OXFORD

Racial bias in media coverage: accounting for structural position and public interest

Eran Shor[1,*] and Arnout van de Rijt[2]

[1]Department of Sociology, McGill University, Montreal, QC H3H 2J2, Canada
[2]Department of Political and Social Sciences, European University Institute, Fiesole, FI 50014, Italy
*Corresponding author. Email: eran.shor@mcgill.ca

Is media coverage racially biased? Past studies documenting differences in the quantity of coverage are small scale or anecdotal. In this article, we investigate whether Blacks receive less coverage than Whites who have reached similarly prominent positions and enjoy similar public interest. We analysed 200 million newspaper references in English-language media to about 32,000 prominent Black and White individuals, predominantly US born. The results do not support the bias hypothesis: Blacks overall receive systematically more coverage than Whites in comparable structural positions and their coverage is on par with that of select Whites who attract equal public interest.

Introduction

Media attention is of consequence. On the one hand, it can greatly aid individuals' success and well-being in various social and occupational domains, from sales to political campaigns to book royalties (Sorensen, 2007; Solomon, Soltes and Sosyura, 2014), while in other domains (e.g. crime), it may be harmful. Do mainstream media have a quantitative coverage bias against ethnic and racial minorities? The answer to this longstanding question is not straightforward partly because the definition of bias in previous research has been quite loose. In many former studies (e.g. Falk, 2010; Cooky, Messner and Hextrum, 2013), the term *bias* has been used simply to indicate newsmakers' tendency to write and publish more stories featuring White men. However, we argue that individuals from certain social groups appearing more frequently in the media than those from other groups relative to their population is not sufficient to establish coverage bias. At least two straightforward sources of coverage differentiation must also be considered in this regard.

The first differentiating factor is real-world structural and occupational inequalities in prominent and visible positions: Certain social groups are underrepresented in social organizations, institutions, and networks. White men dominate most social and employment categories. They are more likely both to enter occupations and social domains that receive greater media attention and to hold the most senior, powerful, and prestigious positions in these domains (Gundemir *et al.*, 2014), leading mainstream media to report more on them. We therefore probe a stronger form of bias, whereby Blacks are less likely to garner media attention even when compared with Whites who have reached similarly prominent positions and achievements.

Second, inequalities in coverage might reflect unequal public interest: Media consumers have varying levels of curiosity for news subjects. Audiences may prefer news about successful Blacks because they find them more exotic or because only very talented and charismatic Black individuals manage to make it into power positions. Alternatively, Whites may generate greater interest even when occupying similar positions because they are assigned more important roles, tend to stay longer in their positions or because of cultural ideologies that consider them more deserving and capable. Media organizations and newsmakers may tailor their programming and coverage to this differential public interest. Our interest is in the excess coverage media give Blacks or Whites compared to the coverage levels one would expect if media simply tailored to public interest.

The various notions of bias, of which we pursue the strongest, derive from three theoretical approaches. First, a long tradition of scholarship in the sociology of the media has looked at the close relationship between societal-level inequalities and media production and reporting patterns. Unequal occupational opportunities sort Whites and Blacks into positions of differential public prominence and access to media coverage (measured here as the number of articles that mention and discuss an individual or topic). Structural constraints and organizational practices endemic to the media industry determine the sorts of societal positions that enjoy regular media attention, positions that are heavily White dominated. Failing to account for these structural inequalities risks mislabelling real-world inequalities as media bias. The second approach emphasizes differences in public interest in African-Americans and Whites who have achieved the same structural positions. Failing to account for these differences in interest would risk mislabelling race-dependent public interest as media bias. The third approach, promoted by sociology of race and ethnicity scholars, has focused on media-level explanations—the role of racial ideologies and newsroom cultures, and the biased coverage choices of journalists, editors, and publishers. This strongest form of media bias can be defined as any excess media attention to individuals from one ethnic or racial above and beyond attention levels one would expect on the basis of structural position and public interest. The goal of our article is the identification of this form of bias.

The identification of this stronger form of media bias requires data on many individuals from many media sources. Our empirical strategy in this study closely follows the strategy employed by Shor, van de Rijt and Fotouhi (2019) in their analysis of strong-form gender bias in news coverage. We combine large-scale data from two separate sources to address our research questions. We measure public interest using the English Wikipedia page views ('hits') of about 32,000 prominent Black and White individuals. We use Wikipedia occupational category membership for capturing structural positions. Since Wikipedia hits are highly correlated with search-engine queries (Yoshida *et al.*, 2015), they provide us with a natural measure of expressed public interest (Shor, Van de Rijt and Fotouhi, 2019). Second, we use the Lydia text analysis system to compare the coverage rates of these individuals in about 2,000 English-language newspapers and online news websites. We regress news coverage on the subject race, controlling for our measures of structural position and public interest. While these methods and measures are not without problems, they nonetheless allow for a more systematic and comprehensive assessment of media bias than what has been possible before.

Potential sources of difference in the media coverage of Blacks and Whites

Below, we consider three major theoretical approaches to racial inequalities in coverage. We maintain that quantitative racial inequalities in media coverage may be attributed to one of three major classes of explanations: (i) real-world structural inequalities, (ii) unequal public interest, and (iii) media-related inequalities, such as newsmaker predispositions and biases. The role of the theories is to conceptually separate three forms of bias in increasing order of strictness:

(i)   Racial coverage differences due to inequalities in structural positions which should vanish once structural position is controlled, presenting a null hypothesis against which H1 is tested;
(ii)   Racial differences in public interest (H2), which provide a null hypothesis against which H3 is tested;
(iii) Media bias in coverage, which yields H3.

Structural inequalities in the workforce and in high-visibility positions

One fundamental driver of media coverage-volume differences is inequalities in the workforce and in access to high-visibility positions. Black individuals are underrepresented in leadership positions in various occupations and domains that attract media attention (e.g. politics, the business world, and various high-status professions, such as medicine and law), while being overrepresented in some other domains (e.g. certain team sports or high-profile criminal activities). These structural inequalities derive from among other things institutional racism and differential opportunities available to Blacks and Whites.

The importance of such 'real-world' factors is supported by a host of research findings. On the one hand, research by sociologists, political scientists, and economists highlights significant disadvantages in the promotion patterns of Black individuals into organizational leadership positions. In politics, studies have reported an under-representation of African-Americans in political leadership positions at the municipal, state, and federal levels (Gundemir et al., 2014). In governmental organizations and in the private sector, Blacks are less likely to be promoted to various positions (Powell and Butterfield, 1997). They are consequently underrepresented in organizational managerial and leadership positions (Stainback and Tomaskovic-Devey, 2010).

These disparities result from a host of differentiating mechanisms at both the individual and organizational level. First, the literature highlights multiple barriers to African Americans' entrance into the workplace. These include disparities in hiring practices (Pager, 2007; Pager, Western and Bonikowski, 2009; Gaddis, 2015), as well as outright discrimination (Couch and Fairlie, 2010). Such disparities in hiring are especially salient among those with college education (Zhang, 2008) and in sectors of the economy with higher status and salaries (Smith and Elliott, 2002; Stainback and Tomaskovic-Devey, 2010; Gaddis, 2015), where employers often use social networks and perceived personal and cultural similarities as hiring criteria (Rivera, 2012).

Within jobs, Black employees face a 'glass ceiling' that hinders their vertical career development, similar to the one that has been documented for women (Maume, 1999). Their progress in the workforce, in politics, and in other professional and public domains is obstructed by persistent institutional barriers, cultural ideologies, and long-held stereotypical perceptions (Baldi and Branch McBrier, 1997; Smith, 2005). For example, multiple studies have documented that leadership traits and managerial competence are more strongly associated with White-majority group members than they are with ethnic minorities (Ridgeway and Kricheli-Katz, 2013; Gundemir et al., 2014), resulting in unfavourable promotion prospects for these ethnic minorities.

The research reviewed above suggests that Blacks are likely to be underrepresented in many high-status positions, less publicly visible as a group, and consequently less likely to appear in the media. However, such under-representation is not uniform across all social and occupational domains. Sports is one field where high-level and high-status African-American athletes may actually be overrepresented relative to their share of the population, in particular in certain high-profile team sports such as basketball and American football, as well as in certain Olympic sports with high visibility, most notably athletics. Criminological research has similarly suggested that Blacks have been heavily overrepresented in reported statistics on criminal offenses (Phillips, 2002). These differences in reported crime rates are influenced by a variety of social factors, including socioeconomic and immigration status, parental marital stability, racial segregation, and various neighbourhood effects (Krivo, Peterson and Kuhl, 2009), as well as racial profiling and differential policing and sentencing practices (Davenport, Soule and Armstrong, 2011). Still, the fact that African Americans are overrepresented in crime statistics might suggest that the media would be more likely to cover Black individuals when writing about crime.

The research reviewed above suggests that over-representation or under-representation of African Americans in the media may, at least in part, be the result of structural inequalities in highly visible positions. However, it remains unclear whether such real-world inequalities are powerful enough to account in full for differences in the coverage volume of Black and White individuals or whether other factors (discussed in the following sections) are also at play. One way towards examining this proposition is to compare pairs of Black and White individuals who hold specific equivalent structural positions or who have reached equivalent achievements, such as state senators (politics), company founders (business), Olympic medallists (sports), or, conversely, murderers (crime). If Whites hold a coverage advantage even when compared with Blacks in equivalent occupations and those with similar achievements or notoriety, then we may conclude that coverage differentials are not entirely due to racially-structured real-world inequalities.

H1: Blacks receive less media coverage than Whites who occupy the same structural positions.

Inequalities in public interest for equivalent positions

A systematic comparison of Blacks and Whites who have reached equivalent positions in various social and occupational domains permits a stronger test of bias than the comparisons of basic coverage rates used in many past studies. Still, it does not account for potential differences in public interest in Blacks and Whites. Popular media outlets need to consider readership and viewership ratings and they face commercial pressures and constraints (Iyengar and Hahn, 2009; Sui et al., 2018). Journalists and editors often ask themselves whether a given topic or individual is of sufficient public interest to merit coverage; that is, to what extent are readers and viewers interested in knowing more about this topic or individual? We may therefore ask: Are there systematic racial differences in public interest?

There are several reasons to believe that racial minorities who reach visible positions might actually attract *more* public interest than White equivalents. First, in many cases, due to the relative rarity of minorities, their stories might be deemed journalistically more exotic and unique. Second, it is quite possible that African Americans who were able to overcome the workplace barriers mentioned above are simply more talented, ambitious, hard-working, and accomplished than White individuals who have reached similar positions without the need to struggle as much. For example, Gaddis (2015) conducted an audit study showing that when applying to new positions, particularly those with higher prestige, Black applicants

from elite universities only do as well as White candidates from less selective universities. Similarly, Pager's work (Pager, 2007; Pager, Western and Bonikowski, 2009) demonstrates that Black job candidates without a criminal record do not do better than White candidates who do hold such a record. As Black employees also find it harder to advance in the organizational hierarchy, it appears likely that those who are eventually hired and promoted are overall more qualified, talented, and unique when compared with their White counterparts. This, in turn, might also make them more likely to attract public and media attention.

While this is a plausible prediction, the sociological literature on labour inequalities suggests a set of mechanisms that might lead to lower visibility and public interest in successful Black individuals. While our review of the workplace literature in section 'Structural inequalities in the workforce and in high-visibility positions' suggests a lower presence of Blacks in the most prominent positions, here we focus on the part of this literature suggesting that Black individuals might also generate less public interest even when they do reach top positions. First, minority employees in senior positions tend to receive inferior work assignments, which are less challenging and glamorous, entail less authority and work responsibilities, and result in lower visibility (Williams and Multhaup, 2018). In addition, racial minorities tend to suffer from shorter career spans, especially when they reach managerial and leadership positions. Black executives are more likely to accept promotions to risky positions and positions that stifle long-term mobility, out of fear that this will be their 'first and only' opportunity (Collins, 1997). Consistent with 'glass cliff' theories, which predict greater instability and higher turnover rates for women and minorities, Cook and Glass (2013) also found that Black managers are more likely to be promoted to CEOs of weakly performing firms. When the performance of these firms further declines, they are then likely to be replaced by White men. Other studies have reported that women and minorities, particularly those in senior positions, are more likely to lose their jobs when firms are downsizing (Couch and Fairlie, 2010; Wilson, Roscigno and Huffman, 2013). This tendency again results in shorter career spans, which lead to shorter periods of public visibility.

The literature on status beliefs and cultural stereotypes further suggests that successful Black individuals might attract less public interest even if they reach leading social and professional positions, succeed in them, and retain them for extended periods. This literature (Lamont, 2012; Ridgeway, 2013) emphasizes the role of status and cultural beliefs in maintaining gender and racial inequalities. Entrenched racial and cultural beliefs about social categories and the types of

individuals who deserve special respect shape people's expectations of themselves and others. For example, common stereotypical beliefs that Blacks are cognitively inferior (Miller, 1998; Harrison and Lawrence, 2004) and that they are less competent to perform leadership and managerial positions (Ridgeway and Kricheli-Katz, 2013; Gundemir et al., 2014) may result in more than just unfavourable promotion decisions and prospects. Rather, these deeply entrenched beliefs may also lead to public perceptions that Black individuals who have reached these positions are merely tokens and do not merit the attention and respect that come with their position. Such perceptions might also influence judgements about who is more worthy of attention in the sports field. The common perception that Black athletes are naturally gifted and do not need to work as hard for their success (Harrison, Azzarito and Burden, 2004) might yield a public notion that White individuals who made it despite their physical limitations are more worthy of praise and attention.

Taken together, these workplace mechanisms and cultural stereotypes suggest that even when African Americans do reach positions of prominence they may still be considered less interesting than White individuals who have reached similar positions. We therefore propose the following:

H2: Blacks attract less public interest than Whites who occupy the same structural positions.

Media-related factors: bias and discrimination in media coverage

Following sections 'Structural inequalities in the workforce and in high-visibility positions' and 'Inequalities in public interest for equivalent positions', we may define a stronger form of racial coverage bias as any unequal media coverage of Blacks and Whites who occupy the same structural positions *and* attract the same degree of public attention. Such bias might either favour or disfavour African Americans. On the one hand, journalists and editors may practice tokenism, celebrating the success of Blacks who made it to the top, and seeking to promote them as future role models for young Black kids. Indeed newspapers covering the campaigns of African-American political candidates tend to frame their candidacy as 'unique' or 'historical' (Zilber and Niven, 2000b; McIlwain, 2011). Others may want to highlight Blacks' successes and accomplishments in order to prove that they are not biased against them, mitigating criticism from academics and the public (e.g. Rivas-Rodriguez et al., 2004). Consequently, they may be motivated to prioritize the coverage of successful or prominent Blacks. Following a similar logic, some journalists may also wish to underplay the involvement of racial minorities in committing violent or criminal

acts, seeking to counteract common stereotypes about these groups.

On the other hand, the literature on racial inequalities, which we reviewed in the previous section, is also relevant for understanding the media's treatment of African Americans. Mass communication scholars have long argued that media practices significantly exacerbate and artificially magnify inequalities between racial minority and majority individuals (Iyengar, 1990; Entman and Rojecki, 2001). The historical dominance of White men in newsrooms and in particular in editorial positions has created a power structure that favours issues associated with White men, their actions, and their preferences, while treating these norms as objective professional routines (Pritchard and Stonbely, 2007). Consequently, even when the proportion of non-White journalists increases, as it has in recent decades (American Society of News Editors [ASNE] 2018), reporters' norms about what constitutes professional journalistic identity may override racial and ethnic identities (Husband, 2005). These norms about professional journalism are, in turn, firmly grounded in historical perceptions of journalistic neutrality, what constitutes 'important' stories, and who are 'newsworthy' individuals.

The idea that the media might award unequal attention to Blacks and Whites in different fields appears to be supported by some anecdotal evidence, as well as a few systematic analyses. Most of these studies have focused on three main fields: politics, entertainment, and crime. In politics, some research has reported that Black politicians and political candidates tend to receive less coverage than their White counterparts (Canon, 1999; Tolley, 2015), while others failed to find such differences (Terkildsen and Damore, 1999; Zilber and Niven, 2000a). It therefore remains questionable whether there are indeed broad racial differences in the media coverage of politicians.

Studies examining the field of entertainment have found that Black characters traditionally appeared almost exclusively in comedies and action films (Eschholz, Bufkin and Long, 2002) and that about half of the films underrepresented Blacks relative to their share in the population (Smith and Searles, 2014). However, on television, while Blacks were underrepresented and appeared mainly as musicians or entertainers until the 1980s (MacDonald, 1983), by the 1990s and 2000s their share among television characters had actually somewhat surpassed their relative share in the general US population (Mastro and Tukachinsky, 2012).

Finally, crime is the field where the largest number of studies on the potential relationship between race and media coverage have been conducted. Early studies (Entman, 1992; Entman and Rojecki, 2001; Dixon, Azocar and Casas, 2003; Dixon, 2008a) have shown that African Americans tend to be overrepresented in television news as perpetrators of criminal acts, not only relative to their population share, but also relative to their representation in official arrest statistics. However, in more recent work, Dixon reported that depictions have changed with the years and Blacks are now proportionally represented in television news as perpetrators, victims, and officers (Dixon and Williams, 2015; Dixon, 2017). We therefore test the following proposition:

H3: When compared with Whites who (i) have reached the same structural positions, and (ii) draw similar public interest, Blacks receive less media coverage.

While the studies reviewed above are informative and instructive, they leave many questions open and share several shortcomings, addressed by the present study. First, most of the previous research has examined the question of whether minorities are underrepresented relative to their general share of the population. As we note above, this question is problematic in the context of real-world inequalities, with uneven rates of promotion, attainment, and participation in various domains. We therefore examine a different question here: Do African Americans who have managed to reach prominent positions receive as much coverage as their White counterparts? Second, most former studies have examined the coverage in a single domain (e.g. crime), relying on a relatively smaller sample of media outlets. Such research designs render generalizations problematic and increase the risk for anecdotal and selective evaluations. Indeed, Hier (2009) has argued that studies on minority representation in the mass media are often conducted by 'like-minded scholars' who pose a limited set of questions, which produce predictable conclusions about the prevalence of racism. While this is certainly not true for all of the research we review above, a careful systematic and large-scale analysis is better equipped to circumvent some of these pitfalls.

Finally, former studies have not evaluated the stronger form of bias we seek to test here. Zilber and Niven (2000b) argue that differential racial coverage for political candidates in the media does not necessarily reflect a double standard. Rather, it might reflect the reality of what they say or do. African-American politicians might systematically differ from White politicians in the kind of images they attempt to project and this, in turn, might influence their coverage. Here we argue that a compelling investigation of racialized coverage requires controlling for both structural position *and* public interest (Iyengar and Hahn, 2009; Sui et al., 2018).

Data and methods

Sampling individual names by race and domain

We obtained a large and systematic sample of names by collecting data from the English Wikipedia categories databases (see https://en.wikipedia.org/wiki/Wikipedia:Categorization). We first constructed an initial list of larger social and occupational domains, based on common newspaper categorizations and domains of interest. These include politics, business, crime, entertainment, sports, arts/literature, and science. We then identified prominent sub-domains within each of these larger domains (e.g. actors, singers, dancers, and directors within entertainment). For example, to assemble our list of Black singers, we extracted all names appearing under the category 'African-American female singers' and all names appearing under the category 'African-American male singers'.. Next, to assemble our comparison group, we extracted all names appearing under the category 'American female singers' and all names appearing under the category 'American male singers'.

There are typically no unique categories for White individuals in Wikipedia, consistent with sociological observations about tendencies to 'unmark' and neglect 'ordinary' phenomena (Berkhus, 1998). We therefore used the general category and then eliminated from it the names included in one of the other categories, marking them as Black, Asian, Latino, Native American, or Middle Eastern. We then manually validated the accuracy of these categorizations, by examining a random sample of 1,369 names: 100 names from the pool categorized as 'Black' for each of our seven main categories except for science (where only 69 names were available) and 100 additional names for each of these seven categories from the pool now categorized as 'White'. Judging by the Wikipedia background descriptions of individuals, as well as by their Wikipedia and Google image photos, categorization was nearly perfect for 'Black' individuals (only one case of probable misidentification) according to common North American standards of racial classification. Categorization was slightly less accurate for 'White' categories, as these were constructed from the general Wikipedia categories. Still, it was at least 90% accurate for all categories and at least 97% accurate for five of the seven categories, suggesting only marginal error in our comparisons (see online appendix for the full list of 1,369 random names and their racial verifications).

While these Wikipedia lists provide the most comprehensive and up-to-date data we could locate of individuals identified by race and across various occupational and social categories, a few limitations must be acknowledged. First, some of these categories are comprehensive and complete, or nearly complete, such as 'US House representatives', 'Cabinet Secretaries', 'presidential candidates', 'Olympic medalists', and 'basketball players'. Other categories, however, are clearly incomplete (e.g. 'murderers', 'singers', 'dancers', and 'painters', although they mostly still include the large majority of well-known American individuals in each sub-domain. However, this is not the case for all of them.

More specifically, in some domains, the Wikipedia lists appear to be quite exclusive and should therefore be treated with extra caution. Perhaps the most prominent example is the category of businesspeople. Beyond being non-specific (i.e. we were unable to analyse sub-categorizations within the larger category), this category is also very exclusive, failing to include most American businesspeople. However, the criteria by which individuals were selected into this category, and even before that, the 'notability' bar used for deciding which individuals merit a Wikipedia entry, remain uncertain. This is especially problematic given research showing that visible minority editors are underrepresented on Wikipedia, which may have led to an under-representation of Black historical and contemporary figures (Skiena and Ward, 2014; Cassano, 2015; Adams, Brückner and Naslund, 2019). It therefore seems plausible that, particularly given the high selectivity of the category, for a Black businessperson to receive a Wikipedia page (or for this page to not be deleted) and then for this page to be included in the category, that person needs to meet a higher bar of success and fame when compared with a White businessperson (see Tripodi, 2021 for a comparable argument and empirical analysis for gender). Indeed, a look through the names included in the category 'Black Businesspeople' reveals that many of them have made their fame as highly successful athletes, actors, and music performers while most of their White counterparts are known only as CEOs, entrepreneurs, and wealthy individuals.

These selection issues in the construction of some Wikipedia categories may pose a particular challenge to analyses that only control for age and gender. In addition, analyses that control only for gender and age may lack sensitivity to important differences within domains. For example, comparing US senators with local judges in the domain of 'politics' or rap singers (more often Black) and country singers (more often White) in the domain of 'singers' may be problematic.

Our use of a measure of public interest as a control variable tackles this issue. We explain below why we believe that both the imperfections in Wikipedia's compiling and categorization and the lack of sensitivity to differences within domains are substantially tempered when one considers, and controls for, public interest. We therefore argue that control is tightest, comparisons

are cleanest, and standard for bias are strictest in our final analysis.

Altogether, we assembled an initial database containing 39,983 Black and White names. We then sought to account for potential imbalances between the ways Black and White lists were originally assembled in each domain and sub-domain. First, many of the sub-domains included no Black individuals born prior to 1920. We eliminated all names of individuals born prior to 1920, restricting our analysis to persons (likely) alive at the time of coverage (2004–2009). This reduced the sample to 31,724 names. Moving the cut-off year of 1920 forward or backward a decade or lifting this restriction entirely is inconsequential for the results.

Second, previous research has shown that minorities need greater levels of public interest to be included in Wikipedia (Skiena and Ward, 2014). Our control for public interest adjusts for such imbalance.[1] Beyond the theoretical importance of controlling for public interest, this strategy also helps minimize the substantial potential limitations and biases in our categories data. One such problem might stem from the Wikipedia entries' selection and categorization process, discussed above, that is the unequal standards for inclusion in both Wikipedia itself and potentially also in each of the categories. When comparing say businesspeople of equal public interest, we substantially reduce the potential bias introduced by such differential selection criteria, with public interest serving as a proxy for relative success and fame. Our use of public interest as control also reduces the danger of 'comparing apples to oranges' due to some occupational categories being too general (e.g. 'entertainers'), as it eliminates the option of a very visible individual (e.g. an Oscar winner) being compared with a more marginal individual (e.g. an obscure B movie cast member) who are both in the same category.

By relying on Wikipedia categories, we assume that categorization is accurate, that domains are meaningful, and that classification by Wikipedians reflects public perception. The dataset is too large to verify this for each case, but in section 3.4, we show how the US House of Representatives lacks any misclassification. A careful inspection of the random sample of 1,369 names described above shows that other categories also appear valid, apart from the business category, which often contains members of other categories, as mentioned above. A limitation of using the English-language Wikipedia is that the names we analyse are very disproportionately sampled from US domains. While limiting the primary scope of our analysis to the United States, this focus at the same time increases our confidence that the racial categorization corresponds with common public perceptions.

Measuring media coverage

Our measure of media coverage relies on data collected by the Lydia news analysis system. Lydia provides time-stamped records of occurrences of person-names in the scanned and digital records of over 2,000 newspapers, magazines, and online news sources (most of them American) between 2004, when Lydia's data collection started, and 2009, after which most online newspapers placed the majority of their content behind paywalls. This timeframe offered a unique opportunity to access very large amounts of data from a very large and highly diverse number of news sources. These sources include both large national newspapers and smaller local ones, as well as weekly magazines and online websites of major TV news stations, and several prominent news websites, portals, and blogs.

One possible challenge to the validity of our analysis is common names, particularly since Blacks are more likely to choose unique first names (Lieberson and Mikelson, 1995). When multiple well-known individuals share the same name (e.g., John Smith), the Lydia system counts them all as the same person, thus artificially inflating the presence of the common name. We therefore excluded from our analyses all names that appeared in the Wikipedia categories followed by parentheses, signifying multiple Wikipedia Webpages for the same name.

Our opening paragraph emphasizes the effect of media coverage on careers, which suggests a causal arrow running back from media coverage to structural position, posing the threat of collider bias when controlling for structural position (Elwert and Winship, 2014). It is therefore important to clarify that our analysis has no ambition to identify a causal effect of structural position on news coverage. Rather, our interest is entirely descriptive in wishing to document the extent to which African Americans receive excess coverage compared to what their coverage would be if it reflected their structural positions only. Our exercise is analogous to an empirical evaluation of racial hiring discrimination that adjusts for prior work experience, which itself may have resulted from prior discrimination.

News coverage is approximately lognormally distributed. Figure 1 shows for all major news categories of kernel density estimates of the logarithm of news coverage with base 10. Dashed curves indicate normal densities. The kernel density estimates generally follow the normal densities closely. In our later analysis, we will use this logarithmically transformed measure of news coverage.

Measuring public interest

Internet search-engine queries are considered a dominant expression of public interest. When people wish
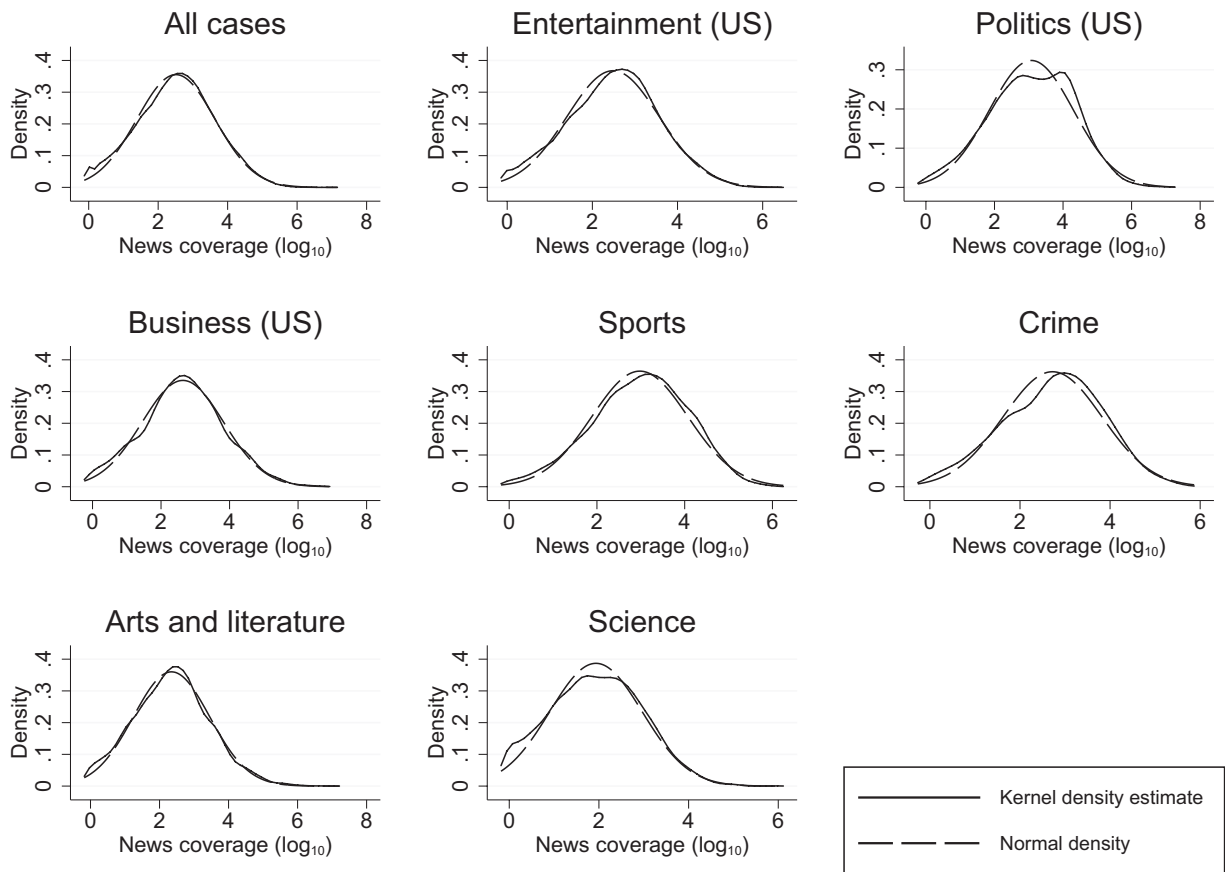
Figure 1 Density plots of news coverage (log10) by major categories. Kernel density is estimated using a Gaussian kernel function and Silverman's (1986) optimal bandwidth.

to know more about a politician, an artist, or a businessperson, they type their name in a search engine. Unfortunately, most major search engines restrict access to longitudinal records of search queries, making it impossible to obtain such quantities for large volumes of names within a reasonable timeframe. However, when querying names of well-known individuals, their Wikipedia page almost always comes up as one of the first-listed search results.

Moreover, Yoshida et al. (2015) found that monthly search-figures obtained from Google Trends correlate strongly with monthly Wikipedia page views. They concluded that, at least for high-frequency names, Wikipedia page views serve as an effective proxy for web search behaviour, and consequently public interest. Unlike search queries, the frequencies of Wikipedia page views are publicly available, unrestricted, and free. Therefore, we use this measure as a powerful and readily available proxy for public interest.[2]

While the relative frequency of Wikipedia page views is not a flawless measure of public interest (we discuss

some of its limitations below), it has several important features. First, it is a behavioural measure, allowing the public's revealed preferences to determine which figures are more interesting or newsworthy. Indeed, when looking only at more conventional measures of 'importance', such as reaching a certain structural position (e.g. Governor) or a tangible achievement (e.g. winning certain prizes or breaking an athletic record), one runs the danger of overemphasizing criteria that may have little to do with public preferences and tastes.

A second important advantage of Wikipedia page views as a measure of interest is that it accounts for some of the unmeasured differences and inequalities between Blacks and Whites, which we discussed in our review of the labour market and racial pay gap literature. Consider, for example, the proposition that Blacks are often assigned roles and tasks of lesser importance and lower visibility in organizations. Whenever that is the case, we would expect these individuals to also draw lesser public interest, be queried less in search engines, and looked up on Wikipedia less frequently. Similarly, if Blacks suffer from greater professional

instability and turnover rates once they reach high-end positions, and consequently spend less time in office compared with White colleagues (the glass cliff), then this should be reflected in reduced public interest in these Black individuals once they exit the position. Of note, the Wikipedia page views measure also exhibits both high face validity and high criterion validity. Scholars commonly agree that the number of web searches reflects the degree of public interest in this subject matter or individual (Ripberger, 2011; Yoshida et al., 2015; Nghiem et al., 2016).

Use of our measure of public interest—median monthly Wikipedia page views—has one important limitation: Public interest is informed by media coverage itself. It shares this drawback with any other reasonable measure of public interest. Media coverage does not only reflect public interest; it also shapes this interest (McCombs and Shaw, 1972; Gans, 1979). When newspapers and other media frequently mention certain individuals, media consumers would be more likely to search for these individuals online (including Wikipedia). Conversely, journalists and editors are often sensitive to public interests and may be more likely to write about individuals who receive higher online traffic (e.g. through name searches on various websites). This partial co-determinacy of interest and coverage means that search queries and page views are measures of actual public interest, not of what public interest would have looked like in the absence of any agenda-setting effects of media coverage. In a causal analysis framework, controlling for public interest would risk endogenous selection bias due to conditioning on a collider variable (Elwert and Winship, 2014). Our approach is explicitly descriptive, with the sole aim of measuring a strict form of bias, defined as racial differences in coverage net of public interest, even if that public interest may itself have been evoked by prior news coverage. Our intention is to examine whether there is excess Black or excess White coverage compared to what coverage would have been if the news merely reflected public interest.

Like our measure of news coverage, our measure of public interest is also reasonably approximated by a lognormal distribution. Figure 2 shows for all major news categories kernel density estimates of the logarithm of public interest with base 10. Dashed curves indicate normal densities. The kernel density estimates tend to have excess kurtosis and left skew, with disproportionate numbers of individuals who attract minimal public interest. As a test of face validity, we checked that individuals with very low public interest are indeed obscure and attract relatively low news coverage as well. A random sample of 10 names with a median number of Wikipedia page visits per month under 10 is composed of names few will have heard of: Debra

Wexler, David Farley, Erik Compton, Isaac Kramnick, Ann Devroy, Edith Diaz, Christopher Farnsworth, Curtis Allgier, Lewis Billings, and Lynette Diaz. Despite these deviations from normality, the logarithmically transformed measure is vastly better approximated by a normal variable than by the untransformed variable with a very high right skew, and we will use the transformed variable throughout as measure of public interest in our regression analysis.

Descriptive statistics

Table 1 presents the names, news coverage, public interest, race, gender, and birth year of a random sample of members of the US House of Representatives. The reader can verify that the race and gender classifications are accurate and the news coverage and public interest measures are higher for more prominent members of the House.

Table 2 shows the descriptive statistics of these variables for all 31,724 individuals in our sample. There are no missing values. The table omits the many structural position variables whose membership numbers can be found in later tables.

Findings

Domain-stratified media supply

To test H1, we run through all structural positions on which we collected data and regress news coverage on race, including only individuals in the respective structural position. All regression models estimate heteroscedasticity-robust standard errors. Per structural position, three regression models are estimated: one with only race as predictor, one controlling for gender and a cubic polynomial of age, and one additionally controlling for public interest. Each row of Table 3 presents the race coefficients and standard errors from three such regression models (columns 5, 6, and 7) for the sub-category indicated (columns 1 and 2). The table also reports the numbers of Blacks and Whites in each category (columns 3 and 4). Because of very small numbers of Black names in some of the sub-categories, reported tests suffer from low statistical power and sensitivity to deviations from normality. Of note, in some categories, such as House representatives, cabinet secretaries, and presidential candidates, we are analysing the full population, rather than just a sample.

The first row of Table 3 shows that Blacks with a Wikipedia page on average receive more than double ($10^{.352} \sim= 2.25$) the news coverage that Whites with a Wikipedia page receive. The most naïve, unadjusted estimate of racial coverage bias is thus that it favours Blacks. This observation is in line with what one obtains when one considers raw total counts of news mentions of all Blacks and all Whites, which
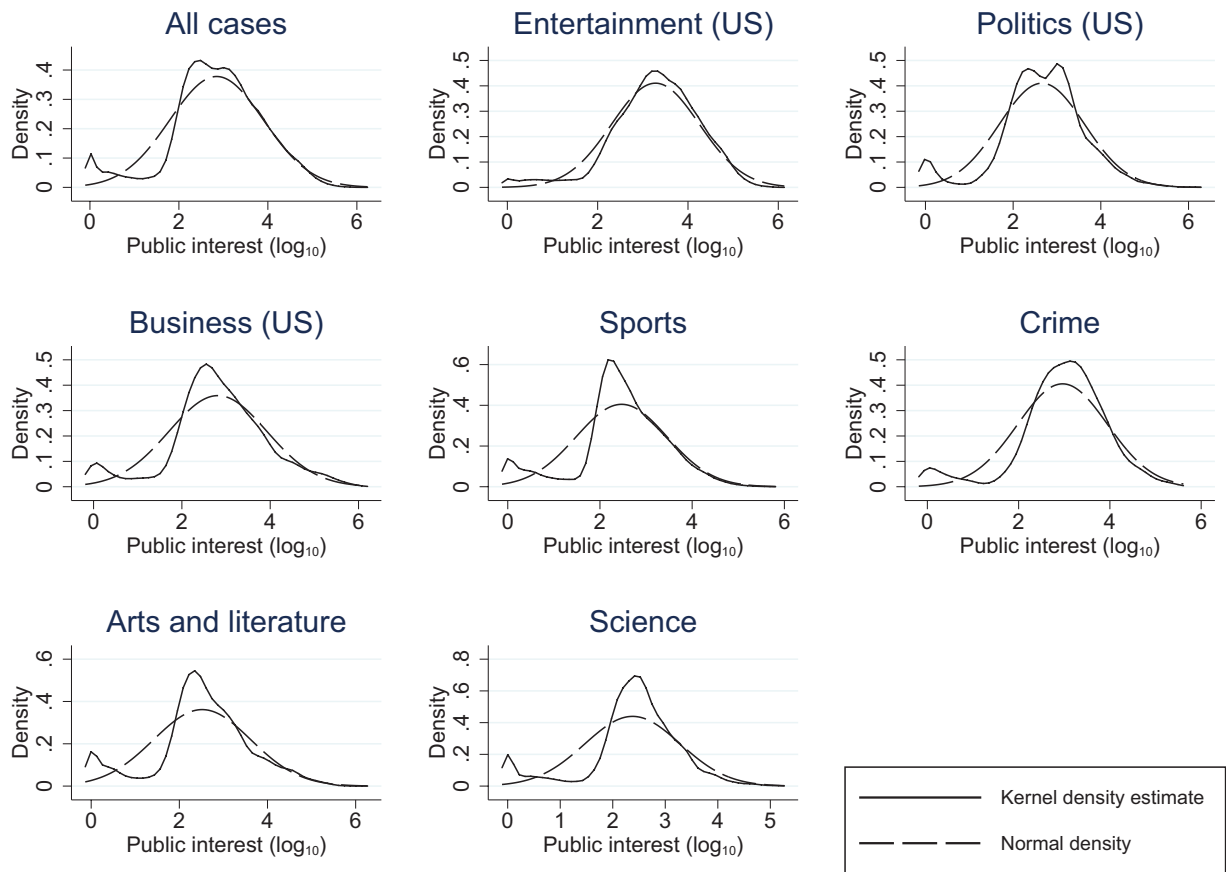
Figure 2 Density plots of public interest (log10) by major categories. Kernel density is estimated using a Gaussian kernel function and Silverman's (1986) optimal bandwidth

are 55,462,880 and 163,124,557, respectively. This Black coverage ratio of roughly 1:3 exceeds the ratios of their US population shares, which are 13.6% and 63.7%, respectively, in the 2010 US Census. Raw estimates of news coverage limited to individuals whose race can be identified through their Wikipedia page therefore suggest that Blacks are overrepresented in media coverage.

Table 3 shows that Blacks also overall receive systematically more news coverage than Whites who occupy the same structural positions (column 5), including when we control for gender and age (column 6). Across all categories the race coefficient is either significantly positive or insignificant, with just one exception: US House representatives. Many coefficients exceed 0.3, corresponding to a doubling of coverage ($10^{.3} = 2$) and some even 0.6, a quadrupling of coverage. Figure 3 shows that these findings are reflected in the coverage distributions of Blacks and Whites, with politics being the clear exception to what is overall a Black coverage advantage. H1 is therefore rejected.

Domain-stratified public interest

The analyses presented in columns 5 and 6 of Table 3 take an important step toward examining media bias by stratifying Blacks and Whites according to social and occupational domains and accounting for some notable achievements. Still, they fail to capture the actual degree of public interest in a certain individual during the time of the study. In Table 4, we present an analysis of public interest analogous to the news coverage analysis. We again report the numbers of cases in each domain and sub-domain (columns 3 and 4). Columns 5 and 6 report race coefficients and standard errors from regression models predicting public interest from race alone (column 5) and with gender and age controls (column 6).

Table 4 demonstrates that Blacks in our sample generally received more views of their Wikipedia page when compared with Whites occupying similar structural positions. This suggests greater overall public interest in Black individuals who have a Wikipedia

Table 1 News coverage, public interest, race, gender, and birth year of US House members. Random sample of size 100

Name	News coverage	Public interest	Black	Woman	Birth year
Jim Talent	17,469	1,989	0	0	1956
Duncan Hunter	77,371	6,198	0	0	1948
Kirsten Gillibrand	45,187	34,152	0	1	1966
Dana Rohrabacher	25,766	2,774.5	0	0	1947
Carolyn Cheeks Kilpatrick	7,120	2,026.5	1	1	1945
John Boozman	11,202	954.5	0	0	1950
Vern Buchanan	57,901	1329	0	0	1951
Mac Thornberry	8,507	849.5	0	0	1958
Les Aspin	554	2,632.5	0	0	1938
David Dreier	61,075	2,666	0	0	1952
Greg Laughlin	799	439	0	0	1942
Corrine Brown	5,797	2,803	1	1	1946
Julia Carson	10,711	1,886.5	1	1	1938
Olympia Snowe	90,274	12,712	0	1	1947
Ernest Istook	9,190	951.5	0	0	1950
Ric Keller	5,845	844.5	0	0	1964
Fred Upton	8,447	1,260.5	0	0	1953
Shelley Berkley	17,312	1,549	0	1	1951
John E. Sununu	10,717	9,575.5	0	0	1964
Anna Eshoo	9,686	1,894.5	0	1	1942
Cass Ballenger	432	375	0	0	1926
Marilyn Musgrave	24,520	1,781.5	0	1	1949
Nathan Deal	5,075	914.5	0	0	1942
Lindsey Graham	145,507	9313	0	0	1955
Anne Northup	14,772	573.5	0	1	1948
Joe Donnelly	11,197	1,209	0	0	1955
Mickey Leland	497	1,301.5	1	0	1944
John Conyers	60,858	9,127	1	0	1929
Don Cazayoux	8,612	1,106	0	0	1964
Joan Kelly Horn	13	215	0	1	1936
John Sarbanes	1,327	1625	0	0	1961
Nita Lowey	6,525	1,455.5	0	1	1937
Norman Sisisky	10	212	0	0	1927
George Radanovich	8,317	999.5	0	0	1955
Henry Hyde	29,062	4,453.5	0	0	1924
George Gekas	479	427.5	0	0	1930
Joe Hoeffel	3,338	1,037.5	0	0	1950
Debbie Halvorson	2,717	1,531.5	0	1	1958
Robin Hayes	12,857	869.5	0	0	1945
Rick Santorum	97,721	17,551	0	0	1958
Tammy Baldwin	10,116	4,984.5	0	1	1962
Ed Whitfield	8,074	740	0	0	1943
Joe Sestak	21,412	2,078.5	0	0	1951
Rod Grams	3,498	955	0	0	1948
George Nethercutt	5,393	858	0	0	1944
Jerrold Nadler	10,718	2,318.5	0	0	1947

Table 1. Continued

Name	News coverage	Public interest	Black	Woman	Birth year
Edolphus Towns	4,604	1,025	1	0	1934
Karen McCarthy	1,919	410	0	1	1947
Chip Pickering	5,045	1,168.5	0	0	1963
Steven Schiff	396	210	0	0	1947
Jim Moran	11,053	2,593	0	0	1945
Bob Filner	10,587	1,345	0	0	1942
Jon Porter	23,032	938	0	0	1955
Jo Bonner	2,672	1,071.5	0	0	1959
Richard Pombo	49,650	1,277.5	0	0	1961
Paul Ryan	28,962	1,844.5	0	0	1970
John Linder	4,850	1490	0	0	1942
Martin Olav Sabo	421	467.5	0	0	1938
Debbie Stabenow	42,130	5,077.5	0	1	1950
David Funderburk	302	213	0	0	1944
Lynn Schenk	203	390.5	0	1	1945
Sonny Bono	13,736	38,145	0	0	1935
Frank Mascara	30	247	0	0	1930
Pat Danner	11	266	0	1	1934
Gwen Moore	4,578	1,476.5	1	1	1951
Mark Udall	39,524	6,102	0	0	1950
Devin Nunes	13,151	1,253	0	0	1973
Cedric Richmond	3,500	3	1	0	1973
Jay Dickey	1,355	316.5	0	0	1939
Heather Wilson	35,343	2,468	0	1	1960
Zack Space	12,696	1,476	0	0	1961
Enid Greene	296	326	0	1	1958
Chris Shays	12,299	1,241	0	0	1945
Alan Wheat	252	287	1	0	1951
Kweisi Mfume	15,186	3,099	1	0	1948
Steve Largent	6,658	4,269.5	0	0	1954
Marcia Fudge	1,962	2,031.5	1	1	1952
Wes Watkins	1,161	342	0	0	1938
Dave McCurdy	4,163	622.5	0	0	1950
Jim Bunning	52,880	11,814	0	0	1931
Mario Diaz-Balart	7,378	1,788	0	0	1961
Lois Capps	14,684	1,323.5	0	1	1938
Roy Blunt	26,906	4,648.5	0	0	1950
Shirley Chisholm	10,533	27,774	1	1	1924
Sandy Adams	1,275	76	0	1	1956
Ralph Hall	5,853	1290	0	0	1923
Bill Zeliff	313	234.5	0	0	1936
Glenn English	576	292.5	0	0	1940
Mike Conaway	8,378	698	0	0	1948
Lloyd Doggett	11,841	1,549.5	0	0	1946
Ben Cardin	13,377	4,561.5	0	0	1943
Thelma Drake	1,871	952	0	1	1949

Table 1. Continued

Name	News coverage	Public interest	Black	Woman	Birth year
Gary Miller	23,724	1878	0	0	1948
Peter Roskam	12,271	1863	0	0	1961
Gil Gutknecht	6,423	543.5	0	0	1951
Melissa Bean	13,070	2,137.5	0	1	1962
Douglas Applegate	40	153.5	0	0	1928
Jo Ann Emerson	7,844	1474	0	1	1950
Marge Roukema	171	412.5	0	1	1929
Niki Tsongas	6,258	2,078.5	0	1	1946

Table 2 Descriptive statistics of variables. Omitted are descriptive statistics of category and sub-category variables, whose case numbers can be found in Tables 3 and 4, columns 3 and 4, broken down by race. News coverage and public interest are logged in the analysis

Variable	Mean	SD	Min	Max	N
News coverage	6,889	86.8×10^3	1	1.08×10^7	31,724
Public interest	6,740	25.9×10^3	0	1.37×10^6	31,724
Black (vs. White)	0.143	n.a.	0	1	31,724
Woman (vs. Man)	0.319	n.a.	0	1	31,724
Birth year	1955	18	1920	2000	31,724

page. Across categories and sub-categories, the regression coefficient for race is either significantly in favour of Blacks or insignificant. There are two exceptions: The larger politics category shows negative coefficients, in favour of Whites, both with and without controls for gender and age. However, in all politics sub-categories effects are either insignificant or in favour of Blacks, suggesting an ecological fallacy with less important sub-categories of politics (e.g. mayors) having greater African American representation. The other exception is crime. This category mostly consists of murderers, for whom the negative effect turns insignificant once age and gender are controlled. The overall result falsifies H2, which suggested that White individuals would generate more public interest when compared with structurally equivalent Black individuals. Although we can only speculate on the drivers of greater public interest in Blacks, these results are consistent with the literature about greater selectivity of Blacks in prominent positions. They are also consistent with higher standards for Blacks' inclusion in Wikipedia categories, apart from the politics categories that contain all White and Black members. Table 4 stresses the importance of accounting for public interest in the assessment of racial bias in news coverage.

Domain-stratified media supply, accounting for public interest

The final column of Table 3 assesses racial differences in media coverage while accounting for both structural position and public interest. It reports race coefficients for regression models controlling for gender, age, and public interest. Overall, we do not find support for H3. The excess coverage documented in columns 5 and 6 of Table 3 largely goes away in column 7, suggesting that it was mostly due to the excess public interest in Blacks evidenced in Table 4. In some categories, significant differences remain, but these are as often in favour of Blacks as in favour of Whites, and some differences are necessarily incidental due to multiple comparisons. Most racial coverage differences in entertainment disappear, except for a negative race effect for film directors and a positive race effect for radio personalities. In politics, differences become smaller, with presidential candidates and House representatives still showing a White advantage, while the coverage of mayors favours Blacks. If anything, business coverage favours Blacks, but low numbers prevent confident interpretation and the Wikipedia selection for this category, as we note above, is particularly problematic. Coverage of sports favours White soccer players but Black Olympians. Crime and science show no race differences, and in

Table 3 OLS regression of *newspaper coverage* ($\log_{10}$) on race, gender, age (cubic), and public interest ($\log_{10}$) with heteroscedasticity-robust standard errors. Each row is a separate regression model including only cases from the corresponding sub-category. Only race effects (Black vs. White) are reported

Category	Sub-category	Number of cases		Model 1: No controls	Model 2: Gender, age	Model 3: Gender, age, interest
		Black	White	β_{-Black} (SE)	β_{-Black} (SE)	β_{-Black} (SE)
All categories	All names	4,534	27,190	+0.352 (0.018)***	+0.283 (0.018)***	+0.239 (0.016)***
Entertainment (United States)	All entertainers	1,321	12,713	+0.191 (0.032)***	+0.187 (0.031)***	−0.021 (0.025)
	Film actors	804	7,787	+0.223 (0.040)***	+0.195 (0.040)***	+0.010 (0.032)
	TV actors	789	7,659	+0.163 (0.040)***	+0.144 (0.040)***	−0.026 (0.032)
	Film directors	77	1,505	+0.004 (0.138)	+0.004 (0.133)	−0.199 (0.097)*
	Singers	246	2,485	+0.251 (0.070)***	+0.314 (0.069)***	+0.087 (0.054)
	Pop singers	55	527	+0.368 (0.168)*	+0.383 (0.174)*	+0.020 (0.141)
	Dancers	61	410	+0.533 (0.166)**	+0.504 (0.173)**	+0.078 (0.133)
	TV personalities	145	710	+0.061 (0.104)	−0.007 (0.104)	−0.113 (0.084)
	Radio personalities	48	454	+0.451 (0.145)**	+0.446 (0.142)**	+0.283 (0.108)**
	Oscar nominees	41	454	+0.043 (0.166)	−0.096 (0.157)	−0.039 (0.080)
	Emmy nominees	38	516	+0.071 (0.123)	+0.023 (0.117)	−0.002 (0.090)
Politics (United States)	All politicians	603	1,672	−0.557 (0.054)***	−0.519 (0.052)***	−0.284 (0.041)***
	Governors	3	299	−0.560 (1.22)	−0.579 (0.994)	−1.08 (0.558)
	Mayors	101	335	+0.547 (0.120)***	+0.517 (0.113)***	+0.211 (0.096)*
	Cabinet secretaries	14	67	−0.688 (0.372)	−0.438 (0.363)	−0.176 (0.195)
	Presidential candidates	23	105	−0.366 (0.316)	−0.216 (0.318)	−0.513 (0.217)*
	Judges	120	293	−0.005 (0.106)	−0.049 (0.122)	−0.045 (0.107)
	House representatives	74	572	−0.449 (0.123)***	−0.276 (0.123)*	−0.402 (0.101)***
	Senators	4	178	+0.719 (0.644)	+0.757 (0.636)	−0.550 (0.346)
Business (United States)	All business people	122	1,279	+0.660 (0.112)***	+0.452 (0.109)***	+0.011 (0.080)
	Chief executives	14	353	+0.917 (0.314)**	+0.931 (0.336)**	+0.131 (0.226)
	Business executives	11	89	+0.704 (0.277)*	+0.641 (0.306)*	+0.533 (0.228)*
	Founders	2	67	+1.31 (0.669)	+1.04 (0.832)	+0.540 (0.813)
	Billionaires	2	282	+1.39 (0.636)*	+1.78 (0.672)**	+1.05 (0.250)***
Sports	All athletes	1,885	2,395	+0.191 (0.034)***	+0.043 (0.032)	−0.258 (0.028)***
	Tennis GS champions	7	113	+0.504 (0.405)	+0.285 (0.346)	−0.076 (0.171)
	Soccer players (US)	62	360	+0.124 (0.145)	−0.087 (0.128)	−0.218 (0.097)**
	Basketball players (US)	419	1,686	+0.354 (0.059)***	+0.104 (0.056)	−0.029 (0.050)
	Golf players (US)	21	1,238	−0.253 (0.164)	−0.041 (0.211)	−0.188 (0.143)
	Olympic athletes (US)	99	277	+0.581 (0.130)***	+0.582 (0.128)***	+0.309 (0.085)***
Crime	All criminals	144	535	+0.054 (0.101)	+0.001 (0.103)	+0.060 (0.100)
	Murderers	139	450	+0.116 (0.103)	+0.064 (0.105)	+0.101 (0.103)
Arts and literature	All artists and writers	425	4,203	+0.325 (0.055)***	+0.312 (0.054)***	+0.142 (0.043)**
	Painters	34	783	+0.508 (0.145)***	+0.490 (0.149)**	+0.433 (0.143)**
	Photographers	14	709	+0.223 (0.176)	+0.226 (0.164)	+0.143 (0.146)
	Sculptors	6	217	+0.149 (0.134)	+0.123 (0.153)	+0.144 (0.180)
	Writers	317	2,243	+0.197 (0.068)***	+0.194 (0.067)**	+0.033 (0.051)
	Pulitzer prize winners	8	244	+0.348 (0.182)	+0.276 (0.179)	+0.029 (0.155)
Science	All scientists	69	3,232	+0.281 (0.136)*	+0.242 (0.134)	+0.139 (0.105)
	Nobel prize winners	4	149	+1.87 (0.989)	+1.41 (1.19)	−0.258 (0.334)

*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ (two-tailed tests).
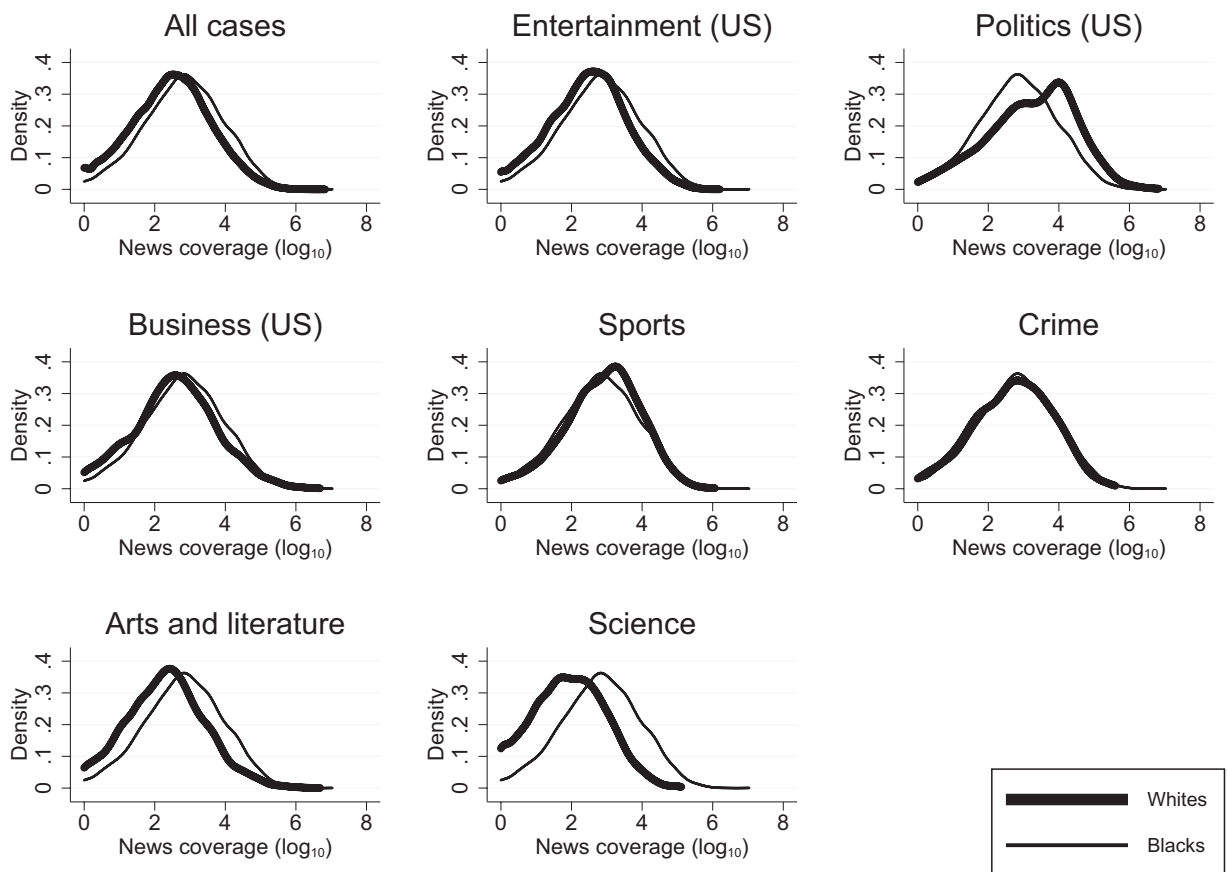
Figure 3 Density plots of news coverage (log10) by race and major categories

the arts and literature category, only painters maintain a Black coverage advantage once public interest is controlled.

Discussion and conclusion

In this article, we reassessed common explanations for inequalities in the media coverage of White and Black individuals, using a novel methodological approach. More specifically, we explored the possibility that newspaper coverage might vary by race among populations of equal structural position, as well as public interest. Comparing about 32,000 prominent Black and White individuals by gender, age, domain, structural position, and public interest, we found that newspapers overall awarded prominent Black individuals with a larger coverage volume when compared to White counterparts. These results contrast with a similar analysis for gender, in which we found that prominent women received significantly less coverage when compared to male counterparts despite women enjoying greater public interest (Shor et al., 2019). Naturally, having

examined a great number of distinct structural positions, the results of the present study varied by domain. We refrain from attaching confident interpretations to results pertaining to any one sub-category due to the potential for incidental findings when conducting multiple comparisons and category-specific idiosyncrasies.

Even when examining the results for the entire sample we should be cautious in interpreting the overall coverage advantage we found for prominent Black individuals and the lack of clear evidence for bias once racial differences in public interest are corrected for. Our sample, while diverse, might systematically fail to include some meaningful publicly visible individuals, as its composition is largely dictated by the availability of category membership lists on Wikipedia. The lack of evidence for systematic bias against Blacks might therefore be the result of both higher career selectivity and barriers to inclusion on Wikipedia. In addition, one may speculate that journalists and editors, aware of both racial injustices and academic and public critiques of the media, consciously write more about prominent Black individuals. They may either do that in a sincere

Table 4 OLS regression of *public interest* (log$_{10}$) on race, gender, and age (cubic) with heteroscedasticity-robust standard errors. Each row is a separate regression model including only cases from the corresponding sub-category. Only race effects (Black vs. White) are reported

Category	Sub-category	Number of cases		Model 1: No controls	Model 2: Gender, age
		Black	White	β_{-Black} (SE)	β_{-Black} (SE)
All categories	All names	4,534	27,190	+0.116 (0.017)***	+0.087 (0.017)***
Entertainment (United States)	All entertainers	1,321	12,713	+0.343 (0.027)***	+0.319 (0.027)***
	Film actors	804	7,787	+0.292 (0.033)***	+0.251 (0.033)***
	TV actors	789	7,659	+0.261 (0.033)***	+0.232 (0.032)***
	Film directors	77	1,505	+0.275 (0.122)*	+0.287 (0.122)*
	Singers	246	2,485	+0.336 (0.064)***	+0.374 (0.062)***
	Pop singers	55	527	+0.591 (0.136)***	+0.540 (0.144)***
	Dancers	61	410	+0.735 (0.164)***	+0.708 (0.171)***
	TV personalities	145	710	+0.217 (0.085)*	+0.195 (0.087)*
	Radio personalities	48	454	+0.396 (0.150)**	+0.309 (0.147)*
	Oscar nominees	41	454	+0.046 (0.109)	–0.043 (0.101)
	Emmy nominees	38	516	+0.049 (0.087)	+0.024 (0.086)
Politics (United States)	All politicians	603	1,672	–0.356 (0.046)***	–0.293 (0.046)***
	Governors	3	299	+0.454 (0.479)	+0.462 (0.410)
	Mayors	101	335	+0.585 (0.114)***	+0.584 (0.117)***
	Cabinet secretaries	14	67	–0.342 (0.258)	–0.237 (0.247)
	Presidential candidates	23	105	+0.192 (0.244)	+0.266 (0.250)
	Judges	120	293	+0.011 (0.100)	–0.006 (0.132)
	House representatives	74	572	+0.043 (0.085)	+0.126 (0.076)
	Senators	4	178	+1.09 (0.422)*	+1.41 (0.499)**
Business (United States)	All business people	122	1,279	+0.833 (0.131)***	+0.703 (0.138)***
	Chief executives	14	353	+1.37 (0.462)**	+1.52 (0.496)**
	Business executives	11	89	+0.255 (0.401)	+0.305 (0.419)
	Founders	2	67	+1.07 (0.145)***	+0.844 (0.232)**
	Billionaires	2	282	+1.08 (0.665)	+0.996 (0.678)
Sports	All athletes	1,885	2,395	+0.554 (0.029)***	+0.502 (0.029)***
	Tennis GS champions	7	113	+0.469 (0.275)	+0.348 (0.247)
	Soccer players (United States)	62	360	+0.288 (0.133)*	+0.182 (0.110)
	Basketball players (United States)	419	1,686	+0.373 (0.045)***	+0.184 (0.044)
	Golf players (United States)	21	1,238	+0.256 (0.211)	+0.251 (0.180)
	Olympic athletes (United States)	99	277	+0.381 (0.120)**	+0.336 (0.114)**
Crime	All criminals	144	535	–0.278 (0.089)**	–0.190 (0.093)*
	Murderers	139	450	–0.242 (0.089)**	–0.142 (0.093)
Arts and literature	All artists and writers	425	4,203	+0.304 (0.056)***	+0.298 (0.056)***
	Painters	34	783	+0.150 (0.169)	+0.146 (0.171)
	Photographers	14	709	+0.205 (0.205)	+0.206 (0.206)
	Sculptors	6	217	+0.051 (0.343)	–0.072 (0.365)
	Writers	317	2,243	+0.264 (0.064)***	+0.253 (0.065)***
	Pulitzer prize winners	8	244	+0.228 (0.294)	+0.309 (0.293)
Science	All scientists	69	3,232	+0.212 (0.143)	+0.226 (0.141)
	Nobel prize winners	4	149	+1.11 (0.731)	+1.41 (0.838)

*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ (two-tailed tests).

attempt to reduce public bias and promote racial diversity or they may simply try to show that they hold no racial prejudice and that those who truly 'deserve' the attention receive coverage regardless of their race.

Some but not all of the coverage patterns are consistent with racial stereotypes. The coverage advantage in sports is in line with sociological and mass communication research on sports that has repeatedly noted stereotypical perceptions among journalists and the public that Black athletes would possess a *natural* physical superiority over White athletes (Entine, 2000; Buffington, 2005), making them more likely to succeed in many sports. Another common 'positive' stereotype about Blacks is that they are very musical and rhythmic, excelling in singing, dancing, and other types of physical performance (Madon, Guyll and Aboufadel, 2001; Czopp and Monteith, 2006). The partial evidence for a Black coverage advantage in entertainment when public interest is not controlled permits the interpretation that such stereotypes guide media coverage patterns, with journalists celebrating the performativity of the Black body and its 'natural' inclination for rhythm and movement.

Our null findings regarding the relative coverage of Black and White criminals should be treated cautiously given the limited sample size in this category. Still, these findings do not correspond with traditional images of Blacks as wild, untamed, violent, and criminal (Hughey and Hernandez, 2013; Dixon, 2017) and with the findings of earlier research suggesting coverage bias (Entman, 1992; Entman and Rojecki, 2001; Dixon, Azocar and Casas, 2003; Dixon, 2008b). However, our finding are in line with later studies, reporting that in recent years Blacks have become proportionally represented in television news as both perpetrators and victims of crimes (Dixon and Williams, 2015; Dixon, 2017).

The mirror image of stereotypes about Blacks' tendency to be more violent and their natural physical/athletic/rhythmic superiority is a belief in Blacks' moral, cognitive, and leadership inferiority (Miller, 1998; Harrison and Lawrence, 2004; Harrison, Azzarito and Burden, 2004; Buffington, 2005). However, we find little evidence that such stereotypical beliefs lead to reduced coverage of Black politicians or businesspeople who have reached prominent positions in their respective fields. Despite any tendency to associate political leadership and managerial competence with Whites (Ridgeway and Kricheli-Katz, 2013; Gundemir et al., 2014), Black leaders in politics and business receive overall more news coverage than their White counterparts, although these differences mostly disappear when controlling for public interest. An exception is US House representatives, where we do find evidence that Black members on average receive less coverage than White members of equal public interest.

Our study contributes not only to the mass communication literature on racial coverage differentials in the media, but also to the larger literature on cultural and occupational racial inequalities. Our findings demonstrate that theoretical propositions highlighting either structural inequalities or individual differences in charisma, motivation, effort, or talent are insufficient when trying to account for the racial coverage gap. While inequalities and discrimination in hiring and in promotion practices are an important part of the story (Pager, 2007; Couch and Fairlie, 2010; Ridgeway and Kricheli-Katz, 2013), our findings suggest that they cannot fully explain coverage differences between White and Black individuals. Neither can explanations that focus on potential differences between Blacks and Whites in charisma, personal characteristics, motivations, and career choices fully account for differential coverage patterns. We have argued that such differences, if and when they exist, should be, at least in part, manifested in levels of public interest. However, even when accounting for measures of public interest, we continue to observe some racialized coverage patterns.

While our findings offer important contributions to the understanding of media coverage dynamics, they leave open many additional interesting questions. First, our news coverage data become too thin when broken down by time. Future research endeavours might attempt to use a time-sensitive research design that would match individuals who appear similar at t1 and track changes in their media coverage over time to better establish causality. Another important task for future research would be to examine more closely the content of the coverage. Our analyses are unable to inspect the differential ways in which Black and White individuals are portrayed in the news, which has been the subject of ample previous research. A large-scale investigation of the sentiment associated with Black and White individuals in the news, potentially using a similar controlled design, could further enrich our understanding of racial media coverage patterns.

Notes

1. See a discussion of the Wikipedia page views measure in section 4.3.2.
2. The median monthly number of Wikipedia page views is a relatively straightforward measure of public interest. Other potential measures, such as article length or the mere number of articles in Wikipedia may be a function of those who commonly write and edit Wikipedia pages—predominantly men.

Supplementary Data

Supplementary data are available at *ESR* online.

Author contributions

Eran Shor (Conceptualization-Lead, Data curation-Lead, Formal analysis-Supporting, Funding acquisition-Equal, Investigation-Lead, Methodology-Supporting, Resources-Lead, Software-Equal, Supervision-Lead, Validation-Equal, Visualization-Equal, Writing – original draft-Lead, Writing – review & editing-Lead), Arnout van de Rijt (Conceptualization-Supporting, Data curation-Supporting, Formal analysis-Lead, Funding acquisition-Equal, Investigation-Equal, Methodology-Lead, Project administration-Supporting, Resources-Equal, Software-Equal, Supervision-Supporting, Validation-Supporting, Visualization-Equal, Writing – original draft-Supporting, Writing – review & editing-Supporting).

Data availability

The data and Stata code for reproducing the results reported in this article are available at https://osf.io/bdcek/.

References

Adams, J., Brückner, H. and Naslund, C. (2019). Who counts as a notable sociologist on Wikipedia? Gender, race, and the 'professor test'. *Socius*, **5**, 1–14.

American Society of News Editors (ASNE) (2018). Minority percentages at participating print and digital newspapers. https://members.newsleaders.org/diversity-survey-percent.

Baldi, S. and Branch McBrier, D. (1997). Do the determinants of promotion differ for blacks and whites? Evidence from the U.S. Labor Market. *Work and Occupations*, **24**, 478–498.

Berkhus, W. (1998). A sociology of the unmarked: redirecting our focus. *Sociological Theory*, **16**, 34–51.

Buffington, D. (2005). Contesting race on Sundays: making meaning out of the rise in the number of black quarterbacks. *Sociology of Sport Journal*, **22**, 19–37.

Canon, D. (1999). *Race, Redistricting, and Representation: The Unintended Consequences of Black Majority Districts*. Chicago: Chicago University Press.

Cassano, J. (2015). Black history matters, so why is Wikipedia missing so much of it? *Fast Company*. https://www.fastcompany.com/3041572/black-history-matters-so-why-is-wikipedia-missing-so-much-of-it.

Collins, S. (1997). Black mobility in white corporations: up the corporate ladder but out on a limb. *Social Problems*, **44**, 55–68.

Cook, A. and Glass, C. (2013). Glass cliffs and organizational saviors: barriers to minority leadership in work organizations?. *Social Problems*, **60**, 168–187.

Cooky, C., Messner, M. and Hextrum, R. (2013). Women play sport, but not on TV: a longitudinal study of televised news media. *Communication & Sport*, **1**, 203–230.

Couch, M. and Fairlie, R. (2010). Last hired, first fired? Black–White unemployment and the business cycle. *Demography*, **47**, 227–247.

Czopp, A. and Monteith, M. (2006). Thinking well of African Americans: measuring complimentary stereotypes and negative prejudice. *Basic and Applied Social Psychology*, **28**, 233–250.

Davenport, C., Soule, S. and Armstrong, D. (2011). Protesting while Black? The differential policing of American Activism, 1960 to 1990. *American Sociological Review*, **76**, 152–178.

Dixon, T. (2008a). Crime news and racialized beliefs: understanding the relationship between local news viewing and perceptions of African Americans and crime. *Journal of Communication*, **58**, 106–125.

Dixon, T. (2008b). Network news and racial beliefs: exploring the connection between national television news exposure and stereotypical perceptions of African Americans. *Journal of Communication*, **58**, 321–337.

Dixon, T. (2017). Good guys are still always in white? Positive change and continued misrepresentation of race and crime on local television news. *Communication Research*, **44**, 775–792.

Dixon, T., Azocar, C. and Casas, M. (2003). The portrayal of race and crime on television network news. *Journal of Broadcasting & Electronic Media*, **47**, 498–523.

Dixon, T. and Williams, C. (2015). The changing misrepresentation of race and crime on network and cable news. *Journal of Communication*, **65**, 24–39.

Elwert, F. and Winship, C. (2014). Endogenous selection bias: the problem of conditioning on a collider variable. *Annual Review of Sociology*, **40**, 31–53.

Entine, J. (2000). *Taboo: Why Black Athletes Dominate Sports and Why We're Afraid to Talk about It*. New York: Public Affairs.

Entman, R. (1992). Blacks in the news: television, modern racism and culture change. *Journalism Quarterly*, **69**, 341–361.

Entman, R., and Rojecki, A. (2001). *The Black Image in the White Mind: Media and Race in America*. Chicago, IL: University of Chicago Press.

Eschholz, S., Bufkin, J. and Long, J. (2002). Symbolic reality bites: Women and racial/ethnic minorities in modern film. *Sociological Spectrum*, **22**, 299–334.

Falk, E. (2010). *Women for President: Media Bias in Nine Campaigns*. Champaign, IL: University of Illinois Press.

Gaddis, M. (2015). Discrimination in the credential society: An audit study of race and college selectivity in the labor market. *Social Forces*, **93**, 1451–1479.

Gans, H. (1979). *Deciding What's News*. New York: Pantheon Books.

Gundemir, S. *et al*. (2014). Think leader, think white? Capturing and weakening an implicit pro-white leadership bias. *PLoS One*, **9**, e83915.

Harrison, K. and Lawrence, S. (2004). College students' perceptions, myths, and stereotypes about African American athleticism: a qualitative investigation. *Sport, Education and Society*, **9**, 33–52.

Harrison, L., Azzarito, L. and Burden, J. (2004). Perceptions of athletic superiority: a view from the other side. *Race, Ethnicity and Education*, **7**, 149–166.

Hier, S. (2009). Beyond vernacular commentary: a response to Minelle Mahtani. *Canadian Journal of Communication*, **35**, 173–178.

Hughey, M. and Hernandez, M. (2013). Black, Greek, and read all over: newspaper coverage of African American fraternities and sororities, 1980–2009. *Ethnic and Racial Studies*, **36**, 298–319.

Husband, C. (2005). Minority ethnic media as communities of practice: professionalism and identity politics in interaction. *Journal of Ethnic and Migration Studies*, **31**, 461–479.

Iyengar, S. (1990). Framing responsibility for political issues: the case of poverty. *Political Behavior*, **12**, 19–40.

Iyengar, S. and Hahn, K. (2009). Red media, blue media: evidence of ideological selectivity in media use. *Journal of Communication*, **59**, 19–39.

Krivo, L., Peterson, R. and Kuhl, D. (2009). Segregation, racial structure, and neighborhood violent crime. *American Journal of Sociology*, **114**, 1765–1802.

Lamont, M. (2012). Toward a comparative sociology of valuation and evaluation. *Annual Review of Scoiology*, **38**, 201–221.

Lieberson, S. and Mikelson, K. (1995). Distinctive African American names: an experimental, historical, and linguistic analysis of innovation. *American Sociological Review*, **60**, 928–964.

MacDonald, F. (1983). *Blacks and white TV: African Americans in television since 1948/J. Fred MacDonald*. Chicago, IL: Nelson-Hall Publishers.

Madon, S. *et al.* (2001). Ethnic and national stereotypes: the Princeton trilogy revisited and revised. *Personality and Social Psychology Bulletin*, **27**, 996–1010.

Mastro, D. and Tukachinsky, R. (2012). Cultivation of perceptions of marginalized communities. In Morgan, M., Shanahan, J. and Signorielli, N. (Eds.), *Living with Television Now: Advances in Cultivation Theory and Research*. New York: Peter Lang, pp. 38–60.

Maume, D. (1999). Glass ceilings and glass escalators: occupational segregation and race and sex differences in managerial promotions. *Work and Occupations*, **26**, 483–509.

McCombs, M. E. and Shaw, D. L. (1972). The agenda-setting function of mass media. *Public Opinion Quarterly*, **36**, 176–187.

McIlwain, C. (2011). Racialized media coverage of minority candidates in the 2008 democratic presidential primary. *American Behavioral Scientist*, **55**, 371–389.

Miller, P. (1998). The anatomy of scientific racism: racialist responses to black athletic achievement. *Journal of Sport History*, **25**, 119–151.

Nghiem, L. *et al.* (2016). Analysis of the capacity of Google trends to measure interest in conservation topics and the role of online news. *PLoS One*, **11**, e0152802.

Pager, D. (2007). *Marked: Race, Crime, and Finding Work in an Era of Mass Incarceration*. Chicago, IL: University of Chicago Press.

Pager, D., Western, B. and Bonikowski, B. (2009). Discrimination in a low-wage labor market: a field experiment. *American Sociological Review*, **74**, 777–799.

Phillips, J. (2002). White, Black, and Latino homicide rates: why the difference? *Social Problems*, **49**, 349–373.

Powell, G. and Butterfield, A. (1997). Effect of race on promotions to top management in a federal department. *Academy of Management Journal*, **40**, 112–128.

Pritchard, D. and Stonbely, S. (2007). Racial profiling in the newsroom. *Journalism & Mass Communication Quarterly*, **84**, 231–248.

Ridgeway, C. (2013). Why status matters for inequality. *American Sociolgical Review*, **79**, 1–16.

Ridgeway, C. and Kricheli-Katz, T. (2013). Intersecting cultural beliefs in social relations: gender, race, and class binds and freedoms. *Gender & Society*, **27**, 294–318.

Ripberger, J. (2011). Capturing curiosity: using internet search trends to measure public attentiveness. *Policy Studies Journal*, **39**, 239–259.

Rivas-Rodriguez, M. *et al.* (2004). Minority journalists' perceptions of the impact of minority executives. *Howard Journal of Communications*, **15**, 39–55.

Rivera, L. (2012). Hiring as cultural matching: the case of elite professional service firms. *American Sociological Review*, **77**, 999–1022.

Shor, E., Van de Rijt, A. and Fotouhi, B. (2019). A large-scale test of gender bias in the media. *Sociological Science*, **6**, 526–550.

Silverman, B. (1986). *Density Estimation for Statistics and Data Analysis*. London: Chapman & Hall.

Skiena, S. and Ward, C. (2014). *Who's Bigger? Where Historical Figures Really Rank*. Cambridge: Cambridge University Press.

Smith, G. and Searles, K. (2014). Who let the (attack) dogs out? New evidence for partisan media effects. *Public Opinion Quarterly*, **78**, 71–99

Smith, R. (2005). Do the determinants of promotion differ for white men versus women and minorities? *American Behavioral Scientist*, **48**, 1157–1181

Smith, R. and Elliott, J. (2002). Does ethnic concentration influence employees' access to authority? An examination of contemporary urban labor markets. *Social Forces*, **81**, 255–279.

Solomon, D., Soltes, E. and Sosyura, D. (2014). Winners in the spotlight: media coverage of fund holdings as a driver of flows. *Journal of Financial Economics*, **113**, 53–72.

Sorensen, A. (2007). Bestseller lists and product variety. *The Journal of Industrial Economics*, **55**, 715–738.

Stainback, K., Tomaskovic-Devey, D. and Skaggs, S. (2010). Organizational approaches to inequality: inertia, relative power, and environments. *Annual Review of Sociology*, **36**, 225–247.

Sui, M. *et al.* (2018). The role of minority journalists, candidates, and audiences in shaping race-related campaign news coverage. *Journalism & Mass Communication Quarterly*, **95**, 1079–1102.

Terkildsen, N. and Damore, D. (1999). The dynamics of racialized media coverage in congressional elections. *The Journal of Politics*, **61**, 680–699.

Tolley, E. (2015). Racial mediation in the coverage of candidates' political viability: a comparison of approaches. *Journal of Ethnic and Migration Studies*, **41**, 963–984.

Tripodi, F. (2021). Ms. categorized: gender, notability, and inequality on Wikipedia. *New Media & Society*, 1–21.

Williams, J. and Multhaup, M. (2018). For women and minorities to get ahead, managers must assign work fairly. *Harvard Business Review*, **5**.

Wilson, G., Roscigno, V. and Huffman, M. (2013). Public sector transformation, racial inequality and downward occupational mobility. *Social Forces*, **91**, 975–1006.

Yoshida, M., Arase, Y., Tsunoda, T. and Tamamoto, M. (2015). Wikipedia page view reflects web search trend. In Proceedings of the ACM Web Science Conference **65**, pp. 1–2. https://doi.org/10.1145/2786451.2786495.

Zhang, L. (2008). Gender and racial gaps in earnings among recent college graduates. *Review of Higher Education*, **32**, 51–72.

Zilber, J. and Niven, D. (2000a). *Racialized Coverage of Congress: The News in Black and White*. Westport, CT: Praeger.

Zilber, J. and Niven, D. (2000b). Stereotypes in the news: media coverage of African-Americans in Congress. *The International Journal of Press/Politics*, **5**, 32–49.