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

Claire Duquennois and Jiangnan Zeng*

November 3, 2026

Abstract

How does prejudice evolve with repeated exposure to out-group leaders? By studying two decades of National Football League Black quarterbacks and four measures of racial animus we show that visible Black leadership can increase animus in the short-run through affect-driven asymmetric retaliation: emotional losses involving any Black quarterback differentially raise anti-Black hate crimes by 43%, with parallel rises in online hate-speech indicators and implicit bias; comparable wins do not lower animus. In contrast, performance-based updating reduces animus in the long-run: over a season, successful Black-quarterback-led teams produce lasting reductions in hate crimes (34%), racial-slur searches, and implicit bias.

1 Introduction

What shapes prejudice, and under what conditions does it change? A large literature documents the prevalence and consequences of prejudice and discrimination (Bertrand and Duflo 2017; Neumark 2018; Lang and Kahn-Lang Spitzer 2020; Kline et al. 2022; Chetty et al. 2020). Work in the social sciences and in economics has sought to understand their sources (Allport 1954; Becker 1957; Bordalo et al. 2016), what causes them to increase (Glaeser 2005; DellaVigna and La Ferrara 2015; Zhuravskaya et al. 2020; Müller and Schwarz 2021, 2023), and, critically, how prejudice can be reduced (Allport 1954; Boisjoly et al. 2006; Corno et al. 2022; Beaman et al. 2009; Paluck et al. 2021). 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 out-group leaders' successes and failures are repeatedly and publicly

*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. We thank Ramez Elprince for excellent research assistance.

tied to in-group gains and losses, and whether prejudice tracks those outcomes or the leaders' performance. Research on such dynamics has focused on individual-level discrimination rather than population-level animus, and has studied settings in which an out-group member's failure necessarily implies an in-group loss (Sarsons 2017a; Egan et al. 2022). In contrast, we study how reactions to prominent out-group leaders extend to their broader group, in a setting where out-group leaders' successes and failures are highly emotional, repeatedly observed, and associated with both in-group gains and losses. This allows us to disentangle the role of in-group utility shocks from that of leader performance. We show that animus is driven primarily by the in-group utility shock rather than by the out-group leader's success or failure itself, and that the response differs by time horizon. In the short run, salient losses associated with an out-group leader activate blame toward the out-group and raise animus, even if these are the result of the out-group leader's success. In the longer run, sustained out-group leader successes that benefit the in-group erode biases, leading to persistent reductions in animus. 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 explore the dynamic responses to Black leadership by studying the impacts of Black quarterbacks in the National Football League (NFL), a historically racialized, uniquely visible, media-saturated, and emotionally intense context. Quarterbacks are team leaders and well-known public figures, routinely praised or blamed for their team's success or failure. The position was historically dominated by white players due to racialized stereotypes about Black intelligence, leadership, and decision-making (Rhoden 2007; Reid 2022), making contemporary Black quarterbacks especially salient in public judgments about Black leadership and competence. This setting lets us examine how racial animus responds when a visible Black leader is associated with a win or loss, and how the response depends on a game's emotional salience and player performance. We combine NFL data with multiple measures of racial animus: anti-Black hate crimes, online hate-speech moderation, racial-slur searches, and implicit bias, to assess how effects generalize across different manifestations of animus. We measure responses at two time horizons: the short run (the days surrounding games); and the longer run (over the course of a season), documenting different patterns and mechanisms.

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 thus 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 comparable wins involving Black quarterbacks. These effects are driven by areas highly engaged with their NFL teams, and concentrated after highly emotional close-scoring upset games. While prior work has documented racial animus increases following a specific game — England's Euro 2020 final defeat (Nägel et al. 2024) — our 5562 games show that this pattern

generalizes, let us identify the conditions that produce it, and allow us to observe asymmetry in effects: a loss involving a Black quarterback inflames animus toward their broader group while an equally salient win produces no short-run reduction.

This asymmetry 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). Our findings make two important contributions to this literature. We provide the first causal evidence of this attribution asymmetry shaping real-world animus spillovers towards an entire racial group. Second, we are able to separately identify the impact of the out-group actor's success or failure from the utility shock experienced by the in-group. Current empirical evidence comes from settings where evaluator's interests are tied to the out-group member's performance: physicians care about patient outcomes after surgeon referrals (Sarsons (2017a)), firms bear the costs of their employees' misconduct (Egan et al. (2022)), departments bear the costs of tenuring under-performing faculty (Sarsons 2017b). In these settings, it is difficult to disentangle what evaluators are responding to: out-group performance or the correlated impact on the evaluators' interests. Our setting separates the two: a local loss can result from either a local Black quarterback's defeat or an opposing Black quarterback's victory. We observe similar increases in animus in both cases, consistent with in-group agents reacting to the harmful outcome associated with the out-group member (e.g. a loss blamable on a Black quarterback), rather than to their performance.

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 after 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 observe 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. Here too, effects are driven by areas highly engaged with their NFL teams. 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 (Allport 1954; Rao 2013; Lowe 2021; Corno et al. 2022; Mousa 2020), and more specifically parasocial contact theory which 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). Only a few paper have documented prejudice reductions through parasocial contact (Paluck 2009; Blouin and Mukand 2019; Duquennois and Zeng 2025; Zussman 2023). Closest to our setting are Lago and Lago-Peñas (2021), who document a correlation between the share of foreign players on winning Spanish soccer league teams and immigration attitudes; and Alrababa'h et al. (2021), who show that anti-Muslim hate crimes and speech in Liverpool fell after 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) by contrasting quarterback effects with estimates for Black head coaches we show that 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 the NFL, one of America's most visible nonpolitical institutions, shapes 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 and intergroup relations. 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. Pinto (2025) extends this to behavior, showing that xenophobic violence rises in the days following a defeat of the German national team. We extend this work in two respects. First, we study majority-group attitudes toward a co-national racial minority rather than toward foreign immigrants or refugees. Second, the relevant intergroup division in our setting is racial rather than national, and animus responds to the salient performance of an individual minority athlete, rather than to a national team's performance. In doing so, we show that domestic sports impact 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. Section 3 outlines our model. Section 4 presents our data. Section 5 presents our short-run estimation strategy and results. Section 6 presents our season estimation strategy and results. Section 7 examines the effect of head coaches, minority leaders that differ sharply from quarterbacks in public visibility. Section 8 examines heterogeneity by 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

[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)).

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 A1). 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.

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 In the Short run: Affect-driven asymmetric racial retaliation

In the short-run, we model a process of affect-driven asymmetric racial retaliation whereby majority group animus toward out-group bystanders is triggered asymmetrically by negative affect shocks associated with salient out-group members. This model is most consistent with the empirical patterns we observe. We then discuss alternative mechanisms and their associated predictions: performance-based asymmetric attribution and status threat.

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.

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;[2] and (iii) whether the game is won or lost:

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

[2]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)).

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 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

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

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; 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 ($\pi^L > \pi^W$). Note that 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 $\pi^W = \pi^W_l \, l_{ng} + \pi^W_o \, o_{ng}$ and $\pi^L = \pi^L_l \, l_{ng} + \pi^L_o \, o_{ng}$, where l_{ng} and o_{ng} indicate a Black quarterback on the local and opposing team respectively.

> ***Game-day predictions for affect driven asymmetric racial retaliation:*** *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 $\frac{\partial |r|}{\partial Win} \leq \frac{\partial |r|}{\partial Loss}$.

The predicted direction of change in racial animus for the four types of games ($\pi^W_l, \pi^W_o, \pi^L_l$ and π^L_o) are summarized in the first column of Table 1.

3.1.1 Alternative mechanisms

Performance-driven asymmetric racial retaliation: Racial animus may instead respond to observing a Black quarterback succeed or fail at leading his team to a win, rather than to affect shocks. Under this model, engagement still matters, but game intensity and the win/loss outcome do not. Rather, what matters are the parameters on Black quarterback failure (π^F) and success (π^S), which capture how much observing failure (success) raises (lowers) animus. If attribution asymmetry holds in this context, failures may raise animus more than successes lower it ($\pi^F > \pi^S$). These parameters capture average effects regardless of whether the Black quarterback plays for the local or opposing team. Thus, π^F applies to local Black quarterback losses (π_l^L); and wins against an opposing Black quarterback (π_o^W), while π^S applies to local Black quarterback wins (π_l^W); and losses to an opposing Black quarterback (π_o^L). Predicted directions of change for the four game types are summarized in the second column of Table 1.

Prior work on attribution asymmetry has examined contexts where evaluators' interests are tied to the out-group member's performance: physicians care about patient outcomes after surgeon referrals (Sarsons (2017a)), firms bear the costs of their employees' misconduct (Egan et al. (2022)), departments bear the costs of tenuring under-performing faculty (Sarsons 2017b). In these contexts it is difficult to differentiate what drives the effect: an asymmetric evaluation of performance or an asymmetric response to the correlated utility. In our context, incentives are similarly aligned for areas with Black quarterback led teams. In contrast, as summarized in Table 1, the affect and performance driven models generate opposing predictions for games where the opposition team is led by a Black quarterback, allowing us to examine whether animus is responding to the utility shock or to observed performance.

Status threat mechanisms: Quarterback is a high-status position historically dominated by white players. Observing Black players in this position could activate a status threat mechanism whereby dominant-group members respond with increased out-group animosity when their relative position is challenged (Blumer (1958), Blalock et al. (1967), Willer et al. (2016), Mutz (2018)). A generalized version of such a mechanism, where just observing a Black player occupying the position triggers a status-threat response, would predict an increase in racial animus after all games where a Black quarterback was on the field (column 3 of Table 1). Alternatively, status threat could only be activated after π_l^W and π_o^L games when a Black quarterback is observed outperforming a white quarterback (column 4 of Table 1).

3.2 In the long run: Performance-contingent updating

Over the course of a season, the emotion-driven game-day affect shocks 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 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

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

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 (Lowe 2021). Whether exposure to opposition quarterbacks reduces or reinforces animus remains an empirical question.

> ***Season predictions for performance-contingent updating:*** *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.[3] We combine local DMA ratings, $HomeRatings^{team}$, which capture how viewed a team is where its fans are most heavily concentrated, with a Google search index for the team, $Google^{team}_{DMA}$, that measures each DMA's interest in the team over the 2004-2025 period relative to the top-searching– and likely local–DMA.[4] For each DMA-team combination we calculate

$$Engagement^{team}_{DMA} = HomeRatings^{team} \times Google^{team}_{DMA}.$$

Each DMA is then assigned to the team it is most engaged with. The resulting catchment areas are mapped in Figure 1(a). Figure 1(b) limits our map to DMAs with above-median engagement levels.[5] These DMAs drive our results (see sections 5.2 and 6.1.2). Our map reflects television and internet search engagement with teams — technologies that are widely used and available to sports fans. Alternative delineations of NFL team catchment areas have been constructed using more selected forms of fan expression: social-media and ticket-sales data.[6] The high engagement areas that drive our results are almost all consistently matched to the same team across all of these maps suggesting that our findings are insensitive to these mapping choices.

NFL data: Using public APIs, [7] we collect play-by-play data for every NFL game in the 2002 through 2024 seasons, recording every on-field action for every player. These data cover 6214 games and 736 season-teams.[8] Only one quarterback is on the field at a time, generally the team's main "starting" quarterback. Multiple quarterbacks may play during a game if the starter is injured or replaced. We define a team-

[3]Full DMA-level ratings are not publicly available. We estimate team's home market ratings, $HomeRatings^{team}$, as the average of local ratings reported for five seasons (2011, 2012, 2013, 2020, 2021) in the Sports Business Journal (Karp (2014); Ourand (2022)).

[4]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.

[5]The distribution of engagement levels across DMAs is plotted in Figure A2.

[6]Several widely circulated maps have been generated using social media data for specific time periods (Sorenson 2024; NFL.com 2014; Wagstaff 2014) as well as ticket-sales data (Priceonomics 2018). We opt for our approach to generate a measure less selected on social media usage or income.

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

[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.

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.[9] 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 had a Black quarterback. To focus on contrasting reactions to European-American and African-American quarterbacks, we exclude 345 games and 23 season-teams where the lead quarterback is one of our data's two non-Black Hispanic or two Samoan quarterbacks. For short-run estimates, which examine outcomes in five-day windows centered on game days, we 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 from exposure to other games in close temporal proximity. Finally, though our results are robust to their inclusion, our main game-level sample omits 307 games featuring two Black quarterbacks, which are few in number and confound multiple effects, complicating interpretation. 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 Black quarterbacks.

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 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. Appendix Table A1 provides descriptive statistics for our main outcomes measured at both the day (panel (a)) and season (panel (b)) levels.

Anti-Black hate crimes: We use the FBI's Uniform Crime Reporting Hate Crime Statistics for 2002–2025 (Federal Bureau of Investigation (2023)).[10] 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)).[11] The incident-level files report the calendar date and location of each incident, a bias-motivation category identifying the

[9] 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)).

[10] 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.

[11] 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)).

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 100,000).

Since hate crimes are rare and many cells contain zero incidents, we check that our main findings are robust to alternative hate crime measurements and other specifications sometimes used in the literature: an extensive-margin indicator for any incident, a zero-adjusted log specification, linear and Poisson specifications on incident counts, inverse hyperbolic sine estimates and robustness to omitting population weights and limiting the sample to DMA-years where hate-crimes are observed at least once.[12] We also estimate effects on the reported non-hate offenses and on hate crimes targeting other groups to better understand mechanisms and rule out possible confounds.

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 2021, 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.[13] 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.[14] This count captures submissions moderated for violating NFL subreddits' prohibitions on hate speech and harassment, and those removed for other violations such as spam or off-topic posts. Because Reddit does not systematically report the substantive reason for each removal, we cannot distinguish between these cases. Our moderation measure is thus a proxy for hateful speech rather than a direct measure of it. Changes in this proxy will signal changes in hate speech in our empirical settings where there is no reason to suspect that non-hate violations such as spam or off-topic posts are changing differentially.

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 et al. 2020).[15] We further fine-tune the

[12]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 2021), Poisson and count specifications for hate-crime outcomes (Anderson et al. 2020), population-scaled hate-crime rates (Pettis et al. 2022) and inverse hyperbolic sine transformations (Bray et al. 2024; Bellemare and Wichman 2020).

[13]All the subreddits and comments data used in this paper comes from Academic Torrents: https://academictorrents.com/

[14]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.

[15]HATEBERTA is trained on a broad set of hate-speech categories, including speech targeting racial and ethnic groups such as Black,

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* (Xu et al. (2024)). *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)).[16] 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.

Race IAT test-takers on Project Implicit are a selected sample. Anecdotally, tests are often part of class assignments, consistent with the 43.5% of respondents who report ages of 17–23. Although the sample is not representative, whether and when students are assigned an IAT is likely orthogonal to NFL game composition and outcomes. As a check, appendix Tables A6 and A13 examine whether daily and monthly DMA-level response counts, along with the age and gender composition of respondents, vary with the presence of Black quarterbacks. We find no evidence of any selection effects.

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 queries.[17] This measure, first developed by Stephens-Davidowitz (2014), has been widely used in the literature as a measure of racial

Asian, and Latino communities. HateBERT was developed and tuned on 2010s social-media text. Analysis with submissions since 2015 mitigates distribution shift between the training corpora and our corpus.

[16]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)).

[17]Specifically, we request indices for searches that include the terms "n___+n___s"—henceforth "the slur query." Following prior work, we use a single search term to maximize geographic coverage as Google does not return indices for geographic-time units in which the absolute number of searches matching the requested term is below an undisclosed threshold. Because we study small areas (DMAs) over short intervals (two months), more specific queries (such as the slur combined with "jokes" or a quarterback name) would fall below this threshold in too many units. Moreover, even if some of these searches reflect unfamiliarity with the slur, such searches will still track changes in the local animus environment as this shapes how often the slur is encountered.

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.[18]

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.[19] 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.

5 Short-run Responses to Games: Affect-driven 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 has shown that emotionally intense games impact behaviors (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 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 immediately following a game based on the racial composition of the quarterbacks that played, the game's result, and 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 ($AnyBlack = 0$).[20] Heterogeneity in effects for opposition and local Black quarterbacks is examined subsequently in section 5.3.

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

[19] 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-two-month bins, mostly in the earliest years of the data and low-population DMAs).

[20] 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 meaningfully alter results.

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 if the game was won (Win_{gn}).[21] We estimate

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

Y_{angd} is an outcome for media market a, team n, on day d, around game g. Day of the week fixed effects, ω_d, control for weekly trends in outcomes. Importantly, a $DMA \times game$ fixed effect, η_{ag}, restricts identifying variation to outcomes measured around the same game in the same media market (and thus the same team).[22] This fixed effect absorbs baseline racial animus which we treat as fixed in the short-run.[23]

With this specification, ρ_2^b identifies how the change in outcomes after games differs for lost 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 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. 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 games with a black quarterback 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).

This effect is driven by emotionally intense games. Figure 2, examines heterogeneity along two dimensions that shape a game's emotional salience: whether the final score was close and whether the outcome was an upset. Figure 2(a) shows that the effects are largest after close games, and Figure 2(b) shows that the effects are largest after upset losses. Motivated by these patterns, we define "close-upset" games as those

[21] Ties are uncommon in the NFL. In our game data, only 14 of 6214 games resulted in a tie. These are coded as $Win_{gn} = 0$.
[22] Geographic identifiers are not available for Reddit outcomes, for which we use η_{ng}, a team-game fixed effect.
[23] For figures we estimate equivalent event time specifications. Day coefficients are estimated relative to the day before the game.

that were (i) close, with a final score differential of 7 points or fewer (i.e., a single touchdown could have changed the outcome), and (ii) upsets, in that the pre-game spread favored the opposing team by 3 points or more. This definition isolates highly emotional games while preserving a sufficiently large sample. Figure A3 shows that ρ_2^b estimates are not discontinuously sensitive to these threshold choices.

Column 2 of table 2 reports estimates for close-upset games. As predicted by the model, effects are larger after these emotionally intense close-upset games. 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 4 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 reports results for moderately intense games (either close scoring or upset games), and column 4 reports estimates on low-intensity games (that were neither close scoring nor upsets). 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 in line with prior work that has examined how sporting events impact hate crimes. Pinto (2025) finds that a national German team soccer loss raises expected attacks on immigrants by roughly 55% while Nägel et al. (2024) document a 30% increase in racial hate crimes in London following England's soccer loss in the Euro 2020 final when three Black English players missed their penalty kicks.

Consistent with model predictions, the effects are also suggestively larger in DMAs that are more engaged with their NFL team. Figure 3(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, and little to no change after wins or low-intensity games—are robust to alternative measures of hate crimes (extensive indicator, adjusted log, Poisson, etc.) and alternative combinations of fixed effects (appendix Tables A2 and A3). We also find no evidence that differential reporting explains this pattern. Column 8 of Table A2 re-estimates our main specification dropping DMA-years in which agencies reported no anti-Black hate crimes. Consistent with agency non-reporting generating attenuation bias, we estimate slightly larger effects in this sample. To examine possible impacts on victims' willingness to report hate crimes, Table A5 estimates our main specification on all non-hate crime offenses and for hate offenses targeting other groups. We find no evidence that Black quarterbacks affect the general crime rate, or hate crimes targeting white or other racial groups. Hate crimes targeting Hispanics respond to games involving Black quarterbacks, but the effects are smaller—plausibly reflecting the racial diversity of the Hispanic population, which includes individuals who identify or are perceived as Black. In sum, for differential reporting to drive our results, it would have

to affect only the reporting of anti-Black hate crimes; and even then, it is difficult to rationalize why victims' reporting behavior would shift only after losses involving Black quarterbacks.

To better understand the nature of these increases in anti-Black hate crimes, Table A4 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).

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 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). Because there is no a priori reason for non-hate violations such as spam to increase differentially after close-upset losses involving Black quarterbacks, we interpret this increase as signaling a rise in hateful submissions. Event-day estimates are illustrated in Figure 4: 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 Figure A4(c) examines 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 A4(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 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 4. Figure 3(b) shows that effects are concentrated in above-median engagement DMAs, with little to no effect in below-median engagement DMAs. Appendix Table A6 shows no evidence of differential changes around game day in the number of white IAT test takers or their age and gender composition, suggesting that selection into testing is unlikely to confound our results.

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. 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.

5.3 An affect driven asymmetric racial retaliation mechanism

The loss-win 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 population-level anti-Black animus towards a broader racial group; second, we are able to examine whether this asymmetric attribution is performance or affect-driven, as outlined in section 3.1.

Several pieces of evidence support an affect driven asymmetric racial retaliation mechanism. First, across outcomes, Tables 2 and 3 show larger changes in animus after emotionally intense games, a pattern consistent with an affect driven response. Second, we estimate equation 5 disaggregating ρ_2^b and ρ_4^b to separately estimate the effects for local and opposition Black quarterbacks ($\rho_2^o, \rho_2^l, \rho_4^o$ and ρ_4^l), allowing us to examine whether animus is responding to affect shocks or to observed performance. Table 4 shows that animus increases for all three measures after losing to an opposition Black quarterback. Despite the opposition Black quarterback's success, each ρ_2^o estimate is positive, large in magnitude, with p-values $\in$[0.056; 0.107]: racial animus increases after a loss even when a Black quarterback succeeded on the field. In contrast, though the opposition Black quarterback failed, we fail to reject the null of no change in animus after wins against an opposition Black quarterback. As outlined in Table 1, this pattern is inconsistent with performance-driven racial retaliation or status threat mechanisms.

Finally, we also find that our estimates are not meaningfully altered when we directly control for how well

a quarterback played. 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, roughly a third of quarterback-games. To separately examine the effect of local and opposition quarterback performance, 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. For hate crimes and IAT scores (columns 1-4), we cannot reject the null of no differential effect for high-EPA Black quarterbacks. These outcomes differentially increase after losses involving Black quarterbacks, regardless of whether they played well. 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. 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 contrast, measures capturing animus in populations less attuned to individual player skill show no such mitigation effect, responding first and foremost to affect.

5.4 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 could reflect racialized evaluations, we use recorded on-field behavior. Quarterback-games are classified as mobility-integrated ($mobQB = 1$) when quarterbacks meaningfully contribute as runners, and as pocket-oriented otherwise ($mobQB = 0$).[24] Using this measure, 50% of games played by Black quarterbacks are mobility-integrated, compared to 13% for white quarterbacks. Table A7 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.

[24] 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.

6 Season responses: Performance-contingent updating

Section 5 shows 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 examine a more consequential and socially important effect: the persistent effect 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 whose teams are led by Black quarterbacks and areas whose teams are led by white quarterbacks, conditional on the teams finishing the season at the same level of performance.

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 categorize each team-season into one of three season performance record 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). 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.

Our main specification estimates the differential evolution of racial animus in areas with Black quarterback teams as compared to areas with white quarterback teams that end the season in the same season record tier

(R_{ns}). We estimate

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

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 common seasonal variation among teams that eventually reach the same season record tier (R_{ns}) , regardless of quarterback race. A $month \times season$ fixed effect, ϕ_{ms}, absorbs common aggregate shocks in a given month.[25] 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 that end the season in the same season record tier R_{ns}.[26]

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 A8 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.

[25] These fixed effects are adjusted for certain outcomes due to data format and availability. 2 month bins are used instead of months for the Google slur index and team subreddits are used instead of DMAs for the submission removal share.

[26] For example, $(\mu_{Sept-Oct}^r + \mu_{Sept-Oct}^c)$ estimates the Sept-Oct difference in animus between areas with Black quarterback teams that will eventually reach the conference championship and areas with white quarterback teams that will eventually reach the conference championship. The $month \times R_{ns}$ fixed effects restrict the comparison to teams that end the season in the same season record tier which, in effect, proxies for within-season performance under the assumption that these teams are in expectation performing at comparable levels over the course of the season.

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 during-season ($\hat{\mu}_d$) and 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 estimates of $\hat{\mu}_d^r$ and$\hat{\mu}_a^r$ implie that these DMAs experience no significant differences 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 both during the season ($\hat{\mu}_d^r + \hat{\mu}_d^c < 0$) and after the season ($\hat{\mu}_a^r + \hat{\mu}_a^c < 0$) relative to DMAs with championship-level white quarterback teams. During the season, their anti-Black hate crime rate is differentially lower by -0.017 (p-value = 0.001), a decrease of approximately 33%, an effect that persists at -0.018 (p-value = 0.037) into the months after the season's end. 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 drop for championship-level teams begins in September, when the NFL season starts, persisting through the NFL season and into the spring after the season ends. The effect of championship-level Black quarterbacks is large but not implausible. This 33% decrease is larger but broadly comparable to the 16% drop in anti-Muslim hate crimes 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.

These results are robust to alternative specifications of our anti-Black hate crime measure (extensive

indicator, adjusted log, Poisson, etc.) and different combinations of fixed effects (appendix Tables A9 and A10). We also find no evidence that these patterns are driven by changes in crime reporting. Dropping DMA-years with no reported anti-Black hate crimes (Table A9, column 8) yields slightly larger effects, consistent with agency non-reporting generating attenuation bias. We also examine the non-hate crime rate and hate crimes targeting other racial groups in appendix Table A12. We find no evidence of effects on these outcomes, for championship-level Black-quarterback teams, either during or after the season, suggesting that the results do not reflect broader shifts in crime reporting. Finally, we check whether this effect is driven by a particular team or quarterback. Appendix Figure A5(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 the shaded 95% confidence intervals.

To better understand the nature of these reductions in anti-Black hate crimes, Table A11 examines impacts by offense type and perpetrator count. Like our short-run effects, the reduction is driven by more common offense types: offenses against persons or property, committed by a single or unknown number of perpetrators. Changes in socially coordinated multi-perpetrator crimes are not driving these results.

Google searches for racial slurs: Table 6, column 2, 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 ($\hat{\mu}_d^r > 0$ and $\hat{\mu}_a^r > 0$). Relative to similarly performing teams led by white quarterbacks, in-season racial slur searches were approximately 2.3% higher (p-value = 0.012) in DMAs whose Black-quarterback team missed the playoffs, an effect that persists at 2.5% (p-value = 0.038) into the months after the season's end. In contrast, and consistent with model predictions, racial slur searches display a persistent decrease for championship-level teams led by Black quarterbacks. Figure 5(b) and Table 6, column 2, show that relative to DMAs with similarly performing teams led by white quarterbacks, in-season searches are approximately 4.6% lower (p-value = 0.006) in these DMAs, an effect that persists at 3.7% (p-value = 0.053) into the months after the season's end. The statistically insignificant effect on playoff-only teams falls in between and we fail to reject the null of no differential effect for these teams. As predicted by the model, this difference in effects is driven by DMAs with high levels of team engagement. Figure 6(b) plots the persistent 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 A5(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; 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 championship estimate is therefore 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 A5(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.[27] 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 IAT 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)). During the season, their IAT scores are differentially lower by -0.018 SD (p-value=0.031), with the reduction reaching -0.040 SD

[27] 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 A6).

(p-value=0.002) in 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 A5(d) 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 A13 examines 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 or their age and gender composition in DMAs exposed to Black quarterback teams.

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 (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 bias reductions of 0.065 to 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 racial animus indicators consistently show that high-performing, championship-level teams with Black quarterbacks reduce animus. This implies 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 ($(\theta_L(R_{ns} = c) - \bar{\theta}_{ns}) > 0$). In contrast, and consistent with the prediction that $(\theta_L(R_{ns} = c) - \bar{\theta}_{ns}) > (\theta_L(R_{ns} = r) - \bar{\theta}_{ns})$, the long-run effects of exposure to Black quarterbacks on low-performing teams 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 thus either matches or falls slightly below baseline perceptions ($\theta_L(R_{ns} = r) \leq \bar{\theta}_{ns}$).

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* × *season*, *month* × *DMA*, *month* × *record*, and *month* × *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 time trends. Focusing on white quarterback teams, we estimate versions of

$$Y_{ansm} = \delta_0^E + \sum_{T \neq t_0} \delta_T^E \, E_{ns} \times \mathbb{1}\{m = T\} + \tau_{as} + \kappa_{am} + \psi_{Rm} + \phi_{ms} + \epsilon_{ansm}, \tag{7}$$

where E_{ns} is 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 opposition Black quarterbacks 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 set $K_{ns} = 1$ for teams knocked out of the playoffs by an opposition Black quarterback (with $K_{ns} = 0$ for other 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 specifications correspond 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

The long-run aggregate effects of Black quarterbacks could differ if local reductions are offset by persistent increases in racial animus in opposing teams' catchment areas. 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.

Overall, we find no evidence that facing opposing Black quarterbacks on the field leads to persistent changes in racial animus. 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. Coefficients for knockout games are imprecisely estimated, given the smaller playoff sample, and also show no consistent evidence of a persistent impact. The 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 increase anti-Muslim sentiment among the comparison group of rival-club fans.

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

Like Black quarterbacks, Black head coaches are also consequential team leaders (Berry and Fowler (2019)) in a historically racialized leadership position with persistent barriers to entry and differential evaluation (Fanning Madden and Ruther (2011), Pitts et al. (2024), Salaga and Juravich (2020)). Yet they 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 are less visible as on-field actors and their influence on game outcomes is less obvious. They receive substantially less media attention and fame rankings show they are less recognized by the general public.[28]

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. Estimates for Black head coaches are reported in Table A14 for the short-run and Table A15 for the long-run. In general, estimates are small and statistically insignificant and we cannot reject the null of no effect on outcomes in either the short-run and long-run when compared to white head coach teams. Note that consistent with NFL subreddit users being more engaged and familiar with head coaches than the wider public, estimates for reddit moderation are suggestively in predicted directions. Having a championship-level Black head coach is associated with a 2.3 ppt reduction (p-value=0.13) in the share of moderated submissions during the season and the short-run estimates on reddit moderation are directionally consistent with a small response, though the coefficients are small and far from statistically significant.

Taken together, the overall pattern of null effects suggests that leadership alone is not sufficient to generate the 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 leadership positions, 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

[28]Vince Lombardi, the legendary 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 like Tom Brady (87%) and Peyton Manning (82%).

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 A16 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 A9(a) plots the differential hate-crime effects for championship-level Black-quarterback teams in below-median and above-median 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 A16 reports the short-run estimates, and Figure A9(b) presents season-level coefficients. Again, 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 prejudice when prominent out-group leaders' successes and failures are repeatedly and publicly associated with in-group gains and losses. By examining Black quarterbacks in the NFL, we show that exposure to Black leadership is animus-activating in the short-run, triggering affect driven asymmetric racial retaliation, and animus-ameliorating in the longer run, through performance contingent updating. 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 loss–win asymmetry extends evidence on asymmetric attribution from individual-level discrimination to population-level animus toward unrelated members of a broader racial group. Importantly, the estimated post-loss increases are similar whether the Black quarterback plays for the local or opposing team: animus rises even when the Black quarterback succeeds at the local team's expense. This evidence advances our understanding of attribution asymmetry by separating the leader's success or failure from the outcome experienced by the audience. Animus responds to salient local losses associated with highly visible Black athletes, not to the athlete's performance.

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 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 rather than fleeting, their cumulative consequences for intergroup relations are large. These seasonal findings advance evidence on parasocial contact and minority role models by moving beyond the study of a single celebrity and showing how effects are performance-contingent. The lack of comparable effects for Black head coaches further identifies the importance of public media visibility and attribution in shaping these effects.

Finally, 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. Greater visibility of successful minority leadership reduces prejudice, but the same visibility can also expose minority groups to racialized blame. Future work should examine how media and institutional environments shape the way minority performance is observed, attributed, and discussed, and how this influences these countervailing effects. Another promising avenue is to test whether these dynamics extend to other arenas where minority leaders are highly visible and publicly judged, including politics, business, education, and other forms of 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. (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.

Bellemare, M. F. and Wichman, C. J. (2020). Elasticities and the inverse hyperbolic sine transformation. Oxford Bulletin of Economics and Statistics, 82(1):50–61.

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.

Boisjoly, J., Duncan, G. J., Kremer, M., Levy, D. M., and Eccles, J. (2006). Empathy or antipathy? the impact of diversity. American Economic Review, 96(5):1890–1905.

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

Bray, K., Braakmann, N., and Wildman, J. (2024). Austerity, welfare cuts and hate crime: Evidence from the uk's age of austerity. Journal of Urban Economics, 141:103439.

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.

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.

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.

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.

Federal Bureau of Investigation (2002–2023). Hate crime statistics. United States Department of Justice. Uniform Crime Reporting Program; accessed May 20, 2025.

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.

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.

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.

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.

Lago, I. and Lago-Peñas, C. (2021). The glories of immigration: How soccer wins shape opinion on immigration. Migration Studies, 9(3):466–489.

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

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.

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. (2021). 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.

Mutz, D. C. (2018). Status threat, not economic hardship, explains the 2016 presidential vote. Proceedings of the National Academy of Sciences, 115(19):E4330–E4339.

Nägel, C., Kros, M., and Davenport, R. (2024). Three lions or three scapegoats: Racial hate crime in the wake of the Euro 2020 final in London. Sociological Science, 11:579–599.

Neumark, D. (2018). Experimental research on labor market discrimination. Journal of Economic Literature, 56(3):799–866.

NFL.com (2014). Twitter breaks down NFL fan bases across the country. Accessed: October 8, 2026.

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.

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., 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.

Pinto, G. (2025). Sports nationalism and xenophobia: When cheering turns into violence. Journal of peace research, 62(3):595–612.

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.

Priceonomics (2018). Where do NFL fans live? mapping football fandom across the U.S. Accessed: October 8, 2026.

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.

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.

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.

Sorenson, M. (2024). NFL follower map. Accessed: October 8, 2026.

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.

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

Wagstaff, K. (2014). Which NFL teams do Americans 'like?' Facebook map has the answer. Accessed: October 8, 2026.

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

Willer, R., Feinberg, M., and Wetts, R. (2016). Threats to racial status promote tea party support among white americans. Available at SSRN 2770186.

Xu, K., Nosek, B. A., and Greenwald, A. G. (2002–2024). Race implicit association test (IAT) data. Project Implicit Demo Website Datasets, Open Science Framework. accessed May 20, 2025.

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.

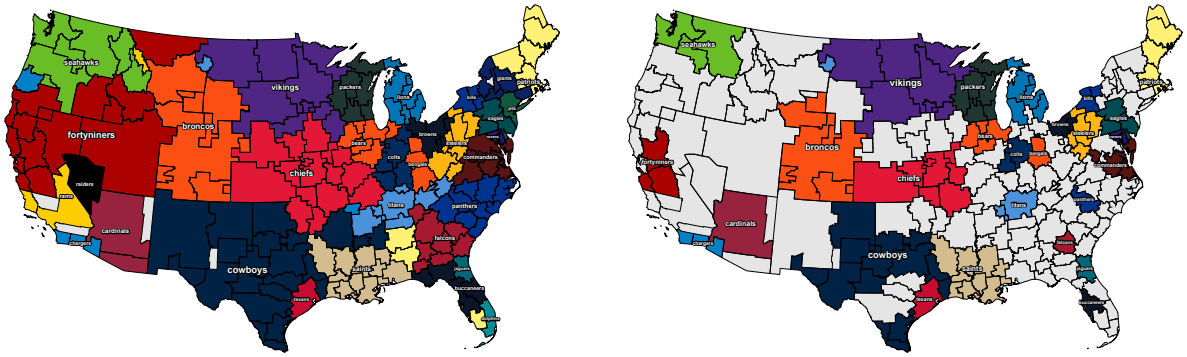
Zhuravskaya, E., Petrova, M., and Enikolopov, R. (2020). Political effects of the internet and social media. Annual Review of Economics, 12:415–438.

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 and Figures

Table 1: Predicted change in racial animus by mechanism for the four types of games

Game type	Local team	Black QB	Mechanism (1) Affect-driven asymmetric racial retaliation	Mechanism (2) Performance-driven asymmetric racial retaliation	Mechanism (3) Generalized status threat	Mechanism (4) Performance-based status threat
π_l^L Local Black QB-led team loses	Loss	Fail	↑	↑	↑	0
π_l^W Local Black QB-led team wins	Win	Success	↓ or 0	↓ or 0	↑	↑
π_o^L Local team loses to a Black QB opposition team	Loss	Success	↑	↓ or 0	↑	↑
π_o^W Local team wins against a Black QB opposition team	Win	Fail	↓ or 0	↑	↑	0



(a) Local NFL teams across all media markets

(b) Media markets with above-median local team engagement

Figure 1: NFL team catchment areas

Note: NFL team catchment areas assign each designated market area (DMA) to the team with the highest engagement. Our engagement measure combines each team's local home-market TV ratings with DMA-level Google Search interest in that team over 2004–2025. Panel (a) shows assignments for all DMAs. Panel (b) restricts the map to DMAs with above-median engagement levels, which drive our empirical results.

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.

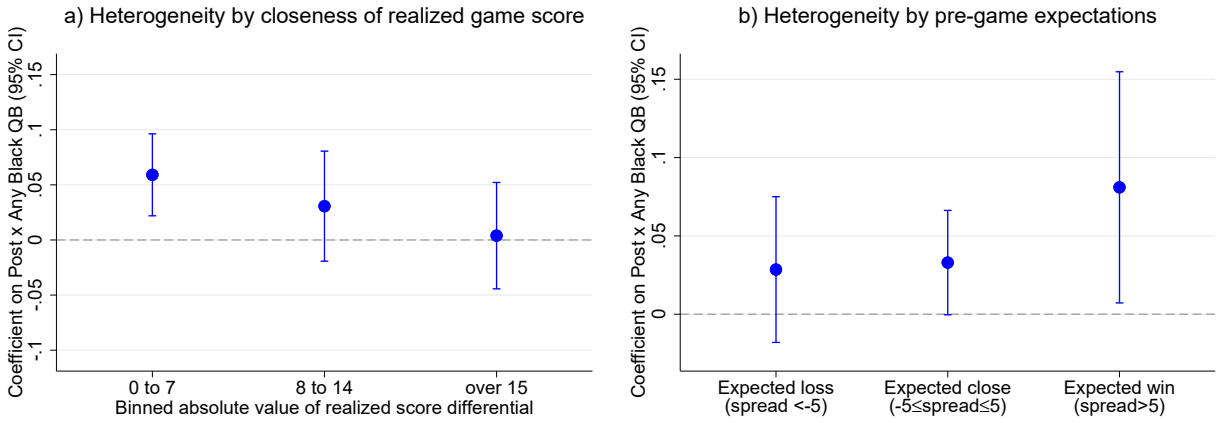


Figure 2: Effects after Black quarterback losses – heterogeneity by game's emotional salience

Note: The figures plot the ρ_2^b coefficient from equation 5 for different samples of games. Figure (a) estimates ρ_2^b separately for three samples defined by the range of realized game point-spreads. Figure (b) estimates ρ_2^b separately for three samples defined by pre-game spread expectations. Estimates include DMA × game and day of the week fixed effects. 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.

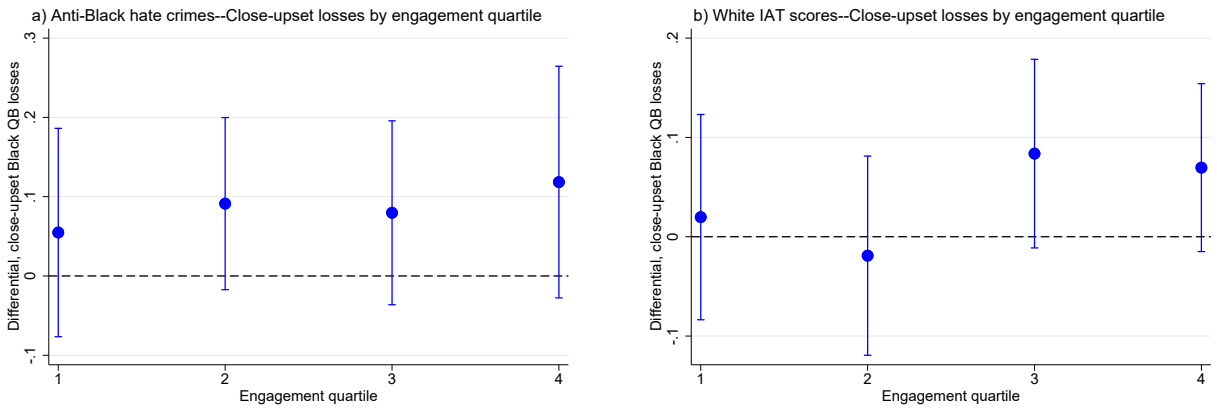


Figure 3: Effects after Black quarterback close-upset losses by engagement quartile

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.

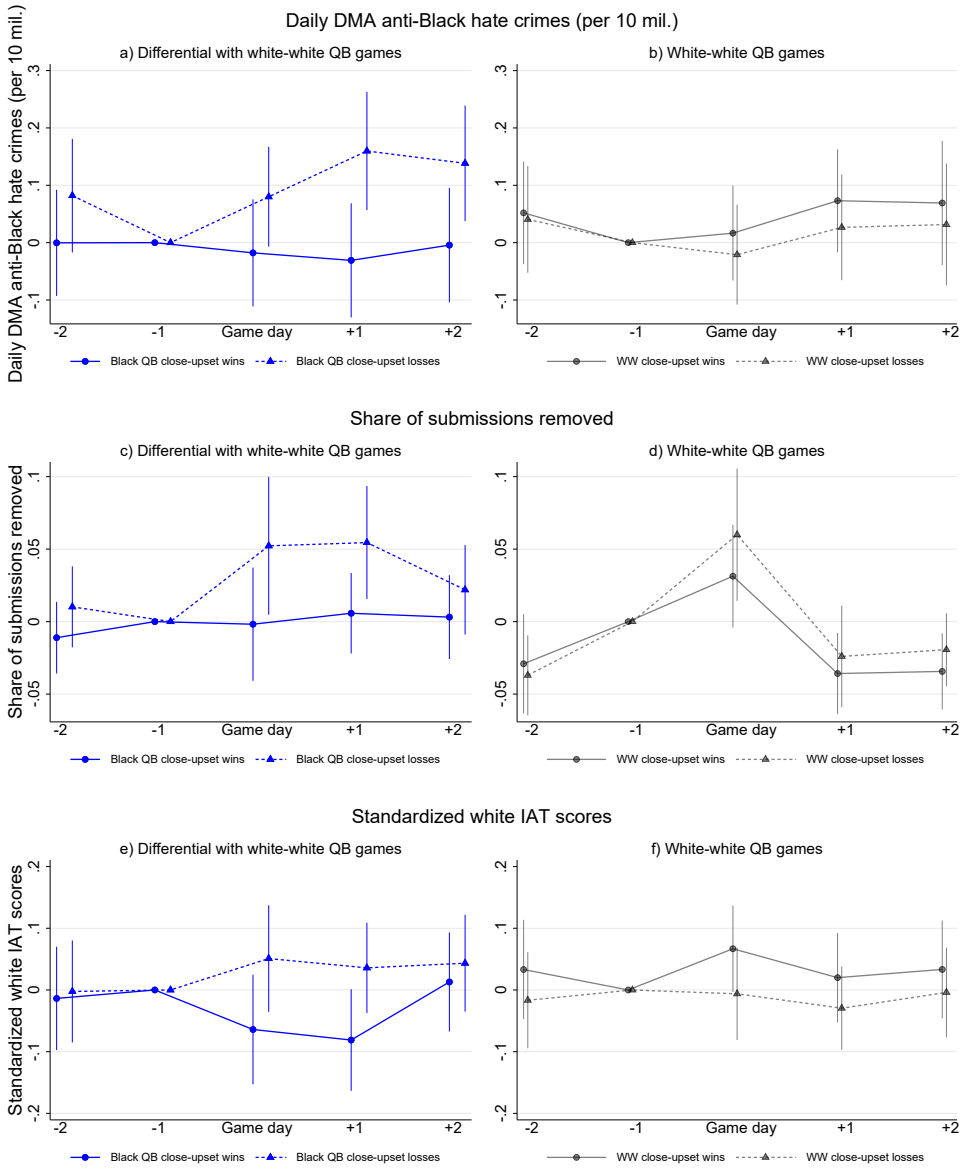


Figure 4: Game-day effects of Black quarterbacks in close-upset games

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 (or subreddit$\times$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.

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. Fixed effects are adjusted due to data availability using subreddits instead of DMAs for the submission removal share. 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)
Coefficient sums and significance:			
$\rho_2^l + \rho_4^l$	0.014	-0.001	-0.041
$\rho_2^o + \rho_4^o$	-0.048	0.010	-0.022
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. Fixed effects are adjusted due to data availability using subreddits instead of DMAs for the submission removal share. 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. Fixed effects are adjusted due to data availability using subreddits instead of DMAs for the submission removal share. 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
During season				
Not in playoffs				
Main QB is Black × During season	-0.00110	0.0227**	0.00409	-0.00570
	(0.00277)	(0.00902)	(0.00871)	(0.00447)
In playoffs only				
Main QB is Black × During season × In playoffs only	-0.00475	-0.00986	-0.000324	0.00461
	(0.00472)	(0.0143)	(0.0114)	(0.00756)
In championships				
Main QB is Black × During season × In championship	-0.0160***	-0.0690***	-0.00458	-0.0121
	(0.00587)	(0.0185)	(0.0309)	(0.00938)
After season				
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:				
In playoffs only teams				
During season differential ($\mu_d^r + \mu_d^p$)	-0.006	0.013	0.004	-0.001
After season playoff differential ($\mu_a^r + \mu_a^p$)	-0.006	-0.000	0.004	0.001
Championship teams				
During season differential ($\mu_d^r + \mu_d^c$)	-0.017***	-0.046***	-0.000	-0.018**
After season differential ($\mu_a^r + \mu_a^c$)	-0.018**	-0.037*	0.009	-0.040***
Dependent mean	0.0530	3.518	0.103	0.168
Observations	52044	15414	2927	3584160
Team × seasons	682	613	244	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: 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.

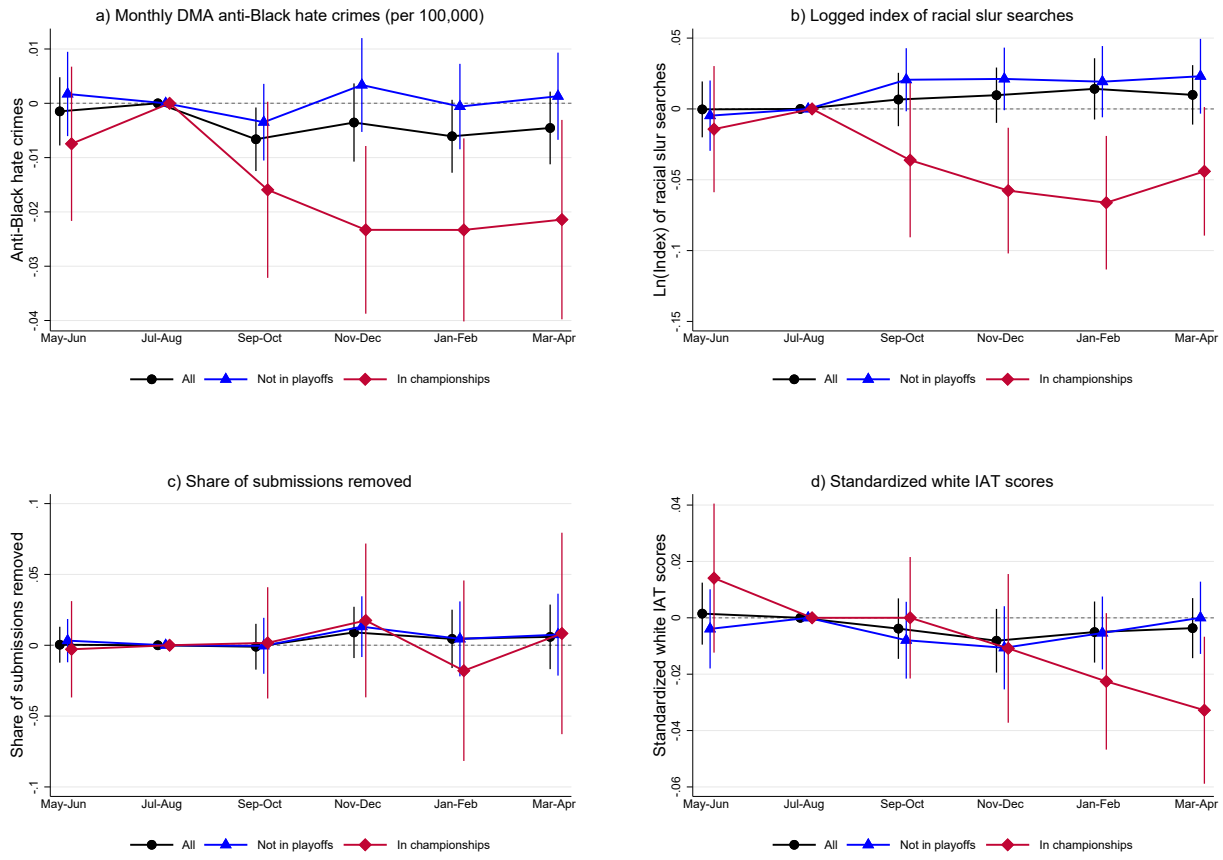


Figure 5: Season effects of local Black quarterbacks by team's season performance

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. 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. 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.

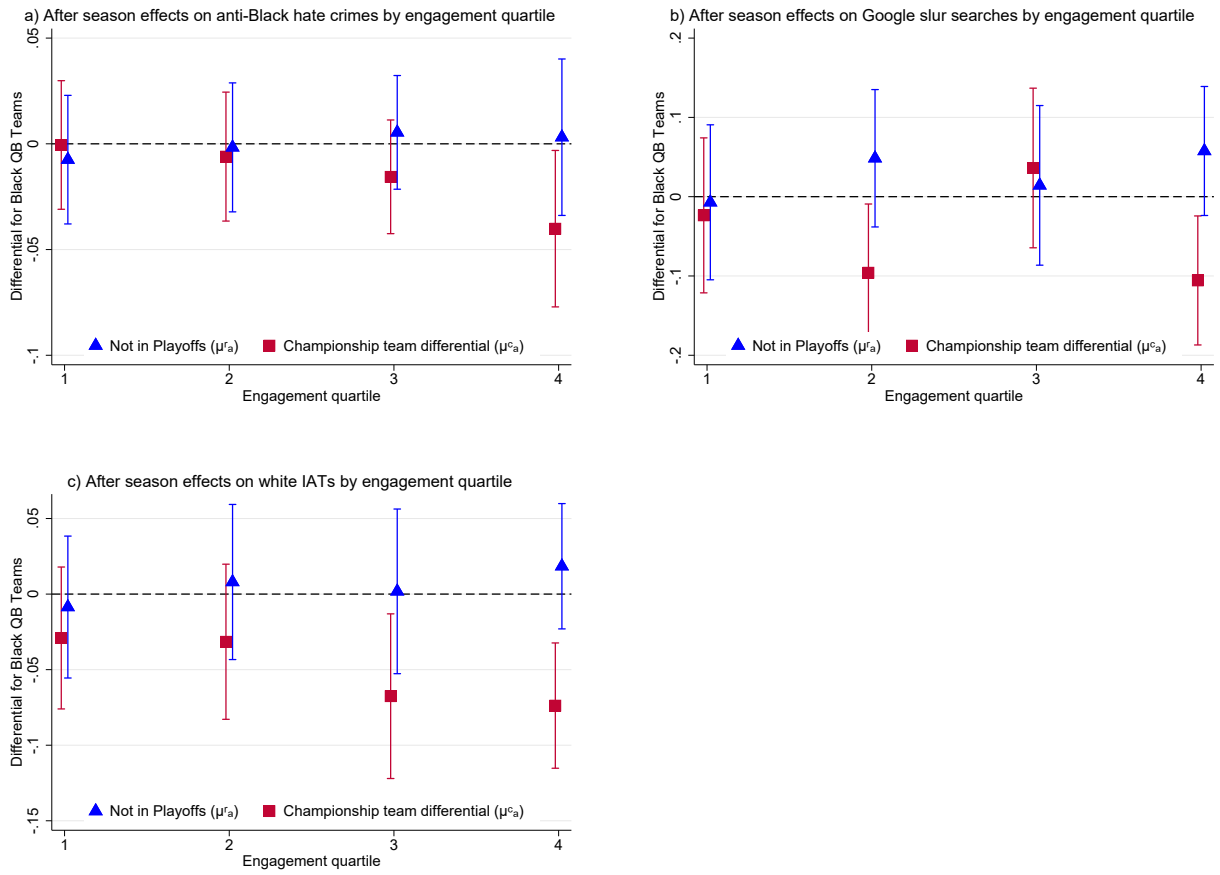


Figure 6: After-season effects of Black quarterbacks by engagement quartile

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. Fixed effects are adjusted due to data formats using 2 month bins (instead of monthly) for the Google slur index. 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.

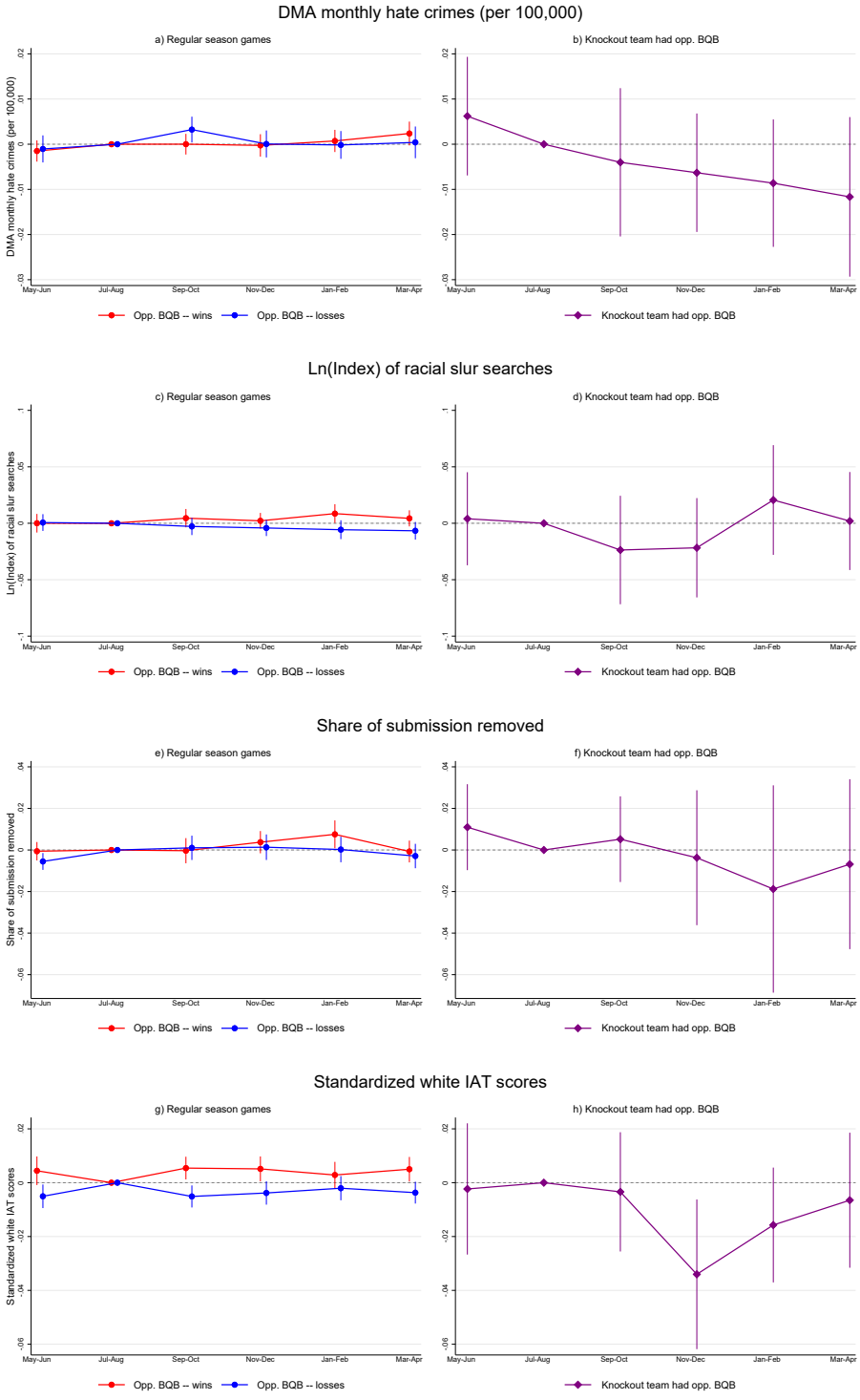


Figure 7: Season effects of opposition Black quarterbacks

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. 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. 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

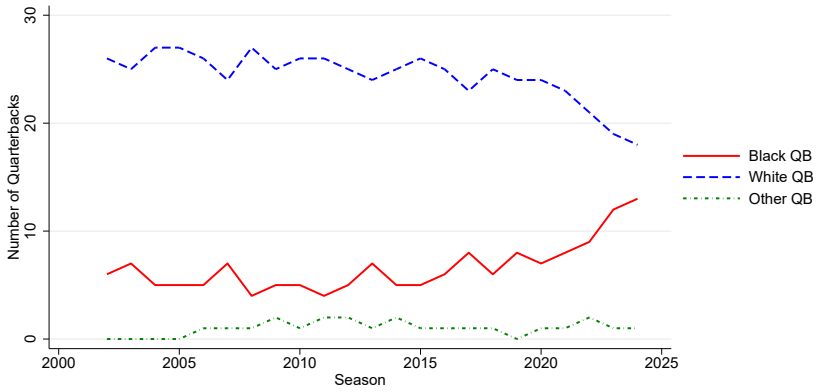


Figure A1: Quarterbacks playing the majority of regular season games in the NFL

Table A1: 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
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

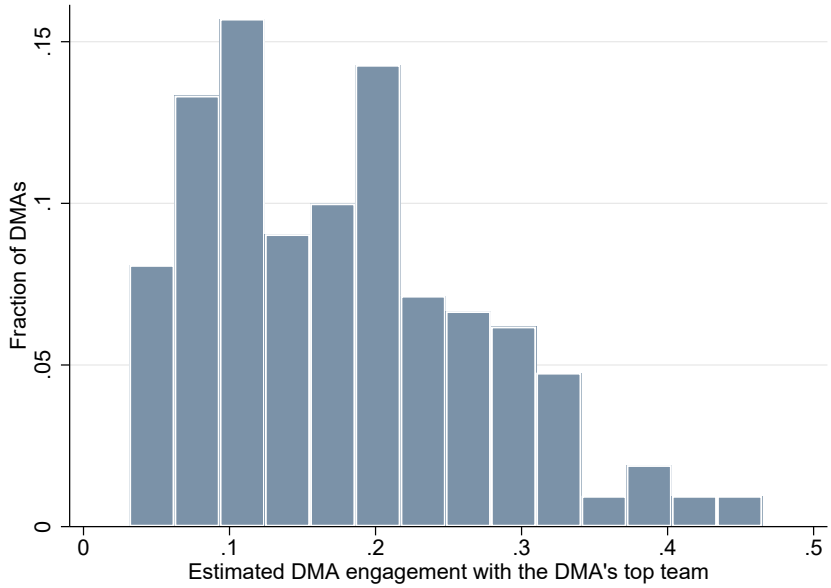


Figure A2: Distribution of DMA engagement levels with their top team

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.

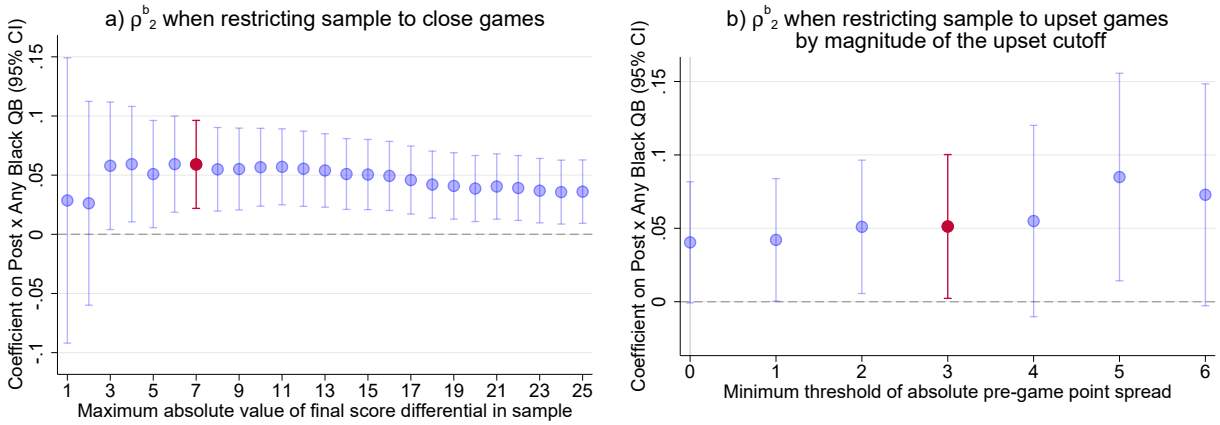


Figure A3: Effects after Black quarterback losses – threshold sensitivity of ρ_2^b

Note: The figures plot the ρ_2^b coefficient from equation 5 for different samples of games. Figure (a) shows how ρ_2^b changes as the sample is restricted to closer games. The 7-point cutoff used in our close-upset sample is shown in red. Figure (b) estimates ρ_2^b on samples of upset games, varying the magnitude of the upset. The 3-point cutoff used in our close-upset sample is shown in red. Estimates include DMA × game and day of the week fixed effects. 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.

Table A2: Robustness to alternative outcome measures of daily DMA anti-Black hate crimes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	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)	Inverse hyperbolic sine (incidents per 10 mil.)	Unweighted daily DMA rate (incidents per 10 mil.)	Reporting sample (incidents per 10 mil.)
Panel a: All games								
Post × *Any Black QB*	0.0364***	0.00436**	0.00355***	0.00571***	0.176***	0.00457***	0.0718***	0.0381***
	(0.0130)	(0.00171)	(0.00129)	(0.00202)	(0.0646)	(0.00165)	(0.0207)	(0.0139)
Post × *Any Black QB* × *Win*	-0.0207	-0.00213	-0.00197	-0.00344	-0.102	-0.00255	-0.0279	-0.0224
	(0.0179)	(0.00231)	(0.00172)	(0.00269)	(0.0904)	(0.00220)	(0.0298)	(0.0193)
Dependent mean	0.200	0.027	0.020	0.030	0.288	0.025	0.165	0.212
Observations	306625	306625	306625	306625	31755	306625	306625	234218
Panel b: Close-upset games								
Post × *Any Black QB*	0.0863***	0.0107***	0.00798***	0.0121***	0.464***	0.0102***	0.171***	0.0913***
	(0.0324)	(0.00404)	(0.00298)	(0.00455)	(0.172)	(0.00382)	(0.0596)	(0.0350)
Post × *Any Black QB* × *Win*	-0.104**	-0.0127**	-0.00958**	-0.0148**	-0.544**	-0.0123**	-0.199**	-0.115**
	(0.0469)	(0.00620)	(0.00455)	(0.00694)	(0.233)	(0.00583)	(0.0773)	(0.0504)
Dependent mean	0.202	0.027	0.020	0.030	0.281	0.025	0.179	0.232
Observations	44715	44715	44715	44715	4710	44715	44715	34090
Panel c: Neither close nor upset games								
Post × *Any Black QB*	0.0191	0.00257	0.00209	0.00324	0.0783	0.00270	0.0283	0.0209
	(0.0200)	(0.00266)	(0.00203)	(0.00323)	(0.0973)	(0.00261)	(0.0335)	(0.0213)
Post × *Any Black QB* × *Win*	-0.00944	-0.000909	-0.000963	-0.00185	-0.0380	-0.00125	0.00982	-0.0108
	(0.0273)	(0.00349)	(0.00263)	(0.00415)	(0.139)	(0.00337)	(0.0436)	(0.0293)
Dependent mean	0.199	0.027	0.020	0.030	0.291	0.025	0.160	0.205
Observations	134220	134220	134220	134220	13800	134220	134220	102680

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 and 8. Column 8 drops DMA-years with no reported hate crimes. 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: 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 A4: 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 A5: Game-day impacts on hate crimes targeting other groups and non-hate crimes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Hate crimes				Non-hate crimes			
	Anti-Black	Anti-white	Anti-Hispanic	Anti-other race	All	Against Persons	Against Property	Against Society
White-white games:								
Post	-0.0121	-0.0216	0.0126	-0.00224	-1.281	3.600*	-2.486	-2.396
	(0.0368)	(0.0227)	(0.0175)	(0.0230)	(7.242)	(2.019)	(5.281)	(1.483)
Post × *Win*	0.0348	0.000238	0.0300**	-0.00334	-2.589	-0.515	-1.542	-0.531
	(0.0297)	(0.0160)	(0.0135)	(0.0191)	(5.416)	(1.484)	(3.747)	(1.209)
Any Black quarterback games:								
Post × *Any Black QB*	0.0863***	0.00797	0.00525	-0.000923	-4.686	-1.849	-2.554	-0.283
	(0.0324)	(0.0178)	(0.0138)	(0.0199)	(7.629)	(2.220)	(5.264)	(1.373)
Post × *Any Black QB* × *Win*	-0.104**	0.00741	-0.0450**	0.0230	-4.357	-0.372	-3.076	-0.908
	(0.0469)	(0.0253)	(0.0212)	(0.0273)	(10.64)	(3.118)	(7.298)	(1.995)
$\rho_2^b + \rho_4^b$	-0.018	0.015	-0.040**	0.022	-9.04	-2.22	-5.63	-1.19
Dependent mean	0.202	0.061	0.044	0.066	590.016	130.894	385.164	73.957
Observations	44715	44715	44715	44715	44715	44715	44715	44715
Game × teams	1401	1401	1401	1401	1401	1401	1401	1401
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 of equation 5 on other types of hate crimes and non-hate crimes for close-upset games. Column headers indicate the type of 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.

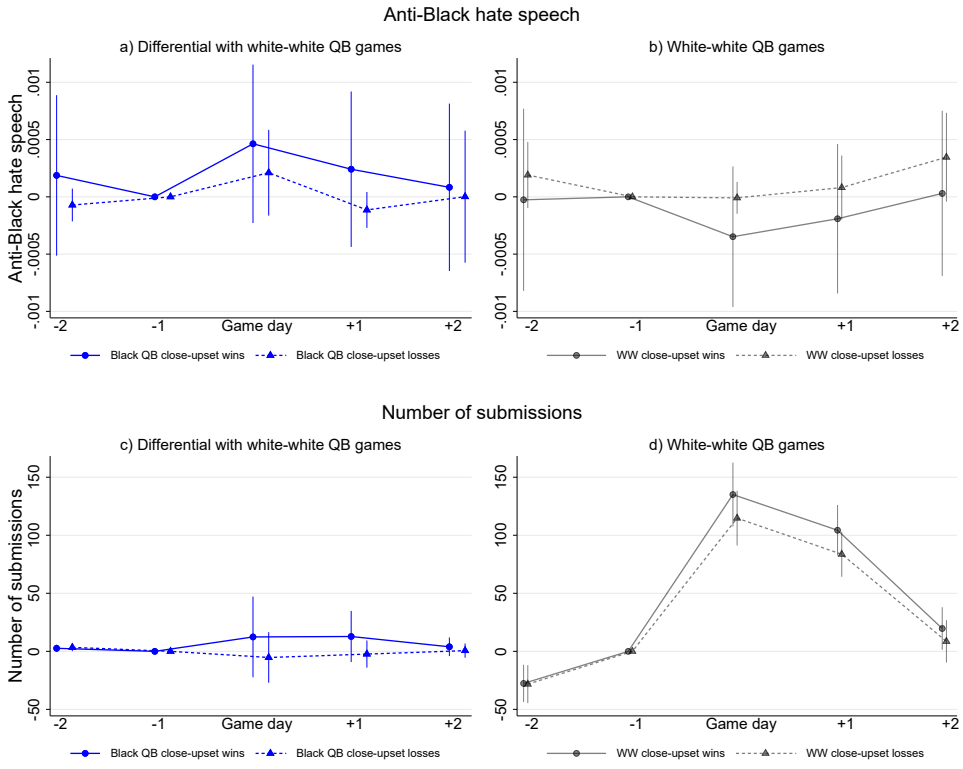


Figure A4: Game-day effects of any Black quarterbacks

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 subreddit $\times$ 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. All plots show 95% confidence intervals with standard errors clustered at the team-game level.

Table A6: Game-day impacts on IAT test taking – Close-upset games

	(1)	(2)	(3)
	Number of white IAT test takers	Age of IAT test takers	Share female of IAT test takers
White-white games:			
Post	0.169	-0.703	0.0213
	(0.208)	(0.658)	(0.0159)
Post × *Win*	-0.230	-0.318	0.00485
	(0.192)	(0.645)	(0.0138)
Any Black quarterback games:			
Post × *Any Black QB*	-0.210	0.329	0.0207
	(0.179)	(0.508)	(0.0146)
Post × *Any Black QB* × *Win*	0.237	-0.473	-0.0283
	(0.256)	(0.862)	(0.0216)
Dependent mean	2.290	29.572	0.626
Observations	44715	68311	71688
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 the number of IAT test takers, as well as their age and gender composition. 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 in column 1 and individual tests in columns 2 and 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.

Table A7: 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	0.157	0.187
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. Fixed effects are adjusted due to data availability using subreddits instead of DMAs for the submission removal share. 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 A8: 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 A9: 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)	(7) Inverse hyperbolic sine (incidents per 100k)	(8) Reporting sample (incidents per 100k)
During season								
Not in playoffs								
Main QB is Black × During season	-0.00110	0.00369	-0.00180	-0.0306	0.0109	-0.00526	-0.00183	-0.00115
	(0.00277)	(0.0124)	(0.0131)	(0.0440)	(0.0443)	(0.00397)	(0.0168)	(0.00307)
In playoffs only								
Main QB is Black × During season × In playoffs only	-0.00475	0.00699	-0.0000110	-0.0700	-0.0579	-0.00118	0.00137	-0.00428
	(0.00472)	(0.0211)	(0.0217)	(0.0753)	(0.0736)	(0.00769)	(0.0280)	(0.00509)
In championships								
Main QB is Black × During season × In championship	-0.0160***	-0.0312	-0.0585**	-0.191**	-0.260***	-0.0120	-0.0750**	-0.0166***
	(0.00587)	(0.0273)	(0.0274)	(0.0878)	(0.0802)	(0.00936)	(0.0352)	(0.00611)
After season								
Not in playoffs								
Main QB is Black × After season	0.000445	-0.0135	-0.0116	-0.00736	0.0250	-0.0106**	-0.0154	0.000166
	(0.00355)	(0.0170)	(0.0167)	(0.0574)	(0.0557)	(0.00535)	(0.0215)	(0.00395)
In playoffs only								
Main QB is Black × After season × In playoffs only	-0.00618	-0.0109	-0.00710	-0.101	-0.0620	-0.00234	-0.00699	-0.00462
	(0.00701)	(0.0273)	(0.0307)	(0.108)	(0.0994)	(0.00917)	(0.0395)	(0.00760)
In championships								
Main QB is Black × After season × In championship	-0.0181**	-0.0318	-0.0564	-0.211	-0.285**	-0.00958	-0.0724	-0.0188*
	(0.00921)	(0.0412)	(0.0441)	(0.139)	(0.136)	(0.0133)	(0.0568)	(0.00973)
Coefficient sums and significance:								
In playoffs only teams								
During season differential ($\mu^r_d + \mu^p_d$)	-0.006	0.011	-0.002	-0.101	-0.047	-0.006	-0.000	-0.005
After season playoff differential ($\mu^r_a + \mu^p_a$)	-0.006	-0.024	-0.019	-0.108	-0.037	-0.013*	-0.022	-0.004
Championship teams								
During season differential ($\mu^r_d + \mu^c_d$)	-0.017***	-0.028	-0.060**	-0.222***	-0.249***	-0.017**	-0.077**	-0.018***
After season differential ($\mu^r_a + \mu^c_a$)	-0.018**	-0.045	-0.068*	-0.219*	-0.260**	-0.020*	-0.088*	-0.019**
Dependent mean	0.0629	0.335	0.374	0.939	1.265	0.0525	0.0525	0.0671
Observations	52044	52044	52044	52044	38629	52044	52044	40152
FE: DMA × Season	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE: Month × Record	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE: Month × DMA	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE: Month × Season	Yes	Yes	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.

Table A10: 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)					
During season						
Not in playoffs						
During season		-0.0136***	-0.0121***			
		(0.00165)	(0.00135)			
Main QB is Black × During season	-0.00110	0.00213	0.000409	0.00214	0.00225	-0.000801
	(0.00277)	(0.00366)	(0.00238)	(0.00367)	(0.00298)	(0.00355)
In playoffs only						
During season × In playoffs only		0.00142	-0.00183			
		(0.00301)	(0.00257)			
Main QB is Black × During season × In playoffs only	-0.00475	-0.00206	0.00219	-0.00206	-0.00574	-0.000397
	(0.00472)	(0.00569)	(0.00455)	(0.00569)	(0.00473)	(0.00553)
In championships						
During season × In championship		0.00561*	-0.00157			
		(0.00288)	(0.00259)			
Main QB is Black × During season × In championship	-0.0160***	-0.0202***	-0.00971*	-0.0202***	-0.0185***	-0.0185***
	(0.00587)	(0.00712)	(0.00510)	(0.00712)	(0.00567)	(0.00676)
After season						
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:						
In playoffs only teams						
During season differential ($\mu_d^r + \mu_d^p$)	-0.006	0.000	0.003	0.000	-0.003	-0.001
After season playoff differential ($\mu_a^r + \mu_a^p$)	-0.006	-0.002	0.000	-0.002	-0.006	-0.001
Championship teams						
During season differential ($\mu_d^r + \mu_d^c$)	-0.017***	-0.018***	-0.009**	-0.018***	-0.016***	-0.019***
After season differential ($\mu_a^r + \mu_a^c$)	-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. 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.

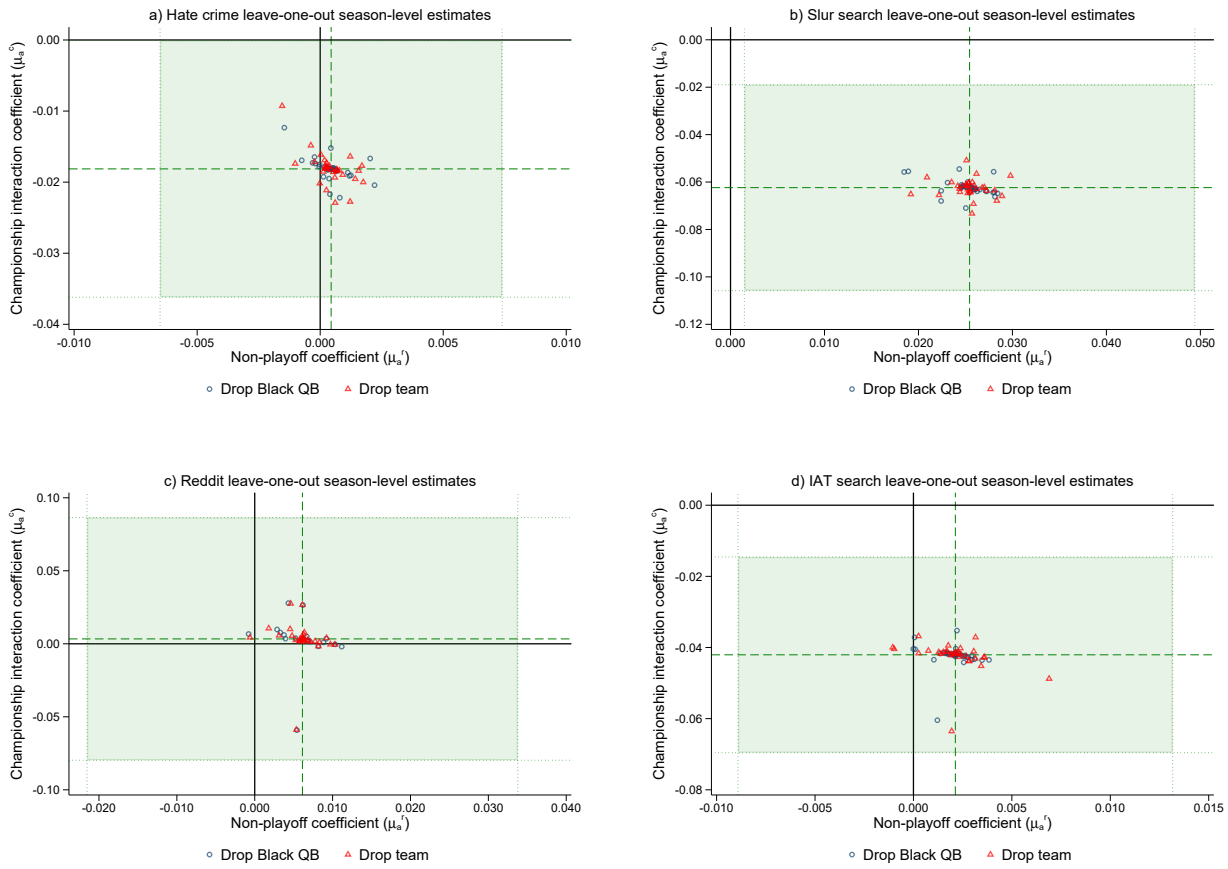


Figure A5: Estimate robustness to leaving out observations for a single team or Black quarterback

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 A11: 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
During season							
Not in playoffs							
Main QB is Black × During season	-0.00149	-0.000611	0.00100	-0.000321	-0.00112	0.000231	0.0000146
	(0.00162)	(0.00193)	(0.000779)	(0.00193)	(0.00169)	(0.000279)	(0.0000941)
In playoffs only							
Main QB is Black × During season × In playoffs only	0.000107	-0.00160	-0.00325**	-0.00382	-0.00174	-0.000291	-0.0000396
	(0.00258)	(0.00304)	(0.00138)	(0.00354)	(0.00276)	(0.000653)	(0.000193)
In championships							
Main QB is Black × During season × In championship	-0.00432	-0.00825**	-0.00346	-0.0114**	-0.00344	-0.000299	0.0000333
	(0.00310)	(0.00399)	(0.00214)	(0.00468)	(0.00322)	(0.000847)	(0.000203)
After season							
Not in playoffs							
Main QB is Black × After season	-0.00366*	0.00331	0.000791	0.00269	-0.00283	-0.000112	0.000214
	(0.00202)	(0.00255)	(0.00120)	(0.00277)	(0.00194)	(0.000328)	(0.000165)
In playoffs only							
Main QB is Black × After season × In playoffs only	-0.00178	-0.00300	-0.00140	-0.00455	-0.00227	-0.0000429	-0.000376*
	(0.00357)	(0.00482)	(0.00197)	(0.00542)	(0.00352)	(0.000931)	(0.000228)
In championships							
Main QB is Black × After season × In championship	-0.00392	-0.0108	-0.00340	-0.00986	-0.00619	-0.000566	-0.000240
	(0.00401)	(0.00658)	(0.00283)	(0.00753)	(0.00408)	(0.000994)	(0.000269)
Coefficient sums and significance:							
In playoffs only teams							
During season differential ($\mu_d^r + \mu_d^p$)	-0.001	-0.002	-0.002*	-0.004	-0.003	-0.000	-0.000
After season playoff differential ($\mu_a^r + \mu_a^p$)	-0.005*	0.000	-0.001	-0.002	-0.005	-0.000	-0.000
Championship teams							
During season differential ($\mu_d^r + \mu_d^c$)	-0.006**	-0.009**	-0.002	-0.012***	-0.005	-0.000	0.000
After season differential ($\mu_a^r + \mu_a^c$)	-0.008**	-0.008	-0.003	-0.007	-0.009**	-0.001	-0.000
Dependent mean	0.0140	0.0310	0.00800	0.0360	0.0170	0.00100	0
Observations	52044	52044	52044	52044	52044	52044	52044
Team × seasons	682	682	682	682	682	682	682
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.

Table A12: Season effects of local Black quarterbacks by team performance on other hate crimes and non-hate crimes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Hate crimes				Non-hate crimes			
	Anti-Black	Anti-white	Anti-Hispanic	Anti-other race	All	Against Persons	Against Property	Against Society
During season								
Not in playoffs								
Main QB is Black × During season	-0.00110	-0.00127	0.000519	-0.000714	2.377	-0.179	1.956	0.600
	(0.00277)	(0.00157)	(0.00107)	(0.00134)	(2.981)	(0.809)	(1.998)	(0.547)
In playoffs only								
Main QB is Black × During season × In playoffs only	-0.00475	-0.00111	-0.00224	-0.00307	-9.156*	-1.550	-6.027	-1.580*
	(0.00472)	(0.00300)	(0.00184)	(0.00232)	(5.293)	(1.214)	(3.866)	(0.830)
In championships								
Main QB is Black × During season × In championship	-0.0160***	0.000861	0.00211	-0.00124	-4.973	-0.466	-2.731	-1.776
	(0.00587)	(0.00331)	(0.00236)	(0.00295)	(4.766)	(1.160)	(3.189)	(1.139)
After season								
Not in playoffs								
Main QB is Black × After season	0.000445	-0.000484	-0.000623	-0.00320	5.065	0.627	2.577	1.862
	(0.00355)	(0.00198)	(0.00149)	(0.00239)	(7.095)	(1.713)	(4.591)	(1.382)
In playoffs only								
Main QB is Black × After season × In playoffs only	-0.00618	-0.00130	0.00264	0.000234	-14.49	-2.301	-8.739	-3.451
	(0.00701)	(0.00387)	(0.00267)	(0.00391)	(11.92)	(2.782)	(8.068)	(2.169)
In championships								
Main QB is Black × After season × In championship	-0.0181**	0.000944	0.00214	0.00174	-7.836	-0.261	-5.651	-1.924
	(0.00921)	(0.00456)	(0.00319)	(0.00500)	(10.12)	(2.473)	(6.504)	(2.442)
Coefficient sums and significance:								
In playoffs only teams								
During season differential ($\mu_d^r + \mu_d^p$)	-0.006	-0.002	-0.002	-0.004*	-6.78	-1.73*	-4.07	-0.980*
After season playoff differential ($\mu_a^r + \mu_a^p$)	-0.006	-0.002	0.002	-0.003	-9.43	-1.67	-6.16	-1.59
Championship teams								
During season differential ($\mu_d^r + \mu_d^c$)	-0.017***	-0.000	0.003	-0.002	-2.60	-0.644	-0.775	-1.18
After season differential ($\mu_a^r + \mu_a^c$)	-0.018**	0.000	0.002	-0.001	-2.77	0.366	-3.07	-0.062
Dependent mean	0.0530	0.0230	0.0100	0.0150	225.3	50.05	142.4	32.85
Observations	52044	52044	52044	52044	54516	54516	54516	54516
Team × seasons	682	682	682	682	713	713	713	713
FE: DMA × Season	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE: Month × Record	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE: Month × DMA	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE: Month × Season	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports estimates of equation 6 on other types of hate crimes and non-hate crimes. Column headers indicate the type of 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.

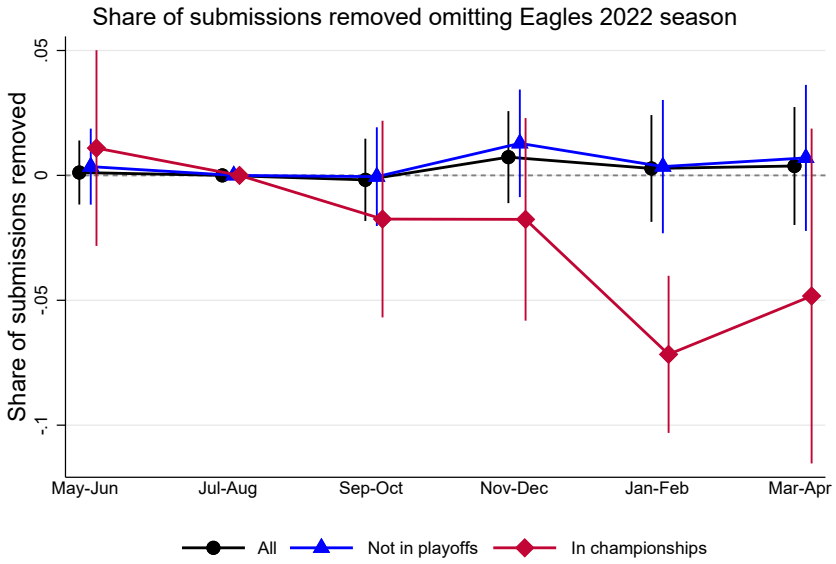


Figure A6: Season effects of local Black quarterbacks on subreddit moderation— Eagle's 2022 season omitted

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 $subreddit \times season$, $month \times subreddit$, $month \times record$, 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.

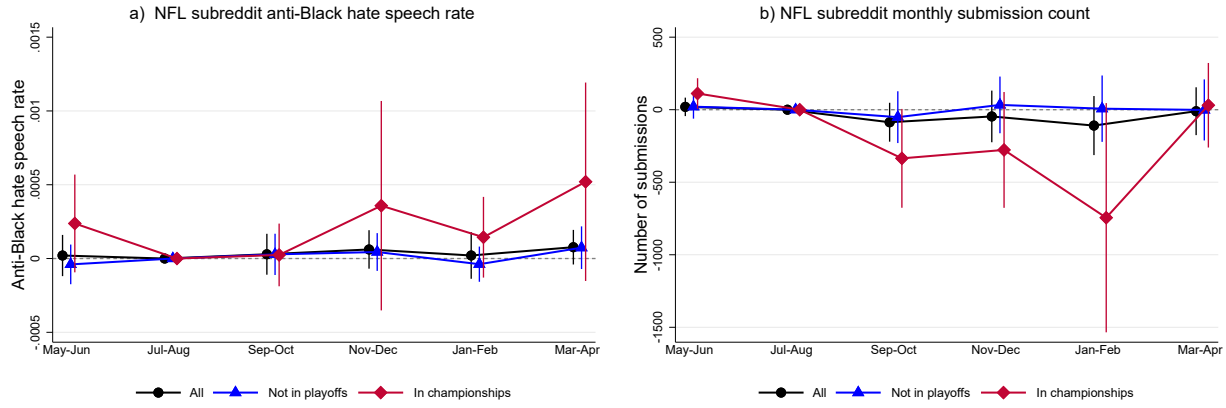


Figure A7: Season effects of local Black quarterbacks by performance level – appendix outcomes

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 $subreddit \times season$, $month \times subreddit$, $month \times record$, 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). All plots show 95% confidence intervals with standard errors clustered at the team-season level.

Table A13: Season effects of IAT test taker characteristics

	(1)	(2)	(3)
	Number of white IAT test takers	Age of IAT test takers	Share female of IAT test takers
During season			
Not in playoffs			
Main QB is Black × During season	-0.0916	0.351**	0.00334
	(0.237)	(0.164)	(0.00306)
In playoffs only			
Main QB is Black × During season × In playoffs only	0.0960	-0.368	0.00698
	(0.471)	(0.262)	(0.00620)
In championships			
Main QB is Black × During season × In championship	0.896*	-0.793**	-0.00714
	(0.525)	(0.380)	(0.00729)
After season			
Not in playoffs			
Main QB is Black × After season	-0.0795	0.367	-0.00438
	(0.229)	(0.227)	(0.00444)
In playoffs only			
Main QB is Black × After season × In playoffs only	-0.261	-0.0520	0.0143**
	(0.479)	(0.350)	(0.00669)
In championships			
Main QB is Black × After season × In championship	0.995*	-0.698	-0.00780
	(0.588)	(0.572)	(0.00982)
Coefficient sums and significance:			
In playoffs only teams			
During season differential ($\mu_d^r + \mu_d^p$)	0.004	-0.017	0.010*
After season playoff differential ($\mu_a^r + \mu_a^p$)	-0.340	0.315	0.010*
Championship teams			
During season differential ($\mu_d^r + \mu_d^c$)	0.804	-0.442	-0.004
After season differential ($\mu_a^r + \mu_a^c$)	0.916	-0.331	-0.012
Dependent mean	4.610	30.88	0.623
Observations	816024	3432152	3516886
Team × seasons	713	618	618
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 and the number and demographic characteristics 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 are at the month-DMA level in Column 1 and individual tests in columns 2 and 3.

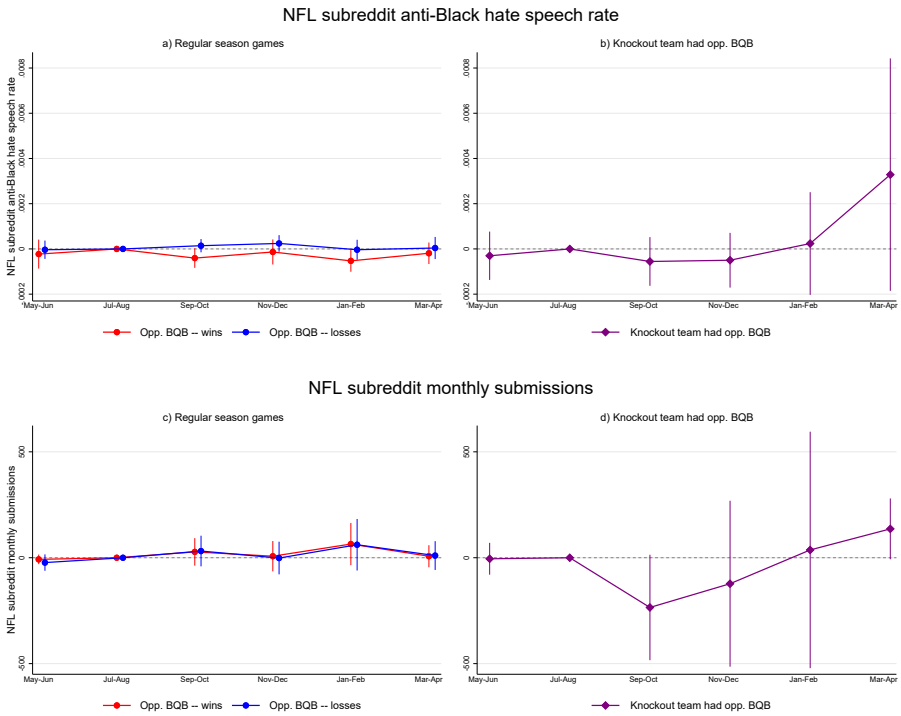


Figure A8: Season effects of opposition Black quarterbacks – appendix outcomes

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 a Black quarterback. Estimates control for $subreddit \times season$, $month \times subreddit$, $month \times record$, 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. All plots show 95% confidence intervals with standard errors clustered at the team-season level.

Table A14: 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. Fixed effects are adjusted due to data availability using subreddits instead of DMAs for the submission removal share. 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 A15: 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
During season				
Not in playoffs				
Head coach is Black × During season	0.00460	0.00195	0.00524	0.00595
	(0.00294)	(0.00929)	(0.00967)	(0.00529)
In playoffs only				
Head coach is Black × During season × In playoffs only	-0.00566	0.0166	0.00445	0.000667
	(0.00474)	(0.0196)	(0.0140)	(0.0102)
In championships				
Head coach is Black × During season × In championship	0.000604	0.0231	-0.0280**	-0.00715
	(0.00585)	(0.0228)	(0.0141)	(0.0150)
After season				
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:				
In playoffs only teams				
During season differential ($\mu_d^r + \mu_d^p$)	-0.001	0.019	0.010	0.007
After season playoff differential ($\mu_a^r + \mu_a^p$)	0.006	0.013	-0.005	-0.001
Championship teams				
During season differential ($\mu_d^r + \mu_d^c$)	0.005	0.025	-0.023	-0.001
After season differential ($\mu_a^r + \mu_a^c$)	-0.001	-0.013	-0.029	-0.001
Dependent 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.

Table A16: 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.

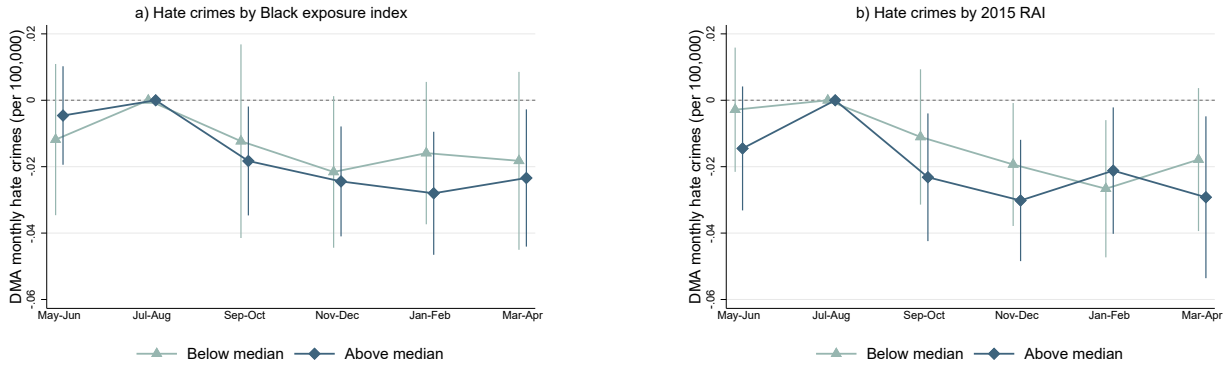


Figure A9: Heterogeneity in the hate crime season effects of Black quarterbacks

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.