Race below the fold: Race-evasiveness in the news media's coverage of student loans

Dominique J. Baker
Southern Methodist University

Lauren Mena Shook
University of Texas at Austin

Jaime Ramirez-Mendoza
University of California, Davis

Christopher T. Bennett
RTI International

The media discourse on student loans plays a significant role in the way that policy actors conceptualize challenges and potential solutions related to student debt. This study examines the racialized language in student loan news articles published in eight major news outlets between 2006 and 2021. We found that 18% of articles use any racialized language, though use has accelerated since 2018. This increase appears to be driven by terms that denote groups of people instead of structural problems, with 8% of articles mentioning "Black" but less than 1% mentioning "racism." These findings emphasize the importance of treating the media as a policy actor capable of shaping the salience of racialization in discussions about student loans.

Race below the fold: Race-evasiveness in the news media's coverage of student loans

Dominique J. Baker, *Southern Methodist University*
Lauren Mena Shook, *University of Texas at Austin*
Jaime Ramirez-Mendoza, *University of California, Davis*
Christopher T. Bennett, *RTI International*

Abstract:
The media discourse on student loans plays a significant role in the way that policy actors conceptualize challenges and potential solutions related to student debt. This study examines the racialized language in student loan news articles published in eight major news outlets between 2006 and 2021. We found that 18% of articles use any racialized language, though use has accelerated since 2018. This increase appears to be driven by terms that denote groups of people instead of structural problems, with 8% of articles mentioning "Black" but less than 1% mentioning "racism." These findings emphasize the importance of treating the media as a policy actor capable of shaping the salience of racialization in discussions about student loans.

Race below the fold: Race-evasiveness in the news media's coverage of student loans

Student loans play a critical role in helping students afford a college education. Yet the loan system is beset by long-standing inequities—developed over centuries of deliberate policymaking by federal, state, and local governments (e.g., Katznelson, 2005; Shermer, 2021)—that disproportionately affect students of color. Total U.S. student loan debt reached almost $1.6 trillion at the end of 2020 (Federal Reserve Bank of New York, 2021), but this student loan reliance is not evenly spread across the country. Systemic racism creates an environment in which students of color, especially Black students, disproportionately rely on and struggle with student loans (e.g., Addo et al., 2016; Baker, 2019; Seamster & Charron-Chénier, 2017). Compounding the inequities in students' and families' ability to afford college are structural barriers related to student loan repayment. As a result, students of color face a greater risk of defaulting on their student loans, having their credit adversely affected, having their wages garnished, and even losing Social Security benefits.

An aspect of student loan policy that has not yet been explored is how the news media shapes the student loan policy environment, particularly along racialized dimensions. Literature from political communication conceives of the media as a policy actor in its own right, as an entity that behaves separately from both the public and from policymakers (Curran et al., 2022) and actively shapes how policy issues are defined (Wolfe, 2012; Woodly, 2015). Mass media does not function primarily as a medium for communication, but rather as an actor which exerts power through, over, and within ideas (Curran et al., 2022). The media possesses power in the policymaking process through persuasion, gatekeeping policy issue definitions, and determining which information becomes a part of normative policy discourse. As a result, media "shap[es] the contours through which policy moves" (Wolfe, 2012, p. 123) by exerting influence at every

stage of policymaking, from agenda setting to policy formation and policy evaluation (Liu et al., 2019).

Scholars have found that the media's framing of policy issues influences how the public and policy actors socially construct policy challenges and potential solutions (de Benedictis-Kessner & Hankinson, 2022; Djourelova, 2023; Gerstl-Pepin, 2007; Wolfe, 2012; Woodly, 2015). The media's portrayal of policy issues, such as student loans, is not value-neutral. Scholars have found clear relationships between the media's discourse and the maintenance of elite, often White and wealthy, interests (Curran et al., 2022; Grossman, 2022; Soroka et al., 2013), including the structural maintenance of racism (van Dijk, 1989). Given that the media can reify systems of oppression (in both what is said and what is *not* said) and directly impact the policymaking environment, it is important to understand the existing frames used in media coverage of student loans so that we can evaluate and, if needed, improve the conversation about race and student loan policy.

To our knowledge, there has been no comprehensive study of the media's racialized framing of student loans, much less how that framing has shifted over time or by the source (e.g., national vs. regional outlets based in predominantly Black markets). Analysis on these topics will allow education policy scholars, the media, and the public to better understand how news policy communication can accurately reflect the racialized dimension of student loan debt and support racially just policies to address systemic oppression.

In this study, we examine the racialized framing of student loans in the media using Natural Language Processing (NLP) techniques applied to fifteen years' worth of newspaper articles (2006 to 2021), covering the resurgence in loan reliance through the nationwide protests for racial justice. The collection of news articles, referred to as a corpus, allows us to explore

temporal shifts in: a) specific words and phrases used in articles, and b) the context in which writers use those words. This study addresses the following research questions:

1) To what extent do news articles discuss student loans in ways that are racialized? How does this behavior shift over time?

2) When racialized language is present in news articles focused on student loans, are there systematic differences in how terms are used? How does this behavior shift over time?

This paper provides evidence on a critical area of education research for historically and presently marginalized students. Too often, education policy research focuses solely on the causal impact of adopted policies. This choice ignores other important facets of the policy process, such as policy communication and the social construction of policy problems and solutions, in the shaping of policies, their implementation, and effectiveness. By ignoring these other areas of the policy process, stakeholders may create a decontextualized and race-neutral understanding of student loan debt—one that blames certain students for making poor choices instead of focusing on the failures of public policy and potential avenues for improvement.

Literature Review

In this study, we build upon two different bodies of research to better understand how the media covers, or fails to cover, the racialization of student loans. The first concerns race, racism, and student loans, which demonstrates that student loan policy is a racialized issue disproportionately affecting borrowers of color, especially Black borrowers. The second strand examines literature on the role of media in policy formation, particularly how norms dictate coverage of complex policy issues and prime audiences for particular policy solutions. Our study contributes to these bodies of literature by identifying how the media frames student loans in race-evasive or race-conscious ways and how this varies across time and media outlet.

Race, racism, and student loans

Research demonstrates that student loans are a raced and gendered issue (Houle & Addo, 2019; Seamster & Charron-Chenier 2017). Data on bachelor's degree earners highlights the racialized nature of reliance on student loans. In the most recent data, 86% of Black bachelor's degree earners borrow, compared to 59%, 67%, and 70% for Asian, Hispanic, and White recipients, respectively (Baum et al., 2019). While often overlooked due to smaller population shares, American Indian/Alaska Native and Native Hawaiian/Other Pacific Islander students also borrow at above-average rates (76% and 90% for bachelor's degree earners, respectively, compared to an overall average of 69%) (Taylor et al., 2020). When compared to their White peers, students of color also tend to have higher rates of default (Taliaferro et al., 2021). Further, while on average women are more likely to borrow than men, one study conducted at the state level found that women of color—and Black women in particular—drive these findings (Baker, 2019). Due to the disproportionate burden carried by Black borrowers, and the unique contribution of anti-Blackness in student loan policy (Mustaffa & Dawson, 2022), this paper has a special focus on Black borrowers while still recognizing that other minoritized groups experience systemic racism as well.

The disproportionate impact of student loans on Black students has significant consequences and is fundamentally a racial justice issue. Black borrowers are more likely to take on debt in part due to the racial wealth gap, and this student debt in turn contributes to the maintenance of that disparity (Addo et al. 2016; Houle & Addo, 2019). Student loan debt's role in perpetuating the racial wealth gap continues to grow across students' lifetimes, impacting labor market outcomes and credit opportunities (Houle & Addo, 2019). Martin and Dwyer's (2021) study of student loan debt and financial stress during the Great Recession suggests that

high levels of student debt create high economic vulnerability for Black households, while finding no relationship for White households with similar debt (after controlling for socioeconomic resources).

These challenges are not due to some innate issue within Black people; instead, student loan burden disproportionately falls on these students due to centuries of structural forces and government policies (e.g., Addo et al., 2016; Mustaffa & Dawson, 2022; Seamster & Charron-Chénier, 2017). For example, structural racism has denied Black families the ability to build wealth to pay for college (e.g., Seamster & Charron-Chénier, 2017), while also shaping residential and K-12 school experiences (e.g., Persico, Figlio, & Roth, 2020; Rothstein, 2017). As a result, Black students are more likely to enroll in either systematically underfunded (e.g., Harris, 2021) or outright predatory colleges (e.g., Cottom, 2017) that can saddle them with high debt, little chance of earning a degree, or both. Those same students face a discriminatory labor market that pays them less than their peers for the same amount of education and must navigate other structures that impede wealth accrual (Ards & Myers Jr., 2001).

Although student loans provide many Black students with funds needed to access higher education, researchers identify the current student loan system as an example of *predatory inclusion*, meaning the system profits off of students under the guise of opening access (Cottom, 2020; Seamster & Charron-Chénier, 2017). Mustaffa and Dawson (2022) additionally argue that the racial logics in U.S. government and education systems coalesce into a *racial debt trap* which harnesses predatory inclusion and Black educational aspiration for profit.

The disproportionate impact of student loans on students of color, particularly Black borrowers, is rooted in complex, overlapping systems of oppression. Although these issues are systemic, the dominant discourse tends to frame student loans as an issue of individual

responsibility (e.g., Taliaferro et al., 2021). As discussed in the following section, these particular complexities often mediate the coverage and framing of policy issues.

Media and Education Policy

Research about the relationship between media and education policy is somewhat limited, but literature from the field of political communication demonstrates the importance of media in policy formation at large. As mentioned in the introduction, the media is considered a distinct policy actor rather than merely a conduit of information (Curran et al., 2022). Media actively alters the policy context in the creation, formation, and evaluation of policymaking and performs a variety of functions at each stage (Liu et al., 2019). In the agenda-setting phase, media directs attention to specific policy issues at the expense of others (Soroka et al., 2013), though this direction is often characterized by short bursts of attention rather than consistent strategy (Boydstun, 2013). In the policymaking phase, media sets the parameters of discourse on a given policy. Liu et al. (2019, p. 191) posit that media provides a platform for public discussion "aimed at optimizing policy plans," though more critical perspectives note that this public discussion is characterized by gatekeeping and journalistic norms which favor elites (Curran et al., 2022; Grossman, 2022; Soroka et al., 2013). During the implementation and evaluation phase of the policy process, the media provides feedback messaging and serves to legitimate the policy in question (Liu et al., 2019). Throughout policy formation and implementation, the media is not merely a mode of information transmission. It actively shapes policy possibilities and outcomes.

Mass media primes both the public and policymakers for a constrained set of policy solutions by socially constructing policy problems. How problems are defined and framed determines whether policy perspectives are accepted into normative political discourses and adopted on a mass scale (Woodly, 2015). Grossman (2022) argues that the government uses its

power and privileged access to mass media to concentrate conflict on advantageous issues, while marginal players seek to expand the conflict by attracting attention and weakening the government's narrative. The expansion of conflict might include the addition of new actors or new perspectives, such as the racialized dimensions of student loans. When media chooses to expand conflict, it "[redefines] how decision makers understand policy by reweighing the importance of some issues and attributes over others" (Wolfe, 2012, p.110). This redefinition can both create new directions and slow policy action, establishing an equilibrium around the status quo (Wolfe, 2012). Therefore, scholars recognize that setting policy definitions is an act of power that elicits policy conflict and power struggles.

As an example of the news media's ability to influence perceptions of policy problems and solutions, we highlight a recent study focused on immigration. Djourelova (2023) used the Associated Press's (AP) decision to ban the phrase "illegal immigrant" in 2013 to examine how the media's language decisions affected the broader public. The author exploited variation in local news outlets' reliance on text from the AP in order to causally examine the effect of a reduction in the phrase "illegal immigrant." The study found that residents of counties with daily newspapers more reliant on the AP, and therefore more exposed to the ban, had a nearly one percentage point lower support for policies to restrict immigration. This research provides some of the clearest evidence that language decisions in the media have a direct effect on the public's social constructions of policy problems and solutions.

Additionally, journalistic norms around policy coverage often drive preference for short-term, politically expedient solutions that neglect the underlying complexities of policy.[1] Soroka and colleagues (2013) outline how journalistic norms often prime audiences for specific policy

---

[1] Journalism is a collective enterprise, and published work reflects the input of editors, publishers, and other actors, in addition to journalists themselves.

solutions in problematic ways. Journalists tend to present "both sides" of an issue regardless of supporting evidence, utilize conflict frames, neglect complexity in favor of compelling individualistic storylines, and report on *political* issues rather than *policy* issues. These norms orient audiences towards policy solutions that are based on misleading, decontextualized policy problems. They may also focus the audience on policy options deemed "viable" by politicians rather than ones which address complex root causes. These norms are additionally reinforced by underlying logics and power structures within media institutions (Curran et al., 2022). An example of how these norms may impact education policy is Ertas and McKnight's (2020) content analysis of editorials and op-eds on charter schools. Their study found that media narratives about charter schools created novel definitions of educational policy problems and were superficial, lacking in depth, and missing important policy context.

Coverage norms, in addition to journalistic norms, can impact policy possibilities by creating a scattershot picture of policy problems. Boydstun (2013) found that newspaper coverage of policy issues is marked by short bursts of attention rather than being guided by policy significance. As a result, policy issues move cyclically in and out of public consciousness, transitioning through periods of high and low salience regardless of policies' consistent impact. Soroka and colleagues (2013) argue that journalists struggle to write about long-term, complex policy issues without a triggering event to make the impact of these issues salient, resulting in inconsistent coverage. Additionally, research suggests that the media is more likely to engage with policy issues where they are low in complexity and high in drama (Soroka et al., 2013). This literature suggests that long-term, complex, and deeply ingrained policy problems, like systemic racism in student loans, are neglected by media until moments of crisis or other high-profile developments.

Scholars have consistently found an important link between the media and the crafting of education policy. For example, in 2004, the U.S. Department of Education spent approximately $700,000 to assess how the media was representing then-President Bush and his signature education policy, No Child Left Behind, in order to craft their communication strategy to increase support for the policy (Goldstein, 2011). Prior research on education policy and the media typically focuses on the discourse or framing primarily found within news articles (e.g., Goldstein, 2011). Scholars have less frequently examined the racialized language choices in articles, even though political scientists have found that racialized language can play an important role in the policy process (e.g., Gillion, 2016). Scholars have also found that there are real power differentials across and within media outlets (e.g., Gamson & Wolfsfeld, 1993). Different outlets have varying levels of influence on both the public and policy actors, while certain individuals, such as editors, have more power than others to determine how information will be presented in the popular press (Grossman, 2022). Therefore, research focused on the media benefits from contextualized analysis within the publication outlet.

To our knowledge, no published research has examined news media coverage of student loans. Since media shapes both the public's and policy actors' social construction of policy issues, it is vital to examine how race is discussed in the context of student loan coverage. Understanding the race-conscious or race-evasive framing of student loans in the media, along with changes over time and shifts based on the outlet, can help inform racially just policy communication and policymaking.

Theoretical Framework

We use color-evasive racism as the theoretical framework.[2] Color-evasive racism details how, beginning in the 1960s, efforts to maintain a societal privileging of White people shifted from overt actions to more covert behaviors (Bonilla-Silva, 2014). Bonilla-Silva (2014) notes that racial ideology—how the author refers to racism— "is produced and reproduced in communicative interaction" (p. 11). Therefore, one of the clearest ways to examine racism is through communication, like the news articles that are the focus of this study. While racism can be expressed in interpersonal interactions between individuals, this theory is focused on exploring how larger societal structures covertly reinforce racial hierarchy. Because of this emphasis, Bonilla-Silva's color-evasive racism is not interested in individuals' intentions; instead, it is invested in uncovering how behaviors align with larger patterns across society.

In accordance with the structural emphasis of the theory, we are less interested in individual writers' aims and instead wish to focus on the larger patterns across articles and outlets that can unveil larger truths about society's beliefs on student loans, race, and racism. Newsrooms are primarily comprised of White men (particularly in leadership roles), have cultures often described as outright racially hostile, and frequently push for "objectivity" as an ideal for articles as a thinly veiled method of silencing writers of color (e.g., Gray, 2020). Additionally, media's proximity to economic and political powers, combined with internal factors like ownership and journalistic norms, result in the maintenance of current power structures (Grossman, 2022; Soroka et al., 2013). Therefore, regardless of the intentions of individual writers, it is likely that news articles produced in these environments exhibit signs of color-evasive racism.

---

[2] We use the term color-evasive racism to cover Bonilla-Silva's (2014) theory while removing ableist references to sight and emphasizing color-evasive racism as an active, not passive, phenomenon (Annama et al., 2017).

Scholars have used race-evasiveness or color-evasiveness as an interpretive lens in studies of state higher education equity policy documents (Felix & Trinidad, 2020), federal and institutional policies for Hispanic Serving Institutions (Aguilar-Smith, 2021; Vargas & Villa-Palomino, 2019), and anti-affirmative action legal discourse (Ward, 2020). However, as Taliaferro, Taylor, and Wheatle (2021) note in their essay collection, it has taken decades for the public and policy actors to acknowledge the role racism plays in student loan reliance and repayment. Prior to that shift, the public and policy actors, aided by narratives crafted in the media, minimized or completely ignored structural concerns. Therefore, we propose a novel use of the framework for the study of student loans in media discourse.

Applied to media coverage of student loans, this framework can help explain how color-evasive rhetoric perpetuates the maintenance of racism. According to race-evasive racism, refusal to engage in race-conscious rhetoric is not merely a passive act but also a form of power (Annama et al., 2017). The U.S. Department of Education's Digest of Education Statistics shows significant racial disparities in loan reliance and borrowing since at least 2006, when Black students borrowed at a rate nearly 8 percentage points higher than the next racial group.[3] As a result, scarce attention to race in the context of student loan news articles would suggest race-evasive coverage.

Guided by the theoretical framework, we had two hypotheses for the study. First, due to the overwhelming pressures to keep racism covert and pretend that it does not exist, articles written toward the beginning of the analytical time period will rarely use any racialized discourse. Second, in light of the need to minimize racism, when racialized discourse does

---

[3] Authors' calculations of data from the U.S. Department of Educations' Digest of Education Statistics for 2006, Table 323 (percentage of undergraduates receiving aid, by type and source of aid and selected student characteristics).

appear, news articles will use language that focuses on generalized group differences or choices (e.g., "Black students choose to take out larger student loans") instead of structural oppression for racially minoritized individuals (e.g., "Due to structural racism in wealth accrual, Black students rely on student loans at a higher rate").

Using a large corpus of news articles, we build upon prior research on education policy and the media to explore the discourse and framing that news outlets use in articles focused on student loans. While scholars have investigated racialized discourse focused on other topics, such as rental advertisements for housing (Kennedy et al., 2021), few research studies focused on media and education policy explore racialized discourse. Since the media plays a critical role in crafting the public and policy actors' perception of a policy issue, this gap represents a barrier to fully understanding the policymaking process for student loans. The current article helps to fill this significant gap in the literature, while examining trends over time and across news outlets.

Data and Methods

Data Source

We included news articles from 2006 to 2021.[4] This analytical time period includes news articles from one year before the official start of the Great Recession, when there was a resurgence in student loan reliance, through the end of 2021, allowing us to include any potential shifts in articles during the media's self-professed improvements in the coverage of race and racism after the murder of George Floyd. As noted above, prior research emphasizes that not all media outlets have the same amount of power in communicating with the public and that actors within media organizations play a critical role in the production of published news articles (e.g.,

---

[4] Before we began data collection of the news articles, and throughout the analysis and writing phases of this paper, the lead author communicated with several professionals within the news media to ensure that the research team was thinking appropriately about how articles go from ideas, through the drafting stage, and ultimately arrive as the finalized, published article.

Gamson & Wolfsfeld, 1993; Grossman, 2022). Therefore, we used a purposive sample of eight media outlets to allow for the context of each media organization to influence data analysis and the interpretation of the findings. We included the top five national newspaper publications by circulation: *Wall Street Journal, New York Times, USA Today, Washington Post,* and *Los Angeles Times* (Turvill, 2021). Due to our focus on racialized framing and the reality that student loan debt impacts Black individuals in a unique manner, we also included the major newspaper from the five metropolitan areas with the largest Black populations. This second criterion resulted in the addition of three newspapers—the *Atlanta Journal-Constitution, Chicago Tribune,* and *Philadelphia Inquirer*—since the other metropolitan areas in the top five are already represented in the newspapers we selected based on the circulation criterion (Tamir et al., 2021).

We created a corpus of articles focused on student loans by conducting a search for news articles at each of the eight outlets using ProQuest for the analytical time period that include the terms: a) college*, universit*, or "higher education", and b) loan*, debt*, or borrow*.[5] Whenever ProQuest included multiple versions of an outlet, we used the East Coast or Eastern print edition.[6] Our search resulted in a sample of 87,983 articles with full text available. For each article, we collected metadata including the author, title, publication, language, and date of publication. We prepared the full text for the main analysis by removing all punctuation

---

[5] An * represents any combination of additional characters. Therefore, "college*" in a search would return results that included "college" as well as "colleges."

[6] The publication ids in ProQuest for each newspaper were: Atlanta Journal-Constitution-26032, Chicago Tribune-46852, Los Angeles Times-46999, New York Times, Late Edition (East Coast)-11561, Philadelphia Inquirer-45611, USA Today-15008, Wall Street Journal, Eastern edition-10482, and Washington Post-10327. Several of these outlets added online versions of their newspapers that include additional articles during the analytical time period (e.g., ProQuest adds an online version of the *Washington Post* in 2016). In order to streamline the sources of articles, we solely rely on articles that appeared in the newspaper edition available throughout the entire analytical time period. Since these articles were also available in the online editions of some of the newspapers, this means that we did not include online-only articles.

(replacing hyphens with spaces and all other punctuation with nothing) and extra spaces.[7] We then removed all exact duplicates (615) and near-duplicates (318) of the full text.[8]

To ensure that the articles were actually about student loans, we conducted three rounds of iterative review of the articles to create additional inclusion parameters. Following Stryker and colleagues (2006), we first created a list of parameters for student loan relevance based on prior research. Then, we randomly sampled 240 articles, 30 per newspaper. Two authors independently reviewed the articles for whether they were relevant based on the *a priori* parameters created for software analysis and then whether the individual person found them to be relevant to student loans according to their personal judgement. After the authors completed the review, the entire research team met to come to consensus on whether articles were relevant and then to compare those decisions to the *a priori* parameters. The research team then came to consensus on how the computational parameters needed to be adjusted based on the human assessment of student loan relevance. We repeated this process three times, stopping once we felt that the parameters, which could be assessed using statistical analysis software, were adequately approximating human judgement. As we conducted these rounds of review, we realized that loan relevance had two layers: a broad layer focused on articles that mentioned student loans in some fashion and a narrow layer where one of the primary foci of the article was student loans. We therefore began coding for both broad relevance and narrow relevance in the articles until we felt confident that the decision rules for the statistical software were appropriate.[9]

---

[7] For the part-of-speech tagging, we used a version of the full text that retained all punctuation.

[8] See Appendix A for a description of how we identified near-duplicates of full text.

[9] We defined broad relevance as news articles that included certain "debt" terms (loan, debt, and borrow) multiple times along with clear indications that this debt was student loan debt. We defined narrow relevance as a substantial portion of an article using "debt" terms. See Appendix B for full information on the criteria for broad and narrow relevance and the decision rules coded in the statistical software.

We used measures of recall and precision (Stryker et al., 2006) to verify the accuracy and applicability of our relevance decision rules. Recall estimates the probability that a relevant text will be retrieved, while precision estimates the probability that a retrieved article is in fact relevant. These measures tend to be inversely proportional, though skillful relevance decision rules may achieve high scores in both (Stryker et al. 2006). Our final decision rules had a 52% recall and 93% precision for broadly relevant articles and a 95% recall and 77% precision measure for narrowly relevant articles, indicating the soundness of our recall decision criteria.[10]

We retained all articles that met our broad relevance criteria as part of the analytical sample and created an indicator for whether an article met our narrow relevance criteria (all narrowly relevant articles are also, by construction, broadly relevant). Once we completed this process, we removed 78,213 articles, leaving 8,843 broadly relevant articles. We again evaluated the set of articles for duplicates and found an additional 6 articles to remove. Therefore, our final analytical data set includes 8,837 articles. Table 1 provides purposive selection criteria for each outlet and shows how the analytical data set breaks down across newspapers. The table additionally highlights that we removed approximately 90% of the articles from each newspaper due to the broad relevance restrictions (except for *USA Today*, where we removed 79%).

Beyond the articles themselves, we collected information on each outlet, such as demographic characteristics of the newsrooms and authors' colleges, and how they have changed over time. Given the importance of institutional actors in media policy coverage (Grossman,

---

[10] We prioritized recall for narrowly relevant articles and precision for broadly relevant articles. Said another way, we wanted to ensure that our decision rules did the best possible job at retaining articles that actually focused on student loans (narrowly relevant) while at the same time excluding articles without a single genuine mention of student loans (not broadly relevant). Therefore, the 93% precision for broad relevance and 95% recall for narrow relevance are ideal.

2022), we collected information on the percentage of people of color in the newsroom.[11] We generally used annual reports from the American Society of News Editors' Newsroom Employment Diversity Survey through 2019 for this data. After 2019, this survey did not receive enough responses to be published. For this reason, we used individual newspapers' calculations of the racial diversity in the newsroom instead for 2020 and 2021.[12] In the years when newspapers had missing data, we used mean imputation unless it was in the final year of the panel (when this occurred, we carried forward the last year's percentages).

We incorporated characteristics of authors' colleges in order to examine the context in which authors likely gained formative knowledge of student loans generally and their racialization in particular. We examined all authors in our dataset and focused on anyone who was listed as an author at least three times since we could argue those authors were ones more likely to be reporters who frequently wrote about student loans (instead of idiosyncratic opinion piece writers). This gave us a sample of 595 authors (from the 2,704 in the entire dataset).[13] We were able to identify 563 of those authors' undergraduate institutions, with 552 having attended a US institution.[14] For articles written by those 552 authors, we created measures of whether all "loan authors" attended an Ivy Plus institution,[15] whether any loan author attended an Ivy Plus or

---

[11] While we would have preferred to use more granular measures of racial diversity in the newsroom, we could not consistently find that data for a sufficient portion of the panel. For this reason, we focus on all employees in the newsroom not racialized as White.

[12] Newspapers reported these numbers; the research team did not assess the racial group for members of newsrooms.

[13] Authors we deemed to be "loan authors" wrote 5,661 articles or 64% of the total analytical sample.

[14] We included any institution that a loan author attended, not graduated from, in this analysis. That means we could have multiple undergraduate institutions for a single author. When an author had multiple institutions, we created measures based on whether an author "ever" experienced the categorical measures (e.g., Ivy Plus institution) or a more inclusive institution for continuous measures (e.g., higher decile of Black undergraduate students, higher decile of loan reliance). We attempted to only include degree-seeking undergraduate attendance (e.g., excluding an institution when an author appeared to only be attending for workshops or study abroad).

[15] Ivy Plus institutions included all institutions in the Ivy League athletic association (Brown University, Columbia University, Cornell University, Dartmouth College, Harvard University, University of Pennsylvania, Princeton University, and Yale University). We supplemented this list with other peer institutions in admissions and resources. We relied on prior literature (e.g., Chetty et al., 2017) to add the following peer institutions: Duke University, Massachusetts Institute of Technology, Stanford University, and the University of Chicago. We also included

public flagship institution (excluding author teams that all attended Ivy Plus institutions, as captured in the previous measure), the decile of the percentage of students at the institution who borrow, and the decile of the percentage of undergraduates who were Black.[16] We created these measures due to the distribution of where authors enrolled as undergraduates (e.g., only 2 of the 552 authors attended a community college while 179 attended an Ivy Plus institution). Given the significant percentage of the sample that attended an Ivy Plus institution (approximately 32%), we created two separate Ivy Plus measures. More specifically, if an article was written by two loan authors, and one attended Northwestern University and the other attended the University of Virginia, this article would have a 0 for the all Ivy Plus variable (since the entire author team did not attend an Ivy Plus institution) and a 1 for the Ivy Plus and public flagship variable. Once we had those estimates, we aggregated them up to the newspaper-year level.

Combined, we created analytical covariates for each newspaper that included annual measures of: the percentage of articles with all Ivy Plus loan authors, the percentage of articles with Ivy Plus or public flagship loan authors, the average decile of loan reliance, the average decile of Black student enrollment, and the percentage of people of color in the newsroom.

Racialized Language Terms

We selected the racialized language for this study based on the theoretical framework of color evasiveness, prior linguistics research (e.g., Alim et al., 2016; Baker et al., 2022), and our

---

Northwestern University given its prominence in educating graduates who go on to work in elite media outlets. (Northwestern University and Columbia University are generally accepted as the most prestigious journalism programs in the United States [Benton, 2021].)

[16] For the loan authors we could find college data for, authors earned degrees from 1959 to 2020. Given that IPEDS data did not exist in 1959, we decided to use decile ranks of the time-variant institutional characteristics. We compared the decile ranks for loan reliance, Black student share of undergraduate enrollment, and total undergraduate enrollment for 1998 (the first year loan reliance data is available), 2008, and 2018. We found strong correlations across the measures (smallest correlations for each: loan reliance-0.77, total enrollment-0.93, share Black enrollment-0.82), so we use deciles of 2018 IPEDS data throughout the paper.

review of the AP Style Guide over the 15 year analytic period.[17] We then immersed ourselves in

the text data and read a random sample of 400 articles (stratified by newspaper) to identify any

other words or phrases we should include as part of the racialized language terms.

Once we completed this review, we created a dictionary holding our final list of

racialized language terms. As the theory suggests that individuals will be hesitant to directly

mention race, we focus on ethnoracial categories that could appear in news media during the

analytical time period, including "White", "Caucasian", "Black", "African American", "Asian",

"Native Hawaiian", "Pacific Islander", "Native American", "Alaskan Native," "American

Indian", "Indigenous", "Hispanic", "Latino", "Latina", "Latinx," "Chicano," "Chicana," "South

Asian," "east Asian," "Middle Eastern North African," "biracial," and "multiracial." We also

include collective phrases that indicate race (e.g., "minority" and "of color") and racialized

institutions (e.g., "Minority Serving Institutions," "Historically Black Colleges and Universities"

[HBCU]). Finally, the theoretical framework necessitates that we investigate words or phrases

that point to larger, systemic issues, including "racial inequity," "racial inequality," and

"racism." This final set of words allowed us to examine when articles use racialized discourse

that explicitly references structural issues (counteracting pressures from color-evasive racism).[18]

Analytic Strategy

We used Natural Language Processing (NLP) methods to analyze racialized discourse in

student loan news articles. These methods were attractive because they allowed us to analyze a

---

[17] The AP style guide provides style guidelines for journalists, including acceptable racialized language, much the same way the APA style guide outlines the rules for published social science research.

[18] See Appendices C and D for a full list of racialized language terms used in the study and the decision criteria used for Black, White, and minority to ensure that we captured uses of these terms that were racialized. As noted in Appendix C, we also search for each HBCU by individual institution name and include those mentions as part of the overarching HBCU measure. We do this for two reasons. One, as we have noted throughout this paper, we have a particular focus on Black borrowers' experiences with student loans. Two, we noticed when reviewing the 400 random articles for racialized language that news articles did not always note that individual institutions were HBCUs, yet it seemed likely that readers would be processing these institutions as racialized.

substantial number of documents (significantly more than we could do by hand). Still, we acknowledge that these analytical tools should be used carefully and as a complement to traditional methods of analysis given that NLP and machine learning strategies are not free of bias or error (e.g., Noble, 2018).

Analysis for RQ1 focused on the actual racialized language terms themselves, while RQ2 examined the words that modify those terms. When sample size allowed, we conducted analyses described below within each of the media outlets, which allowed for a contextual exploration within and across news outlets.

RQ1 tracked the use of racialized language terms over time in the corpus of news articles using term frequency. We restricted this first analysis only to the racialized language terms to explore writers' word choice. We created binary variables for each racialized term indicating whether an article mentions this term. We then converted these indicators into the percentage of articles published annually that mention each term within each outlet (count of articles mentioning a term divided by total number of articles), allowing us to control for the number of articles published each year. These simple percentages allowed us to descriptively explore how the frequency of different terms shifted over time and across outlets.

We used the measures we collected about each outlet to predict the frequency of racialized term use over time. We included the covariates and year fixed effects, allowing the analysis to explore the correlation between time-varying outlet characteristics and term frequency while controlling for time-invariant characteristics. We clustered standard errors at the newspaper level. The findings from RQ1, a descriptive exploration and regression analysis, allowed us to examine the extent to which the news media has changed its use of words

signifying race, racialized institutions, or racism with relation to student loans over time and to test the first and second hypothesis.

RQ2 examined whether there were systematic differences in how racialized language terms were used by examining the frequency of modifying words. We identified 3,959 sentences that contained one of the top five most frequently used racialized language terms. Next, we applied part-of-speech tags to of the words in each sentence, thereby categorizing the grammatical use of each modifying word.[19] This portion of the analysis is a computational corollary to diagramming the parts of speech in sentences (in fact, the algorithms that are used to create part-of-speech tags have been trained on a significant corpus of human-annotated documents). From these tags, we counted the frequency of each adjective and noun used for the most frequent racialized language terms overall, over time, and within each outlet. Grouping these modifying words by racialized language term allowed us to compare how articles discussed each term. The results from RQ2 provide evidence on the ways different terms are used when racialized language is present in student loan news articles and to test the second hypothesis.

Results

We provide a descriptive overview of racialized language trends in the student loan article data. Then we explore regression estimates of the relationship between racialized-language use and newspaper outlet characteristics. We conclude this section with the most frequent modifying words for the most frequent racialized language terms.

Descriptive Overview

Overall. Figure 1 shows the number of student loan articles that we found for each year of the panel, for both broad- and narrow-relevance articles (3,809 articles are narrowly relevant).

---

[19] We used *udpipe* in R to conduct part-of-speech tagging. We applied the "english-ewt" dictionary given the extensive size and variety of its corpus.

There are spikes in 2007, 2012, and 2019, though no year goes above 800 articles across the eight outlets. The newspapers included in our study have published fewer student loan articles in recent years.

Turning to Figure 2, we show the annual percentage of articles that mention any racialized language term. Here, unlike the number of articles, we see a sharp increase in the share of student loan articles that use racialized language, whether the article is broadly or narrowly relevant. This sharp increase was especially prominent after 2018. Prior to 2015, articles with racialized language consistently compromised less than 16% of all broadly relevant articles (narrowly relevant articles that used racialized language generally did not go above 9%, except for 2010). Yet, in 2019, student loan articles using racialized language made up approximately 30% of broadly relevant articles and 26% of narrowly relevant articles. By the end of our panel, nearly 40% of broadly relevant articles used some form of racialized language, as did nearly 27% of all narrowly relevant articles. This finding provides evidence supporting our first hypothesis, that early in our panel student loan news articles rarely used racialized language.

While there is a clear increase in the use of racialized language in student loan news articles, that does not necessarily mean that articles are discussing structural issues like racism. Indeed, when we explore the top five racialized language terms in narrowly relevant articles, we found that they were, in order: Black, HBCU, minority, White, and racial.[20] Figure 3 shows the trends in the use of these terms in narrowly relevant articles. Even though there are unique trends for each term (e.g., "Black" is decreasing from 2020 to 2021 while during the same time period "racial" is increasing), in general the percentage of articles using these terms are small until

---

[20] The top five terms for broadly relevant articles were: Black, minority, HBCU, White, and African American. Given the overlap, and that we are particularly interested in articles focused on student loans, we chose to focus on the most frequent terms in narrowly relevant articles.

around 2019. In Figure 2, interestingly, there is a spike in the percentage of narrowly relevant articles using any racialized language in 2010. We have explored the titles for these articles and find that this spike was being driven by policy discussion around for-profit institutions and the switch from the Federal Family Education Loan Program to the Direct Loan program.[21] Figure 3 shows that this 2010 spike is primarily being driven by uses of "Black" and "minority," which makes sense given a number of references to the Congressional Black Caucus in discussion of the Direct Loan program switch.

Still, even with that spike in some racialized language term use in 2010, we found suggestive evidence for our second hypothesis. In fact, while approximately 18% of all articles in our data use at least one racialized language term, nearly 8% use "Black" while less than 1% use "racism." To deepen our analysis, we explicitly explored the frequency of the structural racialized language terms (i.e., racism, racial equity, racial equality). Table 2 shows the percentage of articles that use the terms for all years and for 2006 to 2019 (removing the years when newspapers might be reacting to more recent murders of Black people). We include the percentages for "Black" as a useful comparison for the magnitude of the structural terms' percentages. We find consistent evidence that structural racialized language is used less frequently overall and especially before 2020. Therefore, while there has been an increase in the use of racialized language, this has primarily been driven by descriptive language about groups of people or institutions, instead of a focus on structural issues like racism.

Within newspapers. Given the ways that newspaper outlets differ in their internal structure and power dynamics, we also explore racialized language term use across newspapers.

---

[21] From the 29 narrowly relevant articles in 2010, 14 article titles were focused on for-profit institutions (a mix of articles on for-profits as institutions and others on gainful employment, a federal accountability policy that solely targets for-profit institutions). An additional 11 article titles discussed the switch to the Direct Loan program.

Figure 4 shows two heatplots of the annual percentage of articles with any racialized language term (panel A includes the broadly relevant articles and panel B includes the narrowly relevant). Focusing on panel A, the year is on the x-axis and newspapers are on the y-axis. Each square represents the annual percentage of broadly relevant articles that use any racialized language term. The legend on the right shows that darker blues and purples represent smaller annual percentages while lighter greens and yellows represent larger annual percentages. It is clear that different newspapers have different trends in racialized language use, even if generally the annual percentages increase over time. Of particular note, the *Atlanta Journal-Constitution* has the highest annual percentage for racialized language use in the entire panel (which occurs in 2021).[22] The *Atlanta Journal-Constitution* rose from using any racialized language in 29% of broadly relevant student loan articles in 2006 to 74% in 2021. The changes across our panel for each of the remaining newspapers are 54 percentage points (pp) for the *Philadelphia Inquirer* (10% to 64%), 35 pp for the *New York Times* (10% to 45%), 33 pp for *USA Today* (20% to 53%), 23 pp for the *Washington Post* (18% to 41%), 26 pp for the *Los Angeles Times* (24% to 50%), 17 pp for the *Chicago Tribune* (10% to 27%), and 14 pp for the *Wall Street Journal* (6% to 20%). Descriptively, some newspapers start shifting their racialized language use around 2018-2019 while others, particularly the *Chicago Tribune* and *Wall Street Journal*, show minimal change (these two newspapers also happen to have the lowest percentages in the final year of the panel).

When we investigate Figure 4, panel B, we see a similar trend for narrowly relevant student loan articles, though a smaller percentage of articles used racialized language at the

---

[22] Given that the Atlanta area has a significant number of HBCUs, we also explored these trends not including the individual names of each HBCU and the *Atlanta Journal-Constitution* still published articles with the highest percentage.

beginning of the panel. (Some newspapers had no articles that used racialized language.) The *Atlanta Journal-Constitution* continues to have one of the highest percentages in the final year of the panel along with the *Philadelphia Inquirer* (67% of narrowly relevant articles for each). We still see that some newspapers began to shift their racialized language use in 2018 and 2019, while the *Chicago Tribune* and the *Wall Street Journal* continued to be outliers with the small percentage of articles using racialized language in 2021 (12% and 14%, respectively).

Given the focus of this paper and the prevalence of articles using "Black" from Figure 3, we also explore newspapers' use of this term over time. Figure 5 shows heatplots similar to Figure 4 that report the annual percentages for the use of "Black" for broadly and narrowly relevant articles (panel A and B, respectively). Figure 5, panel A shows a much darker heatplot overall, indicating a smaller percentage of broadly relevant articles use "Black" than any racialized language, which is logical. We still find that some newspapers began shifting their use of the term "Black" around 2018 and 2019, the *Atlanta Journal-Constitution* has the highest percentage in 2021 (52%), and the *Chicago Tribune* and the *Wall Street Journal* have the lowest percentages (nearly 12% for both newspapers). These trends are also repeated for narrowly relevant articles, though the *Philadelphia Inquirer* has the largest percentage at nearly 67%. Therefore, we consistently found evidence that, while racialized language use increased overall, these trends were highly uneven across the newspapers. The *Atlanta Journal-Constitution* and the *Philadelphia Inquirer*, both included due to their surrounding areas' number of Black residents, are the newspapers that generally had the largest growth over time (measured in pp) as well as the highest percentages in 2021. The *Chicago Tribune* (chosen due to the number of Black residents) and the *Wall Street Journal* (chosen due to the number of Black residents and

circulation numbers) consistently had the smallest growth over time (measured in pp) and the smallest percentages in 2021.

Regression Estimates

Turning to the regression analysis, we explore how newspaper characteristics relate to the use of racialized language. Table 3 includes the regression results for the outcome percentage of articles using any racialized language (first four columns) and the outcome percentage of articles using "Black" (last four columns).[23] For each outcome, we present the estimates for the set of covariates for broadly relevant articles, those covariates plus year fixed effects, for narrowly relevant articles, and then those covariates plus year fixed effects (see Appendix E Table E1 for summary statistics on the covariates and outcomes). We found that a 1 pp increase in people of color in the newsroom correlates with a 1.367 pp increase in the percentage of broadly relevant student loan articles using any racialized language, when controlling for other newspaper characteristics. When we added year fixed effects, controlling for time invariant characteristics of the newspapers, we found a similar relationship between the share of people of color in the newsroom and use of racialized language in broadly relevant articles (though the point estimate is smaller). We also found that a one decile increase in the share of Black student undergraduate enrollment at authors' undergraduate institutions was associated with a 3.355 pp increase in the percentage of articles using any racialized language. We note that the difference between deciles can be relatively small given that the first decile cutpoint is 0.87% and the next is 2.30% (see Appendix E Table E2 for all decile cutpoints).

The next columns report the estimates for the narrowly relevant student loan articles. When solely including the newspaper characteristic covariates, we again found a positive

---

[23] We specifically examine the term "Black" given the focus on Black people's experiences with student loans, outlined at the beginning of this paper, and its position as the most popular racialized language term in the analysis.

relationship between the share of people of color in the newsroom and any use of racialized language. However, when we added year fixed effects, we found that only the estimate for decile of Black student enrollment maintains statistical significance. It may be that narrowly relevant articles had less variation that we could explore, as these articles had a significantly smaller sample of articles. (The annual average per newspaper was approximately 30 narrowly relevant articles, while the average for broadly relevant was 69.)

We found a qualitatively similar story when we switched the outcome to the percentage of articles that use "Black." The share of people of color in the newsroom continued to have a positive relationship with the use of "Black," though the point estimates are smaller. We also found no statistically significant relationships in the models including narrowly relevant news articles. Together, we took this to mean that the relationship between use of any racialized language and both authors' college Black student enrollment and the share of people of color in the newsroom was not solely about which newsroom characteristics predict articles mentioning the term "Black."

Modifying Words for Most Frequent Racialized Language Terms

Our final analysis explored the modifying words for the most frequently used racialized language terms. Table 4 shows the five most frequently used nouns and adjectives for each of the five most frequent racialized language terms (Black, minority, White, racial, HBCU). Within each modifying word category (i.e., noun, adjective), we list each modifying word in order of overall frequency in the data. We found that sentences including racialized uses of "Black" primarily used nouns describing statistics ("percent"), institutions ("college" and "university"), people ("student"), and student loans ("debt"). This trend is generally true for the nouns in

sentences with other racialized language terms, though there are a few differences. Sentences that included either "White" or "racial" more frequently included the noun "wealth."

The most frequent adjectives for these terms show the dominance that using "Black" plays in racialized discussions in student loan articles. For all of the racialized language terms except Black, the most frequent adjective was "Black." (We cannot say for certain that these uses were racialized uses of "Black" though it seems likely given that the sentences already contain other racialized language usage.) The adjective "federal" appeared frequently in sentences that included "Black," "HBCU," or "minority." This trend likely aligned with the significant focus on federal student loan debt in the media, though the trend was not replicated for sentences including "White" or "racial."

We also explored how the racialized language modifying words changed over time and across newspapers.[24] We found expected trends for macropolitical events (e.g., "Obama" appeared as the second most frequent noun for sentences using "Black" and most frequent for sentences using "White" in 2008, when former President Barack Obama ran his first presidential campaign). Over time, we found that adjectives like "low," "diverse," and "urban" were used in sentences with "minority" at a rate not seen with the other more frequent racialized language terms (though "low" is one of the top adjectives in sentences with "Black" for one year). We also found that sentences that include "Black" frequently include "White" and vice versa. We took this to suggest that racialized use of "Black" often occurs when comparing Black people to White people.

Similar to the over time analysis, when we explored modifying words across newspapers, we found that "low" was frequently used as an adjective in sentences that used "minority." One

---

[24] Given the length of this paper, tables available upon request.

of the top two adjectives for each newspaper's sentences using "Black" was "White" and vice versa, reinforcing that these terms are likely being used to compare the two groups of people. The *Los Angeles Times* added "Latino" as one of the most frequent adjectives in sentences using "Black."

The modifying words findings provide clues as to how the articles were actually using the racialized language terms. Based on the overall trends, trends over time, and trends across newspapers, it appears that the most frequently used racialized language terms were used to compare different groups of people to each other. While useful to better understand how people experience the world differently, it is not clear that discussions of barriers created by white supremacy and structural racism were being incorporated. The modifying word findings, coupled with the distinct absence of structural racialized language terms we document above, provides suggestive evidence for our second hypothesis.

Discussion

We sought to identify whether and how the media discussed race in the context of student loan coverage. Using data from fifteen years of student loan coverage from eight newspapers, we found that, until recently, the majority of news media coverage of student loans was almost entirely race-evasive. We did find an increase in the use of racialized language over time, though this increase accelerated around 2018—not 2020, as might be posited given the news media's pledges to shift their coverage of race and racism in the wake of systematic state violence against Black people in the United States. When newspapers in our study began including racialized language, this language was primarily focused on talking about racialized groups of people or institutions, instead of focusing on larger systemic issues. Given these findings, and the

literature, we argue that race-evasive framing of student loans contributes to a public discourse that ignores systemic causes of unjust outcomes and precludes race-conscious policy solutions.

Individuals at media outlets make deliberate decisions in their language and framing to convey their arguments. Yet, the framing of a policy issue and discourse about it are not solely driven by the media (e.g., the media is influenced by the public and vice versa). Ultimately, this reality means that our analysis of racialized language in student loan news articles can reveal how decision-making, whether conscious or unconscious, influences the larger discourse around student loans. That discourse then relates to the ways policy actors and the public view the issue.

Our descriptive explorations revealed fairly consistent trends in the use of racialized language. We also explored how the characteristics of individual newspapers correlated with the percentage of student loan articles that used any racialized language. We found evidence that the racial composition of the newsroom was a significant predictor of racialized language in articles. It could be that this finding was being driven by newspapers being more likely to hire more people of color in recent years when the newspapers were also more likely to produce articles with racialized language. We explored the potential for this relationship in Figure 6, which includes a heatplot of the annual percentage of people of color in the newsroom. While some newspapers do show a steady increase in people of color (e.g., *Atlanta Journal-Constitution*, *Los Angeles Times*), there were several newspapers that had quite small changes over the panel. Therefore, there is not a clear, significant trend of increasing the people of color in the newsroom across all the papers in our study. This reality, coupled with the year fixed effects we include in some of our models, suggest that there is a relationship between newsrooms having additional people of color and the presence of racialized language in student loan news articles.

We also found evidence that the share of undergraduate students who were Black at loan authors' undergraduate institutions occasionally correlated with the use of racialized language in articles (especially for broadly relevant articles). When we explored the annual average decile of Black students at loan authors' colleges, we found no clear temporal trend (see Figure 7). For example, focusing on the *Atlanta Journal-Constitution*, the year where loan authors' colleges had the highest decile of Black students was 2014. That is the same year when *USA Today* had one of its lowest deciles. Given that the relationship between the share of Black undergraduate students attending authors' undergraduate institutions and racialized language in those same authors' articles appears primarily in broadly relevant articles, our findings suggest that this link may be more prevalent in articles that were not primarily focused on student loans (and therefore may be authored by people who less regularly write about the subject).[25]

We acknowledge that this study uses a purposive sample of newspapers. We do not argue that the findings from this study are applicable to all newspapers, or all news media more broadly. However, given the variation in prestige for newspaper outlets (which we have noted throughout this paper), we have assembled a sample of particularly powerful newspapers across the country. We wish that we could have examined all newspapers across the United States, but we would not have been able to collect the newspaper and author characteristics that help deepen our exploration. We encourage other scholars to expand our work by examining racialization and student loans in other regional newspapers, other forms of news media, and broader media discussions, including social media. We also encourage policy scholars to explore how the media discusses other areas of education policy, the "quality" of this discourse (given that our analysis techniques cannot assess if mentions of racialized language were harmful or affirming), and the

---

[25] We further explore the colleges attended by student loan authors in a separate paper.

ways these discussions lead to and stem from other policy actors' discourse (e.g., how shifts in the use of racialized language in news articles relate to shifts in language in the Congressional Record or policy think tank reports). Even if causal relationships cannot be estimated, there is too little empirical evidence on the role of the media in the formation, implementation, and evaluation of education policy.

Our findings support both of our hypotheses that we posited based on our theoretical framework from Bonilla-Silva's (2014) race-evasive racism. The literature on media and policy suggests that journalistic norms and the punctuated equilibrium of coverage in mass media (e.g., Boydstun, 2013) help explain the race-neutral framing of student loan policy. The racialized impacts of student loan debt are complex, long-term policy problems without simple, politically expedient solutions. Consistent with this literature, we found that the media largely avoids the racialized impacts of student loan policies. Additionally, events such as racialized discourse in the Trump era, the demonstrations following the murder of George Floyd and other Black people, and the financial instability caused by the pandemic may partially explain the increase of coverage connecting racialization and student loans in recent years.

Further, the power of journalistic norms and the role of information gatekeeping may help explain the relationship between racialized term use and newsroom racial diversity measures. It is not the responsibility of journalists of color to pressure newsrooms to incorporate more racialized language in news articles. It is also unlikely that the mere presence of journalists of color automatically increases the share of articles that include racialized language. Political communications literature suggests that the ability of journalists of color to cover racialized policy issues may be mitigated by powerful journalistic norms (Soroka et al., 2013) and the institutional gatekeeping methods from editors, publishers, and owners (Grossman, 2022).

Therefore, it is more likely that our findings point to the potential that newsrooms that hire more people of color are also more likely to have standards in place that produce more articles that use racialized language. This could be why we found consistent evidence that certain newspapers (e.g., *Atlanta Journal-Constitution* and *Philadelphia Inquirer*) increased their use of racialized language at a different rate than other newspapers (e.g., *Chicago Tribune* and *Wall Street Journal*). We emphasize that the newspapers that have consistent evidence of increasing their racialized language use also were included in this study solely due to their metro area's number of Black residents. It may be that more "regionally" focused newspapers representing a significant number of Black residents were more responsive to updating how they discussed student loans and racialization (though we did not see this trend for the *Chicago Tribune*, highlighting that all regional newspapers did not incorporate racialized language in the same manner).

The media's race-evasive approach to student loans has significant implications for student loan policy formation. Race-neutral framing of student loan policies primes the public and policymakers for race-neutral policy solutions. Avoiding the specific impacts of student loan debt on communities of color defines student loan policy in a particular way. It sets policymakers up for simple, politically expedient solutions, but it removes communities of color from the policy discourse.

The recent addition of race, and occasionally racism, to the news discussion of student loan policy expands the public's conception of the problem and introduces long-neglected considerations necessary for racially just policy formation. Given the body of literature, we might expect this broadened conception to also "hit the brakes" on student loan policy changes (Wolfe, 2012). This is not necessarily the case. Interestingly, the increase of racialized terms in

student loan media coverage has been quickly followed by the Biden administration's student loan cancellation effort in 2022 (White House, 2022). In fact, policymakers specifically focused attention on the impacts of student loans on communities of color when crafting loan cancellation and added a provision to the policy (an additional $10,000 of cancellation for Pell grant recipients) to ensure an increase in benefits for borrowers of color, and Black borrowers more specifically (Blom, 2021; White House, 2022).

We are not arguing that the direction of a potential causal relationship is that the news media includes more racialized language which directly leads to the federal government creating new policies. There is clear evidence that increased attention alone does not improve people's lives (e.g., Cox & Edwards, 2022). Instead, we are asserting that the news media's inclusion of racialization when writing about student loans helps to expand the public's and policymakers' understanding of the "problem" of student loans, which can lead to real changes in policy decisions when crafting solutions. For instance, while applications for the loan cancellation policy were paused due to lawsuits (the policy is currently under review at the Supreme Court of the United States), reporters found that ZIP codes with a majority of residents of color submitted applications at a higher rate per capita than ZIP codes with a White majority (Stratford et al., 2023).

Mass media is a powerful actor in policymaking and is a worthy subject of exploration for education policy research. Research and theory focused on media communication can help explain the connection between race-evasive policy coverage and race-evasive policy solutions in education. Consideration of how the media frames student loan policy is key to advancing racially just policy solutions in discourse and practice.

References

Addo, F. R., Houle, J. N., & Simon, D. (2016). Young, Black, and (still) in the red: Parental wealth, race, and student loan debt. *Race and Social Problems 8*(1), 64–76.

Aguilar-Smith, S. (2021). Seeking to serve or $erve? Hispanic-Serving Institutions' race-evasive pursuit of racialized funding. *AERA Open*, *7*. https://doi.org/10.1177/23328584211057097

Alim, H. S., Rickford, J. R., & Ball, A. (2016). *Raciolinguistics: How language shapes our ideas about race.* New York: Oxford University Press.

Annamma, S. A., Jackson, D. D., & Morrison, D. (2017). Conceptualizing color-evasiveness: Using dis/ability critical race theory to expand a color-blind racial ideology in education and society. *Race Ethnicity and Education*, *20*(2), 147-162.

Ards, S. D., & Myers Jr, S. L. (2001). The color of money: Bad credit, wealth, and race. *American Behavioral Scientist*, *45*(2), 223-239.

Baker, D. J. (2019). When average is not enough: A case study examining the variation in the influences on undergraduate debt burden. *AERA Open*, *5*(2).

Baker, D. J., Ford, K. S., Viano, S., & Johnston-Guerrero, M. P. (2022). Racial category usage in education research: Examining the publications from AERA journals (EdWorkingPaper: 22-596). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/r9dg-kd13

Baum, S., Ma, J., Pender, M. & Libassi, C. J. (2019). *Trends in Student Aid 2019.* New York: College Board.

de Benedictis-Kessner, J., & Hankinson, M. (2022). How the identity of substance users shapes public opinion on opioid policy. *Political Behavior*, 1-21.

Benton, J. (2021, Oct. 20). It's time to create an alternative path into a journalism career. *Nieman Lab*. https://www.niemanlab.org/2021/10/its-time-to-create-an-alternative-path-into-a-journalism-career/

Blom, E. (2021, Apr. 12). A more targeted approach to student loan forgiveness. *Urban Institute*. https://www.urban.org/urban-wire/more-targeted-approach-student-loan-forgiveness

Bonilla-Silva, E. (2014). *Racism without racists: Color-blind racism and the persistence of racial inequality in the United States (Fourth Edition)*. Maryland: Rowman & Littlefield Publishers.

Boydstun, A. E. (2013). *Making the news: Politics, the media, and agenda setting*. University of Chicago Press.

Chetty, R., Friedman, J. N., Saez, E., Turner, N., & Yagan, D. (2017). *Mobility report cards: The role of colleges in intergenerational mobility* (No. w23618). National Bureau of Economic Research.

Cottom, T. M. (2020). Where platform capitalism and racial capitalism meet: The sociology of race and racism in the digital society. *Sociology of Race and Ethnicity*, *6*(4), 441-449. https://doi.org/10.1177/2332649220949473

Cottom, T. M. (2017). *Lower ed: The troubling rise of for-profit colleges in the new economy*. New York: The New Press.

Cox, K., & Edwards, K. (2022, Aug 30). Black Americans are pessimistic about their position in U.S. society. *Pew Research Center*. https://www.pewresearch.org/race-ethnicity/2022/08/30/black-americans-are-pessimistic-about-their-position-in-u-s-society/

Curran, D., Gillanders, R., & Mahmalat, M. (2022). Policymaking, ideational power and the role of the media. *Political Studies Review*, *20*(2), 250–264. https://doi.org/10.1177/1478929920968348

van Dijk, T. A. (1989). Mediating racism: The role of the media in the reproduction of racism. *Language, power and ideology*, *7*, 199-226.

Djourelova, M. (2023). Persuasion through Slanted Language: Evidence from the Media Coverage of Immigration. *American Economic Review*, *113*(3), 800-835.

Ertas, N., & McKnight, A. N. (2020). A narrative policy framework analysis of charter school editorials in local media. *Critical Questions in Education*, *11*(1), 1-20.

Federal Reserve Bank of New York. (2021) *Quarterly report on household debt and credit*. Washington, DC: Federal Reserve Bank of New York.

Felix, E. R., & Trinidad, A. (2020). The decentralization of race: Tracing the dilution of racial equity in educational policy. *International Journal of Qualitative Studies in Education*, *33*(4), 465-490.

Gamson, W., & Wolfsfeld, G. (1993). Movements and media as interacting systems. *Annals of the American Academy of Political and Social Science*, 528, 114-125.

Gerstl-Pepin, C. (2007). Introduction to the special issue on the media, democracy, and the politics of education. Peabody Journal of Education, 82(1), 1-9.

Gillion, D. Q. (2016). *Governing with words: The political dialogue on race, public policy, and inequality in America*. Cambridge University Press.

Goldstein, R. A. (2011). Imaging the frame: Media representations of teachers, their unions, NCLB, and education reform. *Educational Policy, 25*(4), 543-576.

Gray, K. (2020, Sept 25). The racial divide on news coverage, and why representation matters. *Knight Foundation*. https://knightfoundation.org/articles/the-racial-divide-on-news-coverage-and-why-representation-matters/

Grossman, E. (2022). Media and policy making in the digital age. *Annual Review of Political Science*, *25*, 443–461. https://doi.org/10.1146/annurev-polisci-051120-103422

Harris, A. (2021). *The state must provide: Why America's colleges have always been unequal – and how to set them right*. New York: HarperCollins.

Houle, J. N., & Addo, F. R. (2019). Racial disparities in student debt and the reproduction of the fragile black middle class. *Sociology of Race and Ethnicity*, *5*(4), 562-577. https://doi.org/10.1177/2332649218790989

Katznelson, I. (2005). *When affirmative action was white: An untold history of racial inequality in twentieth-century America*. WW Norton & Company.

Kennedy, I., Hess, C., Paullada, A., & Chasins, S. (2021). Racialized discourse in Seattle rental ad texts. *Social Forces, 99*(4), 1432-1456.

Liu, S., Cai, L., & Zhao, X. (2019). The role of mass media in education policies: A Chinese case study. *Journal of Higher Education Policy and Management*, *41*(2), 186–203. https://doi.org/10.1080/1360080X.2018.1554548

Martin, E. C., & Dwyer, R. E. (2021). Financial stress, race, and student debt during the Great Recession. *Social Currents*, *8*(5), 424–445. https://doi.org/10.1177/23294965211026692

Mustaffa, J. B., & Dawson, C. (2021). Racial capitalism and the Black student loan debt crisis. *Teachers College Record*, *123*(6), 1-28.

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism.* New York: New York University Press.

Persico, C., Figlio, D., & Roth, J. (2020). The developmental consequences of Superfund

> sites. *Journal of Labor Economics*, *38*(4), 1055-1097.

Rothstein, R. (2017). *The color of law: A forgotten history of how our government segregated

> America*. Liveright Publishing.

Seamster, L., & Charron-Chénier, R. (2017). Predatory inclusion and education debt: Rethinking

> the racial wealth gap. *Social Currents*, *4*(3), 199-207.

Shermer, E. T. (2021). Indentured students: How government-guaranteed loans left generations

> drowning in college debt. Cambridge, MA: *Harvard University Press.*

Soroka, S. N., Lawlor, A., Farnsworth, S., & Young, L. (2013). Mass media and policymaking.

> In W. Xun, M. Ramesh, M. Howlett, S. Fritzen, & E. Araral (Eds.), *Routledge Handbook
>
> of Public Policy* (pp. 1–15). Routledge.

Stratford, M., McMinn, S., & Thomas, T. M. (2023, Feb. 16). What we know about the 25m

> Americans who signed up for Biden's student debt relief. *Politico*.
>
> https://www.politico.com/news/2023/02/16/joe-biden-student-debt-relief-00083243

Stryker, J. E., Wray, R. J., Hornik, R. C., & Yanovitzky, I. (2006). Validation of database search

> terms for content analysis: The case of cancer news coverage. *Journalism & Mass
>
> Communication Quarterly*, *83*(2), 413-430.

Taliaferro, W., Taylor, T., & Wheatle, K. (2021). *Changing the narrative on student borrowers

> of color*. Washington, DC: Lumina Foundation. https://www.luminafoundation.org/wp-
>
> content/uploads/2021/02/borrowers-of-color-2.pdf

Tamir, C., Budiman, A., Noe-Bustamante, L., & Mora, L. (2021). Facts about the U.S. Black

> population. *Pew Research Center*. https://www.pewresearch.org/social-trends/fact-
>
> sheet/facts-about-the-us-black-population/#geography

Taylor, M., Turk, J. M., Chessman, H. M., & Espinosa, L. L. (2020). *Race and ethnicity in higher education: 2020 supplement.* Washington, DC: American Council on Education.

Turvill, W. (2021, Aug 25). Top 25 US newspapers by circulation: America's largest titles have lost 20% of print sales since Covid-19 hit. *Press Gazette.* https://www.pressgazette.co.uk/biggest-us-newspapers-by-circulation/

Vargas, N., & Villa-Palomino, J. (2019). Racing to serve or race-ing for money? Hispanic-Serving Institutions and the colorblind allocation of racialized federal funding. *Sociology of Race and Ethnicity*, *5*(3), 401–415. https://doi.org/10.1177/2332649218769409

Ward, L. W. (2021). Radical affirmative action: A call to address hegemonic racialized themes in US higher education race-conscious admissions legal discourse. *International Journal of Qualitative Studies in Education*, *34*(4), 315-329.

White House. (2022, Aug. 24). Fact sheet: President Biden announces student loan relief for borrowers who need it most. *White House*. https://www.whitehouse.gov/briefing-room/statements-releases/2022/08/24/fact-sheet-president-biden-announces-student-loan-relief-for-borrowers-who-need-it-most/

Wolfe, M. (2012). Putting on the brakes or pressing on the gas? Media attention and the speed of policymaking. *Policy Studies Journal*, *40*(1), 109–126. https://doi.org/10.1111/j.1541-0072.2011.00436.x

Woodly, D. R. (2015). *The politics of common sense: How social movements use public discourse to change politics and win acceptance*. Oxford University Press.

Figure 1. Annual number of student loan articles.
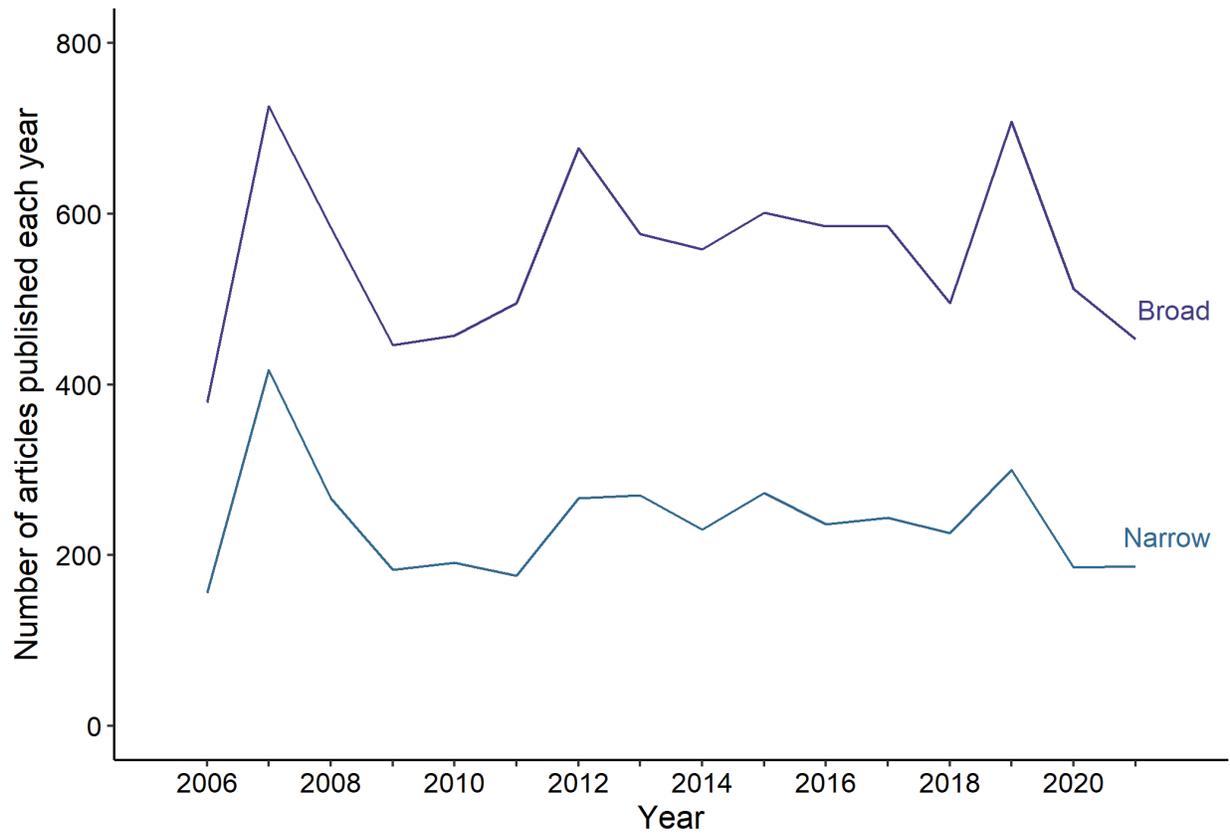
Note: Broad line is for articles that are broadly relevant (include a real mention of student loans at least once in the article). Narrow line is for articles that are narrowly relevant (focus of the article is on student loans).

Figure 2. Annual share of mentions of any racialized language term.
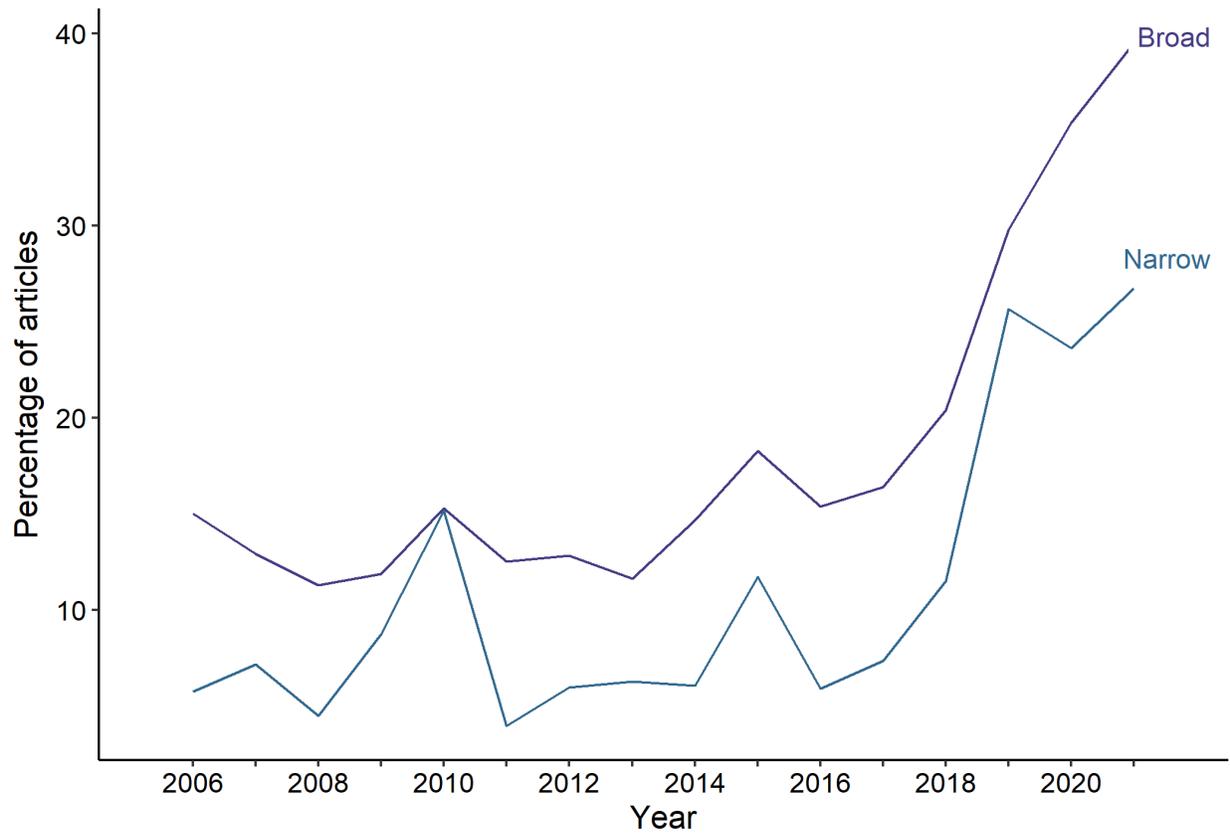
Note: Broad line is for articles that are broadly relevant (include a real mention of student loans at least once in the article). Narrow line is for articles that are narrowly relevant (focus of the article is on student loans).

Figure 3. Annual share of mentions of top five racialized language terms for narrowly relevant articles.
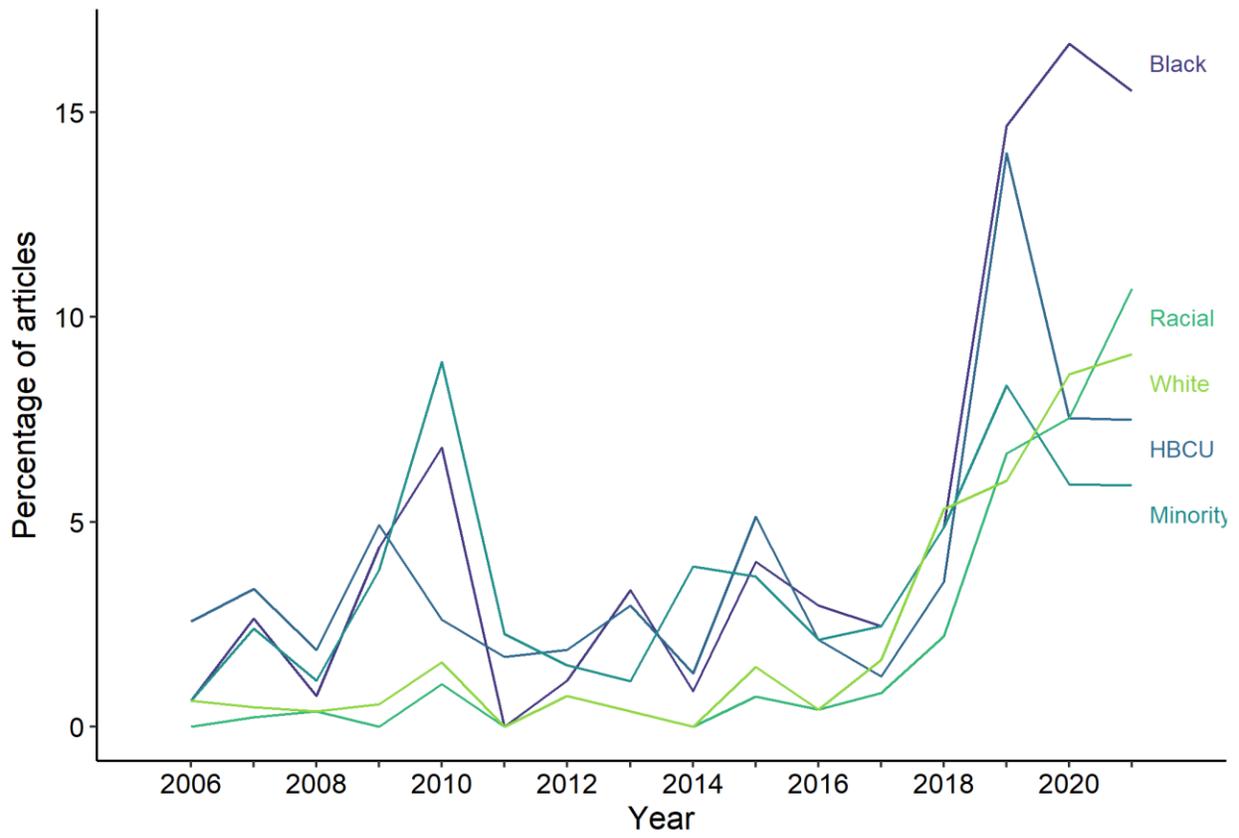

Note: Narrowly relevant articles are ones where the focus of the article is on student loans.

Figure 4. Annual percentage of articles with any racialized language term.
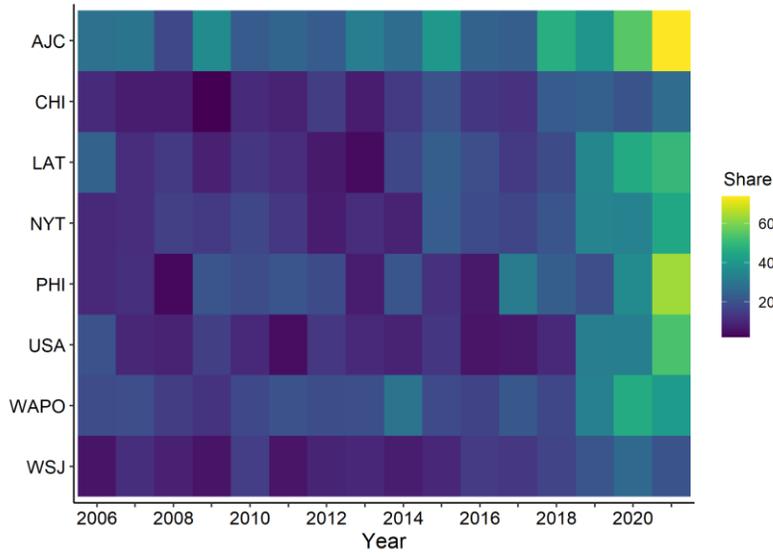
Panel A. Broadly relevant articles.
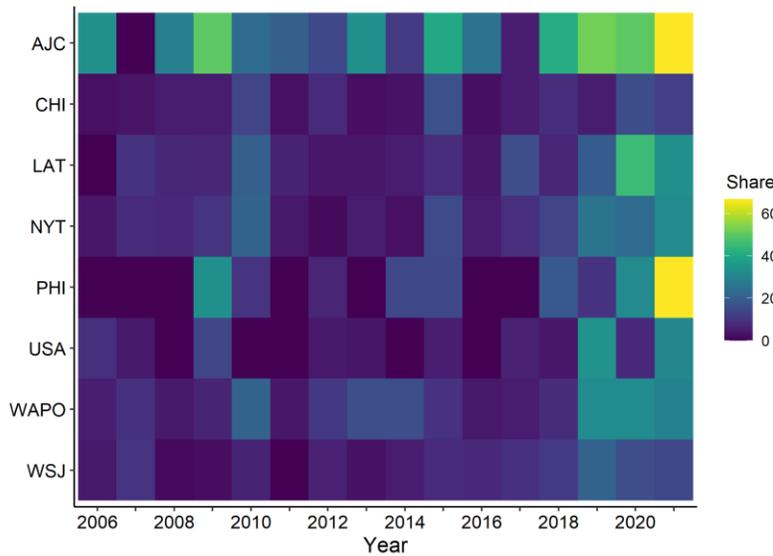
Panel B. Narrowly relevant articles.

Note: Panel A includes the annual percentages of articles that are broadly relevant (mention student loans at least once). Panel B includes the annual percentages of articles that are narrowly relevant (focus of the article is on student loans). More purple/blue cells represent years with a smaller percentage of articles using racialized language terms. More yellow/green cells represent years with a larger percentage of articles using racialized language terms. AJC = Atlanta Journal-Constitution, CHI = Chicago Tribune, LAT = Los Angeles Times, NYT = New York Times, PHI = Philadelphia Inquirer, USA = USA Today, WAPO = Washington Post, WSJ = Wall Street Journal

Figure 5. Annual percentage of articles with Black racialized language terms.
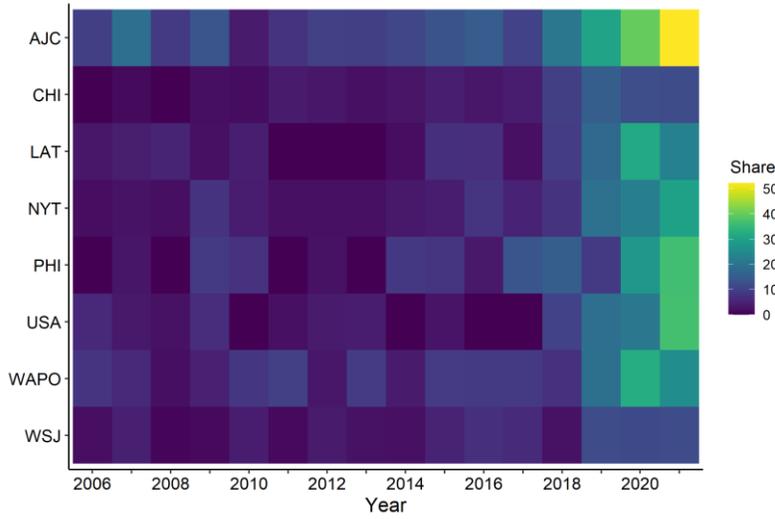

Panel A. Broadly relevant articles.
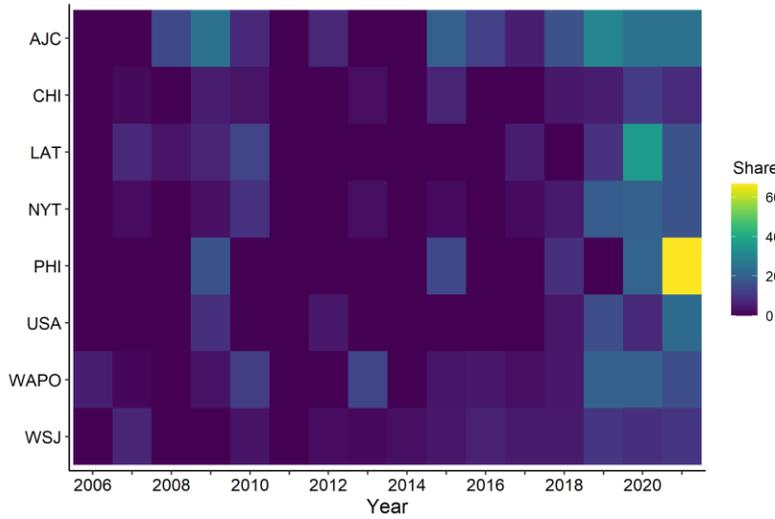
Panel B. Narrowly relevant articles.

Note: Panel A includes the annual percentages of articles that are broadly relevant (mention student loans at least once). Panel B includes the annual percentages of articles that are narrowly relevant (focus of the article is on student loans). More purple/blue cells represent years with a smaller percentage of articles using the term "Black." More yellow/green cells represent years with a larger percentage of articles using the term "Black." AJC = Atlanta Journal-Constitution, CHI = Chicago Tribune, LAT = Los Angeles Times, NYT = New York Times, PHI = Philadelphia Inquirer, USA = USA Today, WAPO = Washington Post, WSJ = Wall Street Journal

Figure 6. Annual percentage of people of color in the newsroom.
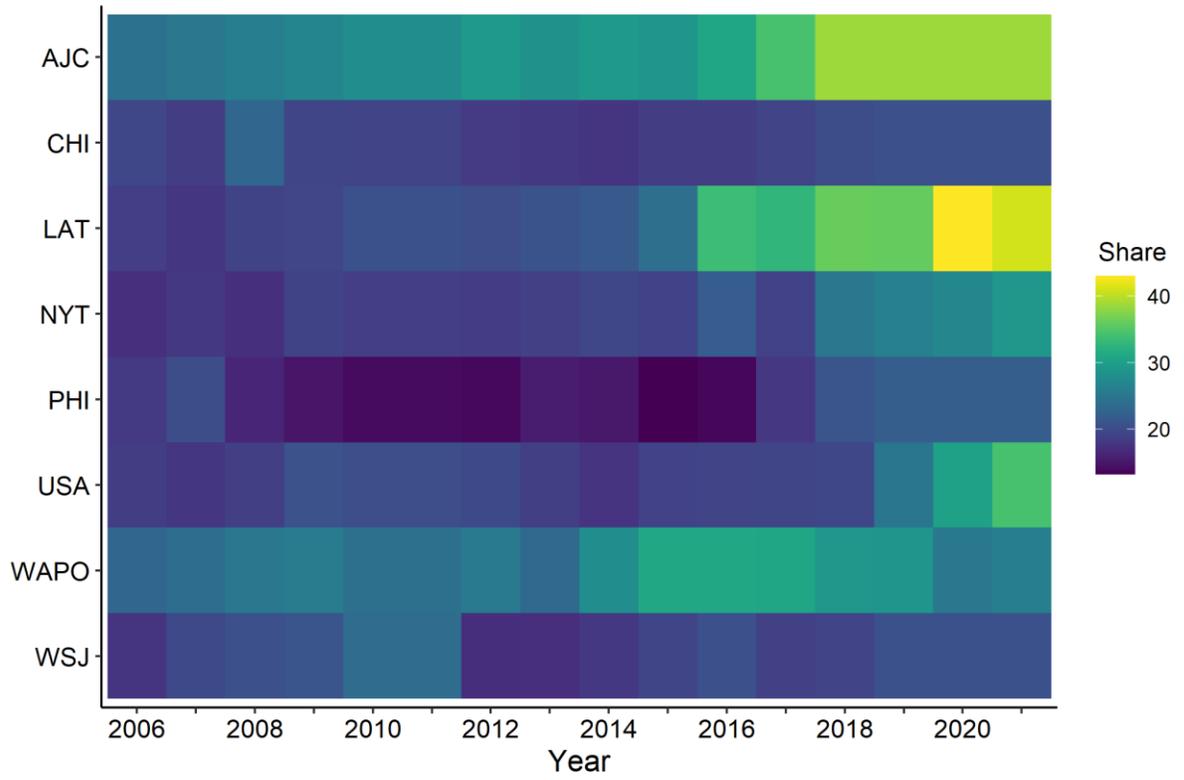
Note: Figure includes data for broadly relevant student loan news articles (mention student loans at least once). More purple/blue cells represent years with a smaller percentage of people of color in the newsroom. More yellow/green cells represent years with a larger percentage of people of color in the newsroom. AJC = Atlanta Journal-Constitution, CHI = Chicago Tribune, LAT = Los Angeles Times, NYT = New York Times, PHI = Philadelphia Inquirer, USA = USA Today, WAPO = Washington Post, WSJ = Wall Street Journal

Figure 7. Annual average decile of Black students at authors' colleges.
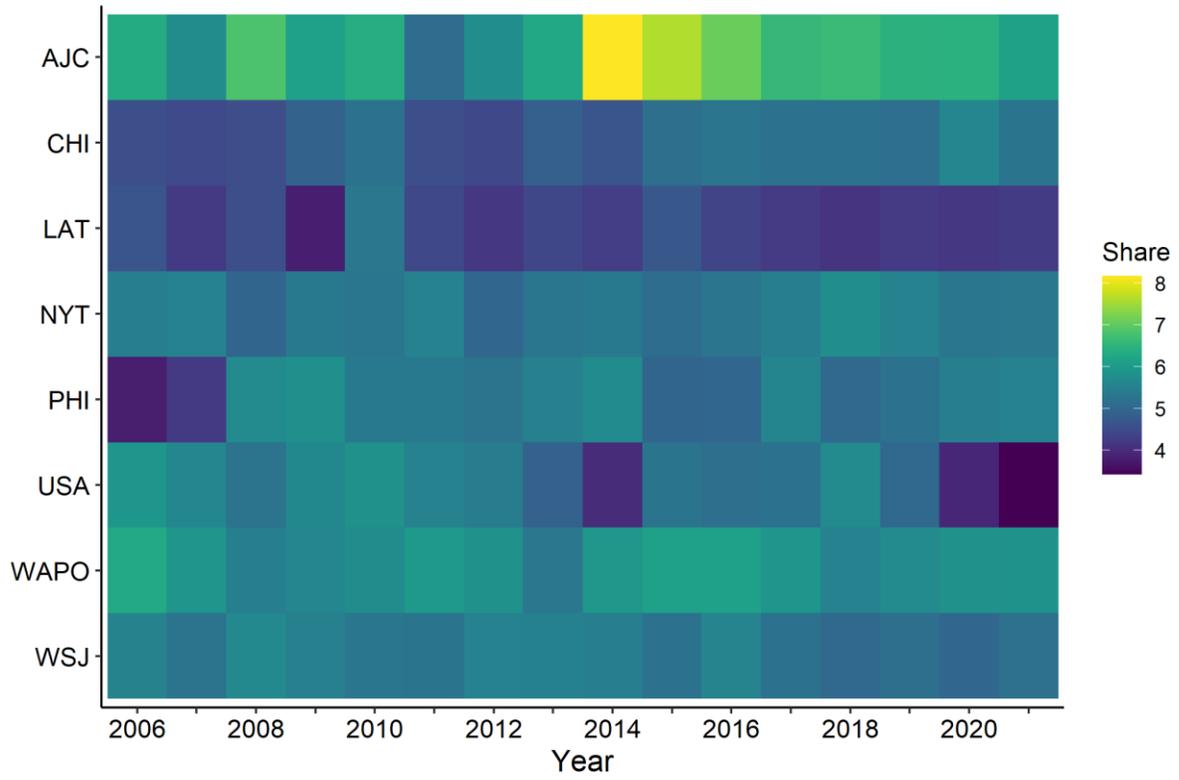
Note: Figure includes data for broadly relevant student loan news articles (mention student loans at least once). More purple/blue cells represent years with a smaller average decile of Black students at authors' colleges. More yellow/green cells represent years with a larger average decile of Black students at authors' colleges. AJC = Atlanta Journal-Constitution, CHI = Chicago Tribune, LAT = Los Angeles Times, NYT = New York Times, PHI = Philadelphia Inquirer, USA = USA Today, WAPO = Washington Post, WSJ = Wall Street Journal

Table 1. Newspaper inclusion criteria and final count of relevant articles.

Newspaper	Largest Circulation	Largest Number of Black Residents	Percent Removed for Broad Relevance	Final Count
Atlanta Journal-Constitution		X	91	500
Chicago Tribune		X	87	1,190
Los Angeles Times	X		90	764
New York Times	X	X	90	1,890
Philadelphia Inquirer		X	91	471
USA Today	X		79	711
Wall Street Journal	X	X	89	1,724
Washington Post	X	X	87	1,587
Total				8,837

Note: The first column lists the eight newspapers we purposively selected for the proposed project. Those with an X under the "Largest Circulation" column are newspapers selected due to their circulation size (the top five in the United States). Those with an X under the "Largest Number of Black Residents" column are newspapers selected due to their metropolitan area's number of Black residents (top five in the United States). The Percent Removed column shows the percentage of articles removed from each newspaper based on the broad relevance criteria. The Final Count column shows the final number of articles from each newspaper for the analytical data.

Table 2. Percentage of articles with structural racialized language.

	Black		Racism		Racial Equity		Racial Equality	
	All	Pre-2020	All	Pre-2020	All	Pre-2020	All	Pre-2020
Atlanta Journal-Constitution	16.60	13.79	0.60	0.44	0.20	0.00	0.00	0.00
Chicago Tribune	4.29	3.51	0.59	0.37	0.00	0.00	0.00	0.00
Los Angeles Times	6.28	4.53	0.92	0.42	0.13	0.14	0.00	0.00
New York Times	8.15	5.29	1.16	0.74	0.26	0.00	0.11	0.06
Philadelphia Inquirer	8.07	5.66	1.70	0.94	0.21	0.00	0.00	0.00
USA Today	7.17	4.86	1.13	0.47	0.14	0.00	0.28	0.00
Wall Street Journal	4.81	4.02	0.12	0.00	0.12	0.00	0.06	0.00
Washington Post	10.59	8.03	1.89	1.08	0.32	0.14	0.57	0.29

Note: "All" columns include all analytical years (2006 to 2021) and "Pre-2020" columns use the same set of years, removing the more recent years following the publicizing of the murders of several Black people (2006 to 2019). Each cell includes the percentage of articles (published across the noted time period) that use the indicated racialized language term.

Table 3. Relationship between newspaper characteristics and racialized language use.

	Any				Black			
	Broad		Narrow		Broad		Narrow	
Authors' undergraduate institutions								
All "Ivy Plus"	-0.028	0.007	-0.020	0.006	-0.028	-0.002	-0.018	0.009
	(0.147)	(0.107)	(0.065)	(0.056)	(0.088)	(0.050)	(0.046)	(0.040)
"Ivy Plus" or public flagship	-0.103	-0.037	-0.057	-0.005	-0.086	-0.033	-0.047	-0.007
	(0.069)	(0.048)	(0.047)	(0.045)	(0.051)	(0.028)	(0.037)	(0.036)
Loan reliance (decile)	0.837	0.719	-1.130	-1.098	0.717	0.537	-0.181	-0.198
	(2.009)	(1.494)	(1.606)	(1.325)	(1.388)	(0.820)	(1.022)	(0.695)
Black enrollment (decile)	1.832	3.355**	2.654	3.368*	0.753	2.224*	0.305	1.053
	(1.878)	(0.725)	(1.755)	(1.305)	(1.771)	(0.871)	(1.211)	(0.626)
POC in newsroom	1.367**	0.845**	1.185**	0.732	0.980**	0.492*	0.688**	0.266
	(0.276)	(0.197)	(0.329)	(0.396)	(0.238)	(0.162)	(0.158)	(0.174)
Year fixed effect included	X		X		X		X	
	128	128	128	128	128	128	128	128

Note: The first four columns have the percentage of articles with any racialized language as the outcome and the second four columns have the percentage of articles using the word "Black" as the outcome. "Broad" columns are the percentage of articles that are broadly relevant (at least one real mention of student loans). "Narrow" columns are the percentage of articles that are narrowly relevant (focus of the article is on student loans). All variables without decile in the title are percentages. All "Ivy Plus" is the annual percentage of articles with education authors who attended Ivy Plus institutions (Ivy League athletic association plus Duke University, Massachusetts Institute of Technology, Northwestern University, Stanford University, and the University of Chicago). "Ivy Plus" or public flagship is the annual percentage of articles with at least one author who attended an Ivy Plus or public flagship institution (excluding author teams that are solely Ivy Plus enrollees). POC in newsroom is the annual percentage of people of color who are in the newsroom. Standard errors are clustered at the newspaper level.

Table 4. Modifying words in sentences containing most frequent racialized language terms.

	Rank	Black	HBCU	Minority	White	Racial
Noun	1	percent	college	college	percent	wealth
	2	college	university	income	college	student
	3	student	student	percent	wealth	debt
	4	debt	percent	education	debt	percent
	5	university	school	student	student	college
Adjective	1	white	black	black	black	black
	2	other	federal	low	likely	white
	3	many	other	other	american	economic
	4	new	new	many	african	new
	5	federal	financial	federal	many	other

Note: Within each racialized language term column, we rank the most frequent nouns and adjectives (from 1 to 5). Modifying words are all lowercase given that we prepared our data by removing capitalization and therefore treat capitalized and uncapitalized words the same.

Appendix A
Identification of Near Duplicates

For every newspaper, we compared the full text of each article to articles published one week before and one week after the focal article. For every article pairing (focal article and another article published within the time window), we calculated a similarity score. To calculate that score, we first used *stringdist* in R to estimate the Levenshtein distance between the two articles. The Levenshtein distance is the number of changes that needs to happen for one string to become another (via insertion, deletion, or substitution). For example, the Levenshtein distance for "teacher" and "preacher" is two given that you insert a "p" and substitute "t" for "r." Once we have estimated that distance, we created a similarity percentage using the following model:

$$100 * \left(1 - \frac{LV(focal\ article,\ comparison\ article)}{characters(focal\ article,\ comparison\ article)} \right)$$

In other words, we estimated the Levenshtein distance, then divided it by the number of characters in the larger of the two articles (either the focal or comparison article). We then subtracted the result from one (to obtain the inverse of the distance, a rough measure of how similar the articles were) and multiplied it by 100 to create a percentage.

We reviewed all articles with at least an 80% similarity score and removed any duplicates (retaining the first version of the article available in the data). For articles with more than 33,000 characters (461), which we could not analyze using the method outlined above due to size restrictions with *stringdist*, we visually verified that there were no duplicate articles published one week before or after.

Appendix B
Decision Rules for Broad and Narrow Relevance

Articles were immediately excluded from broad relevance if their title included "obituar*", "death notice", "corrections", or "area member of congress voted" (which was a list of yes and no votes from members of Congress). Once those articles had been removed, we identified articles as having broad relevance if they met two criteria:

1. Article must mention debt term ("loan*", "debt*", "borrow*") at least twice
2. Either the article has:
   a. "Student" within 4 words of debt terms
   b. Aid term ("financial aid*", "grants and aid*", "student aid*", "college aid*", "public service loan forgiveness*", "stafford loan*", "perkins loan*", "pell grant*", "parent plus loan*", "grad plus loan*", "income driven repayment*") mentioned at least twice

From within articles that we defined as broadly relevant, we created a binary measure that equaled 1 if the percentage of "debt" terms in the articles was greater than 1%.

For the recall and precision estimates, we used Stryker et al.'s (2006) equations. Recall was defined as:

$$Recall = \frac{\#\ Articles\ humans\ found\ relevant\ that\ we\ retain\ with\ decision\ rules}{\#\ Articles\ humans\ found\ relevant}$$

Precision was defined as:

$$Precision = \frac{\#\ Articles\ humans\ found\ relevant\ that\ we\ retain\ with\ decision\ rules}{\#\ Articles\ the\ decision\ rules\ retain}$$

Appendix C: Racialized Language Terms

As noted in the main article, we also removed all punctuation when preparing the data which means that "non-tribal" in an article would be "non tribal" in our analytical data. Additional detail on the coding for "White," "Black," and "minorit*" can be found in Appendix D.

Racialized Language Category	Terms Included**
Ethnoracial Labels	White, Caucasian, Black, African American, Asian, Native Hawaiian, Pacific Islander, Native American, Alaskan Native, American Indian, Indigenous, Hispanic, Latino, Latina, Latinx, Chicano (including Chicano and Chicana), south Asian, east Asian, Middle Eastern and North African (including Middle Eastern North African), biracial (including biracial, bi racial, multiracial, and multi racial)
Collective Phrases Indicating Race	minorit*, underrepresented racial, of color
Racialized Institutions	Minority Serving Institutions, Historically Black Colleges and Universities (including Historically Black College, Historically Black University, and each individual institution name¥), Tribal Colleges and Universities (including Tribal College and Tribal University), Predominantly Black Institutions, Asian American and Native American Pacific Islander Serving Institutions, Hispanic Serving Institutions, Native American Serving Non-Tribal Institutions (including spelling this as Non Tribal and NonTribal)
Racialized Organizations	National Association for the Advancement of Colored People, United Negro College Fund, Excelencia, Thurgood Marshall College Fund, Black Lives Matter
Systemic Issues	racial inequity, racial inequality, racism
Overarching	racial, ethnic, ethnicity, racialized

**For each of these terms, we included plural, singulars, and acronyms (e.g., we searched for both "National Association for the Advancement of Colored People" and "NAACP").

¥ During the research team's review of 400 randomly selected newspaper articles (stratified by newspaper), we found that several HBCUs, and only HBCUs, were listed by the institution's name without any reference to their status as an HBCU. Given the particular focus of this study on Black borrowers' experiences with student loans, we also included the name of each individual HBCU. We used the 2021 data collection from the Integrated Postsecondary Education Data System (IPEDS) to construct the list of all HBCU names. We obtained the 102 names, removed all punctuation (ampersands, apostrophes, and periods were replaced with nothing and hyphens were replaced with spaces aligning with our preparation of the full text), and created alternative versions in case of typos. For example, "Alabama A & M University" in IPEDS became the following options: "alabama a m university" and "alabama am university."

Appendix D: Differentiating Black, White, and Minority from Non-Racialized Uses

We note that the terms used in this appendix are frequently lowercase given that the data is lowercase when we conducted this portion of the data preparation. As noted in the main article, we also removed all punctuation when preparing the data which means that "you're" in an article would be "youre" in our analytical data.

To differentiate between racialized and non-racialized use of White, Black, and minority, the authors conducted three iterative rounds of coding using a random sample of 90 articles, 30 for each term. We reviewed the samples for systematic patterns and to create decision rules. We created a list of common words that precede or follow racialized uses (e.g., "Black man", "who is White"), removed compound words (e.g., "whitehead") and specific phrases (e.g., "House Minority Leader"). In our last round of checking, our resulting decision rules accurately located 90% of racialized uses of Black, 92% of racialized uses of White, 100% of racialized uses of the words minority or minorities, and, most importantly, did not incorrectly identify any non-racialized uses as being racialized.

For "minority," we included any article that used the string "minorit" as long as that string was not followed or preceded by the following words:

Not Followed By	Not Preceded By
leader	in the
shareholder	
shareholders	
partner	
partners	
of	

For "Black" and "White," we included any article that used either of those terms followed by this set of words:

actor, actors, actress, actresses, adult, adults, agenda, alumni, alumnus, america, american, americans, and low income, applicant, applicants, associates degree holder, associates degree holders, audience, audiences, babies, baby, baccalaureate degree holder, baccalaureate degree holders, bachelors degree holders, bachelors degree holder, borrower, borrowers, boy, boys, business, businessman, businessmen, businesswoman, businesswomen, buyer, buyers, campus, candidate, candidates, caucus, citizen, citizens, classmate, classmates, college, colleges, communities, community, congressman, congressmen, congresswoman, congresswomen, consumer, consumers, counselor, counselors, counterpart, counterparts, county commissioner, county commissioners, couple, couples, critic, critics, crowd, crowds, culture, customer, customers, death row inmate, death row inmates, debt, democrat, democratic, democrats, doctor, doctors, driver, drivers, educator, educators, electorate, elevator operator, elevator operators, employment, ex congressman, ex congressmen, ex congresswoman, ex congresswomen, faculty, families, family, female, females, financial institution, financial institutions, founder, founders, four year, freshman, freshmen, girl, girls, graduate, graduates, graduation, group, groups, guy, guys, hbcu, hbcus, historical figure, historical figures, homeowner, homeowners, homeownership, household, households, individual,

individuals, institution, institutions, investor, investors, junior, juniors, justice, kid, kids, leader, leaders, listener, listeners, lives, male, males, man, mayor, mayors, medical school, medical schools, men, middle income, millenial, millenials, mob, mobs, mortgage borrower, mortgage borrowers, neighborhood, neighborhoods, observer, observers, one, ones, owned, parent, parents, participant, participants, patient, patients, peer, peers, people, person, physician, physicians, player, players, politics, population, power, president, presidents, primary care doctor, primary care doctors, progressive, progressives, race, ran, recipient, recipients, representative, representatives, republican, republicans, researcher, researchers, resident, residents, run, scholar, scholars, school, schools, science and engineering phd recipients, senate, senator, senators, senior, seniors, shoe shiner, shoe shines, skin, society, sophomore, sophomores, state school, state schools, student, students, supremacy, surgical laboratory supervisor, surgical laboratory supervisors, teacher, teachers, teammate, teammates, television, televisions, trailblazer, trailblazers, tv, union, unions, universities, university, voice, voices, vote, voter, voters, votes, wall street, wealth, woman, women, worker, workers, working class, young adult, young adults, youth.

For example, we searched to see if any article used the phrase "black worker" and, if it did, we tagged that as a racialized use of the term "Black." Given the way we coded the search for this syntax, we also needed to do a separate search for "Black" that included if any of these nouns were preceded by "black and white" or "black or white." As an example, this means we also searched for whether an article mentioned "black and white students." We added this additional step because, if we had not done this, the prior sentence's example would have only been tagged as a racialized use of the term "White."

We also included any use of those two terms that was preceded by the following words: who are, who is, percent, historically, youre, association of, society of, alliance of.

Finally, we also included a separate set of words that was particular to either "Black" or "White." We selected the majority of these sets based on prior linguistic research and the racialized language review of 400 randomly selected articles (stratified by newspaper) outlined above and in the main article.

	Included Words
Black	blacks, blackness, black ish, blackpac, black and brown, black brown, black and latino, black and hispanic, black and latinx, black latino, black hispanic, black latinx, black white, nonblack, black hbcu, black or brown, black or latino, black or hispanic, black or latinx
White	whites, whiteness, white and asian, white asian, white black, nonwhite, white or asian, white passing

Appendix E

Appendix E includes additional information on the regression models. It includes summary statistics on all covariates and the outcomes (Table E1) and the decile cutoffs for loan reliance and the Black student share of undergraduate enrollment (Table E2).

Table E1. Descriptives of key newspaper characteristics

	Broad	Narrow
Number of articles	69.04	29.76
	(38.64)	(19.44)
HBCU author (%)	1.88	2.09
	(7.29)	(10.22)
All "Ivy Plus" authors (%)	22.45	19.96
	(14.30)	(18.98)
"Ivy Plus" or public flagship author (%)	29.40	30.21
	(18.10)	(24.03)
Loan reliance (decile)	5.05	5.06
	(0.88)	(1.11)
Black student enrollment (decile)	5.32	5.41
	(0.74)	(0.94)
People of color in newsroom (%)	22.76	22.76
	(6.22)	(6.22)
Outcomes		
Any racialized language (%)	19.68	12.89
	(12.87)	(13.80)
Black (%)	8.53	6.08
	(9.44)	(9.36)

Note: "Broad" column is for broadly relevant student loan articles. "Narrow" column is for narrowly relevant student loan articles. Each cell includes the mean with the standard deviation in parentheses. All "Ivy Plus" is the annual percentage of articles with education authors who attended Ivy Plus institutions (Ivy League athletic association plus Duke University, Massachusetts Institute of Technology, Northwestern University, Stanford University, and the University of Chicago). "Ivy Plus" or public flagship is the annual percentage of articles with at least one author who attended an Ivy Plus or public flagship institution (excluding author teams that are solely Ivy Plus enrollees). For loan reliance, we calculated the decile rank for the 2018 percentage of undergraduate students receiving federal student loans. For Black undergraduate student share, we calculated the decile rank for the 2018 percentage of total undergraduate student enrollment who identified as Black. For both measures, we only included public and private not-for-profit institutions (e.g., each decile rank is of not-for-profit institutions in the United States).

Table E2. Decile cutpoints.

Deciles	Loan reliance	Black undergraduate student share
1	0	0.87
2	6	2.30
3	16	3.77
4	28	5.23
5	39	7.06
6	48	9.57
7	56	13.33
8	66	19.24
9	75	32.14

Note: Each cell includes the value at each decile cutpoint. For loan reliance, we calculated the decile rank for the 2018 percentage of undergraduate students receiving federal student loans. For Black undergraduate student share, we calculated the decile rank for the 2018 percentage of total undergraduate student enrollment who identified as Black. For both measures, we only included public and private not-for-profit institutions (e.g., each decile rank is of not-for-profit institutions in the United States).