Short-run Retaliation, Long-run Updating: Black Leadership and Racial Animus

Claire Duquennois and Jiangnan Zeng*

June 16, 2026

Abstract

This paper studies the dynamics of racial animus when visible minority leaders are repeatedly tied to both success and failure. We focus on Black quarterbacks in the National Football League, a setting in which leadership is highly visible, historically racialized, emotionally salient, and publicly linked to team performance. Combining game- and season-level data with four different measures of racial animus, we document two opposing responses. First, relative to games where both quarterbacks were white, emotionally intense losses involving Black quarterbacks increase racial animus in subsequent days: anti-Black hate crimes increase by 43%, online hate-speech indicators rise by 29%, and implicit bias increases by 0.043 SDs; comparable wins do not reduce animus. Second, over the course of a season, successful teams led by Black quarterbacks generate persistent reductions in hate crimes (34%), racial-slur searches (3.7%), and implicit bias (-0.040 SDs). Opposition and low-performing Black quarterback teams have limited impacts. Black head coaches produce no comparable effects, highlighting the importance of media visibility and public attribution. These results show that minority representation is not uniformly prejudice-reducing: in the short run, highly visible minority leaders can activate animus through racialized blame but, when successful, reduce it through performance-contingent updating over longer horizons.

*Duquennois: Department of Economics, University of Pittsburgh, Pittsburgh, PA 15260 (email: ced87@pitt.edu) Zeng: Gordon S. Lang School of Business and Economics, University of Guelph, Guelph, ON N1G 1M8 (email: jiangnan@uoguelph.ca). This research was supported by a grant from the Russell Sage Foundation.

1 Introduction

What shapes racial prejudice, and under what conditions does it change? A large literature documents that racial prejudice and discrimination are prevalent, persistent, and consequential across social and economic settings (Pager and Shepherd 2008; Bertrand and Duflo 2017; Lang and Kahn-Lang Spitzer 2020; Small and Pager 2020). Work in the social sciences and in economics has sought to understand their sources (Becker 1957; Glaeser 2005; Bordalo et al. 2016; Onuchic 2022), what causes them to increase (DellaVigna and La Ferrara 2015; Zhuravskaya et al. 2020), and, critically, how they can be reduced (Paluck et al. 2021; Paluck and Green 2009). Existing empirical work, however, typically isolates a single force acting in one direction: a shock that inflames hostility or an intervention that reduces it. Less is known about how prejudice evolves when the same minority group is repeatedly and publicly tied to both success and failure, where repeated successes may erode stereotypes and shift social norms while salient failures activate racialized blame. Research on such dynamics has largely focused on individual-level discrimination rather than population-level animus (Bohren et al. 2019; Glover et al. 2017). We address this gap by studying a single real-world setting in which prominent minority leaders' performance is repeatedly observed, showing that exposure to Black leadership can be both animus-activating and animus-ameliorating, with the direction depending on performance and time horizon. We trace these responses across four measures of racial animus that differ in visibility and severity, ranging from private expression and implicit bias to public hate speech and hate crimes.

We study these dynamic responses to Black leadership in a recurring context that is historically racialized, uniquely visible, media-saturated, and emotionally intense: quarterbacks in the National Football League (NFL). As team leaders, quarterbacks are well-known public figures who are routinely praised or blamed for their team's success or failure. Because the position was historically dominated by white players due to racialized stereotypes about Black intelligence, leadership, and decision-making (Rhoden 2007; Reid 2022), contemporary Black quarterbacks are especially salient figures for judgments about Black leadership and competence. These features let us examine how local racial animus responds when a visible Black leader is publicly tied to success or failure. We combine game and season data with multiple measures of racial animus: anti-Black hate crimes, online hate-speech moderation, racial-slur searches, and implicit bias, allowing us to examine how these effects generalize across different manifestations of animus.

We examine how Black quarterbacks impact both short-run game-day responses and longer-

run seasonal changes in animus. In the short-run, we use an event-study design to investigate how racial animus changes in the five days surrounding games. We find evidence of an affect-driven racial retaliation effect: racial animus increases when Black quarterbacks are salient, and therefore blamable, for an emotionally intense negative outcome. Relative to similar games where both quarterbacks were white, emotionally intense close-scoring upset losses involving a Black quarterback are followed by a 43% increase in the anti-Black hate crime rate, a 29% increase in online hate speech indicators on Reddit, and a 0.043 SD rise in implicit racial biases. In contrast, animus indicators are unchanged after emotionally intense wins involving Black quarterbacks. This asymmetry, in which an individual's failures inflame animus toward their broader group while successes do not attenuate it, mirrors the attribution asymmetry (Pettigrew 1979) documented at the individual level in evaluation, credit, and punishment (Sarsons 2017a,b; Egan et al. 2022) and in experimental studies on racial animus spillovers (Bursztyn et al. 2022; Bauer et al. 2023). We provide the first evidence of this attribution asymmetry impacting real-world spillovers on animus against an entire racial group.

In the long run we turn to examining the persistent effects of exposure to local and opposition Black quarterbacks over the course of an NFL season. The short-run effects are concentrated in a small set of emotionally engaging games. Our season-level analysis examines a different and potentially more consequential margin: whether repeated exposure to Black quarterbacks changes baseline racial animus beyond the immediate emotional window surrounding a game. For local Black quarterbacks, we compare how outcomes evolve in areas with Black and white quarterback teams over the course of an NFL season. In the longer run, we find a pattern of performance-contingent updating: sustained exposure to high-performing Black leadership reduces animus while Black quarterbacks on low performing teams have limited effects. In areas where Black-quarterback teams advance to the championships, hate crime rates are 34% lower, Google searches for racial slurs decline by 3.7%, and implicit racial biases drop by -0.040 SDs, patterns that persist into the months following the end of the NFL season. For opposition Black quarterbacks, we find no evidence that increased exposure to opposing Black quarterbacks leads to persistent change in racial animus.

These persistent animus reductions are consistent with contact theory, though its mechanisms emphasize in-person contact between individuals of similar status (Allport et al. 1954; Rao 2013; Lowe 2021; Mousa 2020). Work on minority leadership has often focused on same-group role-model effects (Riley 2024; Gershenson et al. 2022; Lim and Meer 2020; Fairlie et al. 2014; Dee 2005),

with some evidence on animus impacts in the political setting (Sakong 2026; Beaman et al. 2009). Parasocial contact theory is closer to our setting as it proposes that media exposure to positively represented out-group role models can reduce prejudice (Schiappa et al. 2005). While a number of studies examine how the media impacts racial animus, existing work has overwhelmingly focused on how media can inflame it (DellaVigna and La Ferrara 2015; Yanagizawa-Drott 2014; Adena et al. 2015; Ang 2023; Wang 2021; Müller and Schwarz 2023; Cao et al. 2023; Petrova et al. 2019). Often noted in reviews of the contact theory and role-model literatures (Paluck et al. 2021; Kearney and Levine 2020; Bertrand and Duflo 2017), evidence that parasocial contact can reduce animus is far rarer (Paluck 2009; Blouin and Mukand 2019; Duquennois and Zeng 2025; Zussman 2023). Closest to our setting is Alrababa'h et al. (2021), who document reduced anti-Muslim hate crimes and speech in Liverpool following the arrival of the Muslim soccer star Mohamed Salah and his team's subsequent success. Our seasonal findings advance this work in several ways: (i) we move beyond the study of a single celebrity to estimate the impacts of 68 Black quarterbacks observed over two decades; (ii) we examine effects by team performance revealing that effects are performance-contingent; (iii) we show these animus spillovers occur only when paired with high personal media visibility; and iv) we can contrast the seasonal effects with game-day increases after losses, shedding light on the temporal dynamics of racial animus.

Finally, this paper provides the first causal evidence that one of America's most visible non-political institutions—the NFL—can shape racial animus and social cohesion within the United States. This contributes to the literature on the economic and social impacts of sporting events. Research has shown that sporting events generate powerful emotional shocks with important social consequences, including effects on domestic violence (Card and Dahl 2011; Matsuzawa and Arnesen 2024), other violent crimes (Rees and Schnepel 2009; Munyo and Rossi 2013; Montolio and Planells-Struse 2016; Ivandić et al. 2024), judicial decisions (Eren and Mocan (2018)), financial markets (Edmans et al. (2007)), fertility (Stoecker et al. (2016)), and infant health (Duncan et al. (2017)). Our short-run results show that these emotional shocks can be racialized: games can generate retaliatory behavior toward out-groups who share an identity with salient athletes. A couple of existing papers have examined the impact of international sporting tournaments on national identity building. Depetris-Chauvin et al. (2020) show that international soccer tournaments strengthen national identities and reduce civil conflict, and Rosenzweig and Zhou (2021) document a similar effect, with a concurrent increase in negative views towards foreign refugees. We extend this work by examining how sports affects within-country intergroup relations, showing that sports

can shape social cohesion by shaping majority-group attitudes toward minorities.

The remainder of this paper is organized as follows. Section 2 presents contextual information on the NFL and Black quarterbacks in the NFL. Section 3 outlines our model. Section 4 presents our data. Section 5 focuses on our short-run analysis, presenting first our estimation strategy followed by our short-run results. Section 6 develops our season analysis beginning with the estimation strategy and then presenting results. Section 7 examines the effect of head coaches, a form of minority leadership that differs sharply from quarterbacks in public visibility. Section 8 examines how the effects of exposure to Black quarterbacks vary with an area's racial context. Section 9 concludes.

2 The NFL and Black quarterbacks

The National Football League (NFL) is the highest level of professional American football and the most popular sporting league in the United States, receiving high levels of media attention. It is estimated that Super Bowl games, which determine the season's winning team, are watched by over 40% of the US population. Yet even regular season games are highly visible events averaging over 16 million viewers. The league is composed of 32 teams who are strongly rooted in local identity with fans overwhelmingly following their local team. The regular season, which typically runs from September to January, features weekly games. These often occur on Sundays (85% of games), though teams occasionally play on Mondays (7%), Thursdays (5%), and Saturdays (4%). The regular season consisted of 16 games until the 2021 season when it was increased to 17 games per season. During the regular season, games are determined by a scheduling formula. After the regular season, the best teams qualify for the playoffs (12 teams prior to the 2020 season and 14 teams thereafter). The playoffs, typically running throughout January, are a knockout tournament that leads to the "semi-final" conference championship games (4 teams), and finally to the Super Bowl, usually held in the first two weeks of February.

Games are highly structured, with frequent stops in play, highly specialized positions, and players who take the field only for offense or defense. At the center of each team's offense is its primary "starting" quarterback, who receives the ball at the start of most plays, makes rapid decisions under pressure, and implements the team's strategy. Consequently, the quarterback is

often the team's most visible and scrutinized player, receiving a disproportionate share of both media attention and public credit or blame for the team's performance. The position is widely regarded as a "high IQ" role and carries deep symbolic associations with leadership and masculinity in American culture.

Though the NFL began reintegrating Black players in 1946, for decades Black athletes were largely excluded from the quarterback position due to racist stereotypes questioning their intelligence, leadership, and decision-making ability (Rhoden (2007), Reid (2022)). In 1953, Willie Thrower broke that barrier as the first Black quarterback to play in an NFL game, and in 1968, Marlin Briscoe became the first to start as quarterback in the modern era. Though Black athletes have long comprised the majority of NFL players–roughly 70% since 1990–they remained severely underrepresented at quarterback for decades, with the number of Black starting quarterbacks rising only gradually from the late 1970s onward (Marquez-Velarde et al. (2023)).[1] This history of exclusion has made Black quarterbacks culturally marked figures whose presence at the position remains especially salient. As the player most publicly praised and blamed for collective outcomes, a Black quarterback becomes a prominent and racialized focal point for judgments about competence and leadership.

In our period of study, about 20% of starting quarterbacks were Black between 2002 and 2020, rising notably to 37% by 2024 (Figure 1). The position is occupied almost exclusively by players of European or African-American descent. Almost all NFL players are American, allowing us to focus on responses to racial identity unconfounded by immigration status.

3 Modeling Black athletes' impact on racial animus

We model how exposure to Black quarterbacks affects racial animus in both the short-run, right around game-day, and over the course of a season. Below we discuss potential mechanisms and their respective predictions before presenting our empirical findings in sections 5 and 6.

[1] Racial discrimination within professional sports has been studied fairly extensively. We know that racial biases have influenced referees, umpires' calls, and player assessments (Price and Wolfers (2010), Parsons et al. (2011), Gallo et al. (2013), Kamel and Woo-Mora (2023)) and that the racist behavior of fans impacts player performance (Caselli et al. (2023), Caselli et al. (2024)). In the NFL, several studies have explored the effects of player race on careers and pay with mixed findings, reflecting the wide variety of seasons, positions, and approaches employed for analysis (most recently Berri et al. (2023), Gregory-Smith et al. (2023), Keefer and Kniesner (2023), Keefer (2013), Burnett and Van Scyoc (2015), Ducking et al. (2014), Ducking et al. (2017)). With respect to quarterbacks, it is widely acknowledged that the position was historically much less accessible to Black players. In more recent periods, some evidence suggests Black quarterbacks experience differential treatment in benching and pay (Volz (2017), Berri and Simmons (2009), Berri et al. (2023)).

Let A_{das} denote racial animus on day d of season s in area a. Observed outcomes — hate crimes, online hate speech, racial slur searches, and implicit bias scores — are increasing functions of A_{das}. A_{das} is shaped by baseline animus, $\bar{A}_{as}$, a slow-moving state variable that evolves between seasons, capturing long-run attitude change. In the short-run, animus on day d around game g deviates from this baseline by r_{dag}, a game-induced shock, such that in expectation:

$$E[A_{dag}] = \bar{A}_{as} + r_{dag}. \tag{1}$$

3.1 *Racial Retaliation*: Short-run emotion-driven retaliatory racial animus

In the short-run, we model a process of *racial retaliation* whereby majority group animus toward out-group bystanders is triggered by affect shocks associated with salient out-group members. Racial animus, r_{dag}, is impacted in a two-step process. First, an event generates an affect shock. This in turn impacts racial animus if (i) the affect shock is attributed to a Black actor and (ii) this attribution impacts anti-Black animus aimed at the wider Black community.[2]

For our context, we model the affective response a game generates in team catchment areas as dependent on three key factors: (i) how engaged area a is in team n's outcomes; (ii) the emotional intensity of the game g [3]; and (iii) whether the game is won or lost:

$$E[Affect_{ang}] = Engagement_{an} \cdot (\gamma_0 + \gamma_1 Intensity_{ng}) \cdot (\alpha_1 Win_{ng} - \alpha_2 Loss_{ng}). \tag{2}$$

In expectation, the affect generated by game g is increasing in wins and decreasing in losses ($\alpha_1, \alpha_2 > 0$). Furthermore, these affect changes will be greater in magnitude after (i) more intense games ($\gamma_1 > 0$); and (ii) in more engaged areas ($\frac{\partial |Affect|}{\partial Engagement} > 0$).

These affect shocks inversely translate into racial animus shocks ($\frac{\partial r}{\partial Affect} < 0$) when (i) the shock is attributed to a Black quarterback; and (ii) this attribution spills over into broader anti-Black animus. We model this as

$$E[r_{ang}] = (-1) \cdot E[Affect_{ang}] \cdot \left[\pi^W \cdot Win_{ng} \cdot BlackQB_g + \pi^L \cdot Loss_{ng} \cdot BlackQB_g\right], \tag{3}$$

[2]While other mechanisms may operate—such as status threat, whereby dominant-group members respond with increased out-group animosity when their relative position is challenged (Blumer (1958), Blalock et al. (1967))—our empirical findings do not consistently align with status threat predictions, and are largest following emotionally intense games, pointing to affect shocks as the primary driver.

[3]We follow prior work on sporting events which uses game characteristics such as upsets and point spreads to identify emotionally salient games (Card and Dahl (2011); Eren and Mocan (2018); Healy et al. (2010)).

where $BlackQB_g = 1$ if either quarterback for game g was Black. Of interest are the π parameters, where π^L (π^W) captures the degree to which negative (positive) affect is attributed to a Black quarterback and translates into increased (decreased) anti-Black animus.

We model the translation of affect into anti-Black animus with deliberate flexibility, as the direction and magnitude of this translation are theoretically ambiguous and may depend on specific conditions. First, the translation of affect to animus may not be symmetric for positive and negative shocks. A consistent finding across experimental and field settings is that majority group members impose harsher penalties on out-group individuals for equivalent failures; withhold credit for positive outcomes; and disproportionately attribute negative outcomes to out-groups — even when those individuals are not causally responsible (Pettigrew (1979); Sarsons (2017a); Sarsons (2017b); Egan et al. (2022); Bursztyn et al. (2022); Bauer et al. (2023)). As a result, it is plausible that the animus-increasing effect of losses involving Black quarterbacks is larger than the animus-reducing effect of wins due to this attribution asymmetry ($\pi^L > \pi^W$).

Second, even conditional on a win or loss, effects may further differ for local and opposing Black quarterbacks. The $\pi^{L,W}$ terms capture the average effect of having a loss or win involving any Black quarterback, regardless of whether they are playing for the local or the opposition team. The model can be refined by setting $\pi^W = \pi_l^W \, l_{ng} + \pi_o^W \, o_{ng}$ and $\pi^L = \pi_l^L \, l_{ng} + \pi_o^L \, o_{ng}$, where l_{ng} and o_{ng} indicate a Black quarterback on the local and opposing team respectively. This allows the modeling of separate parameters for local and opposition quarterbacks. Social identity theory suggests that fan reactions to players are shaped by group membership and allegiance (Tajfel and Turner 1979; Wann and Branscombe 1993). Local quarterbacks may attract more discussion and blame following a loss, yet shared identity, familiarity, and repeated exposure may attenuate negative reactions and reduce the salience of race. Opposing quarterbacks may face less direct blame, yet their lack of in-group status could leave them more susceptible to othering and demonization, with race potentially more salient (Wann et al. 2003). Their relative effects on racial animus, captured by π_l^L and π_o^L (and likewise π_l^W and π_o^W), are theoretically ambiguous. The relative magnitude of these terms, and how they might vary with quarterback performance, is therefore evaluated empirically.

Game-day predictions: *Relative to games with two white quarterbacks, games involving a Black quarterback may generate area-level racial animus inversely related to the affect shock generated by the game. These shocks will be greater in high engagement areas and for high intensity games. If attribution asymmetry holds, the magnitude of changes following losses will be greater than after wins.*

- $\frac{\partial r}{\partial Win} \leq 0$ and $\frac{\partial r}{\partial Loss} \geq 0$;
- $\frac{\partial |r|}{\partial Engagement} > 0$;
- $\frac{\partial |r|}{\partial Intensity} > 0$;
- If attribution asymmetry holds, $\pi^L > \pi^W$ such that $\frac{\partial |r|}{\partial Win} \leq \frac{\partial |r|}{\partial Loss}$.

3.2 *Performance-contingent updating*: Season-level adjustments to baseline animus

Over the course of a season, the emotion-driven affect shocks that characterize short-run game-day responses give way to slower updating of baseline animus $\bar{A}_{as}$. Parasocial contact theory (Schiappa et al. 2005) predicts that repeated media-based exposure to positively characterized out-group members reduces prejudice, with evidence from soccer documenting animus-reducing effects of exposure to a high-performing Muslim player (Alrababa'h et al. 2021). The quarterback position carries deep cultural associations with intelligence, leadership, and decision-making—traits that have historically been the basis for excluding Black athletes from the role (Rhoden 2007). Repeated exposure to Black quarterbacks occupying this counter-stereotypical position may therefore shift racial attitudes by altering beliefs about competence, leadership, and social norms, with Black quarterback performance acting as a salient diagnostic signal for updating beliefs about Black competence and leadership.

We model baseline animus as evolving between seasons as the population partially adjusts its baseline perceptions, $\bar{\theta}_{ns}$, toward season-specific signals received from local and opposition Black quarterbacks. Season-teams with local Black quarterbacks ($L_{ns} = 1$) receive signal θ_L which is an increasing function of the team's season performance record (R_{ns}) such that $\theta'_L(R_{ns}) > 0$. Season-teams that engage with opposition Black quarterbacks ($E_{ns} > 0$) receive a signal θ_O which may be a function of the outcome of these engagements (Ω_{ns}). All of these processes are scaled by area engagement such that

$$E[\bar{A}_{n,s+1} - \bar{A}_{ns}] = (-1) \cdot Engagement_{an} \cdot \left[\lambda_L L_{ns}(\theta_L(R_{ns}) - \bar{\theta}_{ns}) + \lambda_O E_{ns}(\theta_O(\Omega_{ns}) - \bar{\theta}_{ns})\right], \quad (4)$$

with $\lambda_L, \lambda_O \geq 0$ governing updating rates.

Local Black quarterbacks can either increase or decrease animus, depending on the sign of $(\theta^L(R_{ns}) - \bar{\theta}_{ns})$. Animus-reducing positive updating occurs when $\theta^L(R_{ns}) > \bar{\theta}_{ns}$. Because better

team performance raises $\theta^L(R_{ns})$, positive updating is more likely after high-record seasons. The direction of updating after low-record seasons is ambiguous: the mere presence of a Black quarterback in a counter-stereotypical position may generate a positive signal, but poor team performance may outweigh it. Which effect dominates after low-record seasons is ultimately an empirical question.

Even if a team does not have a Black quarterback, exposure to opposing Black quarterbacks ($E_{ns} > 0$) could generate season-level adjustments to baseline animus. The direction of the updating signal received from this exposure, $(\theta_O(\Omega_{ns}) - \bar{\theta}_{ns})$, is ambiguous and may depend on the outcomes of those interactions (Ω_{ns}). While opposition Black quarterbacks increase exposure to counter-stereotypical Black leadership, this exposure is less frequent and occurs in a markedly different context: media portrayals are less likely to be favorable, and the competitive setting lacks a shared local identity, potentially heightening social boundaries and reinforcing group distinctions rather than attenuating them. Whether exposure to opposition quarterbacks reduces or reinforces animus remains an empirical question.

Season predictions: *Local Black quarterbacks on high-record teams reduce racial animus, with smaller reductions–or even increases– for low-record teams. Effects are larger in high-engagement areas.*

- $(\theta_L(R_{ns} = \textit{high}) - \bar{\theta}_{ns}) \geq 0$ and $(\theta_L(R_{ns} = \textit{high}) - \bar{\theta}_{ns}) \geq (\theta_L(R_{ns} = \textit{low}) - \bar{\theta}_{ns})$;
- $\frac{\partial |\Delta \bar{A}|}{\partial Engagement} > 0$.

4 Data

4.1 Measuring exposure to Black quarterbacks

To measure exposure to Black quarterbacks, we proceed in two steps. First, we delineate each NFL team's geographic catchment area. Second, using play-by-play game data, we identify the games and seasons in which a Black quarterback served as the team's primary quarterback.

Defining NFL teams' geographic catchment areas: We match U.S. Designated Market Areas (DMAs), which delineate local media markets, to NFL teams by estimating DMAs' engagement levels with each of the 32 teams.[4] We combine local DMA ratings, $HomeRatings^{team}$, which

[4]Full DMA-level ratings data are not publicly available. We can estimate each team's home market ratings,

capture how popular a team is where its fans are most heavily concentrated, with a Google search index for the team, $Google_{DMA}^{team}$, that measures each DMA's interest in the team over the 2004-2025 period relative to the top-searching– and likely local– DMA.[5] For each DMA-team combination we calculate

$$Engagement_{DMA}^{team} = HomeRatings^{team} \times Google_{DMA}^{team}.$$

Each DMA is then assigned to the team it is most engaged with. The resulting catchment areas are mapped in Figure 2. Figure A1 plots the distribution of engagement levels across DMAs.

NFL data: Using public APIs [6], we collect play-by-play data for every NFL game in the 2002 through 2024 seasons, recording every on-field action for every player. Although only one quarterback can be on the field at a time, teams can field multiple quarterbacks within a game if their starting quarterback is injured or removed from play. We define a team-game's lead quarterback as the quarterback with the greatest share of plays. A quarterback's race is then identified by a team of research assistants.[7] For our season-level analysis, a team-season is classified as having a Black quarterback if the majority of the team's regular season games are classified as having a Black quarterback.

These data cover 6214 games and 736 season-teams.[8] We exclude 345 games and 23 season-teams in which the lead quarterback is one of the two non-Black Hispanic or two Samoan quarterbacks in our data, so as to focus on contrasting reactions to European-American and African-American quarterbacks. For our short-run estimates, which examine outcomes in five-day windows centered on game days, we further restrict the sample to team-games occurring at least seven days after the team's previous game and three days before the subsequent one to avoid contamination

$HomeRatings^{team}$, as the average of local ratings reported for five seasons (2011, 2012, 2013, 2020, 2021) in the Sports Business Journal (Karp (2014); Ourand (2022)).

[5] We use DMA-level Google Trends indexes for each team over the 2004–2025 period. These are normalized to 1 for the highest-searching DMA so that the index for all other DMAs is interpretable as the search intensity relative to that market. Since Google Trends indexes are calculated from a varying sample–rather than the universe–of searches, we collect 8 separate requisitions and average across them to construct a stable measure. This average correlates at 1.00 with the average of the first 7 requisitions, confirming stability.

[6] We use nfl_data_py, a Python library for interacting with NFL data sourced from nflfastR, nfldata, dynastyprocess, and Draft Scout.

[7] Players are coded as Black if they are visibly perceived as Black or mixed-race, consistent with a constructivist perspective that highlights social perceptions of race as the relevant construct for understanding discriminatory responses to race (Rose (2023)).

[8] Each of the 32 teams played 16 regular season games between 2002 and 2020, and 17 between 2021 and 2024. Playoffs consisted of 11 games between 2002 and 2019 and 13 games between 2020 and 2024. We drop the 2022 week 17 game between the Buffalo Bills and the Cincinnati Bengals that was canceled early in the first quarter due to Bills player Damar Hamlin's medical emergency.

from exposure to other games in close temporal proximity. Finally, our main game-level sample omits 307 games featuring two Black quarterbacks, which are few in number and confound multiple effects, complicating interpretation, though our results are robust to their inclusion. Our final game-level sample covers 5562 games: 3502 with two white quarterbacks and 2060 with a Black quarterback on the field. At the season level, 152 of 713 season-teams are classified as having a Black quarterback.

4.2 Measures of racial attitudes and hate

Our empirical strategy examines impacts on four complementary measures of racial attitudes and hate: anti-Black hate crimes, reflecting overt offline criminal behavior; Reddit moderation rates and hate speech, capturing online hate speech within NFL fan communities; IAT scores, measuring implicit racial bias; and the Racial Animosity Index, capturing private searches for racial slurs on Google. These measures differ in severity, visibility, and the costs of expression, spanning implicit attitudes, private search behavior, public online expression, and overt criminal action, allowing us to assess how impacts generalize across different manifestations of racial animus.

Anti-Black hate crimes: We use the FBI's Uniform Crime Reporting Hate Crime Statistics for 2002–2025.[9] These data are the most comprehensive standardized police-reported measure of hate-crime incidents available in the United States and are widely used in hate-crime research (Müller and Schwarz 2023; Anderson et al. 2020; Mulholland 2013). Nevertheless, it is important to note that reported incidents documented in these data are only a small share of all hate crimes. Many hate crimes are never reported to law enforcement, and many local law enforcement agencies do not reliably report hate crimes to the FBI as reporting is voluntary (Dharmapala and Huq (2024)).[10] The incident-level files report the calendar date and location of each incident, a bias-motivation category identifying the targeted group, and the reporting agency. We map incidents to counties using the locations of over 24,000 reporting agencies. We then calculate the count of anti-Black incidents on each day for each DMA to generate the daily DMA incident rate (per 10 million). For our season-level analysis, day counts are aggregated to generate monthly DMA incident rates (per

[9] In these data, a hate crime is defined as a criminal offense motivated, in whole or in part, by the offender's bias against a protected characteristic such as race.

[10] Between 2011 and 2015, 54 percent of hate crimes identified in the National Crime Victimization Survey were not reported to the police (Dharmapala and Huq (2024)).

100,000). Throughout, we check that our main findings are robust to alternative measurements and specifications of hate crimes.[11]

Reddit data: We introduce the share of submissions removed by content moderators as a novel outcome measure and our primary indicator of engagement in online hate speech. The existing literature predominantly treats moderation as a platform intervention to be evaluated (Chandrasekharan et al. 2017, 2022), or studies the effects of online hate speech that remains publicly observable (Müller and Schwarz 2021b, 2023). In contrast, we use the moderation rate as an outcome that captures online expressions of animus severe enough to trigger removal. Our Reddit data begins in 2015 as this is when our corpus of Reddit submissions starts to include flags for submissions that were moderated or otherwise removed from the platform. Our sample is composed of 32 team-specific subreddits — online communities dedicated to discussion of each NFL team. For each subreddit-day we calculate the daily count of submissions and moderated submissions.[12] While moderation also captures other violations such as spam and off-topic posts, NFL team subreddits explicitly prohibit hate speech and harassment. Importantly, we show below that moderation rates respond systematically to quarterback race, supporting its validity as a measure of hateful expression and sentiment. In addition to these platform-wide policies, NFL team subreddits enforce their own community-specific content rules, which commonly prohibit hate speech and harassment. It is worthwhile to note that Reddit does not systematically report the substantive reason for each removed submission. We therefore interpret the share of removed submissions as a moderation-based proxy for hateful speech, rather than as a direct measure of racial hate speech.

For submissions with visible content—those neither removed by moderators nor deleted by authors—we construct a direct text-based measure of anti-Black hate speech. We begin with HATEBERTA, a transformer-based classifier trained to detect abusive content on Reddit (Caselli

[11]Since hate crimes are rare and many cells contain zero incidents, we also estimate common alternatives in the hate-crime literature: an extensive-margin indicator for any incident, a zero-adjusted log specification, linear and Poisson specifications on incident counts and the inclusion of population weights. These specifications follow prior work that uses log transformations of hate-crime counts (Müller and Schwarz 2023), extensive-margin outcomes for any hate-crime incident (Müller and Schwarz 2021a), Poisson and count specifications for hate-crime outcomes (Anderson et al. 2020), and population-scaled hate-crime rates (Pettis et al. 2022).

[12]We focus on Reddit submissions rather than comments for several reasons. First and foremost, comments associated with submissions that were moderated off of the platform are unobserved, resulting in a selected sample. Second, submissions are more likely to reflect deliberate and sustained expression, whereas comments frequently capture rapid reactions and conversational back-and-forth, often consisting of exclamations or fragmentary expressions with limited semantic information. This substantially increases measurement error when estimating hate-speech intensity using text-based classifiers.

et al. 2020).[13] We further fine-tune the model using ToxiGen, a large-scale dataset of targeted hate speech (Hartvigsen et al. 2022), to better capture anti-Black hate speech specifically. For each visible submission, the model produces a score between zero and one, which is interpreted as the predicted likelihood that the submission contains anti-Black hate speech. Anti-Black hate speech in submissions that survive moderation is rare: the mean estimated probability is 0.019% and only 0.04% of submissions score above zero. This is consistent with prior work finding similarly low levels of hate speech among soccer fans (Alrababa'h et al. (2024)). Given this limited signal, we treat the moderation share as our primary indicator and report estimates using the ToxiGen score in the appendix. After matching the Reddit data with NFL games, our final sample spans the 2015 to 2024 seasons, covering 4002 unique game-teams and 244 season-teams.

Implicit association test (IAT) scores: Measures of implicit biases come from the publicly available data on implicit association tests collected online by *Project Implicit*. *Project Implicit* allows participants to complete implicit association tests (IATs), generating a measure of their implicit associations with individual characteristics such as race. IATs are widely used in psychology, and increasingly in economics, as a way of measuring implicit attitudes (Glover et al. (2017), Carlana (2019), Corno et al. (2022), Lowes et al. (2015)).[14] After matching micro-level IAT survey responses to DMAs and NFL games, our final sample spans the 2002 to 2024 seasons, covering 6019 unique game-teams and 618 season-teams. Since participation in *Project Implicit* is voluntary and self-selected, we also examine whether the number of daily and monthly DMA responses varies with quarterback race, finding no evidence of a selection effect.

Racial Animosity Index (RAI) using Google search indices: We construct a location- and time-varying measure of anti-Black racial animus using the index of Google searches for racial slur

[13] HATEBERTA is trained on a broad set of hate-speech categories, including speech targeting racial and ethnic groups such as Black, Asian, and Latino communities. HATEBERTA was developed and tuned on 2010s social-media text. Analysis with submissions since 2015 mitigates distribution shift between the training corpora and our corpus.

[14] When completing the race IAT, respondents are sequentially presented with images of Black and white individuals, and words that have positive or negative connotations. Respondents complete a series of rapid sorting exercises grouping together the images with words. The IAT is meant to measure the strength of a respondent's association between individual characteristics and word connotations, as sorting is easier when associated items are sorted together. IAT measures have been the subject of increased scrutiny in the psychology literature (Ratliff and Smith (2021)). There is general agreement that these measures are relevant and predictive, particularly for socially sensitive topics (Greenwald et al. (2009), Bertrand and Duflo (2017)), and aggregate regional measures (Hehman et al. (2019)). Disagreements have centered on the validity of the *implicit* bias construct, whether it differs from *explicit* biases, and whether the low test-retest reliability is due to high measurement error, or the construct itself being time-varying (Gawronski (2019), Schimmack (2021), Connor and Evers (2020)).

queries.[15] This measure, first developed by Stephens-Davidowitz (2014), has been widely used in the literature as a measure of racial animosity (Chetty et al. (2020), Kline et al. (2022), Derenoncourt (2022)). The slur query is most frequently searched for alongside the term "joke(s)", which returns offensive websites. Searches that include the slur query are not uncommon — occurring at rates comparable to common everyday search terms.[16]

We follow the approach used in Anderson et al. (2020) so that our regression coefficients can be interpreted as a percentage change in the racist search rate. To capture geographic and temporal variation in slur searches while retaining enough search volume to construct a reliable index, we calculate the index for each DMA in two-month bins. This is calculated as the average of the natural log of ten separate Google Trends samples to reduce sampling noise.[17] After matching this index to NFL seasons, our final sample comprises 613 team-seasons spanning the 2004–2023 seasons, and includes only DMA-seasons for which the index is observed throughout the season. This outcome can only be used in season-level estimations.

Table 1 provides descriptive statistics for our main outcomes measured at both the day (panel (a)) and season (panel (b)) levels.

5 Short-run Responses to Games: Racial Retaliation

In this section, we examine the short-run impact of exposure to Black quarterbacks on racial animus around game days. A well-established literature documents how the emotional intensity of games impacts behavior in the days that follow (Card and Dahl 2011; Montolio and Planells-Struse 2016; Ivandić et al. 2024; Matsuzawa and Arnesen 2024). We build on this literature, finding evidence of a racial retaliation mechanism: beyond the direct effects of the affect shocks generated by wins and losses, racial animus increases differentially following losses in games featuring a Black quarterback—whether local or opposing—consistent with negative affect being attributed to the salient Black player and spilling over into broader anti-Black animus. This effect is robust

[15] Specifically, we requisition indexes for searches that include the terms "n___ +n___s" — henceforth "the slur query."

[16] For example, from 2004 to 2025, such searches are about as common as searches for the terms "cavity+cavities".

[17] For each two-month requisition, the Google index is set to 1 for the DMA with the greatest share of slur queries, so that the index for all other DMAs is interpretable as search intensity relative to that DMA. Since Google Trends indexes are calculated from a varying sample — rather than the universe — of Google searches, we collect ten requisitions per two-month bin and average across them to construct a stable measure. As Google does not return an index when samples are too small, we drop any mean calculated with fewer than two observed requisitions (25.4% of DMA-2month bins, mostly in the earliest years of the data and low-population DMAs).

across all our animus measures, and is more pronounced after emotionally intense games and in high-engagement DMAs. We describe our empirical strategy and present these findings below.

5.1 Short-run empirical strategy

Our short-run design estimates how outcomes change in the days that immediately follow a game based on the racial composition of the quarterbacks that played, the game's result, as well as the game's emotional intensity. We classify game g for NFL team n as having a Black quarterback if a Black player played that position for the majority of the game's plays. For each game-team we define the game's quarterback composition from each team's perspective. Our main specification compares the effect of games where any quarterback (opposition or local) is Black ($AnyBlack = 1$), relative to games where both quarterbacks are white ($AnyBlack = 0$).[18] Heterogeneity in effects for opposition and local Black quarterbacks is examined subsequently in section 5.3.

We evaluate how outcomes change in five-day windows centered around game days ($d = 0$) using a pre-post specification where $Post_{dg}$ is set to 1 for the game day and the two subsequent days, and 0 for the two days prior. $Post_{dg}$ is interacted with $AnyBlack_g$ and whether the game was won (Win_{gn}).[19] We estimate

$$\begin{aligned} Y_{agnd} =& \rho_0 + \rho_1^w Post_{dg} + \rho_2^b Post_{dg} \times AnyBlack_g \\ &+ \rho_3^w Post_{dg} \times Win_{gn} + \rho_4^b Post_{dg} \times AnyBlack_g \times Win_{gn} \\ &+ \eta_{ag} + \omega_d + \epsilon_{agnd}, \end{aligned} \tag{5}$$

where Y_{angd} is an outcome for media market a, team n, on day d, around game g. A day of the week fixed effect, ω_d, controls for weekly trends in outcomes. Importantly, we include a $DMA \times game$ fixed effect, η_{ag}, which restricts identifying variation to outcomes measured around the same game in the same media market (and thus the same team).[20] This fixed effect absorbs baseline racial animus which we treat as fixed in the short-run.[21]

With this specification, ρ_2^b identifies how the change in outcomes after games differs for lost

[18] For clarity and brevity of analysis we omit games where both quarterbacks are Black. Two-Black-quarterback games are less common preventing precise estimates of their differential impacts. Including them in the analysis does not substantially alter our results.

[19] Ties are uncommon in the NFL during this period. In our game data, only 14 of 6214 games resulted in a tie. These are coded as $Win_{gn} = 0$.

[20] Geographic identifiers are not available for Reddit outcomes, for which we use η_{ng}, a team-game fixed effect.

[21] For figures we estimate an equivalent event time specification where coefficients are estimated for each day relative to the day before the game.

games involving a Black quarterback, as compared to losses where both quarterbacks were white. ($\rho_2^b + \rho_4^b$) gives this differential for wins. These are the reduced-form effects of the racial retaliation mechanism modeled in section 3.1, $\frac{\partial r}{\partial Loss}$ (and $\frac{\partial r}{\partial Win}$), with the model predicting $\rho_2^b \geq 0$ for losses (and $\rho_2^b + \rho_4^b \leq 0$ for wins).

We estimate equation 5 for all games. We also estimate effects separately for games associated with three different levels of emotional salience to verify that the magnitude of the effects increases with game intensity ($\frac{\partial |r|}{\partial Intensity} > 0$), consistent with affect shocks driving the racial retaliation mechanism. We identify emotionally intense "close-upset" games where (i) the final score differential was close at 7 or less (i.e. one touchdown determined the game's outcome), and (ii) the outcome was an upset as the pre-game spread favored the losing team by 3 points or more. These are compared to the effects around moderately intense games (that were either close, or upset games), and non-intense games (that were neither close, nor upset games). Finally, we also estimate equation 5 by quartile of team engagement to test the prediction that the magnitude of the effects is greater in DMAs that are more engaged with the NFL team ($\frac{\partial |r|}{\partial Engagement} > 0$).

5.2 Short-run effects of Black quarterbacks: increased racial animus after losses

Anti-Black hate crimes: Table 2, rows 3 and 4, reports our main short-run estimates for local Black quarterback games on DMAs' daily hate crime incident rates (per 10 million).

Column 1 reports estimates of equation 5 on all games in our analysis sample. Consistent with our model predictions, the anti-Black hate crime rate differentially increases by 0.036 (p-value=0.005), an 18% increase, following losses involving a Black quarterback, relative to losses where both quarterbacks were white — which have no significant impact on hate crimes (row 1). The differential effect following wins ($\rho_2^b + \rho_4^b$) is much smaller at 0.016, and we cannot reject the null of no effect (p-value=0.20).

As predicted by the model, effects increase with game intensity. Effects are driven by emotionally intense close-upset games. This confirms affect as central to the mechanism generating these effects. Column 2 restricts the sample to close-upset games, finding substantially larger effects. After a close-upset loss involving a Black quarterback, the anti-Black hate crime rate differentially increases by 0.086 (p-value=0.008), a 43% increase. In contrast, the ($\rho_2^b + \rho_4^b$) point estimate of -0.018 suggests a reduction in the hate crime rate following wins involving Black quarterbacks, though this is not statistically significant (p-value=0.60). Figure 3 presents event-day estimates for close-upset games. Panel (b) plots daily estimates for white-white close-upset losses and wins,

showing no detectable impact. Panel (a) plots the differential effects for close-upset losses and wins involving Black quarterbacks, showing large effects that begin on game day and persist for two days. Columns 3 and 4 split the sample into moderately intense games (close scoring or upset outcomes, column 3), and low-intensity games (neither close scoring nor upsets, column 4). The point estimate for losses in low-intensity games is positive but small and statistically insignificant at 0.019 (p-value=0.34), confirming that effects are concentrated in emotionally salient games. A 43% increase after close-upset losses involving a Black quarterback is large but not implausible. Previous work shows that specific events can generate substantial spikes in hate crimes. Trump's anti-Muslim and "Chinese virus" tweets have been linked to very large increases in hate incidents, orders of magnitude larger than those we observe (Müller and Schwarz 2023; Cao et al. 2023). Closer to our magnitudes are studies documenting increases in hate crime following the Brexit referendum, US elections, and terror attacks, with estimated increases ranging from 15 to 30% (Carr et al. 2020; Guha 2024; Hanes and Machin 2014).

Consistent with model predictions, the effects are also larger in DMAs that are more engaged with their NFL team. Figure 4(a) plots the post-game effect for close-upset losses involving a Black quarterback (ρ_2^b) by quartile of DMA levels of team engagement. Point estimates increase with engagement, from 0.055 in low engagement DMAs, to 0.118 in high engagement DMAs.

These patterns—large increases after emotionally intense losses involving Black quarterbacks, with little to no change following other games, whether wins or low-intensity games—are robust to alternative specifications of our measure of hate crimes and fixed effects combinations (see appendix tables A1 and A2).

To better understand the nature of these increases in anti-Black hate crimes, Table A3 examines impacts by crime type. Effects are driven by the most common categories — crimes against persons and property—and by crimes committed by a single or unknown number of perpetrators. The absence of effects on multi-perpetrator crimes suggests that game exposure triggers isolated, emotion-driven responses rather than socially coordinated hate (Glaeser 2005). Appendix Table A4 examines hate targeting other racial groups. Consistent with a racial retaliation mechanism, we find no evidence of game-day changes in hate crimes targeting white or other racial groups. Hate crimes targeting Hispanics show some responsiveness to Black quarterback presence, though effects are smaller—plausibly reflecting the racial heterogeneity of the Hispanic population, which includes individuals who identify as or present as Black.

Hate speech and IAT scores: Table 3 repeats this analysis for our other day-level measures of racial animus: the share of Reddit submissions moderated off of NFL team subreddits, and white respondents' IAT scores. Both outcomes display patterns consistent with our model predictions and strikingly similar to those observed for hate crimes: differential increases following close-upset losses involving Black quarterbacks (row 3 of columns 2 and 6), little to no differential effect following wins, with effects increasing in both game intensity and levels of DMA engagement with the NFL team.

For submission removals, moderation rates show a general increase after all games (rows 1 and 2), reflecting heightened activity after game days. This increase is substantially larger following emotionally intense close-upset losses involving a Black quarterback: the moderation share differentially increases by 0.045 (p-value=0.010), a 29% increase, relative to close-upset losses where both quarterbacks were white. We find no differential change following close-upset wins (0.006; p-value=0.65) or around low-intensity games (column 4). Event-day estimates are illustrated in Figure 3: panel (d) shows the general increase in moderation share around game days, while panel (c) shows the large differential increase following losses involving Black quarterbacks. Appendix Table A5 and Figure A2(c) examine whether the number of submissions is differentially affected by quarterback race. As expected, submission activity on team subreddits increases after games, but we find no evidence of a differential effect around Black quarterback games. Figure A2(a) shows no differential change in the predicted probability of anti-Black hate speech among submissions that remain visible after moderation, suggesting that moderation policies likely removed the most problematic content.

IAT scores show a similar pattern, though estimates are somewhat underpowered. White respondents' implicit bias scores differentially increase by 0.043 SDs (p-value=0.06) following emotionally intense close-upset losses involving a Black quarterback, with no statistically significant differential change following wins (-0.031 SDs; p-value=0.21) or after low-intensity games (column 8). Event-day estimates are shown in the last two panels of Figure 3. Figure 4(b) shows that effects are concentrated in above-median engagement DMAs, with little to no effect in below-median engagement DMAs. Appendix Table A5 and Figure A2 confirm no differential change in the number of white IAT test takers around game day, suggesting selection into testing is not a confound.

Overall, these findings are consistent with the game-day predictions outlined in Section 3.1. The model predicts that losses involving Black quarterbacks may increase racial animus ($\frac{\partial r}{\partial Loss} \geq 0$),

and that wins may reduce it ($\frac{\partial r}{\partial Win} \leq 0$), with attribution asymmetry predicting a larger effect for losses ($\frac{\partial |r|}{\partial Loss} > \frac{\partial |r|}{\partial Win}$ as $\pi^L > \pi^W$). Both effects are predicted to be larger following intense games ($\frac{\partial |r|}{\partial Intensity} > 0$) and in DMAs more engaged with their NFL team ($\frac{\partial |r|}{\partial Engagement} > 0$). We find strong support for the loss prediction: racial animus differentially increases after losses involving Black quarterbacks ($\frac{\partial r}{\partial Loss} > 0$); consistent with these losses being attributed to Black quarterbacks and thereby generating anti-Black animus ($\pi^L > 0$). Anti-Black racial animus increases after Black quarterback losses across all three measures—hate crimes, Reddit moderation, and IAT scores—suggesting the effect reflects a genuine increase in racial animus rather than an artifact of any single outcome. Consistent with model predictions, effects are concentrated after emotionally intense close-upset games and in higher-engagement DMAs, confirming the centrality of emotion as the key driver of these short-run effects. In contrast, we find minimal evidence of impacts following wins and cannot reject the null of no effect ($\frac{\partial r}{\partial Win} = 0$), implying that $\pi^W = 0$. Positive affect shocks are not attributed to Black quarterbacks in a way that changes anti-Black animus.

This asymmetry ($\pi^L > 0$; $\pi^W = 0$) is consistent with evidence on differential out-group blame attribution, whereby out-group individuals are disproportionately blamed for negative outcomes but receive less credit for positive ones (Sarsons 2017a,b; Egan et al. 2022). Our results extend this literature in two ways: first, while negative spillovers into broader racial animus have been documented experimentally (Bursztyn et al. 2022; Bauer et al. 2023), we show that these attribution patterns generate real-world anti-Black animus toward people who had no role in the outcome; and second, we show evidence of the loss-win asymmetry in a natural setting at scale.

5.3 A reaction to both local and opposition Black quarterbacks

Estimates in section 5.2 show attribution operating on average after losses ($\pi^L > 0$) but not after wins ($\pi^W = 0$). These π parameters pool both the local and opposition attribution parameters — π_l^L and π_o^L — for losses (π_l^W and π_o^W for wins) to estimate the effect of losses and wins involving any Black quarterback. As discussed in section 3.1, team identity alignment with the Black quarterback could substantially impact these effects in theoretically ambiguous ways, generating significant heterogeneity based on whether the Black quarterback was playing for the local or the opposition team. To examine this, we estimate equation 5 disaggregating ρ_2^b and ρ_4^b to separately estimate $\rho_2^o, \rho_2^l, \rho_4^o$ and ρ_4^l to examine how attribution and blame may differ for wins and losses associated with local (π_l^W, π_l^L) and opposition (π_o^W, π_o^L) Black quarterbacks.

Estimates are reported in Table 4. Across all three outcomes we observe similar patterns. Racial

animus indicators increase following engaging losses but not after engaging wins involving Black quarterbacks, regardless of whether the Black quarterback is playing for the local or opposition team. Though there are some differences in the point estimate magnitudes for local (ρ_2^l) and opposition (ρ_2^o) Black quarterbacks, both estimates are positive. For all three outcomes, we fail to reject that ρ_2^l and ρ_2^o are equal. Similarly, both the ρ_4^o and ρ_4^l estimates are negative, suggesting offsetting effects following wins. For all three outcomes, we fail to reject that $\rho_2^l + \rho_4^l = \rho_2^o + \rho_4^o$. Overall, there is no clear evidence of effect heterogeneity based on whether the Black quarterback is playing for the local or opposition team. This is consistent with a racial retaliation effect where either the local or opposition Black quarterback is "blamable" for the negative affect shock.

5.4 Quarterback performance and racial blame

While fan utility ultimately depends on winning, blame attribution may be responsive to quarterback performance. In this section, we evaluate whether racial animus increases after losses even when the local Black quarterback played well—a pattern that would indicate not only that blame spills over to the broader population, but that the blame itself is misattributed to begin with.

For each game we calculate both quarterbacks' game Expected Points Added (EPA), a common measure of quarterback performance. High performance indicators for the local and opposition quarterbacks are set to 1 for game EPAs above 0.2.[22] To allow the effect of local and opposition quarterback performance to differ, we estimate two modified versions of equation 5 for each outcome, interacting the treatment variables with the local (opposition) performance indicator using games where both quarterbacks were white and where only the local (opposition) quarterback was Black. Results are reported in Table 5.

Heterogeneity estimates suggest that hate crimes and IAT scores are broadly unresponsive to quarterback performance. These outcomes differentially increase after losses involving Black quarterbacks, regardless of whether they played well. In columns 1-4, we cannot reject the null of no differential effect for high-EPA Black quarterbacks. Reddit submission removals tell a different story: differential increases in moderation rates are observed only after low-EPA losses, with no differential increase following high-EPA losses or wins involving Black quarterbacks. This pattern suggests high performance mitigates racial blaming among the highly-engaged NFL fans active on NFL team subreddits—a population better equipped to evaluate quarterback performance. In

[22] About one third of quarterback-games have an EPA above 0.2.

contrast, measures capturing animus in less-NFL focused populations, who may be less attuned to individual player skill, show no such mitigation effect.

5.5 An effect independent of differences in play styles

African-American quarterbacks have historically been more likely to be described as "dual-threat" quarterbacks who combine running and passing ability, in contrast to the traditional "pro-style pocket passer" archetype associated with less mobile styles of play. To check whether the effects reflect quarterbacks' playing style rather than race, we construct a measure of play style and add it as a control. To avoid externally-imposed labels such as "dual-threat" or "pro-style," which may themselves reflect racialized evaluations, we rely on recorded on-field behavior. Quarterback-games are classified as mobility-integrated ($mobQB = 1$) when quarterbacks meaningfully contribute as runners, and as pocket-oriented otherwise ($mobQB = 0$).[23] Using this measure, 50% of games played by Black quarterbacks are mobility-integrated, compared to 13% for white quarterbacks. Table A6 reports estimates of equation 5 with controls interacting this indicator with $Post$ and Win. Across all three outcomes, controlling for play style does not meaningfully alter the estimated effects after close-upset losses involving Black quarterbacks: the coefficients on the Black quarterback interaction terms (rows 3 and 4) are largely unchanged, and the play style controls themselves are not statistically significant. There is no evidence that our measures of racial animus respond to quarterback play style.

6 Season responses: Performance-sensitive updating

The patterns described in Section 5 show that anti-Black racial animus increases for a few days after emotionally intense losses involving a Black quarterback. Though large in magnitude on affected days, the broader impact of these short-run effects is limited: only 21% of games are losses involving a Black quarterback, and only 3.46% involve the close-upset losses that drive the largest effects. In this section we turn to examining what are likely more consequential and socially

[23] A quarterback-game is coded as mobility-integrated if the quarterback records at least three mobility plays and if mobility plays account for at least 10 percent of quarterback style plays. Mobility plays include scrambles and non-scramble quarterback rushes, excluding kneels and spikes. Quarterback style plays are defined as dropbacks plus designed quarterback runs. All remaining quarterback-plays of sufficient duration are classified as pocket-oriented.

important effects: the persistent effects of repeated seasonal exposure to Black quarterbacks. We describe our empirical strategy and present these findings below.

6.1 Local Black quarterbacks in the long-run

6.1.1 Empirical strategy

Our long-run design estimates how racial animus evolves over the NFL season as a function of quarterback race and team performance. The core comparison is between areas served by Black quarterback-led teams and those served by white quarterback-led teams, holding team performance constant.

For the time period under consideration, the NFL regular season begins in early September and ends in early January. This is followed by playoff games in January, culminating in the Super Bowl in early February. We define a season as running from May through April to ensure our data window captures outcomes both before and after NFL activity. For tables, we divide seasons into three time periods T such that $T \in \{b, d, a\}$. The *before* season, b (May-August), is our omitted baseline period t_0. This is compared to outcomes observed *during* the season, d (September - February); and in the *after* season, a (March-April) – our period of interest to observe persistent impacts. For figures, we use six two-month bins – $T \in \{mj, ja, so, nd, jf, ma\}$– to trace the evolution of outcomes more granularly across the season, with July–August as t_0, our omitted baseline period.

We define a team n in season s as having a local Black quarterback ($Black_{ns} = 1$) if the majority of the team's regular season games are classified as being played by a Black quarterback. We categorize each team-season into one of three performance tiers, $R_{ns} \in \{r, p, c\}$, based on playoff advancement: regular season only, our omitted baseline ($Regular_{ns}$ – 450 team-seasons); playoff but not championship ($PlayOnly_{ns}$ – 194 team-seasons); and championship-level teams ($Champ_{ns}$ – 92 team-seasons).

Our main specification estimates the differential effect of having a Black quarterback as compared to a white quarterback for teams performing at the same level. We estimate

$$
\begin{aligned}
Y_{ansm} = \mu_0 &+ \sum_{T \neq t_0} \mu_T^r Black_{ns} \times \mathbb{1}\{m = T\} \\
&+ \sum_{T \neq t_0} \mu_T^p Black_{ns} \times PlayOnly_{ns} \times \mathbb{1}\{m = T\} \\
&+ \sum_{T \neq t_0} \mu_T^c Black_{ns} \times Champ_{ns} \times \mathbb{1}\{m = T\} \\
&+ \tau_{as} + \kappa_{ma} + \psi_{mR} + \phi_{ms} + \epsilon_a nsm.
\end{aligned} \tag{6}
$$

where Y_{ansm} is an outcome for DMA a, team n, observed in month m of season s. For teams that only play the regular season, the $\hat{\mu}_T^r$ coefficients capture how outcomes evolve differentially across the season for teams with Black quarterbacks as compared to those with white quarterbacks. $(\hat{\mu}_T^r + \hat{\mu}_T^p)$ and $(\hat{\mu}_T^r + \hat{\mu}_T^c)$ capture this same differential for playoff only and championship-level teams respectively.

We include four main fixed effects. The $DMA \times season$ fixed effect, τ_{as}, absorbs long-run trends in outcomes within each DMA. The $month \times DMA$ fixed effect, κ_{am}, controls for location-specific seasonal patterns. The $month \times record$ fixed effect, ψ_{Rm}, absorbs seasonal variation in the outcomes attributable to the team's record, regardless of the quarterback's race. A $month \times season$ fixed effect, ϕ_{ms}, absorbs common aggregate shocks in a given month. Together, these fixed effects ensure the $\hat{\mu}$'s identify the differential seasonal evolution of outcomes in areas with Black versus white quarterback-led teams among teams with equivalent performance records.

One may be concerned that the race of a team's quarterback is not randomly assigned. It is plausible that teams in areas with more racial animus are less likely to select a Black quarterback. With our specification, if teams make this choice in response to the level of racial animus in their fan base, the $DMA \times season$ fixed effect controls for this selection as we are estimating within-season changes. Of greater concern is if teams select Black quarterbacks when racial hate in their fan base is trending downward. This could affect our within season estimates, generating downward bias over the course of a season. In Table A7 we check for evidence of this type of selection pattern. Having a Black quarterback in season s is regressed on $Std(HateCrimeRate_{n,s-1} - HateCrimeRate_{n,s-2})$, the standardized change in the hate crime rate in the team's catchment areas over the two prior seasons. Several specifications are tested. We find no evidence that reductions in hate crimes over the prior two seasons are predictive of a team having a Black quarterback.

This specification corresponds to the local quarterback effects of the performance-contingent updating model in Section 3.2. The animus change over the course of the season, $E[\bar{A}_{n,s+1} - \bar{A}_{n,s}]$, is estimated by our *after* season (March-April) parameters ($\hat{\mu}_{T=a}$) which capture the change in animus relative to baseline, when $T = t_0$. For teams with Black quarterbacks, our model predicts that high-record championship-level teams will induce positive updating ($\theta_L(c) > \bar{\theta}$), thereby reducing racial animus ($\hat{\mu}_a^r + \hat{\mu}_a^c \leq 0$). This signal will be increasing in performance ($\theta_L(c) > \theta_L(p) > \theta_L(r)$) such that $\hat{\mu}_a^c < \hat{\mu}_a^p < 0$. The model generates ambiguous predictions on how the signal received from lower performing teams ($\theta_L(p)$ and $\theta_L(r)$) compares to baseline $\bar{\theta}$. Whether these generate positive updating and reduce animus, captured by $\hat{\mu}_a^r$ and ($\hat{\mu}_a^r + \hat{\mu}_a^p$), is estimated empirically. The magnitude of all effects will be larger in high engagement DMAs ($\frac{\partial |\Delta \bar{A}|}{\partial Engagement} > 0$).

6.1.2 Long-run effects of local Black quarterbacks

Anti-Black hate crimes: Table 6, column 1, reports the after-season estimates ($\hat{\mu}_a$) from equation 6 on monthly DMA anti-Black hate crimes per 100,000. For low-performance, non-playoff Black quarterback teams, the small and statistically insignificant estimate of $\hat{\mu}_a^r$ in the first row implies that these DMAs experience no significant or persistent difference in their anti-Black hate crime rates as compared to DMAs with non-playoff white quarterback teams. In contrast, and consistent with model predictions, areas with championship-level Black quarterback teams experience a substantial reduction in the anti-Black hate crime rate ($\hat{\mu}_a^r + \hat{\mu}_a^c < 0$) relative to DMAs with championship-level white quarterback teams. In the months after the season, their anti-Black hate crime rate is differentially lower by -0.018 (p-value = 0.037), a 34% decrease. As predicted by the model, the statistically insignificant effect on playoff-only teams falls in between at -0.006. Figure 5(a) shows the evolution of the differential gap between Black and white quarterback teams for championship teams (in red) and non-playoff teams (in blue); the black line shows the average differential across all performance records. Consistent with a response to team quarterbacks, the anti-Black hate crime rate in DMAs with championship-level teams begins to drop in September, when the NFL season starts, and is then persistently lower through the NFL season and into the spring after the season ends. The effect of championship-level Black quarterbacks is large but not implausible. The 34% decrease we estimate is larger than, but broadly comparable to, the 16% drop in anti-Muslim hate crimes that Alrababa'h et al. (2021) document in Liverpool after Mohamed Salah joined the club and led it to sustained success.

As the model predicts, this effect is driven by DMAs with above-median levels of team engagement ($\frac{\partial|\Delta\bar{A}|}{\partial Engagement} > 0$). Figure 6(a) plots the after-season effect of Black quarterbacks on non-playoff teams ($\hat{\mu}_a^r$) and the differential for championship-level teams ($\hat{\mu}_a^c$) across quartiles of estimated team engagement. While Black quarterbacks on non-playoff teams have no differential effect in any engagement quartile, the magnitude of the effect for championship-level Black quarterbacks grows with DMA engagement: from -0.001 in low-engagement DMAs to -0.040 in high-engagement DMAs.

Alternative specifications of our anti-Black hate crime measure (appendix Table A8) all estimate similar patterns: small effects on anti-Black hate crimes for non-playoff Black quarterback led teams and large reductions for championship teams. This pattern is also robust to alternative estimation approaches using different combinations of fixed effects (appendix Table A9). To check whether this effect is driven by a particular team or quarterback, appendix Figure A3(a) estimates equation 6 omitting (i) all observations for a specific team, and (ii) all observations tied to a particular Black quarterback. Estimates are robust to this leave-one-out test, suggesting the effect is not solely driven by a single quarterback or team. All of the non-playoff ($\hat{\mu}_a^r$) and championship interaction ($\hat{\mu}_a^c$) estimates lie close to our full-sample estimates and fall well within their 95% confidence intervals (the shaded green region).

To better understand the nature of these reductions in anti-Black hate crimes, Table A10 examines impacts by crime type. In contrast to the short-run effects, these long-run effects are driven by a reduction in multi-perpetrator crimes, suggesting that these changes in a DMA's baseline level of anti-Black animus reduce instances of socially coordinated hate (Glaeser 2005). Appendix Table A11 examines hate targeting other racial groups. Consistent with the updating of anti-Black attitudes, we find no evidence of impacts on hate crimes targeting Whites, Hispanics, or other racial groups.

Google searches for racial slurs: Table 6, column 2, row 1, and the blue line in Figure 5(b) show that when a team led by a Black quarterback fails to make the playoffs, racial slur searches increase and remain high beyond the end of the NFL season ($\hat{\mu}_a^r > 0$). Relative to similarly performing teams led by white quarterbacks, racial slur searches in the subsequent after-season are 2.5% higher (p-value = 0.038) in DMAs where a Black quarterback team did not reach the playoffs. In contrast, as shown by the red line in Figure 5(b) and the last row of Table 6, column 2, racial slur searches display a persistent decrease for championship-level teams led by Black quarterbacks ($\hat{\mu}_a^r + \hat{\mu}_a^c < 0$), consistent with model predictions. Relative to DMAs with similarly performing teams led by

white quarterbacks, racial slur searches in the subsequent season are 3.7% lower (p-value = 0.053) in DMAs with championship-level Black quarterback teams. The statistically insignificant effect on playoff-only teams falls in between with a reduction of 0.021%. As predicted by the model, this difference in effects is driven by DMAs with high levels of team engagement. Figure 6(b) plots the after-season effect of Black quarterbacks for non-playoff teams and the differential for championship-level teams for different quartiles of estimated team engagement. We detect no differences in the lowest quartile of team engagement. It is in the highest engagement quartile that Black quarterbacks on non-playoff teams increase searches by 5.8%, while for championship-level teams searches are differentially lower by 11%. Appendix Figure A3(b) presents the leave-one-out test estimates, checking whether this effect is driven by a particular team or quarterback. All of the leave-one-out estimates lie close to our full-sample estimates and fall within their confidence intervals, suggesting the effect is not driven by a single quarterback or team. In terms of magnitudes, Anderson et al. (2020) examine how the economic shock of the Great Recession fueled racial animus, finding that post-recession slur searches increased by a differential 5.5% in states with a 1 SD greater manufacturing share. Similarly, Bestenbostel and Peralta (2026) examine how racial slur searches respond to news of mass layoffs, finding a 1.4% increase. These slur-search responses are smaller in magnitude than the corresponding effects on hate crimes, reflecting the differences between the two measures. Hate crimes are rare, severe, and high-cost actions, so small changes in the number of incidents translate into large percentage effects, whereas racial-slur searches are private, low-cost, and far more common, making them a broader measure of shifts in racial animus. Anderson et al. (2020) also observe this: the one-standard-deviation increase in manufacturing share associated with a 5.5% increase in slur searches was associated with a 55% increase in anti-Black hate crimes.

Hate speech on Reddit: Table 6, column 3, reports estimates for the share of submissions removed. We find no statistically significant difference in the share removed for Black- versus white-quarterback teams at any level of performance. Figure 5(c) plots the corresponding two-month estimates separately for non-playoff and championship-level Black-quarterback teams, together with the average across all teams (in black); none show evidence of differential effects. For championship-level teams, however, these estimates are severely underpowered and sensitive to sample composition. We observe Reddit moderation only for the full 2016–2023 seasons, over which just three Black quarterbacks led teams to a combined eight championship-level finishes.

The full-sample championship estimate is correspondingly uninformative: its 95% confidence interval spans effect sizes from an 89% increase to an 81% decrease in the share removed. The leave-one-out test in appendix Figure A3(c) underscores this fragility. Most estimates lie close to the full-sample values. However, omitting the Eagles, or their quarterback Jalen Hurts, moves the championship estimate substantially, flipping its sign to imply a reduction in subreddit moderation, in line with our other outcomes.[24] We therefore read the Reddit results for championship teams as uninformative rather than as evidence against an effect: the sample is too small to separate the reductions we observe in other measures from no effect at all. By contrast, estimates for low-performing Black-quarterback teams are far more precise, with a 95% confidence interval that rules out increases in the share removed above 3.37 percentage points (32.7%).

Implicit racial biases: Table 6, column 4, reports estimates on the Implicit Association Test scores of white test-takers. We see no statistically significant differential changes in the IAT scores of DMAs with non-playoff Black quarterback teams. The effect on playoff-only teams is also null. In contrast, and consistent with model predictions, DMAs with teams led by a Black quarterback that reach the championship games experience a gradual decline in implicit racial biases beginning in the playoff season (Figure 5(d)), reaching a reduction of -0.040 SD (p-value=0.002) that persists into the months after the NFL season ends. This effect increases with a DMA's level of team engagement, as illustrated in Figure 6(c). As the model predicts, the differential effect of championship-level Black quarterbacks grows in magnitude with engagement, from -0.029 in low-engagement DMAs to -0.074 in high-engagement DMAs. Black quarterbacks on non-playoff teams have no differential effect in any engagement quartile. Appendix Figure A3(c) presents the leave-one-out test estimates, checking whether this effect is driven by a particular team or quarterback. All of the leave-one-out estimates lie close to our full-sample estimates and fall within their confidence intervals, suggesting the effect is not driven by a single quarterback or team. Appendix Table A12 and Figure A5 examine whether exposure to Black quarterbacks affects selection into taking the IAT. We cannot reject the null of no differential effect on the number of test-takers in DMAs exposed to Black quarterback teams. The magnitude of this effect is large but not implausible. A few papers examine the persistent effect of out-group exposure on implicit bias, studying exposure to immigrant communities

[24] Jalen Hurts led the Philadelphia Eagles to the championships once during our Reddit sample, in the 2022 season. He lost the Super Bowl, the first in NFL history to feature two Black starting quarterbacks, to Patrick Mahomes. We have no a priori reason to exclude these observations; we note only that the championship estimate rests heavily on the Eagles 2022 season, and that without it the Reddit results align with the reduced-animus pattern found in our other measures (see appendix Figure A4).

(Bursztyn et al. 2024), to Black US troops stationed in Britain during World War II (Schindler and Westcott 2021), and to the racially diverse cast of Sesame Street (Duquennois and Zeng 2025). These studies estimate out-group bias reductions of between 0.065 and 0.075 SDs, slightly larger than the magnitudes we find. A more intensive form of contact, sharing a dorm room with a Black student in South Africa, induced much larger 0.63 SDs reductions in implicit bias (Corno et al. 2022).

Taken together, our estimates across multiple indicators of racial animus consistently suggest that over the course of a season, anti-Black animus responds to exposure to local Black quarterbacks. High-performing, championship-level teams with Black quarterbacks reduce animus across a range of indicators, implying that (i) local Black quarterbacks induce updates in racial attitudes ($\lambda_L \neq 0$), and (ii) exposure to Black quarterbacks on high-performing teams generates an animus-reducing signal ($(\theta_L(Rec_{ns} = c) - \bar{\theta}_{ns}) > 0$). In contrast, and consistent with the prediction that $(\theta_L(Rec_{ns} = c) - \bar{\theta}_{ns}) > (\theta_L(Rec_{ns} = r) - \bar{\theta}_{ns})$, the long-run effects of exposure to Black quarterbacks on low-performing teams that do not advance to the playoffs are generally null, with the exception of a 2.5% increase in Google searches for racial slurs. The signal generated by Black quarterbacks on low-performing teams therefore either matches or falls slightly below baseline perceptions ($\theta_L(Rec_{ns} = r) \leq \bar{\theta}_{ns}$).

6.2 Opposition Black quarterbacks in the long-run

6.2.1 Empirical strategy

To estimate how the evolution of outcomes over the NFL season differs with exposure to opposition Black quarterbacks, we adapt the strategy developed in Section 6.1. We use the same T time periods and the same *DMA* $\times$ *season*, *month* $\times$ *DMA*, *month* $\times$ *record*, and *month* $\times$ *season* fixed effects to control for the general evolution of outcomes over the years, DMA-specific seasonal variation, seasonal variation attributable to NFL team performance, and general time trends. Focusing on teams with white quarterbacks, we estimate versions of

$$Y_{ansm} = \delta_0^E + \sum_T \delta_T^E \, E_{ns} \times \mathbb{1}\{m = T\} + \tau_{as} + \kappa_{am} + \psi_{Rm} + \phi_{ms} + \epsilon_{ansm}, \tag{7}$$

where E_{ns} denotes a measure of team n's exposure to opposition Black quarterbacks in season s. We consider two such measures.

The first decomposes exposure by game outcome. In place of E_{ns}, we enter jointly the number of

regular-season games team n won (GW_{ns}) and lost (GL_{ns}) against an opposition Black quarterback in season s. The resulting coefficients, δ_T^{GW} and δ_T^{GL}, give the effect of one additional regular-season game won (or lost) against an opposition Black quarterback on the outcome in period T, capturing whether the response varies with the signal conveyed by that exposure ($\theta_O(\Omega_{ns})$).

The second focuses on highly salient games, since the average effect of a single regular-season game may be too small to generate persistent long-run impacts. Among teams that reached the playoffs, we define an indicator K_{ns} equal to one for teams knocked out of the playoffs by an opposition Black quarterback (and zero for the remaining playoff teams). The coefficients δ_T^K then give the effect of a playoff elimination by an opposition Black quarterback on the outcome in period T.

These specification corresponds to the opposition-quarterback effects of the performance-contingent updating model in Section 3.2. The change in animus over the course of the season, $E[\bar{A}_{n,s+1} - \bar{A}_{n,s}]$, is estimated by our *after*-season (March–April) parameters ($\hat{\delta}_{T=a}$), which capture the change in animus relative to the baseline period ($T = t_0$). For teams facing opposition Black quarterbacks, the model's predictions are ambiguous: opposition Black quarterbacks increase exposure to counter-stereotypical Black leadership, but (i) this exposure is less frequent and (ii) the competitive setting may heighten social boundaries and reinforce group distinctions rather than attenuate them. Given this ambiguity, the sign of the opposition signal is theoretically indeterminate and must be established empirically.

6.2.2 Long-run effects of opposition Black quarterbacks

One team's local Black quarterback is, by definition, an opposition Black quarterback to a different team. In Section 6.1 we found that championship-level local Black quarterbacks generated a long-run reduction in racial animus in their teams' catchment areas. In aggregate, the long-run effects of Black quarterbacks could be quite different if this reduction were offset by persistent increases in racial animus in opposing teams' catchment areas. Figure 7 examines this question using two approaches to categorizing exposure to opposition Black quarterbacks. Focusing on teams with white quarterbacks, Figure 7 reports the δ_T^{GW} and δ_T^{GL} estimates from equation 7. These estimate the long-run effects of an additional regular-season win (in red) and loss (in blue) against an opposition Black quarterback. The right column of Figure 7 reports the δ_T^K's from equation 7 estimated with the knockout indicator K_{ns}. These estimate the long-run effect of being knocked out of the playoffs by an opposition Black quarterback.

Using these approaches, we find no evidence that facing opposing Black quarterbacks on the field leads to persistent changes in racial animus. The estimated coefficients for regular-season games in the left column of Figure 7 are small in magnitude, generally statistically insignificant, and show no consistent or persistent pattern over the course of the NFL season. The coefficients for knockout games are imprecisely estimated, given the smaller playoff sample, and also show no consistent evidence of a persistent impact. This absence of a persistent opposition effect suggests one of three explanations: (i) exposure to opposition Black quarterbacks does not lead to updating ($\lambda_O = 0$); (ii) opposition exposure is too small relative to that generated by local Black quarterbacks to produce comparable updating ($E_{ns} \ll L_{ns}$); or (iii) the (theoretically ambiguous) signals from interactions with opposition Black quarterbacks are largely uninformative, leaving perceptions near baseline ($\theta_O(\Omega_{ns}) \approx \bar{\theta}_{ns}$). This finding is consistent with Alrababa'h et al. (2021), who document that Mohamed Salah did not generate a rise in anti-Muslim sentiment among the comparison group of rival-club fans.

7 Star power and media visibility: Black quarterbacks vs. Black coaches

Like quarterbacks, head coaches are also consequential team leaders (Berry and Fowler (2019)). Black head coaches also occupy historically racialized leadership positions and have faced persistent barriers to entry and differential evaluation (Madden (2004), Fanning Madden and Ruther (2011), Pitts et al. (2024), Salaga and Juravich (2020)). Yet head coaches are not as publicly salient as quarterbacks. Quarterbacks receive substantial media attention both during games and off the field through lucrative sponsorship deals and advertising visibility. They are involved in nearly every offensive play, and their successes and failures are easily recognized by audiences. Head coaches, by contrast, are less directly visible as on-field actors and their influence on game outcomes is less obvious. They receive substantially less media attention and are less recognizable to the general public — a gap reflected in fame rankings.[25]

To compare these two forms of minority leadership that differ sharply in public visibility, we repeat our analysis to examine the effect of the race of NFL head coaches. If visibility is critical to the effects we document for Black quarterbacks, we would expect the corresponding estimates for Black head coaches to be smaller. Table A13 reports our main short-run estimates for Black

[25]Vince Lombardi — the legendary NFL coach and namesake of the Super Bowl trophy — is outranked by 9 NFL quarterbacks on YouGov's all-time sports personality fame ranking (YouGov 2026). Even highly successful contemporary coaches such as Bill Belichick and Andy Reid have fame scores of 64% and 53%, well below those of leading quarterbacks such as Tom Brady (87%) and Peyton Manning (82%).

head coaches on hate crime incidents, Reddit moderation, and IAT scores. For all three outcomes, we cannot reject the null of no differential post-game change after losses involving a Black head coach, relative to losses where both head coaches are white. Estimates are small and statistically insignificant, as are the corresponding win effects. Table A14 reports our main long-run estimates for Black head coaches on hate crime incidents, Reddit moderation, racial slur searches, and IAT scores. Across all outcomes and performance levels, we cannot reject the null of no effect.

Taken together, these null estimates suggest that leadership alone is not sufficient to generate the short-run racial retaliation and long-run attitude-updating effects we document for Black quarterbacks. Black head coaches affect wins, losses, and season outcomes, but without comparable media visibility or individualized public attribution. This distinction underscores the role of the media in shaping when minority leadership influences public attitudes: minority leaders affect racial biases when they are not only in positions of authority, but are also highly visible, repeatedly discussed, and individually credited or blamed for emotionally meaningful outcomes.

8 Heterogeneity by racial context

How do the effects of exposure to Black quarterbacks differ with an area's racial context? A priori, the direction of any such interaction is ambiguous. Racial context could shape both how entrenched baseline perceptions $\bar{\theta}$ are and how far new signals fall from them, and these forces need not push the same way: more entrenched perceptions imply less updating, while a larger gap between perceptions and new signals implies more. Because we cannot sign the net effect ex ante, we assess it empirically, examining how the hate-crime effects vary along two dimensions: (i) the white–Black exposure index, a DMA-level measure of white residents' exposure to Black residents, and (ii) baseline racial animosity, captured by the frequency of Google slur searches in the DMA in 2015. For each measure, we define an indicator equal to one for above-median DMAs and zero otherwise, and estimate modified versions of our main specifications that interact the treatment variables with these indicators.

Column 1 of Table A15 reports the short-run heterogeneity estimates for emotionally intense close-upset games by the white-Black exposure index. The first two rows report the differential effect of Black-quarterback games in below-median exposure DMAs, while rows three and four capture the additional differential effect in above-median exposure DMAs. For the full sample, we cannot reject the null of no heterogeneity by the exposure index. A similar pattern emerges in

the season-level results. Figure A7(a) plots the differential hate-crime effects for championship-level Black-quarterback teams in below-median (blue) and above-median (red) exposure DMAs. Here too, we find no statistically significant difference across the two groups. We next examine heterogeneity by the 2015 RAI. Column 2 of Table A15 reports the short-run estimates, and Figure A7(b) presents the season-level coefficients. Again, for the full sample, we cannot reject the null of no heterogeneity by baseline RAI.

These null results are difficult to interpret. The white–Black exposure index and the baseline RAI do not isolate exogenous contextual features of an area. Both are likely to correlate with a broad set of DMA characteristics, including the Black population share, urbanicity, local attachment to the NFL team, media consumption patterns, policing and reporting environments, and the cost of committing hate crimes, among many others. As a result, the absence of strong heterogeneity by these measures does not necessarily imply that racial context is unimportant. Rather, it suggests that these reduced-form proxies may bundle offsetting channels, making it hard to map the estimates cleanly onto a single theoretical mechanism.

9 Conclusion

This paper studies the dynamics of racial animus in a setting where the same minority group is repeatedly and publicly tied to both success and failure. Existing empirical work typically isolates a single force acting in one direction: a shock that inflames hostility or an intervention that reduces it. By contrast, we study a single real-world setting in which prominent Black leaders' performance is repeatedly observed and publicly evaluated. Using Black quarterbacks in the NFL, we show that exposure to Black leadership can be both animus-activating and animus-ameliorating, with the direction depending on performance and time horizon. We trace these responses across four measures of racial animus that differ in visibility and severity: anti-Black hate crimes, online hate-speech moderation, racial-slur searches, and implicit bias.

In the short run, emotionally intense losses involving Black quarterbacks generate racial retaliation. Relative to comparable games where both quarterbacks were white, close-scoring upset losses involving a Black quarterback are followed by a 43% increase in the anti-Black hate crime rate, a 29% increase in online hate-speech indicators on Reddit, and a 0.043 SD increase in implicit racial bias. Comparable wins do not generate corresponding reductions in animus. This asymmetry provides evidence that individual-level attribution bias can spill over into population-level animus:

when a salient Black leader is associated with failure, hostility can extend beyond the individual to unrelated members of the broader racial group. The results also show that the emotional shocks generated by sporting events can be racialized, producing retaliatory behavior toward out-groups who share an identity with salient athletes.

Over longer horizons, we find a different and more consequential pattern. The short-run effects, however large, are transient and concentrated in the days around a handful of emotionally intense games. The seasonal effects, by contrast, shift baseline racial animus and persist well beyond the end of the season. Repeated exposure to successful local Black quarterbacks reduces racial animus, while exposure to low-performing Black-quarterback teams has limited effects. In areas where Black-quarterback teams reach the championship games, hate crime rates are 34% lower, racial-slur searches decline by 3.7%, and implicit racial bias falls by -0.040 SDs. Because these reductions are sustained and population-wide rather than fleeting, their cumulative consequences for intergroup relations are likely larger than those of the short-run spikes. These seasonal findings also advance evidence on parasocial contact and minority role models: we move beyond a single celebrity by studying 68 Black quarterbacks over more than two decades, show that effects are performance-contingent, and contrast longer-run reductions in animus with game-day increases after losses, highlighting the temporal dynamics of racial animus.

The comparison with Black head coaches helps clarify the mechanism. Black head coaches are consequential leaders, but they are less visible during play and less individually credited or blamed for game outcomes. We find no comparable short- or longer-run effects of Black head coaches on racial animus. This suggests that minority leadership alone is not sufficient. Visibility, repeated public attention, and individualized attribution are central conditions under which minority leaders shape majority-group prejudice.

More broadly, this paper provides the first causal evidence that one of America's most visible nonpolitical institutions—the NFL—can shape racial animus and social cohesion within the United States. Representation can reduce animus when minority leaders are repeatedly associated with success, but it can also impose costs on the broader minority group when salient failures trigger racialized blame. This suggests that efforts to reduce prejudice through representation depend not only on increasing minority visibility, but also on the media and institutional environments that shape how minority performance is observed, attributed, and discussed. A natural next step for future research is to test whether the same dynamics arise in other arenas where minority leaders are highly visible and publicly judged, from politics and business to teaching and entertainment.

References

Adena, M., Enikolopov, R., Petrova, M., Santarosa, V., and Zhuravskaya, E. (2015). Radio and the rise of the Nazis in prewar Germany. The Quarterly Journal of Economics, 130(4):1885–1939.

Allport, G. W., Clark, K., and Pettigrew, T. (1954). The nature of prejudice.

Alrababa'h, A., Marble, W., Mousa, S., and Siegel, A. A. (2021). Can exposure to celebrities reduce prejudice? The effect of Mohamed Salah on Islamophobic behaviors and attitudes. American Political Science Review, 115(4):1111–1128.

Alrababa'h, A., Marble, W., Mousa, S., and Siegel, A. A. (2024). Are minorities punished more harshly for underperformance? Evidence from Premier League soccer. OSF Preprints. Working Paper.

Anderson, D. M., Crost, B., and Rees, D. I. (2020). Do economic downturns fuel racial animus? Journal of Economic Behavior & Organization, 175:9–18.

Ang, D. (2023). The birth of a nation: Media and racial hate. American Economic Review, 113(6):1424–1460.

Bauer, M., Cahlíková, J., Chytilová, J., Roland, G., and Želinskỳ, T. (2023). Shifting punishment onto minorities: Experimental evidence of scapegoating. The Economic Journal, 133(652):1626–1640.

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., and Topalova, P. (2009). Powerful women: Does exposure reduce bias? The Quarterly Journal of Economics, 124(4):1497–1540.

Becker, G. S. (1957). The Economics of Discrimination. University of Chicago Press, Chicago.

Berri, D., Farnell, A., and Simmons, R. (2023). The determinants of Black quarterback pay in the National Football League. Managerial and Decision Economics, 44(3):1491–1503.

Berri, D. J. and Simmons, R. (2009). Race and the evaluation of signal callers in the National Football League. Journal of Sports Economics, 10(1):23–43.

Berry, C. R. and Fowler, A. (2019). How much do coaches matter? In Sports Analytic Conference, MIT Sloan, Boston, MA.

Bertrand, M. and Duflo, E. (2017). Field experiments on discrimination. Handbook of Economic Field Experiments, 1:309–393.

Bestenbostel, A. and Peralta, A. (2026). WARNings – The impact of negative economic news on racial animus. Journal of Economic Behavior & Organization, 243:107399.

Blalock, H. M. et al. (1967). Toward a theory of minority-group relations, volume 325. Wiley, New York.

Blouin, A. and Mukand, S. W. (2019). Erasing ethnicity? Propaganda, nation building, and identity in Rwanda. Journal of Political Economy, 127(3):1008–1062.

Blumer, H. (1958). Race prejudice as a sense of group position. Pacific Sociological Review, 1(1):3–7.

Bohren, J. A., Imas, A., and Rosenberg, M. (2019). The dynamics of discrimination: Theory and evidence. American Economic Review, 109(10):3395–3436.

Bordalo, P., Coffman, K., Gennaioli, N., and Shleifer, A. (2016). Stereotypes. The Quarterly Journal of Economics, 131(4):1753–1794.

Burnett, N. J. and Van Scyoc, L. J. (2015). Compensation discrimination for defensive players: Applying quantile regression to the National Football League market for linebackers and offensive linemen. Journal of Sports Economics, 16(4):375–389.

Bursztyn, L., Chaney, T., Hassan, T. A., and Rao, A. (2024). The immigrant next door. American Economic Review, 114(2):348–384.

Bursztyn, L., Egorov, G., Haaland, I., Rao, A., and Roth, C. (2022). Scapegoating during crises. In AEA Papers and Proceedings, volume 112, pages 151–155. American Economic Association.

Cao, A., Lindo, J. M., and Zhong, J. (2023). Can social media rhetoric incite hate incidents? Evidence from Trump's "Chinese Virus" tweets. Journal of Urban Economics, 137:103590.

Card, D. and Dahl, G. B. (2011). Family violence and football: The effect of unexpected emotional cues on violent behavior. The Quarterly Journal of Economics, 126(1):103–143.

Carlana, M. (2019). Implicit stereotypes: Evidence from teachers' gender bias. The Quarterly Journal of Economics, 134(3):1163–1224.

Carr, J., Clifton-Sprigg, J., James, J., and Vujic, S. (2020). Love thy neighbour? Brexit and hate crime. Technical report, IZA Discussion Papers.

Caselli, M., Falco, P., and Mattera, G. (2023). When the stadium goes silent: How crowds affect the performance of discriminated groups. Journal of Labor Economics, 41(2):431–451.

Caselli, M., Falco, P., and Somekh, B. (2024). Inside the NBA bubble: How Black players performed better without fans. Journal of Population Economics, 37(2):39.

Caselli, T., Basile, V., Mitrović, J., and Granitzer, M. (2020). HateBERT: Retraining BERT for abusive language detection in English. arXiv preprint arXiv:2010.12472.

Chandrasekharan, E., Jhaver, S., Bruckman, A., and Gilbert, E. (2022). Quarantined! Examining the effects of a community-wide moderation intervention on Reddit. ACM Transactions on Computer-Human Interaction (TOCHI), 29(4):1–26.

Chandrasekharan, E., Pavalanathan, U., Srinivasan, A., Glynn, A., Eisenstein, J., and Gilbert, E.

(2017). You can't stay here: The efficacy of Reddit's 2015 ban examined through hate speech. Proceedings of the ACM on Human-Computer Interaction, 1(CSCW):1–22.

Chetty, R., Hendren, N., Jones, M. R., and Porter, S. R. (2020). Race and economic opportunity in the United States: An intergenerational perspective. The Quarterly Journal of Economics, 135(2):711–783.

Connor, P. and Evers, E. R. (2020). The bias of individuals (in crowds): Why implicit bias is probably a noisily measured individual-level construct. Perspectives on Psychological Science, 15(6):1329–1345.

Corno, L., La Ferrara, E., and Burns, J. (2022). Interaction, stereotypes, and performance: Evidence from South Africa. American Economic Review, 112(12):3848–75.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? American Economic Review, 95(2):158–165.

DellaVigna, S. and La Ferrara, E. (2015). Economic and social impacts of the media. In Handbook of Media Economics, volume 1, pages 723–768. Elsevier.

Depetris-Chauvin, E., Durante, R., and Campante, F. (2020). Building nations through shared experiences: Evidence from African football. American Economic Review, 110(5):1572–1602.

Derenoncourt, E. (2022). Can you move to opportunity? Evidence from the Great Migration. American Economic Review, 112(2):369–408.

Dharmapala, D. and Huq, A. (2024). Imputing unreported hate crimes using Google search data. Journal of Law & Empirical Analysis, 1(2):232–254.

Ducking, J., Groothuis, P. A., and Hill, J. R. (2014). Compensation discrimination in the NFL: An analysis of career earnings. Applied Economics Letters, 21(10):679–682.

Ducking, J., Groothuis, P. A., and Hill, J. R. (2017). Compensation discrimination: An analysis of linebackers, defensive linemen, and defensive backs in the National Football League. The Review of Black Political Economy, 44(3-4):251–260.

Duncan, B., Mansour, H., and Rees, D. I. (2017). It's just a game: The Super Bowl and low birth weight. Journal of Human Resources, 52(4):946–978.

Duquennois, C. and Zeng, J. (2025). Casting roles, casting votes: Lessons from Sesame Street on media representation, racial biases, and voting. Working paper.

Edmans, A., Garcia, D., and Norli, Ø. (2007). Sports sentiment and stock returns. The Journal of Finance, 62(4):1967–1998.

Egan, M., Matvos, G., and Seru, A. (2022). When Harry fired Sally: The double standard in

punishing misconduct. Journal of Political Economy, 130(5):1184–1248.

Eren, O. and Mocan, N. (2018). Emotional judges and unlucky juveniles. American Economic Journal: Applied Economics, 10(3):171–205.

Fairlie, R. W., Hoffmann, F., and Oreopoulos, P. (2014). A community college instructor like me: Race and ethnicity interactions in the classroom. American Economic Review, 104(8):2567–91.

Fanning Madden, J. and Ruther, M. (2011). Has the NFL's Rooney Rule efforts "leveled the field" for African American head coach candidates? Journal of Sports Economics, 12(2):127–142.

Gallo, E., Grund, T., and James Reade, J. (2013). Punishing the foreigner: Implicit discrimination in the Premier League based on oppositional identity. Oxford Bulletin of Economics and Statistics, 75(1):136–156.

Gawronski, B. (2019). Six lessons for a cogent science of implicit bias and its criticism. Perspectives on Psychological Science, 14(4):574–595.

Gershenson, S., Hart, C. M., Hyman, J., Lindsay, C. A., and Papageorge, N. W. (2022). The long-run impacts of same-race teachers. American Economic Journal: Economic Policy, 14(4):300–342.

Glaeser, E. L. (2005). The political economy of hatred. The Quarterly Journal of Economics, 120(1):45–86.

Glover, D., Pallais, A., and Pariente, W. (2017). Discrimination as a self-fulfilling prophecy: Evidence from French grocery stores. The Quarterly Journal of Economics, 132(3):1219–1260.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., and Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. Journal of Personality and Social Psychology, 97(1):17.

Gregory-Smith, I., Bryson, A., and Gomez, R. (2023). Discrimination in a rank order contest: Evidence from the NFL Draft. Journal of Economics, Race, and Policy, 6(3):215–231.

Guha, I. (2024). Elections and hate crimes in the United States. Available at SSRN 4802809.

Hanes, E. and Machin, S. (2014). Hate crime in the wake of terror attacks: Evidence from 7/7 and 9/11. Journal of Contemporary Criminal Justice, 30(3):247–267.

Hartvigsen, T., Gabriel, S., Palangi, H., Sap, M., Ray, D., and Kamar, E. (2022). ToxiGen: A large-scale machine-generated dataset for adversarial and implicit hate speech detection. arXiv preprint arXiv:2203.09509.

Healy, A. J., Malhotra, N., and Mo, C. H. (2010). Irrelevant events affect voters' evaluations of government performance. Proceedings of the National Academy of Sciences, 107(29):12804–12809.

Hehman, E., Calanchini, J., Flake, J. K., and Leitner, J. B. (2019). Establishing construct validity evidence for regional measures of explicit and implicit racial bias. Journal of Experimental Psychology: General, 148(6):1022.

Ivandić, R., Kirchmaier, T., Saeidi, Y., and Blas, N. T. (2024). Football, alcohol, and domestic abuse. Journal of Public Economics, 230:105031.

Kamel, D. and Woo-Mora, L. G. (2023). Skin tone penalties: Quasi-experimental evidence on colorism in football. SSRN Working Paper.

Karp, A. (2014). Sports business journal: NFL teams saw local market ratings gain in '13: Chiefs, Seahawks up big. Accessed: 2025-11-28.

Kearney, M. S. and Levine, P. B. (2020). Role models, mentors, and media influences. The Future of Children, 30(1):83–106.

Keefer, Q. A. and Kniesner, T. J. (2023). Injury risk, concussions, race, and pay in the NFL. Journal of Risk and Uncertainty, 67(2):107–136.

Keefer, Q. A. W. (2013). Compensation discrimination for defensive players: Applying quantile regression to the National Football League market for linebackers. Journal of Sports Economics, 14(1):23–44.

Kline, P., Rose, E. K., and Walters, C. R. (2022). Systemic discrimination among large US employers. The Quarterly Journal of Economics, 137(4):1963–2036.

Lang, K. and Kahn-Lang Spitzer, A. (2020). Race discrimination: An economic perspective. Journal of Economic Perspectives, 34(2):68–89.

Lim, J. and Meer, J. (2020). Persistent effects of teacher–student gender matches. Journal of Human Resources, 55(3):809–835.

Lowe, M. (2021). Types of contact: A field experiment on collaborative and adversarial caste integration. American Economic Review, 111(6):1807–1844.

Lowes, S., Nunn, N., Robinson, J. A., and Weigel, J. (2015). Understanding ethnic identity in Africa: Evidence from the implicit association test (IAT). American Economic Review, 105(5):340–345.

Madden, J. F. (2004). Differences in the success of NFL coaches by race, 1990-2002: Evidence of last hire, first fire. Journal of Sports Economics, 5(1):6–19.

Marquez-Velarde, G., Grashow, R., Glass, C., Blaschke, A. M., Gillette, G., Taylor, H. A., and Whittington, A. J. (2023). The paradox of integration: Racial composition of NFL positions from 1960 to 2020. Sociology of Race and Ethnicity, 9(4):451–469.

Matsuzawa, K. and Arnesen, E. (2024). Sports betting legalization amplifies emotional cues &

intimate partner violence. Available at SSRN.

Montolio, D. and Planells-Struse, S. (2016). How time shapes crime: The temporal impacts of football matches on crime. Regional Science and Urban Economics, 61:99–113.

Mousa, S. (2020). Building social cohesion between Christians and Muslims through soccer in post-ISIS Iraq. Science, 369(6505):866–870.

Mulholland, S. E. (2013). White supremacist groups and hate crime. Public Choice, 157(1):91–113.

Müller, K. and Schwarz, C. (2021a). Fanning the flames of hate: Social media and hate crime. Journal of the European Economic Association, 19(4):2131–2167.

Müller, K. and Schwarz, C. (2021b). Fanning the flames of hate: Social media and hate crime. Journal of the European Economic Association, 19(4):2131–2167.

Müller, K. and Schwarz, C. (2023). From hashtag to hate crime: Twitter and antiminority sentiment. American Economic Journal: Applied Economics, 15(3):270–312.

Munyo, I. and Rossi, M. A. (2013). Frustration, euphoria, and violent crime. Journal of Economic Behavior & Organization, 89:136–142.

Onuchic, P. (2022). Recent contributions to theories of discrimination. arXiv preprint arXiv:2205.05994.

Ourand, J. (2022). Sports business journal: Gaining yardage: Local TV ratings rise for 21 of NFL's 32 teams including key Los Angeles market. Accessed: 2025-11-28.

Pager, D. and Shepherd, H. (2008). The sociology of discrimination: Racial discrimination in employment, housing, credit, and consumer markets. Annual Review of Sociology, 34(1):181–209.

Paluck, E. L. (2009). Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda. Journal of Personality and Social Psychology, 96(3):574.

Paluck, E. L. and Green, D. P. (2009). Prejudice reduction: What works? A review and assessment of research and practice. Annual Review of Psychology, 60(1):339–367.

Paluck, E. L., Porat, R., Clark, C. S., and Green, D. P. (2021). Prejudice reduction: Progress and challenges. Annual Review of Psychology, 72:533–560.

Parsons, C. A., Sulaeman, J., Yates, M. C., and Hamermesh, D. S. (2011). Strike three: Discrimination, incentives, and evaluation. American Economic Review, 101(4):1410–1435.

Petrova, M., Bursztyn, L., Egorov, G., and Enikolopov, R. (2019). Social media and xenophobia: Evidence from Russia. Technical report, National Bureau of Economic Research.

Pettigrew, T. F. (1979). The ultimate attribution error: Extending Allport's cognitive analysis of

prejudice. Personality and Social Psychology Bulletin, 5(4):461–476.

Pettis, R. W., Valencia, Z., and Williams, B. J. (2022). Pride and prejudice: Same-sex marriage legalization announcements and hate crimes. The Journal of Law and Economics, 65(4):811–835.

Pitts, J. D., Evans, B., and Johnson, J. D. (2024). Race and the probability of becoming a head coach for NFL coordinators since the introduction of the Rooney Rule: Why isn't Eric Bieniemy a head coach yet? The Review of Black Political Economy, 51(1):17–43.

Price, J. and Wolfers, J. (2010). Racial discrimination among NBA referees. The Quarterly Journal of Economics, 125(4):1859–1887.

Rao, G. (2013). Familiarity does not breed contempt: Generosity, discrimination and diversity in Delhi schools. Manuscript, Univ. California, Berkeley.

Ratliff, K. A. and Smith, C. T. (2021). Lessons from two decades of Project Implicit. In The Cambridge Handbook of Implicit Bias and Racism, page 4. Cambridge University Press.

Rees, D. I. and Schnepel, K. T. (2009). College football games and crime. Journal of Sports Economics, 10(1):68–87.

Reid, J. (2022). The Rise of the Black Quarterback: What It Means for America. Disney Electronic Content.

Rhoden, W. C. (2007). Third and a Mile: The Trials and Triumph of the Black Quarterback. ESPN Books.

Riley, E. (2024). Role models in movies: The impact of Queen of Katwe on students' educational attainment. Review of Economics and Statistics, 106(2):334–351.

Rose, E. K. (2023). A constructivist perspective on empirical discrimination research. Journal of Economic Literature, 61(3):906–923.

Rosenzweig, L. R. and Zhou, Y.-Y. (2021). Team and nation: Sports, nationalism, and attitudes toward refugees. Comparative Political Studies, 54(12):2123–2154.

Sakong, J. (2026). Do Black electoral victories shift racial bias? Evidence from close elections. Journal of Public Economics, 256:105594.

Salaga, S. and Juravich, M. (2020). National Football League head coach race, performance, retention, and dismissal. Sport Management Review, 23(5):978–991.

Sarsons, H. (2017a). Interpreting signals in the labor market: Evidence from medical referrals. Job market paper.

Sarsons, H. (2017b). Recognition for group work: Gender differences in academia. American Economic Review, 107(5):141–145.

Schiappa, E., Gregg, P. B., and Hewes, D. E. (2005). The parasocial contact hypothesis. Communication Monographs, 72(1):92–115.

Schimmack, U. (2021). The Implicit Association Test: A method in search of a construct. Perspectives on Psychological Science, 16(2):396–414.

Schindler, D. and Westcott, M. (2021). Shocking racial attitudes: Black GIs in Europe. The Review of Economic Studies, 88(1):489–520.

Small, M. L. and Pager, D. (2020). Sociological perspectives on racial discrimination. Journal of Economic Perspectives, 34(2):49–67.

Stephens-Davidowitz, S. (2014). The cost of racial animus on a Black candidate: Evidence using Google search data. Journal of Public Economics, 118:26–40.

Stoecker, C., Sanders, N. J., and Barreca, A. (2016). Success is something to sneeze at: Influenza mortality in cities that participate in the Super Bowl. American Journal of Health Economics, 2(1):125–143.

Tajfel, H. and Turner, J. C. (1979). An integrative theory of intergroup conflict. In Austin, W. G. and Worchel, S., editors, The Social Psychology of Intergroup Relations, pages 33–47. Wadsworth, Monterey, CA.

Volz, B. D. (2017). Race and quarterback survival in the National Football League. Journal of Sports Economics, 18(8):850–866.

Wang, T. (2021). Media, pulpit, and populist persuasion: Evidence from Father Coughlin. American Economic Review, 111(9):3064–92.

Wann, D. L. and Branscombe, N. R. (1993). Sports fans: Measuring degree of identification with their team. International Journal of Sport Psychology.

Wann, D. L., Haynes, G., McLean, B., and Pullen, P. (2003). Sport team identification and willingness to consider anonymous acts of hostile aggression. Aggressive Behavior: Official Journal of the International Society for Research on Aggression, 29(5):406–413.

Yanagizawa-Drott, D. (2014). Propaganda and conflict: Evidence from the Rwandan genocide. The Quarterly Journal of Economics, 129(4):1947–1994.

YouGov (2026). The most famous all-time sports personalities in America. YouGov Ratings. Data collection period: First quarter of 2026. Accessed April 20, 2026.

Zhuravskaya, E., Petrova, M., and Enikolopov, R. (2020). Political effects of the internet and social media. Annual Review of Economics, 12:415–438.

Zussman, A. (2023). Discrimination in times of crises and the role of the media. American Economic

Journal: Applied Economics, 15(4):422–451.

10 Tables

Table 1: Summary statistics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel a: Game-day data									
	Units	Obs. level	Weights	Year of game	Obs.	Games x teams	Season x teams	Mean	Std. Dev
Anti-Black hate crimes	Crimes per 10 mil.	Date x DMA	DMA population	2002 - 2025	306625	9709	729	0.165	2.011
IAT score	Std. dev.	Individual tests		2003 - 2024	513614	6019	664	0.176	0.952
Share removed	Share	Date x Subreddit	Submission count	2015 - 2024	19734	4002	316	0.120	0.115
Appendix outcomes									
Submissions	Count	Date x Subreddit		2015 - 2024	19805	4002	316	65	83
Anti-Black hate speech	ToxiGen score	Date x Subreddit	Submission count	2015 - 2024	19664	4002	316	0.00008	0.00164
IAT count	Count	Date x DMA		2002 - 2025	4575065	9709	729	0.2	1.2
Panel b: Season data									
	Units	Obs. level	Weights	Complete seasons	Obs.		Season x teams	Mean	Std. Dev
Hate crimes	Crimes per 100,000	Month x DMA	DMA population	2002 - 2023	52044		682	0.053	0.134
Google slur searches	Ln(Index)	2 Months x DMA	DMA population	2004 - 2023	15432		613	3.518	0.377
IAT score	Std. score	Individuals tests		2004 - 2023	3584169		618	0.168	0.949
Share removed	Share	Month x Subreddit	Submission count	2016 - 2023	2928		244	0.103	0.073
Appendix outcomes									
Submissions	Count	Month x Subreddit		2016 - 2023	2928		244	1102	979
Anti-Black hate speech	Toxigen score	Month x Subreddit	Submission count	2016 - 2023	2916		244	0.00007	0.00038
IAT count	Count	Month x DMA		2002 - 2024	816024		713	5	22

Table 2: Game-day impacts on anti-Black hate crime rates

	(1)	(2)	(3)	(4)
	Daily DMA anti-Black hate crime rate (per 10 mil.)			
	All games	Close-upsets	Either	Neither
White-white games:				
Post	-0.00445	-0.0121	-0.00215	-0.00645
	(0.0168)	(0.0368)	(0.0209)	(0.0275)
Post × *Win*	-0.00150	0.0348	-0.00361	0.000879
	(0.0115)	(0.0297)	(0.0154)	(0.0174)
Any Black quarterback games:				
Post × *Any Black QB*	0.0364***	0.0863***	0.0482***	0.0191
	(0.0130)	(0.0324)	(0.0171)	(0.0200)
Post × *Any Black QB* × *Win*	-0.0207	-0.104**	-0.0281	-0.00944
	(0.0179)	(0.0469)	(0.0238)	(0.0273)
Coefficient sums and significance:				
$\rho_2^b + \rho_4^b$	0.016	-0.018	0.020	0.010
Dependent mean	0.200	0.202	0.201	0.199
Observations	306625	44715	172405	134220
Game × teams	9709	1401	5471	4238
FE: DMA × Game	Yes	Yes	Yes	Yes
FE: Day of week	Yes	Yes	Yes	Yes

Note: This table reports estimates from equation 5. Standard errors are reported in parentheses clustered at the game-team level, with the following significance indicators: * p<0.1, ** p<0.05 and *** p<0.01. Observations are at the date × DMA level. All estimations use DMA population weights. The outcome measure is the daily anti-Black DMA hate crime rate per 10 million DMA inhabitants. Close games are those with a score differential of 7 points or less. Upset games are those where the pre-game spread predicted the losing team would win by 3 points or more. The summed coefficient row reports the sum of the indicated coefficients with significance stars.

Table 3: Game-day impacts on Reddit moderation and white IAT scores

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Share of submissions removed				Standardized white IAT scores			
	All	Close-upsets	Either	Neither	All	Close-upsets	Either	Neither
White-white games:								
Post	0.0594***	0.0346	0.0595***	0.0553***	-0.0124	0.00837	-0.00589	-0.0235
	(0.00852)	(0.0232)	(0.0116)	(0.0107)	(0.0127)	(0.0287)	(0.0171)	(0.0175)
Post × *Win*	-0.0371***	-0.0234	-0.0398***	-0.0328***	-0.00740	0.0169	-0.00499	-0.0101
	(0.00539)	(0.0149)	(0.00758)	(0.00719)	(0.00837)	(0.0218)	(0.0114)	(0.0122)
Any Black quarterback games:								
Post × *Any Black QB*	0.000631	0.0455**	0.00534	-0.00545	0.0000162	0.0430*	0.00719	-0.00907
	(0.00651)	(0.0177)	(0.00918)	(0.00865)	(0.00930)	(0.0228)	(0.0124)	(0.0140)
Post × *Any Black QB* × *Win*	-0.00564	-0.0395*	-0.00656	-0.00514	-0.00166	-0.0743**	-0.0141	0.0152
	(0.00803)	(0.0215)	(0.0112)	(0.0111)	(0.0133)	(0.0338)	(0.0178)	(0.0199)
Coefficient sums and significance:								
$\rho_2^b + \rho_4^b$	-0.005	0.006	-0.001	-0.011	-0.002	-0.031	-0.007	0.006
Dependent mean	0.145	0.155	0.147	0.143	0.176	0.188	0.177	0.174
Observations	19701	3009	11252	8449	513614	73129	286180	227434
Game × teams	3960	605	2260	1700	6019	883	3363	2656
FE: DMA × Game	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE: Day of week	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports estimates from equation 5. Standard errors are reported in parentheses clustered at the game-team level, with the following significance indicators: * p<0.1, ** p<0.05 and *** p<0.01. Observations are at the day-subreddit level for columns 1-4 and are individual tests for columns 5-8. Subreddit estimations are weighted by the number of subreddit submissions. Close games are those with a score differential of 7 points or less. Upset games are those where the pre-game spread predicted the losing team would win by 3 points or more. The summed coefficient row reports the sum of the indicated coefficients with significance stars.

Table 4: Close-upset game-day impacts for local and opposition Black quarterbacks

	(1)	(2)	(3)
	Close-upset games		
	Daily DMA anti-Black hate crime rate	Share of submission removals	Standardized white IAT scores
Local Black quarterback games:			
Post × *Only local QB is Black*	0.0951**	0.0530**	0.0340
	(0.0401)	(0.0217)	(0.0242)
Post × *Only local QB is Black* × *Win*	-0.0816	-0.0543**	-0.0752*
	(0.0567)	(0.0271)	(0.0404)
Opposition Black quarterback games:			
Post × *Only opp. QB is Black*	0.0757*	0.0366*	0.0540
	(0.0395)	(0.0196)	(0.0335)
Post × *Only opp. QB is Black* × *Win*	-0.124**	-0.0261	-0.0761*
	(0.0613)	(0.0252)	(0.0456)
Tests of equality:			
p-value of $\rho_2^l = \rho_2^o$	0.68	0.46	0.57
p-value of $\rho_2^l + \rho_4^l = \rho_2^o + \rho_4^o$	0.26	0.58	0.63
Dependent mean	0.179	0.126	0.188
Observations	44715	3009	73129
Game × teams	1401	605	883
Controls: $Post$; $Post \times Win$	Yes	Yes	Yes
FE: DMA × Game	Yes	Yes	Yes
FE: Day of week	Yes	Yes	Yes

Note: This table reports estimates from a modified version of equation 5. Standard errors clustered at the game-team level are reported in parentheses, with the following significance indicators: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Observation levels and weights are: day-DMA level weighted by DMA population in Column 1; day-subreddit level weighted by subreddit submission counts in Column 2; and at the individual test level in Column 3. All columns are restricted to close-upset games with a score differential of 7 points or fewer in which the pre-game spread predicted the losing team to win by 3 or more points. Tests for equality rows report p-values for the indicated coefficient equality tests. Coefficients on $Post$ and $Post \times Win$ are included in all estimations but are not reported for brevity.

Table 5: Game-day impacts by quarterback performance – Close-upset games

	(1)	(2)	(3)	(4)	(5)	(6)
	Close-upset games					
	Daily DMA anti-Black hate crime rate		Standardized white IAT scores		Share of submission removals	
Local Black quarterback games:						
Post × *Only local QB is Black*	0.0987**		0.0224		0.0662**	
	(0.0446)		(0.0281)		(0.0265)	
Post × *Only local QB is Black* × *Win*	-0.0314		-0.0847*		-0.0677**	
	(0.0740)		(0.0504)		(0.0330)	
Post × *Only local QB is Black* × *High EPA local QB*	-0.00747		0.0580		-0.0652	
	(0.0994)		(0.0542)		(0.0422)	
Post × *Only local QB is Black* × *High EPA local QB*× *Win*	-0.120		0.00356		0.0628	
	(0.126)		(0.0835)		(0.0532)	
Opposition Black quarterback games:						
Post × *Only opp. QB is Black*		0.0530		0.0691*		0.0692***
		(0.0495)		(0.0402)		(0.0256)
Post × *Only opp. QB is Black* × *Win*		-0.113		-0.0799		-0.0622**
		(0.0731)		(0.0548)		(0.0294)
Post × *Only opp. QB is Black* × *High EPA opp. QB*		0.0614		-0.0404		-0.0805**
		(0.0821)		(0.0692)		(0.0357)
Post × *Only opp. QB is Black* × *High EPA opp. QB*× *Win*		-0.00980		0.00675		0.0936*
		(0.136)		(0.0969)		(0.0555)
Coefficient sums and significance:						
Win low EPA	0.067	-0.060	-0.062	-0.011	-0.001	0.007
Loss high EPA	0.091	0.114*	0.080*	0.029	0.001	-0.011
Win high EPA	-0.060	-0.009	-0.001	-0.044	-0.004	0.020
Dependent mean	0.196	0.206	0.194	0.196	0.155	0.153
Observations	35885	34580	58653	57440	2347	2349
FE: DMA × Game	Yes	Yes	Yes	Yes	Yes	Yes
FE: Day of week	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports estimates from a modified version of equation 5 to examine heterogeneity by quarterback performance. Effects are estimated separately for local and opposition Black quarterback games in odd and even columns respectively. Standard errors clustered at the game-team level are reported in parentheses, with the following significance indicators: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Observation levels and weights are: day-DMA level weighted by DMA population in Columns 1-2; at the individual test level in Columns 3-4; and at the day-subreddit level weighted by subreddit submission counts in Columns 5-6. All columns are restricted to close-upset games with a score differential of 7 points or fewer in which the pre-game spread predicted the losing team to win by 3 or more points. The summed coefficient row reports the sum of the coefficients active for the type of game indicated with significance stars.

Table 6: Season effects of local Black quarterbacks by team performance

	(1)	(2)	(3)	(4)
	Anti-Black hate crime rate (month-DMA per 100,000)	Log of Google slur index	Share of submission removals	Standardized white IAT scores
Not in playoffs				
Main QB is Black × After season	0.000445	0.0254**	0.00613	0.00213
	(0.00355)	(0.0122)	(0.0141)	(0.00563)
In playoffs only				
Main QB is Black × After season × In playoffs only	-0.00618	-0.0256	-0.00205	-0.00109
	(0.00701)	(0.0181)	(0.0152)	(0.0101)
In championships				
Main QB is Black × After season × In championship	-0.0181**	-0.0624***	0.00335	-0.0421***
	(0.00921)	(0.0222)	(0.0424)	(0.0140)
Coefficient sums and significance:				
Playoffs only level differential ($\mu_T^r + \mu_T^p$)	-0.006	-0.000	0.004	0.001
Championship level differential ($\mu_T^r + \mu_T^c$)	-0.018**	-0.037*	0.009	-0.040***
Dep. mean	0.0530	3.518	0.103	0.168
Observations	52044	15414	2927	3584160
Team × seasons	682	613	244	618
Black QB team × seasons	139	124	64	126
Unique main Black quarterbacks	40	38	21	38
FE: DMA × Season	Yes	Yes	Yes*	Yes
FE: Month × Record	Yes	Yes*	Yes	Yes
FE: Month × DMA	Yes	Yes*	Yes*	Yes
FE: Month × Season	Yes	Yes*	Yes	Yes

Note: Standard errors clustered at the team-season level are reported in parentheses, with the following significance indicators: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Observation levels and weights are: month-DMA level, weighted by DMA population for Column 1; 2 month bin-DMA level weighted by DMA population for Column 2; month-subreddit level weighted by the number of subreddit submissions for Column 3; and at the individual-test level for Column 4. Fixed effects are adjusted due to data formats using 2 month bins (instead of monthly) for the Google slur index and subreddits (instead of DMAs) for the submission removal share. During season ($T = d$) estimates are included in all estimations but are not reported for brevity.

11 Figures

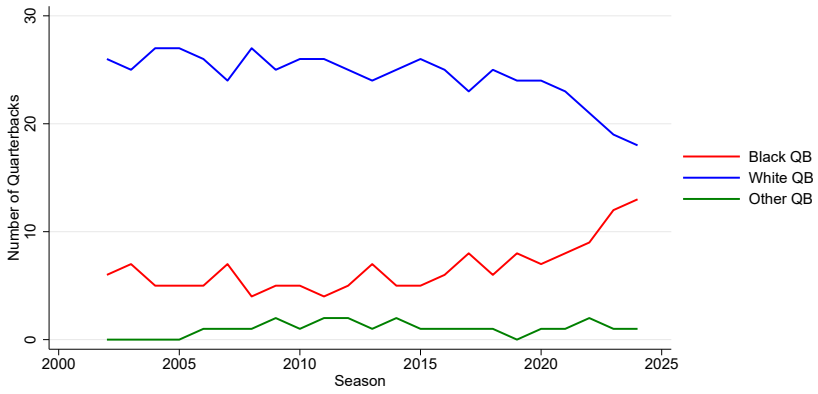

Figure 1: Quarterbacks playing the majority of regular season games in the NFL

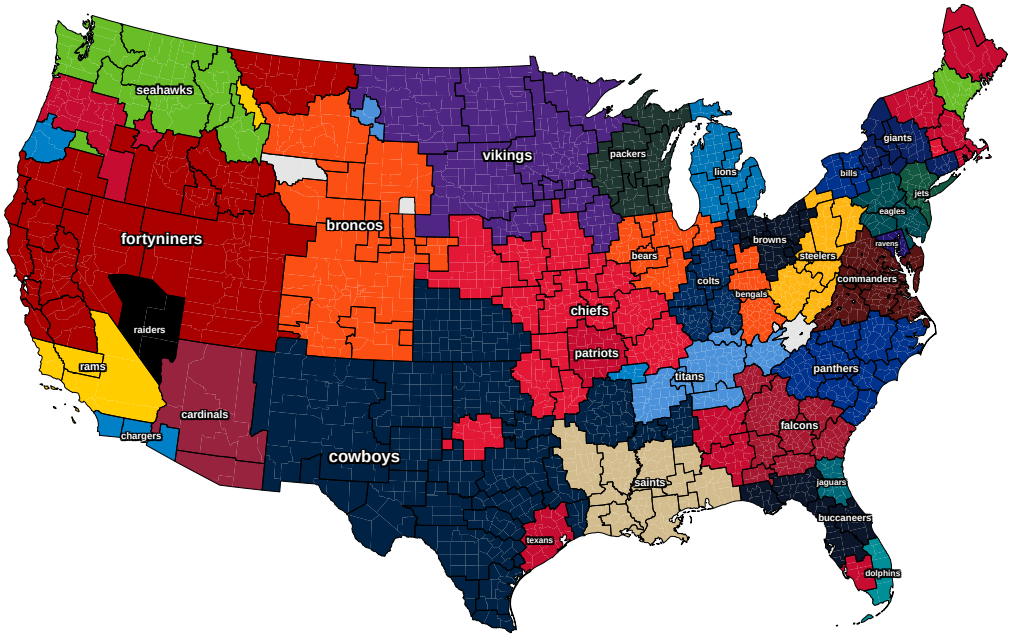

Figure 2: NFL team catchment areas

Note: NFL team catchment areas are defined by matching DMAs to the team they are most engaged with, using our engagement measure that combines a team's local home-market TV ratings with the DMA's Google Search interest for each team over 2004–2025.

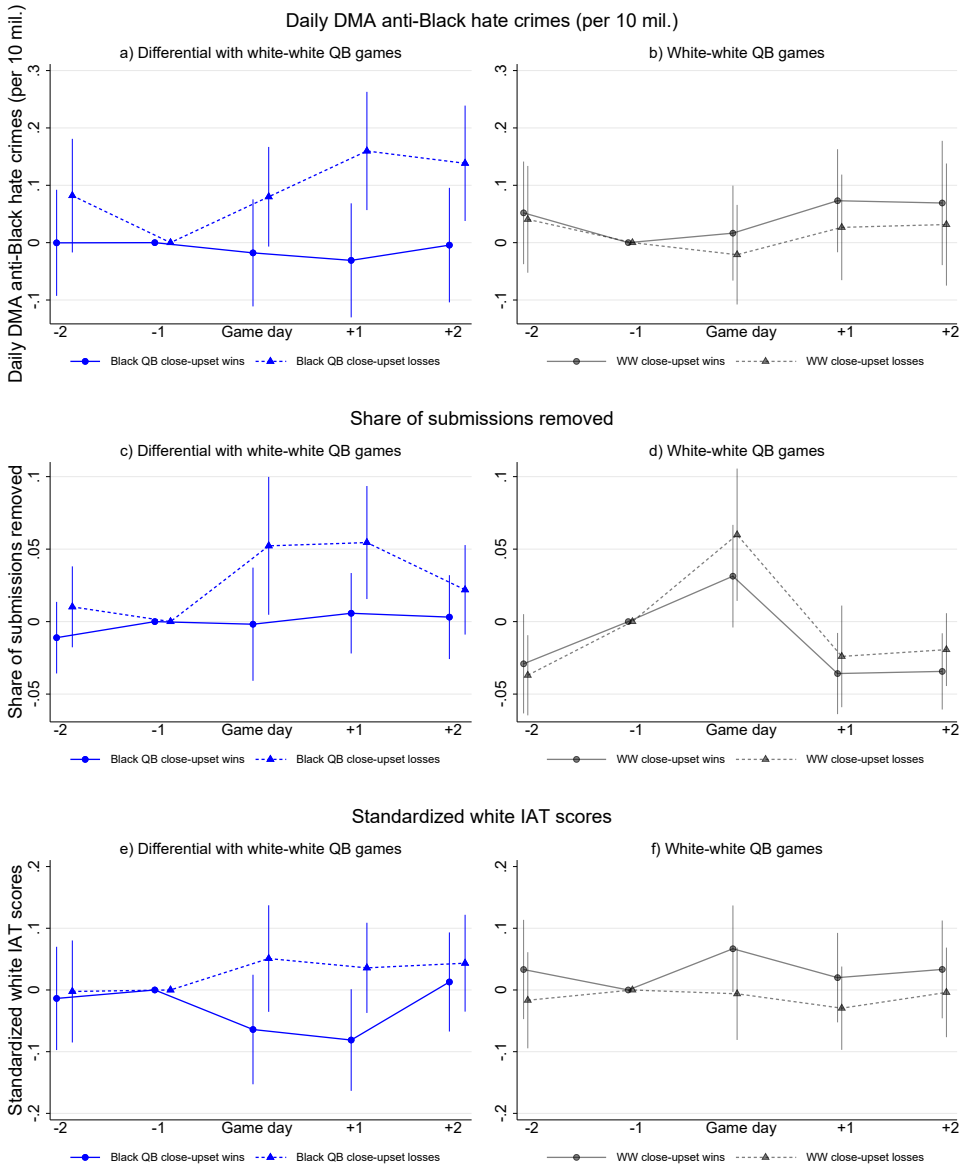

Figure 3: Game-day effects of Black quarterbacks in close-upset games

Note: Figures on the left plot the differential effect of Black quarterbacks for close-upset games, as estimated by a daily specification of equation 5 with the day before game day as the omitted category. Daily ρ_2^b's capture close-upset losses and daily $\rho_2^b + \rho_4^b$'s capture close-upset wins. Figures on the right plot the corresponding effects around white-white quarterback matchups: daily ρ_1^w's for close-upset losses and daily $\rho_1^w + \rho_3^w$'s for close-upset wins. Estimates include DMA $\times$ game and day of the week fixed effects. Close-upset games are defined as those with a score differential of 7 points or fewer in which the pre-game spread predicted the losing team to win by 3 or more points. Observations are at the DMA-day level weighted by DMA population in row 1, subreddit-day level weighted by total daily submissions in row 2, and at the individual level in row 3. All plots show 95% confidence intervals with standard errors clustered at the team-game level.

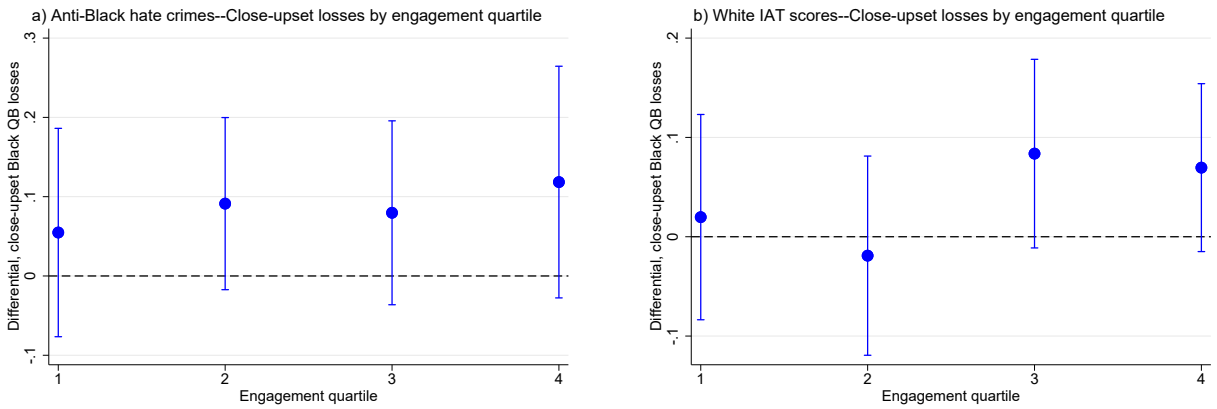

Figure 4: Effects after Black quarterback close-upset losses by engagement quartile

Note: Figures plot the ρ_2^b coefficient from equation 5, estimated on close-upset games, where each estimate uses the sample of DMAs in a different engagement quartile. Estimates include DMA $\times$ game and day of the week fixed effects. Close-upset games are defined as those with a score differential of 7 points or fewer in which the pre-game spread predicted the losing team to win by 3 or more points. Observations are at the DMA-day level weighted by DMA population. Plots show 95% confidence intervals with standard errors clustered at the team-game level.

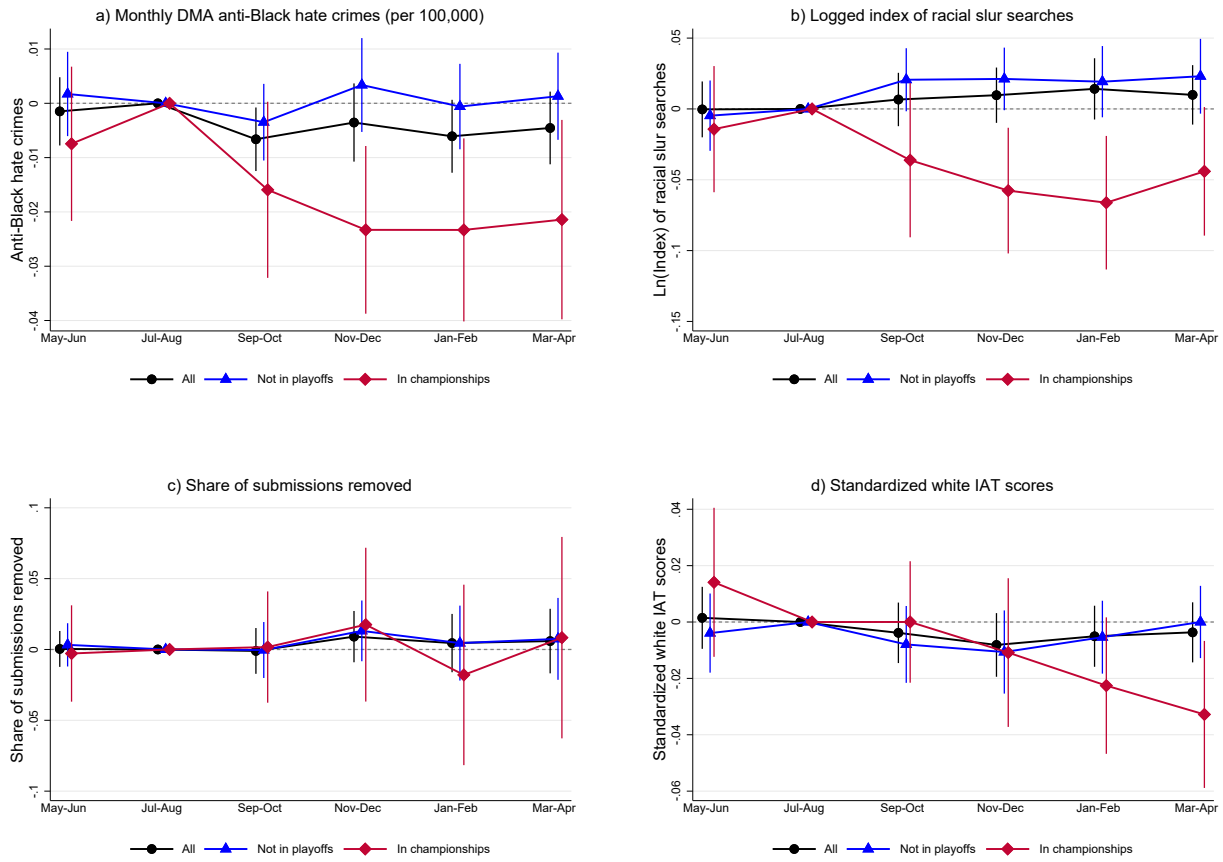

Figure 5: Season effects of local Black quarterbacks by team's season performance

Notes: Figures plot μ_T^r and $\mu_T^r + \mu_T^c$ from equation 6 in blue and red respectively using six two-month time bins with July-August as the omitted category. Aggregate estimates for all performance levels are plotted in black. Estimates control for $DMA \times season$, $month \times DMA$, $month \times record$, and $month \times season$ fixed effects. Observation levels and weights are: month-DMA level, weighted by DMA population for panel (a); 2 month bin-DMA level weighted by DMA population for panel (b); month-subreddit level weighted by the number of subreddit submissions for panel (c); and at the individual-test level for panel (d). All plots show 95% confidence intervals with standard errors clustered at the team-season level.

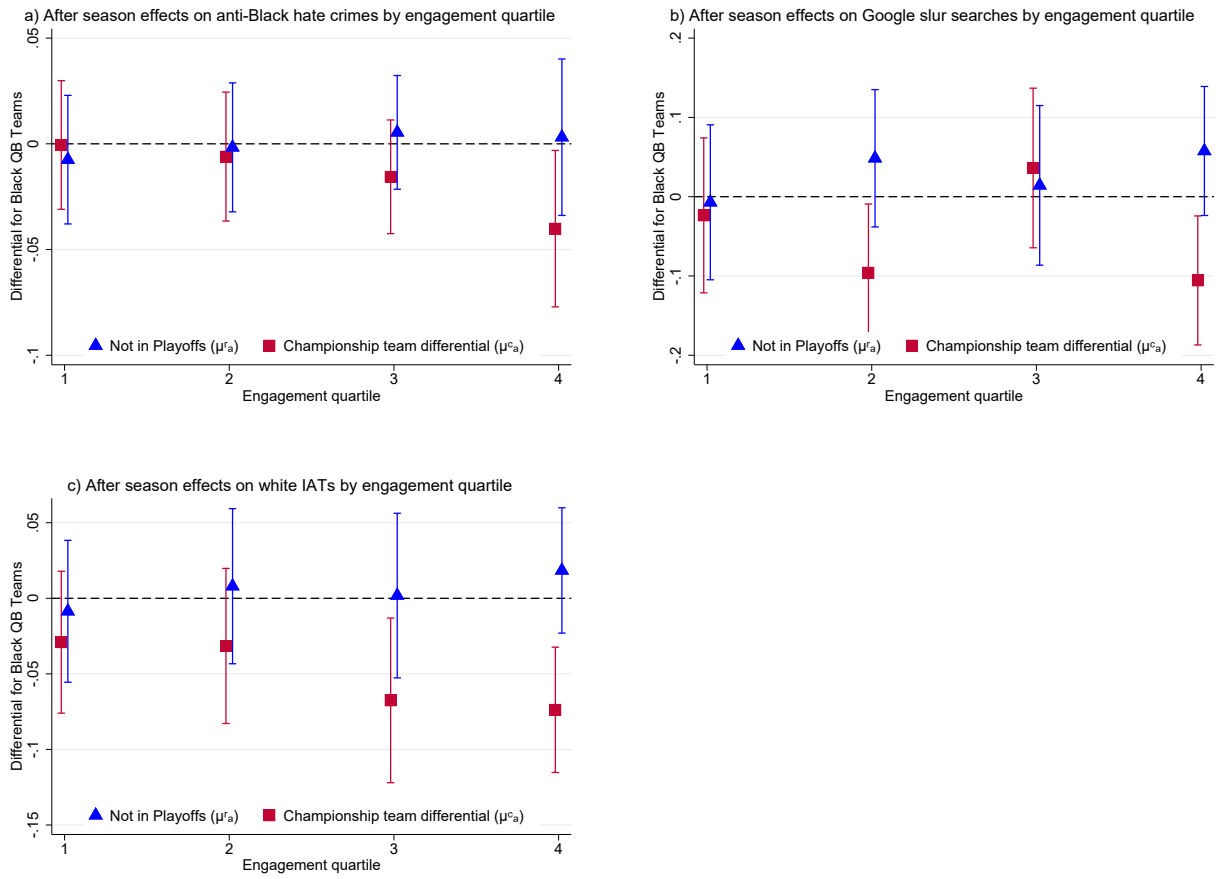

Figure 6: After-season effects of Black quarterbacks by engagement quartile

Note: Figures plot the after-season estimates, μ_a^r and μ_a^c, from equation 6 in blue and red respectively, where each estimate uses the sample of DMAs in a different engagement quartile. Estimates control for $DMA \times season$, $month \times DMA$, $month \times record$, and $month \times season$ fixed effects. Observation levels and weights are: month-DMA level, weighted by DMA population for panel (a); 2 month bin-DMA level weighted by DMA population for panel (b); and at the individual-test level for panel (c). All plots show 95% confidence intervals with standard errors clustered at the team-season level.

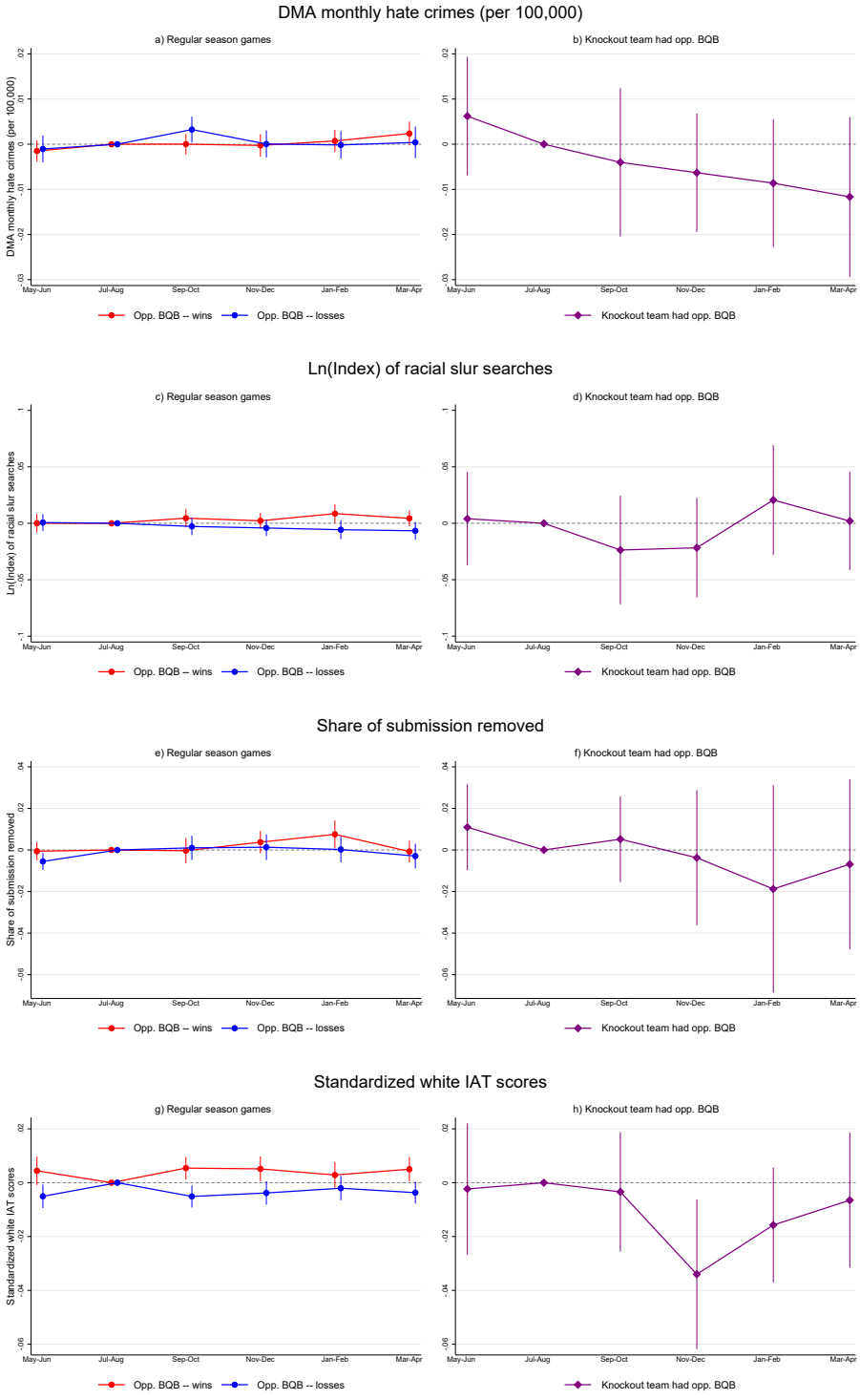

Figure 7: Season effects of opposition Black quarterbacks

Notes: For all figures, the sample is limited to team-seasons where the main quarterback was white. Figures on the left plot δ_T^E from equation 7: the differential effect of an additional regular season game played against opposition Black quarterbacks by game outcome using six two-month time bins with July-August as the omitted category. Figures on the right plot the differential effect of being knocked out of the playoffs by a team with a Black quarterback. Estimates control for $DMA \times season$, $month \times DMA$, $month \times record$, and $month \times season$ fixed effects. Observation levels and weights are: month-DMA level, weighted by DMA population in row 1; 2 month bin-DMA level weighted by DMA population in row 2; month-subreddit level weighted by the number of subreddit submissions in row 3; and at the individual-test level in row 4. All plots show 95% confidence intervals with standard errors clustered at the team-season level.

Online Appendix

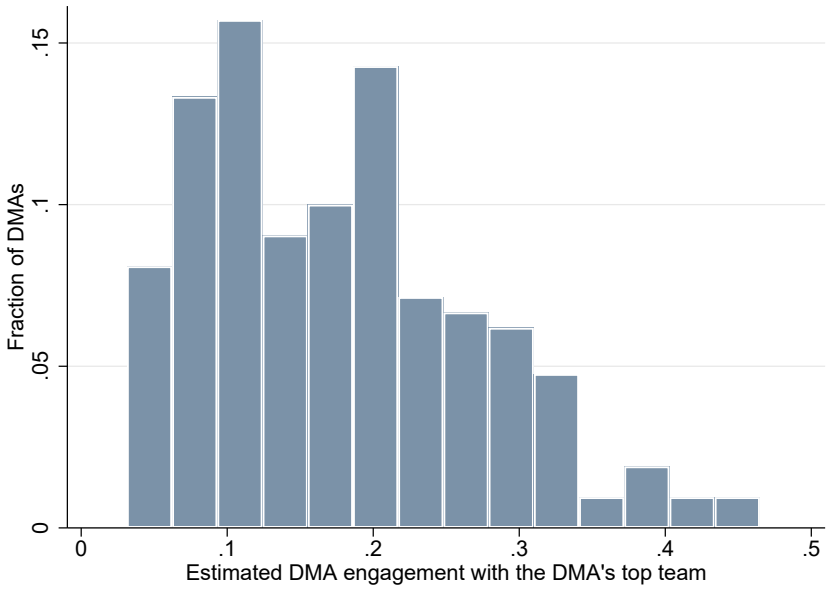

Figure A1: Distribution of DMA engagement levels with their top team

Notes: Engagement levels are calculated by combining a DMA's Google searches for teams with team ratings in their home DMA as described in section 4. The figure plots the engagement level of the DMA with the team it is most engaged with.

Table A1: Robustness to alternative outcome measures of daily DMA anti-Black hate crimes

	(1)	(2)	(3)	(4)	(5)	(6)
	Daily DMA Rate (incidents per 10 mil.)	Extensive indicator (1 if has incidents)	Zero adjusted log log(incidents+1)	Linear raw (incident count)	Poisson raw (incident count)	Unweighted daily DMA rate (incidents per 10 mil.)
Panel a: All games						
Post × *Any Black QB*	0.0364***	0.00436**	0.00355***	0.00571***	0.176***	0.0718***
	(0.0130)	(0.00171)	(0.00129)	(0.00202)	(0.0646)	(0.0207)
Post × *Any Black QB* × *Win*	-0.0207	-0.00213	-0.00197	-0.00344	-0.102	-0.0279
	(0.0179)	(0.00231)	(0.00172)	(0.00269)	(0.0904)	(0.0298)
Dependent mean	0.200	0.027	0.020	0.030	0.288	0.165
Observations	306625	306625	306625	306625	31755	306625
Panel b: Close-upset games						
Post × *Any Black QB*	0.0863***	0.0107***	0.00798***	0.0121***	0.464***	0.171***
	(0.0324)	(0.00404)	(0.00298)	(0.00455)	(0.172)	(0.0596)
Post × *Any Black QB* × *Win*	-0.104**	-0.0127**	-0.00958**	-0.0148**	-0.544**	-0.199**
	(0.0469)	(0.00620)	(0.00455)	(0.00694)	(0.233)	(0.0773)
Dependent mean	0.202	0.027	0.020	0.030	0.281	0.179
Observations	44715	44715	44715	44715	4710	44715
Panel c: Neither close nor upset games						
Post × *Any Black QB*	0.0191	0.00257	0.00209	0.00324	0.0783	0.0283
	(0.0200)	(0.00266)	(0.00203)	(0.00323)	(0.0973)	(0.0335)
Post × *Any Black QB* × *Win*	-0.00944	-0.000909	-0.000963	-0.00185	-0.0380	0.00982
	(0.0273)	(0.00349)	(0.00263)	(0.00415)	(0.139)	(0.0436)
Dependent mean	0.199	0.027	0.020	0.030	0.291	0.160
Observations	134220	134220	134220	134220	13800	134220

Note: Column headers report the alternative measurement or specification used to estimate impacts on daily-DMA anti-Black hate crimes as in 5. Coefficients on $Post$ and $Post \times Win$ are included in all estimations but are not reported for brevity. Standard errors are reported in parentheses clustered at the game-team level, with the following significance indicators: * p<0.1, ** p<0.05 and *** p<0.01. Observations are at the day-DMA level, weighted by DMA population for columns 1-5. Panel a uses the full sample of games. Panel b limits the sample to close-upset games with a score differential of 7 points or less and a pre-game spread predicted the losing team would win by 3 points or more. Panel c limits the sample to games that are neither close nor upset games.

Table A2: Robustness of impacts on anti-Black hate crimes to fixed effects

	(1)	(2)	(3)	(4)
	Daily DMA anti-Black hate crime rate (per 10 mil.)			
Panel a: All games				
Post × *Any Black QB*	0.03641***	0.03668***	0.03668***	0.03668***
	(0.01301)	(0.01301)	(0.01301)	(0.01301)
Post × *Any Black QB* × *Win*	-0.02067	-0.02062	-0.02063	-0.02062
	(0.01794)	(0.01795)	(0.01795)	(0.01795)
Dependent mean	0.200	0.200	0.200	0.200
Observations	306625	306625	306625	306625
Panel b: Close-upset games				
Post × *Any Black QB*	0.08628***	0.08652***	0.08652***	0.08652***
	(0.03237)	(0.03225)	(0.03225)	(0.03233)
Post × *Any Black QB* × *Win*	-0.10421**	-0.10320**	-0.10320**	-0.10320**
	(0.04690)	(0.04686)	(0.04686)	(0.04697)
Dependent mean	0.202	0.202	0.202	0.202
Observations	44715	44715	44715	44715
Panel c: Neither games				
Post × *Any Black QB*	0.01915	0.01966	0.01967	0.01966
	(0.01999)	(0.01999)	(0.01999)	(0.02001)
Post × *Any Black QB* × *Win*	-0.00944	-0.00940	-0.00941	-0.00941
	(0.02735)	(0.02735)	(0.02735)	(0.02737)
Dependent mean	0.199	0.199	0.199	0.199
Observations	134220	134220	134220	134220
FE: DMA × Game-team	Yes	Yes	No	No
FE: Day of the week	Yes	No	No	No
FE: DMA	.	.	No	Yes
FE: Game-team	.	.	Yes	Yes

Note: Columns estimate equation 5 with alternative combinations of fixed effects. Coefficients on $Post$ and $Post \times Win$ are included in all estimations but are not reported for brevity. Standard errors are reported in parentheses clustered at the game-team level, with the following significance indicators: * p<0.1, ** p<0.05 and *** p<0.01. Observations are at the day-DMA level, weighted by DMA population. Panel a uses the full sample of games. Panel b limits the sample to close-upset games with a score differential of 7 points or less and a pre-game spread predicted the losing team would win by 3 points or more. Panel c limits the sample to games that are neither close nor upset games.

Table A3: Game-day impacts by type of anti-Black hate crime – Close-upset games

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Perpetrator count			Crime type			
	Unknown	Single	Multiple	Against Persons	Against Property	Against Society	Not specified
White-white games:							
Post	-0.0176	0.0109	-0.00536	0.00780	-0.0202	0.00466**	0.000590
	(0.0173)	(0.0295)	(0.0127)	(0.0320)	(0.0198)	(0.00222)	(0.000705)
Post × *Win*	0.0225	0.000290	0.0120	0.00379	0.0342**	-0.00575**	-0.0000963
	(0.0165)	(0.0222)	(0.0111)	(0.0242)	(0.0170)	(0.00284)	(0.00126)
Any Black QB games:							
Post × *Any Black QB*	0.0458**	0.0405*	-0.0000577	0.0481*	0.0465**	-0.00512	-0.00269
	(0.0184)	(0.0245)	(0.0116)	(0.0255)	(0.0190)	(0.00321)	(0.00200)
Post × *Any Black QB* × *Win*	-0.0565**	-0.0264	-0.0213	-0.0490	-0.0617**	0.00740	0.00186
	(0.0251)	(0.0352)	(0.0173)	(0.0379)	(0.0258)	(0.00454)	(0.00217)
Dependent mean	0.045	0.103	0.031	0.128	0.051	0.003	0.001
Observations	44715	44715	44715	44715	44715	44715	44715
FE: DMA × Game	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE: Day of week	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports estimates from equation 5 on close-upset games. Column headers indicate the subset of anti-Black hate crimes used as an outcome variable (in daily DMA rate per 10 million). Standard errors are reported in parentheses clustered at the game-team level, with the following significance indicators: * p<0.1, ** p<0.05 and *** p<0.01. Observations are at the day-DMA level weighted by the DMA population. Close-upset games are those with a score differential of 7 points or less and where the pre-game spread predicted the losing team would win by 3 points or more.

Table A4: Game-day impacts on hate crimes targeting other groups

	(1) Anti-Black	(2) Anti-white	(3) Anti-Hispanic	(4) Anti-other race
White-white games:				
Post	-0.0121	-0.0216	0.0126	-0.00224
	(0.0368)	(0.0227)	(0.0175)	(0.0230)
Post × *Win*	0.0348	0.000238	0.0300**	-0.00334
	(0.0297)	(0.0160)	(0.0135)	(0.0191)
Any Black quarterback games:				
Post × *Any Black QB*	0.0863***	0.00797	0.00525	-0.000923
	(0.0324)	(0.0178)	(0.0138)	(0.0199)
Post × *Any Black QB* × *Win*	-0.104**	0.00741	-0.0450**	0.0230
	(0.0469)	(0.0253)	(0.0212)	(0.0273)
$\rho_2^b + \rho_4^b$	-0.018	0.015	-0.040**	0.022
Dependent mean	0.202	0.061	0.044	0.066
Observations	44715	44715	44715	44715
Game × teams	1401	1401	1401	1401
FE: DMA × Game	Yes	Yes	Yes	Yes
FE: Day of week	Yes	Yes	Yes	Yes

Note: This table reports estimates of equation 5 on other types of hate crimes for close-upset games. Column headers indicate the type of hate crime examined (in daily DMA rate per 10 million). Standard errors are reported in parentheses clustered at the game-team level, with the following significance indicators: * p<0.1, ** p<0.05 and *** p<0.01. Observations are at the day-DMA level weighted by the DMA population. Close-upset games are those with a score differential of 7 points or less and where the pre-game spread predicted the losing team would win by 3 points or more.

Table A5: Game-day impacts on Reddit submission numbers, Reddit anti-Black speech and IAT test taking – Close-upset games

	(1)	(2)	(3)
	NFL subreddit daily submission count	NFL subreddit anti-Black hate speech rate	Number of white IAT test takers
White-white games:			
Post	90.59***	0.00000669	0.169
	(11.23)	(0.0000360)	(0.208)
Post × *Win*	16.93**	-0.000206	-0.230
	(6.748)	(0.000182)	(0.192)
Any Black quarterback games:			
Post × *Any Black QB*	-4.729	0.0000999	-0.210
	(5.651)	(0.000126)	(0.179)
Post × *Any Black QB* × *Win*	12.15	0.000127	0.237
	(11.93)	(0.000223)	(0.256)
Coefficient sums and significance:			
$\rho_2^b + \rho_4^b$	7.42	0.000	0.027
Dependent mean	68.543	0.000	2.290
Observations	3025	2997	44715
Game × teams	609	602	1401
FE: DMA × Game	Yes	Yes	Yes
FE: Day of week	Yes	Yes	Yes

Note: This table reports estimates of equation 5 for close-upset games on appendix outcomes: the number of daily Reddit submissions, the rate of anti-Black hate speech in observed Reddit submissions, and the number of IAT test takers. Standard errors are reported in parentheses clustered at the game-team level, with the following significance indicators: * p<0.1, ** p<0.05 and *** p<0.01. Observation levels and weights are: day-subreddit level weighted by subreddit submission counts in Column 1; day-subreddit level in Column 2; and day-DMA level in Column 3. Close-upset games are those with a score differential of 7 points or less and where the pre-game spread predicted the losing team would win by 3 points or more. The summed coefficient row reports the sum of the indicated coefficients with significance stars.

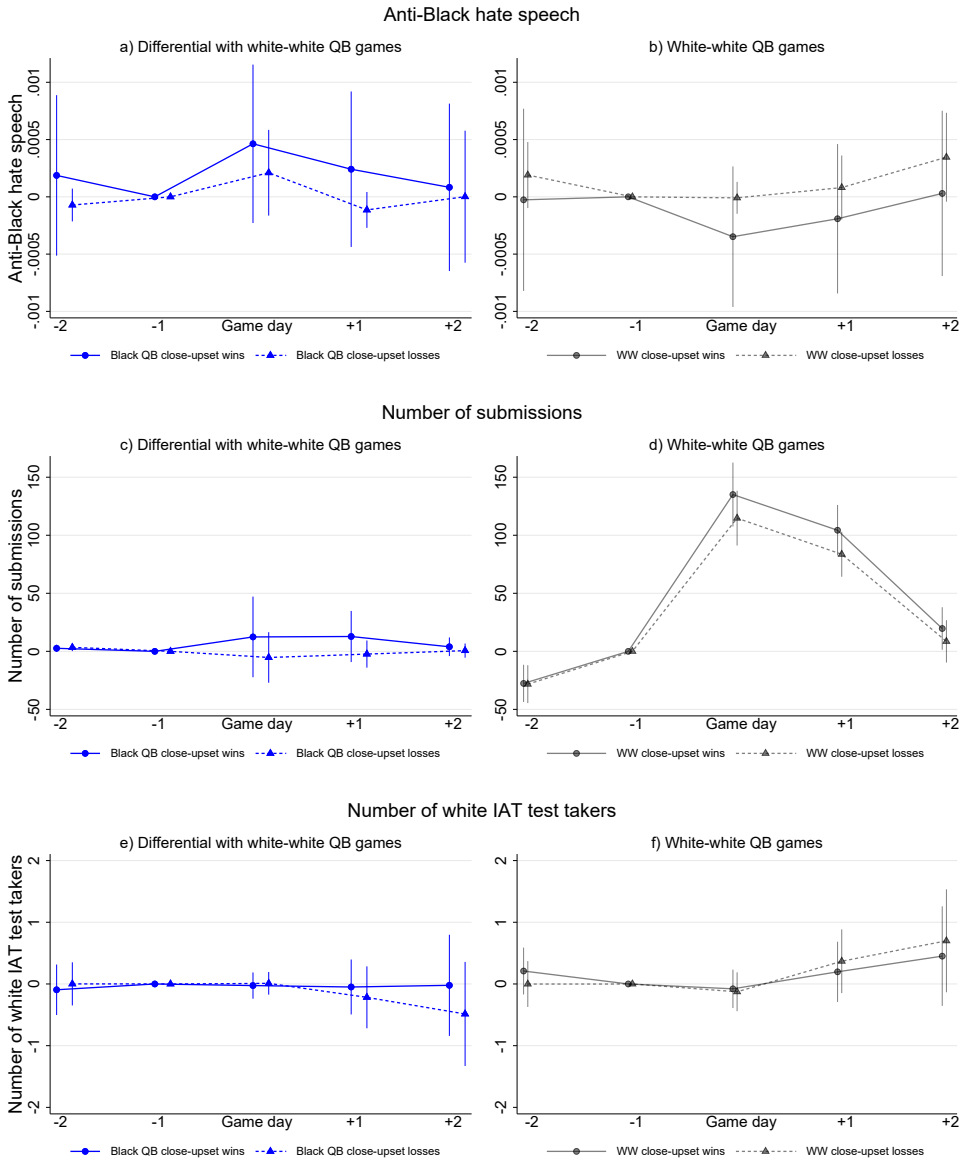

Figure A2: Game-day effects of local Black quarterbacks

Note: Figures on the left plot the differential effect of Black quarterbacks for close-upset games, as estimated by a daily specification of equation 5 with the day before game day as the omitted category. Daily ρ_2^b's capture close-upset losses and daily $\rho_2^b + \rho_4^b$'s capture close-upset wins. Figures on the right plot the corresponding effects around white-white quarterback matchups: daily ρ_1^w's for close-upset losses and daily $\rho_1^w + \rho_3^w$'s for close-upset wins. Estimates include DMA $\times$ game and day of the week fixed effects. Close-upset games are defined as those with a score differential of 7 points or fewer in which the pre-game spread predicted the losing team to win by 3 or more points. Observations are at the subreddit-day level weighted by total daily submissions in row 1, the subreddit-day level in row 2, and day-DMA level in row 3. All plots show 95% confidence intervals with standard errors clustered at the team-game level.

Table A6: Game-day impacts controlling for quarterback play style – Close-upset games

	(1)	(2)	(3)
	Close-upset games		
	Daily DMA anti-Black hate crime rate	Share of submission removals	Standardized white IAT scores
White-white games:			
Post	-0.0203	0.0342	0.00696
	(0.0381)	(0.0241)	(0.0295)
Post × *Win*	0.0327	-0.0220	0.0169
	(0.0317)	(0.0166)	(0.0231)
Any Black quarterback games:			
Post × *Any Black QB*	0.0725**	0.0494***	0.0405
	(0.0337)	(0.0178)	(0.0251)
Post × *Any Black QB* × *Win*	-0.0922*	-0.0410*	-0.0753**
	(0.0493)	(0.0221)	(0.0359)
Any mobility-integrated QB games:			
Post × *Any mob QB*	0.0587	-0.0117	0.00494
	(0.0402)	(0.0170)	(0.0313)
Post × *Any mob QB* × *Win*	-0.00852	0.00270	-0.0135
	(0.0570)	(0.0223)	(0.0443)
Dependent mean	0.197	2.796	2.435
Observations	43340	2879	70527
FE: DMA × Game	Yes	Yes	Yes
FE: Day of week	Yes	Yes	Yes

Note: This table reports estimates from a modified version of equation 5 that includes controls for quarterback play style. Standard errors clustered at the game-team level are reported in parentheses, with the following significance indicators: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Observation levels and weights are: day-DMA level weighted by DMA population in Column 1; day-subreddit level weighted by subreddit submission counts in Column 2; and at the individual test level in Column 3. All columns are restricted to close-upset games with a score differential of 7 points or fewer in which the pre-game spread that predicted the losing team to win by 3 or more points.

Table A7: Changes in a team's catchment area's hate crime rate do not predict team's quarterback race

	(1)	(2)	(3)	(4)
	Main quarterback is Black			
Std. change in hate crimes from s-2 to s-1	-0.0162	0.00152	0.000746	-0.000931
	(0.0182)	(0.0154)	(0.0152)	(0.0139)
s-1 season QB was Black		0.472***	0.382***	
		(0.0430)	(0.0485)	
s-2 season QB was Black		0.120***	0.103**	
		(0.0433)	(0.0429)	
s-1 season made championships			-0.0614	
			(0.0479)	
s-1 season made playoffs only			-0.0408	
			(0.0355)	
s-1 season made championships with Black QB			0.433***	
			(0.114)	
s-1 season made playoffs only with Black QB			0.211***	
			(0.0787)	
Observations	586	586	586	437
FE: Team	Yes	Yes	Yes	Yes
FE: Season	Yes	Yes	Yes	Yes
Subsample: Had a white QB in s-1 and s-2				Yes

Note: This table regresses the indicator for the team's main quarterback being Black on the two-year change in a team's catchment areas' annual anti-Black hate crime rate (standardized). Standard errors are reported in parentheses, with the following significance indicators: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Observations are at the season-team level.

Table A8: Robustness to alternative outcome measures of monthly DMA anti-Black hate crime

	(1) **Monthly DMA rate (incidents per 100k)**	(2) Extensive indicator (1 if has incidents)	(3) Zero adjusted log log(incidents+1)	(4) Linear raw (incident count)	(5) Poisson raw (incident count)	(6) Unweighted monthly DMA rate (incidents per 100k)
Not in playoffs						
Main QB is Black × After season	0.000445	-0.0135	-0.0116	-0.00736	0.0250	-0.0106**
	(0.00355)	(0.0170)	(0.0167)	(0.0574)	(0.0557)	(0.00535)
In playoffs only						
Main QB is Black × After season × In playoffs only	-0.00618	-0.0109	-0.00710	-0.101	-0.0620	-0.00234
	(0.00701)	(0.0273)	(0.0307)	(0.108)	(0.0994)	(0.00917)
In championships						
Main QB is Black × After season × In championship	-0.0181**	-0.0318	-0.0564	-0.211	-0.285**	-0.00958
	(0.00921)	(0.0412)	(0.0441)	(0.139)	(0.136)	(0.0133)
Coefficient sums and significance:						
Championship level differential ($\mu_T^c + \mu_T^c$)	-0.018**	-0.045	-0.068*	-0.219*	-0.260**	-0.020*
Dependent mean	0.063	0.335	0.374	0.939	1.265	0.053
Observations	52044	52044	52044	52044	38629	52044
FE: DMA × Season	Yes	Yes	Yes	Yes	Yes	Yes
FE: Month × Record	Yes	Yes	Yes	Yes	Yes	Yes
FE: Month × DMA	Yes	Yes	Yes	Yes	Yes	Yes
FE: Month × Season	Yes	Yes	Yes	Yes	Yes	Yes

Note: Column headers report the alternative measurement or specification used to estimate impacts on monthly-DMA anti-Black hate crimes as in 6. Standard errors clustered at the team-season level are reported in parentheses, with the following significance indicators: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Observations are at the month-DMA level, weighted by DMA population for columns 1-5. During season ($T = d$) estimates are included in all estimations but are not reported for brevity.

Table A9: Season effects of local Black quarterbacks by team performance – robustness to alternative specifications

	(1)	(2)	(3)	(4)	(5)	(6)
	Monthly DMA anti-Black hate crime rate (per 100,000)					
Not in playoffs						
After season		-0.00933***	-0.00780***			
		(0.00219)	(0.00188)			
Main QB is Black × After season	0.000445	0.00121	-0.000514	0.00121	0.00181	0.000146
	(0.00355)	(0.00422)	(0.00322)	(0.00422)	(0.00358)	(0.00426)
In playoffs only						
After season × In playoffs only		0.00263	-0.000630			
		(0.00366)	(0.00328)			
Main QB is Black × After season × In playoffs only	-0.00618	-0.00342	0.000828	-0.00342	-0.00769	-0.00122
	(0.00701)	(0.00735)	(0.00570)	(0.00735)	(0.00659)	(0.00761)
In championships						
After season × In championship		0.00924**	0.00205			
		(0.00454)	(0.00317)			
Main QB is Black × After season × In championship	-0.0181**	-0.0215**	-0.0110*	-0.0215**	-0.0179**	-0.0229**
	(0.00921)	(0.00921)	(0.00641)	(0.00922)	(0.00810)	(0.00980)
Coefficient sums and significance:						
Championship level differential	-0.018**	-0.020**	-0.011**	-0.020**	-0.016**	-0.023**
Observations	52044	52044	52044	52044	52044	52044
FE: DMA × season	Yes	Yes	No	Yes	Yes	Yes
FE: Month × season	Yes	No	No	No	Yes	No
FE: DMA × month	Yes	No	No	No	No	Yes
FE: Record × month	Yes	No	No	Yes	Yes	Yes
FE: DMA	.	.	Yes	.	.	.
FE: Season	.	.	Yes	.	.	.

Note: Columns estimate alternative versions of equation 6 with different combinations of fixed effects. During season ($T = d$) estimates are included in all estimations but are not reported for brevity. Standard errors are reported in parentheses clustered at the season-team level, with the following significance indicators: * p<0.1, ** p<0.05 and *** p<0.01. Observations are at the month-DMA level, weighted by DMA population.

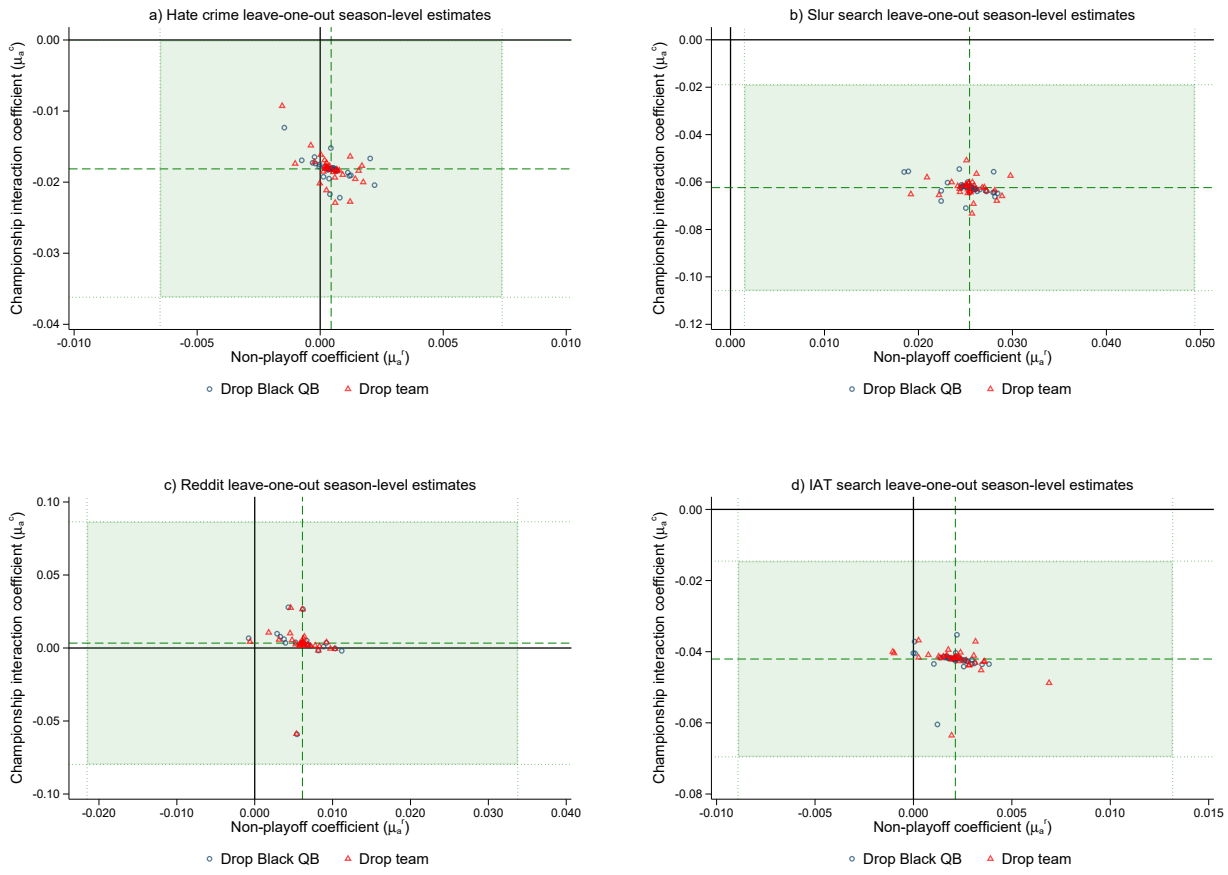

Figure A3: Estimate robustness to leaving out observations for a single team or Black quarterback

Note: Each point plotted in the figure reports the after-season μ_a^r and μ_a^c estimates from equation 6 estimated with the omission of one team (red triangles), or one Black quarterback (blue circles). The dashed green lines show the full-sample estimates with the shaded green area highlighting the 95% confidence intervals, with standard errors clustered at the team-season level, around the full-sample estimates.

Table A10: Season impacts by type of anti-Black hate crime

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Perpetrator count			Crime type			
	Unknown	Single	Multiple	Against Persons	Against Property	Against Society	Not specified
Main QB is Black × After season	-0.00382*	0.000443	0.00371	0.00275	-0.00294	-0.0000978	0.000232
	(0.00206)	(0.00119)	(0.00256)	(0.00276)	(0.00196)	(0.000353)	(0.000168)
Main QB is Black × After season × In playoffs only	-0.000982	-0.00109	-0.00348	-0.00485	-0.00149	-0.0000979	-0.000404*
	(0.00361)	(0.00193)	(0.00485)	(0.00541)	(0.00350)	(0.000968)	(0.000231)
Main QB is Black × After season × In championship	-0.00344	-0.00321	-0.0113*	-0.00981	-0.00607	-0.000633	-0.000298
	(0.00404)	(0.00280)	(0.00653)	(0.00740)	(0.00411)	(0.00108)	(0.000267)
Dependent mean	0.013	0.008	0.031	0.035	0.017	0.001	0.000
Observations	50604	50604	50604	50604	50604	50604	50604
Team × seasons	635	635	635	635	635	635	635
FE: DMA × Season	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE: Month × Record	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE: Month × DMA	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE: Month × Season	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports estimates from equation 6. Column headers indicate the subset of anti-Black hate crimes used as an outcome variable (in monthly DMA rate per 100,000). Standard errors are reported in parentheses clustered at the season-team level, with the following significance indicators: * p<0.1, ** p<0.05 and *** p<0.01. Observations are at the month-DMA level weighted by the DMA population. During season ($T = d$) estimates are included in all estimations but are not reported for brevity.

Table A11: Season effects of local Black quarterbacks by team performance on other hate crimes

	(1) Anti-Black	(2) Anti-white	(3) Anti-Hispanic	(4) Anti-other race
Not in playoffs				
Main QB is Black × After season	0.000445	-0.000484	-0.000623	-0.00320
	(0.00355)	(0.00198)	(0.00149)	(0.00239)
In playoffs only				
Main QB is Black × After season × In playoffs only	-0.00618	-0.00130	0.00264	0.000234
	(0.00701)	(0.00387)	(0.00267)	(0.00391)
In championships				
Main QB is Black × After season × In championship	-0.0181**	0.000944	0.00214	0.00174
	(0.00921)	(0.00456)	(0.00319)	(0.00500)
Dependent mean	0.0530	0.0230	0.0100	0.0150
Observations	52044	52044	52044	52044
FE: DMA × Season	Yes	Yes	Yes	Yes
FE: Month × Record	Yes	Yes	Yes	Yes
FE: Month × DMA	Yes	Yes	Yes	Yes
FE: Month × Season	Yes	Yes	Yes	Yes

Note: This table reports estimates of equation 6 on other types of hate crimes. Column headers indicate the type of hate crime examined (in monthly DMA rate per 100,000). Standard errors are reported in parentheses clustered at the season-team level, with the following significance indicators: * p<0.1, ** p<0.05 and *** p<0.01. Observations are at the month-DMA level weighted by the DMA population. During season ($T = d$) estimates are included in all estimations but are not reported for brevity.

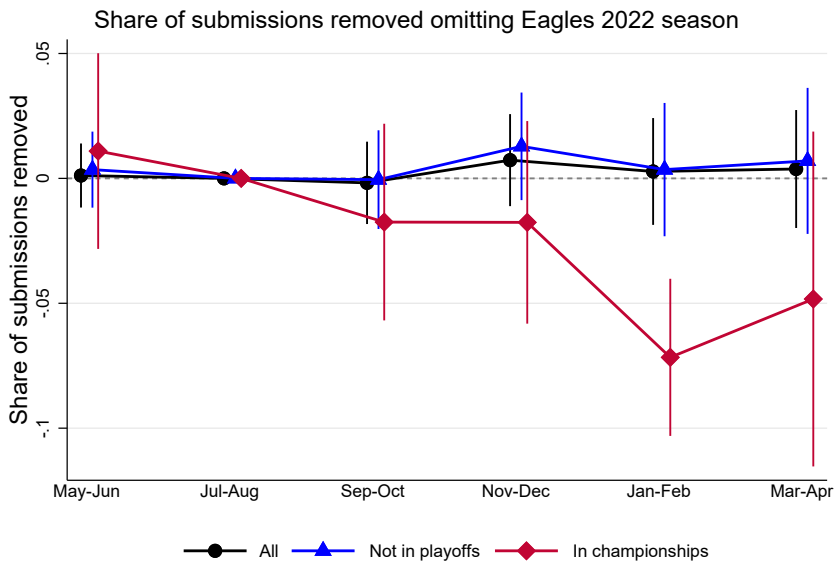

Figure A4: Season effects of local Black quarterbacks on subreddit moderation— Eagle's 2022 season omitted

Notes: This figure replicates the analysis reported in Figure 5 omitting observations pertaining to the Eagle's 2022 season. μ_T^r and $\mu_T^r + \mu_T^c$ from equation 6 are plotted in blue and red respectively using six two-month time bins with July-August as the omitted category. Aggregate estimates for all performance levels are plotted in black. Estimates control for $DMA \times season$, $month \times DMA$, $month \times record$, and $month \times season$ fixed effects. Observation levels and weights are at the month-subreddit level weighted by the number of subreddit submissions. The plot show 95% confidence intervals with standard errors clustered at the team-season level.

Table A12: Season effects of local Black quarterbacks by team performance – appendix outcomes

	(1)	(2)	(3)
	NFL subreddit monthly submission count	NFL subreddit anti-Black hate speech rate	Number of white IAT test takers
Not in playoffs			
Main QB is Black × *After season*	-88.40	0.0000900	-0.0795
	(204.3)	(0.0000707)	(0.229)
In playoffs only			
Main QB is Black × *After season* × *In playoffs only*	-97.96	-0.000149	-0.261
	(200.9)	(0.0000992)	(0.479)
In championships			
Main QB is Black × *After season* × *In championship*	60.63	0.000329	0.995*
	(354.4)	(0.000311)	(0.588)
Coefficient sums and significance:			
Championship level differential	-27.8	0.000	0.916
Dependent mean	1102.5	0	4.610
Observations	2927	2916	816024
Team × seasons	244	244	713
FE: DMA × Season	Yes	Yes	Yes
FE: Month × Record	Yes	Yes	Yes
FE: Month × DMA	Yes	Yes	Yes
FE: Month × Season	Yes	Yes	Yes

Note: This table reports estimates of equation 6 on appendix outcomes: the number of monthly Reddit submissions, the rate of anti-Black hate speech in observed Reddit submissions, and the number of IAT test takers. Standard errors are reported in parentheses clustered at the season-team level, with the following significance indicators: * p<0.1, ** p<0.05 and *** p<0.01. Observation levels and weights are: month-subreddit level in Column 1; month-subreddit level weighted by subreddit submission counts in Column 2; and month-DMA level in Column 3. The summed coefficient row reports the sum of $\mu_a^r + \mu_a^c$ with significance stars. During season $(T = d)$ estimates are included in all estimations but are not reported for brevity.

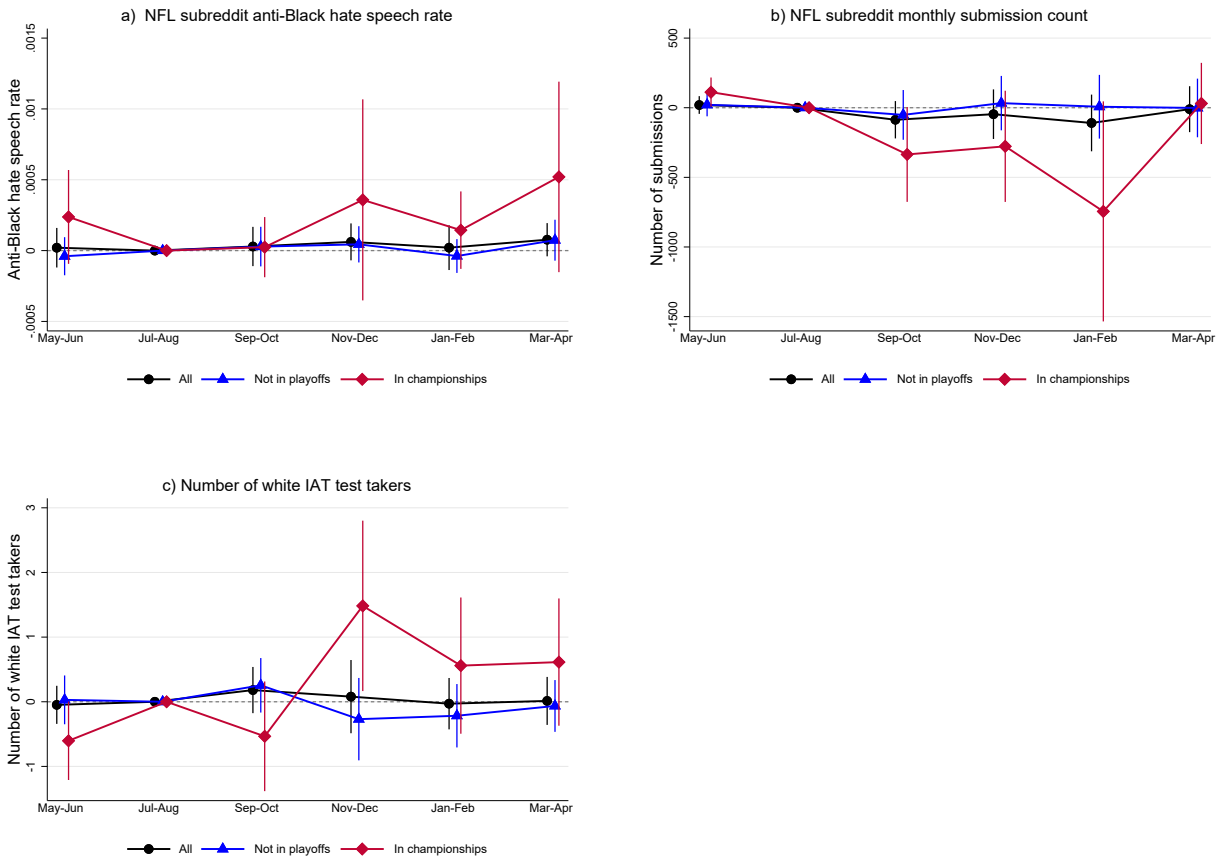

Figure A5: Season effects of local Black quarterbacks by performance level – appendix outcomes

Figures plot μ_T^r and $\mu_T^r + \mu_T^c$ from equation 6 in blue and red respectively using six two-month time bins with July-August as the omitted category. Aggregate estimates for all performance levels are plotted in black. Estimates control for $DMA \times season$, $month \times DMA$, $month \times record$, and $month \times season$ fixed effects. Observation levels and weights are: month-subreddit level weighted by subreddit submission counts in panel (a); month-subreddit level in panel (b); and month-DMA level in panel (c). All plots show 95% confidence intervals with standard errors clustered at the team-season level.

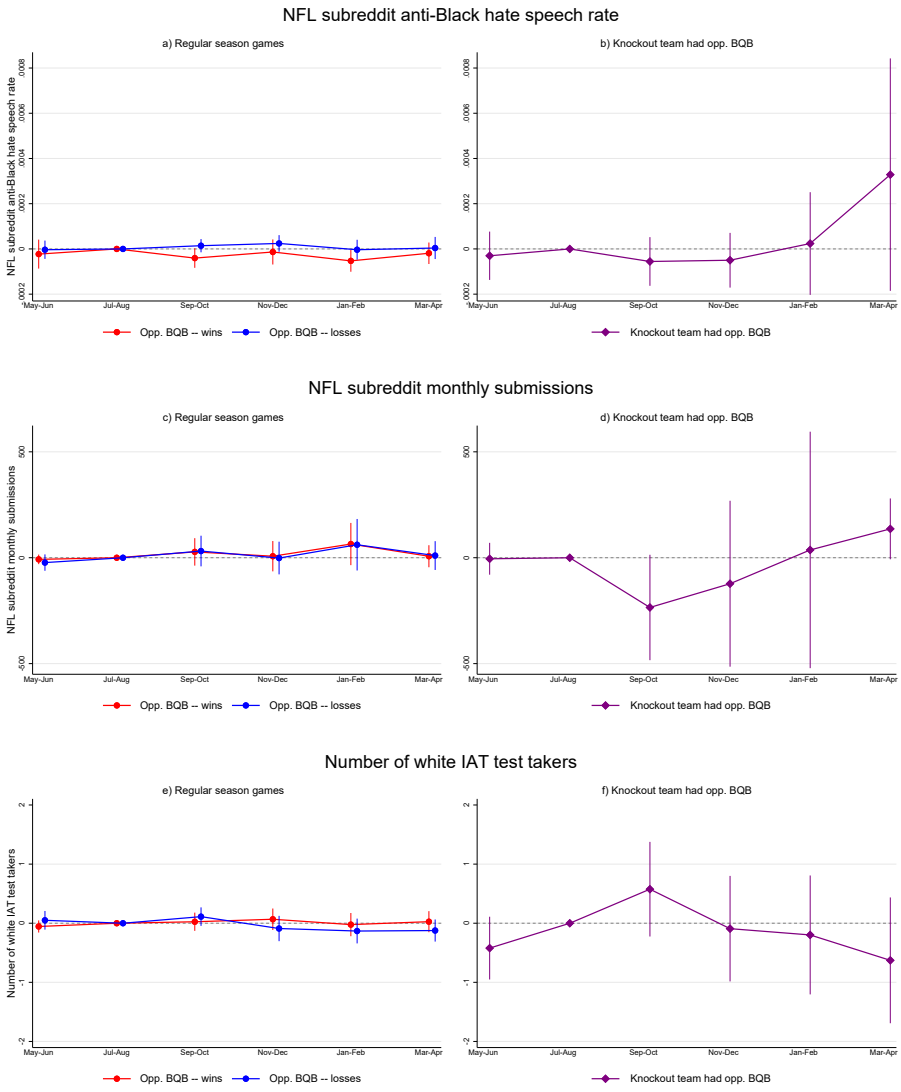

Figure A6: Season effects of opposition Black quarterbacks – appendix outcomes

Notes: For all figures, the sample is limited to team-seasons where the main quarterback was white. Figures on the left plot δ_T^E from equation 7: the differential effect of an additional regular season game played against opposition Black quarterbacks by game outcome using six two-month time bins with July-August as the omitted category. Figures on the right plot the differential effect of being knocked out of the playoffs by a team with an Black quarterback. Estimates control for $DMA \times season$, $month \times DMA$, $month \times record$, and $month \times season$ fixed effects. Observation levels and weights are: month-subreddit level weighted by subreddit submission counts in row 1; month-subreddit level in row 2; and month-DMA level in row 3. All plots show 95% confidence intervals with standard errors clustered at the team-season level.

Table A13: Close-upset game-day impacts of Black head coaches

	(1)	(2)	(3)
	Close-upset games		
	Daily DMA anti-Black hate crime rate	Share of submission removals	Standardized white IAT scores
White-white coach games:			
Post	0.0510	0.0516**	0.0276
	(0.0363)	(0.0224)	(0.0273)
Post × *Win*	-0.0257	-0.0413***	-0.0171
	(0.0294)	(0.0149)	(0.0214)
Any Black coach games:			
Post × *Any Black coach*	-0.0481	0.0166	-0.0203
	(0.0398)	(0.0179)	(0.0257)
Post × *Any Black coach* × *Win*	0.0706	-0.0188	-0.00448
	(0.0553)	(0.0224)	(0.0398)
Observations	38865	2679	60535
FE: DMA × Game	Yes	Yes	Yes
FE: Day of week	Yes	Yes	Yes

Note: This table reports estimates from a modified version of equation 5 where the indicator for a Black quarterback is replaced with an indicator for a Black coach. Standard errors clustered at the game-team level are reported in parentheses, with the following significance indicators: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Observation levels and weights are: day-DMA level weighted by DMA population in Column 1; day-subreddit level weighted by subreddit submission counts in Column 2; and at the individual test level in Column 3. All columns are restricted to close-upset games with a score differential of 7 points or fewer in which the pre-game spread predicted the losing team to win by 3 or more points.

Table A14: Season effects of a local Black head coach by team performance

	(1)	(2)	(3)	(4)
	Anti-Black hate crime rate (month-DMA per 100,000)	Log of Google slur index	Share of submission removals	Standardized white IAT scores
Not in playoffs				
Head coach is Black × *After season*	0.00425	0.00247	-0.0000791	0.00136
	(0.00449)	(0.0154)	(0.00990)	(0.00622)
In playoffs only				
Head coach is Black × *After season* × *In playoffs only*	0.00189	0.0106	-0.00457	-0.00211
	(0.00776)	(0.0206)	(0.0213)	(0.0110)
In championships				
Head coach is Black × *After season* × *In championship*	-0.00538	-0.0152	-0.0287	-0.00259
	(0.00771)	(0.0276)	(0.0232)	(0.0165)
Coefficient sums and significance:				
Championship level differential	-0.001	-0.013	-0.029	-0.001
Dep. mean	0.0510	3.517	0.103	0.168
Observations	54912	16266	3023	3584160
Team × seasons	704	635	252	618
FE: DMA × Season	Yes	Yes	Yes*	Yes
FE: Month × Record	Yes	Yes*	Yes	Yes
FE: Month × DMA	Yes	Yes*	Yes*	Yes
FE: Month × Season	Yes	Yes*	Yes	Yes

Note: This table reports estimates from a modified version of equation 6 where the indicator for a Black quarterback is replaced with an indicator for a Black coach. Standard errors clustered at the team-season level are reported in parentheses, with the following significance indicators: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Observation levels and weights are: month-DMA level, weighted by DMA population for Column 1; 2 month bin-DMA level weighted by DMA population for Column 2; month-subreddit level weighted by the number of subreddit submissions for Column 3; and at the individual-test level for Column 4. Fixed effects are adjusted due to data formats using 2 month bins (instead of monthly) for the Google slur index and subreddits (instead of DMAs) for the submission removal share. During season ($T = d$) estimates are included in all estimations but are not reported for brevity.

Table A15: Heterogeneity in the hate crime game-day effects of Black quarterbacks– close-upset games

Heterogeneity characteristic:	(1) Exposure index	(2) Std RAI
Below median characteristic areas		
Post × *Any Black QB*	0.0845	0.0887**
	(0.0618)	(0.0414)
Post × *Any Black QB* × *Win*	-0.0858	-0.101*
	(0.0958)	(0.0581)
Differential for above median characteristic areas		
Post × *Any Black QB*× *Het*	0.00232	-0.0292
	(0.0723)	(0.0692)
Post × *Any Black QB* × *Win*× *Het*	-0.0220	0.00192
	(0.109)	(0.0990)
Win below median	-0.001	-0.012
Loss above median	0.087**	0.060
Win above median	-0.021	-0.040
Observations	44715	36755
Controls: $Post$; $Post \times Win$; $Post \times Het$; $Post \times Win \times Het$	Yes	Yes
FE: DMA × Game	Yes	Yes
FE: Day of week	Yes	Yes

Note: This table reports estimates from a modified version of equation 5 that includes interaction terms with area characteristics. Standard errors clustered at the game-team level are reported in parentheses, with the following significance indicators: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Observation levels and weights are: day-DMA level weighted by DMA population. All columns are restricted to close-upset games with a score differential of 7 points or fewer in which the pre-game spread predicted the losing team to win by 3 or more points.

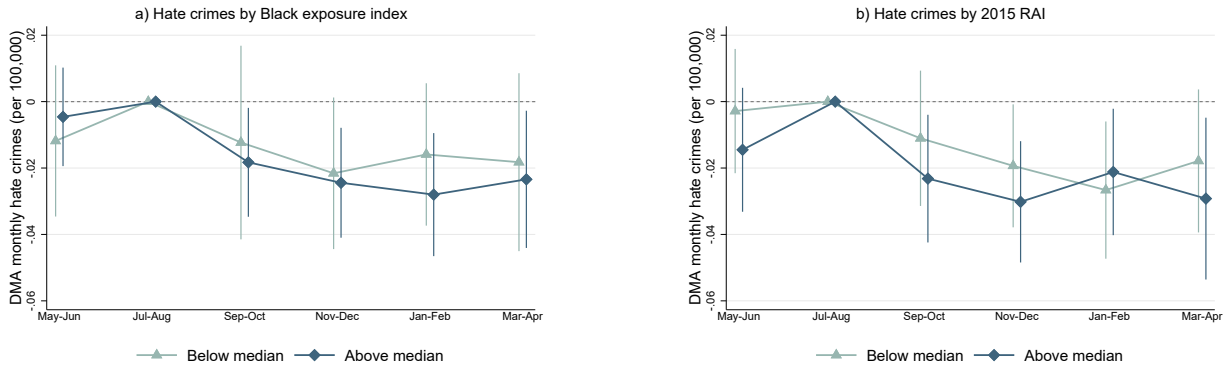

Figure A7: Heterogeneity in the hate crime season effects of Black quarterbacks

Notes: Figures plot $\mu_T^r + \mu_T^c$ from equation 6 for below and above median heterogeneity characteristic using 6 two month time bins with July-August as the omitted category. Estimates control for $DMA \times season$, $month \times DMA$, $month \times record$, and $month \times season$ fixed effects. Observations are at the month-DMA level weighted by DMA population. All plots show 95% confidence intervals with standard errors clustered at the team-season level.