Racial biases and market outcomes: "White men can't jump," but would you bet on it?

Igan, Deniz and Pinheiro, Marcelo and Smith, John

International Monetary Fund, Research Department, Economic Consulting Services, LLC, Rutgers University-Camden, Department of Economics

19 January 2012

Racial Biases and Market Outcomes: "White Men Can't Jump," But Would You Bet On It?*

Deniz Igan† Marcelo Pinheiro‡ John Smith§

July 24, 2012

Abstract

We identify an otherwise efficient market in which racial biases affect market outcomes. In particular, we examine data on point spreads for NBA games over the 15 seasons from 1993-94 to 2007-08. We find evidence that teams with more black players tend to face a larger point spread and that these teams perform worse against the spread. These biased outcomes are significantly large and persistent so that we are able to identify profit opportunities. We also find evidence that the biased spread is set by the bookmakers rather than being moved as a result of excessive betting on the more black team. These findings are consistent with information-based discrimination where mistaken beliefs persist even though they are financially disadvantageous, and, more importantly, recognizable and correctable.

Keywords: Market efficiency, Racial biases, Belief-based discrimination

JEL Classification: D03, G00, J15

*This paper was previously circulated under the name "'White Men Can't Jump,' But Would You Bet On It?" We would like to thank Gadi Barlevy, Stijn Claessens, Gianni De Nicolo, Marcelo Fernandes, Luc Laeven, Tim Larsen, Joseph Price, Michel Robe, Justin Wolfers, Jack Worrall, and the participants in seminars at IMF Research Department, American University, and Queen Mary, University of London for insightful comments and suggestions. Heejin Kim provided valuable research assistance. All remaining errors are our own. This research was supported by Rutgers University Research Council Grant #202344. The views expressed in this paper are those of the authors and do not necessarily represent those of the IMF or its policy.

†International Monetary Fund, Research Department, 700 19th St, NW, Washington, DC 20431. Phone: 202-623-4743. Fax: 202-623-4740. E-mail: digan@imf.org.

‡Economic Consulting Services, LLC; Email: marcelo@alumni.princeton.edu.

§Rutgers University-Camden, Department of Economics; Email: smithj@camden.rutgers.edu.

"Billy, listen to me, white men can't jump."

Sydney Deane

1 Introduction

Beliefs shaped by psychological dispositions and social norms or interactions are commonly recognized as important determinants of economic decision making and market outcomes. Becker (1957) and Arrow (1972) provide models depicting such beliefs in the context of discrimination while Akerlof (1980) and Romer (1984) study the persistence of customs. Most of the evidence on the effects of beliefs formed upon a psychological and social basis, or "biases", comes from the studies aiming to detect discrimination in labor markets (e.g., Bertrand and Mullainathan, 2004), in access to services (e.g., Page, 1995) and in access to resources, most notably credit (e.g., Munnell, Tootell, Browne, and McEneaney, 1996; Pope and Sydnor, 2011). But the study of psychological and social biases go beyond discrimination: a related literature documents systematic deviations from standard assumptions underlying economic behavior and links them to psychological and social factors.[1] Challenges remain in both strands of the literature: documenting whether discrimination exists rather than the observed differences stemming from unobserved heterogeneity, distinguishing information-based discrimination from taste-based discrimination, and understanding whether and how behavioral biases carry over from laboratory experiments to real markets as well as whether and how they persist instead of market forces eliminating such biases.

This paper offers new insights into these challenges by studying the effects of psychologically-based, socially-reinforced beliefs in a financial market setting. In particular, we examine the relationship between National Basketball Association (NBA) betting outcomes and the race of the participants in order to uncover how biases originating from psychological dispositions and social norms affect market outcomes. This is an ideal setting to expand our knowledge on the economics of biases for several reasons.

First, the NBA betting markets provide advantages that other settings, including other financial market settings, cannot. Specifically, bettors pay for their biased tastes à la Becker.

[1] See Camerer, Loewenstein, and Rabin (2004) and DellaVigna (2009) for a review.

Hence, decisions based on mistaken beliefs are punished with direct pecuniary losses. This is in contrast to studies of the impact of beliefs in psychology and sociology literatures, where most evidence relies on experiments or surveys with no immediate, explicit, pecuniary gains or losses for the participants. In addition, unlike most financial markets, the sports betting markets contain well-defined prices, well-defined outcomes, readily accessible information, and a finite time horizon. Therefore, usual caveats associated with measurement problems (e.g., defining the horizon over which returns should be measured) and asymmetric information do not apply. Moreover, the actions and outcomes are repeated with a significant degree of frequency, providing an opportunity to test whether the impact of biases persists or disappears as market participants learn about the bias and compete to grasp arbitrage opportunities.

Second, the beliefs in the market we study are easily recognized since some of the most deeply held ideas about race and racial difference are expressed in our beliefs about sports and athletic ability, creating one of the most well-known stereotypes: the natural black athlete, and especially, the black basketball star. The common perception that black people are better at basketball than people of other races or ethnicities is so evident that the term "the black game" was coined to refer to the sport (George, 1999). What makes it so difficult to counter the argument that blacks have an innate ability to play basketball is that there appears to be evidence to support it: roughly 70% of NBA players are black. As a result, the idea that blacks are better basketball players and the evidence that seemingly supports this idea can have far-reaching consequences. For economists, an interesting question arises when these observations become unwavering, subconscious attitudes that support the *mistaken* belief that athletic ability is inextricably tied to race and these attitudes affect economic decision making in a predictable manner, thus challenging the rationality tenet in its standard form.

Hence, in our setting, market outcomes are objective, common knowledge, determined within a finite time, and repeated regularly, and there exists a widely-familiar, biased view of the participants. Our data consists of the outcomes of NBA games and the Las Vegas point spreads on these games, from the 1993-94 season through the 2007-08 season.[2] Betting on NBA basketball generally involves a point spread wager, where the bet wins based on the

[2] Note that the *Professional and Amateur Sports Protection Act of 1992 (PASPA)* imposes a federal ban on sports betting in all states with the exception of Delaware, Nevada, Montana, and Oregon. These four states already had sports betting laws on their books when the Congress passed PASPA and were permitted to offer parlay-type sports betting. Nevada, however, exclusively allows all types of sports betting, statewide, on any professional or amateur sports games, in any capacity.

relationship between the final score and the point spread. We say that the team *covers* the spread if a bet on the team pays. To illustrate, if the spread is +3.5 for the home team, an $11 bet on the home team would pay $21 if either the home team won the game or lost by 3 points or less. An $11 bet on the visiting team would pay $21 only if the visiting team won by 4 points or more. In this setting, the point spread is a market-based estimate of the actual margin at the end of the game.

Our analysis provides evidence that biases embodied as stereotypes about the relative ability of a certain group of players have an impact on financial decisions by examining how the point spread and the performance against the spread in NBA betting markets varies with the racial composition of the teams. We ask whether the belief that black players are better than their white counterparts affects the point spread and hence the likelihood of beating the spread.[3] In other words, do "more black" basketball teams look better than "less black" teams in the sense that, all things equal, bettors are more inclined to bet on the former? If this is the case, then the spread on a more black team will be higher than it should be, leading to a negative relationship between the fraction of black players and the performance against the spread.

We find that the point spreads are higher for teams with a relatively higher fraction of black players. We also find that the probability of beating the spread decreases as the fraction of black players increases. Our results are robust to alternate measures of the racial composition of the team: the number of black players starting the game, the number of black players on the roster, and the minutes played by black players in recent games.

One important assumption in interpreting the finding that it is less likely for black teams to beat the spread concerns the efficiency of NBA betting markets. In other words, it is implicitly assumed that the spread incorporates all relevant and available information about the game. We confirm that this is indeed the case, by showing that the actual margin minus the spread is normally distributed with a mean of zero. Hence, unconditionally, any game has an equal probability of ending up with a score on either side of the spread.

Further, the notion that the ability difference is indeed a bias, and is not real, is confirmed by the empirical observation that performance measures for black and white players are not statistically different. In fact, white players tend to be taller and more efficient in the

[3] Perhaps with slight abuse of the term, we use "white" to refer to all non-black players.

sense that they score almost the same points as black players despite playing fewer minutes. However, there is no significant difference in terms of scoring ability between black players and their white counterparts. Similarly, there is no evidence of a robust relationship between the relative blackness of a team and its chances of winning the game. Hence, even in the absence of statistical evidence supporting the belief that black players are better, the bias exists and affects decisions made by agents.

There are two hypotheses for the cause of the relationship between race and the point spread. It could be that biased bettors place more money on the more black team, thus causing the spread to move from an unbiased spread to a biased spread. Or it could be the case that the bookmakers are aware of the bias of bettors, and set the spread in order to extract more surplus. Why would bookmakers set a biased spread? Levitt (2004) shows that bookmakers can increase their earnings if bettors have a bias. This is because the bookmakers can set the point spread in a manner such that more than half of the money is bet on the outcome which wins less than half of the time. In order to distinguish between these hypotheses, we use a second data set containing the opening and closing point spreads, for the 2003-04 season through the 2009-10 season.

Our results show that the opening spread does not move *at all* a quarter of the time and the difference between the closing spread and opening spread is normally distributed around zero. Moreover, the movement of the spread is not related to the racial composition of the teams in a statistically significant and robust manner. Hence, it appears to be the case that the bookmakers know of the bias towards more black teams and consider this when they set the spread. This is further supported by evidence that more money is bet on the more black team.

To gain intuition on our results, consider two teams which are exactly as good as each other, consequently each team will win with a probability of 0.5. However, one team is "more black" than the other. Therefore, some people will have a bias that the black team is better and deem their probability of winning to be greater than 0.5, even though "the truth" is 0.5. To exploit this bias, rather than setting the spread as a pick-em (spread of 0), the bookmaker sets the spread in favor of the black team at a value different from 0. This means that (all things equal) the black team will cover the spread with a probability less than 0.5, making this a worse bet. This reasoning still holds when the teams are not as good as each other. In this case, there is a "true spread" which each team will cover with probability 0.5. But the bookmakers do not set the spread at the true spread but rather the true spread adjusted by a few points for the black team. Again, the black team covers with probability less than 0.5.

Let us return to the case where both teams are equally good, so the expected final margin is zero. Further, let us assume that the spread at which an even amount of money would be placed both sides of the bet would be -3 for the more black team. In this case, since half of the money is on either side of the bet, the bookmakers' expected payoff is determined exclusively by the betting cost: for every \$11 bet, the winner gets \$21, that is, a return of \$10 and not \$11. Similarly if the spread is set at 0, the bookmakers' expected payoff is again determined exclusively by the betting cost.[4] The profit-maximizing spread is somewhere between 0 and -3. So, the bookmakers set the spread at, say, -2 and more money is bet on the black team because bettors think the spread should be -3. Since more than half of the money is bet on the outcome which occurs less than half the time, the bookmakers earn extra profits.

Our results imply that biases can indeed influence behavior in financial settings. Hence, we contribute to the literature by providing evidence that economic decision making is altered by conscious or subconscious categorization based on observable characteristics, e.g., race and gender. Additionally, the association between the point spread ("the price") and the racial composition of the teams (a variable that is not systematically related to the winning ability of a team and is observable prior to the bets being placed) creates profitable opportunities that involve betting on the "whiter" team. In other words, the bias is sufficiently large and persistent that we are able to identify a means of profiting from the biased market outcomes. Moreover, our findings are consistent with such biases being more likely to stem from information-based motives than from taste-based motives. This is because the bookmakers seem to incorporate these biases into prices and there is no reason to expect these particular agents to be of a different racial composition than the others. Also, if one presumes that bettors from a particular geographical region would be more inclined to bet on the team based in their region and the racial profile of bettors resembles the demographics of the region, then preference-based explanations would imply a negative relationship between the "black cities" and the probability of the more black home teams beating the spread. We do not, however, find evidence of such a relationship.

[4] For more on this, see Levitt (2004).

The rest of the paper is organized as follows. Section 2 discusses the related literatures on discrimination and financial markets, including sports betting. Section 3 provides an overview of the data. Section 4 presents the results and Section 5 concludes.

2 Background

Our paper relates to several strands of literature. The first of these strands examines biases and their impact on economic outcomes. A large number of studies look at discrimination, i.e., where outcomes depend on characteristics such as gender and race. Because of the inefficiencies discrimination may create and the potential policy implications, it is of great interest to identify settings in which biases exist and, once they are identified, to specify the mechanism which is causing the bias.[5] Often, it is very difficult to find unambiguous evidence of biased outcomes, mostly due to the omitted variables problem. That is, ruling out the possibility that the observed variation may be a consequence of unobserved heterogeneity, which is also correlated with the object of study (in many cases, gender or race), is a difficult task. In response to this problem, some researchers have used audit studies whereby the investigators send identical treatments into the field, with exception that they differ on the basis of, say, race. Then the researchers seek to observe differences in behavior on the basis of race. For instance, in their influential study, Bertrand and Mullainathan (2004) sent otherwise identical resumes to potential employers, where some applicants had "white" names and some had "black" names. The authors found that applicants with black names were less likely to receive a callback for an interview than were the applicants with white names. Audit studies, such as this one, have proven to be useful in identifying biased settings.[6] There are however, some drawbacks with audit studies.[7] First, it is argued (Heckman, 1998) that these studies overstate the effect of discrimination because they do not account for the effects of the unbiased people on the market outcomes. In other words, these audit studies can identify that some behave in a biased fashion, however it is possible that the unbiased people can behave in a way which mitigates the effects of the behavior of the biased people. Our paper is not vulnerable

[5] For more on discrimination literature, see, among others, Altonji and Blank, 1999, Ross and Yinger, 2002, and Charles and Guryan, 2008.

[6] See Ayers and Siegelman (1995) for another example of this type of technique.

[7] For more on the difficulties with audit studies, see Yinger (1998).

to this criticism because the object of our study is not individual behavior but rather market outcomes and we find that the market outcomes are biased. Second, to our knowledge, audit studies are not repeated whereby the decision maker can learn about the unobserved heterogeneity of the subject. Again, we are not vulnerable to this objection because we have a considerable number of observations for the same players and teams, whereby the effects of the unobserved heterogeneity could be learned. Therefore, we find that the market is persistently biased even with the opportunity to learn otherwise, and even though there are pecuniary costs to behaving based on these mistaken beliefs.

The phenomenon we study is a product of "positive stereotypes" and can perhaps be more accurately labeled as "reverse discrimination" since the group which is deemed to be superior faces odds that are more difficult to overcome. In other words, the belief that black basketball players are better creates a bias for betting on the more black team and it becomes more difficult for the black team to beat the spread.[8] Still, one could think of this phenomenon in terms of the main theories of discrimination in the microeconomics literature. In the first of these theories, differential behavior towards a certain group of individuals is driven by the preference for not interacting with them (Becker, 1957; Arrow, 1973). In other words, individuals have a "taste" for their own kind or a distaste for the other kind. In the second theory, agents take race to be a signal for unobserved or costly information about skill levels and mistaken beliefs can survive if they create self-fulfilling outcomes (Phelps, 1972). In our context, information-based explanations would be more relevant if bets reflected the prior belief that blacks are better at basketball while the findings would fit the taste-based explanations if white bettors bet against more black teams. Unfortunately, we do not have information on the race of the individual bettors but we use the demographic characteristics of the region which hosts the team to indirectly address this issue. The lack of a significant relationship between the racial composition of the host communities and that of the teams supports the information-based, rather than taste-based, explanations.

Our paper relates to the literature documenting and explaining market anomalies in finance as well. Closely related to our premise of studying the impact of a common perception in a financial market setting, Hong and Kacperczyk (2009) and Hong and Kostovetsky (2012)

[8] Cheryan and Bodenhausen (2000) provide evidence that stereotypes can lead to such a "choking effect" by looking at the performance of Asian-American women in math tests.

look at case of the "sin stocks" and political values in investment decisions. Wolfers (2006a) examines the stock market returns of companies with female CEOs. Sports betting markets, in particular, provide an attractive ground for testing market efficiency because, unlike most financial markets, the sports betting markets contain well-defined prices, well-defined outcomes and a finite time horizon. In particular, sports betting markets have outcomes which are realized within a short time frame, are observable by all market participants, and are unambiguous (no measurement error or uncertainty about the horizon over which outcomes should be measured). Finally, these markets are unlikely to have uninformed traders due to the widespread availability of information. Therefore, the questions related to the efficiency of the sports betting markets are of interest to economists in order to test market efficiency hypotheses in general. Echoing findings in other financial markets, several studies have found inefficiencies in the sports betting markets.[9] For instance, studies have found that bettors erroneously place bets for sentimental reasons (Avery and Chevalier, 1999; Braun and Kvasnicka, 2012; Forrest and Simmons, 2008), on teams which are deemed *hot* (Brown and Sauer, 1993; Camerer, 1989), and on teams which are favorites (Golec and Tamarkin, 1991; Grey and Grey, 1997). Levitt (2004) finds, using data on the wagers placed by bettors as part of a handicapping contest offered at an online sports book during the 2001-02 NFL season, that the amount of money placed on each side of the bet is not equal and this imbalance is related to observable information. In particular, Levitt finds that the proportion of money bet is higher for favorites and road teams. Interpreting this finding, he argues that the bookmakers set the spread in order to exploit common biases: people like favorites and people do not sufficiently account for the home field advantage.[10]

The particular bias we study, i.e., the common belief that black players are better than their white counterparts, has been the subject of experimental investigation of the effects of stereotypes on judgments in sports.[11] For instance, Stone, Perry and Darley (1997) directed subjects to listen to an audio clip of a basketball game after viewing a picture of the player

[9] See Barberis and Thaler (2002) for a general overview and Sauer (1998) for applications in sports betting.

[10] Paul and Weinbach (2011) corroborate this finding using the percentage of bets actually placed on NFL games. Our analysis shows that the bets on NBA games are also distorted by racial stereotypes. Also see Kuypers (2000). Snowberg and Wolfers (2010) discuss the evidence that, in the odds betting of horse racing, bettors have a bias towards betting on longshots rather than on favorites.

[11] For more on the stereotype of the athletic black man, see Biernat and Manis (1994), Sailes (1996), and Stone, Lynch, Sjomeling, and Darley (1999).

whom they were instructed to judge. The subjects who were shown a picture of a black player rated the performance as better than those subjects who were shown a picture of a white player. While existing experiments are suggestive of biases in judgments involving race and athletic performance, since the accuracy of these judgments are not related to the material incentives of the subjects, it can be difficult to interpret these results. However, our study is not vulnerable to this critique because obviously betting on the outcome of a basketball game is indeed related to a person's material incentives.

Others have also looked at the effect of race on outcomes in sports. Again, this literature is significant beyond the sports context because it involves decisions which exhibit large incentives for success or accuracy and the outcomes can be objectively measured. Price and Wolfers (2010) find a negative relationship between the personal fouls assessed against NBA players and the number of own-race referees who officiated the game. Relatedly, Parsons, Sulaeman, Yates and Hamermesh (2011) find that the likelihood of a called strike in baseball is related to the agreement of the pitcher's and umpire's race. Although these judgments are made by well-trained and experienced professionals, they are also made under great duress and must be made almost instantaneously. Therefore, it is possible that these judgments, while obviously of great significance, would be attenuated if they were made under different circumstances. By contrast, the judgments which comprise the data we provide are made by individuals who have the opportunity to reflect on the merits of their decisions. Hence, our findings imply that racial stereotypes may affect decisions even when they are made under an extended period of deliberation.

In a similar vein, Larsen, Price, and Wolfers (2008) find that the relationship between race and fouls documented in Price and Wolfers (2010) is significant enough so that, given information about the race of the referees and the relative racial composition of the teams, one could improve their chances of placing a winning bet against the spread. By contrast, we focus primarily on the racial composition of the teams. Hence, the bias we examine emerges from the simple observation that many believe black players to be better, rather than the less visible notion that the referees exhibit an own-race bias. In addition, we analyze the opening and closing point spreads and we find evidence that the bookmakers are aware of the bias, thus suggesting that the phenomenon is more likely to be driven by information-based motivations rather than by taste-based explanations. In other words, we provide evidence consistent with the assertion that the bettors may be taking the racial composition of the teams as a signal to guide their betting decisions (statistical discrimination à la Phelps). The evidence also supports the claim that the bookmakers incorporate all relevant information which may not be reflected in the racial discrepancy between the teams and set the spread so that they can exploit the information-based bias of the bettors. Finally, as does Larsen, Price, and Wolfers (2008), we offer an analysis of a simple betting strategy. The simple betting strategy proposed by Larsen, Price, and Wolfers (2008) involves the interaction of the differences in the race of the teams and the referees, and in our case it is exclusively a function of the racial composition of the teams. Hence, arguably, our strategy requires less information and is less computationally-intensive than theirs. Our betting strategies prove to be at least as profitable, and often more so, than the ones analyzed in Larsen, Price, and Wolfers (2008). However, despite the differences between Larsen, Price, and Wolfers (2008) and the present paper, we view our work as offering a complementary investigation into the relationship between race and betting markets.

3 Data

Our baseline dataset combines box score information on all regular season NBA games played from the 1993-94 season to the 2007-08 season. We exclude the playoff games since the outcomes for these games tend to be path-dependent not only across games in the same series but also across rounds, thus accentuating the survivorship bias in the sense that the number of player or team observations would closely depend on their past performance. The box score information is obtained at the player-game level from www.basketball-reference.com, which also keeps track of draft picks and other background information of the players, such as the height and weight. The ultimate team-game level dataset is constructed from these player-by-player observations, obtained from www.basketball-reference.com.

One crucial variable, however, for our analysis that is missing from the www.basketball-reference.com website is the race of the players. In some cases (mostly for players who are still active), a picture of the player accompanies the statistics but this happens only at a small fraction of the overall player universe during our sample period. Hence, we conduct an extensive search to obtain information on the race of the players, navigating www.nba.com, www.hoopedia.nba.com, www.draftreview.com, and images found via Google. This information enables us, by visual inspection, to characterize the racial membership of the players. Admittedly, we use a rather coarse definition of race by assigning players (and coaches) into two broad categories of black and white, where white includes Caucasians, Asians, and Latinos. Yet, we use several measures of the racial composition of the team in order to ensure robustness of the results involving a variable as subjective as a player's race. Further, we also double-check our classification of the racial membership of the players against that used in Price and Wolfers (2010). The discrepancy between the racial classification exists for a mere 31 out of 1128 matched players. This difference corresponds to only 2.5 percent of the more than a quarter of a million player-game observations used in our dataset.

The data for the point spreads are obtained from www.goldsheet.com. We verify the accuracy of the spreads from this source against other sources commonly-used in the academic studies of sports betting, such as www.covers.com, and find no significant discrepancies. In fact, information on the ultimate outcomes of the games tends to be more accurate in www.goldsheet.com than it is in www.covers.com: of the 41 cases when a discrepancy between the two sources exists, the cross-check with www.espn.com confirms that the former has the correct information 80 percent of the time. In the absence of an obvious third source to check the point spreads against, we ultimately use the two data sources as cross-checks against each other in constructing our final dataset and eliminate the observations in which a discrepancy between the two sources exist. We complement this information on the closing spreads with information on the opening spreads and the percent of bets placed on each side of the bet.[12]

A total of 18,450 regular-season games were played during the sample period. After excluding games for which there is a missing box score or racial composition data, we are left with 17,178 games. Further, after excluding games for which there was either no betting information or contradictory betting information, or where the betting outcome was a push, leading to cancellation of all bets (which occurs approximately 1.3 percent of the time), we are left with 14,785 games in the sample. Before we move to the formal analysis, we present some descriptive statistics of this final dataset.

Of the 1021 players who were active in the NBA during our sample period, 71.8 percent are black. Black players are even more over-represented in the starting line-up of the teams:

[12] These data are available, at a fee, from www.sportsbetting.com.

on average, only one out of five starters is white. In a typical game, each team deploys 9 to 11 players, 8 of which are, on average, black. As a result, at the player-game level, 76.7 percent of the observations are identified as being associated with a black player. These statistics confirm the casual observation of the dominance of black players in the NBA, not only by sheer number but also by the visibility they obtain by playing more minutes in more games.

Table 1 provides a summary of the data used in our analysis at the player-game level, and Table 2 summarizes the data at the team-game level. At the player-game level, there are some statistically significant differences between black and white players. However, it is not always the case that black players have "more desirable qualities" and the magnitudes of these differences are not very meaningful. For instance, while, on average, black players score roughly two points more than their white counterparts, they are not as efficient as demonstrated by their slightly lower field goal percentages. According to these metrics, black players overall do not appear to be much better than their white peers. If one assumes that the team is a sum or reflection of the skill levels of individual players, there seems to be no obvious statistical reason to deem more black teams to be of better quality.

At the team-game level, on which we conduct the empirical analysis, we summarize the information on betting spreads and the racial composition of the teams. Racial composition is measured by three alternative metrics: the number of black starters, the number of black players on the team roster regardless of whether they actually play in a game, and the minutes played by black players. This final metric is calculated as the average of the past five games the team has played and is expressed as a percentage of the total minutes in the game. To avoid duplication, all variables are expressed from the home team's perspective. Simple statistics point to a slight advantage for the home team as they win the game 60 percent of the time by an average margin of around 4 points. Point spreads seem to take this into account at least partially because the home team is the favorite about 70 percent of the time and beats the spread 51 percent of the time. Note that the *partial* offset of the home court advantage is in line with earlier studies showing a similar bias in NFL betting markets (Levitt, 2004).

4 Analysis

4.1 Accuracy of point spreads and the link between race and winning probability

In order to demonstrate the relationship between performance against the spread and the racial makeup of the teams, we estimate the following regression:

$$P(\text{home team beats the spread})_{it} = \alpha + \beta\Delta black_{it} + \gamma X_h + \varphi Y_s + \phi(X_h * Y_s) + \varepsilon_{it}$$

where the dependent variable is a dummy variable which takes on the value 1 if the home team beats the spread on game i played at date t and 0 otherwise, $\Delta black_{it}$ is the difference between the "blackness" of the home team and the visiting team, X_h and Y_s are (home) team and season (during which date t is included) fixed effects, respectively. As noted in the previous section, the blackness of a team is measured by the various metrics (number of black starters, number of black players on the team roster, and the fraction of minutes played by black players in the previous five games).

Thus, our empirical approach rests on a baseline specification where the probability that the home team beats the spread is a function of the racial composition of the team relative to its opponent. This relies on two assumptions and, before moving on to the main analysis, we confirm that these assumptions hold.

The first assumption is that basketball betting markets are, in general, efficient, in that any observable information should be reflected in the spread. So, we begin our analysis by looking at the accuracy of point spreads in forecasting the game outcome. Figure 1 shows the distribution of "forecast errors," defined as the actual margin (or realized spread) minus the point spread on a game. Indeed, the errors closely resemble a normal distribution with zero mean.[13] Figure 2 formally verifies this statement by plotting the forecast error against a normal distribution.[14] We find that the NBA betting markets are, in general, efficient in the

[13] See Wolfers (2006b) who examines the distribution of errors in college basketball games and finds evidence of point shaving in games with a large point spread.

[14] Kolmogorov-Smirnov equality-of-distributions test as well as skewness and kurtosis test for normality further verify that forecast errors are normally distributed. Results of these tests are available from the authors upon request.

sense that the distribution of the difference between the winning margin and the point spread is not distinguishable from a normal distribution. In line with this, when plotted against the realized winning margins, one can see that the point spread is an accurate forecast of the actual game outcomes (Figure 3).

The second assumption is that the probability of winning a game does not increase in the relative blackness of the teams. Table 3 presents the results of a regression analysis where the more black team in a match-up is shown not to have a systematically higher probability of winning a game. The sign on the variables of interest, i.e., blackness of the home team relative to the visiting team, varies from one specification to the next and is not always significant and positive when the dependent variable is the realized margin on the game (upper panel in Table 3). Therefore, there is little evidence of a positive association between the blackness of the teams and the decisiveness of the final scores. A quick glance at the table would suggest a somewhat robust *negative* relationship between the blackness of the teams and the probability of winning (lower panel in Table 3).[15] It should be noted that this is not necessarily a sign of lower quality or generally worse performance of teams composed of more black players against teams with more white players. Rather, in these specifications, the relative blackness of a team may be capturing the effect of other factors which determine the performance of one team against another. Indeed, once factors such as the record of the team up to a specific game in a season is controlled for, the magnitude and significance of this coefficient is weakened.[16] In summary, our assumption that the probability of winning a game does not *increase* with the differences in racial composition towards blackness has support in the data.

With the two assumptions verified, we now proceed to the regression analysis of the point spread and actual game outcomes.

4.2 Race and point spreads

Table 4 presents our main findings. In a nutshell, our analysis shows that teams which are more black tend to face a higher point spread and that these teams exhibit a worse performance

[15] Notice that the team with more black starters is likely to have a larger realized margin but a lower probability of winning. While this seems a bit curious, it is consistent with a few outliers where the team with more black starters had a blowout when they won. Also note that the association between the differences in the blackness of the teams and the realized margin of the game is not robust as the positive significant coefficient disappears when alternative measures of blackness are used.

[16] These results are not presented here for sake of brevity but are available from the authors upon request.

against the spread. Note that, in each regression, team fixed effects and season fixed effects as well as team-season interactions are employed. Hence, neither the time-invarying team characteristics nor the team-invarying time effects are driving the results.[17]

In the upper panel, the dependent variable is the spread faced by the home team. According to our three measures of the racial differences between the teams, we see that there is a positive relationship between the spread and these measures. In the middle panel, the dependent variable is the realized margin of the home team minus the spread. Based on the three measures of the racial differences between the teams, we see that there is a negative relationship between the blackness of the team and the realized margin minus spread. In the lower panel, the dependent variable is a dummy which takes the value of 1 if the home team beats the spread and zero otherwise. Again, according to our three measures of the racial differences between the teams, we find a negative relationship between the blackness of the team and the probability that they cover the spread. To summarize, we find evidence that teams which are more black tend to face a larger point spread and that these teams perform worse against the spread. The evidence so far supports part of the conjecture we introduced at the beginning: point spreads, even as they control for all relevant and available information on the two teams facing each other, are disproportionately higher for more black teams, consistent with the belief that they are better than those with more white players.

4.3 Biased bettors or biased bookmakers?

A natural question then is, what is driving the relationship between the racial composition of the teams and the performance against the spread? There are two main competing hypotheses. The first hypothesis is that the bookmakers are aware of the racial bias of bettors and they set the spread in such a way to exploit the bias à la Levitt (2004). The second hypothesis is that the bookmakers are unaware of the bias of the bettors and set the spread to be the expected final score of the game and the relationship found above is caused by bettors who systematically bet on the more black team, thus moving the spread. In order to distinguish between these hypotheses, we investigate whether there is a relationship between the movement of the spread

[17] The results presented in the tables are estimated using probit when the dependent variable is a binary variable, e.g., the probability of beating the spread. To ensure that the results do not suffer from the incidental parameters problem, we also estimate these specifications using ordinary least squares. The sign and significance of the coefficients of interest are indeed robust to the choice of estimation method.

and the racial composition of the teams. Figure 4 demonstrates that the movement of the spread is normally distributed with a mean of zero.

Table 5a presents the results of our regressions involving the movement of the spread.[18] In our first specification, we do not account for team- and season-specific factors, or team-season interaction terms. There we find a significant relationship between race and the movement of the spread. However, for the three specifications in which we do account for these fixed effects, we do not find a significant relationship between the race of the teams and movement of the spread. Moreover, from Table 4, we know that there is a systemic relationship between the relative racial composition of the teams and the spread itself. Hence, the spread reflects the belief that teams with more black players should be placed as favorites. Indeed, more (less) money appears to be bet on the home team if it has more (less) black players than the visiting team, as shown in Figure 5. To further this argument, we investigate whether the money bet on the home team is related to the relative racial composition of the teams. Table 5b shows the results of this exercise. There appears to be a positive, albeit statistically weak, relationship between the racial composition of the team and the percent of money bet in favor of the home team. In other words, the spread is set in a way that the resulting bets are skewed in favor of the more black team. Based on the regressions in Tables 5a and 5b, we favor the explanation that bookmakers are aware of the bias of bettors and set the spread to exploit this bias.

4.4 Robustness checks

How robust are our results? We perform several robustness checks where we control for the race of the referees, the race of the coaches, and the racial composition of the location of the home team.

First, one concern is that our results no longer hold when one accounts for the racial composition of the referee crew. For instance, Larsen, Price, and Wolfers (2008) find that the racial composition of the referee crew, together with the racial composition of the teams, is relevant and can affect the probability of a team winning the game and, hence, beating the spread. Specifically, the authors find that teams can become disadvantaged when the

[18] In the remainder of the tables, for the sake of brevity, we only show the results involving the difference in black starters. The results are virtually identical when the other two metrics are used and are available from the authors upon request.

racial composition of the referee crew differs from the racial composition of the team. We perform a series of regressions with the dependent variable as the probability of beating the spread, however, we restrict attention to the following categories: an all-white crew, a crew with at least one black referee, a crew with at least one white referee, and an all-black crew. We also consider the case where the crew is neither all black nor all white. Finally, we add the proportion of white referees as an additional control variable in our baseline specification. Table 6a presents the results of these regressions. Even when accounting for the racial composition of the referee crew, our results remain significant in each case, with the exception of an all-black crew. However, note that an all-black crew is an extremely rare occurrence as it accounts for only 126 games out of 14,694 in our sample. Hence, in the majority of the games in our sample, it remains true that it is more difficult for the more black team to beat the spread.

Second, another important factor could be the race of the coaches. One could imagine that when the home team has more black players and, according to the beliefs of some, is more talented, there can be an additional bias if bettors also think that a more black team led by a black head coach should do even better than they would when led by a white head coach. We follow a similar approach as when controlling for the racial composition of referees and split our sample by the difference in the race of the head coaches of the two teams. Again, we also run a regression where the difference between race of the home team's head coach and that of the visiting team's head coach is introduced as an additional control variable. Results of these regressions are presented in Table 6b. Interestingly, more black home teams actually have a better chance of beating the spread when their head coach is white but the visiting team is led by a black head coach. Regardless, note that when the coaches are of the same race (a majority of the cases), the negative relationship between the blackness of the team and the probability of beating the spread prevails. Finally, our finding from the baseline regression also holds when the difference between the race of the head coaches is controlled for as an additional regressor in the specification.

It is also possible that the biases found above are related to the racial composition of the bettor or the racial composition of the location of the basketball team. While we cannot account for the race of the bettors, we can control for the racial composition of the location of the teams. Therefore, to the extent that a person living in the location of the team is more likely to bet on the team, we can test whether our results are driven by the characteristics of the population at the locations of the teams. We run a pair of probits with probability of beating the spread as the dependent variable. We include the difference in black starters as an independent variable, while accounting for the racial composition of the location of the teams. In particular we account for the difference in the proportion of blacks in the city and difference in the proportion of blacks in the state. In Table 6c, we present the results of these regressions. In all specifications, the difference in black starters remains significant, and neither of the terms accounting for the racial composition of the location are significant. As a result, we do not find evidence that the racial composition of the locations are related to the bias found above.

On a related question about the race of the *audience* and the team, Kanazawa and Funk (2001) find that the television ratings of games are positively related to the fraction of white players on the teams. This is seemingly at odds with the evidence that white teams are perceived to be worse at basketball. Presumably spectators attend basketball games or watch on television in order to see "good basketball." If this was the case, there would be a negative relationship between the white composition of the teams and television ratings for that game. Yet, this finding could be explained if the majority of NBA fans were white and, while white fans thought that black players were better, they still prefered to watch the white players, leading to a "premium" for white players. This is in line with the own-race preference, which would predict that white audiences choose to watch white players as they derive utility from associating with them even if they perceive the overall quality of the basketball played by these players to be inferior. We do not, however, find an analogous relationship between betting on the more black home teams in locations with a higher proportion of black population. This may imply that, when taking financial decisions directly associated with basketball, audiences stick to the stereotypes, perhaps relying on them as informative signals.

Finally, our finding may not survive if performance criteria of the teams or the factors that may be affecting each team's performance against specific opponents are explicitly included in the specification. As noted earlier, in our baseline, we control for time-invarying team characteristics and team-invarying time effects. But the performance and, relatedly, the morale of a team may vary through a given season or when faced with a particular opponent, e.g., because their game strategies are similar or because the bettors perceive a match between two specific teams differently from others. Another issue could be that bookmakers correct any systemic mistakes which may occur in setting the spread as the same two teams face each other again and again. Table 6d presents the results obtained when the difference in the records and winning streaks (i.e., the number of games the team won out of the last five games played) of the teams are added to the specification and fixed effects for specific team pairings are included. Our finding that the more black team has a lower probability of beating the spread is robust to these checks.

4.5 Profit opportunities

So far, we have presented evidence that there is a negative relationship between the relative blackness of a team and its probability of beating the spread. The question then is whether there are profitable strategies which consistently yield returns over the break-even hurdle. Accounting for the cost of betting, the break-even hurdle requires a winning percentage higher than 52.4 percent. We consider three simple strategies in Table 7: betting on the team with more black players, betting on the team with more white players, and betting on the home team only when it has more white players than the visiting team.[19] The reason for distinguishing between the home and visiting team in the last strategy is to exploit the possibility that the home court advantage may not be fully accounted for by the bettors. Yet, betting on the whiter team regardless of where the game is delivers with a winning percentage of 54 percent and a return of 3 percent. Applying the "bet on the white" strategy while also reflecting on the home court advantage, improves the returns: The bets placed this way, on average, win 58 percent of the time for an overall average return of 10 percent.

Across seasons, the profits obtained by following the "bet on the white home team" strategy are persistent over time. In the 2007-08 season, as a result of the strategy of only betting on the home team when it has 1, 2, 3, or 4 more white players in the starting line-up than the visiting team, we observe the probability of a winning bet to be as high as 75 percent and net returns (accounting for the cost of betting) to range from 8 percent to 43 percent. Table 8 presents the results of adopting this strategy over our whole sample period. Indeed,

[19] One could, of course, design betting strategies based on the roster of the rival teams or the minutes played by black players in the last few games. We obtain similar results using such strategies but prefer to present the results with the black starters on each team because this is the statistic which is the most easily-accessible and the least computationally-intensive.

this strategy of betting on the home team when it has more white players delivers positive net returns that not only increase with the starkness of the racial difference between the two teams but also broadly persist from one season to the next. This may suggest that, although learning opportunities abound, the bettors' behavior still reflects the mistaken belief that teams with more black players are better, akin to the finding in Pope and Schweitzer (2011).

5 Conclusion

This paper examines the impact of the positive stereotype of the black basketball star on financial decisions, as revealed in the market outcomes, using evidence from sports betting markets. We find evidence of a bias in NBA betting markets based on race. We also find evidence that this bias is exploited by the bookmakers. This finding can be explained by bettors taking race as a signal of skill level in deciding on which team to bet and bookmakers setting the point spreads higher for more black teams. An implication of our findings is that stereotypes, and biases in general, may affect financial decisions and, hence, market outcomes.

These findings add to the literature showing the importance of biases in economic decision making. In particular, we demonstrate that market makers process the available information efficiently but at the same time, when setting the prices, allow for the fact that the participants have a bias, which they do not correct, even though not doing so leads to direct pecuniary losses. Information-based discrimination is more likely to explain the phenomenon in the case we study since the prices do not adjust much after being initially set. We also provide evidence that biases do indeed carry over from laboratory experiments to real markets even when stakes are high and the agents have the time and opportunity to learn.

What do these findings mean for other economic markets? If we find persistently-mistaken, financially-disadvantageous beliefs in a market with obviously- and immediately-realized financial costs and many opportunities to learn, then we would expect there to be such in other markets. Most straightforwardly, do presumptions about intellectual or athletic ability based on stereotypes increase or decrease the odds of success for certain groups in certain fields? Another, perhaps a socially and politically uncomfortable question which may arise from this analysis is, if people are prone to making suboptimal sports betting decisions due to racial stereotypes, do people make similar costly judgment errors in other economic decisions? For instance, do employers hire engineers with a background from a particular region presuming that they have an innate ability for quantitative tasks? Is provision of health, education, and other social services affected by subconscious attitudes towards some groups? These and other interesting questions are left for future research.

References

[1] Akerlof, George A, 1980, "A Theory of Social Custom, of Which Unemployment May be One Consequence," *Quarterly Journal of Economics*, Vol. 94, No. 4, pp. 749-775.

[2] Altonji, Joseph and Rebecca Blank, 1999, "Race and Gender in the Labor Market," *Handbook of Labor Economics*, Vol. 3, pp. 3143-3259.

[3] Antonovics, Kate and Brian G. Knight, 2009, "A New Look at Racial Profiling: Evidence from the Boston Police Department," *Review of Economics and Statistics*, Vol. 91, No. 1, pp. 163-177.

[4] Arrow, Kenneth J., 1973, "The Theory of Discrimination," in Orley Ashenfelter and Albert Rees, eds. *Discrimination in Labor Markets,* Princeton, NJ: Princeton University Press, pp. 3-33.

[5] Avery, Christopher and Judith Chevalier, 1999, "Identifying Investor Sentiment from Price Paths: The Case of Football Betting," *Journal of Business*, Vol. 72, No. 4, pp. 493-521.

[6] Ayres, Ian and Peter Siegelman, 1995, "Race and Gender Discrimination in Bargaining for a New Car," *American Economic Review*, Vol. 85, No. 3, pp. 304-321.

[7] Becker, Gary, 1957, *The Economics of Discrimination*, Chicago, IL: University of Chicago Press.

[8] Bertrand, Marianne and Sendhil Mullainathan, 2004, "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *American Economic Review*, Vol. 94, No. 4, pp. 991-1013.

[9] Biernat, Monica and Melvin Manis, 1994, "Shifting Standards and Stereotype-Based Judgments," *Journal of Personality and Social Psychology*, Vol. 66, No. 1, pp. 5-20.

[10] Braun, Sebastian and Michael Kvasnicka, 2012, "National Sentiment and Economic Behavior: Evidence From Online Betting on European Football," forthcoming in the *Journal of Sports Economics.*

[11] Brown, William O. and Raymond D. Sauer, 1993, "Fundamentals or Noise? Evidence from the Professional Basketball Betting Market," *Journal of Finance*, Vol. 48, No. 4, pp. 1193-1209.

[12] Camerer, Colin F., 1989, "Does the Basketball Market Believe in the 'Hot Hand'?," *American Economic Review*, Vol. 79, No. 5 (Dec., 1989), pp. 1257-1261.

[13] Camerer, Colin F., George Loewenstein, and Matthew Rabin, eds. 2004, *Advances in Behavioral Economics*, Princeton, NJ: Princeton University Press.

[14] Charles, Kerwin K. and Jonathan Guryan, 2008, "Prejudice and Wages: An Empirical Assessment of Becker's The Economics of Discrimination," *Journal of Political Economy*, Vol. 116, No. 5, pp. 773-809.

[15] Cheryan, Sapna and Galen V. Bodenhausen, 2000, "When Positive Stereotypes Threaten Intellectual Performance: The Psychological Hazards of Model Minority Status," *Psychological Science*, Vol. 11, No. 5, pp. 399-402.

[16] Darity Jr., William A. and Patrick L. Mason, 1998, "Evidence on Discrimination in Employment: Codes of Color, Codes of Gender," *Journal of Economic Perspectives*, Vol. 12, No. 2, pp. 63-90.

[17] DellaVigna, Stefano, 2009, "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, Vol. 47, No. 2, pp. 315–72.

[18] Entine, Jon, 2000, *Taboo: Why Black Athletes Dominate Sports And Why We Are Afraid To Talk About It*. New York, NY: Public Affairs.

[19] Forrest, David and Robert Simmons, 2008, "Sentiment in the Betting Market on Spanish Football," *Applied Economics*, Vol. 40, pp. 119-126.

[20] George, Nelson, 1999, *Elevating the Game: Black Men and Basketball*, Lincoln, NE: University of Nebraska Press.

[21] Golec, Joseph and Maurry Tamarkin, 1991, "The Degree of Inefficiency in the Football Betting Market," *Journal of Financial Economics*, Vol. 30, pp. 311-323.

[22] Grey, Philip K. and Stephen F. Grey, 1997, "Testing Market Efficiency: Evidence From The NFL Sports Betting Market," *Journal of Finance*, Vol. 52, No. 4, pp. 1725-1737.

[23] Heckman, James J., 1998, "Detecting Discrimination," *Journal of Economic Perspectives*, Vol. 12, No. 2, pp. 101-116.

[24] Hong, Harrison and Marcin Kacperczyk, 2009, "The Price of Sin: The Effects of Social Norms on Markets," *Journal of Financial Economics*, Vol. 93, pp. 15–36.

[25] Hong, Harrison and Leonard Kostovetsky, 2012, "Red and Blue Investing: Values and Finance," forthcoming in the *Journal of Financial Economics*.

[26] Kanazawa, Mark T. and Jonas P. Funk, 2001, "Racial Discrimination in Professional Basketball: Evidence from Nielsen Ratings," *Economic Inquiry*, Vol. 39, No. 4, pp. 599-608.

[27] Kuypers, T., 2000, "Information Efficiency: An Empirical Study of a Fixed Odds Betting Market," *Applied Economics*, Vol. 32, pp. 1353-1363.

[28] Larsen, Tim, Joseph Price and Justin Wolfers, 2008, "Racial Bias in the NBA: Implications in Betting Markets," *Journal of Quantitative Analysis in Sports*, Vol. 4, No. 2, pp. 1-19.

[29] Levitt, Steven D., 2004, "Why are Gambling Markets Organised so Differently from Financial Markets?," *Economic Journal*, Vol. 114, No. 495, pp. 223-246.

[30] Levitt, Steven D., 2004, "Testing Theories of Discrimination: Evidence from Weakest Link," *Journal of Law and Economics*, Vol. 47, No. 2, pp. 431-453.

[31] Munnell, Alicia H., Geoffrey M. B. Tootell, Lynn E. Browne and James McEneaney, 1996, "Mortgage Lending in Boston: Interpreting HMDA Data," *American Economic Review*, Vol. 86, No. 1, pp. 25-53.

[32] Page, Marianne, 1995, "Racial and Ethnic Discrimination in Urban Housing Markets: Evidence from a Recent Audit Study," *Journal of Urban Economics*, Vol. 38, No. 2, pp. 183-206.

[33] Parsons, Christopher A., Johan Sulaeman, Michael C. Yates and Daniel S. Hamermesh, 2011, "Strike Three: Discrimination, Incentives and Evaluation," *American Economic Review*, Vol. 101, No. 4, pp. 1410-1435.

[34] Paul, Rodney J. and Andrew P. Weinbach, 2011, "NFL Bettor Biases and Price Setting: Further Tests of the Levitt Hypothesis of Sportsbook Behavior," *Applied Economics Letters*, Vol. 18, No. 2, 193-197.

[35] Phelps, Edmund S., 1972, "The Statistical Theory of Racism and Sexism," *American Economic Review*, Vol. 62, pp. 659-61.

[36] Pope, Devin G. and Maurice E. Schweitzer, 2011, "Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes," *American Economic Review*, Vol. 101, pp. 129–157.

[37] Pope, Devin G. and Justin R. Sydnor, 2011, "What's in a Picture? Evidence of Discrimination from Prosper.com," *Journal of Human Resources*, Vol. 46, No. 1, pp. 53-92.

[38] Price, Joseph and Justin Wolfers, 2010, "Racial Discrimination among NBA Referees," *Quarterly Journal of Economics*, Vol. 125, No. 4, pp. 1859-1887.

[39] Romer, David, 1984, "The Theory of Social Custom: A Modification and Some Extensions," *Quarterly Journal of Economics*, Vol. 99, No. 4, pp. 717-727.

[40] Ross, Stephen L. and John Yinger, 2002, *The Color of Credit: Mortgage Discrimination, Research Methodology, and Fair-Lending Enforcement*, Cambridge, MA: MIT Press.

[41] Sailes, G., 1996, "An Investigation of Campus Stereotypes: The Myth of the Black Athletic Superiority and the Dumb Jock Stereotype," in R. E. Lapchick, ed. *Sport in Society: Equal Opportunity or Business as Usual?* Thousand Oaks, CA: Sage Publishers, pp. 193-202.

[42] Sauer, Raymond D., 1998, "The Economics of Wagering Markets," *Journal of Economic Literature*, Vol. 36, No. 4, pp. 2021-2064.

[43] Snowberg, Erik and Justin Wolfers, 2010, "Explaining the Favorite-Longshot Bias: Is it Risk-Love or Misperceptions?," *Journal of Political Economy*, Vol. 118, No. 4, pp. 723-746.

[44] Stone, Jeff, Zachary W. Perry and John M. Darley, 1997, "'White Men Can't Jump': Evidence for the Perceptual Confirmation of Racial Stereotypes Following a Basketball Game," *Basic and Applied Social Psychology*, Vol. 19, No. 3, pp. 291-306.

[45] Stone, Jeff, Christian I. Lynch, Mike Sjomeling and John M. Darley, 1999, "Stereotype Threat Effects on Black and White Athletic Performance," *Journal of Personality and Social Psychology*, Vol. 77, No. 6, pp. 1213-1227.

[46] Wolfers, Justin, 2006a, "Diagnosing Discrimination: Stock Returns and CEO Gender," *Journal of the European Economic Association*, Vol. 4, No. 2-3, pp. 531-541.

[47] Wolfers, Justin, 2006b, "Point Shaving: Corruption in NCAA Basketball," *American Economic Review*, Vol. 96, No. 2, pp. 279-283.

[48] Yinger, John, 1998, "Evidence on Discrimination in Consumer Markets," *Journal of Economic Perspectives*, Vol. 12, No. 2, pp. 23-40.

Figure 1: Distribution of Forecast Errors

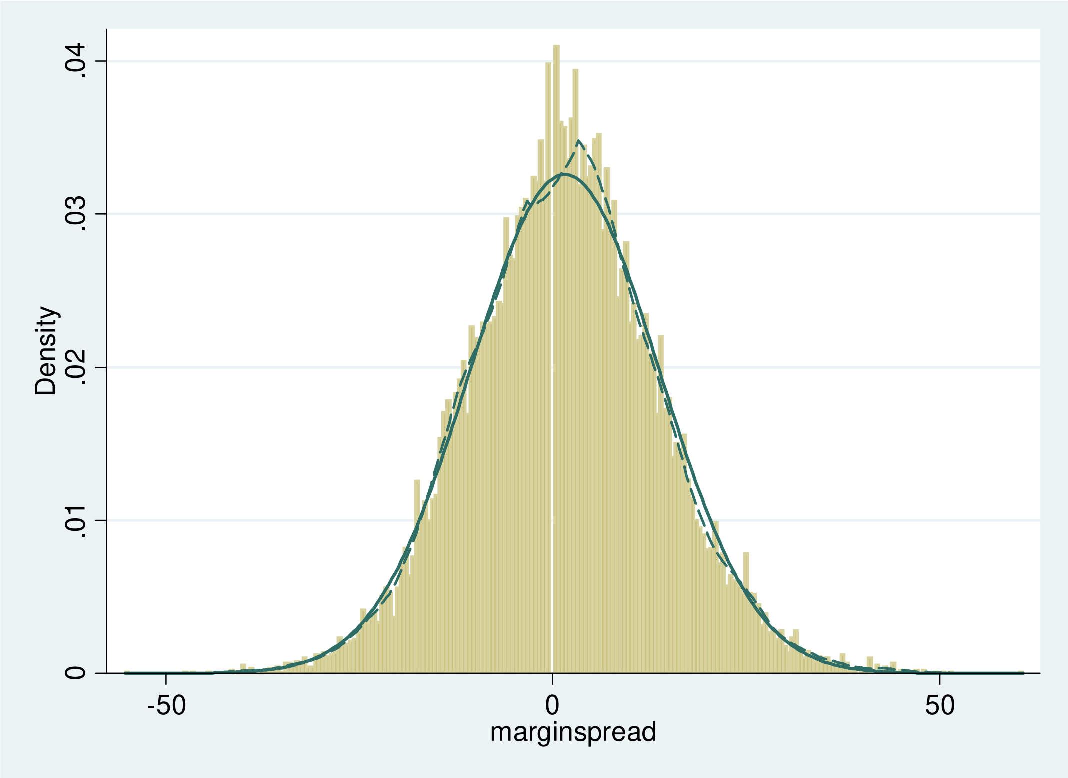

Figure 2: Difference between the Winning Margin and the Point Spread against Normal Distribution

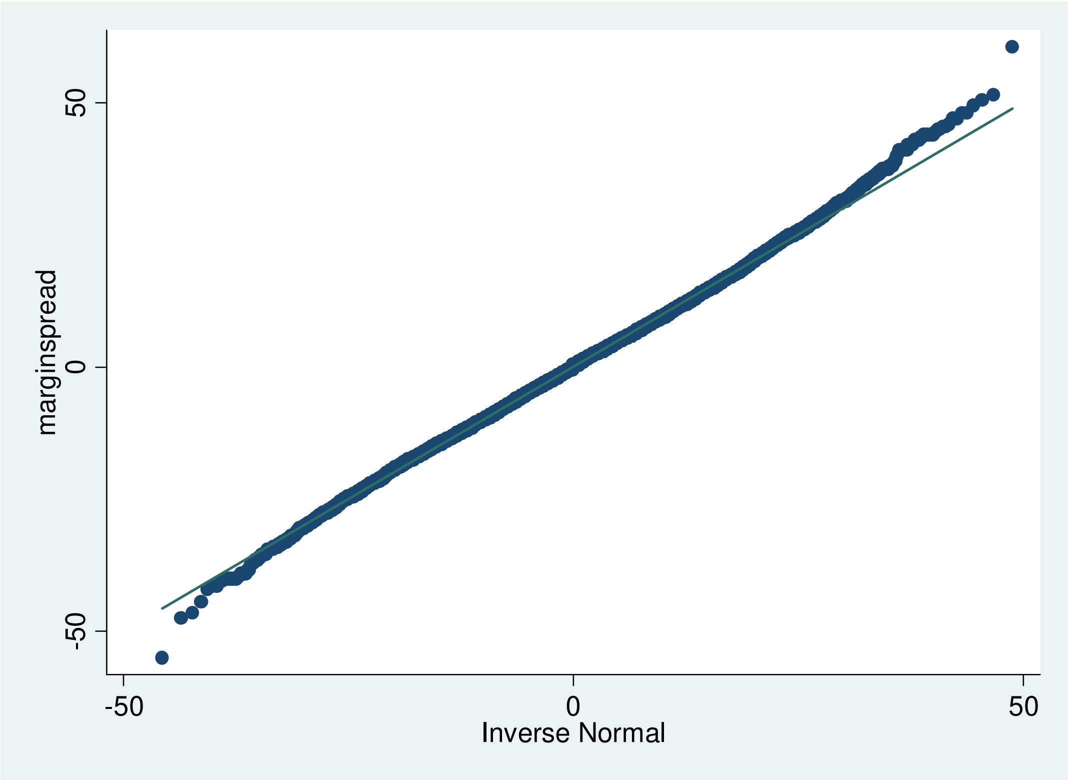

Figure 3: Accuracy of Point Spreads

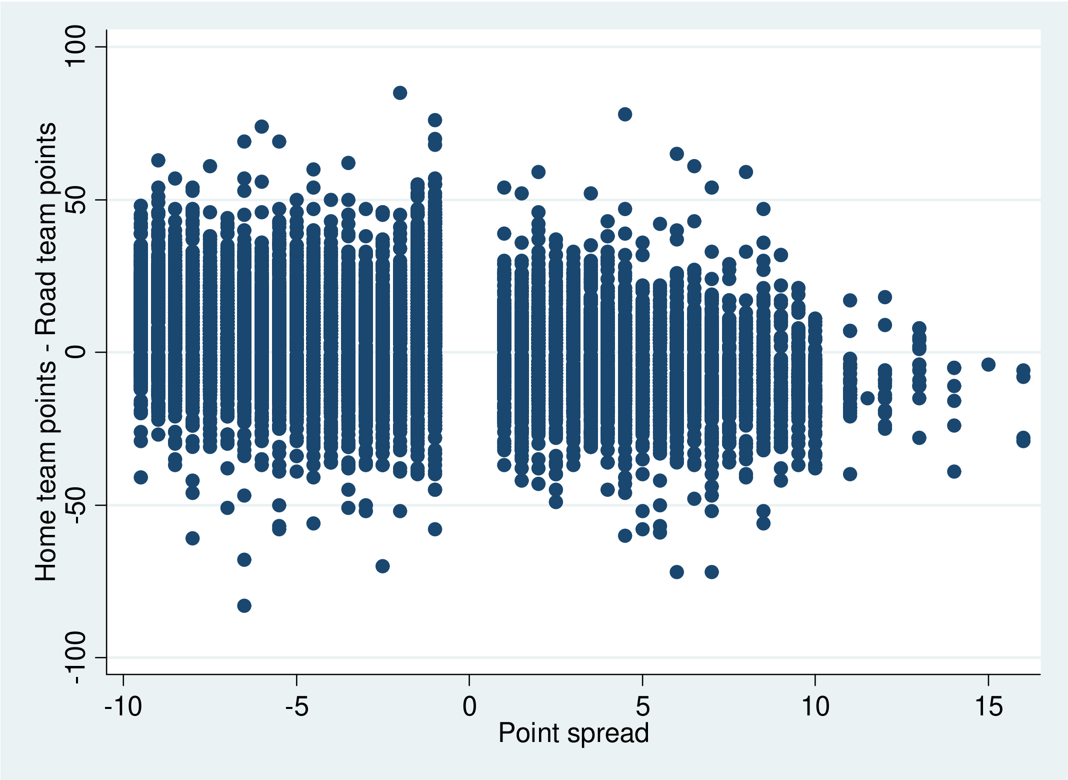

Figure 4: Distribution of the Difference between the Opening and Closing Spreads

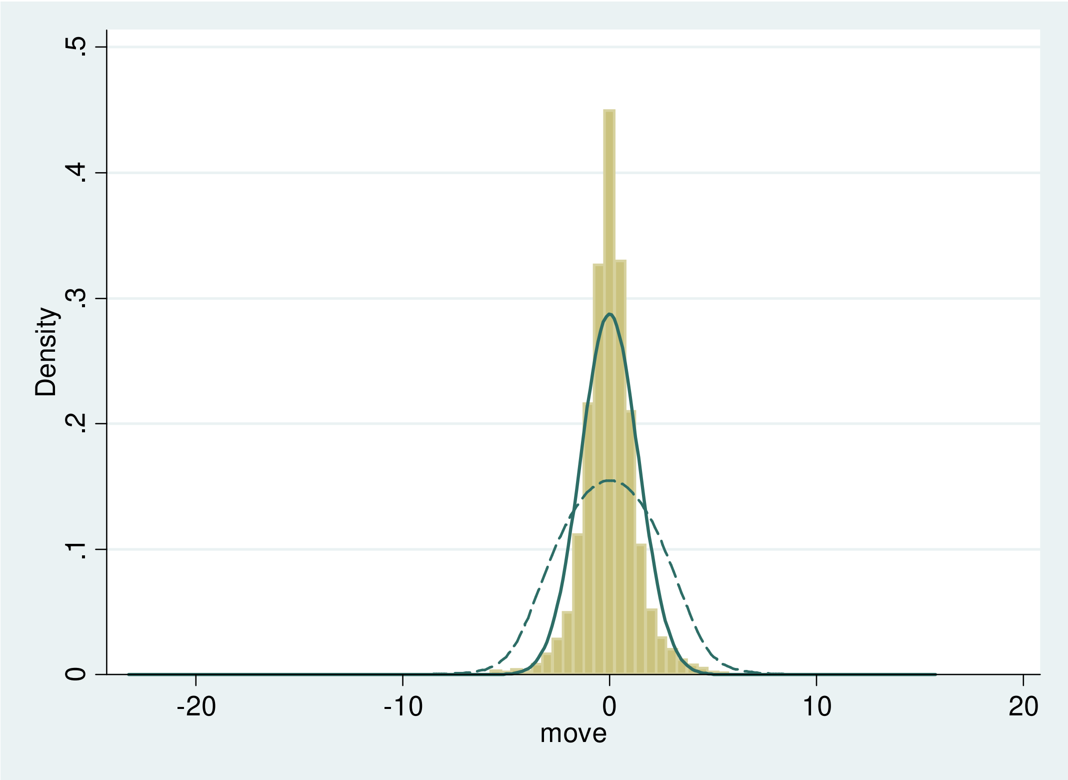

Figure 5: Distribution of Money Bet

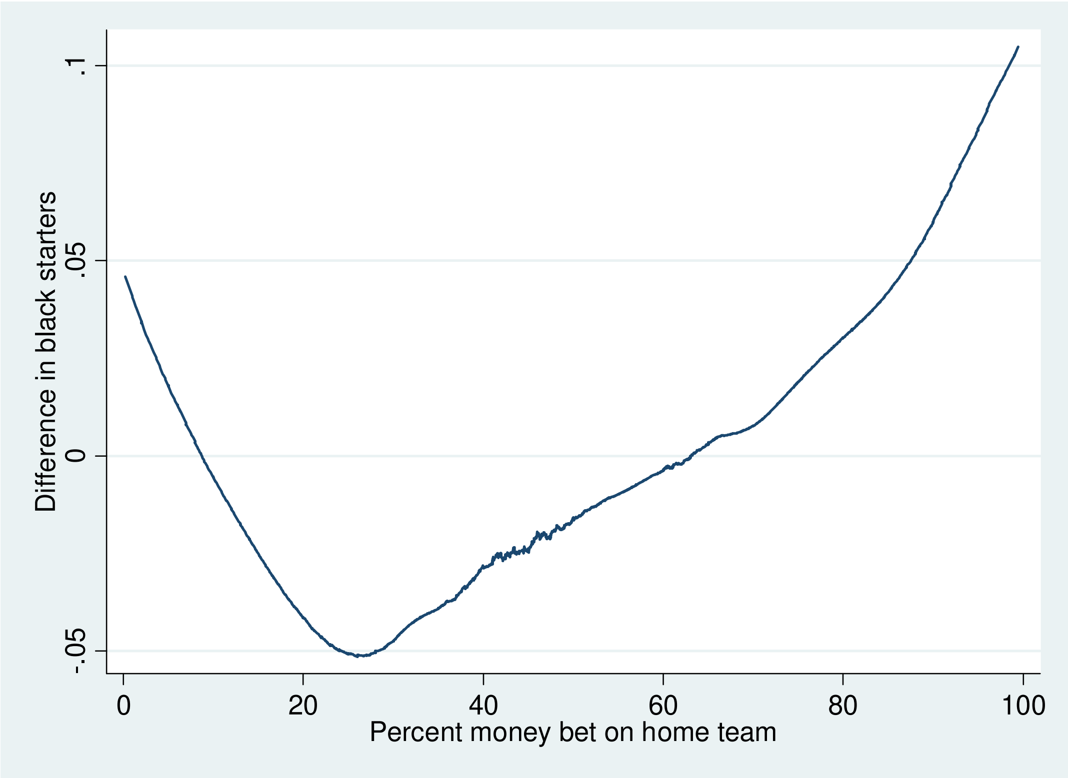

Table 1. Summary Statistics at Player-Game Level: Adjusted

	Black players			White players			Significance
Variable	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	of difference
Personal fouls	323178	4.99	5.07	98028	5.61	6.14	>0.99
Points	323176	18.04	12.07	98028	17.13	12.79	<0.01
Free throws attempted	323179	4.89	6.32	98028	4.38	6.69	<0.01
Free throws made	323179	3.55	4.85	98028	3.24	5.14	<0.01
Free throw percentage	206288	1.53	1.86	55807	1.76	2.33	>0.99
Field goals attempted	323179	15.61	8.07	98028	14.71	8.39	<0.01
Field goals made	323179	6.80	4.90	98028	6.44	5.16	<0.01
Field goal percentage	308710	1.15	2.27	91836	1.35	2.73	>0.99
Two point shots attempted	323179	12.94	7.62	98028	11.78	7.99	<0.01
Two point shots made	323179	5.92	4.70	98028	5.43	4.92	<0.01
Two point shot percentage	303829	1.18	2.25	89363	1.37	2.67	>0.99
Three point shots attempted	323179	2.67	4.04	98028	2.92	4.52	>0.99
Three point shots made	323179	0.88	1.89	98028	1.01	2.22	>0.99
Three point shot percentage	155412	0.64	1.39	46544	0.80	1.83	>0.99
Offensive rebounds	323179	2.48	3.55	98028	2.66	3.95	>0.99
Defensive rebounds	323179	5.66	4.93	98028	6.20	5.52	>0.99
Total rebounds	323179	8.15	6.52	98028	8.85	7.17	>0.99
Assists	323179	4.09	4.46	98028	3.92	4.77	<0.01
Steals	323179	1.57	2.38	98028	1.39	2.44	<0.01
Blocks	323179	0.98	2.14	98028	1.09	2.40	>0.99
Turnovers	323179	2.99	3.56	98028	2.90	3.87	<0.01
Win score	323176	6.74	11.70	98028	7.28	12.88	>0.99

Notes: The last column shows the p-values from t-tests with the null hypothesis that the statistic for black players is greater than the statistic for white players. Adjusted statistics are calculated by multiplying the raw statistics by 48 (the total number of minutes in a regular game, i.e., no overtime) and then dividing by the actual number of minutes played by that player in that game.

Table 2. Summary Statistics at Team-Game Level

	Obs	Mean	St. Dev.
Point spread	14785	-1.86	5.07
Realized margin	17178	3.36	14.15
Realized margin - spread	14785	0.35	11.53
Probability of beating the spread	14785	0.51	0.50
Black starters	17179	3.90	1.05
Difference in black starters	17178	0.01	1.41
Black players on the roster	17179	7.60	1.63
Difference in black players on the roster	17178	-0.01	2.12
Black minutes	17022	0.78	0.16
Difference in black minutes	16982	0.001	0.21

Notes: Point spread is the quoted spread on a game as of the closing time for bets, expressed from the home team's perspective. Realized margin is the actual difference between the home team score and the visiting team score at the end of the game. Probability of beating the spread is a dummy that is 1 if a bet on the home team wins. Black starters is the number of black players in the starting line-up. Black players on the roster is the number of black players on the team roster. Black minutes is the proportion of minutes played by black players to the total minutes in the game, calculated over the past five games the team has played. These measures of blackness of a team refer to the home team. Difference in black starters is calculated as the number of black players in the starting line-up (number of black players on the roster, proportion of black minutes) of the home team minus the number of black players in the starting line-up (number of black players on the roster, proportion of black minutes) of the visiting team.

Table 3. Winning the Game

	Realized margin		
Difference in black starters	0.290***		
	[0.105]		
Diff. in black players on the roster		0.091	
		[0.066]	
Difference in black minutes			-4.356***
			[0.675]
Team fixed effects	yes	yes	yes
Season fixed effects	yes	yes	yes
Team-season interactions	yes	yes	yes
Observations	17178	17178	16982
R-squared	0.14	0.14	0.14
	Probability of winning		
Difference in black starters	-0.023***		
	[0.009]		
Diff. in black players on the roster		-0.029***	
		[0.006]	
Difference in black minutes			-0.399***
			[0.065]
Team fixed effects	yes	yes	yes
Season fixed effects	yes	yes	yes
Team-season interactions	yes	yes	yes
Observations	17178	17178	16982

Notes: The dependent variable in the upper panel is the realized margin in the game, computed as the home team score minus the visiting team score. The dependent variable in the lower panel is the probability of winning, which is a dummy that is 1 if the home team won the game. The regressions are estimated using ordinary least squares for the winning margin, and using probit for the probability of winning. Difference in black starters is calculated as the number of black players in the starting line-up (number of black players on the roster, proportion of black minutes over the past five games) of the home team minus the number of black players in the starting line-up (number of black players on the roster, proportion of black minutes over the past five games) of the visiting team. Robust standard errors are in square brackets. ***, **, and * denote statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 4. Beating the Spread

	Point spread		
Difference in black starters	0.364***		
	[0.032]		
Diff. in black players on the roster		0.143***	
		[0.021]	
Difference in black minutes			3.161***
			[0.224]
Team fixed effects	yes	yes	yes
Season fixed effects	yes	yes	yes
Team-season interactions	yes	yes	yes
Observations	14784	14784	14631
R-squared	0.33	0.32	0.33
	Realized margin - spread		
Difference in black starters	-0.239***		
	[0.083]		
Diff. in black players on the roster		-0.275***	
		[0.053]	
Difference in black minutes			-1.427**
			[0.596]
Team fixed effects	yes	yes	yes
Season fixed effects	yes	yes	yes
Team-season interactions	yes	yes	yes
Observations	14784	14784	14631
R-squared	0.06	0.06	0.06
	Probability of beating the spread		
Difference in black starters	-0.021**		
	[0.009]		
Diff. in black players on the roster		-0.033***	
		[0.006]	
Difference in black minutes			-0.116*
			[0.066]
Team fixed effects	yes	yes	yes
Season fixed effects	yes	yes	yes
Team-season interactions	yes	yes	yes
Observations	14784	14784	14631

Notes: The dependent variable in the upper panel is the point spread quoted on the game, expressed from the home team's perspective. The dependent variable in the middle panel is the difference between the realized margin (the actual outcome of the game) and the point spread. The dependent variable in the lower panel is the probability of beating the spread, which is a dummy that is 1 if a bet on the home team wins. The regressions are estimated using ordinary least squares for the point spread and the difference between the realized margin and the spread, and using probit for the probability of beating the spread. Difference in black starters is calculated as the number of black players in the starting line-up (number of black players on the roster, proportion of black minutes over the past five games) of the home team minus the number of black players in the starting line-up (number of black players on the roster, proportion of black minutes over the past five games) of the visiting team. Robust standard errors are in square brackets. ***, **, and * denote statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 5a. Moving the Point Spread

	Closing spread - Opening spread			
Difference in black starters	0.014**	0.008	0.008	0.006
	[0.007]	[0.008]	[0.008]	[0.009]
Team fixed effects	no	yes	yes	yes
Season fixed effects	no	no	yes	yes
Team-season interactions	no	no	no	yes
Observations	7977	7977	7977	7977
R-squared	0.00	0.01	0.01	0.05

Notes: The dependent variable is the difference between the closing and opening values of the spread on the game, showing how much the point spread moves from the start of betting until all bets close. The regressions are estimated using ordinary least squares. Difference in black starters is calculated as the number of black starters on the home team minus the number of black starters on the visiting team. Robust standard errors are in square brackets. ***, **, and * denote statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 5b. Bias in Bets

	Money bet on home team			
Difference in black starters	0.185	0.243*	0.243*	0.285*
	[0.118]	[0.140]	[0.140]	[0.158]
Team fixed effects	no	yes	yes	yes
Season fixed effects	no	no	yes	yes
Team-season interactions	no	no	no	yes
Observations	8011	8011	8011	8011
R-squared	0.00	0.00	0.01	0.01

Notes: The dependent variable is the money placed as bets on the home team, expressed as a percentage of the total bets placed on the game. The regressions are estimated using ordinary least squares. Difference in black starters is calculated as the number of black starters on the home team minus the number of black starters on the visiting team. Robust standard errors are in square brackets. ***, **, and * denote statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 6a. Robustness: Referees

	Probability of beating the spread	
	At least one black referee	At least one white referee
Difference in black starters	-0.015*	-0.022***
	[0.011]	[0.009]
Team fixed effects	yes	yes
Season fixed effects	yes	yes
Team-season interactions	yes	yes
Observations	11244	14460
	All-black crew	All-white crew
Difference in black starters	-0.098	-0.044**
	[0.128]	[0.021]
Team fixed effects	yes	yes
Season fixed effects	yes	yes
Team-season interactions	yes	yes
Observations	126	3464
	Neither all-black nor all-white crew	Referee race as additional control
Difference in black starters	-0.017*	-0.019**
	[0.011]	[0.009]
Proportion of white referees		-0.150***
		[0.046]
Team fixed effects	yes	yes
Season fixed effects	yes	yes
Team-season interactions	yes	yes
Observations	10911	14694

Notes: The regressions are estimated using probit. Difference in black starters is calculated as the number of black starters on the home team minus the number of black starters on the visiting team. The race composition of referees are taken into account by splitting the sample by the proportion of black referees in the 3-person crew. Alternatively, the proportion of white referees is included as a control variable. Robust standard errors are in square brackets. ***, **, and * denote statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 6b. Robustness: Coaches

	Probability of beating the spread	
	Black (H), white (V)	White (H), black (V)
Difference in black starters	-0.015	0.044*
	[0.020]	[0.026]
Team fixed effects	yes	yes
Season fixed effects	yes	yes
Team-season interactions	yes	yes
Observations	3011	2797
	Both black or both white	Coach race as additional control
Difference in black starters	-0.033***	-0.017*
	[0.012]	[0.009]
Difference in coaches' race		-0.064***
		[0.024]
Team fixed effects	yes	yes
Season fixed effects	yes	yes
Team-season interactions	yes	yes
Observations	8859	14784

Notes: The regressions are estimated using probit. Difference in black starters is calculated as the number of black starters on the home team minus the number of black starters on the visiting team. The race of the coach is taken into account by splitting the sample by the races of both the home and visiting teams' coaches. Alternatively, the proportion of white referees is included as a control variable. In the first column, "Black (H), white (V)" indicates that only observations where the coach of the home team is black and the coach of the visiting team is white are included. In the second column, "White (H), black (V)" indicates that only observations where the coach of the home team is white and the coach of the visiting team is black are included. In the third column, either both coaches are black or both coaches are white. In the last column, the difference in coaches' race is calculated by first creating a dummy for the coach of each team (1 if the coach is black) and then subtracting the visiting team's dummy from the home team's. Robust standard errors are in square brackets. ***, **, and * denote statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 6c. Robustness: Population in Host Location

	Probability of beating the spread	
Difference in black starters	-0.020**	-0.019**
	[0.009]	[0.009]
Diff. in proportion of blacks in the city	-0.0001	
	[0.0001]	
Diff. in proportion of blacks in the state		-0.001
		[0.001]
Team fixed effects	yes	yes
Season fixed effects	yes	yes
Team-season interactions	yes	yes
Observations	14705	14705

Notes: The regressions are estimated using probit. Difference in black starters is calculated as the number of black starters on the home team minus the number of black starters on the visiting team. The difference in proportion of blacks in the city (state) is computed by subtracting the percent of black population, as of 2000, in the visiting team's host city (state) from the percent of black population in the home team's host city (state). Robust standard errors are in square brackets. ***, **, and * denote statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 6d. Robustness: History of Teams

	Probability of beating the spread				
Difference in black starters	-0.019**	-0.023**	-0.023**	-0.026***	-0.033***
	[0.009]	[0.009]	[0.010]	[0.010]	[0.012]
Difference in records	0.033***				
	[0.007]				
Difference in recent games		0.041			
		[0.033]			
Margin on the teams' last match			0.002**		
			[0.001]		
Spread on the teams' last match				0.0001	
				[0.003]	
Team fixed effects	yes	yes	yes	yes	yes
Season fixed effects	yes	yes	yes	yes	yes
Team-season interactions	yes	yes	yes	yes	yes
Match fixed effects	no	no	no	no	yes
Observations	14781	14631	14041	13437	13324

Notes: The regressions are estimated using probit. Difference in black starters is calculated as the number of black starters on the home team minus the number of black starters on the visiting team. In the first column, the difference in records is calculated as the difference between the number of wins the home team had in a particular season prior to the game in consideration and the number of corresponding wins for the visiting team. In the second column, the difference in recent games is calculated as the difference between the number of wins the home team had in a particular season over the five previous games before the game under consideration and the number of corresponding wins for the visiting team. In the third column, the margin on the teams' last match is computed as the difference between home team's score and the visiting team's score obtained the last time the two teams played against each other (irrespective of the location). In the fourth column, the spread on the teams' last match is the point spread quoted on the last game the two teams faced each other (irrespective of the location and expressed from the home team's perspective). Note that the margin/spread is equal to the margin/spread from the last match-up in the previous season when the game in consideration is the first time the two teams face each other in a given season. Robust standard errors are in square brackets. ***, **, and * denote statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 7. Chances of Winning with Simple Strategies

Strategy 1	Bet on the team that has X more black starters ...											
	1			2			3			4		
	Bets	Win %	Return %	Bets	Win %	Return %	Bets	Win %	Return %	Bets	Win %	Return %
Average	6518	49.3	-5.9	2670	47.2	-9.9	783	47.6	-9.0	254	40.3	-23.0
Average for all	10225	46.1	-12.0									
Strategy 2	**Bet on the team that has X more white starters ...**											
	1			2			3			4		
	Bets	Win %	Return %	Bets	Win %	Return %	Bets	Win %	Return %	Bets	Win %	Return %
Average	6518	50.7	-3.1	2670	52.8***	0.8***	783	52.4	0.0	254	59.7***	14.0***
Average for all	10225	53.9***	2.9***									
Strategy 3	**Bet on the home team only when it has X more white starters than the visiting ...**											
	1			2			3			4		
	Bets	Win %	Return %	Bets	Win %	Return %	Bets	Win %	Return %	Bets	Win %	Return %
Average	3224	51.5	-1.7	1331	55.4***	5.8***	399	58.3***	11.4***	118	65.3***	24.7***
Average for all	5072	57.6***	10.0***									

Notes: The table shows the outcome of bets placed on a team when it has X (taking on values of 1, 2, 3, or 4) more white/black starters than the opposing team, as defined by the strategy. Bets show the number of games that satisfy the condition in a given season and the strategy would require an \$11 bet being placed on the team. Win % is the proportion of bets that the betted-on team would beat the spread and the bettor would receive \$21. Return % denotes the return on the betting strategy, computed as the total money earned on the bets as a proportion of the money spent on placing the bets. Bets are expressed in units; win % and return % are in percent terms. The first row reports the total number of bets and the simple average for the wins and returns over all the seasons in the sample. The last row reports the total number of bets and the simple average for the wins and returns for the strategy considered as a whole. *** indicates that the win % (return %) is significantly higher than 52.4% (0.0%), i.e., the win percentage required to break even, at the 1% level.

Table 8. Chances of Winning with a Simple Strategy: Season by Season

	Bet on the home team only when it has X more white starters than the visiting ...											
	1			2			3			4		
	Bets	Win %	Return %	Bets	Win %	Return %	Bets	Win %	Return %	Bets	Win %	Return %
2007-08	243	56.4	7.6	141	57.4	9.7	41	58.5	11.8	8	75.0	43.2
2006-07	270	48.5	-7.4	110	53.6	2.4	40	52.5	0.2	10	70.0	33.6
2005-06	244	54.9	4.8	126	54.0	3.0	50	56.0	6.9	8	50.0	-4.5
2004-05	214	50.9	-2.8	123	57.7	10.2	38	52.6	0.5	20	35.0	-33.2
2003-04	209	51.2	-2.3	80	63.8	21.7	47	63.8	21.9	27	55.6	6.1
2002-03	217	49.3	-5.9	101	59.4	13.4	38	60.5	15.6	11	63.6	21.5
2001-02	192	54.2	3.4	99	59.6	13.8	28	53.6	2.3	11	63.6	21.5
2000-01	202	51.5	-1.7	122	54.1	3.3	29	51.7	-1.3	13	46.2	-11.9
1999-00	227	53.7	2.6	101	59.4	13.4	18	61.1	16.7	1	100.0	90.9
1998-99	141	53.2	1.5	57	43.9	-16.3	21	61.9	18.2	4	50.0	-4.5
1997-98	201	55.2	5.4	73	58.9	12.5	12	58.3	11.4	4	75.0	43.2
1996-97	227	45.4	-13.4	52	65.4	24.8	16	81.3	55.1			
1995-96	235	49.8	-5.0	42	45.2	-13.6	2	50.0	-4.5			
1994-95	199	49.7	-5.0	60	41.7	-20.5	7	71.4	36.4			
1993-94	203	48.8	-6.9	44	56.8	8.5	12	41.7	-20.5	1	100.0	90.9
Average	3224	51.5	-1.7	1331	55.4***	5.8***	399	58.3***	11.4***	118	65.3***	24.7***
Average for all	5072	57.6***	10.0***									

Notes: The table shows the outcome of bets placed on the home team only when the home team has X (taking on values of 1, 2, 3, or 4) more white starters than the visiting team. Bets show the number of games that satisfy the condition in a given season and the strategy would require an $11 bet being placed on the home team. Win % is the proportion of bets that the home team would beat the spread and the bettor would receive $21. Return % denotes the return on the betting strategy, computed as the total money earned on the bets as a proportion of the money spent on placing the bets. Bets are expressed in units; win % and return % are in percent terms. The row before last reports the total number of bets and the simple average for the wins and returns over all the seasons in the sample. The last row reports the total number of bets and the simple average for the wins and returns if one bets anytime the home team has more white starters than the visiting team. *** indicates that the win % (return %) is significantly higher than 52.4% (0.0%), i.e., the win percentage required to break even, at the 1% level.