

Racial Bias in Expert Quality Assessment: A Study of Newspaper Movie Reviews

Lona Fowdur, Vrinda Kadiyali and Jeffrey Prince *

October 2009

Abstract

Newspaper critics' movie reviews are often used by potential movie viewers as signals of expert quality assessment. In this paper, we assess if there is any racial bias in these critics' reviews, and if so, what impact these biases have on viewer demand. To do this, we develop a dataset that tracks ratings from 68 popular movie critics for 566 movies released in the U.S. between 2003 and 2007. The data also include measures of movie production costs, marketing expenditures, type of movie (i.e. genre, MPAA rating, etc.), actor and director quality measures, audience tastes and critics' gender, experience and race. Despite inclusion of all these controls for movie quality and other drivers of critic ratings, we find that ratings for movies with a black lead actor and all white supporting cast are approximately 6% lower than for other racial compositions. These results appear consistent with implicit discrimination. Using estimates of the impact of critics' ratings on movie revenues, we find that lower critic ratings for black lead-white support movies translate into lost revenues of up to 4% or about \$2.57 million on average. In sum, prejudice concerning race roles (e.g., the race of the leader versus supporters/followers) can have a direct impact on critic quality assessment, and thereby alter market outcomes.

Key words: racial bias, quality assessment, expert ratings, movies.

* This paper is based on Fowdur's doctoral dissertation. Lona Fowdur is at Economists Inc., Vrinda Kadiyali is a Professor of Marketing and Economics and Jeffrey Prince is an Assistant Professor of Economics, both at Cornell University. They can be reached at l_fowdur@yahoo.com, kadiyali@cornell.edu and jtp35@cornell.edu. We are especially grateful to Melissa Thomas-Hunt for directing us to relevant psychology literature in stereotyping and race. We also thank Daniel Benjamin, George Jakobson, Anirban Mukherjee, Vishal Narayan, Daniel Simon and Manoj Thomas for comments.

1. Introduction

Many products and services are experience goods, where consumers cannot fully observe quality before making a purchase. In such markets, consumers can utilize the insights of experts to assess quality before purchasing the product. Examples of experience goods markets with experts include automobiles, restaurants, financial stocks, books, and movies. Expert opinions can influence consumer demand in these experience goods markets, so it is important to understand whether there exists any inherent bias in their assessments. In this paper, we test for a particular form of bias in expert opinion - racial bias - of the quality of movies when they are released in theaters. The U.S. motion picture industry is large (revenues of nearly \$10 billion in the United States in 2008), and one where critics' reviews constitute expert opinion, and are widely disseminated and read.

There is evidence that racial discrimination exists in various economically pertinent situations. Examples include differences in wages (Chandra 2000), prices paid in consumer markets (Ayres and Siegelman 1995; Yinger 1998), credit availability (Blanchflower et al. 2003), judicial sentencing (Abrams et al. 2006) and even in sports (Price and Wolfers 2007). To our knowledge, this is the first study to test for racial bias in expert quality assessment, where consumers make their choices subsequent to, and conditional on, this assessment.

That said, there are at least two reasons why critical movie reviewing is an unlikely context for racial discrimination. First, movie critics are employed by the news industry with a documented liberal bent (Gentzkow and Shapiro 2006). Second, critics' ratings are read and assessed by a very broad audience, including other critics and movie studios, who can potentially detect bias. However, certain types of discrimination may be more subtle than simply penalizing a movie for having a black¹ lead or all-black cast, and so can be harder to detect anecdotally. Further, discrimination may not even occur at a conscious level (e.g. Bertrand et al. 2005). Consequently, racial bias can be hard to detect not just for readers of the newspaper (who might not follow reviews for all movies), but even for studios or watchdog groups who might be more vigilant in their reading of critic reviews².

¹ We use the term black instead of African-American throughout simply because not all black actors in our sample are American. Also, for simplicity, we refer to both actors and actresses as actors.

² Studios devote considerable resources to the management of the review process. Therefore one would expect that studios would quickly detect overt racism in the review process. Examples of studio management of the review

In determining whether critic reviews are racially biased, an important factor to consider is whether they are influenced by racism in casting patterns in Hollywood. The U.S. movie industry has a history suggestive of racial discrimination in casting. For several decades the majority of blacks cast in movies were seen in subservient or stereotypical roles. Only in the last decade have black actors made some progress in the variety of roles available, and in recognition of their performance via awards. For example, in its eighty years of existence, the Academy Awards have only awarded 11 Oscars to blacks, six of which were awarded between 2002 and 2009 (Khatami 2009). Robinson (2006) documents that Hollywood studios often specify a preferred race for particular roles and overwhelmingly favor white male actors for leading roles, leaving only a small proportion of roles open to blacks, and even fewer roles that are “cross-over” roles, or roles that can be played by an actor of any race. Racial stereotypes often apply to Latino actors as well; Latinos are almost always portrayed as expendable characters or as “foils for largely white characters” who define the movie (Powalski 2009). It is useful to see whether historical stereotypes affect (viewer and) critic rankings of movies where these stereotypes are violated or reinforced.³

To conduct our analysis of whether newspaper critics exhibit racial bias, we obtain newspaper critics’ reviews for 566 movies released between 2003 and 2007. We focus on 68 critics from a set of critics referred to as “cream-of-the-crop” in movie industry parlance (metacritic.com; rottentomatoes.com), because they are employed by some of the most illustrious newspapers in the U.S. Our main explanatory variable is the racial composition of the cast. In addition, to control for whether critics are tailoring their reviews to their audience preferences, we select those who publish in newspapers for which we can observe readership characteristics. We also collect an extensive set of controls, including the type of movie (genre, MPAA rating), production cost, advertising expenditure, quality of the actors, the race of the director, critical acclaim awards and nominations for the actors, director, and movie, demographics of the newspaper audience and time controls. These controls are important in our analysis since some

process include studios excerpting positive reviews in movie advertising, and sometimes delaying or foregoing advance screenings if they anticipate bad reviews. An extreme example was when Sony Corporation invented fictitious film critic David Manning, who consistently gave good reviews to Sony movies. Further, studios sometimes invite and host critics at lavish movie premieres, presumably in an attempt to coerce good reviews.

³ Dovidio et al. (1997) show that when confronted with evidence that does not fit racial stereotypes (e.g. observing a black Secretary of state or a black President or a black CEO of a Fortune 500 company), whites construct subtypes (e.g. black professionals) to fit the new evidence rather than alter their basic stereotyping schema.

of the movie characteristics that comprise quality may also be correlated with the racial composition of the cast. Additionally, we include controls for reviewer experience, gender, and race. Finally, we obtain data on movie box office revenues, allowing us to measure the impact of critics' reviews on them. Using our estimates, we can attempt to determine the cost of critics' bias to movies' theatrical revenues due to lower ratings.

We find that movies with black leads and all-white supporting casts (Bw henceforth) receive lower critic ratings – more than 6% lower – compared to movies with other racial compositions. This finding suggests a bias by critics not against black actors or even black lead actors per se, but instead a bias against movies with black leads and white support. Notably, these movies are the ones that break the trend of blacks playing subservient or subordinate (to whites) characters. In contrast, we find director (whose role is not as visible as an actor's) race has no impact on critics' ratings.

Using estimates from our movie box office revenues regressions, we establish an upper bound on the impact that critic reviews can have on viewer choices (through total revenues) for the movies in our sample. We find that a movie with a black lead and all-white support suffers up to a 4% loss in box office revenues due to lower ratings, compared to a movie with an all-white cast. The average revenue for movies in our sample is \$64.2 million, and therefore the 4% loss in box office is on average about \$2.57 million. The measured loss is similar when comparing to all-black casts (henceforth Bb) and casts with white leads and at least one black support (henceforth Wb). Since theatrical revenues often correlate with sales in subsequent distribution channels like DVD, pay-per-view etc., the total revenue loss across these various outlets might be even larger.

This paper primarily contributes to two streams of literature. First, it adds to the literature on impartiality of expert judgments. Extant literature highlights the role of critical reviews in determining market outcomes for several experience goods: financial stocks (Goh & Ederington 1993), sports betting (Avery & Chevalier 1999), movies (Eliashberg & Shugan 1997), etc. In the movie industry, newspaper critics are perceived to be unaffiliated with studio producers and therefore perceived as impartial. Research shows that more than one third of consumers indicate that they actively seek out critics' reviews, and about one third of filmgoers attest to choosing movies on the basis of positive reviews (Eliashberg & Shugan 1997). Given

the importance of positive critical reviews for movies, our finding of a systematic bias in these reviews is of particular relevance.

Second, our paper contributes to the literature on discrimination. Following this literature (e.g., Arrow 1998 and Bertrand et al. 2005), there are three types of discrimination: statistical discrimination, taste-based discrimination, and implicit discrimination. Statistical discrimination is based on the evaluator's inability to perfectly observe quality of the candidate, and therefore the evaluator resorts to statistical differences in ability of races as a measure of any one candidate's ability. Taste-based discrimination relies directly on racial prejudice of the evaluator. Both statistical and taste-based discrimination are based on conscious thought processes of the evaluator. In implicit discrimination, agents non-consciously discriminate on the basis of existing stereotypes. As we discuss further below, our results are consistent with implicit discrimination on the part of newspaper critics, and we provide two psychological mechanisms for non-conscious or implicit discrimination.

The rest of this paper is organized as follows: In section 2, we discuss various theories behind critics' influence and mechanisms for racial discrimination. We discuss data in section 3. In section 4, we present our empirical results and provide discussion of their implications. We offer conclusions in section 5.

2. Determinants and Impact of Critical Reviews

In this section, we discuss how critics' reviews influence demand. We then discuss key factors that can influence critical reviews, including the mechanisms through which racial discrimination may play a role.

There are two broad routes through which critic reviews might correlate with movie demand: the influencer role and the predictor role. As influencers, critics act as opinion leaders who steer consumers' movie selection decisions, especially in the early weeks of a movie's release when alternative sources of movie information (e.g. word-of-mouth) are scant. As predictors, critics' ratings capture movie characteristics that appeal to their audience rather than influence audience preference. Reinstein and Snyder (2005) summarize the difference between the two roles by defining the influence effect as the causal effect of reviews on demand, and the prediction effect as the spurious correlation between reviews and demand, induced by their

mutual correlation with quality. After differencing out the spurious correlation between reviews and demand, they find a significant influence effect for the reviewers they study. Any critical bias in reviews is problematic since it distorts the quality signal in this influence role.

In principle, the only factor that should determine critical reviews is movie quality, e.g., a well-written story, compelling acting, or state-of-the-art special effects. However, other factors may play a role in determining movies' scores. Mullainathan and Schleifer (2005) offer theoretical evidence showing that it is optimal for media outlets to slant information in the direction of their audience's political biases. Gentzkow and Shapiro (2006) empirically confirm this theory by estimating the optimal political slant of different media outlets as a function of the political biases of their respective audiences. The notion of optimal slant can logically be extended to critical movie reviews. That is, it may be optimal for profit-maximizing firms to provide incentives to employ critics who reflect the tastes of their readers. For this reason, readers' tastes may affect critical reviews.

Inherent critic bias can also play a role in determining critical reviews. Our focus is on the possibility that racial discrimination is a source of bias. Below, we discuss two possible types of discrimination: explicit discrimination, which involves a conscious decision to discriminate, and implicit discrimination, where the decision to discriminate is at the non-conscious level.

Arrow (1998) highlights the differences between two explanations for explicit discrimination: taste-based discrimination and statistical discrimination. Taste-based discrimination can be attributed to a disutility that an evaluator attaches to contact with someone from another group. Statistical discrimination, on the other hand, arises when, on average, differences exist between race groups in human capital, productivity, risk assessment, etc., but the evaluator cannot assess these traits for specific individuals. In such cases, evaluators use average characteristics of a race group as a surrogate for unobservable characteristics of an individual. Both taste-based and statistical discrimination generate sustained differences in market outcomes for individuals from different race groups.

The high visibility of movie reviews combined with the general liberal bent of newspapers suggests that explicit discrimination is unlikely⁴. Therefore, we dedicate the remainder of this section to a discussion of implicit discrimination. Bertrand et al. (2005) propose that an alternative explanation for racial discrimination is implicit discrimination, which is unintentional and non-conscious. Here, racial discrimination arises because of natural stereotypes held by people about particular race groups. To show this, the authors measure the response time differential in computerized testing where subjects are asked to pair words that connote “African American” with words that connote “good” or “bad.”⁵

Research in psychology on mechanisms of non-conscious discrimination includes three that are particularly relevant for our setting: expectancy violation, shifting-standards theory and status leakage. Expectancy violation theory (Burgoon 1978, Burgoon and Burgoon 2001) sees communication, and more broadly human interaction, as an exchange of information. In an evaluation context, when the evaluator receives information about the appraised target that violates her expectation, or when the evaluator is faced with a target’s unexpected behavior, the evaluator initiates cognitive appraisals of this violation of her expectation. If the target’s behavior is more positive than expected, the evaluation is more favorable than those for behavior that matches expectation. Negative violations result in evaluations less favorable than when behavior matches expectation. In our context this theory would predict that movies with cast racial compositions generating a negative (positive) expectancy violation may receive inferior (superior) ratings as compared to those without such a violation, *ceteris paribus*. Historical casting patterns in Hollywood have been predominantly Ww or Wb in some cases. Therefore, Bb and Bw movies are most likely to violate expectations of the evaluator. We further explore in section 4.3 the relative ranking of Bb versus Bw movies.

Shifting-standards theory is described in Foschi (1996) and Biernat and Kobrynowicz (1997). They show that lower performance is expected of low-status groups. Therefore, there is a tendency to downplay even good performance since the *ex-ante* expectations are low. For

⁴ For a contrarian view on possible racism despite being liberal, see Gaertner & Dovidio (1986)’s theory of aversive racism. They discuss how even whites with liberal values have trouble reconciling these values with their inherent racism. For example, in a field experiment, they found that (white) liberals were less likely to help a black (rather than a white), stranger in trouble; (white) conservatives were equally unlikely to help whites and blacks.

⁵ This study also documents an intriguing feature of implicit attitudes that can be manipulated: exposure to photographs of admired African Americans like Bill Cosby led to a decrease in implicit discrimination against African Americans.

example, in our movie context, performance might be evaluated as “He acted well, considering he’s black”. However, in absolute rankings, this individual (of a low-status group) might still be ranked lower than someone from a higher status group with the same ability. That is, a low-status individual is still held to a higher absolute standard than a higher-status individual. In our context, this theory would generally predict lower ratings for casts consisting of blacks (i.e. Bb); however, it says little about other aspects of the cast’s racial composition, in particular the race of the lead vs. race of supporting actors.

The final mechanism we consider for implicit discrimination is status leakage (Podolny, 2005). This theory was originally developed for groups of entities e.g. firms in an industry, workers in an organization, members of a team, etc. It is especially appropriate for us, given the movie quality is being judged by the performance of the entire team of actors (and the entire production team’s efforts), and there might be some impact of the lead actor’s status on the perceived performance of the support cast. According to this theory, the status of the leader of a team can leak over and affect the status of the other members of the team. For example, the lower status of the leader of a team can cause lower rankings for the rest of the team as well, even if the rest of the team is objectively of higher quality or higher status. In our application, despite equally good performance by lead actors and supporting cast, white support cast in a Ww team will be ranked higher than white support cast in a Bw team, because the lower status of the leader in Bw leaks over to the white supporting cast. However, at first blush, it is not obvious why a Bb would be judged better than Bw; more on that in section 4.3 below.

Summarizing, each of the three mechanisms of non-conscious or implicit discrimination discussed above might be useful in interpreting our results.

3. Data

In this section we discuss the collection and content of our data on the critics and their audience, and then data on movies.

3.1. Data on Critics and Their Newspaper Audience

We begin this subsection by discussing our data on critics’ movie ratings. Individual critics’ reviews for our sample time period are available from www.Metacritic.com (Metacritics). We only consider reviews for widely released movies (movies released in more than 600 theaters

nationally), because they have measures of actor and director quality that will prove to be useful for controlling movie quality (discussed further below). Metacritics compiles reviews from a group of 232 critics referred to as “cream-of-the-crop” given their clout in the industry. Since different critics use different rating metrics (e.g. 4 star, 5-star, letter grade, etc), Metacritic converts each critic’s review into a score on a zero-to-100 scale. Alternative compilations of individual critic reviews are available from sources like [www.rottentomatoes](http://www.rottentomatoes.com) (Rotten Tomatoes) www.yahoo.com/movies (Yahoo! Movies). We choose to use Metacritic ratings because the site offers more comprehensive coverage of critics’ reviews (all reviews available online are compiled as opposed to just a sub-sample of available reviews, as is the case with Yahoo! Movies). Also, Metacritic uses a more granular scale (Rotten Tomatoes uses a thumbs-up/thumbs-down appraisal format and Yahoo! Movies uses a letter grading system).

We limit our analysis to the set of 68 critics who are employed by one of the 11 daily newspapers for which we observe audience characteristics. For each of these critics, we observe experience levels (proxied by a count of their individual reviews compiled on MetaCritic), race, and gender (determined using pooled information from MetaCritic, Rotten Tomatoes and the popular press). For five critics in our sample we observe a job transfer to a different newspaper market, and for each of these critics we attribute reviews to the market for which they wrote, given the date of the job transfer and the date at which the review would have been published⁶. We provide summary statistics for critics and critic reviews in Table 1.

[Table 1 about here]

A potential issue with critic ratings is whether there is random assignment of critics to the movies they review. Since all newspapers in our sample employ a team of movie critics, it may be a concern that critics who are most racially biased self-select into reviewing movies against which they might be prejudiced. While we were unable to obtain information on how movies are assigned to critics within newspapers, we do not believe any sort of (self- or other) selection of critics to movies to be reviewed would distort our results. If any selection is occurring, it would

⁶ Since newspaper critics’ reviews are meant to serve as a quality signal prior to consumers having seen the movie, it is reasonable to assume that reviews will be published shortly before a movie’s release, or shortly after release at the latest. In any case, only a few critics transfer from one paper to another.

not disprove any findings of racial bias; it would only imply an overstatement of its effect if we attempted to make predictions in a world where reviews were randomly assigned. However, we believe the proper context to evaluate our results is under the current system of review assignment (selected or not) since the counterfactual of truly random assignment is not likely to occur.

As mentioned in Section 2, critics might attempt to tailor their ratings towards the preferences of their audience. To account for this possibility, we gather information on audience demographics. Our demographic profile of newspaper audiences comes from Reader Profile reports, commissioned by the Audit Bureau of Circulations (ABC). ABC audits newspaper circulation and maintains a database of average reader demographics for most major US newspapers⁷. These reports are compiled from annual phone surveys of consumers in each newspaper's market. From these reports, we get information for each newspaper's readership's racial composition (percentage of blacks and whites), average age, education (proportion of high-school and college educated), income (proportion of households earning under \$35,000 and above \$100,000), and geographic location (East, South, West, Midwest).

While it is conceivable that not all readers who attest to reading a newspaper necessarily read the movie reviews section, the characteristics of the overall audience is likely to be correlated with the characteristics of the readership segment that does read movie reviews. Further, critics may not have precise information (beyond average reader characteristics), about which segments of newspaper readers read their reviews. Hence we consider Reader Profiles to be a reasonable measure for the mean demographic characteristics of the average reader of movie reviews as perceived by the movie critic. Meta Critic reviews for 'cream-of-the-crop' critics are available for 11 newspapers for which we have Reader Profiles. Table 2 summarizes the demographic characteristics of each of these newspapers from 2003 to 2007.

[Table 2 about here]

⁷ Previous literature (e.g., George & Waldfogel 2003) use MSA (Metropolitan Statistical Area) demographics as a proxy for characteristics of the average reader, but we rely on ABC reader profiles instead since the assumption that all demographic segments have an equal predisposition to read newspapers may be flawed.

3.2. Data on Movie Characteristics

Our data on movie characteristics come from several sources. Yahoo! Movies has a synopsis page for each movie that also lists the names of the five main actors (i.e., the actors who played the most consequential parts of the movie) in the order of their role's importance for the movie plot. Specifically, the name of the lead actor is listed first followed by the names of supporting actors, from the most significant role played to the least. While this order is sometimes lead actor, romantic interest, side-kick, villain etc., it can also be lead actor, co-lead, side-kick etc. For example, the movie *My Baby's Daddy* (2004) is about two black and one white friends growing up together, and all three have pregnant girlfriends. In this movie, the first two characters listed in Yahoo! are the two black friends, with their white friend listed at number three. Therefore, the order reflects more the significance of roles rather than the type of role (main character, romantic interest, etc.)

We rely on cast information from Yahoo! Movies because it uses a harmonized system to define top-five roles across movies. Other movie sites, like imdb.com, only list actors in the two main roles and the rest of the cast (including all cameo appearances) in alphabetical order. This does not allow us to easily infer the identities and corresponding race groups of the main cast members.

For our sample, we track the top-five cast members (as listed in Yahoo! Movies) for 566 wide-release movies screened in theaters from 2003 to 2007. For each of these movies, we investigate which of the top-five cast members is black; we follow a similar procedure to determine the race of the director. To do this, we start with a list of black actors from Wikipedia.org. While Wiki may not always be reliable as a general source of information, we believe it to be for our purposes. This is because Hollywood stars tend to be popular, and information about them is readily available and can be readily verified and altered if incorrect. Also, given the open access format of Wikipedia, it is reasonable to assume that any wrong information about figures as popular as Hollywood stars would be quickly corrected or disputed by fans, actor's managers, etc. Then, in the next step, we searched for on-line photographs to verify the race of actors; the information came from various sources including actors' own webpages, fan pages, African-American magazines, etc. For the purposes of this study, we

classify all other race-groups as white since the number of actors in other minority groups (Asian, Hispanic, Arab, etc.) is much lower.⁸ As such, if the industry discriminates against other non-white race groups as well, our results will likely underestimate the total extent of discrimination.

Given our race definitions, we observe 64 (of 566) movies which cast blacks in lead roles (either Bb or Bw). Further, 182 movies had at least one black cast member (Bb, Bw, or Wb), and 384 movies had an all-white (Ww) top-5 cast. Table 3a shows the racial breakdown of each of the top five cast members given their respective positions in the cast.

[Table 3a about here]

To control for quality differences in actors and directors, we obtain a count of BFCA (Broadcast Film Critics Association) awards and nominations that each actor and director had received for all movies prior to focal movie⁹. We supplemented these measures by obtaining actors' "bankability" scores, which are designed to measure the commercial viability of each movie star. These measures were computed by The Hollywood Reporter in 2002 and are therefore invariant to the commercial success of movies in our sample period. Consequently, they may not capture the evolution of star-bankability from 2002 to 2007 (the end of our sample period). Therefore, these measures are proxies for the actual bankability of the actors at the time they were observed¹⁰.

Table 3b shows the average bankability scores of cast members for all movies in our sample by race and by position in the cast. We observe that bankability scores for lead actors are on average higher for all movies and trend down between the lead and the fourth supporting role. Average bankability of black actors in the lead position is on average higher than that of their white counterparts. In supporting positions, black actors have higher average bankability for the two most important supporting roles.

⁸ Jewish actors collectively constitute an exception; we leave an evaluation of critics' treatment of this and other categories (like gender) for further study.

⁹ We considered other awards counts as well (e.g., Oscars). However, for reasons discussed below, we use movie BFCA awards and nominations as another control, so for consistency, we use these measures here as well.

¹⁰ We believe these are proper proxies since they are almost certainly correlated with bankability of the actors at the time of the observation (e.g., if an actor had a high bankability in 2002, it appears likely it was high in 2005 also), and they should be uncorrelated with critic reviews conditional on bankability at the time of the observation.

[Table 3b about here]

We include several other measures of movie quality. First, we include production budget. We also include advertising expenditure (called “print and advertising” in industry parlance); this might capture expected quality of the movie and might also influence critics’ ratings if a movie is expected to be a commercial success. These data are obtained from Paul Kagan and Associates. We also include MPAA rating (i.e. whether a movie is rated G, PG-13, R, etc.), as well as movie genre and production studio dummies, to control for any systematic differences in quality along these dimensions. For example, it is possible that G-rated movies have lower artistic quality given their primary audience is children, or that critics systematically prefer the Drama genre to the Horror genre. Similarly, some studios might have higher quality than others. Our genre controls include: Animation, Action and Adventure, Comedy, Drama, Horror and Thriller, Romance and Science Fiction. These data are from www.boxofficemojo.com (BoxOfficeMojo).

The above controls are important, especially if they correlate with the racial composition of the casts. For example, lower budgets could be a consequence of black leads being paid less on average (which should imply a higher quality of acting, conditional on any level of budget). Similarly, lower advertising expenditures could be a consequence of more targeted marketing of movies with black leads to the black community (e.g. in *Ebony* magazine). It is possible that if movies with black cast members are not as well-promoted, critics do not form a positive opinion of such movies as a result.

Our final measure of quality is the movie’s artistic appeal. Holbrook (2005) argues that movies are either commercial successes or artistically high quality, with viewers preferring the former (and hence they are commercial successes), and critics preferring the latter. To capture the artistic quality of a movie, we use a summary statistic comprised of total BFCA nominations and wins (with nominations weighted by 0.5) for a movie across all categories (movie overall, acting, directing, writing, music, etc.). We also considered alternative measures of artistic appeal including Academy Awards (Oscars) and Golden Globe Awards. The Academy awards are awarded by a community of experts and non-experts. The Golden Globes are awarded by a community of foreign critics who may not appraise movie quality in the same way critics in our

sample do, especially because they are writing for a different audience. In contrast, the BFCAs are voted upon by a large community of U.S. based critics who are members of the Broadcast Film Critics Association, and who cast independent and anonymous ballots for the movies of their choice. In voting for the BFCAs, critics are not constrained by the profit-maximizing objectives of the newspapers which publish their reviews. Therefore, these awards are likely to be reflective of quality¹¹. Of the 566 movies in our sample, 97 won at least one BFCA nomination. Of these 97 movies, 70 featured an all-white top-five cast and 7 starred a black actor in the lead position (these latter movies were nominated but did not win).

In Table 4 below, we present summary statistics for all our movie-level measures.

[Table 4 about here]

Finally, to measure the impact of critic ratings movie revenues, we use box office revenue data collected from two internet sites: www.Boxofficejo.com (BM) and www.imdb.com (IMDB), for years 2003-2007. A quick glance at this variable shows that the average total box office revenue for the movies in our sample was \$64.2 million, with significant variance despite all these movies being wide-release only.

4. Empirical Results

4.1. Critic Reviews

In this subsection, we present the results of our regressions of critic ratings on cast racial composition, which include various critic-, movie-, and newspaper-level (including newspaper fixed effects) controls discussed in section 3. Additionally, it is possible that average review scores may change over time, with seasons, or due to extraneous events. Therefore, we also include fixed effects for every month in our sample.

We specify the following equation for critics' ratings:

¹¹ We note here that, it is possible that our award measure is negatively correlated with cast racial composition. In fact, we do find a negative correlation in our data. However, this does not induce any bias in our results (since both are used as covariates) or pose a collinearity problem (since the correlation is not large, -0.07).

$$(1) \quad R_{ijkt} = \alpha_0 + \alpha_1 \text{BlkLead}_j + \alpha_2 \text{MovieChar}_j + \alpha_3 \text{NPDemog}_{kt} + \alpha_4 \text{CritChar}_i + \alpha_5 \text{Month}_t + \varepsilon_{ijkt}$$

Here, R_{ijkt} is the individual review of critic i for movie j in newspaper market k in time period t . BlkLead_j is a dummy variable which equals one if the lead actor is black. MovieChar_j is a vector of movie j 's characteristics, including genre, MPAA rating, quality measures for actors and directors, director race, and movie awards (nominations and wins). NPDemog_{kt} is a vector of demographics for newspaper market k in time period t , including newspaper fixed effects. CritChar_i is a vector of critic characteristics for critic i including race, gender and experience. Month_t represents fixed effects corresponding to the timing of the critic's review. Finally ε_{ijkt} is an idiosyncratic error term.

Our regression results are in Table 5 below, which contains the primary findings of this paper. In column 1, we regressed critic ratings only on BlkLead_j . Here we see that movies with black leads receive scores that are nearly four points lower than those without black leads; this represents about a 6% difference given the average rating is approximately 58 (on the 0-100 scale used by Metacritic). The race of the director does not influence critic ratings. Taken by itself, this result is suggestive of a racial bias against black lead actors. However, as noted in section 3, there are many movie characteristics that could be correlated both with critic ratings and cast racial composition. Column 2 presents the results when we regress ratings on the dummy for black lead along with our full set of controls following equation (1) above. Here we see the effect is notably smaller (less than half the size), and insignificant.

[Table 5 about here]

The results in columns 1 and 2 of Table 5 suggest that critics may exhibit a small bias against black leads (once we include the controls¹²). However, as mentioned in the Introduction, it seems unlikely that such an overt practice of discrimination could be occurring among movie critics. In column 3, we allow for a finer partitioning of movies along racial lines. Specifically,

¹² These results are robust to inclusion of interaction terms between % of newspaper audience that is black and the dummy for the lead being black.

we consider four groupings: Bw (black lead with all-white support), Bb (black lead with at least one black support), Ww (white lead with all-white support), and Wb (white lead with at least one black support). This grouping allows us to determine whether lower scores for movies with black leads depend upon the racial composition of the supporting cast. This new specification is as follows:

$$(2) \quad R_{ijkt} = \alpha_0 + \alpha_1 \text{BlkLead}_j + \alpha_2 \text{BlkSupp}_j + \alpha_3 \text{BlkLead} * \text{BlkSupp}_j + \alpha_4 \text{MovieChar}_j + \alpha_5 \text{NPDemog}_{kt} + \alpha_6 \text{CritChar}_i + \alpha_7 \text{Month}_t + \varepsilon_{ijkt}$$

Here, BlkSupp_j is a dummy equal to one if one of the supporting (1-4) actors is black. The third term is an interaction term allowing for different effects of having a black lead depending on the racial composition of the supporting cast.

The results of this regression (reported in column 3 of table 5) are striking. In particular, we find that Bw movies receive significantly lower scores than movies of any other composition ($\alpha_1 = -3.40$). Using a Ww cast as our base group, we find that Bw movies are scored 3.40 points lower (about a 6% effect). In contrast, Bb movies are scored just 0.49 ($= -3.40 - 0.75 + 3.66$) points lower, and Wb movies are scored just 0.75 points lower – neither being statistically significant¹³. The race of the director continues to not influence critics’ ratings.

In column 4 of Table 5, we rerun our results instrumenting for advertising expenditure. Of all of our controls, advertising seems the most likely to be endogenous since this is adjustable in the short-run and could depend on unobserved quality (at least as perceived by the producers). In fact, this is the only control variable we use that can be adjusted in response to a movie’s perceived quality after production is finished. As an instrument for advertising, for each given movie, we construct a variable that is the average advertising expenditure on all other movies by the given movie’s studio during that movie’s run in movie theaters. Studios generally do not engage in cross promotions for their movies, so it is reasonable to assume that advertising

¹³ These results are robust to interactions between cast racial composition and reviewer race. We do not report these results because the estimates of the interaction parameters are extremely noisy due to the small number of black reviewers (6% of reviews in our sample, see table 1).

expenditures on other movies by the studio should not affect a given movie's rating. We expect our instrument to be positively correlated with advertising since all movies in a given season are likely to see similar demand shocks (e.g., peak demand such as during the Holidays or off-peak such as February). Further, changes in studio advertising budgets may depend on recent previous successes (or failures) by the studio; if the budget increases, then expenditures on all movies released near that time should rise, after controlling for quality differences between them. Our first stage results strongly confirm this intuition (the coefficient on studio advertising for other concurrent movies is positive with a t-stat over 6). The results in column 4 show that our coefficient on advertising does change; however, our primary results for our race variables are virtually identical.

In columns 5 and 6, we consider alternative ways to measure "leader" vs. "follower" roles to evaluate the robustness of our results. As mentioned previously, movies can have two co-leads, or two actors with about the same level of importance to plot, e.g. the two black friends in *My Baby's Daddy* (2004) mentioned in section 3. Consequently, in column 5, we let our "black lead" dummy capture whether the actor listed as the lead *or* as #2 is black, and then define the support as actors in #3-#5 positions. In column 6, we take a similar approach, except here the "black lead" dummy captures whether the actor listed as the lead *and* as #2 are both black. The results for both of these regressions corroborate our main findings. In column 5, we see very similar estimates to our primary regression in column 4. Further, in column 6, we see an even stronger, qualitatively identical effect. We also reran the equation in column (4) redefining *BlackSupp_j* to equal one only if one of the first three supporting cast members is black (as opposed to four). Here again we found very similar results to those in column (4). This gives us confidence that our results are not driven by quirks in movie themes/plots (e.g. movies with leads/co-leads versus those with lead/romantic interest) or in Yahoo!'s definitions of lead versus support #2.

4.2. Impact of Critics' Ratings on Movie Sales

Having established that critic ratings appear to depend on cast racial compositions, we attempt to calculate a rough estimate of the revenue impact from lower ratings for Bw movies. To do this, we utilize our data on total box office revenues and run the following regression:

(3)

$$\ln(TotRev_j) = \gamma_0 + \gamma_1 CriticRating_j + \gamma_2 MovieChar_j + \gamma_3 OpenTheat_j + \gamma_4 Month_j + \xi_{jt}$$

Here, $CriticRating_j$ is the average critic rating for movie j , $MovieChar_j$ contains all information as in (2) above (including racial composition of cast and director), $OpenTheat_j$ is the number of theatres in which the movie opened, $Month_j$ is a dummy for the month during which the movie opened, and ξ_{jt} is an idiosyncratic error term.

[Table 6 about here]

We present the results of this regression in Table 6. Column 1 contains OLS estimates, and in column 2 we again instrument for advertising expenditures as we did in section 4.1¹⁴. For both regressions, we find a similarly positive and significant estimate for average critic rating. Our estimates for the impact of cast racial composition are all statistically insignificant. These results suggest that the cast racial composition of movies does not directly affect revenues, at least in the aggregate. In the absence of individual viewing data, we are unable to assess further why this is the case. We conjecture three possibilities: (1) viewers are not influenced by racial composition of the cast. So our sample of critics is not representative of the population at large; (2) only a segment of the population is influenced by cast racial composition, and this segment is counterbalanced by others that are not, or are influenced by racial composition of the cast in opposite ways. (3) cast racial composition does influence quality assessment by viewers, but quality assessment only occurs after seeing the movie and so cannot directly influence the purchase decision (they must rely on critic reviews and word of mouth from those who have seen the movie to determine quality ex ante)¹⁵.

¹⁴ The IV results suggest that the coefficient on advertising is negatively biased in our OLS revenue regressions (in column 1), as opposed to the apparent positive bias in our OLS ratings regressions in columns 1-3 of Table 5. This may be due to studios choosing to advertise more for higher quality movies that are initially underperforming at the box office.

¹⁵ Consequently, in future research, it would be especially informative to gather individual viewers' assessments of movie quality after seeing the film.

While these results indicate race does not affect revenues directly, they do indicate that it affects revenues via its impact on critic ratings. Using our estimate for average critic ratings from this (IV) regression and our estimate for the rating penalty for Bw movies (column 4 of Table 5), we find that biased ratings can translate into approximately a 4% loss ($3.4 * 1.3\%$) in total box office revenues. The average revenue for movies in our sample is \$64.2 million, and therefore the 4% loss in box office is about \$2.57 million.

Following arguments in Reinstein & Snyder (2005), we recognize that our estimates for the impact of critic ratings on revenues is rough. Specifically, we are not instrumenting for the critic review variable, and so these estimates likely confound critics' influencer role with their predictor role, despite the fact that we include a substantially larger number of controls. For this reason, these estimates likely yield an upper bound for the influencer effect of critic reviews. Given that Reinstein & Snyder (2005) find that critics do influence consumer demand for a movie (the influencer role), we conclude that the box office revenue cost of the critic bias we find is strictly bounded away from zero on the left and 4% on the right (i.e., $0 < \text{revenue effect} < 4\%$). However, an increasingly large fraction of studio profits are made in secondary channels (DVD sales and rentals, pay-per-view, television syndication, etc.), and these sales are highly correlated with theatrical sales. Therefore, the total revenue effect of racially biased critic reviews (including all secondary channels) could be even larger.

4.3. Discussion

The nature of our results (lower reviews for black lead, white support movies) along with the common view that newspapers are generally left-leaning and the high visibility of reviews, lead us to believe that explicit discrimination (taste-based or statistical) is unlikely in this case. In particular, had movies with black leads per se (both Bb and Bw) or movies with a significant number of black actors (Bb) received lower scores, these lower scores could have been indicative of a systematic distaste for black actors or some form of statistical discrimination. It seems unlikely that professional critics consciously decide they like a movie less because it has a Bw configuration (due to distaste for such a lineup or statistical belief that this lineup will produce lower quality. Note that the statistical discrimination argument is weak to begin with given critics view the movie, observe its quality and then write a review). Consequently, we revisit the two relevant mechanisms of non-conscious discrimination from literature in psychology

discussed earlier (expectancy violation, and shifting-standards theory) and their predictions concerning our results in turn.

We begin with expectancy violation. The result of higher critic rankings for Bb movies compared to Bw movies) suggests that there is a negative expectancy violation in seeing blacks in lead roles when all of the “followers” are white. This is not surprising given blacks have traditionally starred in supporting roles in movies, and there is a disproportionately lower number of blacks in several “lead roles” in society, relative to their percentage of the population.

As an exploratory exercise, we analyzed whether lead roles in Bw movies were cross-over roles, i.e. roles that are not specific to blacks. A cross-over role played by a black actor is likely to cause more expectancy violation for the critic than a role specifically designed to be played by a black. To measure whether roles are cross-over, we used two methods. First, using Bureau of Labor Statistics’ data on black employment in various professions, we measured if blacks were over-represented or under-represented in any profession relative to their percentage in the working population of the U.S. For example, the movie “Catwoman” (2004) is about a shy black graphic designer who works for a cosmetics company, and transforms into a woman with magical powers¹⁶. The percentage of blacks in “Arts and related fields” is 2.6% of total employed, and in “Design” is 4.0%, compared to blacks comprising 11% of total employed population. Therefore, this role could easily have been played by a white actress; in fact, having a white actress in the role would be more representative of racial employment patterns. However, not all roles can be classified using national statistics on industry employment by race. In some cases, we have to use plot details to classify whether the lead role is black-specific or cross-over. Consider the movie “Barbershop 2: Back in Business” (2004) set in a black barbershop in South Chicago. This movie can be expected to have a black lead with some (black and some) white supporting cast. Therefore, we classify the lead role as being black-specific because the movie plot demands that it be so. The movie Freedomland (2006) is about a white woman who claims a black man kidnapped her son, and the case is investigated by a black detective. While blacks might be under-represented in higher ranks in the police force, the story line calls for a black

¹⁶ Incidentally, Halle Berry, the actress who played the lead role, received the Golden Raspberry or Razzie award for the Worst Actress for playing this role. This award was instituted to (dis)honor the worst performances in movies in any year. She had previously won the Oscar for Best Actress in 2001. Halle Berry’s father is African-American and mother Caucasian.

police lead detective. Therefore, we classify this lead role as also being black-specific. Following this classification scheme, we find that 82% of critics' ratings of black-lead movies in our data are for cross-over lead roles. Interestingly, we find the less black-specific the lead role, the lower the critics' rating (correlation = 0.29). That is, critics rate higher movies where lead roles are meant to be played by blacks, but rank lower movies where blacks are cast in roles that should, statistically speaking, be played by whites

As mentioned in section 2, shifting-standards theory predicts the following: while black actors might more easily receive a passing grade when cast in lead roles, because of the prior lower expectations of them, black actors will still be held to a higher absolute standard than (higher-status) white actors. Therefore, we would have expected Bb movies to be lowest ranked, if shifting standards are applied to each team member, and Bw to be ranked higher.

Consider now the status leakage theory (Podolny, 2005). As mentioned in section 2, the lower status of the leader of a team can cause lower rankings for the rest of the team as well, even if the rest of the team is objectively of higher quality or higher status. Therefore, Ww will be ranked above Bw, because of the contamination of status of whites in a Bw movie. To understand why Bb is ranked above Bw, it would have to be the case that the contaminated higher status of white support in Bw results in lower status than the (to begin with) low status of black support in Bb. Consider the work of Cabrera et al. (2008). They find that performance expectations for a team were more favorable when the leader's gender was congruent with the industry's gender-typing. For example, if a team works for a male investment banker, people might judge them to be better at their jobs than if the team worked for a female banker. In our industry, if black lead roles are cross-over roles and not black-specific, critics might evaluate the whole cast and movie as lower. Seeing a black lead in a movie with black support is less likely to violate expectations for the race of the leader ("There are several black actors in the movie, it's not unusual that they have a black leader"), and therefore status leakage is not triggered. This then could be another mechanism (in addition to expectancy violation theory) through which Bb movies are rated more highly than Bw movies.

Summarizing, literature in psychology suggests the mechanism for implicit discrimination in our context might be as follows: seeing black leads with white supporting cast members violates both the expectation of who critics expect to be starring in non-race specific

roles and presents a situation where the lower status of the lead in (the unexpected) lead role leaks to the white supporting cast.

We conclude this section by discussing a couple possible caveats for our findings. While we have attributed our empirical findings to implicit discrimination on the part of critics, it is possible that our results could be driven by discrimination and/or selection at the production level. Specifically, it is possible that black leads get roles after these roles have been turned down by white leads (with greater market power), perhaps because of weaker scripts and other quality dimensions not captured by our measures of quality. However, this phenomenon is likely true of non-lead roles as well; that is, for any cross-over support role, a black actor might get cast after a white one turns down the role. By this logic, there is no reason why a Bb movie should be ranked higher than a Bw movie. Also, if this explanation were true, we would expect to see a sales penalty for (the lower quality) Bw movies in the box office; results from regression (4) do not support this. Another explanation for our results is the longer history of discrimination in casting in Hollywood, mentioned in the introduction. Therefore, it is possible that producers or casting directors are less willing to cast black leads in movies they believe to be of higher quality, i.e. that only the lower quality movies are available to blacks. Once again, if this were true, then the lowest quality movies should be cast Bb; this is inconsistent with our result of Bb having greater critical ranking than Bw. Additionally, we do not observe a sales penalty for a Bw movie; this fact is inconsistent with a Bw movie being lower quality. .

5. Conclusion

Despite accounting for a large number of controls, we find that critics' movie ratings have a systematic relationship with cast racial compositions. Specifically, movies with black leads and all-white supporting casts (Bw) appear to suffer about a 6% penalty in their ratings compared to movies with other cast racial compositions (Bb, Ww and Wb). This finding is robust to alternative measures of "lead" vs. "support" roles, and even appears stronger when "black lead" is defined as the top two cast members being black. Aside from the potential social welfare costs resulting from this bias, we find that this critic ratings penalty can result in theatrical box office revenue costs of up to 4% for the average film.

Given our findings and the industry we study, our results seem most consistent with implicit (or non-conscious, in psychology terms) discrimination, as opposed to taste-based or statistical discrimination (both of which are conscious decisions). The two possible mechanisms for this implicit discrimination are expectancy violation theory and/or status leakage theory. According to these theories, seeing black leads with white supporting cast members violates the expectation of who critics expect to be starring in non-race specific roles, and presents a situation where the lower status of the lead leaks to the white supporting cast, respectively. In either case, the result would be a lower overall evaluation of the film.

There are several ways in which this study could be extended. A logical extension would be to investigate critical bias in favor or against other minority groups including female leads. It might also be instructive to see how black newspapers rate movies, given they are writing for a different audience than black critics in our sample who are writing for general newspapers. Various experiments have found that blacks are subject to negative stereotypes about themselves (see Clark and Clark 1940 for the original study of black children preferring white dolls to black dolls; see Edney 2006 for similar and more recent results). Unfortunately, we were unable to find any viable data on critics' reviews from these newspapers. Alternatively, if sufficient data could be found, it might be instructive to study how the critics in our sample rate foreign films with mixed race cast compositions (e.g. Brazilian movies with mixed races, or Japanese movies with a white actor in it) to understand if racism varies with plot context.

Bibliography

- Abrams, D., M. Bertrand, and S. Mullainathan (2006), "Do Judges Vary in their Treatment of Race?", *working paper*, University of Chicago.
- Arrow, K.J. (1998), "What Has Economics to Say about Racial Discrimination?". *Journal of Economic Perspectives*, 12(2), 91-100.
- Avery, C. and J. Chevalier (1999), "Identifying Investor Sentiment from Price Paths: The Case of Football Betting", *Journal of Business*, 1999, 72(4).
- Ayres, I. and Siegelman, P. (1995), "Race and Gender Discrimination in Bargaining for a New Car", *American Economic Review*, 85(3), 304-321.
- Bertrand, M., D. Chugh, and S. Mullainathan (2005), "Implicit Discrimination", *American Economic Review*, 95 (2), 94-98.
- Biernat, M. and D. Kobrynowicz (1997) "Gender- and Race-Based Standards of Competence: Lower Minimum Standards But Higher Ability Standards For Devalued Groups", *Journal of Personality and Social Psychology*, 72, 544-557.
- Blanchflower, D., P. Levine, and D. Zimmerman (2003), "Discrimination in the Small-Business Credit Market", *Review of Economics and Statistics*, 85(4), 930-943.
- Burgoon, J (1978), "A Communication Model of Personal Space Violations: Explication and An Initial test", *Human Communication Research*, 4, 129-142.
- Burgoon, J and M. Burgoon (2001), "Expectancy Theories", in W. P. Robinson and H. Giles, eds. *The new handbook of language and social psychology*. John Wiley & Sons Ltd. 79-10.
- Cabrera, S, S. Sauer and M. Thomas-Hunt (2008), "The Evolving Manager Stereotype: The Effects of Industry Gender-typing on Performance Expectations for Leaders and Their Teams", forthcoming, *Psychology of Women Quarterly*.
- Chandra, A. (2000), "Labor-Market Dropouts and the Racial Wage Gap: 1940-1990", *American Economic Review*, 90(2), 333-338.
- Clark, K. and M. Clark (1940), "Skin Color as a Factor in Racial Identification of Negro Preschool Children", *Journal of Social Psychology*, 11, 159-169.
- Dovidio, J. K. Kawakami, and S. Gaertner (2002), "Implicit and Explicit Prejudice and Interracial Interaction", *Journal of Personality and Social Psychology*, 82, 62-68.
- Edney, T (2006), "New 'Doll Test' Produces Ugly Results", August 16, *Baltimore Times*.

- Eliashberg, J. and S. Shugan (1997), "Film Critics: Influencers or Predictors?", *Journal of Marketing*, 61 (2), 68-78.
- Foschi M. (1996), "Double Standards In the Evaluation of Men and Women". *Social Psychology Quarterly*, 59, 237-54.
- Gaertner, S. and J. Dovidio (1986), "The aversive form of racism", in J.Dovidio and S.Gaertner (Eds.), *Prejudice, Discrimination and Racism: Theory and Research*. Orlando, FL: Academic Press, 61-89.
- Gentzkow, M. and J. Shapiro (2006), "What Drives Media Slant? Evidence from U.S. Daily Newspapers, *NBER Working Paper* No. W12707.
- George, L. and J. Waldfogel (2003), "Who Affects Whom in Daily Newspaper Markets?", *Journal of Political Economy*. 111, 765-784.
- Goh, J and L. Ederington (1993), "Is a Bond Rating Downgrade Bad News, Good News, or No News for Stockholders?", *Journal of Finance*, 48 (5), 2001-2008.
- Holbrook, M.B. (2005), "The Role of Ordinary Evaluations in the Market for Popular Culture: Do Consumers Have "Good Taste"?", *Marketing Letters*, 16 (2), 75-86.
- Khatami, E. (2009), "Black Actors Still Face Oscar challenges", *CNN article*: <http://www.cnn.com/2009/SHOWBIZ/Movies/02/19/black.actors.oscar/index.html>
- Podolny, J. M. (2005). *Status signals: a sociological study of market competition*. Princeton: Princeton University Press.
- Powalski, A. (2009), "Film Fest Explores Hollywood's Take on Latinos", *CNN article*: <http://www.cnn.com/2009/SHOWBIZ/Movies/05/05/latino.images.film.festival/index.html>
- Price, J. & Wolfers, J. (2007), "Racial Discrimination Among NBA Referees", forthcoming, *Quarterly Journal of Economics*.
- Reinstein, D. and C. Snyder (2005), "The Influence of Expert Reviews on Consumer Demand for Experience Goods: A Case Study of Movie Critics", *Journal of Industrial Economics*, 53, 27-53.
- Robinson, R. (2006), "Hollywood's Race/Ethnicity and Gender-Based Casting: Prospects for a Title VII Lawsuit", UCLA Chicano Studies Center, *Latino Policy Brief Number 14*.
- Yinger, J. (1998), "Evidence on Discrimination in Consumer Markets", *Journal of Economic Perspectives*, 12(2), 23-40.

Table 1: Summary Statistics for Critics and Reviews (N= 4339) ¹⁷

Variable	Mean	Std. Dev.
Rating	58.14	21.45
Gender (Female = 1)	0.26	0.44
Race (Black = 1)	0.06	0.24
Number of Reviews	1062.8	682.3

Table 2: Audience Demographics for the 11 Major Newspapers for Which Critics in the Data Review

Newspaper	% White Readers	% African-American Readers	% readers aged 18-24	% readers aged 55+	% readers with college degree	% readers with income under \$35k	% readers with income over \$100k+	Location
Boston Globe	88.7	5.8	9.7	32.7	44.0	16.3	29.8	East
Charlotte Observer	79.7	17.5	9.4	33.7	30.0	21.8	18.2	South
Chicago Sun Times	64.7	31.3	13.0	31.6	21.5	21.8	22.0	Midwest
Chicago Tribune	82.2	12.6	9.2	35.9	40.0	14.2	31.6	Midwest
Los Angeles Times	78.9	9.3	10.3	33.4	36.4	17.6	28.7	West
Miami Herald	75.4	20.8	9.9	34.2	29.3	23.8	21.6	South
New York Times	84.1	9.7	12.0	34.4	57.2	13.1	38.7	East
Philadelphia Inquirer	77.5	18.2	8.2	38.2	33.3	19.9	24.3	East
San Francisco Chronicle	79.3	5.8	7.3	38.1	46.5	12.7	39.5	West
Seattle Post Intelligence ¹⁸	87.2	3.5	8.3	33.3	37.5	17.5	21.2	West
Washington Post	65.4	28.0	9.1	30.5	46.3	10.3	40.8	East

¹⁷ Note that our summary statistic for rating is at the critic/movie level since each critic reviews many movies. The other two summary statistics are for variables that vary at the critic level.

¹⁸ This newspaper has since shut down its print edition, and is only available online.

Table 3a: Number of movies by race of cast member and position in cast

	Lead	Supporting 1	Supporting 2	Supporting 3	Supporting 4
Black Actor	64	57	49	59	46
White Actor	502	509	517	507	520
Total	566	566	566	566	566

Table 3b: Average bankability by race of cast member and position in cast

	Lead	Supporting 1	Supporting 2	Supporting 3	Supporting 4
Black Actor	46.4	26.8	24.0	11.8	8.0
White Actor	39.3	23.2	16.1	11.2	8.8
All Actors	40.1	23.6	16.8	11.2	8.7

Table 4: Summary Statistics for Movie Characteristics (N = 566)

Variable	Mean	Std. Dev.
Lead is Black	0.11	0.32
At Least One Black Support	0.28	0.45
Budget (\$ Mil.)	51.0	43.0
Advertising (\$ Mil.)	30.0	9.8
Director is Black	0.04	0.21
BFCA Nom. + Awards of Lead	0.61	1.03
BFCA Nom. + Awards of #2	0.34	0.75
BFCA Nom. + Awards of #3	0.21	0.61
BFCA Nom. + Awards of #4	0.11	0.40
BFCA Nom. + Awards of #5	0.10	0.41
Bankability of Lead	40.1	33.6
Bankability of #2	23.6	25.9
Bankability of #3	16.8	21.3
Bankability of #4	11.2	17.8
Bankability of #5	8.7	16.2
BFCA Nominations * (0.5) + Awards	0.25	0.80
Buena Vista	0.09	0.29
Dream Works	0.05	0.23
Fox	0.13	0.33
MGM	0.05	0.21
Miramax	0.03	0.17
New Line	0.07	0.25
Paramount	0.09	0.29
Sony	0.09	0.29
Universal	0.10	0.30
Warner Brothers	0.11	0.32
Animated	0.08	0.28
Action and Adventure	0.22	0.42
Comedy	0.27	0.44
Horror	0.17	0.38
Romantic Comedy	0.08	0.27
Sci-Fi	0.04	0.20
Drama	0.13	0.34
G Rated	0.03	0.18
PG Rated	0.19	0.39
PG13 Rated	0.46	0.50
R Rated	0.32	0.47
Total Box Office Revenue (\$ Mil.)	64.2	67.5

Table 5: OLS and IV Regressions for Critic Reviews on Racial Composition of Cast and Controls (N = 4339) ¹⁹

Covariate	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) IV	(6) IV
	Lead: #1 Supp: #2-#5	Lead: #1 Supp: #2-#5	Lead: #1 Supp: #2-#5	Lead: #1 Supp: #2-#5	Lead: #1 or #2 Supp: #3-#5	Lead: #1 & #2 Supp: #3-#5
Lead is Black	-3.69 (0.99)**	-1.45 (1.16)	-3.30 (1.77)+	-3.43²⁰ (1.78)+	-2.97 (1.13)*	-9.06 (3.34)**
At Least 1 Black Support			-0.71 (0.85)	-0.37 (1.05)	0.28 (1.38)	0.55 (1.06)
Lead is Black * At Least 1 Black Support			3.57 (2.09)+	3.21 (2.52)	2.32 (2.22)	6.79 (3.73)+
Black Director		0.34 (1.55)	0.32 (1.52)	0.53 (1.54)	0.72 (1.45)	1.34 (1.48)
BFCA Awards + Noms*.5		6.82 (0.33)**	6.81 (0.33)**	6.85 (0.39)**	6.91 (0.40)**	6.92 (0.41)**
Budget		0.003 (0.014)	0.004 (0.014)	0.01 (0.05)	0.01 (0.05)	0.02 (0.05)
Advertising		0.10 (0.05)+	0.11 (0.05)*	0.05 (0.42)	0.07 (0.41)	0.002 (0.430)
Awards + Noms for Actor #1		1.05 (0.30)**	1.03 (0.30)**	0.93 (0.32)**	0.91 (0.32)**	0.93 (0.35)*
Awards + Noms for Actor #2		1.77 (0.43)**	1.77 (0.43)**	1.70 (0.51)**	1.70 (0.51)**	1.71 (0.51)**
Awards + Noms for Actor #3		2.15 (0.60)**	2.17 (0.60)**	1.99 (0.64)**	2.01 (0.64)**	1.98 (0.64)**
Awards + Noms for Actor #4		-0.89 (0.85)	-0.84 (0.86)	-0.89 (0.97)	-0.93 (0.97)	-0.89 (0.98)
Awards + Noms for Actor #5		0.92 (0.79)	0.94 (0.80)	1.31 (1.17)	1.37 (1.19)	1.48 (1.25)
Awards + Noms for Director		2.48 (0.62)**	2.44 (0.63)**	2.58 (0.65)**	2.50 (0.65)**	2.64 (0.62)**
R-Square	0.003	0.260	0.261	0.265	0.266	0.265

¹⁹ Additional controls in all but regression (1) i.e. in regressions (2)-(6) include: month/year fixed effects; bankability of actors #1-#5; gender, experience, and race of reviewer; studio fixed effects; genre fixed effects; movie rating (e.g., R, PG, PG-13); newspaper fixed effects; region fixed effects; and proportion of newspaper audience that is: over 55, under 25, college educated, black, has income < \$25K, has income > \$100K. Standard errors are in parentheses and were clustered by critic. ** is significant at 1% level, and * is significant at 5% level.

²⁰ P-value is 0.059.

Table 6: OLS and IV Regressions for Log of Total Revenue on Average Critic Rating and Controls (N = 566) ²¹

Covariate	(1) OLS	(2) IV
Avg. Critic Rating	0.017 (0.002)**	0.013 (0.004)**
Budget	0.001 (0.001)	-0.003 (0.004)
Advertising	0.068 (0.005)**	0.136 (0.071)+
BFCA Awards + Noms*.5	0.139 (0.034)**	0.012 (0.140)
Lead is Black	0.109 (0.085)	0.121 (0.128)
At Least 1 Black Support	0.058 (0.067)	0.012 (0.097)
Lead is Black * At Least 1 Black Support	0.009 (0.141)	0.114 (0.208)
R-Square	0.767	0.682

²¹ Additional controls include month/year fixed effects, awards and nominations of actors #1-#5 and director, bankability of actors #1-#5, cast racial composition, director race, number of opening theatres, studio fixed effects, genre fixed effects, and movie rating (e.g., R, PG, PG 13). Robust standard errors are in parentheses. ** is significant at 1% level, and * is significant at 5% level.