</td>
</tr>
<tr>
<td>Age (years)</td>
<td>28.8</td>
</tr>
<tr>
<td>College educated</td>
<td>.33</td>
</tr>
<tr>
<td>Born in U.S.</td>
<td>.99</td>
</tr>
<tr>
<td>Currently married</td>
<td>.41</td>
</tr>
<tr>
<td>Works in a professional occupation</td>
<td>.36</td>
</tr>
<tr>
<td>Parent with a college degree</td>
<td>.41</td>
</tr>
<tr>
<td>Light skin tone</td>
<td>.99</td>
</tr>
<tr>
<td>Medium skin tone</td>
<td>.01</td>
</tr>
<tr>
<td>Body mass index</td>
<td>28.7</td>
</tr>
</tbody>
</table>

NOTE.—Means of binary/categorical variables encode proportions (e.g., Born in U.S. gives the proportion of respondents that were born in the United States; Light skin tone gives the proportion of respondents with light skin tones, as opposed to medium or dark skin tones). N is White = 1,866, Black = 482, Hispanic = 519.
Race, Gender, and the Distribution of Perceived Physical Attractiveness

Additional descriptive analysis (see fig. 1) shows that socially perceived attractiveness is, as hypothesized, unevenly distributed across race and gender (see hooks 1996; Craig 2002; Tate 2008). Among male respondents, negligible racial differences exist in the distribution of socially perceived attractiveness: male sample populations of all races had average perceived attractive scores of approximately $-0.05$. Among women, however, significant ethnoracial differences in scores were apparent: White and Hispanic women displayed higher average perceived attractiveness ratings (mean score = 0.11 and 0.13, respectively) than Black women (mean score = $-0.05$). Additionally, while approximately 30% of Black females had ratings in the highest one-third of socially perceived attractiveness scores, 43% of White and Hispanic females had scores in this highest tertile. Consonant with prior studies (Hill 2002; Stepanova and Strube 2018) and serving as evidence of the racialization of beauty, further tests reveal that skin tone is a significant predictor of perceived attractiveness (app. table A5).
Returns to Physical Attractiveness

For an initial look at the relationship among income and perceived attractiveness, figure 2 plots earnings against attractiveness scores, with a linear line of best fit, for each race-by-gender group. Overall, this plot demonstrates a general, positive association among earnings and attractiveness: regardless of race and gender, individuals with higher socially perceived attractiveness scores appear to have higher annual incomes than their peers. Nevertheless, especially steep returns to—or penalties for lacking—perceived attractiveness are observed among Black women and men. In fact, for Black women, returns to attractiveness are so pronounced that, at the highest levels of perceived attractiveness, Black women’s earnings converge with, and ultimately cross over with White and Hispanic women’s earnings.

To assess whether the associations implied by figure 2 are significant and robust to the inclusion of controls, we next fit the gamma regression models of earnings described above. Table 3 provides coefficient estimates from models of the full sample. (Note that these models control for an array of factors that may influence earnings and/or be associated with perceived attractiveness. A number of these variables, such as educational attainment or marital status, may be consequences of/effect of perceived attractiveness—and thus might be more properly thought of as post-treatment variables.) In the appendix, we specify models that avoid conditioning on post-treatment variables and thus the potential bias that might arise from doing so. Substantive results from this sensitivity analysis are consistent with the findings presented below.
Parameter estimates from table 3 reinforce the idea that socially perceived attractiveness is a salient predictor of earnings. The $P$ value associated with the attractiveness parameter is significant by conventional standards and, perhaps more important, the AIC of a model that includes an attractiveness term is roughly 19 points lower than a model that excludes AIC. Substantively, our model suggests that moving up 1 point on the socially perceived attractiveness scale—for example, from the sample median score (approximately 0.00) to the 99th-percentile score (approximately 1.00)—multiplies earnings by a factor of $\exp(0.15) \approx 1.16$ (e.g., from $30,000$ to $34,800$). Similarly, a 1-SD increase in perceived physical attractiveness (approximately 0.40 points) multiplies expected annual earnings by a factor of 1.06 (e.g., $30,000$–$31,800$ per year).

The next set of results examines whether the association among perceived physical attractiveness and earnings varies across race and gender subpopulations (see table 4). In model 1, perceived attractiveness is constrained to have a uniform association with earnings among all race-by-gender groups in the data. In model 2, each race-by-gender group is allowed to have distinct attractiveness slopes.

Table 4 shows that the strength of the association among earnings and attractiveness varies across race-by-gender groups. Including an interaction term that allows for perceived attractiveness slopes to vary by race and gender leads to a modest improvement in model fit over a specification that constrains effects to be equal across groups. This modest increase in model fit appears to be driven by the freeing of slopes among Black respondents. While most groups’ attractiveness parameters are indistinguishable
from slopes among White males (the reference group in this case), both Black males and Black females have slopes that are significantly different at a 0.05 level. Indeed, where a 1-point increase in perceived attractiveness multiplies earnings by roughly $\exp(0.10) = 1.10$ among White males, the same 1-point change in attractiveness score yields incomes that are approximately 1.43 and 1.44 times higher among Black females and males, respectively.

For additional clarity on how the association among earnings and perceived attractiveness varies across groups, next, we use the parameters from our interactive model to calculate expected earnings for different hypothetical individuals. Holding all other variables in the model at their means/modes, we predict earnings for individuals from the 5th to the 95th percentile of perceived attractiveness scores. (Note that because skin tone is highly correlated with race, we set skin tone to its groups-specific modes for these predictions.) Figure 3 plots predicted values, while table 5 provides the ratio of predicted earnings across several scenarios of interest.

In both figure 3 and table 5, we observe sizable differences in predicted earnings between individuals near the bottom of the perceived attractiveness distribution and individuals near the very top. For instance, among White males, individuals with physical attractiveness scores of -0.65 (approximately the 5th percentile of all observed ratings) have predicted

### TABLE 3
**Gamma GLMS of income, pooled sample**

<table>
<thead>
<tr>
<th>Term</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>9.99 (.25) .00</td>
<td>9.77 (.26) .00</td>
</tr>
<tr>
<td>Age (years)</td>
<td>.04 (.01) .00</td>
<td>.04 (.01) .00</td>
</tr>
<tr>
<td>Born in U.S.</td>
<td>-.02 (.04) .58</td>
<td>-.03 (.04) .51</td>
</tr>
<tr>
<td>Race: White (ref.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race: Black</td>
<td>-.14 (.04) .00</td>
<td>-.14 (.04) .00</td>
</tr>
<tr>
<td>Race: Hispanic</td>
<td>.08 (.04) .06</td>
<td>.06 (.04) .09</td>
</tr>
<tr>
<td>Female</td>
<td>-.35 (.03) .00</td>
<td>-.38 (.03) .00</td>
</tr>
<tr>
<td>Education: college (ref.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education: HS/AA</td>
<td>-.29 (.03) .00</td>
<td>-.28 (.03) .00</td>
</tr>
<tr>
<td>Education: &lt; than HS</td>
<td>-.54 (.06) .00</td>
<td>-.53 (.06) .00</td>
</tr>
<tr>
<td>Parent’s education: college (ref.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent’s education: HS/AA</td>
<td>-.04 (.03) .14</td>
<td>-.03 (.03) .18</td>
</tr>
<tr>
<td>Parent’s education: &lt; HS</td>
<td>-.11 (.05) .04</td>
<td>-.10 (.05) .06</td>
</tr>
<tr>
<td>Works in professional occupation</td>
<td>.20 (.03) .00</td>
<td>.20 (.03) .00</td>
</tr>
<tr>
<td>Currently married</td>
<td>.11 (.02) .00</td>
<td>.10 (.02) .00</td>
</tr>
<tr>
<td>(log) body mass index</td>
<td>-.10 (.05) .05</td>
<td>-.04 (.05) .48</td>
</tr>
<tr>
<td>Perceived attractiveness score</td>
<td>.15 (.03) .00</td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>3,047.41</td>
<td>3,028.69</td>
</tr>
</tbody>
</table>

*Note.— ref. refers to reference category; HS and AA refer to high school diploma and associate’s degree, respectively.*
earnings of $32,777, while individuals with attractiveness scores of 0.65 (approximately the 95th percentile of ratings) have predicted annual earnings of $37,052. According to these predictions, White males who are perceived to be among the least physically attractive have annual earnings that are 0.88 times the size of White males who are perceived to be among the most physically attractive. Put somewhat differently, White males with very low levels of perceived attractiveness are estimated to earn 88 cents to every dollar likely to be paid to White males who are perceived to possess very high levels of attractiveness. This is similar in magnitude to the canonical Black-White race gap, wherein using the same set of controls we find that a Black person earns 87 cents to every dollar a white person makes.

TABLE 4

| TERM | MODEL 1 | | | | | MODEL 2 | | | |
|-------|---------|----------------------|-------|----------------------|-------|
|       | Estimate | SE | P value | Estimate | SE | P value |
| Intercept | 9.76 | .26 | .00 | 9.79 | .26 | .00 |
| Perceived attractiveness score | .16 | .03 | .00 | .10 | .05 | .05 |
| Age (years) | .04 | .01 | .00 | .04 | .01 | .00 |
| Born in U.S. | −.03 | .04 | .45 | −.03 | .04 | .51 |
| Race × gender: White male (ref.) | | | | | | | | | |
| Race × gender: Black male | −.21 | .06 | .00 | −.22 | .06 | .00 |
| Race × gender: Hispanic male | .03 | .05 | .47 | .04 | .05 | .42 |
| Race × gender: White female | −.41 | .03 | .00 | −.40 | .03 | .00 |
| Race × gender: Black female | −.44 | .06 | .00 | −.44 | .06 | .00 |
| Race gender: Hispanic female | −.31 | .05 | .00 | −.31 | .05 | .00 |
| Education: college (ref.) | | | | | | | | | |
| Education: HS/AA | −.28 | .03 | .00 | −.28 | .03 | .00 |
| Education: < HS | −.52 | .06 | .00 | −.53 | .06 | .00 |
| Parent’s education: College (ref.) | | | | | | | | | |
| Parent’s education: HS/AA | −.04 | .03 | .16 | −.04 | .03 | .16 |
| Parent’s education: < HS | −.11 | .05 | .04 | −.10 | .05 | .05 |
| Works in professional occupation | .20 | .03 | .00 | .19 | .03 | .00 |
| Currently married | .10 | .02 | .00 | .10 | .02 | .00 |
| (log) body mass index | −.04 | .05 | .38 | −.04 | .05 | .37 |
| Skin tone: dark (ref.) | | | | | | | | | |
| Skin tone: medium | .06 | .05 | .30 | .04 | .05 | .42 |
| Skin tone: light | .06 | .05 | .22 | .03 | .05 | .57 |
| Attractiveness score*Black male | | | | | | | | | |
| Attractiveness score*Hispanic male | | | | | | | | | |
| Attractiveness score*White female | | | | | | | | | |
| Attractiveness score*Black female | | | | | | | | | |
| Attractiveness score*Hispanic female | | | | | | | | | |
| AIC | 3,025.95 | | | 3,024.50 | | |

NOTE.—GLMs of income from (1) a specification where attractiveness is constrained to have the same slope across all groups (model 1); and (2) a specification allowing distinct race-by-gender group attractiveness slopes (model 2). ref. refers to reference category, while HS and AA refer to high school diploma and associate’s degree, respectively.
Sizable income disparities are observed among subjects judged to be least and most physically attractive in each other subpopulation analyzed as well. The ratio of predicted earnings of individuals at the 5th percentile of perceived attractiveness compared to individuals at the 95th percentile of perceived attractiveness is 0.83 among White females; 0.78 among Hispanic males; and 0.80 among Hispanic females. Again, note that returns to attractiveness are most pronounced among Black respondents: Black females at the 5th percentile of attractiveness ratings are estimated to earn 63 cents to every dollar of Black females at the 95th percentile of attractiveness. Black males at the 5th percentile of attractiveness are expected to earn 61 cents to every dollar earned by Black males at the 95th percentile of attractiveness.

Across the board, those judged to be physically unattractive appear to earn significantly less than those judged to be physically attractive. In our models, the ratio of expected earnings between individuals near the bottom and top ends of the socially perceived attractiveness scale approximates or even exceeds in magnitude more commonly remarked upon disparities in income—such as the ratio of predicted earnings between men and women (where, according to our models, women are estimated to earn 69 cents for every dollar earned by men) or between Black and White individuals (where Black respondents are estimated to earn 87 cents for every dollar earned by White respondents). Indeed, in the case of Black men and women, the ratio of predicted incomes of individuals at starkly different ends of the perceived attractiveness scale appears to be even larger than either of these more familiar disparities in earnings.

![Figure 3](https://example.com/f3.png)

**Figure 3.**—Model predicted annual income by perceived physical attractiveness, race, and gender. Note 90% and 75% uncertainty intervals are marked.
Our findings emphasize, once again, the importance of capturing within-category inequalities, which are substantial and, as Leicht (2008) argues, growing on the labor market. Moreover, our results suggest that perceived physical attractiveness is a powerful, yet often sociopolitically neglected and underappreciated dimension of social difference and inequality regardless of race and gender. Further still, its consequences are intersectional—the penalties and rewards for perceived attractiveness are strongest among African-American men and women, despite Hamermesh’s (2011) conjecture to the contrary.

**DISCUSSION**

In the past few decades an interdisciplinary body of research has identified a whole range of social and economic benefits associated with perceptions of physical attractiveness (Dion 1972; Webster and Driskell 1983; Hamermesh and Biddle 1994; Conley 2004; Mobius and Rosenblat 2006; Anderson et al. 2010; Conley and McCabe 2011; Mears 2011, 2014; Jæger 2011; Hamermesh 2011; Holla and Kuipers 2015; Wong and Penner 2016). Indeed, similar to research on ethnoracial and gender inequality, studies consistently report that physical attractiveness is an important axis of social stratification, especially in the labor market. One study, for example, even finds that returns to physical attractiveness on the labor market exceed the returns to actual ability (Fletcher 2009). Nevertheless, important questions remain in the literature.

### TABLE 5
**Ratio of Predicted Earnings across Several Scenarios of Interest**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Ratio of Predicted Earnings</th>
</tr>
</thead>
<tbody>
<tr>
<td>At 5th/95th percentile of attractiveness:</td>
<td></td>
</tr>
<tr>
<td>White males</td>
<td>.88</td>
</tr>
<tr>
<td>White females</td>
<td>.83</td>
</tr>
<tr>
<td>Hispanic males</td>
<td>.78</td>
</tr>
<tr>
<td>Hispanic females</td>
<td>.80</td>
</tr>
<tr>
<td>Black males</td>
<td>.61</td>
</tr>
<tr>
<td>Black females</td>
<td>.63</td>
</tr>
<tr>
<td>With average −/+ 1-SD of attractiveness:</td>
<td></td>
</tr>
<tr>
<td>White males</td>
<td>.96</td>
</tr>
<tr>
<td>White females</td>
<td>.94</td>
</tr>
<tr>
<td>Hispanic males</td>
<td>.92</td>
</tr>
<tr>
<td>Hispanic females</td>
<td>.93</td>
</tr>
<tr>
<td>Black males</td>
<td>.86</td>
</tr>
<tr>
<td>Black females</td>
<td>.87</td>
</tr>
<tr>
<td>By race/gender:</td>
<td></td>
</tr>
<tr>
<td>Black/White</td>
<td>.87</td>
</tr>
<tr>
<td>Female/Male</td>
<td>.69</td>
</tr>
</tbody>
</table>

*Note.*—Predicted earnings for the perceived attractiveness score comparisons are calculated from the interactive model displayed in table 4, model 2. Predicted earnings for the race and gender ratios are calculated from the pooled model referenced in table 3, model 2.
Specifically, while this interdisciplinary body of research has identified multiple social and economic benefits of perceived physical attractiveness and has even taken within-category inequalities seriously (see Leicht 2008), this body of literature has seldom examined how race and gender may play a role in shaping perceptions of and returns to perceived physical attractiveness on the labor market (see Hamermesh 2011).

To extend existing research, this study uses nationally representative data and adopts an intersectional approach (see Crenshaw 1989, 1991; McCall 2005; Collins 2015), which examines whether (and to what extent) returns to physical attractiveness on the labor market vary not only by race, but the combination of race and gender. First, we must emphasize, however, that perceived physical attractiveness is a major factor of inequality and stratification regardless of one’s race or gender. In fact, our analyses suggest that the magnitude of the earnings gap among White men along the perceived attractiveness continuum rivals that of the canonical Black-White wage gap and the attractiveness earnings gap among White women actually exceeds, in real dollars, the Black-White wage gap.

This, however, is not at all to say that race and gender do not matter. Quite the contrary. We find that while the returns to perceived physical attractiveness are similar for most race-by-gender combinations, the slope of the returns to perceived physical attractiveness is steepest among Black women and Black men. The returns to attractiveness among Black women, for instance, are so immense, that the earnings of the most attractive Black women appear to converge or even overlap those of white women of similar levels of attractiveness (see figs. 2 and 3). Notably, the returns to perceived physical attractiveness, in terms of real dollars, are similar between White males and Black females even though Black females make significantly less than White males on average. Similarly, the returns to attractiveness among Black men are quite substantial as well, though not enough to see a convergence or cross-over with white men. Among Black women and Black men, the wage penalties associated with perceived physical attractiveness are also so substantial that, taken together, the earnings disparity between the least and most physically attractive exceeds in magnitude both the Black-White wage gap and the gender gap.

These findings line up neatly with insights from theories of capital (Bourdieu 1986) signaling theory (Spence 1973, 2002), and typicality (Maddox 2004). For instance, conceptualizing socially perceived physical attractiveness as a form of bodily capital (e.g., Bourdieu 1984; Mears 2011, 2014; Monk 2015), we restore relationality to physical attractiveness as a form of capital by highlighting the real possibility that from the vantage point of White gatekeepers, in particular, the distribution of physical attractiveness may be unequally distributed across ethnoracial categories. In short, (bodily) capital (e.g., socially physical attractiveness) is not something one either has or does not have. It is
a social relation (Marx 1867; Bourdieu 1986) and its value, in this case, is contingent upon the beholder and the beheld (see Mears 2014).

Following the intuitions of Bourdieu (1984) and Spence (1973), we should expect that members of stigmatized categories, who still manage to be viewed as physically attractive from the vantage point of members of dominant categories who are disproportionately likely to serve as gatekeepers, may be handsomely rewarded by those gatekeepers for being atypical to their stigmatized categories (see Maddox 2004). Consider the finding from Bauldry et al. (2016) that, while people from disadvantaged backgrounds are less likely to be perceived as attractive, attractiveness is actually more strongly associated with educational attainment precisely among this relatively disadvantaged population (i.e., resource substitution).

Thus, contra Hamermesh’s (2011, p. 58) conjecture that attractiveness may matter less among African Americans, the relative scarcity of bodily capital among the stigmatized actually suggests that returns to attractiveness should, in fact, be greatest among them. This is precisely what we find in the case of Black respondents, especially Black women—for whom physical attractiveness is the scarcest among the pool of women in our sample, but also associated with the strongest penalties and returns. In other words, these women may “stand out” as “tokens” on the labor market and enjoy advantages in the labor market (see Kanter 1977; Wingfield and Wingfield 2014), broadly, and wage setting (e.g., initial salaries, returns to performance evaluations, promotions, etc.), specifically.

From another theoretical angle, perceived attractiveness, conceptualized as a status characteristic, does indeed combine with other status characteristics (e.g., race and gender) to modify the effects of attractiveness (Webster and Driskell 1983, p. 159). Given the racialization of beauty, however, the vast majority of ethnoracial minorities, especially those of darker skin, face steep penalties for not “living up to” hegemonic notions of beauty (see Cottom 2018). Not only do they face discrimination based on their ethnoracial category and gender, but also with respect to perceived physical attractiveness. Unlike ethnoracial and gender discrimination, however, lookism is actually legal (Rhode 2010). This means that not only does some appreciable number of whites suffer from labor market inequalities due to physical appearance, but also that ethnoracial and gender inequalities on the labor market may be compounded and exacerbated by lookism. In other words, lookism may operate as a technically legal channel through which ethnoracial and gender stratification on the labor market is maintained. This is not, of course, to state that lookism and its intersections with ethnoracial inequality are always legally defensible. Consider, for example, the $40 million class action lawsuit settlement by Abercrombie and Fitch in 2004 over discriminatory practices linked to their aims to “maintain their brand’s identity” as “classic” and “collegiate,” which very few minorities were seen to “fit” (Mears 2014, p. 1337).
These, however, are exceptions to the usual state of affairs, where lawsuits to address lookism (Rhode 2010) and intersectional claims of discrimination by employers put forth by employees, in general, are rarely heard or successful (Best et al. 2011). Current Civil Rights legislation seems inadequate to serve as a robust buffer against this important dimension of inequality and stratification. To the extent that Civil Rights legislation and secular shifts in ethnoracial attitudes result in opening up more labor market opportunities for Blacks, our findings suggest that the best of these opportunities may be disproportionately awarded to the lightest and “most attractive” Blacks, thereby upholding Eurocentric beauty standards and deeply entrenched ethnoracial biases.

Not only is lookism an important axis of inequality and stratification that receives relatively little scholarly attention within sociology (especially relative to the attention received by ethnoracial and gender inequality and stratification) and little to no serious political mobilization to agitate against it, but it is technically legal in many instances, leaving its victims few viable options or recourses to defend themselves against it or even scholarly literature to more deeply understand its machinations. In this way, lookism reveals itself to be similar to colorism, a phenomenon with which it is intimately linked (see Monk 2014, pp. 1330–31). Taken together, these bodies of research compellingly suggest that the amount of scholarly attention an issue receives need not track alongside its political and popular salience and recognition. After all, the political and popular salience and recognition of colorism and lookism significantly deviate from their empirical consequences (see Hochschild 2006; Monk 2015).

As with any study, however, there are some limitations. Given that we are using observational data, we cannot claim that the relationship between perceived physical attractiveness and earnings is directly causal (though the extensive social psychological research on beauty bias is compelling). While the vast majority of research in this area, including work done by economists (see Hamermesh 2011), tends to interpret the existence of beauty premia and penalties as the result of demand-side discrimination by employers (e.g., taste based or statistical), it remains the case that the role of supply-side factors associated with physical attractiveness may also help explain some nonzero portion of the beauty premier and penalties we find in this study.

For example, it could be the case that individuals who are perceived as physically attractive have enjoyed cumulative advantages such as better treatment in schools by teachers and peers (Umberson and Hughes 1987; Ritts, Patterson, and Tubbs 1992) or even within their own families that may help explain their better labor market outcomes. Bauldry et al. (2016) find that physical attractiveness is associated with educational attainment (e.g., human capital or cultural capital) and Gordon et al. (2013) find that physical attractiveness is associated with social capital (in the context of
the labor market, however, with the importance of recruitment, this would be a mechanism and not necessarily a confounder). Moreover, these advantages could result in increased self-confidence or personality differences that aid not only in attaining better jobs, but also being more willing and able negotiators for wages with employers. At this juncture, however, researchers express little concern over reverse causality playing a role in returns to attractiveness via factors such as self-confidence, which studies find explain only a negligible portion of returns to attractiveness on the labor market (see Mobius and Rosenblatt 2006; Pfeifer 2012). Furthermore, prior studies on beauty premia consistently report that personality traits have little effect on the magnitude of beauty premia (Fletcher 2009; Hamermesh 2011). Consonant with these results we find that our substantive conclusions are robust to controlling for “personality attractiveness” (app. table A5).

Thus, we would caution the reader against believing that our findings are wholly or even mostly reducible to hypothetical supply-side dynamics. After all, our models (similar to previous work on beauty premia) control not only for educational attainment, but also family background and a host of other relevant sociodemographic factors. Moreover, teacher and peer biases and/or biases within families related to perceived attractiveness may simply be part of a long chain of cumulative advantages (and disadvantages with respect to perceived unattractiveness) that act in concert with biases held by current or future employers.

Another potential limitation of this study is the measure of attractiveness in Add Health, which does not explicitly code for some microlevel attributes such as facial symmetry, babyfacedness, body language, hair (Opie and Phillips 2015; Sims, Pirtle, and Johnson-Arnold 2020), eye contact, and facial expressions (e.g., smiling) that may affect attractiveness scores. These attributes may even interact with race/ethnicity and gender and indelibly shape perceptions of attractiveness (see Zebrowitz, Montepare, and Lee 1993; Brescoll and Uhlmann 2008; Livingston and Pearce 2009; Reece 2016). Studies suggest, for instance, that Black men are held to a higher bar for warmth than white men, so Black men who do not smile or have soft facial features (e.g., babyfacedness) may face penalties with respect to perceived attractiveness scores. Similarly, gender stereotypes dictate that women should be nice and warm, consequently, raters may penalize women who are perceived not to be “nice and warm.” This dynamic may be even more acute given the confounding of race/ethnicity and gender such that Black women, in particular, may be relatively more likely to be perceived as masculine (Johnson et al. 2012) and thus less likely to be perceived as stereotypically feminine (e.g., nice and warm). Consequently, future studies and (survey) data collection efforts will profit from seeking to capture even more fine-tuned data about faces and social interactions between respondents and interviewers. Still, it stands to reason that insofar as the present study is concerned, our extensive set of robustness checks
on grooming and “personality attractiveness” should capture some of these potentially important dynamics (see the appendix).

Future research on physical attractiveness and the labor market should seek to render a more granular understanding of how and when in wage-setting processes physical attractiveness is implicated in producing inequalities (e.g., social networks and recruitment, initial hiring stages, performance evaluations, promotions, wage negotiations, etc.). More broadly, however, there remains a lot to be learned about the role of the body and physical appearance in social stratification and inequality in general. In short, so much of our research focuses on politically salient, highly visible social categories, while ignoring how the body signifies these very categories – doing so in ways that are far more complex than analyses of the consequences of mere membership in this or that social category can do justice to. Researchers, then, should continue to explore mechanisms to help us understand the intersections of race, gender, and beauty as interactive axes of social stratification using a variety of methods (e.g., ethnography, interviews, experiments, etc.). This should include examinations of how different bodily attributes interact with race/ethnicity and gender to produce complex patterns of inequality (see Reece 2019).

Nevertheless, while this study shows the unique intersections of perceived attractiveness, race/ethnicity, and gender, we should remain cognizant of the considerable consequences of perceived attractiveness regardless of race/ethnicity or gender. Research that seeks to center the body in modeling social stratification and inequality by conceptualizing various aspects of physical appearance as forms of bodily capital will be an important next step toward deepening our understanding of the processes that produce and reproduce social inequality and stratification. This kind of research will not only enrich our understanding of processes that produce and reproduce inequality and stratification with respect to commonly studied, sociopolitically salient, and officially recognized forms of social difference, but will also involve gazing beyond these commonly studied forms of social difference toward understudied and underappreciated forms of social difference that lack sociopolitical salience and popular and/or official recognition. For as this study shows, sociopolitical salience and popular or even official recognition (e.g., Civil Rights legislation) do not necessarily correspond to the magnitude of a dimension of social difference’s empirical consequences.

APPENDIX
Attractiveness Measurement Model
Respondents’ perceived attractiveness scores generally overlapped across waves. Figure A1 displays the distribution of change in attractiveness ratings across the four waves of Add Health.
Fig. A1.—Difference in perceived attractive scores across various wave pairings
Panel 1 of figure A1, for instance, demonstrates that approximately 50% of respondents received the same attractiveness rating across waves 1 and wave 2 (e.g., “average” in both waves 1 and 2); that 40% were $+/- 1$ category away from their original assessment (e.g., “average” in wave 1 and “un-attractive” or “attractive” in wave 2), and that 10% were $+/- 2$ categories away from their original score (e.g., “average” in wave 1 to “very unattractive” in wave 2). Additionally, analyses show that, across wave comparisons, between 50% and 60% of individuals remained within the same broad category—that is, either “unattractive,” “average,” or “attractive.” (For example, 60% of individuals were either “attractive” or “very attractive” in wave 1 and wave 2; 39%–45% of individuals changed only 1 category—i.e., from “average” to “attractive” or “unattractive”—while less than 3%–5% jumped across the spectrum, from “attractive” to unattractive” or visa-versa.) We opt for a measurement model to combine these indicators, reduce variance across raters, and extract a stable estimate of an individuals’ socially perceived attractiveness. Table A1 provides additional information—factor loadings and goodness of fit metrics—for the measurement model of attractiveness described above.

### TABLE A1
**Measurement Model of Attractiveness**

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Estimate</th>
<th>P value</th>
<th>Standardized Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wave 1 attractive . . . . . .</td>
<td>1.00</td>
<td>.56</td>
<td></td>
</tr>
<tr>
<td>Wave 2 attractive . . . . . .</td>
<td>1.12</td>
<td>.00</td>
<td>.63</td>
</tr>
<tr>
<td>Wave 3 attractive . . . . . .</td>
<td>.74</td>
<td>.00</td>
<td>.42</td>
</tr>
<tr>
<td>Wave 4 attractive . . . . . .</td>
<td>.62</td>
<td>.00</td>
<td>.35</td>
</tr>
</tbody>
</table>

*Note.*—These are factor-loading estimates and goodness-of-fit statistics. Fit measures: CFI 0.99; RMSEA 0.05, 90% CI (0.04, 0.07); SRMR 0.03.

**Wave 1 and Wave 2 Attractiveness Only**

As a sensitivity analysis, we operationalize attractiveness using only wave 1 and wave 2 indicators. As these two indicators are recorded six months to one year apart, we can consider them assessments of an individual’s physical attractiveness at approximately the same point in time (in contrast to, e.g., wave 2 and wave 3 indicators, which were taken five years apart and thus are likely somewhat more representative of real changes in underlying physical traits) and recorded before adult earnings. Notably, these are ratings of attractiveness for respondents from *before* most respondents entered the labor market (i.e., they had no income). To measure attractiveness in this analysis, we take the average of the wave 1 and wave 2 attractiveness scores. Table A2 and figure A2 give results from this analysis. Results point to the same general substantive conclusions as the main text.
TABLE A2
WAVE 1 AND WAVE 2 ATTRACTIVE × RACE-BY-GENDER GROUP ESTIMATES

<table>
<thead>
<tr>
<th>GROUP</th>
<th>Attractiveness × Group Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
</tr>
<tr>
<td>Global slope</td>
<td>.05</td>
</tr>
<tr>
<td>White females</td>
<td>ref.</td>
</tr>
<tr>
<td>Black females</td>
<td>.10</td>
</tr>
<tr>
<td>Hispanic females</td>
<td>.02</td>
</tr>
<tr>
<td>White males</td>
<td>-.02</td>
</tr>
<tr>
<td>Black males</td>
<td>.16</td>
</tr>
<tr>
<td>Hispanic male</td>
<td>.04</td>
</tr>
<tr>
<td>AIC constrain − AIC free</td>
<td>.89</td>
</tr>
</tbody>
</table>

FIG. A2.—Wave 1 and wave 2 attractiveness × race-by-gender group predicted incomes.

Models without Potential Mediators

Though we are not necessarily interested in identifying the casual effect of attractiveness on earnings, we do wish to assess the association of attractiveness and earnings while not controlling for potential mediators of said relationship. For instance, physical attractiveness may operate on earnings by through occupational attainment; if so, controlling for this mechanism could potentially induce posttreatment bias. Conditioning on posttreatment variables via standard regression models greatly reduces accuracy around causal claims by potentially inducing (unbounded/unpredictable) forms of bias (Gelman and Hill 2007; Acharya, Blackwell and Sen 2016).
Beholding Inequality

In this supplemental analysis, we estimate the association among earnings and income while not conditioning on potential mechanisms of that relationship—that is, education, occupation, marital status. This includes not stratifying on employment status as well—as stratifying on a mediator runs the same risk of bias as controlling for one. Results from this analysis are given in table A3 and point to similar substantive conclusions as primary analyses.

### TABLE A3

**Attractiveness by Gender Group Estimates, Net of Potential Mediators**

<table>
<thead>
<tr>
<th>Group</th>
<th>Attractiveness by Group Parameter</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SE</td>
<td>P Value</td>
</tr>
<tr>
<td>global slope</td>
<td>.24</td>
<td>.05</td>
<td>0.00</td>
</tr>
<tr>
<td>White female</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
</tr>
<tr>
<td>Black female</td>
<td>.20</td>
<td>.09</td>
<td>0.06</td>
</tr>
<tr>
<td>Hispanic female</td>
<td>-.03</td>
<td>.09</td>
<td>0.73</td>
</tr>
<tr>
<td>White male</td>
<td>-.01</td>
<td>.07</td>
<td>0.91</td>
</tr>
<tr>
<td>Black male</td>
<td>.36</td>
<td>.13</td>
<td>0.01</td>
</tr>
<tr>
<td>Hispanic male</td>
<td>-.00</td>
<td>.11</td>
<td>.99</td>
</tr>
<tr>
<td>AIC constrain – AIC free</td>
<td>3.23</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Grooming (and Personality Attractiveness)**

Grooming is assessed by interviewers at each wave of the data on a similar scale to attractiveness. Interviewers were asked “How well-groomed is the respondent?” Respondents were assessed as either very poorly groomed, poorly groomed, about average, well groomed, or very well groomed. Grooming and attractiveness assessments are highly correlated within waves. As shown in figure A3, approximately 60% of respondents received the exact same score on their attractiveness and grooming measures (e.g., “average attractiveness” and “average grooming”).

To further interrogate how grooming and attractiveness hang together, we fit a measurement model with two latent variables: (1) a latent attractiveness measure, where each wave’s attractiveness indicator was allowed to load on that latent factor and (2) a latent grooming measure, where each wave’s grooming indicator was allowed to load onto that latent grooming factor. Residual correlations are allowed among each wave’s grooming and attractive indicator. (Failing to do so produces an unidentifiable model structure, given how highly correlated these measures are.) Table A4 summarizes this model.
The two-factor model of grooming and attractiveness fits the data well. As expected from the overlap in observed scores within waves, the grooming and attractiveness measures estimated from the model are highly correlated ($P = 0.82$). Indeed, figure A4 plots attractiveness scores and grooming scores predicted from this model. Because attractiveness and grooming are so highly correlated, their distribution among race and gender groups are similar. Table A5 gives the mean of each race-by-gender group’s attractiveness and grooming scores.
Moreover, because grooming seems to encode the same underlying construct as perceived attractiveness in these data, grooming assessments are contingent on perceived bodily traits in as perceived attractiveness is. For instance, table A6 gives average grooming and attractiveness by race and skin tone. Similarly, social perceptions of one’s grooming are predicted by one’s measured body mass index—particularly among women—as displayed in figure A5.
TABLE A6  
AVERAGE GROOMING AND ATTRACTIVENESS SCORES BY RACE AND SKIN TONE

<table>
<thead>
<tr>
<th>Racial Group</th>
<th>Skin Tone</th>
<th>Mean Attractiveness Score</th>
<th>Mean Grooming Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>Darker</td>
<td>-.11</td>
<td>-.20</td>
</tr>
<tr>
<td>White</td>
<td>Lighter</td>
<td>.01</td>
<td>.01</td>
</tr>
<tr>
<td>White</td>
<td>Medium</td>
<td>.15</td>
<td>.13</td>
</tr>
<tr>
<td>Black</td>
<td>Darker</td>
<td>-.08</td>
<td>-.06</td>
</tr>
<tr>
<td>Black</td>
<td>Lighter</td>
<td>.13</td>
<td>.13</td>
</tr>
<tr>
<td>Black</td>
<td>Medium</td>
<td>-.03</td>
<td>.02</td>
</tr>
<tr>
<td>Hispanic</td>
<td>Darker</td>
<td>-.17</td>
<td>-.20</td>
</tr>
<tr>
<td>Hispanic</td>
<td>Lighter</td>
<td>.01</td>
<td>.01</td>
</tr>
<tr>
<td>Hispanic</td>
<td>Medium</td>
<td>-.05</td>
<td>-.05</td>
</tr>
</tbody>
</table>

Fig. A5.—Plot of grooming scores against log body mass index (with linear smooth)

Because grooming and attractiveness are highly correlated in these data, including both in the same model leads to potential issues with multicollinearity. Nevertheless, when estimating models that control for grooming, substantive conclusions from the main analysis above are generally the same. Table A7 contains the estimated interaction coefficients from a model: earnings ~ f(attractiveness score*race*gender + controls + grooming score).
Finally, analyses examining personality attractiveness—assessed, by interviewers, as either very unattractive personality, unattractive personality, average personality, attractive personality, or very attractive personality—yield similar conclusions. Within wave measures of attractiveness and personality attractiveness are highly correlated (60% of individuals scored the same on physical and personality attractiveness across waves); latent attractiveness and personality attractiveness variables, estimated from a two-factor measurement model, are highly correlated ($P = 0.75$). As was the case with grooming (and perceived attractiveness), perceptions of personality attractiveness are similarly linked to racialized and gendered bodily features, such as skin tone and body mass index.

_Race-by-Gender Stratified Models of Income_

Following the “intercategorical” intersectional framework of McCall (2005), we fit race-by-gender stratified models of the sample to identify attractiveness’ heterogeneous association with income among race-by-gender combinations. Conceptually, this “unpooled” approach allows for each parameter to vary among groups in how in how it predicts income. Such an approach trades precision in income predictions at the cost of greater degrees of uncertainty in estimates. For a sense of whether these stratified estimates are significantly different from one another, we use the test described in Clogg et. al. (1995) and Paternoster et. al. (1998) to compare attractiveness slopes across racial groups, within gender categories. We refer to this test as _Clogg Z-score_. We specifically compare each group’s slope to their same-gender White counterparts. Table A8 provides attractiveness parameter estimates for stratified models of income, for each race-by-gender group.
### TABLE A8
Attraction Parameter Estimates from Race-by-Gender Stratified Models of Income

<table>
<thead>
<tr>
<th>Group</th>
<th>B</th>
<th>SE</th>
<th>(P) value</th>
<th>Clogg Z-Score</th>
<th>Clogg (P) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>White female</td>
<td>.10</td>
<td>.05</td>
<td>.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black female</td>
<td>.28</td>
<td>.08</td>
<td>.00</td>
<td>-1.94</td>
<td>.05</td>
</tr>
<tr>
<td>Hispanic female</td>
<td>.18</td>
<td>.08</td>
<td>.02</td>
<td>-1.22</td>
<td>.36</td>
</tr>
<tr>
<td>White male</td>
<td>.12</td>
<td>.04</td>
<td>.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black male</td>
<td>.40</td>
<td>.12</td>
<td>.00</td>
<td>-2.30</td>
<td>.02</td>
</tr>
<tr>
<td>Hispanic male</td>
<td>.20</td>
<td>.08</td>
<td>.01</td>
<td>-1.94</td>
<td>.35</td>
</tr>
</tbody>
</table>

**Note.**—Clogg Z-scores and corresponding \(P\) values for test of differences in attractiveness slopes from same-gender White counterparts are given as well.

Results from stratified models largely mirror results above: socially perceived attractiveness is a salient predictor of income within each race-by-gender group—with the largest multiplicative changes in income occurring among Blacks. We find evidence that attractiveness slopes among Blacks are significantly different than attractiveness slopes among same-gendered Whites. Predictions from stratified models show that Black females at the 5th percentile of attractiveness have expected earnings of approximately $18,900 per year, while Black females at the 95th percentile of attractiveness have expected earnings of $26,500 per year (i.e., Black women at the 95th percentile of perceived attractiveness earn 40% more than those at the 5th percentile). By comparison, White women at the same intervals are predicted to earn from $22,300 to $25,200 per year (a 13% premium for White women at the 95th percentile of perceived attractiveness relative to those at the 5th percentile). Similar results are seen among men, with Black males at the 5th and 95th percentiles of socially perceived attractiveness earning an estimated $21,000 to $33,800 per year (a 61% premium relative to Black men at the 5th percentile of perceived attractiveness) and White males earning from $31,400 to $36,500 per year (a 16% premium relative to White men at the 5th percentile of perceived attractiveness). In short, while beauty premia are detected across all race-by-gender combinations, these premia (and penalties) are significantly larger among Black men and women.

**REFERENCES**


Beholding Inequality


American Journal of Sociology

Gordon, Rachel A., Robert Crosnoe, and Xue Wang. 2013. “Physical Attractiveness and the Accumulation of Social and Human Capital in Adolescence and Young Adulthood:
American Journal of Sociology


Beholding Inequality

American Journal of Sociology


Beholding Inequality


